Continental Shelf Research 73 (2014) 41–48

Contents lists available at ScienceDirect

Continental Shelf Research

journal homepage: www.elsevier.com/locate/csr

Research papers

An examination of a multi-scale three-dimensional variational data assimilation scheme in the Kuroshio Extension using the naval coastal ocean model

Philip A. Muscarella [a,*], M.J. Carrier [b], H.E. Ngodock [b]

[a] ASEE Post doc, Naval Research Lab, Stennis Space Center, USA
[b] Ocean Data Assimilation & Probabilistic Prediction, Naval Research Laboratory, Stennis Space Center, USA

ARTICLE INFO

Article history:
Received 11 March 2013
Received in revised form
9 October 2013
Accepted 1 November 2013
Available online 25 November 2013

Keywords:
Data assimilation
Multi-scale
Kuroshio Extension
3DVAR

ABSTRACT

The implementation of a multi-scale three dimensional variational (MS3DVAR) data assimilation scheme for use with the Navy Coastal Ocean Model (NCOM) in the Kuroshio Extension western boundary current region is presented here. This work leverages on Li et al. (2013), who initially developed this method. MS3DVAR data assimilation allows for the effective assimilation of both spatially coarse and dense collections of observations. Traditional 3DVAR produces an inherent filtering of dynamical features smaller than the decorrelation length. The MS3DVAR allows for a scale selective background error covariance capable of handling a wider range of ocean scales. Here the MS3DVAR is examined in an energetic coastal regime using simulated and real observations. The results show that the MS3DVAR reduces analysis errors when compared to a traditional 3DVAR scheme. Forecast errors appear to be similar for both systems and are most likely due to the coarse resolution of the surface forceing being applied.

1. Introduction

Thanks to the ever increasing computer capabilities, numerical ocean models are now routinely run at higher and higher resolution, e.g. global hybrid coordinate ocean model (HYCOM) at 1/25° (Metzger et al., 2010; Hurlburt et al., 2011). The spatial resolution in regional models is usually higher than that of the global models providing initial and boundary conditions. When models are run at high resolution, they resolve features and processes (usually of small spatial and short temporal scales) that are absent at lower resolution. These small scale processes may be the realistic result of the dynamics and physics of the model equations, but they are not necessarily synoptic, and need to be validated against observations that are not always available, let alone at high resolution. The lack of high density observations, however, results in ocean analyses that are largely unconstrained and may lead to unsatisfactory forecasts of the ocean state.

And, even if high resolution observations were available, in addition to the sparse observations routinely collected, an assimilation system that is able to take into account the presence of multiple scales in both the model and the observations would still be needed. A natural choice for such an assimilation system would be an ensemble based algorithm, such as the EnKF (Evensen, 2003) in which the interaction between processes of different scales is implicitly accounted for in the ensemble generated background error covariance. However, with models running at high resolution, an ensemble approach would be very costly. An alternative was proposed by Li (2011) and Li et al. (2013) in the form of the multi-scale 3DVAR (MS3DVAR), where the model forecast, the observations and their respective error covariances are decomposed into large and small scale components, and the 3DVAR analysis is carried out for each component. MS3DVAR has been successfully used in assimilating observations from the Southern California Coastal Ocean Observing System (SCCOOS) (Li, 2011), and in a real time nowcast-forecast modeling system for Prince William Sound that was developed and operated in support of the summer 2009 Sound Predictions Field Experiment (Farrara et al., 2013; Li et al., 2012).

A similar algorithm was proposed by Xie et al. (2011) and used for hurricane data assimilation. Another atmospheric application was carried out by Xiao and Tastula (2011) for the problem of hurricane initialization.

The objectives of this study are, first of all to describe the MS3DVAR system being used with NCOM (Martin, 2000). A first order validation of this system using simulated and real observations is presented here as well. Comparison metrics will focus on the assimilation system's ability to improve the model analyses and forecasts. Oceanic applications of the MS3DVAR have been

* Corresponding author. Tel.: +1 228 688 4708.
E-mail address: philip.muscarella.ctr@nrlssc.navy.mil (P.A. Muscarella).

0278-4343/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.csr.2013.11.009

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2014		**00-00-2014 to 00-00-2014**

4. TITLE AND SUBTITLE
An examination of a multi-scale three-dimensional variational data assimilation scheme in the Kuroshio Extension using the naval coastal ocean model

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory,1 Stennis Space Ctr,Stennis Space Center,MS,39529

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

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

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The implementationofamulti-scalethreedimensionalvariational(MS3DVAR)dataassimilationscheme for usewiththeNavyCoastalOceanModel(NCOM)intheKuroshioExtensionwesternboundarycurrent region ispresentedhere.Thisworkleverageson Li etal.(2013), whoinitiallydevelopedthismethod. MS3DVARdataassimilationallowsfortheeffectiveassimilationofbothspatiallycoarseanddense collections ofobservations.Traditional3DVARproducesaninherent filtering ofdynamicalfeatures smaller thanthedecorrelationlength.TheMS3DVARallowsforascaleselectivebackgrounderror covariancecapableofhandlingawiderrangeofoceanscales.HeretheMS3DVARisexaminedinan energetic coastalregimeusingsimulatedandrealobservations.TheresultsshowthattheMS3DVAR reduces analysiserrorswhencomparedtoatraditional3DVARscheme.Forecasterrorsappeartobe similar forbothsystemsandaremostlikelyduetothecoarseresolutionofthesurfaceforceingbeing applied.

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			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**8**	

carried out mostly in small domains with weak to moderate flows. In this study, the MS3DVAR is applied to a domain containing the continental shelf, shelf break, deep ocean and strong western boundary flow. Of particular interest is the ability of the MS3DVAR to handle both the large scale circulation in the region as well as the small scale processes generated in the island wakes and their interaction and with the Kuroshio. These small scale features are present in most shelf regions around the world and a better representation of them in our models is an issue that is addressed with the MS3DVAR.

Background error covariances are essential to any assimilation algorithm. They describe the assumptions made about the errors in the model solution, which are poorly understood and are responsible for spreading the information from the observation locations to everywhere else in the model domain and to all model variables. Li et al. (2013) describes a method for the construction of background error covariances for a MS3DVAR based on the Kronecker product formulation using perturbations derived from an ensemble of multi-year simulations. Here, an empirical orthogonal function (EOF) analysis is performed to partition the variances from the large and small scales, and a diffusion equation is used for the correlation operator in the background error covariance (Carrier and Ngodock, 2010; Derber and Rosati, 1989; Egbert et al., 1994; Weaver and Courtier, 2001; Weaver et al., 2005; Ngodock, 2005) with appropriate spatial correlations for both the large and small scale components of the MS3DVAR.

The next section presents a summary of the formulation of MS3DVAR by Li et al. (2013), followed by a description of the numerical model used in Section 3. Section 4 deals with the application of the MS3DVAR in the Kuroshio Extension using simulated and real observations, and a summary follows in Section 5.

