Research Article

A Potential Density Gradient Dependent Analysis Scheme for Ocean Multiscale Data Assimilation

Hongli Fu,¹ Jinkun Yang,¹ Wei Li,¹ Xinrong Wu,¹ Guijun Han,¹ Yuanfu Xie,² Shaoqing Zhang,^{3,4} Xuefeng Zhang,¹ Yingzhi Cao,¹ and Xiaoshuang Zhang¹

¹Key Laboratory of Marine Environmental Information Technology, State Oceanic Administration, National Marine Data and Information Service, Tianjin 300171, China

²NOAA/Earth System Research Laboratory, Boulder, CO, USA

³Key Laboratory of Physical Oceanography, MOE China, Ocean University of China, Qingdao 266100, China

⁴Qingdao National Laboratory for Marine Science and Technology, Qingdao 266100, China

Correspondence should be addressed to Jinkun Yang; yangjinkun@nmdis.org.cn

Received 21 May 2017; Revised 20 September 2017; Accepted 3 October 2017; Published 5 December 2017

Academic Editor: Stefania Bonafoni

Copyright © 2017 Hongli Fu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study addresses how to maintain oceanic mixing along potential density surface in ocean data assimilation (ODA). It is well known that the oceanic mixing across the potential density surface is much weaker than that along the potential density surface. However, traditional ODA schemes allow the mixing across the potential density surface and thus may result in extra assimilation errors. Here, a new ODA scheme that uses potential density gradient information of the model background to rescale observational adjustment is designed to improve the quality of assimilation. The new scheme has been tested using a regional ocean model within a multiscale 3-dimensional variational framework. Results show that the new scheme effectively prevents the excessive unphysical projection of observational information in the direction across potential density surface and thus improves assimilation quality greatly. Forecast experiments also show that the new scheme significantly improves the model forecast skills through providing more dynamically consistent initial conditions

1. Introduction

Data assimilation makes use of model background to retrieve observational information for generating continuous observation-constrained model states in time and space which serve as model initial conditions for numerical weather and climate predictions. Scientists have been pursuing physically consistent analysis methods to minimize assimilation error and help increase prediction accuracy. For example, the multigrid 3-dimensional variational (3D-Var) algorithm (e.g., [1–3]) has been introduced to avoid the deficiency of insufficient spectra of error correction. Ones also introduce physical constraints or balances to improve the assimilation performance, such as the geostrophic balance [4, 5], the temperature-salinity (T-S) relationship (e.g., [6–9]), and total heat content and salt content conservation in ocean data assimilation (ODA) [10–12].

In this study, we address the importance of maintaining the natural oceanic mixing process in ODA. Given the

isopycnal nature of ocean water movement, a natural oceanic mixing occurs mostly along the potential density surface so that the mixing across the potential density surface is always much weaker than that along the potential density surface. However, when projecting an observational increment onto the neighboring model space, traditional analysis schemes use covariance and a distance-dependent localization factor without much consideration of mixing physics. Given the important role of mixing in ocean circulations, such deficiency can result in extra assimilation errors and limit the quality of ODA.

As a demonstration, we use the following simple case to illustrate the role of maintaining a natural mixing process in ODA. Assuming one observation at point (100°E, 20°N, 100 m) has increment of 5°C and the other observation at point (150°E, 20°N, 300 m) has increment of 1°C, these two observations locate at same potential density layer with 1027 km/m³ value (Figure 1). To some extent, the data

FIGURE 1: Schematic illustration of isopycnal data assimilation. Since oceanic mixing usually occurs along potential density surface, assuming one observation at point (100°E, 20°N, 100 m) has increment of 5°C and the other observation at point (150°E, 20°N, 300 m) has increment of 1°C, these two observations locate at same potential density layer with 1027 km/m³ value and a smart data assimilation should simulate real mixing procedure and result in an analyzed increment value of 3°C at point (125°E, 20°N, 250 m), not at point (125°E, 20°N, 200 m).

assimilation can be regarded as a mixing procedure [11], and the oceanic mixing usually occurs along the potential density surface. However, a simple data assimilation approach without considering the gradient of the potential density may linearly result in an analyzed increment value of 3°C at point (125°E, 20°N, 200 m). But the real mixing procedure should happen on the same potential density surface, and therefore a dynamically consistent data assimilation approach should simulate this procedure and result in an analyzed increment value of 3°C at point (125°E, 20°N, 250 m) and the value at point (125°E, 20°N, 200 m) should be reduced to about 2°C.

