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, and 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 Communications Directorate (0704-0188). Respondents should be aware that notwithstanding any 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 valid 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)*
20-02-2019	Book Chapter	

4. TITLE AND SUBTITLE
Operational Ocean Data Assimilation

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
0602435N

6. AUTHOR(S)
Gregg A. Jacobs, Charlie N. Barron, Cheryl A. Blain, Matthew J. Carrier, Joseph M. D'Addezio, Robert W. Helber, Jackie C. May, Hans E. Ngodock, John J. Osborne, Mark D. Orzech, Clark D. Rowley, Innocent Souopgui, Scott R. Smith, Jay Veeramony, Max Yaremchuk

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
73-6A53-07-5

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
Oceanography Division
Stennis Space Center, MS 39529-5004

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL/BC/7320--18-3761

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
One Liberty Center
875 North Randolph Street, Suite 1425
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Operational ocean data assimilation is necessary to continually correct and maintain accurate ocean forecasts. The primary GODAE objective was prediction of mesoscale eddies, and the science community has successfully addressed this issue. Here, we examine a data assimilation process for this problem, beginning with a generalized solution of 4D variation assimilation (4DVar) so that assumptions will be clear as we reduce to a 3DVar that is often used operationally. The primary difficulty lies in specifying the covariances that relate variables at different locations in space and time. Simplifications are applied to provide covariances that sufficiently describe the relations and are computationally feasible. Some deficiencies are introduced through the assumptions leading to the 3DVar and within the covariances, and this points to areas of future research. Prior assumptions were predicated on the expected observing systems and numerical model capabilities, which were all consistent with prediction of mesoscale features. We believe that numerical models and observations will surpass present capability, and there is strong motivation to move data assimilation forward to achieve prediction at scales not now feasible.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Gregg A. Jacobs
Unclassified	Unclassified	Unclassified	UU	33	19b. TELEPHONE NUMBER *(Include area code)* (228) 688-4720

Reset

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

CHAPTER 18

Operational Ocean Data Assimilation

Gregg A. Jacobs[1], Charlie N. Barron[1], Cheryl A. Blain[1], Matthew J. Carrier[1], Joseph M. D'Addezio[2], Robert W. Helber[1], Jackie C. May[1], Hans E. Ngodock[1], John J. Osborne[4], Mark D. Orzech[1], Clark D. Rowley[1], Innocent Souopgui[3], Scott R. Smith[1], Jay Veeramony[1], and Max Yaremchuk[1]

[1]Oceanography Division, Naval Research Laboratory, Stennis Space Center, Mississippi, USA; [2]University of Southern Mississippi, Stennis Space Center, Mississippi, USA; [3]University of New Orleans, Stennis Space Center, Mississippi, USA; [4]ASEE Postdoc, Naval Research Laboratory, Stennis Space Center, Mississippi, USA

Operational ocean data assimilation is necessary to continually correct and maintain accurate ocean forecasts. The primary GODAE objective was prediction of mesoscale eddies, and the science community has successfully addressed this issue. Here, we examine a data assimilation process for this problem, beginning with a generalized solution of 4D variation assimilation (4DVar) so that assumptions will be clear as we reduce to a 3DVar that is often used operationally. The primary difficulty lies in specifying the covariances that relate variables at different locations in space and time. Simplifications are applied to provide covariances that sufficiently describe the relations and are computationally feasible. Some deficiencies are introduced through the assumptions leading to the 3DVar and within the covariances, and this points to areas of future research. Prior assumptions were predicated on the expected observing systems and numerical model capabilities, which were all consistent with prediction of mesoscale features. We believe that numerical models and observations will surpass present capability, and there is strong motivation to move data assimilation forward to achieve prediction at scales not now feasible.

Introduction

The goal for operational ocean prediction is to enable decisions that address societal problems across time scales from hours to years. For example, flooding and other impacts due to storms require accurate forecasts to determine evacuations, coastal ocean conditions off Louisiana and Texas leading to hypoxia determine the quantity of fertilizer that farmers in Iowa may use, and growth in aquaculture requires consideration of material transport over years as well as its impact on the ocean environment. For all these problems and more, the end measure of success is the accuracy of the ocean prediction.

We must consider the physics governing ocean evolution to understand the application of the assimilation process. Incoming solar radiation drives the global physical environment. The relatively higher energy input per unit of Earth surface at the equator versus the poles results in spatial density and pressure gradients that drive atmosphere flows, ocean flows, surface waves, ice, and a hydrological cycle of rainfall and river runoff. Energy flowing within and between these connected systems converts into a myriad of forms that we see as storms, eddies, and many other individual events. For the ocean, nonlinear instability processes develop into mesoscale eddies,

Jacobs, G.A., et al., 2018: Operational ocean data assimilation. In "*New Frontiers in Operational Oceanography*", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView, 513-544, doi:10.17125/gov2018.ch18.

which is a result of energy exchange from the large-scale potential and kinetic energy fields into the growth of mesoscale features. These eddy features have been observed in situ and remotely to have timescales on the order of ten days and spatial scales from hundreds of km in the tropics to tens of km in higher latitudes (IMAWAKI, 1981; Richman et al., 1977; Le Traon, 1991). Mesoscale perturbations grow exponentially in time (Charney, 1947), and thus errors in a model initial condition grow exponentially through the forecast. The ocean eddy physics limit the time period over which accurate predictions may be made, and these features are described as nondeterministic. The time and location of specific features are of primary interest in operational decisions. This mandates a process to continually correct the model evolution for mesoscale prediction. Operational assimilation provides this process.

This chapter begins with a general formulation of the problem. There are many excellent sources that examine the historical evolution of the data assimilation and forecasting problem (Daley, 1993; Kalnay, 2003). The history provides insight into how considerations have continually extended the data assimilation problem leading from initial simple solutions to the present day complex systems. In this chapter, rather than examine the historical development, we begin with general formulation of a four-dimensional variational (4DVar) data assimilation that incorporates many of the past considerations (Bennett, 2002).

Following the general formulation, we outline simplifications of the 4DVar for the ocean mesoscale prediction problem. In past years, computational resources in the operational environment were insufficient to apply a 4DVar approach. At present, the 4DVar approach may be applied to the ocean in limited regions (Moore et al., 2011; Ngodock and Carrier, 2014; Powell et al., 2008), though the global ocean 4DVar assimilation problem is not presently computationally feasible. Explicit assumptions are made to reduce the 4DVar to a three-dimensional variational solution (3DVar), optimal interpolation, or Kalman Filter. The general assimilation problem contains statistical covariance information relating error of one variable at one space and time location to another variable at a different location, and the relations between all variables are contained in the covariance matrix. The assumptions simplifying the covariance matrix reduce the 4DVar to a 3DVar solution.

The covariance relations are extremely important for ocean predictions. Relative to the scales of eddies, present ocean observations are somewhat sparse. Recent observations are the input information to the data assimilation process. For GODAE mesoscale forecasts, the primary data sources have been satellite observations of sea surface height (SSH), sea surface temperature (SST), and profile observations of temperature and salinity from ARGO amongst others (Le Traon, Rienecker et al. 1999). It has been recognized that even with the range of satellite and in situ global systems, the spatial and temporal coverage is relatively sparse, and efforts to augment the regular observations have been proposed (Testor, Meyers et al. 2010). In particular, for ocean prediction, the data assimilation process must extend the influence of observations from individual points or small area averages (such as satellite footprints) of the ocean state. The covariances provide the relations of errors in variables at observation locations to other variables at other locations in space

and time. For example, the covariances must extend observations such as SSH to the ocean interior temperature and salinity.

When computing the corrections to a model state at a particular time, the covariance relations depend on the present conditions such as the positions of mesoscale features. The covariances between variables across an eddy front will be very different from those within an eddy. The covariances also change over time. Before observations correct the model, there is a background covariance, and once observations correct the model the covariance magnitudes decrease in the analysis covariance. During the time period from one correction to the next, the forecast covariance magnitudes increase until the forecast covariance becomes the background for the next assimilation cycle. There are analytic solutions in a linear system for the covariance decrease in the analysis and increase in the forecast period. However, the high degree of freedom nonlinear ocean system makes the problem intractable.

Present operational systems utilize many simplifying assumptions. Because covariance relations provide the critical function of extending sparse observations, covariance simplifications must be carefully considered due to effects on reduced forecast skill. The covariance matrix in the reduced 3DVar remains daunting, and specifying the statistical information has required further simplifying assumptions. One example we examine assumes an analytic representation of the covariances without considering dependence on feature location, and a simple representation of the time variation is incorporated. Because there are simplifying assumptions to the covariance matrix, these restricts predictive skill.

We must recognize such assumptions because these imply where the frontiers of future research lie. The simplifications reducing the problem from 4DVar to 3DVar also remove time correlated information. Ocean features have long time correlations. The 4DVar problem retains these correlations, and additional considerations of the covariances are required. Approaches such as ensemble methods are intended to incorporate temporally and spatially changing information within the covariances based on the dynamical evolution (Evensen, 2003). All these methods are at the leading edge of research as the next generation of operational systems is developed, and we examine an example of considerations within the 4DVar system.

Operational data assimilation requires considerations beyond just the data assimilation process. We must keep in mind other parts of the overall prediction system. Predicting features in an operational environment must consider a balance of 1) physical understanding and representation in numerical models, 2) computational capability that limits resolution of features in numerical models, 3) systems that are typically insufficient to observe the features high resolution numerical models may represent, and 4) assimilation capability to specify covariances and correct numerical forecasts with the given observations. In all of these areas, significant limitations exist. Understanding physics within just the ocean is incomplete. Computational capacity limits model resolution and, therefore, physics. Observations typically are sparse. It is not possible to fully and accurately specify the full time-evolving covariance. All of these limitations motivate advancing the data assimilation problem.

The measure of success can be traced from perspective of the operational decision. Are decisions made correctly with greater frequency? Is the environmental information on which decisions are based more accurate? Are the environmental features that affect the decisions information more accurately forecast? Following the chain from operational decisions to accuracy in forecasts is a challenging problem. We know that GODAE systems have skill in forecasting mesoscale eddies (Dombrowsky et al., 2009). For many applications, such as search and rescue and oil spill response, the trajectory of objects or material is of high interest. For these problems, trajectory accuracy and forecast placement of ocean features on the order of 10 km is required. We provide a recent example showing that present systems do not meet this challenging requirement, which implies there is much work to be done in the future.

In this chapter, we examine the operational ocean data assimilation problem to understand the considerations. The second section provides a brief general formulation of the problem, and the third section provides simplifying assumptions leading from the general 4DVar solution to a 3DVar solution used operationally. Within any assimilation approach are covariance relations between variables, space, and time. The fourth section provides one example of a covariance implementation in 3DVar that considers maintaining consistency with ocean physics. One simplifying assumption in reaching the 3DVar reduces time correlations, and the considerations and impact when incorporating time correlations are examined in the fifth section. Finally, looking to the future, in the final section, we gauge the performance of a present system prediction to independent observations in a difficult problem of providing drift predictions.

Formulation

Data assimilation corrects operational forecast systems on a regular basis. The frequency varies across GODAE systems from every six hours (Saha et al., 2010b) to daily (Cummings et al., 2009) to weekly (Brasseur et al., 2005). The typical process uses a prior forecast as a background, and the background is subtracted from observations over a time period leading up the present to form a set of innovations. The data assimilation is often a 3DVar that uses the innovations to compute an increment that corrects the model at the present time, and the corrected model state is integrated forward over a forecast interval. We first formulate the problem in a more generalized manner to understand what assumptions lead to this particular procedure and therefore the inherent limitations.

A well-posed forecast problem consists of initial and boundary conditions (IBCs) and dynamical specifications. If we consider the global ocean system with a given initial condition, the boundary conditions are surface momentum flux, surface heat flux, surface turbulence injection, and incoming solar radiation. From the IBCs, an initial forecast is made. For a regional ocean application, lateral boundary conditions of the dynamical system augment the IBCs. Conceptually, specific model implementation information also augments the IBCs. This includes bathymetry, tidal boundary conditions, bottom friction parameterization, lateral river transports, and other coefficients within horizontal and vertical mixing. While typical operational applications do not make regular

corrections of model implementation information, many data assimilation problems address these as optimal parameter estimations. Including these as IBCs in the general formulation allows extension to many problems.