2. Formulation of ms3dvar

The original formulation of the MS3DVAR scheme was proposed by Li et al. (2013). The scheme uses an incremental 3DVAR approach that explicitly treats multiple scales separately, but was written for a two-scale partition of the traditional 3DVAR cost function into a large and a small scale cost function as

$$J_L(\delta x_L)=\frac{1}{2}\delta x_L^T \boldsymbol{B}_L{}^{-1}\delta x_L+\frac{1}{2}(\boldsymbol{H}\delta x_L-\delta_y)^T(\boldsymbol{R}+\boldsymbol{H}\boldsymbol{B}_s\boldsymbol{H}^T)^{-1}(\boldsymbol{H}\delta x_L-\delta_y) \quad (1)$$

$$J_S(\delta x_S)=\frac{1}{2}\delta x_S^T \boldsymbol{B}_S{}^{-1}\delta x_S+\frac{1}{2}(\boldsymbol{H}\delta x_L-\delta_y)^T(\boldsymbol{R}+\boldsymbol{H}\boldsymbol{B}_L\boldsymbol{H}^T)^{-1}(\boldsymbol{H}\delta x_S-\delta_y) \quad (2)$$

where δx is the incremental state variable defined as $\delta x = x - x^b$, with x^b the background state (such as a previous forecast) created by the forward model. The cost function simply measures the weighted sum of squares of distances to the background state and to the observations. $\boldsymbol{B}_L$ and $\boldsymbol{B}_S$ are the background error covariances associated with the large and small scales respectively δy is the innovation vector (data model misfit) defined as $y-\boldsymbol{H}x^b$, where y is the observation vector and $\boldsymbol{H}$ is the operator that maps the model state to the observations and $\boldsymbol{R}$ is the observation's error covariance, according to the unified notations of Ide et al. (1997).

As mentioned by Li et al. (2013) the use of a multi-scale assimilation introduces a new dimension of flexibility not present in the traditional formulation of 3DVAR. From a practical standpoint this flexibility allows for a full utilization of high resolution observations in addition to coarse resolution observations. These observations can be divided into two types: dense (y^d) and coarse (y^c). The first type can be partitioned again into a large and small scale components y_L^d and y_S^d with observation errors covariances $\boldsymbol{R}_L^d$ and $\boldsymbol{R}_S^d$ respectively, while the second cannot be partitioned. After these partitions, the large scale cost function becomes

$$J_L(\delta x_L)=\frac{1}{2}\delta x_L^T \boldsymbol{B}_L{}^{-1}\delta x_L+\frac{1}{2}(\boldsymbol{H}^d\delta x_L-\delta y_L^d)^T(\boldsymbol{R}_L^d)(\boldsymbol{H}^d\delta x_L-\delta y_L^d)$$
$$+\frac{1}{2}(\boldsymbol{H}^c\delta x_L-\delta y^c)^T(\boldsymbol{R}^c+\boldsymbol{H}^c\boldsymbol{B}_S\boldsymbol{H}^{cT})^{-1}(\boldsymbol{H}^c\delta x_L-\delta y^c) \quad (3)$$

with $\delta y_L^d = y_L^d - \boldsymbol{H}^d x_L^b$ as the large scale innovation. Due to the assimilation of y_L^d the representativeness error in the small scale background errors is removed. Therefore the observational error covariance of the large scale component of the dense observations reduces to $\boldsymbol{R}_L^d$. It can be shown that the elimination of the representativeness error will improve the assimilation of the dense observations into the large scale component of the model state variables (Li et al., 2013). Representativeness error is the error associated with an observation not adequately representing the actual state of the ocean on the scale of the assimilation model. It is important to note that spatially continuous observations, such as satellite sea surface temperature measurements, should be denoted as the dense observation type. This is not true for sparse observations, such as moorings or floats, which lack spatial continuity and thus cannot be partitioned into a large and small scale. These types of observations must then be specified as coarse observations.

The implementation of the small scale cost function is slightly more difficult because there are types of high resolution observations that have highly localized and patchy spatial distributions, such as high frequency radar (Paduan and Graber, 1997). The characteristics of these types of observations can produce unrealistic large scale analysis increments around these patchy localized observations. This problem is resolved by only assimilating these types of observations into the small scale component of the MS3DVAR. So the small scale cost function is

$$J_S(\delta x_S)=\frac{1}{2}\delta x_S^T \boldsymbol{B}_S{}^{-1}\delta x_S+\frac{1}{2}(\boldsymbol{H}^d\delta x_S-\delta\hat{y}^d)^T(\boldsymbol{R}_S^d+\boldsymbol{H}^d\boldsymbol{B}_L^a\boldsymbol{H}^{dT})^{-1}(\boldsymbol{H}^d\delta x_S-\delta\hat{y}^d)$$
$$+\frac{1}{2}(\boldsymbol{H}^c\delta x_S-\delta y^c)^T(\boldsymbol{R}^c+\boldsymbol{H}^c\boldsymbol{B}_L\boldsymbol{H}^{cT})^{-1}(\boldsymbol{H}^c\delta x_S-\delta y^c) \quad (4)$$

where $\delta\hat{y}^a = y^d - \boldsymbol{H}^d x_L^a$ and $x_L^a = x^b + \delta x_L^a$. The large scale analysis increment here is created without localized and patchy observations. Also, the assimilation of the small scale component of the dense observations eliminates the small scales representation error, reducing the small scales component of the dense observations error covariance matrix to $\boldsymbol{R}_S^d$ (Li et al., 2013).

2.1. Background error covariances for the large and small scales

The background error covariances for both the large and small scales cost functions in (3) and (4) take the same form

$$B(x,x')=v(x)^{1/2}v(x')^{1/2}\exp\left(\frac{-|x-x'|^2}{2L^2}\right) \quad (5)$$

where $v(x)$ is the variance and L is the decorrelation length. The differences between $\boldsymbol{B}_L$ and $\boldsymbol{B}_S$ lie in the choice of the variance and the length scales. The latter are chosen to be 100 km and 20 km respectively, and the former are derived from an EOF decomposition of a time series of model solutions. The number of modes used in the EOF analysis is determined by the number of total number of time steps. The lower order modes (1)–(5) are qualitatively determined to represent the large scale features. Fig. 1 shows the amount of variance as a function of depth (left panel) where there is a maximum located at approximately 75 m. The middle and right panels show the variance that is attributed to the large and small scales respectively. At the surface the large scale is approximately 90% of the variance. This value drops to 80% within the mixed layer where small scale features become more energetic. This analysis reveals the importance of large scale features over the majority of the domain. The choice of the covariance form in (5) is