The Princeton Ocean Model with a generalized coordinate system (POMgcs, which will be discussed in Section 4.1) is used to show the dynamical evolution of the analyzed increment relative to the background in a 3D-Var case. Figure 2(a) shows the analyzed increment at the initial time, in which the increment shape does not follow the isolines of the potential density, and instead many increment contours run across the isolines. We integrate the model for 5 days starting from the background fields and analyzed fields. We show their difference as the dynamical evolution of the analyzed increments in Figure 2(b). It is seen that the model dynamics tend to push the increments to evolve along the isoline of the potential density, especially in the frontal regions (the Kuroshio extension and the North Equator Current areas, e.g.). This means that if the initial analysis increments are consistent with the distribution of isolines of the potential density, both the assimilation and forecast errors will be reduced as the model forecast initial shock is minimized.

Based on the above discussion, we will design an isopycnal multiscale mapping operator in the original multigrid 3D-Var proposed by Li et al. [2, 13, 14] to implement the mixing consistent ODA scheme. The paper is organized as follows: Section 2 gives the methodology of the new data assimilation scheme. The data assimilation experiments and result analyses are given in Section 3. To further validate this new scheme

in oceanic numerical forecast, “twin” forecast experiments are conducted and analyzed in Section 4. Section 5 gives conclusions and discussions.

2. Methodology

The traditional 3D-Var, using either correlation scales or recursive filters, can only correct certain wavelength errors [1]. Note that the shortwave errors should not be sufficiently corrected until the longwave ones are corrected; otherwise, longwave errors could be mistakenly treated as shortwave errors, resulting in erroneous analyses. To correctly minimize the longwave and shortwave errors, a multigrid 3D-Var data assimilation scheme is proposed in Li et al.’s papers [2, 13, 14]. In this study, we will use the multigrid 3D-Var as an example to design the potential density dependent ODA scheme. Our solution is to introduce a preprocessing in the multigrid 3D-Var system. Before optimization on each level grid, the preprocessing projects the observational residues between the observations and the sum of all the previous level grid analyses at observational point to the present level grid points using a mapping operator with the spatial correlation scale corresponding to the length of this level grid and the threshold of potential density difference. Thus, the assimilated data are not the residues on observation locations but the projected residues on the analysis grid points.

The mapping operator is described as follows. Following the basic idea of the multigrid 3D-Var (e.g., [1]) and the convention in Li et al. [2], let

$$\mathbf{Y}^{(n)} = \mathbf{Y}^{(n-1)} - \mathbf{H}^{(n-1)} \mathbf{X}^{(n-1)} \quad (n = 1, 3, \dots, N), \quad (1)$$

where n represents the n th level grid. $\mathbf{Y}$ represents the observational residue at observational point. Note that $\mathbf{Y}^{(0)}$ is the original observation vector $\mathbf{Y}^{\text{obs}}$ and $\mathbf{X}^{(0)}$ is the model background $\mathbf{X}^b$. $\mathbf{H}^{(n-1)}$ is a linear mapping operator from the $(n-1)$ th grid level to the observation space. And $\mathbf{H}^{(0)}$ projects the background field to observation space. $\mathbf{X}^{(n-1)}$ is the analysis results of the last level grid and N is the final level grid of the multigrid analysis which depends on the observations’ distribution.

The proposed mapping operator based on the correlation scales and the threshold of potential density can be expressed as

$$F_{il}^{(n)} = M_i^{(n)} \cdot \exp \left\{ - \left[\frac{x_i^{(n)} - x_l^{\text{obs}}}{L_x^{(n)}} \right]^2 - \left[\frac{y_i^{(n)} - y_l^{\text{obs}}}{L_y^{(n)}} \right]^2 - \left[\frac{z_i^{(n)} - z_l^{\text{obs}}}{L_z^{(n)}} \right]^2 - \left[\frac{t_i^{(n)} - t_l^{\text{obs}}}{L_t^{(n)}} \right]^2 - \left[\frac{f_i^{(n)} - f_l^{\text{obs}}}{L_f} \right]^2 \right\}, \quad (2)$$

where the subscript i is index of grid point and superscript obs represents observation. $x_i^{(n)}$, $y_i^{(n)}$, $z_i^{(n)}$, and $t_i^{(n)}$ are the

FIGURE 2: Model dynamical evolution of analyzed temperature increment (color shade) relative to background. (a) Analyzed increment; (b) evolution of analyzed increment after 5-day simulation. Contour: isoline of potential density at 300 m.