Define the state vector x as the concatenation of all IBCs and model variables over the time period of the observations and over the forecast period. In addition to the ocean temperature, salinity, velocity, and surface elevation, every intermediary variable calculation in the model solution can be considered as part of the state. This includes intermediary fluxes of momentum and heat, horizontal and vertical diffusivity, and others. For convenience, we assume that the state is ordered in the same order as variables are computed in the model solution. A subset of the state is provided by the IBCs, and the dynamics of a numerical model provides the complementing set. The IBCs can be written in a matrix form as $A_{IBC}x = b_{IBC}$, where each row of the matrix A_{IBC} corresponds to setting one state value corresponding to b_{IBC}. The dynamics are nonlinear, and these can be written in a similar matrix form as $D = b_D$, where coefficients of the ith row of D may depend on state variables with indices less than i. Thus D is lower left triangular. As the problem is well posed, A_{IBC}, D, b_{IBC}, and b_D define a unique solution, and we refer to this solution as the background state x^b. The background state is traditionally the forecast conducted after the prior data assimilation cycle. Together, the IBCs and dynamics compose the model, so that the full model can be cast as:

$$\begin{bmatrix} A_{IBC} \\ D \end{bmatrix} x^b = \begin{bmatrix} b_{IBC} \\ b_D \end{bmatrix}, \text{ or } Mx^b = b_M \tag{1}$$

The observations are y^o , and the operator H acts on the model state to simulate the observing systems. The observation innovation vector is $d = y^o - Hx^b$. In some instances, the observation value may depend on the state, so that H is nonlinear. This is typically not the case in the ocean, but it does occur in examples such as acoustic or optical data. We seek a state that satisfies all the IBCs, dynamics, and matches the observations. However, the problem with only IBCs and dynamics is well-posed. Adding the additional requirement to match the observations over-determines the problem. Thus, our objective is to compute an analysis state that minimizes the squared errors to the background IBCs, the dynamical equations, and the observations:

$$\begin{bmatrix} M \\ H \end{bmatrix} x^a = \begin{bmatrix} b_M \\ y^o \end{bmatrix} + \begin{bmatrix} \eta \\ \varepsilon \end{bmatrix} \tag{2}$$

where the errors to IBCs and dynamics are contained in the model errors η , the observations are y^o , and observations errors are ε. The problem is cast to find a correction δx to the background x^b so that $x^a = x^b + \delta x$. In this context, the dynamical operator is linearized around the background state so that the tangent linear model is M'. The assumption is that the solution time interval is short enough that a linearized dynamical operator is sufficiently accurate. This can be tested by comparing the tangent linear solution to a full nonlinear solution, both with a small

perturbation from the background. The optimal increment should provide a minimum squared error to

$$\begin{bmatrix} M' \\ H \end{bmatrix} \delta x = \begin{bmatrix} 0 \\ d \end{bmatrix} + \begin{bmatrix} \eta \\ \varepsilon \end{bmatrix}, \text{ or simply } A\delta x = b + e. \tag{3}$$

The weighted minimum error variance solution to Eq. 3 is given by $\delta x = \left(A^T W A \right)^{-1} A^T W b$, where the weighting $W = B^{-1}$, or the (pseudo) inverse of the covariance matrix B. The covariance matrix is composed of the values

$$B = \left\langle \begin{bmatrix} \eta \\ \varepsilon \end{bmatrix}, \begin{bmatrix} \eta^T & \varepsilon^T \end{bmatrix} \right\rangle \tag{4}$$

where $\langle , \rangle$ indicates an expected value that is an average over many realizations of the forecast under similar conditions, and we will see that many simplifications are required to estimate this covariance matrix. Assumptions underlying this solution are that the corrections δx are sought in the range of B and are Gaussian-distributed random variables with zero means. For some variables this is appropriate, but there are exceptions to keep in mind. Ice concentration and salinity, for example, cannot have negative values. Therefore, the corrections are not simply Gaussian-distributed with zero mean. Methods to contend with non-Gaussian errors are an active area of research (Bocquet et al., 2010), but for operational applications, we continue with this assumption.

The covariance matrix describes how errors in one IBC, dynamical equation, or observation are related to others. For example, an initial condition temperature error in one location will be correlated to an error in temperature nearby horizontally and vertically. Likewise, an error in the dynamical equation for diffusivity at one location and time will have a spatial and a temporal correlation. In this solution process, we must specify the statistical relation of the errors in every state variable to every other state variable at every location and at every time over the entire forecast trajectory.

When this process is applied to only one system, such as only the ocean or only ice or only the atmosphere, we must realize an assumption has already been made. The state can contain all variables of all the Earth systems together such as $x^T = \begin{bmatrix} x^T_{ocean} & x^T_{atmosphere} & x^T_{ice} & x^T_{waves} \cdots \end{bmatrix}$, and we should recognize that a correction to ocean surface temperature should lead to a correction of near surface atmospheric properties. By applying assimilation to individual systems, we are assuming that the covariances between the systems are zero. If we assume that errors in other dynamical systems are not correlated to errors in the ocean, the covariance matrix becomes block diagonal with vast areas assumed to be zero. Each block may be inverted separately. This assumption allows us to conduct ocean data assimilation separately from other systems. There are coupled fluxes between all these dynamical systems, and the justification for separating them is mainly to reduce the problem to a tractable size rather than considering a scale analysis. The decoupling certainly simplifies the problem, and we should be aware that addressing these covariances within coupled systems remains an area of research (Zhang et al., 2007). There are also intermediary steps toward fully-coupled data assimilation. For example, the computation of fluxes

between the ocean and atmosphere may be considered a separate model, and observations are made that allow estimation of the fluxes. If we include the fluxes as a separate model in the state, we can again assume the error covariance matrix to be block diagonal so that data assimilation may be performed separately on the ocean, atmosphere, and on fluxes between the two.

The optimal correction given by (3) is the solution of a weighted least squares problem, however, the covariance matrix poses two challenges. The first challenge is the size of the matrix that prohibits direct solution by the usual weighted least squares approach. The direct inversion of the large matrices is problematic. If we consider a 1/25°L40 global ocean model with a five-day period prior to the present containing observations and a five-day forecast with a one-minute time step, the state size for just temperature, salinity, velocity, and surface height is about 10^{14} variables. The covariance matrix then contains 10^{28} variables. No computer can store the state in memory, and thus storing the error covariance matrix is orders of magnitude beyond present capacity. The problem is formulated to reduce the size. Here, we start by casting it through a 4DVar approach:

$$\delta x = M'B_4M'^T H^T \left(HM'B_4M'^T H^T + R \right)^{-1} d \tag{5}$$

The covariance matrix in Eq. 4 is separated into two components under the assumption that observation errors ε are not correlated to model errors η. The matrix R is the error covariance of the observations $R = \left\langle \varepsilon, \varepsilon^T \right\rangle$, and the covariance for the 4DVar contains the model error covariances $B_4 = \left\langle \eta, \eta^T \right\rangle$ of dynamics, IBCs and other model input parameters. Both R and the matrix $HM'B_4M'^T H^T$ are the size of the number of observations rather than the size of the full state. So the problem is reduced. The inversion is usually through a conjugate gradient with an appropriate preconditioner so that the actual inverse is never computed. It is only the action of $\left(HM'B_4M'^T H^T + R \right)$ on a vector that is required in the conjugate gradient process. Even when cast in this form, the problem remains very large and must be reduced further.

The second covariance matrix challenge is that there is insufficient observational data on which to compute the statistics represented in the matrix. We must make further simplifying assumptions to enable progress on this problem. We consider further assumptions to address the ocean mesoscale problem in the next section, and further simplifying assumptions are required for the covariance matrix in the subsequent section.

Simplifications for Mesoscale Assimilation

The initial focus of the Global Ocean Data Assimilation Experiment (GODAE) has been the prediction of the mesoscale eddy field. Internal ocean dynamics lead to instabilities that convert potential energy stored in the displacement of isopycnals into kinetic energy. The western boundary currents, such as the Gulf Stream, continually generate very strong eddies, and eddies are pervasive in the interior oceans.

The process outlined in the prior section makes regular corrections to the ocean forecast. The prior forecast is used as the background state x^b, observation differences to the background provide the observation innovations d, and the data assimilation computes an analysis increment δx. In the formulation leading up to the 4DVar in Eq. 5, the analysis increment is computed for every variable over all space and time, as well as the model inputs. This leads to the 4DVar covariance matrix being very large, which requires us to make simplifying assumptions. The temporal representation of corrections can be reduced by assuming a functional form over time, which may be as simple as linear interpolation between correction estimates at times spaced much greater than the model time step. The 3DVar takes this simplification several steps further. The formal assumptions leading from the 4DVar to the 3DVar should be stated explicitly so that we are aware.

The 3DVar first assumes that errors exist only at the times when a new forecast is made, which are the analysis times and that all observations exist only at the analysis times. In addition, assume that the errors are not correlated between analysis times. Finally, assume the dynamical errors are negligible so that $\eta = 0$, which is the strong constraint problem.

The effect of these assumptions is to set areas of the 4DVar error covariance matrix B_4 to zero. Because we have ordered the state vector to correspond to time, the assumption that the errors are not correlated at different analysis times leads to B_4 being zero except at diagonal blocks that represent the covariances of errors at the analysis times. The solution process at each analysis time is now independent, which greatly reduces the problem. The assumption that the dynamical errors are negligible allows reformulating the problem in terms of the Lagrangian multipliers so that the solution from one analysis time to the next exactly satisfies the model dynamics. The assumption that observations are only at the analysis time leads to having to compute the analysis increment only at the analysis time.

At each analysis time, the 3DVar solution is computed:

$$\delta x_3 = B_3 H^T \left(H B_3 H^T + R \right)^{-1} d \tag{6}$$

The analysis increment δx_3 is computed just at the analysis time, and the covariance matrix B_3 is the error covariance at the one analysis time. The dynamical operator in time no longer enters into the assimilation problem. The error covariance matrix B_3 now represents covariances of errors in the background at the analysis time. Because the background field from a prior forecast is typically the initial condition, B_3 is the background error covariance in the 3DVar approach. As in the 4DVar, the matrix $\left(H B_3 H^T + R \right)$ has dimension of the number of observations rather than the state size. At the initial time of GODAE, these assumptions were necessary to reduce the problem to a tractable size.

We still must specify B_3, and this requires some dynamical insight. How are errors of one variable at one (x, y, z) related to another at a different position? This is the most critical part of the problem. If there were full observations of all variables over all (x, y, z), the form of B_3 would not be so important as it would serve to only reduce the observation error from one point to

another. However, the ocean is under-sampled. It is rare that even one aspect of the ocean in one area is sufficiently sampled, let alone over-sampled. This has significant implications. The analysis that serves as the initial condition for the forecast is critically dependent on our ability to specify the error covariances because this matrix greatly extends the influence of observations. This one matrix is the point at which the disciplines of oceanography and data assimilation meet.

As a first example, consider one primary source of observational information for operational ocean prediction, satellite altimeter data. Assume one observation of sea surface height (SSH) occurs at a model grid location. In this case, the observation operator H is a vector of zeros except one element that has a value of one, and the position of that element corresponds to the appropriate observation location. In this case, the matrix $\left(HB_3H^T + R\right)^{-1}$ is a scalar that provides the relative weight of the background error in SSH and the observation error. This scalar multiplies one column of the covariance provided by B_3H^T . The observation operator is selecting one column of B_3 that specifies the relation between SSH errors to all other variables over space. Thus, the correction due to one observation is proportional to one column of the covariance matrix.

How should the background state be changed given SSH observations? Suppose the column of B_3 under consideration specifies that SSH is not correlated to other model variables. The resulting analysis increment would change just the SSH. Prior experiments show that changing only model SSH dissipates quickly as a barotropic wave. The oceanography science community experience is that the primary errors in initial conditions are the positions of mesoscale eddies. The error in eddy feature position leads to errors of internal thermohaline structure that results in SSH errors. Therefore, B_3 must provide the covariances between SSH and the underlying temperature and salinity. Additionally, observations have shown that the dynamical balance between the mesoscale pressure field and the velocity field is primarily geostrophic at low Rossby number (large spatial scale and long time periods). Corrections to temperature, salinity, and SSH that lead to pressure must relate to corresponding changes in the velocity field. The background covariance matrix must specify these properties.

There is no entirely satisfactory method to determine B_3, and there are many proposed. All of these methods have advantages and disadvantages. Two considerations are important when evaluating methods. The first is physical representation and the second is a large sample size to confidently compute statistics. Ideally, we would compute the error covariances based on observations (Helber et al., 2013). This removes the question of any bias of the physical system being used since we would be using the true ocean. However, we require many realizations to compute each element of B_3. There are not sufficient data to compute statistics of errors in all variables and the relation to errors in all other variables at other points in space. A numerical model run can provide many events over a very long period of time to build confident statistics such as those by Oke et al. (2005). However, biases in models can bias the statistics. Relations should be expected to change depending local conditions such as if two variables inside a cyclone, inside an anticyclone, or on opposite sides of a front. The relations also depend on the season since

thermocline structure changes from summer to winter These considerations result in conditional statistics, which are covariance relations given certain conditions such as time of year or information on the local background features. Covariance estimates from observations or historical model runs typically do not include substantial conditional information. Estimates from ensemble-based systems allow the conditional information to be included more naturally (Evensen, 2003). However, the number of ensemble members is typically small, which leads to errors in the estimated covariances and techniques such as localization (Anderson, 2012). At this point, we must make the reader aware that one particular formulation of the covariance will be examined here, but it is an important area of active work with many possible approaches given the references above.