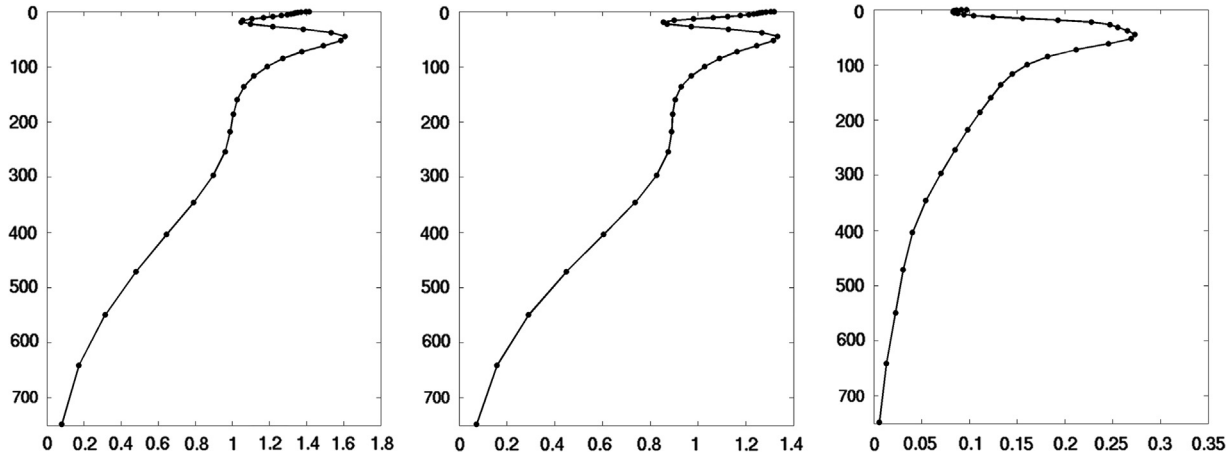

Fig. 1. Mean total temperature variance (left), large scale variance (center), and small scale variance (right). The *y*-axis is depth in meters.

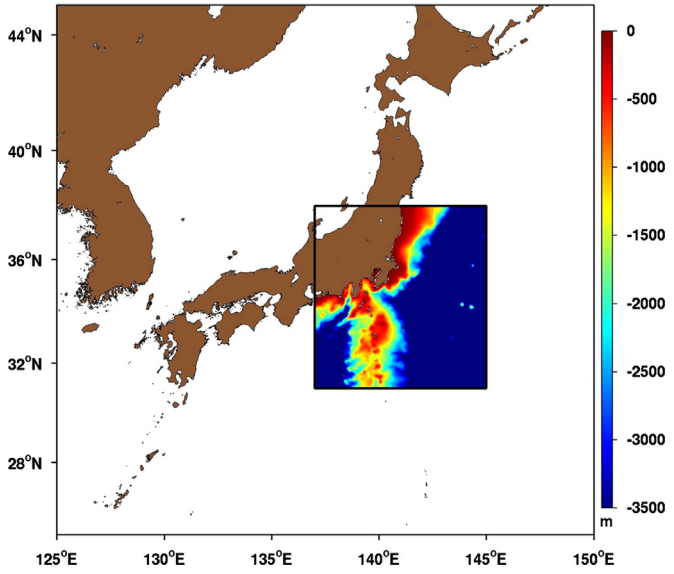

Fig. 2. Western Pacific region including Japan and Korean Peninsula. Relocatable NCOM model domain within black box. Color plot shows model maximum depths in meters.

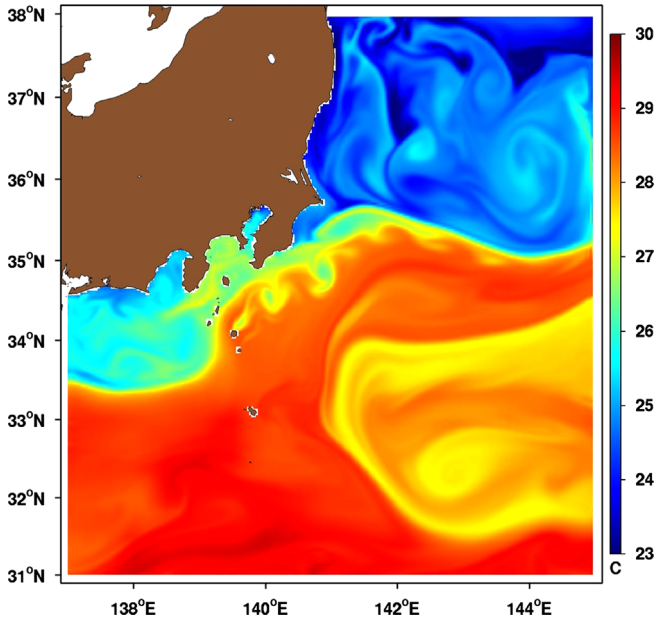

Fig. 3. Model surface temperature in degrees Celsius for 20th September, 2010.

mostly for convenience, because the covariance multiplication during the analysis is obtained as a solution of a suitably initialized diffusion equation as mentioned earlier.

3. Model description

The forward ocean model is the NCOM and produces forecasts of temperature, salinity, sea surface height, and velocity. The model domain spans longitudes 137°E to 145°E and latitudes 31°N to 38°N at 3 km horizontal resolution. The model grid has dimensions of 244×259 with 50 vertical levels. In order to eliminate boundary noise issues the model is nested down from a global grid (1/8 degree resolution) to intermediate resolutions of 9 and 6 km. Each nested grid is inset from its parent grid by 5 grid points. The atmospheric forcing is provided by the Navy Global Atmospheric Prediction System (NOGAPS, Rosmond et al., 2002) with a horizontal resolution of 0.5 degrees which is archived at intervals of 12 h. The model grid location and depth can be seen in Fig. 2.

In this region the Kuroshio appears as a narrow large magnitude current in excess of 30 cm/s which follows the continental margin of the East China Sea and then passes into the deeper North West Pacific over the Izu Ridge south and east of Honshu, seen in Fig. 2 (Niiler et al., 2003). The Kuroshio represents a significant transport mechanism in this region. According to Qu et al. (2001) the eastward flow of this western boundary current at the point where it separates from the main island of Japan (approximately 140° east longitude) is 40 Sv which increases to 51 Sv (144° east longitude). Also present are Kuroshio-induced cold-core eddy trains which have the general form of a von Karman vortex street and are described in detail in Isoguchi et al. (2009). An example of these features can be seen in Fig. 3. The appropriate handling of these features remains difficult for traditional single scale 3DVAR assimilation schemes.

4. Assimilation experiments and results

Here two experiments are examined to study the characteristics and utility of the multi-scale system in a highly energetic region of the coastal ocean. The first experiment uses an idealized

configuration of surface and profile observations taken from a non-assimilative run of the forward model. This experiment will be referred to as the simulated data experiment in which observations consist of surface temperature and subsurface temperature and salinity profiles. The second experiment, referred to as the real data experiment, uses real observations of satellite sea surface temperature (SST) as well as subsurface temperature and salinity profiles. In essence, both experiments assimilate the same types of observations with the exception that one data set is simulated from the model and the other is real. Both of these experiments will be evaluated relative to the analysis time, the 24-h and 48-h forecasts.