spatial and temporal coordinates of i th grid point on n th level grid. x_l^{obs} , y_l^{obs} , z_l^{obs} , and t_l^{obs} are the spatial and temporal coordinates of l th observational point. $L_x^{(n)}$, $L_y^{(n)}$, $L_z^{(n)}$, and $L_t^{(n)}$ represent the correlation scales of zonal, meridional, vertical, and temporal direction, respectively, on n th level grid, and are equal to resolution of the corresponding directions of this level grid. $f_i^{(n)}$ stands for the potential density of model background at i th grid point on n th level grid. f_l^{obs} stands for the potential density of model background at l th observation point. L_f is the threshold value of the potential density difference of model background between the observational point and the grid point and is set to 0.05 kg/m^3 in this study. $M_i^{(n)}$ is a mask value of i th grid point on n th level grid with 1 representing wet grid and 0 representing dry grid. The black dot means multiplication operation. Therefore based on the model instant background potential density information, if the potential density difference between the observational point and the grid point is large (small) compared to the threshold, the projected observational residual value on grid point will be small (large). This character of the mapping operator can effectively inhibit the excessive spread of observation residue across potential density surface. From (2), this mapping operator can in general handle the temporal information for observations and can deal with the 4D data assimilation, but in this study, for simplicity, we omit the temporal term $-[(t_i^{(n)} - t_l^{\text{obs}})/L_t^{(n)}]^2$ and only discuss the 3D data assimilation.

Given totally P observational points, using the above mapping operator, the gridded observational residue $G_i^{(n)}$ on i th grid point of n th level analysis grid is

$$G_i^{(n)} = \frac{\sum_{l=1}^P \{F_{il}^{(n)} \cdot Y_l^{(n)}\}}{\sum_{l=1}^P F_{il}^{(n)}}. \quad (3)$$

Since for a given grid point all observations can make an observational residual contribution to it, the final gridded observational residue should take the form of weighted average of all the observational residues calculated by (1). The black dot means multiplication operation.

It should be noted that the same goal could be achieved by imposing a density gradient constraint on the background error covariance of the traditional 3D-Var data assimilation [4, 15–17]. The traditional 3D-Var using correlation scale method or recursive filter method is a single scale data assimilation method [1]. Although the density gradient constraint can be used in the traditional 3D-Var, it can only resolve single scale information. In this study, the multigrid 3D-Var is used to extract multiscale information from observation, and the density gradient constraint is used in this assimilation scheme. More reasonable analysis results with the multiscale information from observational network can be anticipated.

Following (9) in Li et al. [13] and (5) in Li et al. [14], the cost function of the original multigrid 3D-Var is

$$\begin{aligned} J^{(n)}[\mathbf{X}^{(n)}] &= \frac{1}{2} \gamma \mathbf{X}^{(n)T} \mathbf{S}^{(n)} \mathbf{X}^{(n)} \\ &+ \frac{1}{2} [\mathbf{H}^{(n)} \mathbf{X}^{(n)} - \mathbf{Y}^{(n)}]^T \mathbf{O}^{(n)-1} [\mathbf{H}^{(n)} \mathbf{X}^{(n)} - \mathbf{Y}^{(n)}] \end{aligned} \quad (4)$$

$(n = 1, 2, \dots, N),$

where the matrix $\mathbf{O}$ stands for the observational error covariance matrix in the observation space. Superscripts T and -1 represent matrix transpose and inverse respectively. $(1/2)\gamma\mathbf{X}^{(n)T}\mathbf{S}^{(n)}\mathbf{X}^{(n)}$ is a smoothing term with matrix $\mathbf{S}^{(n)}$ being a smoothing matrix at n th level grid, and γ is a penalty parameter weighing how strong the smoothing term will be. The smoothing matrix $\mathbf{S}^{(n)}$ is introduced in order to remove the local small-scale noise (such as “bull’s-eye”) in the analysis and is set to a static two-order smoothing operator S [14]. If $\mathbf{X}^{(n)}$ is a 2D matrix with (im, jm) dimension, the expression of the smoothing term can be formulated by

$$\begin{aligned} & \frac{1}{2}\mathbf{X}^{(n)T}\mathbf{S}\mathbf{X}^{(n)} \\ &= \frac{1}{2}\sum_{i=2}^{im-1}\sum_{j=1}^{jm}\left(A_iX_{i-1,j}^{(n)} - 2X_{i,j}^{(n)} + B_iX_{i+1,j}^{(n)}\right) \\ &+ \frac{1}{2}\sum_{i=1}^{im}\sum_{j=2}^{jm-1}\left(C_jX_{i,j-1}^{(n)} - 2X_{i,j}^{(n)} + D_jX_{i,j+1}^{(n)}\right), \end{aligned} \quad (5)$$