The approach of Brandt and Zaslavsky (1997) initially separates B_3 into variance S_b and a correlation function C_b as $B_3 = S_b^{1/2} C_b S_b^{1/2}$. A decomposition of C_b into separable functions is then made so that the correlation between two field variables v and v' is given by

$$C_b\left(x,y,z,t,x',y',z',t'\right) = C_{vv'}^{H}\left(x,y,x',y'\right) C_{vv'}^{V}\left(z,z'\right) C_{vv'}^{FDB}\left(x,y,x',y'\right) \qquad (7)$$

The correlation between horizontal locations x, y and x', y' and vertical positions z and z' is the product of the separable components of horizontal $C_{vv'}^{H}$, vertical $C_{vv'}^{V}$ and flow-dependent $C_{vv'}^{FDB}$ correlation functions. This separation of variables is relatively typical since there is not sufficient a priori information to provide the full correlations between the values of $v\left(x,y,z,t\right)$ and $v'\left(x',y',z',t'\right)$. When separated in this manner, there begins to be sufficient observational data for certain aspects as examined in the next section. Such a decomposition is applied in general circulation models (Derber and Rosati, 1989), the Harvard Ocean Prediction System (HOPS; Lozano et al., 1996), MERCATOR (Brasseur et al., 2005), the Forecasting Ocean Assimilation Model (FOAM; Martin et al., 2007) and the Climate Forecast System (CFS) reanalysis (Saha et al., 2010b).

A basic assumption in specifying B_3 is that the covariances between errors should have characteristics similar to the ocean features themselves. Vertical structures of corrections should be similar to the vertical structures of eddies; this vertical structure being critical for operational ocean prediction. The basic dynamics of mesoscale eddies are results of thermocline displacement, which change across the globe and seasonally. With historical expendable bathythermograph, CTD, and ARGO data, sufficient observational information to specify the vertical covariance structure of temperature and salinity from observations is becoming available.

Specifying the Background Covariance

An example of the vertical structure within C_{vb} of Eq. 7 computed from historical observations at one location (275°E, 24°N, which is in the Gulf of Mexico Loop Current just southwest of Florida) during February is shown in Fig. 18.1. The diagonal submatrices from lower-left to upper-right show the correlations over depth for temperature, salinity, and geopotential respectively. These

correlations are computed globally on a 1/2° grid monthly at 72 vertical levels. The data size is reduced by retaining only the 20 leading eigenmodes of the correlation matrices, which is a further simplifying assumption. The results in Fig. 18.1 show diagonal correlations that are slightly less than 1 due to this. The temperature and salinity correlations show distinct layers. Temperatures in depths less than 50 m are correlated to one another more strongly that to those greater than 50 m, while temperatures greater 50 m are correlated to one another more strongly than those less than 50 m. Salinity exhibits a similar pattern with the separation depth at about 150 m. We must relate the thermohaline variations to the pressure field given by the geopotential. Define the vector of temperature and salinity anomalies (deviations from the mean vertical structure during the month) as $X = \left[T_1' \cdots T_N' S_1' \cdots S_N' \right]^T$, where N is the number of vertical depths. The geopotential is a linear transformation of the temperature and salinity given by $\varphi = G\kappa X$. The operator κ provides specific volume anomaly (the operator may be linearized about the monthly mean), and the operator G provides an integral over pressure. The cross correlation of temperature, salinity, and geopotential is then computed as

$$C_{vb} = \begin{bmatrix} \langle XX^T \rangle & \langle XX^T \kappa^T G^T \rangle \\ \langle G\kappa X^T X \rangle & \langle G\kappa XX^T \kappa^T G^T \rangle \end{bmatrix} \tag{8}$$

One interesting aspect of the vertical structure in Fig. 18.1 can be seen by the correlation of surface T, S, or *G* to the values deeper in the water column. The correlation of *G* at the surface to the underlying T and S is much stronger than the correlation of surface T or S to the subsurface. This is a result of geopotential at the surface (which is SSH multiplied by the gravitational acceleration) being an integral of the underlying properties as well as the large vertical correlation scales of T and S over depth. The SSH is a vertical integral of the effects and is strongly related to T and S. This leads to satellite altimeter data being a very influential observation for thermocline variability.

As in the vertical structure, we commonly assume horizontal structures of the estimated corrections are similar to the ocean features themselves. The horizontal correlations $C_{vv'}^H$ in equation (7) are computed using a second order auto-regressive function (Gaspari and Cohn, 1999) with a length scale specified as a fraction of the Rossby radius of deformation (Chelton et al., 1998). An example of the temperature analysis increments at 190 m depth is shown in Fig. 18.2 (top figure). The lines of features are due to the satellite SSH observations, and the vertical covariance provides the relation to the 190 m depth. The horizontal correlation relates increments along the ground tracks to values away from the ground tracks. The length scale varies spatially and has an average value of 21 km in the example, which is smaller than the Rossby radius.

The flow-dependent correlation for each variable is specified as $C^{FDB} = (1+s_f)e^{-s_f}$, where $s_f = \delta p / dh$, δp is the difference in model forecast pressure between two observation locations and *dh* is the specified flow-dependence scale factor. This increases the correlations between points that have little pressure difference and decreases correlation across pressure gradients. The pressure field

is used in the flow-dependent correlation under the assumption that the flow is in geostrophic balance and directed along pressure surfaces due to mesoscale features. Obviously, for some processes such as tides, this assumption does not hold. However, spatial scales for tides are typically well-separated from the mesoscale in the open ocean. Thus, it is not expected to be a significant problem. The flow field itself could be used in the formulation, though some influence of the tidal signal would still contribute. An example of strong flow-dependence influence is provided in Fig. 18.2 (bottom figure). The flow-dependence remains a minor influence on the accuracy of ocean prediction as the major covariance deficiencies lie mainly in the amplitudes as will be shown (Jacobs et al., 2014a).

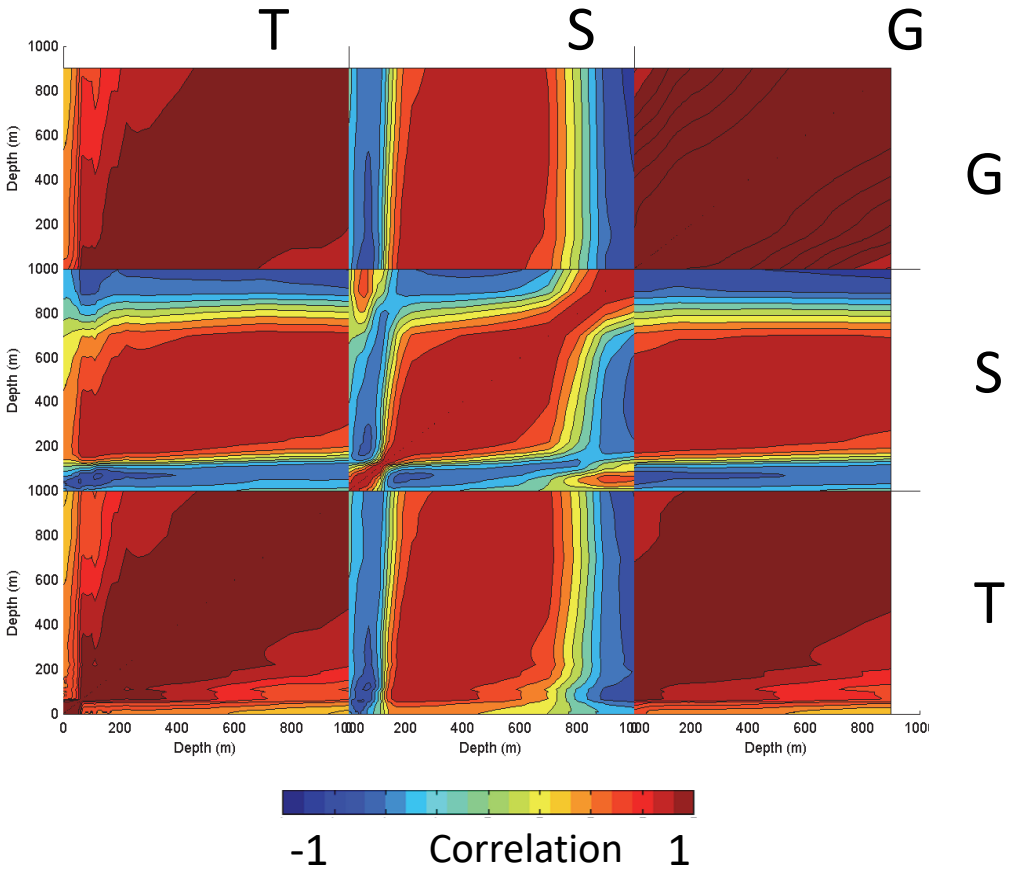

Figure 18.1. Cross correlations between temperature, salinity, and geopotential from the ocean surface to 1000m depth. This is based on historical data in February around the point 275°E, 24°N, which is in the Gulf of Mexico Loop Current just northeast of Cuba. The 20 most significant eigenmodes of the corelation matrix are retained to reduce data storage requirements of the global database, so the diagonal values (particularly in S-S correlation) are slightly less than 1. Distinct layers of correlation are apparent. The layers are different in temperature and salinity.

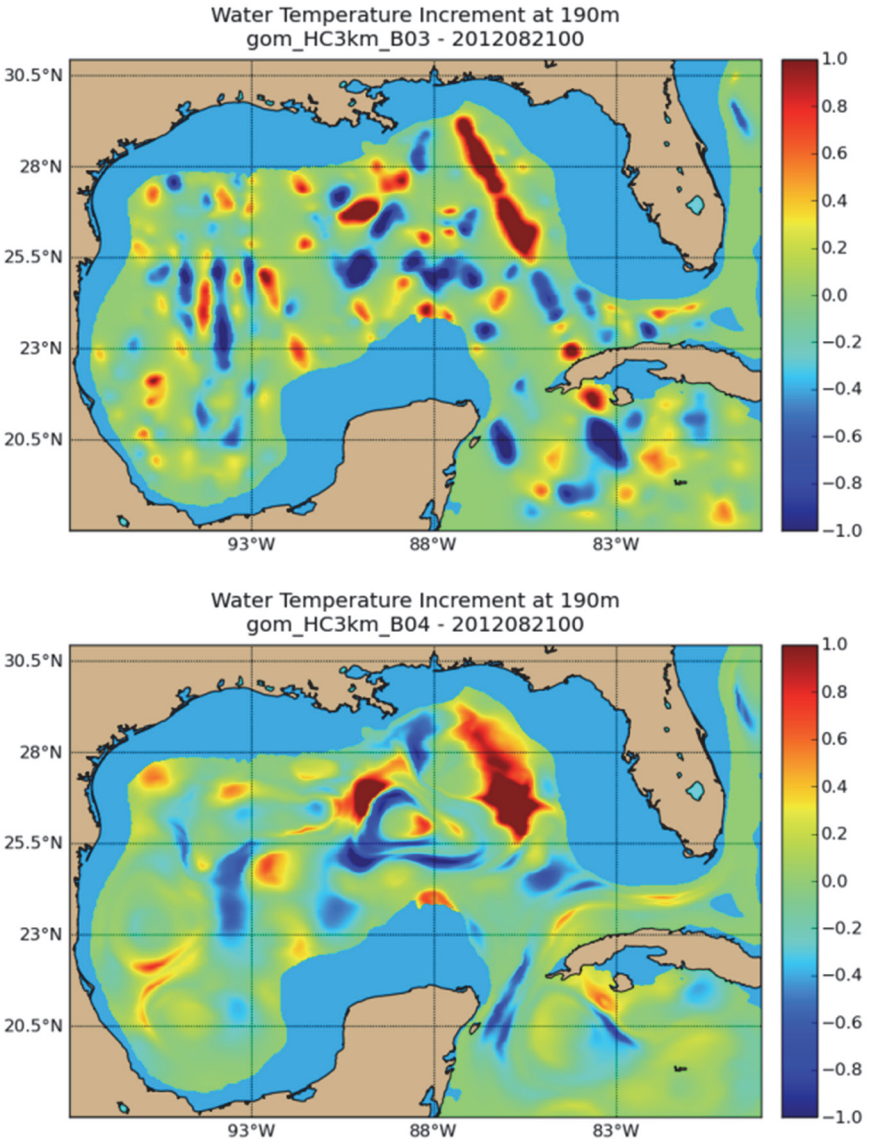

Figure 18.2. The analysis increments during one assimilation cycle of a 3DVar using (top) no flow dependence and (bottom) high flow dependence.