4.1. Simulated data experiment

The simulated data experiment consists of two runs: (1) a non-assimilative model run from 13th September to 13th October from which the simulated observations are sampled and (2) an assimilative model run from 2nd September to 30th September. The offset of dates between the non-assimilative and assimilative runs ensures that the initial background to be corrected is significantly different from the simulated data run, and offers a means of verification for the assimilation's effectiveness in driving the model towards the observations. The simulated observations of surface temperature and subsurface profiles of temperature and salinity are sampled from the non-assimilative run and assimilated at a 24-h interval with the MS3DVAR and a traditional 3DVAR. A 24-h forecast is run using the analysis field as the initial condition; this becomes the background for the next assimilation cycle. The simulated observations are sampled in such a way to mimic the spatial distribution of a field experiment with coarse satellite coverage over the whole domain along with several profiling buoys and a dense collection of observations meant to sample small scale features generated in the island wakes within the core of the Kuroshio. An example of these features, e.g. meanders, can be seen in Fig. 3. The spatial distribution of the simulated observation is seen in Fig. 4.

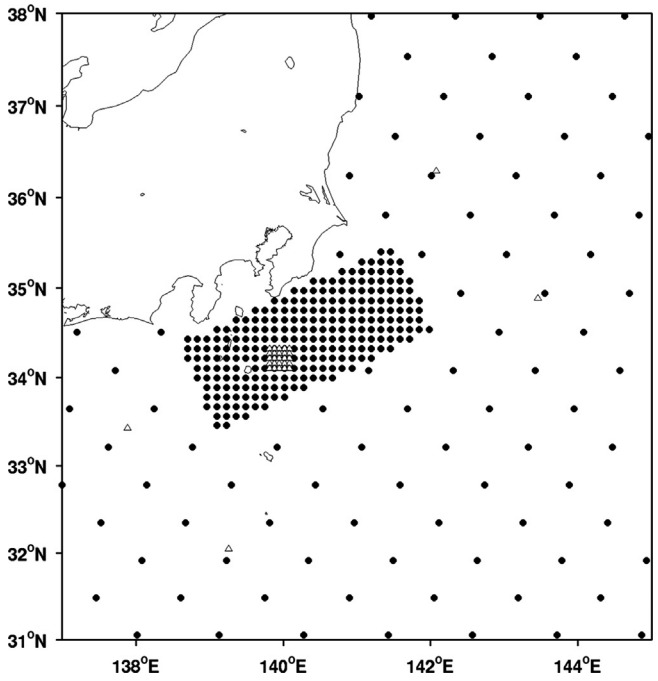

Fig. 4. Simulated observation locations. Surface observations (black circles) and profile locations (white triangles).

4.1.1. Experiment results

To assess the performance of the MS3DVAR comparisons with a traditional 3DVAR system are made. The MS3DVAR partitions the total variance into large (100 km) and small (20 km) scale contributions, whereas the traditional 3DVAR system uses the total variance and has only one decorrelation scale set at 100 km. There is no clear or rigorous separation of scales between 100 km and 20 km. It is understood that the scales of features are better described as a cascade. Attempting to represent all the scales in a cascade would be burdensome and computationally costly and beyond the scope of this study.

The differences between the assimilation and the observations is examined at the observation locations using the fit to the observations metric

$$e_b = \frac{1}{M}\sum_{m=1}^{M}\frac{|y_m - x_m|}{\sigma_m{}^{1/2}} \tag{6}$$

where m is the observation index, x_m is the model state (mapped to the observation locations), y_m is the observation, M is the total number of observations and $\sigma_m^{1/2}$ is the observation standard deviation. Because the assimilation is expected to fit the observations to within one standard deviation, the metric is expected to have values less or equal to 1. When this metric is applied to an analysis or forecast trajectory it can be used to examine the time variability of the errors. Although the latter is not subject to the same expectation as the analysis, the metric still reveals improvements or deteriorations of the forecast.

Fig. 5 shows the comparison of the fit to the observation metric (e_b) for the analysis, 24 h forecast, and 48 h forecast. Obviously the smaller the error metric the better a given assimilation scheme is at representing the observations when compared to the simulated observations. It is clear that the analysis time series comparison shows the MS3DVAR has a significantly lower metric than the traditional 3DVAR. The MS3DVAR has a sharp initial decrease and remains slightly less than 1 until the end of the month where there is another sharp decrease. The traditional 3DVAR does not exhibit the same sharp decrease. Additionally, the errors for both the MS3DVAR and the traditional 3DVAR appear to decrease over time, an indication that the systems have a 'spin-up' period; however, this is not investigated in this study. The 24-h and 48-h forecasts have several peaks in the time series of the error metrics. In general the traditional 3DVAR has a higher maximum value (i.e. large errors) than the MS3DVAR. Also it may be helpful to note that the scales on Fig. 5 are not the same. The errors are higher for the 24-h and 48-h forecasts compared to the analysis time.

Fig. 6 depicts a scatter plot comparison between all the observations and the model analyses, 24-h, and 48-h forecast values for both the MS3DVAR and the traditional 3DVAR. The red points are the traditional 3DVAR while the blue are the MS3DVAR. The differences between the observations and the model are smallest at the analysis time (left panel). The differences between the simulated data run and the model increase as the forecast is pushed out to 24-h and 48-h. The spread at the analysis time is significantly smaller than in the forecasts. The spread shown between the data and the forecast values are similar between the 24-h and 48-h results. Examining the RMS error however indicates that the MS3DVAR produces lower error than the traditional 3DVAR (RMS values shown in Fig. 6).

4.1.2. Three one day experiments

A second comparison experiment between the traditional and the MS3DVAR is carried out, where the density of simulated observations is drastically increased, and the background is chosen to be significantly different from the simulated data run from

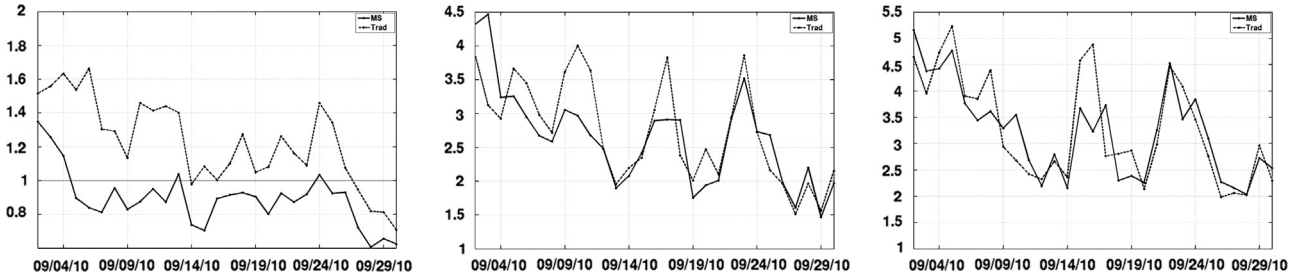

Fig. 5. e_b comparison metric time series between MS3DVAR (solid line) and traditional 3DVAR (dashed line) for September 2010 at the analysis time (left), 24 h forecast (middle) and 48 h forecast (right).