where $X_{i,j}^{(n)}$ is the (i, j) th element of $\mathbf{X}^{(n)}$ and the four coefficients are

$$\begin{aligned} A_i &= \frac{2[\text{lon}(i+1) - \text{lon}(i)]}{[\text{lon}(i+1) - \text{lon}(i-1)]}, \\ B_i &= \frac{2[\text{lon}(i) - \text{lon}(i-1)]}{[\text{lon}(i+1) - \text{lon}(i-1)]}, \\ C_i &= \frac{2[\text{lat}(j+1) - \text{lat}(j)]}{[\text{lat}(j+1) - \text{lat}(j-1)]}, \\ D_i &= \frac{2[\text{lat}(j) - \text{lat}(j-1)]}{[\text{lat}(j+1) - \text{lat}(j-1)]}, \end{aligned} \quad (6)$$

where $\text{lon}(i)$ and $\text{lat}(j)$ represent the longitude and latitude of the (i, j) th model grid, respectively. It is easy to infer that the smoothing matrix S is a quasidiagonal matrix whose magnitude is $O(I)$, where I is the identity matrix. The multigrid 3D-Var data assimilation scheme attempts to obtain the optimal solution of $\mathbf{X}^{(n)}$ through balancing the smoothing term and the observation term. Therefore, the $\mathbf{O}$ matrices are used to balance the smoothing matrices and are different from the observational error covariance matrix in the theory. For simplicity, we set the $\mathbf{O}$ matrices to identity matrix, which is also adopted in the previous studies [2, 13, 14].

Since the observational residues on original observational positions have been projected to the grid points in the preprocessing procedure in our new multigrid 3D-Var scheme, the bilinear interpolation operator $\mathbf{H}^{(n)}$ from model space to observational space is reduced to the identity matrix. Therefore the cost function of the multigrid 3D-Var takes the following form:

$$\begin{aligned} J^{(n)}[\mathbf{X}^{(n)}] &= \frac{1}{2}\gamma\mathbf{X}^{(n)T}\mathbf{S}^{(n)}\mathbf{X}^{(n)} \\ &+ \frac{1}{2}[\mathbf{X}^{(n)} - \mathbf{G}^{(n)}]^T\mathbf{O}^{*(n)-1}[\mathbf{X}^{(n)} - \mathbf{G}^{(n)}] \quad (7) \\ &(n = 1, 2, \dots, N). \end{aligned}$$

$\mathbf{G}$ is the observation projected in the state space. $\mathbf{O}^*$ is the observational error covariance matrix in state space. On one hand this preprocessing using the above mapping operator can effectively improve the condition number of the Hessian matrix of the cost function on every level grid because the observational residues on original observational positions have been projected to the analysis grid points. On the other hand it can restrain excessive incremental analysis on direction cross front and produce more reasonable analysis results, which will be discussed later. Therefore, a critical step for the theoretical success of this data assimilation scheme is the projection of observations given in the observational space to “pseudo observations” in the model space and the consistent specification of matrix $\mathbf{O}^*$.

Similar to the original multigrid 3D-Var, the new multigrid 3D-Var also starts analysis at its coarsest level. After $J^{(n-1)}$ is solved, $\mathbf{G}^{(n)}$ is computed with $\mathbf{X}^{(n-1)}$ and (1)–(3). Then $\mathbf{X}^{(n-1)}$ is projected onto the model grids with the operator $\mathbf{F}^{(n-1)}$. Last, $\mathbf{X}^{(n)}$ can be solved by minimizing $J^{(n)}$. This process is repeated until it reaches the finest grid level. During the analysis, only the half “V” cycle [18] property of the multigrid technique is applied as it continuously refines the resolutions. Then the final analysis is as follows [2, 13, 14]:

$$\mathbf{X}^a = \mathbf{X}^b + \sum_{n=1}^N \mathbf{F}^{(n)}\mathbf{X}^{(n)}. \quad (8)$$

From the above description, it is obvious that the new scheme [hereafter, we use “the new (old) scheme” to briefly denote the new (original) multigrid data assimilation scheme] also has multiscale analysis capability inherited from the old scheme, while the new scheme can make isopycnal data assimilation.

3. Data Analysis Experiments

In this section, we will conduct three data analysis experiments to assess the performance of this new scheme.