The background variance S_b is a diagonal matrix providing the amplitudes of the error variances, but is very difficult to compute. As discussed in the introduction, the 3DVar analysis should use the observations with background variance to compute the analysis covariance. The observations reduce the errors, and the amplitude of elements in the analysis covariance are smaller than those in background. The errors then increase in amplitude to provide the forecast error covariance that becomes the background error covariance for the next assimilation. For small dimensional problems to which Kalman filters are applied, propagating the analysis error covariance forward in time by applying the system dynamical operator and its adjoint is straight forward. However, an ocean prediction system dimension is far too large. Ensemble methods

provide a mechanism to accomplish this (Evensen, 2003). The example here uses a simplified approach. Given a first estimate of the background variance and the observations, the expected analysis error variance S_a is:

$$S_a = \left(I - BH^T \left(HB_3 H^T + R \right)^{-1} H \right) S_b \tag{9}$$

where I is the identity matrix, B_3 and R are as defined in Eq. 6, and only the diagonal values are computed in S_a with off-diagonal values being zero. The forecast variance may be estimated by

$$S_f = \left(1 - 1/\gamma^c \right) S_a + 1/\gamma^c \left(S_c - S_a \right) \tag{10}$$

which increases forecast error variance S_f toward the climatological variability S_c provided by the Generalized Digital Environmental Model (GDEM) database (Carnes et al., 2010) with $\gamma^c = 20$ days. This value generally provides an upper limit on variance amplitude in the thermocline in the absence of observations. In practice, the forecast variance amplitude shows the desired behaviors: observations reduce the background errors, and the errors increase to the upper limit of climatologically observed variance if no observations occur over several weeks. One example of the background error variance is provided in Fig. 18.3. The results are the forecast error standard deviation after the analysis of each day from Aug 21 to Aug 24 of 2012. As new data are received each day, the analysis error decreases according to Eq. 9. On subsequent days, the forecast errors increase in previously observed locations according to Eq. 10.

It is possible to test the performance of the estimated standard deviation $S_b^{1/2}$ by comparing the observation minus background difference over many assimilation cycles to build statistical estimates of the observed standard deviation. Two methods are tested with results shown in Fig. 18.4. The first method (referred to as R1) uses the system variability from one analysis cycle to another to estimate the standard deviation. This provides areas with high variability with higher expected errors. However, because ocean timescales are long, the estimate is too low. The second method (referred to as R2) uses the estimation given by Eqs. 9 and 10. The results over a three-month experiment in the Gulf of Mexico in Fig. 18.4 indicate the forecast error estimate of R2 is much more consistent with the root mean square (RMS) of observation minus background. In addition, the RMS of observation minus background is lower for R2 than for R1 indicating that the background field has lower errors relative to observations. As the background error covariance is a critical part of the data assimilation and thus prediction system, verification of the accuracy as in Fig. 18.4 is a very important part of the process of overall system validation for operational application.

Finally, we must specify how velocities are related to other variables. This is based on the correlation between geopotential and velocity given the horizontal $C_{vv'}^H$ spatial structure as determined by a second order auto-regressive function. The derivative of this spatial function provides a geostrophic velocity balancing geopotential height (Daley, 1993).

All these details in the background error covariance are typical from the beginning of mesoscale ocean prediction. Over time, a range of new assimilation considerations and observations have made

advancements incrementally. However, the results are still constrained by assumptions and limitations that reduced the 4DVar of Eq. 5 to the 3DVar of Eq. 6.

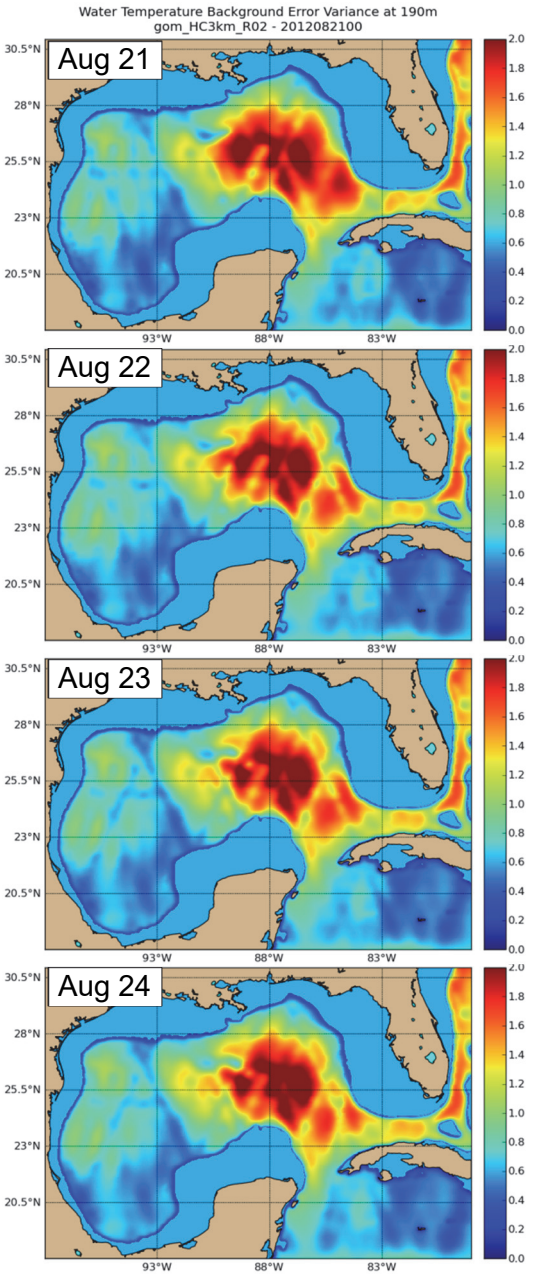

Figure 18.3. Temperature error standard deviation at 190 m temperature on August 21-24, 2012 (top to bottom rows).

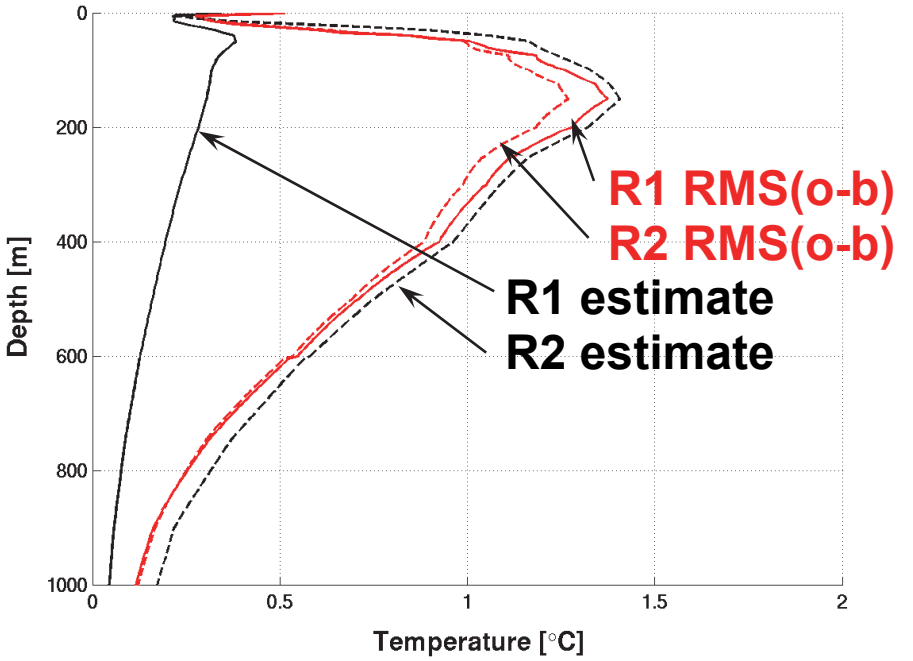

Figure 18.4. The estimated (black lines) temperature standard deviation error over depth from two methods (solid versus dashed lines) is compared to differences in the observations minus background (red lines). The background error estimate using analysis reduction in variance and growth to climatological variance (R2 black line) indicates a closer agreement with the RMS of observations minus the background state (red lines) than the methodology using only the temporal variability of the background state (R1 black line). In addition, the RMS difference between observations and background is smaller for R1 than R2 (red lines) indicating the background is more accurate.

Time Correlation

The temporal correlation of errors is a significant issue. In simplifying the 4DVar to a 3DVar, we assumed errors exist only at the analysis time, errors between analysis times are uncorrelated, observations exist only at the analysis time, and that the dynamical equations have negligible errors. The 3DVar analysis in Eq. 6 computes analysis increments at sequential analysis times independently. Errors in the assumption that observations be at the analysis time are reduced by using the First Guess at Appropriate Time (FGAT) in which the observation operator acts on the background field at the observation time [J A Cummings, 2005]. If the model has skill at predicting deterministic high frequency variability such as wind-driven events or internal tides, the FGAT approach reduces the high frequency signal in the data and the 3DVar corrects the slowly evolving field. However, the temporal influence of the data remains limited.

From the 4DVar perspective in Eq. 5, assuming errors between different analysis times are uncorrelated implies the temporal correlation is a set of Dirac delta functions. Impulsive forcing is not physically realistic. If the initial condition is reset at the analysis time, transients such as inertial oscillations develop even though the analysis may be geostrophically balanced. The source is the ageostrophic flow present in a model system as well as nonlinear advective processes that are not

geostrophic. Often, rather than resetting the initial condition at the analysis time, the analysis increment is inserted over a time interval. The total increment is divided by the time interval, and at every time step the model state is adjusted. This significantly reduces initial transients. The insertion also treats the corrections as though there is a long decorrelation time rather than an impulsive forcing at a single time. The time correlation function when inserting the increment is a boxcar function. The insertion approach recognizes some time correlations in errors, though errors from one analysis time to another remain uncorrelated.

Timescales in the deep ocean are long. Motivated by the assumption that errors have characteristics similar to the ocean mesoscale processes, the errors and subsequent analysis increments should also have long timescales. The improved forecast skill when moving from a time correlation of a delta function to a boxcar function is consistent with errors having long timescales. The assumptions leading to the 3DVar preclude observations from making corrections beyond the system assimilation interval. If the assimilation interval is one day, observations cannot influence subsequent days.

Attempts to take the time correlation into account mainly consist of including observations over a data window T_{obs} that is long. This consideration appears in many GODAE systems. MERCATOR assimilates altimeter data covering the prior seven days in an assimilation cycle that occurs every seven days (Brasseur et al., 2005); FOAM has a daily assimilation cycle with data used over multiple cycles and an error variance increasing linearly with data age (Martin et al., 2007); the Bluelink Ocean Data Assimilation System (BODAS) uses an 11-day T_{obs} to assimilate observations (Oke et al., 2008); and CFS conducts a six-hour assimilation cycle using data in the prior ten days with a weighting based on data age (Saha et al., 2010a). The Global Ocean Forecast System uses a daily assimilation cycle with observations covering the past several days so that data are used more than once (Cummings et al., 2009). Formally, observations should be assimilated only once. Otherwise, observation error variances are not accurate. There are competing considerations when determining the time period between analysis cycles to correct the model. A long time between assimilation cycles allows more data to influence the initial condition, however this implies that data soon after the assimilation could improve the forecast. A one-day assimilation cycle does use newly acquired observations, but it does not allow a long time period correlation. These problems arise because of the assumptions simplifying 4DVar to 3DVar. Early in GODAE, the development of 3DVar was more mature, and infrastructure supported it. There has been strong motivation to enable observations to influence corrections over a long time period, and this motivates a 4DVar approach.

The infrastructure and capability within ocean data assimilation has advanced to the point where 4DVar assimilation is feasible at least in regional nested areas (Moore et al., 2011; Ngodock and Carrier, 2014; Powell et al., 2008). The 4DVar solution in Eq. 5 contains the ocean tangent linear operator M' and its adjoint M'^T. The adjoint provides a way to compute the derivative of the observation innovations (difference between the observation and background field) with respect to all variables over all time and space. We immediately know how to change the observation errors by a change to the state at any time, boundary condition, or any dynamical equation throughout the

solution period. Moving to 4DVar overcomes the inherent limitations of 3DVar. Observations can influence the analysis over long time periods.

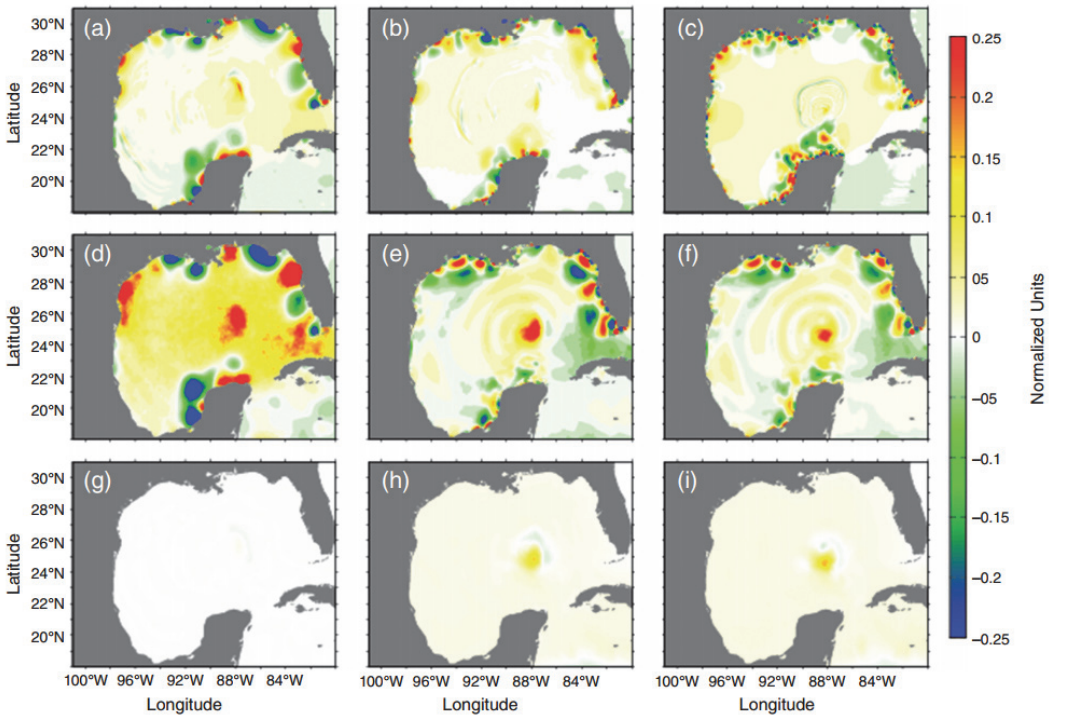

Figure 18.5. An example of gravity waves in the adjoint free-surface solution 96-, 48- and 3-h into the adjoint integration respectively (a–c), and the corresponding free surface for the forward solutions at the corresponding adjoint times, one forced with the adjoint solution (d–f) and the other without adjoint forcing of the free surface (g–i).