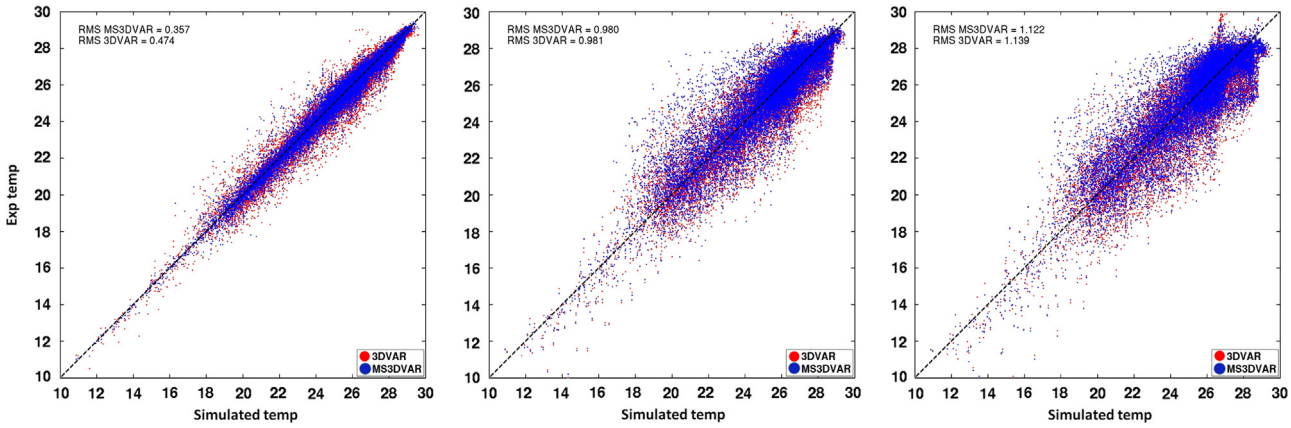

Fig. 6. Scatter plot of temperature for the MS3DVAR (blue points) and traditional 3DVAR (red points) at the analysis time (left), the 24-h forecast (middle) and the 48-h forecast (right) versus simulated observations taken from the simulated data run. The values given in the plot are the root-mean-square (RMS) difference between the analyses and the observations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

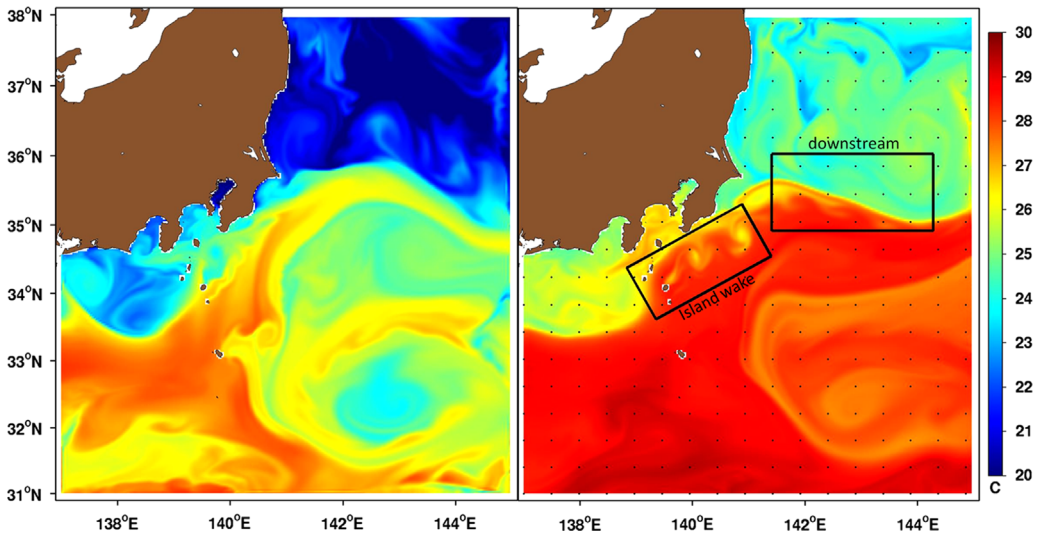

Fig. 7. Surface temperature, in °C, of the background field (left) and simulated data run field (right) regions of interest are denoted by black boxes and coarse observations are the black points.

which the observations are sampled, at a time when the model exhibits small scale features as seen in Fig. 7. The black boxes shown in Fig. 7 represent regions of interest where there are significant differences between the background and the simulated data run. The island wake region is a generation area for small scale eddies and is densely sampled in space. The downstream region is important for evaluating the influence of correcting small scale features in the island wake region and letting them propagate downstream. This experiment aims to highlight the inability of the traditional 3DVAR to handle dense observations accurately. First, a coarse subset of the dense observations is assimilated using the traditional 3DVAR, then the dense observations are assimilated using 3DVAR, and finally the dense observations are assimilated using the MS3DVAR. All three experiments in this section are carried out for a single day and use the same background.

A comparison of the analyses from these three experiments with the simulated data run reveals the limitations of the traditional 3DVAR. Qualitative comparisons using the color plots as well as using RMS differences over the full region, the downstream region, and the island wake region are examined. A direct comparison between the

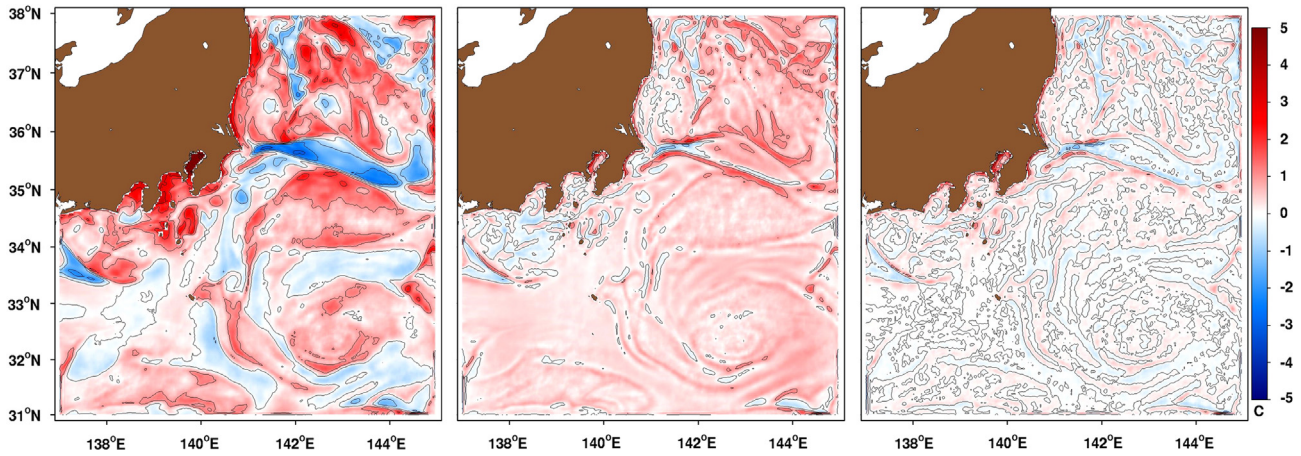

Fig. 8. Surface temperature from simulated data run minus analysis field for coarse observations with traditional 3DVAR experiment (left), dense observations with traditional 3DVAR experiment (middle), and dense observations with MS3DVAR experiment (right).