3.1. Idealized Experiment 1: Single Observation on a Front. The study domain in this experiment is a rectangular flat-bottom area which spans 115° – 135°E in longitude and 0° – 40°N in latitude with water depth being 500 m. From top to bottom, a south-north front is assumed to exist near 125°E . Figure 3 gives the 3D background temperature field, and to make the front clear, only the central longitudes of the horizontal map are shown (Figure 3(a)). For simplicity, salinity in the whole domain is set to be 35 psu homogeneously. One observation with a given increment 4°C compared to the background temperature is just located at point $(125^\circ\text{E}, 20^\circ\text{N}, 300\text{ m})$ in the front. The results of the old scheme are shown in

FIGURE 3: Background temperature field. From top to bottom a south-north front exists in 3-dimensional field. (a) Horizontal map; (b) cross front section; (c) along front section.

Figures 4(a) and 4(b). Figure 4(a) gives the analysis increment relative to the background. It can be seen that, without the constraint of potential density gradient, the increment forms a simple ellipsoid shape with three axial lengths defined by spatial correlation scales. It is obvious that the front structure is destroyed in the analysis field and the gradient of front decreases significantly (see Figure 4(b)). Figures 4(c) and 4(d) show the results of the new scheme. The increment mainly spreads along the potential density front while the cross front increment is squeezed (left panel of Figure 4(c)). The detailed structure can be seen in the vertical section (middle and right panels of Figure 4(c)). Unlike the vertical structure of increment in the old scheme which is simple ellipsoid, the new scheme vertical structure is mainly distributed along the surface of potential density, which is consistent with the isolines of the potential density very well. The increment is almost center-symmetric relative to the observational position and the vertical position of the increment on the cold side of the front is a little bit higher than that on the warm side. Importantly, the strength of the front is conserved and the location is shifted to the cold side slightly in the new scheme (Figure 4(d)), since the only one observational increment on the front has a positive value.

3.2. Idealized Experiment 2: Two-Point Observation on Both Sides of a Front. It is the same as experiment 1 but two observations are assimilated in this experiment. One observation is located at point (123°E, 10°N, 300 m) on the cold side of the front with negative increment -4°C , and the other is located at point (127°E, 30°N, 300 m) on the warm side of the front with positive increment 4°C . The results are shown in Figure 5 but both the center and right panels are of the cross front. Figures 5(a) and 5(b) give the results of old scheme. Horizontally, increments relative to the background are circles around observational points, and the increments on the cold (warm) side can be spread to the warm (cold) side through the front. Therefore the front is destroyed. But in new scheme (Figures 5(c) and 5(d)), the negative and positive increments are constraint to their own side not only horizontally but also vertically. And the strength of the front is conserved.

3.3. Real Data Analysis Experiment. Furthermore a realistic case is carried out in this experiment. Figure 6(a) shows the temperature background at 300 m depth of output of March 24, 2005, from our operational data assimilation and forecast system of China Ocean Reanalysis (CORA) [2, 19–21]. Figure 6(b) shows the in situ observational position

FIGURE 4: (a) Analyzed increment using old scheme (color shade); contour is isoline of potential density; (b) analysis using old scheme; (c) similar to (a) but for new scheme; (d) similar to (b) but for new scheme. Left: horizontal map; middle: cross front section; right: along front section.

FIGURE 5: (a) Analyzed increment using old scheme (color shade); contour is isoline of potential density; (b) analysis using old scheme; (c) similar to (a) but for new scheme; (d) similar to (b) but for new scheme. Left: horizontal map; middle: cross front section at 10°N; right: cross front section at 30°N.

FIGURE 6: (a) The background; (b) observational position (brown points) at 300 m; color shade is temperature increment of observation relative to background using Kriging interpolation; (c) analyzed temperature increment using old scheme; (d) similar to (c) but using new scheme; (e) vertical temperature increment profile at point (134.512°E, 14.454°N); black: observation, blue: result of old scheme, and red: result of new scheme; (f) zoom out of (e) between 0 and 500 m.