4DVar does not alleviate the need for dynamical consideration of the covariance matrix. The covariance matrix in the 4DVar of Eq. 5 contains not only the error covariance of the state at a specific time within 3DVar in Eq. 6 but also the error covariances between all the dynamical equations of the dynamical operator M'. The problem is larger, and the science communities must address the problem more closely. The consideration of errors in the dynamical operator requires further insight. One example is provided by considering the dynamical errors associated with SSH observations in the 4DVar (Ngodock et al., 2016). Consider the difference between observed and background SSH. A number of physical mechanisms may result in the difference. A mesoscale structure could produce a temperature and salinity anomaly deep within the thermocline that changes the specific volume and thus SSH. Alternatively, a barotropic wave generated at a particular location and time (e.g., a tsunami) may arrive at exactly the observation location to produce the observed SSH discrepancy. A sudden wind stress or any ocean process that may change SSH is an equally feasible source of error. The dynamic operator and its adjoint in the 4DVar do not differentiate between these solutions.

An example of this problem is shown in Fig. 18.5. The result of the operator $M'B_4M'^TH^T$ for one observation of SSH in the central Gulf of Mexico is computed by first applying the adjoint

M'^{T} to the observation operator and integrating backward in time (Fig. 18.5, top row). The error covariance and tangent linear model operators are then applied sequentially. The result (Fig. 18.5, middle row) shows the impact that one SSH observation would have on the analysis at the observation time under the assumptions contained in the covariance matrix B_4. The resulting analysis appears outside our normal expectations. The corrections 96 hours prior to the observed SSH (middle row left) indicate features around the coastlines, and if we intend to correct mesoscale features, these are not consistent. However, the analysis is physically correct. Stating that the result is outside our expectations implies that it is statistically anomalous, which implies a problem in the covariance that embodies our expectations of errors.

The covariances in Eq. 5 include the error in the barotropic equations used in the numerical model, which express the time rate of change of the vertically integrated velocities and SSH. The horizontal divergence integrated over the water column is the time rate of change of SSH. The results above included errors to the vertical integration of continuity, even though there is little error expected in this equation. This equation is a result of applying the conservation of mass property. Ocean models are typically constructed in a flux conservative form so that properties are conserved. The experience of physicists and oceanographers is that there is no error to the conservation of mass equation. Thus, the covariance matrix components describing these errors should be set to zero, and the minimization should have Lagrangian multipliers applied. When this is done, SSH observation stimulation of barotropic waves is greatly reduced (Fig. 18.8, bottom row). The data assimilation community has led the development of the 4DVar optimization approach. The expectation of where errors lie must be included in the error covariances of the 4DVar, and this requires the insight and experience of oceanographers. Expanding application of 4DVar to alleviate the time correlation issues inherent in 3DVar remains a challenge. Experiments conducted with 4DVar indicate good advancement beyond the 3DVar, and some of these are examined later on.

A Recent Example and Implications

The Lagrangian Submesoscale Experiment (LASER) as part of the Consortium on Advanced Research for Transport of Hydrocarbons in the Environment (CARTHE) released over 1,000 surface drifters starting January 2016 in the northeastern Gulf of Mexico. The objective was understanding submesoscale features and sharp horizontal buoyancy gradients. The persistence of drifters allowed observation of a wide range of additional features including the Loop Current Eddy (LCE) and associated cyclonic eddies. To highlight the types of features that are not well-predicted, we examine the details of when the assimilative forecast system diverges significantly from the observed drift trajectories. These are important in helping to understand the limits of all the components of the prediction system including the assimilation and its simplifying assumptions. The assimilation system is a 3DVar described by Eq. 6 with error covariances specified by the standard deviation, vertical correlations, and horizontal correlations described in the previous sections.

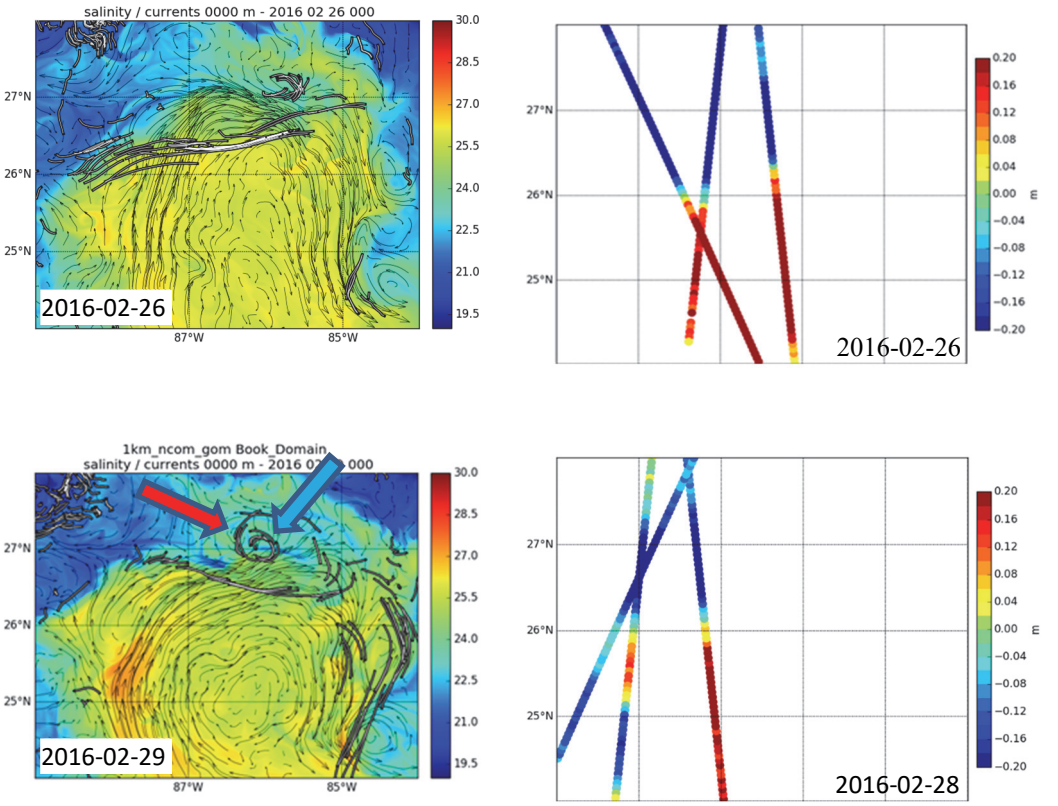

Figure 18.6. On the left are assimilative model results showing surface currents (vectors) and salinity (color) with observed drifter trajectories (white lines). On the right are satellite SSH anomaly observations assimilated into the model. The model shows a Loop Current Eddy (LCE) flow far to the north up to 27°N, which is not observed by the drifters. After assimilating the satellite observations, the position of the front is more correctly located. Also note the cyclonic eddy on the northern edge of the eddy at 27°N on 2016-02-29. The red arrow notes the position in the model, and the blue arrow notes the position observed by drifters.

Figs. 18.6 to18.8 display three events. In each, the LCE covers the lower portion of the domain. The size of this feature is about 300 km in diameter, which is in the targeted features of prediction for GODAE. The figures show surface currents and salinity from the system assimilating all the regular observations but not the LASER drifters. The general location of the LCE is consistent with the satellite SSH anomaly observations shown in the figures as well as the LASER drifters. The main discrepancies are in the position of the eddy front and cyclonic eddies on the periphery of the LCE.

Fig. 18. 6 shows an event starting February 26, 2016 when the model LCE extends up to 27°N to the north. The LASER drifters indicate the LCE front is around 26.3°N. The SSH anomaly observed by the altimeter data in subsequent days shows the LCE front much further to the south as well. After the satellite observations have been assimilated to correct the model, on February 29, 2016 the LCE front is much closer to the observed location February 29, 2016. This is an error in frontal position on the order of 50 to 100 km. At the end time of this event, also note the cyclonic

eddy to the north of the LCE front at about 27°N 86°W. After the assimilation has corrected the frontal position, the forecast model has a cyclonic feature near the cyclonic circulation observed by the LASER drifters. This particular cyclonic feature is the focus of the subsequent two events.

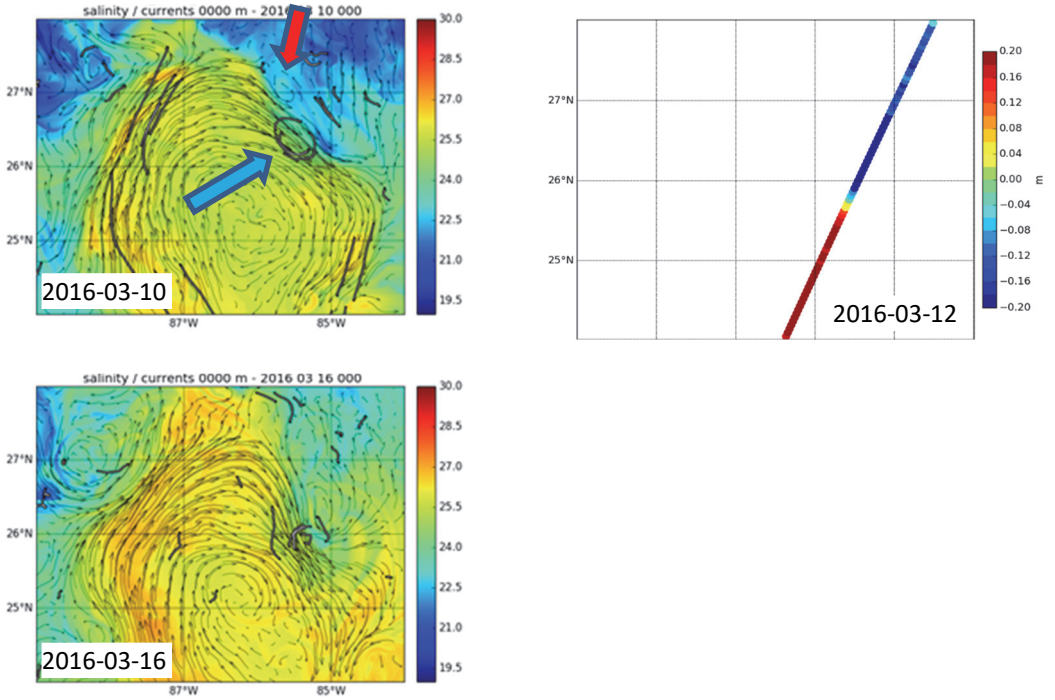

Figure 18.7. On the left are assimilative model results showing surface currents (vectors) and salinity (color) with observed drifter trajectories (white lines). On the right are satellite SSH anomaly observations assimilated into the model. At 2016-03-10, the drifters are entrained in the cyclonic eddy at 26.3°N 85.5°W, though the model indicates cyclonic circulation either at 27°N or at 85°W. The red arrow notes the position in the model, and the blue arrow notes the position observed by drifters. After assimilating the satellite data shown on the right, the model cyclonic circulation is similar to the drifter observed cyclone at 26°N on 2016-03-16.

The second event begins on March 10, 2016 in Fig. 18.7. Note the error in positioning of the cyclonic feature has grown significantly compared to the end time of the first event in Fig. 18.6. The model cyclonic feature is at 27°N while the drifter-observed feature is at 26.3°N. By the end time of the event on March 16, 2016, the model cyclonic feature is much closer to the observed. However, the shape is quite different. The most significant satellite SSH anomaly data available during this interval are shown in Fig. 18.7. This one satellite pass transects the cyclone, and these are the main data affecting the model cyclone position. On March 20, 2016, the model cyclone is at 26°N while the observed cyclone is at 25.3°N (Fig. 18 8). There are several satellite passes observing the area, and by March 27, 2016 the model cyclone position is much better aligned with the observations.