Table 1
Root mean square differences of surface temperature between simulated data run and experiments for three regions in °C.

Data/Assimilation /Length scale (km)	Full	Downstream	Island wake
Coarse/traditional 3DVAR/100	**0.94**	**1.37**	**1.02**
Dense/traditional 3DVAR/100	**0.55**	**0.8**	**0.36**
Dense/traditional 3DVAR/50	**0.33**	**0.36**	**0.31**
Dense/traditional 3DVAR/20	**0.37**	**0.41**	**0.29**
Dense MS3DVAR	**0.3**	**0.33**	**0.27**

simulated data run and the background field shows that the temperatures are higher in the simulated data run and that the placement of the Kuroshio in the downstream region is different. Additionally, there are major differences between the magnitude and shape of the small scale features present in the island wake region.

Fig. 8 shows these differences for each of the three experiments. The largest errors are found in the traditional 3DVAR experiment with the coarse observations (left panel). This is expected and represents a common problem in data assimilation experiments where there are not enough observations to constrain the models. The middle panel of Fig. 8 shows the traditional 3DVAR comparison with the simulated data run when the observation density is greatly increased to mimic the spatial resolution of a geostationary operational environmental satellite (GOES), for this example a spatial resolution of 6 km was used. The right panel shows the same observations as the middle panel but now using the MS3DVAR. Simply put, small scale features cannot be reproduced without high density observations. There also seems to be an inherent smoothing of the small scale features in the traditional 3DVAR experiments due to the long decorrelation length. The MS3DVAR better represents the multiple scales present in these dense observations. This is confirmed by the smaller differences in the experiment with the MS3DVAR when compared to the traditional 3DVAR. To support this qualitative assessment Table 1 shows some RMS results which reveal the dense MS3DVAR experiment to have the lowest errors over the full domain. The RMS errors in the downstream region for the MS3DVAR are comparable to the full domain. This is not the case for the traditional 3DVAR experiments, where the coarse collection of observations lacks the necessary coverage, and the dense collection has too many observations within a decorrelation length of the background error covariance, causing the cooling of the Kuroshio in this region to be over estimated. In the island wake region the MS3DVAR again produces the lowest RMS errors with the dense traditional 3DVAR performing better than the coarse traditional.

It can be argued that setting the decorrelation length scale for the traditional 3DVAR at 100 km puts it at a disadvantage. Several additional experiments with the traditional 3DVAR decorrelation length scale set at 50 km and 20 km show that RMS errors decrease for these runs (seen in Table 1). However, this decrease in length scale may not be appropriate for most basin scale domains which hope to represent the large scale first. Also the use of a small decorrelation length in a traditional 3DVAR will only reduce errors when the observations are extremely dense. For most real world cases a larger length scale is more prudent. Generally, the traditional 3DVAR will represent a single length scale whereas the MS3DVAR represents multiple scales.

4.2. Real data experiment

The same NCOM grid and 3DVAR length scales from Section 4.1 are also used here for the assimilation of real observations, unlike the experiment in the previous section which used simulated observations. The remotely sensed and in-situ ocean data assimilated here include GOES SST, ARGO profiling floats (Roemmich et al., 2001), and in-situ profile buoys. These observations are collected, prepared and quality controlled with the Navy Coupled Ocean Data Assimilation (NCODA) system (Cummings, 2005, 2011; Cummings and Martin, 2013). The experiment uses a 24-h update cycle, assimilating observations at 0000 UTC of each day, and the time period for this experiment is 1st September through 31st October 2010.

Fig. 9 shows an example of the distribution of SST observations for 17th September, 2010. In order to perform the MS3DVAR assimilation the observations need to be partitioned into dense and coarse types. A simple sorting method is used here. For each observation location the distance to the nearest observation is recorded. If that distance is less than 15 km it is flagged as dense. These dense flagged data are then examined again where the distance to the nearest dense observation is recorded and if it is less than 15 km it is again flagged as dense. These double flagged data are determined to be of dense data type while the remaining observations are of the coarse data type. This partitioning is only relevant for the MS3DVAR scheme; the traditional 3DVAR uses all observations. All the profile observations are assigned to the coarse type.

4.2.1. Experiment results

Again, in order to examine the temporal variability of the analysis errors at the observation locations we look at the fit to the observations metric e_b over September and October 2010 (Fig. 10). In general the error metric shows that MS3DVAR has lower values when compared to the traditional 3DVAR. It is important to note there are no observations on 14th October, 2010 as seen by the missing data point. The left panel in Fig. 10 depicts the metric at the analysis time with the majority of the metric values near or less than one. This implies that the analysis is fitting the observations to within a standard deviation. The metric comparison at the analysis time (left panel) shows variability in September which then decreases in October. This variability can be attributed to the atmospheric forcing. Additionally, the middle and right panels show the observation metric for the 24-h and 48-h forecasts respectively. The atmospheric forcing is also responsible for larger values of the observation metric in the forecast fields. Obviously the forcing is most influential at the surface where the majority of the observations are located. Fig. 11 shows that there is an increased amount of variance in the surface temperature in September that then decreases in October thereby confirming the observed decrease in variability of e_b over this time period.

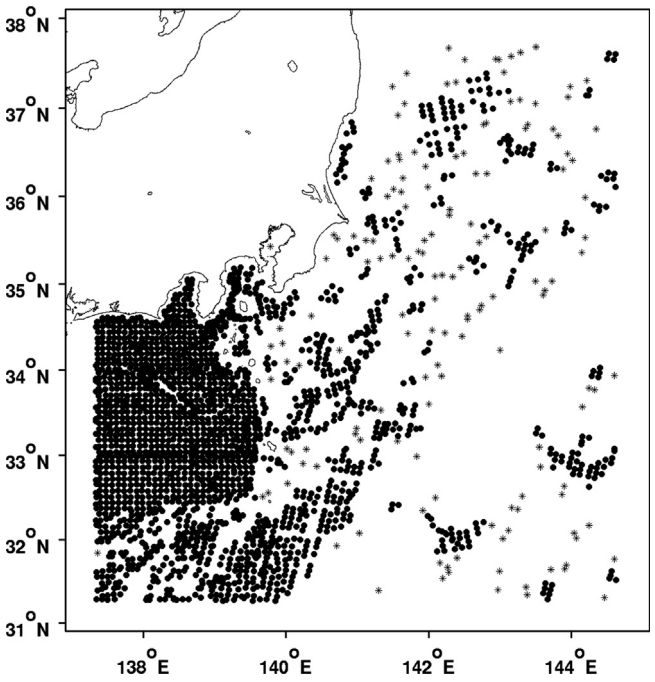

Fig. 9. SST observation locations for 17th Sept, 2010. Dense observations (circles) and coarse locations (asterisks).