(brown points) at 300 m depth in this day and the color shade gives the increment of observation relative to the background using the Kriging interpolation. With the old scheme, the analysis increment relative to the background at 300 m depth is shown in Figure 6(c). As seen, the increment is very similar to that of the Kriging interpolation which does not take care of the fronts existing in the background and makes increments along and across these fronts. Figure 6(d) gives the result of the new scheme. The increments mainly spread along the isolines of the potential density of this layer, and cross front elements of increments are confined to very narrow areas. Figures 6(e) and 6(f) give the increment analysis result of one profile at a given observational point. It is obvious that vertically the old scheme only catches a roughly shape (blue line) of the observational profile (black line), while the new scheme gets more detailed structure, and the result (red line) is closer to the observation than that of the old scheme. In the multigrid 3D-Var, one can implement a full 3D analysis, and the cost function will be minimized to make the analysis field to fit observations as possible as it can in all three spatial directions. Therefore the multigrid 3D-Var just compromises horizontal and vertical analysis accuracy to fit observations. If one directional element, such as the cross front element, is squeezed, the minimization of the cost function will only focus on the other two directions, that is, along front and vertical directions, which makes it easier to fit observational profiles and obtains more detailed vertical information.

4. Twin Forecast Experiments

In this section, one-month twin forecast experiments are set up to further verify the validation of this new scheme in oceanic numerical forecast and examine whether the new scheme can provide more physically consistent initial conditions for model so as to improve model prediction quality.

4.1. Model. The model domain is 99° – 150° E in longitude and 10° S– 52° N in latitude including the Bohai Sea, the Yellow Sea, the East China Sea, the South China Sea, and adjacent seas. The numerical model used in this study is the Princeton Ocean Model with a generalized coordinate system (POMgcs). The model horizontal resolution is $1/4^{\circ}$. The hybrid vertical grid of sigma and z -level is used, with a total of 35 levels and a maximum depth of 5035 m for the configuration. In the area with the local depth being less than 200 m depth, the sigma-level grids are employed, while in other area the z -level grids are used with shaved cell applying to the model bottom grids.

As for the meteorological forcing the wind field is obtained from the US National Centers of Environmental Prediction (NCEP) reanalysis, with a resolution of $1.875^{\circ} \times 1.875^{\circ}$ in space and 6 hours in time. The bulk formula is used to calculate wind stress. The heat and water flux are also taken from NCEP reanalysis. There is no river runoff to be considered in this study. The sea level, temperature, salinity, and currents specified on the open boundary come from the SODA (Simple Ocean Data Assimilation) reanalysis datasets.

4.2. Experiment Setup. A one-month free model run using the above POMgcs from the CORA [19, 21–24] reanalysis datasets on March 16, 2005, as the initial conditions and forced by the above open boundary condition and meteorological forcing is applied to generate TRUTH run results from March 16 to April 16, 2005.

Synthetic temperature and salinity observations are constructed by interpolating the TRUTH run results to some preset positions spatially and temporally. Instead of using highly inhomogeneous real observational positions which are not fit for making it clear to analyze the advantage of the different data assimilation schemes in this study domain, spatially and temporally homogeneous positions are employed. The zonal and meridional distances between neighboring observational positions are 5-degree. The vertical resolution is the same as that of the model. And the synthetic observations are generated with daily frequency. To simulate observations as real as possible, the random error of 0.3°C is imposed to the synthetic temperature observation and random error of 0.02 psu is added to the synthetic salinity observation by using the random white-noise simulation.

Three experiments are setup to demonstrate the performance of this new scheme. EXP1 is a control run similar to the TRUTH run but starting from the climatological March T/S field with the currents and sea surface height being zero. EXP2 is 7-day forecast experiment similar to EXP1 but using the old scheme to assimilate the synthetic T/S observations to construct the initial field. EXP3 is similar to EXP2 but using the new scheme to assimilate the synthetic T/S observations. All three experiments span one month from March 16 to April 16, 2005.

In order to match the resolution of the synthetic observations, 6 level grids are applied to both schemes in EXP2 and EXP3 from $2 \times 2 \times 2$ (only one big cell containing the whole study domain) to $9 \times 9 \times 33$ (the horizontal interval is $6.375^{\circ} \times 7.75^{\circ}$, and the vertical resolution is almost the same as that of the model) which has a similar resolution with the synthetic observations.

4.3. Results. Figure 7 shows the first-day evolution of the temperature increments of EXP2 and EXP3 relative to the control run (EXP1) at 200 m depth, as well as the TRUTH run. Since the horizontal resolution of the synthetic observations is 5-degree, from this observational network only the information with wavelength not less than 5-degree can be resolved by the old scheme. Therefore, compared to the difference (we will call it truth) between the TRUTH run and the EXP1, the difference between EXP2 and the control run only catches some large scale (not less than 5-degree) information of the truth. However, by using the isopycnal mapping operator, some model dynamical information is effectively used to obtain some useful and additive small-scale and detailed information from the same coarse observational network; therefore lots of detailed structures are caught. Interestingly, after one month's data assimilation, the difference between EXP2 and EXP1 and the difference between EXP3 and EXP1 both gradually are close to the truth, and the latter is more and more alike the truth in detail, which can be seen in Figure 8. This can be interpreted as follows. The