These errors are typical and to be expected. The main features on which GODAE systems have focused have been the larger mesoscale eddies. The time and spatial scales of the features are roughly consistent with the observation systems, from satellite to in situ. The data assimilation assumptions of horizontal and vertical correlation structures are consistent as well. The results from the LASER experiment highlight the edge of technical capability within present prediction systems.

The forecast skill is limited by the observation density as well as the prescribed assumptions within the data assimilation systems that have been built to be consistent with the expected observation density. Technology continues to move forward, and we must consider how the observing systems will evolve, how numerical models will evolve, and therefore what changes are necessary to the data assimilation to remain consistent with expected future development.

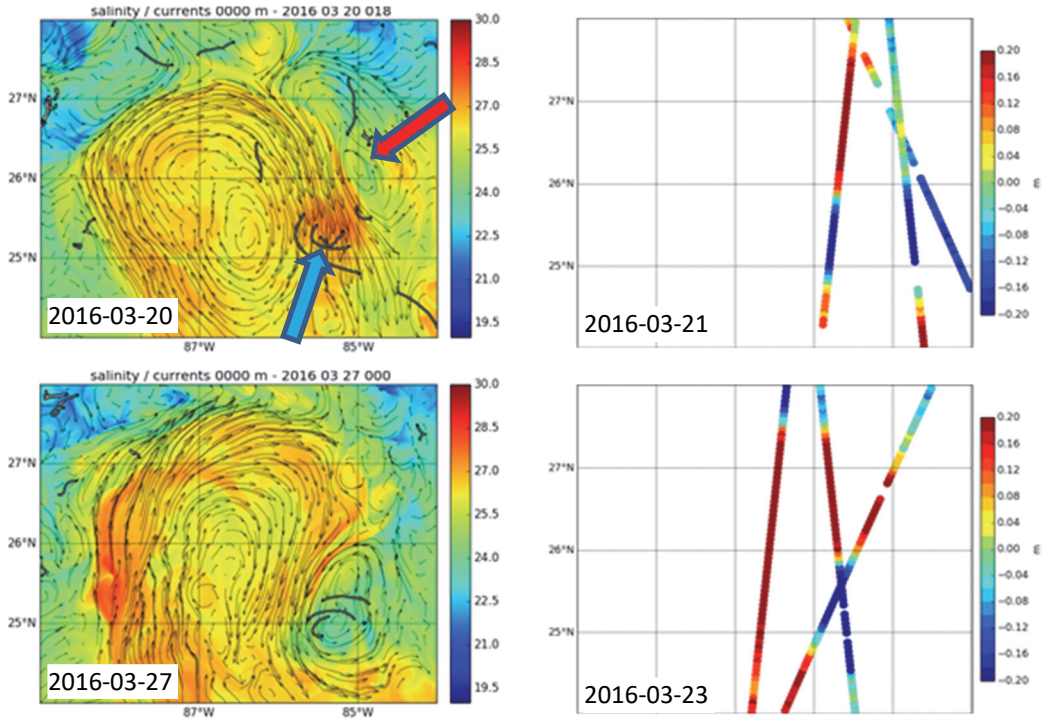

Figure 18.8. On the left are assimilative model results showing surface currents (vectors) and salinity (color) with observed drifter trajectories (white lines). On the right are satellite SSH anomaly observations assimilated into the model. At 2016-03-20, the drifters are entrained in the cyclonic eddy at 25.3°N 85.5°W, though the model indicates cyclonic circulation is at 26°N 85°W. The red arrow notes the position in the model, and the blue arrow notes the position observed by drifters. After assimilation of the satellite data shown on the right, the model cyclonic circulation is similar to the drifter observed cyclone at 25°N on 2016-03-27.

Future Advancements

From the 3DVar perspective of forecasting mesoscale features, the primary model state errors to be corrected are at the analysis time. The 4DVar approach is demonstrated to be superior to the 3DVar (Powell et al., 2008; Smith et al., 2017; A Weaver et al., 2003), and the errors are corrected in both the initial conditions and the state trajectory over time. In addition, the ability to advance The largest impediment to advancing operational ocean prediction is the density of observations. It is important to identify the metric by which forecasts are evaluated as we move forward. Many examinations quantify the error in the SSH of a forecast system. While this was very useful for initially gauging skill, it is not a strongly sensitive measure of accuracy. An example in Fig. 18.9 shows the skill of several ocean variables as a function of the number of altimeter data streams (Jacobs et al., 2014b).

The data were constructed from a series of observation system experiments in which all data were assimilated into one system, which is referred to as the nature run. This is the best solution possible. Subsequent experiments use the same boundary conditions and forcing but varying numbers of altimeter satellite data streams from TOPEX/Poseidon, Jason-1, GFO, and ENVISAT over a 1.5-year period. The initial condition at the beginning of the period for the observation system experiments is different from the nature run to ensure nondeterministic features deviate from the nature run. The spatial anomaly correlation between the observation system experiments and the nature run is averaged over the 1.5 years to produce the results plotted in Fig. 18.9. Anomaly correlations must be 0.6 or greater to achieve skillful prediction. The correlation of steric height relative to 1000 m (a proxy for SSH) with one altimeter is about 0.85 and asymptotically moves toward 1.0 with additional data. This reflects that the observation system experiments have the main mesoscale features in roughly the correct position. The steric height correlation is not greatly sensitive to small errors in mesoscale feature position. Other variables such as the surface divergence or the frontogenesis forcing, Q_1 (Hoskins, 1982; Pollard and Regier, 1992), linearly increase from about 0.25 with one data stream to just under 0.6 with four data streams. The frontogenesis is driven by confluence of water masses, which occurs along the fronts of eddies, and the surface divergence is controlled by this frontal process. Thus, these two variables measure the ability to accurately predict frontal positions. Typical frontal position errors are on the order of 50 -100 km at present, as we have seen in the LASER results (Figs. 18.6 to 18.8).

As observation density increases, the processes associated with fronts become predictable. The present regular observing systems such as altimeter satellites, ARGO, and others are not sufficient to resolve the positions of fronts. There are two approaches to increase the observation density, either by new in situ dense observations or new remote sensing instruments. An example of in situ dense observations is provided by the Grand Lagrangian Deployment (GLAD) during which 200 CODE-like drifters were deployed in the northeastern Gulf of Mexico in July 2012 (Ozgokmen et al., 2013). The forecast of Lagrangian trajectories is one of the most difficult problems we face. Errors grow exponentially in time. Small displacements of ocean features cause trajectory forecasts to rapidly diverge from observations. A pair of assimilation experiments is conducted, in which the first experiment uses the standard 3DVar approach for all regular observations and the second experiment uses the 4DVar approach and infers the velocity from drifter trajectories as observations in the assimilation. Forecast trajectories are then compared to the observed trajectories (Fig. 18.10). The general area of the LCE (around 90°-88°W, 25°-27°N) is similar in both results, but the shape and, therefore, the frontal location is quite different. The 4DVar experiment assimilating the inferred velocity shows forecast trajectories that are much more accurate than the 3DVar that is not assimilating the trajectories

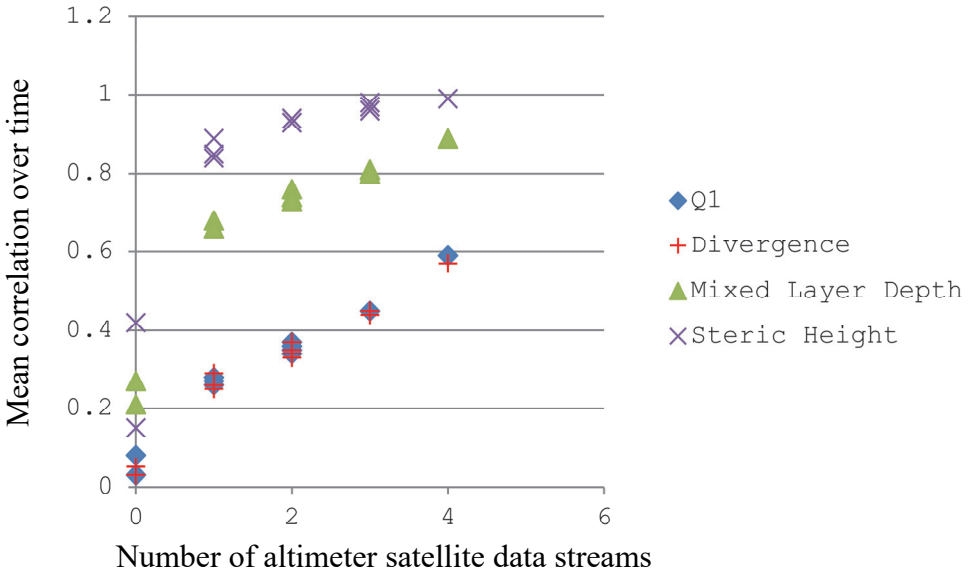

Figure 18.9. The time-average correlation as a function of number of satellite altimeters assimilated. The steric height correlation increases rapidly with just one data stream, and the marginal improvement of additional altimeters is small and decreases with increasing altimeters. The mixed layer depth marginal improvement is relatively constant from 1 to 4 altimeters. Frontogenesis Q_1 and surface divergence (vertical velocity) show a nearly linear increase in correlation as the number of altimeters assimilated increases.

The drifters prove to be a powerful observing system. The sensors are relatively simple and inexpensive. Large quantities can be deployed. In mesoscale features, the drifters persist for weeks while providing data continually, which is contrasted to satellite observations which typically require more than a week to return to the same area. If we assume the drifters are affected by geostrophic currents, an analogy is that the drifters are observing SSH horizontal gradients. Ageostrophic effects should be accounted for as representativeness errors, and these may be primarily a function of wind speed. Thus, the drifters provide information similar to altimeters in a selected area for a persistent period. This makes drifters an ideal observing system at the mesoscale for urgent events during which accurate forecasts become necessary, such as oil spills or search and rescue.

These high-density persistent observations affect considerations from the data assimilation perspective. The spatial scales require relating the observations across time as drifters move along fronts. Therefore, we must retain the time-correlated errors that are contained within the 4DVar in Eq. 5. The spatial scales presently used in the 3DVar are not consistent with the small scales of observations. Much work is under way to construct analysis increments that successively correct the larger-scale and the smaller-scale features ((Brandt and Zaslavsky, 1997; Choi et al., 2008; Haley and Lermusiaux, 2010; Lermusiaux, 2002; (Li, McWilliams et al. 2015)). These approaches become necessary as the dynamical processes of smaller scales change from the larger scales. The dynamical change across scales implies that a change in error covariances is required as well. As we consider new observing systems, the data assimilation of operational prediction systems must advance.

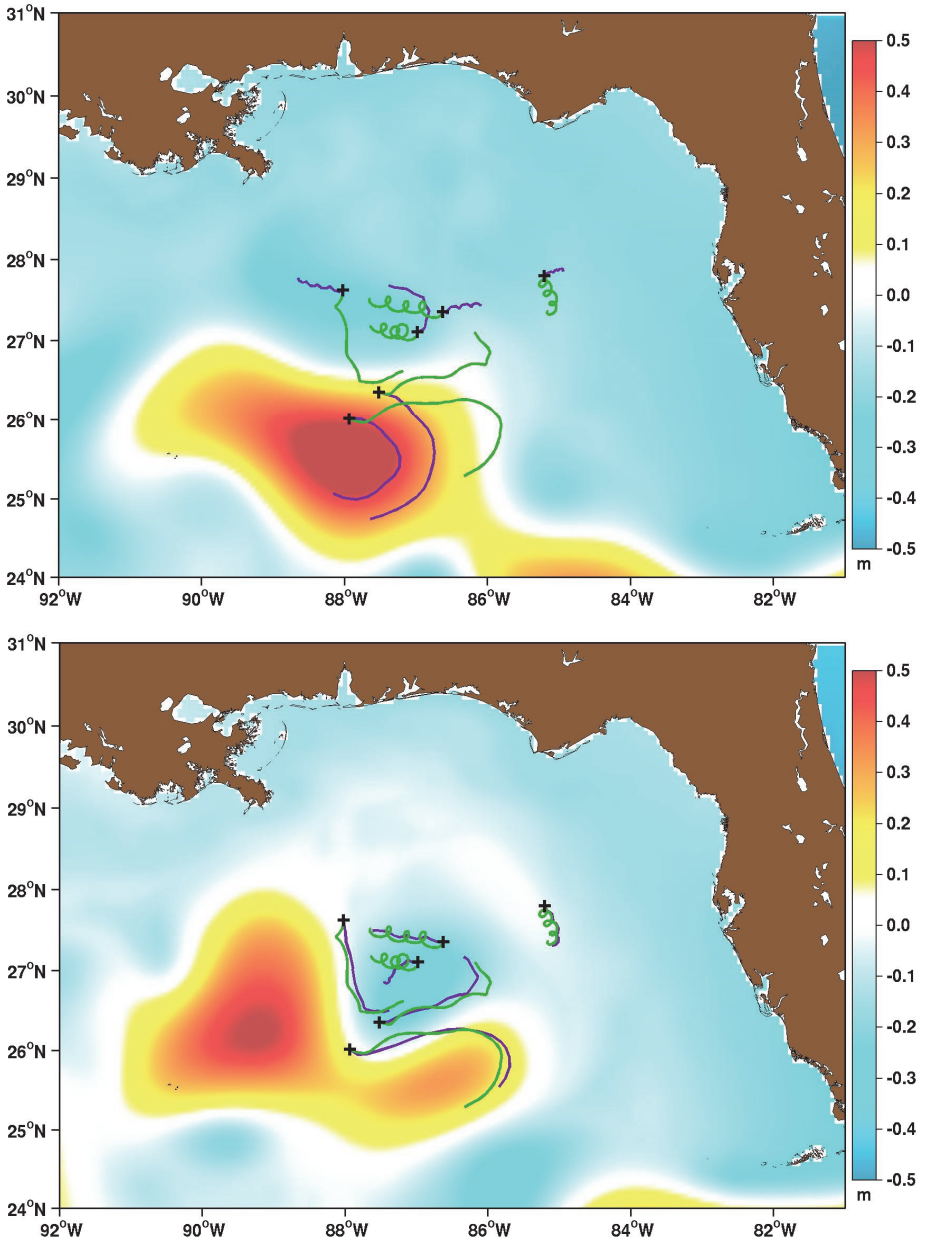

Figure 18.10. Two experiments are conducted during the July 2012 time period when 200 drifters were released in the northeast Gulf of Mexico. The 3DVar experiment (top) does not assimilate the drifters while the 4DVar experiment (bottom) does assimilate the drifters. The examples are shown one month after the assimilation experiments begin. Model forecast trajectories (purple lines) are compared to the observed trajectories (green lines). The background color is the model SSH.