Fig. 12 shows scatter plots between the model and observations for both the MS3DVAR and the traditional 3DVAR at the analysis, 24-h and 48-h forecasts . For all of the time levels the MS3DVAR produces the smallest errors when compared to the traditional 3DVAR and is significantly better than the non-assimilative model run. These figures show that the free run diverges away from the observations much more than the model runs using data assimilation, which is to be expected. The scatter plots for the MS3DVAR and the traditional 3DVAR forecasts are relatively similar with some minor differences relating to lower error values for the MS3DVAR.

5. Summary

An MS3DVAR was developed following Li et al., 2013, and was used in assimilating simulated and real observations with the Navy coastal ocean model in the Kuroshio extension. The main advantage of this MS3DVAR includes the ability to handle dense collections of observations (that would otherwise be redundant according to a single correlation scale) through explicitely accounting for large and small correlation scales within the cost function minimization.

The comparisons between the MS3DVAR and traditional 3DVAR assimilation experiments show decreased analysis and forecast

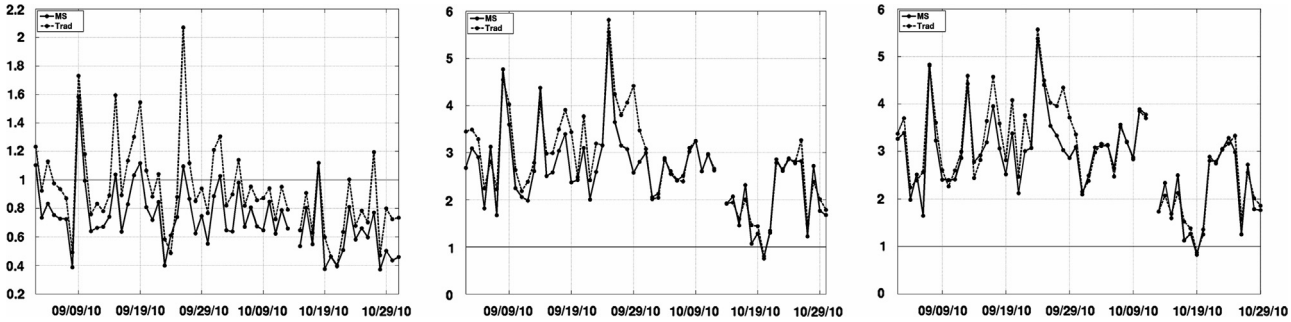

Fig. 10. e_b comparison metric time series between MS3DVAR (solid line) and traditional 3DVAR (dashed line) for September and October 2010 at the analysis time (left), 24 h forecast (middle) and 48 h forecast (right).

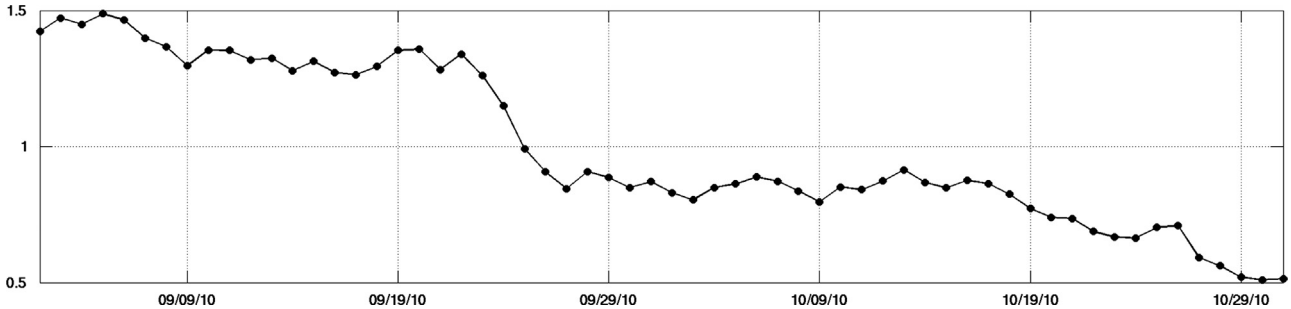

Fig. 11. Mean total variance of temperature time series for September and October 2010.

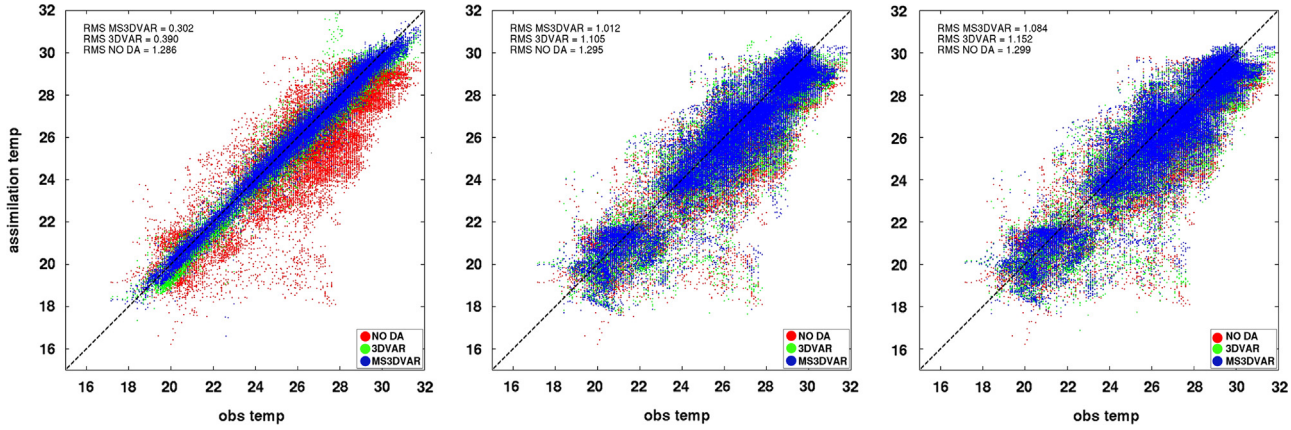

Fig. 12. Scatter plot of temperature for the MS3DVAR (blue), traditional 3DVAR (green) and the non-assimilative model run (red) at the analysis time (left), the 24 h forecast (middle) and the 48 h forecast (right) versus real observations. The values given in the plot are the root-mean-square (RMS) difference between the analyses and the observations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

errors for the multi-scale system. These statistics hold for the simulated and real observations. Results suggest that the MS3DVAR is an improved method of data assimilation when compared to a traditional 3DVAR. The traditional 3DVAR smooths over the small scale features causing them to be under represented. Therefore, the MS3DVAR seems better suited for representing these small scale features allowing for improvements to the analyses and forecasts.