FIGURE 7: First-day's evolution of temperature (top) and salinity (bottom) increment relative to control run (EXP1) at 200 m of the other two experiments (EXP2 and EXP3) and the TRUTH run. (a) and (d) for TRUTH run; (b) and (e) for the old scheme; (c) and (f) for the new scheme.

old scheme can retrieve the multiscale information from the given observational network and construct good analysis for further model integration. Unfortunately, the resolution of the given observational network is very coarse and can not provide small-scale and detailed information. However in the new scheme, besides the multiscale information obtained in the old scheme, using the model dynamical information, the flow-dependent analysis with some additive and useful small-scale and detailed structure can be obtained by the new scheme from the observational network, and this flow-dependent analysis is then input to the model as the initial field. Finally through the model integrating forward, we get more and more reasonable model forecast fields which can further provide more useful model dynamics for the next data assimilation cycle. Therefore the difference between EXP3 and EXP1 is close to the truth more quickly than that between EXP2 and EXP1.

All of the forecast results are compared to the TRUTH run results and the root mean square errors (RMSEs) are calculated. Figure 9 shows the time series of 7-day forecast RMSEs of the different experiments. The RMSE of EXP2 is smaller than that of EXP1, which can attribute to the multiscale analysis capability of the old scheme. The RMSE of EXP3 is the smallest one among all these experiments through the whole forecast experiment time, which can

attribute to the flow-dependent multiscale analysis capability of the new scheme.

Figure 10 gives the vertical distribution of 7-day forecast RMSEs of different experiments. Similar to Figure 9, the RMSE of EXP2 is smaller than that of EXP1 and the RMSE of EXP3 is the smallest one among all these experiments. For the temperature forecast the most improved layer of EXP3 compared to EXP2 occurs at about 100 m level, and for salinity forecast occurs at surface and about 200 m level. For one month average of the whole domain, the forecast temperature (salinity) RMSEs are 0.488°C (0.055 psu), 0.398°C (0.038 psu), and 0.360°C (0.033 psu) in EXP1, EXP2, and EXP3, respectively, which indicates that, compared to the control run, the old scheme can reduce the RMSE by about 18.4% for temperature forecast (31.2% for salinity forecast), while the new scheme can reduce RMSE by about 26.2% for temperature forecast (39.4% for salinity forecast); therefore the improvement of the new scheme over the old one is about 7.8% for temperature forecast (8.2% for salinity forecast).

5. Conclusions and Discussions

A new potential density gradient dependent analysis scheme for ocean data assimilation (ODA) has been designed and evaluated within a multigrid 3D-Var algorithm and a regional

FIGURE 8: Similar to Figure 7 but for one month's evolution.

FIGURE 9: Time series of 7-day forecast RMSEs of different experiments. Black: EXP1 (control run); blue: EXP2 (old scheme); red: EXP3 (new scheme).

ocean model. From data assimilation and model forecast experiments, three remarkable conclusions can be drawn. (1) The new scheme based on the model background potential density information can effectively inhibit the excessive spread of observation residue in the direction of across potential density face. (2) Due to uses of local density information, the new scheme produces more physically consistent analysis

increment structure in 3-dimensional space. (3) The new scheme provides more dynamically and physically consistent initial conditions and thus improves model forecast skills.

Although promising results are obtained in this idealized study, further research work still remains when the new scheme is applied to assimilate real observations. First, the flow-dependent mapping operator is generated from

FIGURE 10: Vertical distribution of 7-day forecast RMSEs of different experiments. Black: EXP1 (control run); blue: EXP2 (old scheme); red: EXP3 (new scheme).

the forecast model potential density fields, and thus the assimilation and forecast performances of this new scheme depend strongly on the accuracy of the model forecast. This feature is also illustrated in Cummings [4]. A possible hybrid multigrid 3D-Var will be considered in the future studies. Furthermore, using a set of ensemble forecasts to construct a flow-dependent background error covariance may further improve assimilation quality. Second, the threshold of potential density gradient used in this new scheme may be model dependent and need further examination when the new scheme is implemented into a new global or regional ocean model for a global or regional ocean reanalysis. In addition, when the number of observations is large or the number of analysis grids is large, the interpolating from observational to model grid points is indeed not trivial. It needs us to further improve our scheme in the future.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was jointly supported by grants of National Key R&D Program of China (2016YFC1401800, 2017YFC1404100, and 2017YFC1404102), National Natural Science Foundation (41376013, 41376015, 41306006, 41506039, 41606039, and 41775100), and National Programme on Global Change and Air-Sea Interaction (GASI-IPOVAI-04) of China.