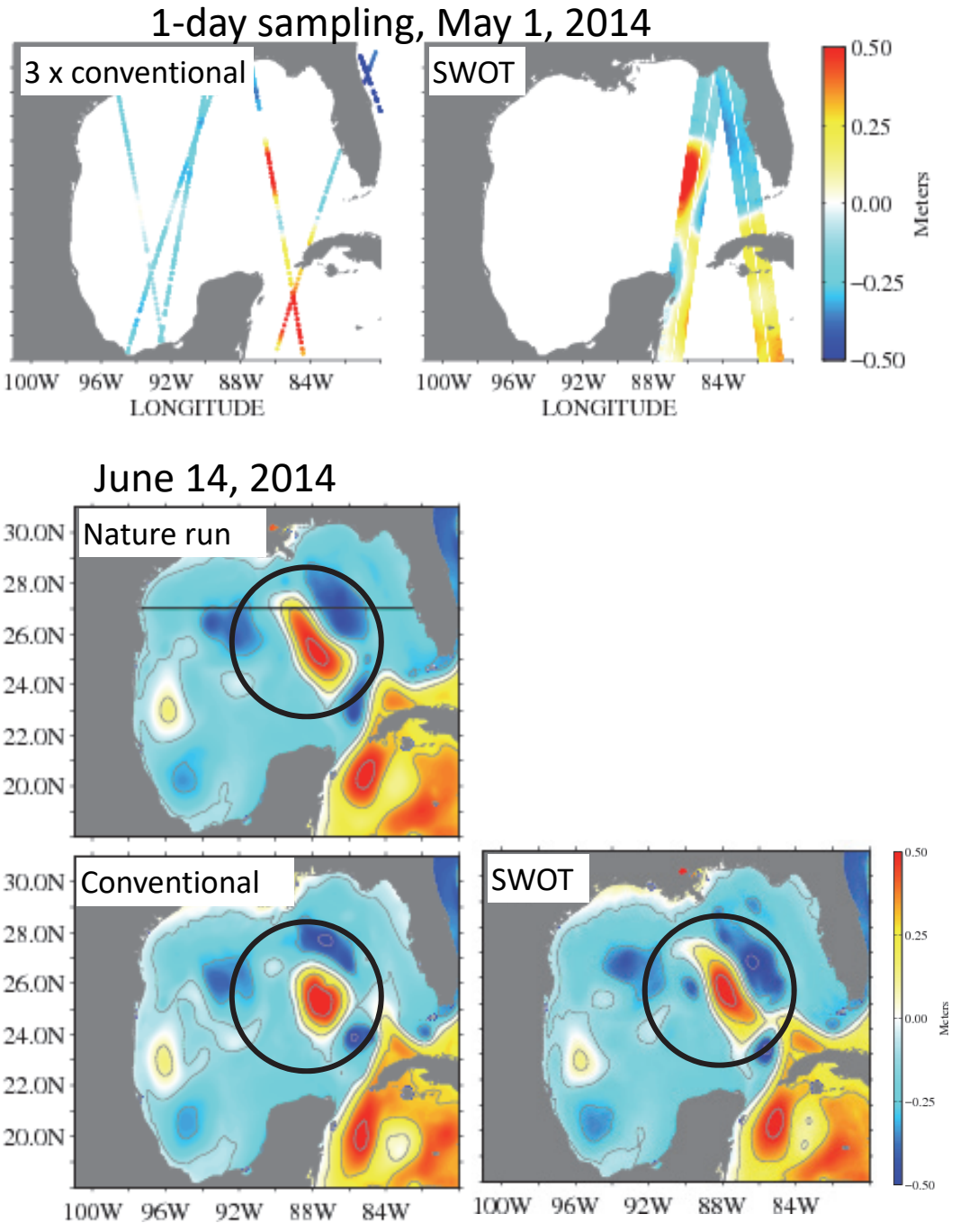

Figure 18.11. Observation System Simulation Experiments (OSSEs) simulate three conventional altimeter satellites and one Surface Water Ocean Topography (SWOT) satellite in terms of predicing the SSH of the LCE. The top row shows one day of sampling from the two OSSEs near the beginning of the experiment. The middle row shows the nature run SSH on June 14 with the LCE circled. The bottom row is the SSH from the conventional (left) and SWOT (right) simulations with the circle in the same location.

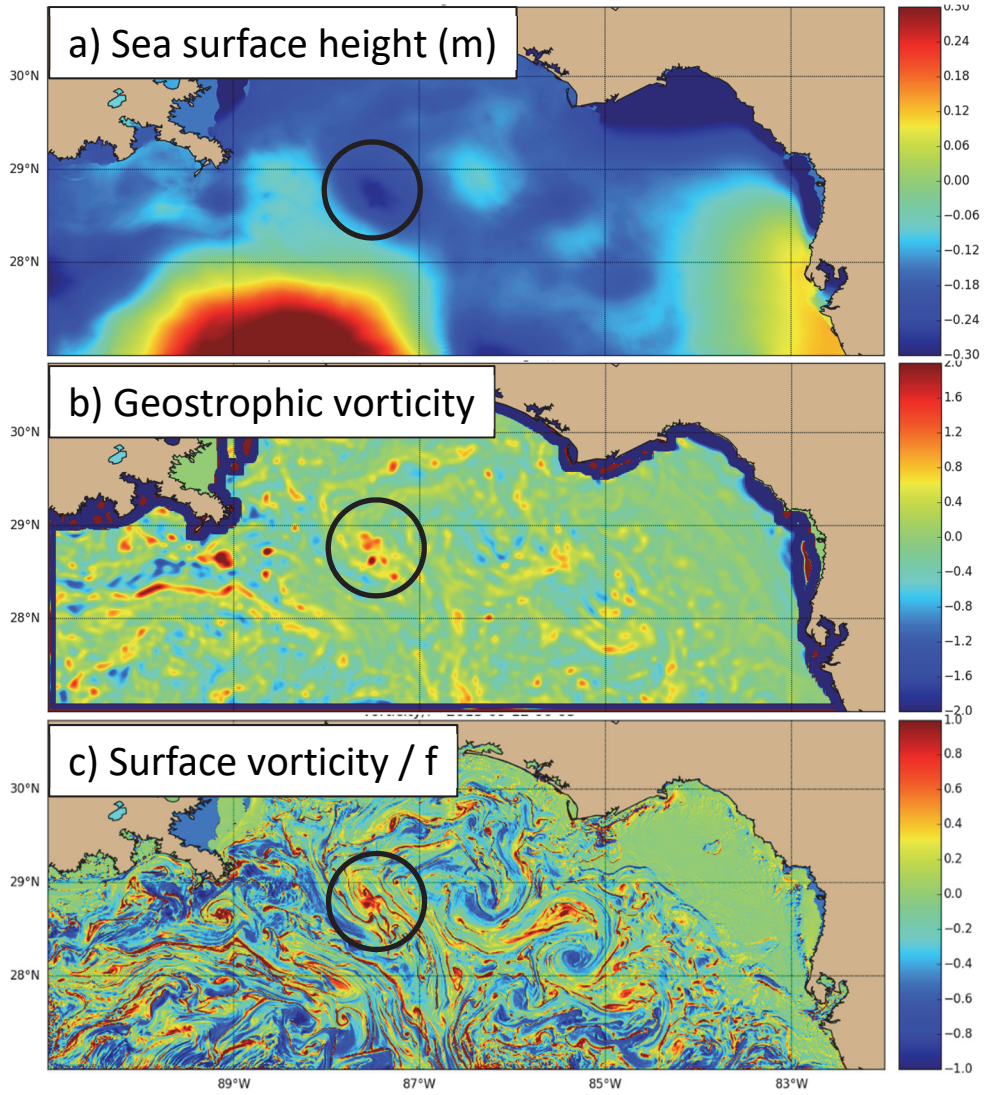

Figure 18.12. A numerical model shows the submesoscale difference from mesoscale. The SSH (top) is dominated by large amplitdue large horizontal scale features. The geostrophic vorticity computed from the SSH and normalized by the Coriollis parameter (middle) indicates small scale features with order one vorticity. The vorticity from the model normalized by the Coriollis parameter (bottom) contains these features, as well as the latteral shear associated with fronts.

A significant advancement in satellite observations will be provided by the Surface Water Ocean Topography (SWOT) mission (Durand et al., 2010). This instrument is an interferometric synthetic aperture radar (INSAR) that measures SSH across a 120 km wide swath (with a gap at the nadir). The ocean data are expected to be about 1 km resolution in the directions both along and across the ground track. Two Observation System Simulation Experiments (OSSEs) were conducted to gauge the influence of SWOT on assimilation and model skill. One experiment was run without any assimilation to provide the nature run that was sampled for the two OSSEs. One sampling scheme is provided by using the ground tracks of three conventional altimeters (Jason-1, AltiKa, Jason-2 in

the interleaved orbit). The second sampling scheme uses the sampling of SWOT by linking the nature run to the SWOT simulator (Gaultier et al., 2015). The sampled data are assimilated into two separate model runs starting from an initial condition that differs from the nature run, and the results are examined one month later (Fig. 18.11). The shape and frontal positions of the LCE in the SWOT OSSE are much closer to the nature run than the features in the conventional altimeter OSSE. We expect three conventional altimeters to be operating during the SWOT time period. Together, we can expect these systems to advance ocean predictive skill significantly.

SWOT also presents a possibility of predicting much smaller scales throughout the globe, which moves from mesoscale to submesoscale prediction. The dynamics change significantly and it requires reevaluation of the assimilation approach. We can examine the Rossby number based on the ratio of vorticity to the local Coriolis parameter. Mesoscale eddies have a Rossby number much less than one and are primarily geostrophically balanced. The main mesoscale variation in ocean structure is due to vertical movement of the thermocline in balance with the SSH (Fig. 18.12a). Spatial scales are on the order of 200 km in this area. The submesoscale has order one or larger Rossby number. The main vertical structure variability is in the mixed layer depth. The submesoscale deviates significantly from the observed historical vertical structure in Fig. 18.1, where the ocean acts mainly as a two-layer fluid with the interface at the thermocline depth. The submesoscale acts with an interface at the mixed layer depth.

Small scale features (Fig. 18.12b,c) are dominated by submesoscale eddies with spatial scales on the order of 10 km. The background error covariance formulated in the previous sections is targeted specifically at mesoscale features. The present operational data assimilation process is oriented to this underlying assumption. The vertical structure, horizontal structure, geostrophic balance, cycling frequency, and application of data all have underlying assumptions that the primary features are mesoscale. As we move to forecast submesoscale features, the dynamics change dramatically, and the features themselves are strongly advected by the mesoscale flow.

Addressing these requires a multiscale approach in which the analysis of the two dynamical regimes is conducted sequentially (Li, McWilliams et al. 2015). A correction is first computed for the mesoscale using appropriate error covariance, and this correction is applied to ensure the mesoscale is as accurate as possible as it will advect the submesoscale field. A second correction is computed for the submesoscale. The error covariances for the submesoscale must be different from the mesoscale, but it would be convenient if the infrastructure for both analyses were as reusable as possible. This leads to a generalization of the assimilation problem.