The improvements to the MS3DVAR forecasts are not as large as the improvements to the analysis times. This is mainly because most of the circulation is dominated by the large scale features, and the small scales account only for a small fraction of the variance. Additionally, atmospheric forcings applied to both the 3DVAR and MS3DVAR may be contributing to the similarity of the forecasts even after corrections are applied at the analysis time. Nevertheless the marginal gains in analysis and forecast accuracy may prove beneficial or even crucial in areas where small scale features play a more significant role.

Scales are not just large and small, but they are a cascade. We have not yet addressed where the line is to be drawn. Addressing this is beyond the scope of this paper, which only attempted to apply the MS3DVAR in a more energetic circulation region than where it has been applied so far.

Acknowledgments

This work was sponsored by the Office of Naval Research as part of the project “A Multiscale approach to assessing predictability of ASW environment.” This paper is the Naval Research Laboratory paper contribution number JA/7320-13-1695.

References

Carrier, M.J., Ngodock, H., 2010. Background-error correlation model based on implicit solution of a diffusion equation. Ocean Modelling 35, 45–53.
Cummings, J.A., 2005. Operational multivariate ocean data assimilation. Q. J. R. Meteorol. Soc. 131, 3583–3604.
Cummings, J.A., 2011. Ocean data quality control. In: Schiller, A., Brassington, G.B. (Eds.), Operational Oceanography in the 21st Century. Springer Science+Business Media B.V., pp. 91–121.
Cummings, J.A., Martin, O.M., 2013. Variational data assimilation for the global ocean. In: Park, S.K., Xu, L. (Eds.), Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, vol. II. Springer, Verlag, Berlin, Heidelberg.
Derber, J., Rosati, A., 1989. A global oceanic data assimilation system. J. Phys. Oceanogr. 19, 1333–1347.
Egbert, G.D., Bennett, A.F., Foreman, M.G.G., 1994. Topex/Poseidon tides estimated using a global inverse model. J. Geophys. Res. 99, 24821–24852.
Evensen, G., 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dyn. 53, 343–367.
Farrara John D., Yi Chao, Zhijin Li, Xiaochun Wang, Xin Jin, Hongchun Zhang, Peggy Li, Quoc Vu, Peter Q. Olsson, Schoch G. Carl, Mark Halverson, Mark A. Moline, Carter Ohlmann, Mark Johnson, James C. McWilliams, Francois A. Colas, 2013. A data-assimilative ocean forecasting system for the Prince William sound and an evaluation of its performance during sound Predictions 2009, Cont. Shelf Res. 63 Suppl., S193–S208, ISSN 0278-4343.
Hurlburt, H.E., Metzger, E.J., Sprinthall, J., Riedlinger, S.N., Arnone, R.A., Shinoda, T., Xu, X., 2011. Circulation in the Philippine Archipelago simulated by 1/12 and 1/25 global hycom and EAS NCOM. Oceanography 24, 28–347.
Ide, K., Courtier, P., Ghil, M., Lorenc, A., 1997. Unified notation for data assimilation: operational sequential and variational. J. Meteorol. Soc. Jpn. 75 (1B), 71–79.
Isoguchi, O., Shimada, M., Sakaida, F., Kawamura, H., 2009. Investigation of Kuroshio-induced cold-core eddy trains in the lee of the Izu Islands using high-resolution satellite images and numerical simulations. Remote Sens. Environ. 113, 1912–1925.
Li, Z., 2011. A multi-scale three-dimensional variational data assimilation scheme and its application to coastal oceans. In: Proceedings of the 9th Workshop on Adjoint Model Applications in Dynamic Meteorology, Cefalu, Sicily, Italy, ⟨http://gmao.gsfc.nasa.gov/events/adjoint_workshop-9/presentations/Li.pdf⟩.
Li, Z., Chao, Y., Farrara, J.D., McWilliams, J.C., 2012. Impacts of distinct observations during the 2009 Prince William sound field experiment: a data assimilation study. Cont. Shelf Res. (in press).
Zhijin Li, Yi Chao, John D. Farrara, James C. McWilliams, 2013. Impacts of distinct observations during the 2009 Prince William Sound field experiment: a data assimilation study, Cont. Shelf Res. 63, Suppl. S209–S222, ISSN 0278-4343.
Martin, P.J., 2000. Description of the Navy Coastal Ocean Model Version 1.0. NRL Rep. NRL/FR/7322/00/9962, 45 p. [Available from NRL, Code 7322, Bldg. 1009, Stennis Space Center, MS 39529-5004].
Metzger, E.J., Hurlburt, H.E., Xu, X., Shriver, J.F., Gordon, A.L., Sprintall, J., Susanto, R.D., Van Aken, H.M., 2010. Simulated and observed circulation in the Indonesian Seas: 1/12 global HYCOM and the INSTANT observations. Dyn. Atmos. Oceans, 275–300.
Ngodock, H., 2005. Efficient implementation of covariance multiplication for data assimilation with the representer method. Ocean Modelling 8, 237–251.
Niiler, P.P., Maximenko, N.A., Panteleev, G.G., Yamagata, T., Olson, D.B., 2003. Near-surface dynamical structure of the Kuroshio Extension. J. Geophys. Res. 108 (C6), 3193.
Paduan, J.D., Graber, H.C., 1997. Introduction to high-frequency radar: reality and myth. Oceanography 10, 36–39.
Qu, T., Mitsudera, H., Qui, B., 2001. A climatological view of the Kuroshio/Oyashio System east of Japan. J. Phys. Oceanogr. 31, 2575–2589.
Roemmich, D., Boebel, O., Desaubies, Y., Freeland, H., Kim, K., King, B., Le Traon, P.Y., Molinari, R., Owens Brechner, W., Riser, S., Send, U., Takeuchi, K., Wijfels, S., 2001. Argo: the global array of profiling floats. In: Koblinsky, C.J., Smith, N.R. (Eds.), Observing the Oceans in the 21st Century. Melbourne Bureau of Meteorology, p. 604.
Rosmond, T.E., Teixeira, J., Peng, M., Hogan, T.F., Pauley, R., 2002. Navy operational global prediction system (NOGAPS): forcing for ocean models. Oceanography 15, 99–106.
Weaver, A., Courtier, P., 2001. Correlation modeling on the sphere using a generalized diffusion equation. Q. J. R. Meteorol. Soc. 127, 1815–1846.
Weaver, A.T., Deltel, C., Machu, E., Ricci, S., Daget, N., 2005. A multivariate balance operator for variational ocean data assimilation. Q. J. R. Meteorol. Soc. 131, 3605–3625.
Xiao, Q., Tastula, E.M., 2011. .
Xie, Y., Koch, S., McGinley, J., Albers, S., Bieringer, P.E., Wolfson, M., Chan, M., 2011. A space-time multiscale analysis system: a sequential variational analysis approach. Mon. Weather Rev.139 (4), 1224–1240.