As discussed earlier, the analysis increment is not constructed by explicitly computing a matrix inverse. Iterative techniques, such as conjugate gradient descent with preconditioners, are often used. In this case, all we require is the action of matrices on vectors. The effects of the covariance operating on state vector can also be applied through means polynomials of the diffusion operators have been routinely used to provide the action (Carrier and Ngodock, 2010; Derber and Rosati, 1989; Weaver and Mirouze, 2013; Yaremchuk et al., 2013). Dynamical balance operators may be generalized similarly. There is an operation that provides density from temperature and salinity by equation of state. Integration of the continuity equation relates density to SSH. Hydrostatic balance

provides pressure from density and SSH, and finally the geostrophic relation provides velocity from the pressure field. The action of these operators maps from one variable to another by the specified dynamics, and the covariance matrix may be specified in terms of these operators (Yaremchuk et al., 2017; Yaremchuk and Martin, 2016). For example, the sequence of computations described above relates temperature and salinity to SSH and velocity by a mesoscale operator:

$$\begin{bmatrix} \zeta \\ \vec{u} \end{bmatrix} = \mathbf{L_M} \begin{bmatrix} T \\ S \end{bmatrix} \tag{11}$$

The full covariance matrix may be provided by:

$$\left\langle \begin{bmatrix} T \\ S \\ \zeta \\ \vec{u} \end{bmatrix} \begin{bmatrix} T \\ S \\ \zeta \\ \vec{u} \end{bmatrix}^T \right\rangle = \begin{bmatrix} B_{TS} & B_{TS}\mathbf{L_M}^{\mathbf{T}} \\ \mathbf{L_M}B_{TS} & \mathbf{L_M}B_{TS}\mathbf{L_M}^{\mathbf{T}} \end{bmatrix} \tag{12}$$

where B_{TS} is the cross-covariance of temperature and salinity as provided in Fig. 18.1. Note the similarity between Eqs. 8 and 12. This generalizes the assimilation problem so that it may be applied to the submesoscale. The only required change is replacing the mesoscale dynamical operator $\mathbf{L_M}$ with an appropriate dynamical operator for the submesoscale $\mathbf{L_S}$. The challenge becomes to provide an operator that appropriately relates ocean properties for the submesoscale. The same solution process may then be used in a sequential multiscale analysis to predict ocean submesoscale features.

Conclusions

GODAE has led to operational data assimilation systems that produce regular forecasts in many centers across the globe. The implementations have primarily been based on 3DVar analysis, and by deriving this from a 4DVar analysis we can explicitly understand the inherent assumptions and impact on the forecast skill. Many of the assumptions were appropriate for the observing systems and numerical model predictions. New observing systems can provide high density observations, both in local settings and globally. GODAE systems predict the general locations of mesoscale eddy features. However, the errors in the frontal positions are on the order of 50 to 100 km. At these scales and smaller, dynamics become more ageostrophic. The continued development of data assimilation approaches promises to extend existing observations and exploit new ones. We expect the present errors in frontal positions to decrease dramatically. The movements toward multiscale analysis and relieving the 3DVar assumptions are required to advance. The frontal positions are critical aspects to many operational problems for fisheries, search and rescue, oil drilling, and aquaculture. With the progression of computational capability in operational centers along with numerical model representation of physics and observing systems, we have the opportunity to extend operational data assimilation to meet many new challenges and applications.

References

Anderson, J. L. (2012), Localization and sampling error correction in ensemble Kalman filter data assimilation, Mon Weather Rev, 140(7), 2359-2371.

Bennett, A. F. (2002), Inverse modeling of the ocean and atmosphere, Cambridge University Press.

Bocquet, M., C. A. Pires, and L. Wu (2010), Beyond Gaussian statistical modeling in geophysical data assimilation, Mon Weather Rev, 138(8), 2997-3023.

Brandt, A., and L. Y. Zaslavsky (1997), Multiscale algorithm for atmospheric data assimilation, Siam J Sci Comput, 18(3), 949-956.

Brasseur, P., et al. (2005), Data assimilation for marine monitoring and prediction: The MERCATOR operational assimilation systems and the MERSEA developments, Q J Roy Meteor Soc, 131(613), 3561-3582.

Carnes, M. R., R. W. Helber, C. N. Barron, and J. M. Dastugue (2010), Validation test report for GDEM4Rep., DTIC Document.

Carrier, M. J., and H. Ngodock (2010), Background-error correlation model based on the implicit solution of a diffusion equation, Ocean Model, 35(1-2), 45-53.

Charney, J. G. (1947), The dynamics of long waves in a baroclinic westerly current, Journal of Meteorology, 4(5), 136-162.

Chelton, D. B., R. A. Deszoeke, M. G. Schlax, K. El Naggar, and N. Siwertz (1998), Geographical variability of the first baroclinic Rossby radius of deformation, J Phys Oceanogr, 28(3), 433-460.

Choi, M. J., V. Chandrasekaran, D. M. Malioutov, J. K. Johnson, and A. S. Willsky (2008), Multiscale stochastic modeling for tractable inference and data assimilation, Comput Method Appl M, 197(43-44), 3492-3515.

Cummings, J., et al. (2009), Ocean Data Assimilation Systems for Godae, Oceanography, 22(3), 96-109.

Cummings, J. A. (2005), Operational multivariate ocean data assimilation, Q J Roy Meteor Soc, 131(613), 3583-3604.

Daley, R. (1993), Atmospheric data analysis, Cambridge university press.

Derber, J., and A. Rosati (1989), A global oceanic data assimilation system, J Phys Oceanogr, 19(9), 1333-1347.

Dombrowsky, E., L. Bertino, G. B. Brassington, E. P. Chassignet, F. Davidson, H. E. Hurlburt, M. Kamachi, T. Lee, M. J. Martin, and S. Mei (2009), GODAE systems in operation, Oceanography, 22(3), 80-95.

Durand, M., L.-L. Fu, D. P. Lettenmaier, D. E. Alsdorf, E. Rodriguez, and D. Esteban-Fernandez (2010), The surface water and ocean topography mission: Observing terrestrial surface water and oceanic submesoscale eddies, Proceedings of the IEEE, 98(5), 766-779.

Evensen, G. (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean dynamics, 53(4), 343-367.

Gaspari, G., and S. E. Cohn (1999), Construction of correlation functions in two and three dimensions, Q J Roy Meteor Soc, 125(554), 723-757.

Gaultier, L., C. Ubelmann, and L.-L. Fu (2015), SWOT Simulator DocumentationRep., Tech. Rep. 1.0. 0, Jet 422 Propulsion Laboratory, California Institute of Technology. 423.

Haley, P. J., and P. F. J. Lermusiaux (2010), Multiscale two-way embedding schemes for free-surface primitive equations in the "Multidisciplinary Simulation, Estimation and Assimilation System", Ocean Dynamics, 60(6), 1497-1537.

Helber, R. W., et al. (2013), Validation test report for the Improved Synthetic Ocean Profile (ISOP) system, Part I: Synthetic profile methods and algorithm, Naval Research Lab Stennis Detachment, Stennis Space Center, MS, Oceanography Div.

Hoskins, B. J. (1982), The Mathematical Theory of Frontogenesis, Annu Rev Fluid Mech, 14, 131-151.

IMAWAKI, S. (1981), Vertical Structure and Horizontal Scales of the Mesoscale Baroclinic Variability in the Western North Pacific.

Jacobs, G. A., B. P. Bartels, D. J. Bogucki, F. J. Beron-Vera, S. S. Chen, E. F. Coelho, M. Curcic, A. Griffa, M. Gough, and B. K. Haus (2014a), Data assimilation considerations for improved ocean predictability during the Gulf of Mexico Grand Lagrangian Deployment (GLAD), Ocean Model, 83, 98-117.

Jacobs, G. A., J. G. Richman, J. D. Doyle, P. L. Spence, B. P. Bartels, C. N. Barron, R. W. Helber, and F. L. Bub (2014b), Simulating conditional deterministic predictability within ocean frontogenesis, Ocean Model, 78, 1-16.

Kalnay, E. (2003), Atmospheric modeling, data assimilation, and predictability, Cambridge university press.

Lermusiaux, P. (2002), On the mapping of multivariate geophysical fields: sensitivities to size, scales, and dynamics, J Atmos Ocean Tech, 19(10), 1602-1637.

Le Traon, P.-Y., (1991), Time scales of mesoscale variability and their relationship with space scales in the North Atlantic, J Mar Res, 49(3), 467-492.

Le Traon, P.-Y., et al. (1999). Operational oceanography and prediction-a GODAE perspective, OceanObs' 09.

Li, Z., J. C. McWilliams, K. Ide, and J. D. Farrara (2015), A multiscale variational data assimilation scheme: formulation and illustration, Mon Weather Rev, 143(9), 3804-3822.

Lozano, C. J., A. R. Robinson, H. G. Arango, A. Gangopadhyay, Q. Sloan, P. J. Haley, L. Anderson, and W. Leslie (1996), An interdisciplinary ocean prediction system: Assimilation strategies ana structured data models, Elsevier Oceanography Series, 61, 413-452.

Martin, A. J., A. Hines, and M. J. Bell (2007), Data assimilation in the FOAM operational short-range ocean forecasting system: A description of the scheme and its impact, Q J Roy Meteor Soc, 133(625), 981-995.

Moore, A. M., H. G. Arango, G. Broquet, B. S. Powell, A. T. Weaver, and J. Zavala-Garay (2011), The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems Part I - System overview and formulation, Prog Oceanogr, 91(1), 34-49.

Ngodock, H., and M. Carrier (2014), A 4DVAR system for the Navy Coastal Ocean Model. Part I: System description and assimilation of synthetic observations in Monterey Bay, Mon Weather Rev, 142(6), 2085-2107.

Ngodock, H., M. Carrier, I. Souopgui, S. Smith, P. Martin, P. Muscarella, and G. Jacobs (2016), On the direct assimilation of along-track sea-surface height observations into a free-surface ocean model using a weak constraint four-dimensional variational (4D-Var) method, Q J Roy Meteor Soc, 142(695), 1160-1170.

Oke, P., A. Schiller, D. Griffin, and G. Brassington (2005), Ensemble data assimilation for an eddy-resolving ocean model of the Australian region, Q J Roy Meteor Soc, 131(613), 3301-3311.

Oke, P. R., G. B. Brassington, D. A. Griffin, and A. Schiller (2008), The Bluelink ocean data assimilation system (BODAS), Ocean Model, 21(1-2), 46-70.

Ozgokmen, T., A. Poje, B. Lipphardt Jr, A. Haza, B. Haus, G. Jacobs, A. Reniers, J. Olascoaga, E. Ryan, and G. Novelli (2013), Grand LAgrangian Deployment (GLAD): Surface Dispersion Characteristics Near the Deepwater Horizon Oil Spill Site, paper presented at EGU General Assembly Conference Abstracts.

Pollard, R., and L. Regier (1992), Vorticity and vertical circulation at an ocean front, J Phys Oceanogr, 22(6), 609-625.

Powell, B., H. Arango, A. Moore, E. Di Lorenzo, R. Milliff, and D. Foley (2008), 4DVAR data assimilation in the intra-Americas sea with the Regional Ocean Modeling System (ROMS), Ocean Model, 25(3), 173-188.

Richman, J. G., C. Wunsch, and N. G. Hogg (1977), Space and time scales of mesoscale motion in the western North Atlantic, Rev Geophys, 15(4), 385-420.

Saha, S., S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, and D. Behringer (2010a), The NCEP climate forecast system reanalysis, B Am Meteorol Soc, 91(8), 1015-1057.

Saha, S., et al. (2010b), The Ncep Climate Forecast System Reanalysis, B Am Meteorol Soc, 91(8), 1015-1057.

Smith, S., H. Ngodock, M. Carrier, J. Shriver, P. Muscarella, and I. Souopgui (2017), Validation and Operational Implementation of the Navy Coastal Ocean Model Four Dimensional Variational Data Assimilation System (NCOM 4DVAR) in the Okinawa Trough, in Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III), edited, pp. 405-427, Springer.

Testor, P., et al. (2010). Gliders as a component of future observing systems, OceanObs' 09.

Weaver, A., J. Vialard, and D. Anderson (2003), Three-and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal diagnostics, and consistency checks, Mon Weather Rev, 131(7), 1360-1378.

Weaver, A. T., and I. Mirouze (2013), On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation, Q J Roy Meteor Soc, 139(670), 242-260.

Yaremchuk, M., M. Carrier, S. Smith, and G. Jacobs (2013), Background error correlation modeling with diffusion operators, in Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II), edited, pp. 177-203, Springer.

Yaremchuk, M., P. Martin, G. Panteleev, C. Beattie, and A. Koch (2017), Adjoint-Free 4D Variational Data Assimilation into Regional Models, in Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III), edited, pp. 83-114, Springer.

Yaremchuk, M., and P. J. Martin (2016), Implementation of a balance operator in NCOMRep., Naval Research Lab Stennis Detachment, Stennis Space Center, MS, Ocean Dynamics and Prediction Branch.

Zhang, S., M. Harrison, A. Rosati, and A. Wittenberg (2007), System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies, Mon Weather Rev, 135(10), 3541-3564.