References

- [1] Y. Xie, S. Koch, J. McGinley et al., "A sequential variational analysis approach for mesoscale data assimilation," in *Proceedings of the 21st Conf. on Weather Analysis and Forecasting/17th Conf. on Numerical Weather Prediction*, Washington, D.C., USA, 2005.

- [2] W. Li, Y. F. Xie, Z. J. He et al., "Application of the multi-grid data assimilation scheme to the China seas' temperature forecast," *Journal of Atmospheric and Oceanic Technology*, vol. 25, no. 11, pp. 2106–2116, 2008.
- [3] Y. Xie, S. Koch, J. McGinley et al., "A space-time multiscale analysis system: a sequential variational analysis approach," *Monthly Weather Review*, vol. 139, no. 4, pp. 1224–1240, 2011.
- [4] J. A. Cummings, "Operational multivariate ocean data assimilation," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, pp. 3583–3604, 2005.
- [5] G. Burgers, M. A. Balmaseda, F. C. Vossepoel, G. J. Van Oldenborgh, and P. J. Van Leeuwen, "Balanced ocean-data assimilation near the equator," *Journal of Physical Oceanography*, vol. 32, no. 9, pp. 2509–2519, 2002.
- [6] A. Troccoli and K. Haines, "Use of the temperature-salinity relation in a data assimilation context," *Journal of Atmospheric and Oceanic Technology*, vol. 16, no. 12, pp. 2011–2025, 1999.
- [7] J. O. S. Alves, K. Haines, and D. L. T. Anderson, "Sea level assimilation experiments in the tropical Pacific," *Journal of Physical Oceanography*, vol. 31, no. 2, pp. 305–323, 2001.
- [8] A. Troccoli, M. A. Balmaseda, J. Segschneider et al., "Salinity adjustments in the presence of temperature data assimilation," *Monthly Weather Review*, vol. 130, no. 1, pp. 89–102, 2002.
- [9] K. Haines, J. D. Blower, J.-P. Drecourt et al., "Salinity assimilation using S(T): Covariance relationships," *Monthly Weather Review*, vol. 134, no. 3, pp. 759–771, 2006.
- [10] F. L. Yin and E. S. Sarachik, "An efficient convective adjustment scheme for ocean general circulation models," *Journal of Physical Oceanography*, vol. 24, no. 6, pp. 1425–1430, 1994.
- [11] P. C. Chu and C. Fan, "A conserved minimal adjustment scheme for stabilization of hydrographic profiles," *Journal of Atmospheric and Oceanic Technology*, vol. 27, no. 6, pp. 1072–1083, 2010.

- [12] X. Wang, P. C. Chu, G. Han, W. Li, X. Zhang, and D. Li, "A fully conserved minimal adjustment scheme with (T, S) coherency for stabilization of hydrographic profiles," *Journal of Atmospheric and Oceanic Technology*, vol. 29, no. 12, pp. 1854–1865, 2013.
- [13] W. Li, Y. F. Xie, S.-M. Deng, and Q. Wang, "Application of the multigrid method to the two-dimensional Doppler radar radial velocity data assimilation," *Journal of Atmospheric and Oceanic Technology*, vol. 27, no. 2, pp. 319–332, 2010.
- [14] W. Li, Y. Xie, and G. Han, "A theoretical study of the multigrid three-dimensional variational data assimilation scheme using a simple bilinear interpolation algorithm," *Acta Oceanologica Sinica*, vol. 32, no. 3, pp. 80–87, 2013.
- [15] A. T. Weaver, J. Vialard, and D. L. T. Anderson, "Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal diagnostics, and consistency checks," *Monthly Weather Review*, vol. 131, no. 7, pp. 1360–1378, 2003.
- [16] M. J. Carrier and H. Ngodock, "Background-error correlation model based on the implicit solution of a diffusion equation," *Ocean Modelling*, vol. 35, no. 1-2, pp. 45–53, 2010.
- [17] M. Yaremchuk, M. Carrier, S. Smith, and G. Jacobs, "Background error correlation modeling with diffusion operators," *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II)*, pp. 177–203, 2013.
- [18] W. L. Briggs, V. E. Henson, and S. McCormick, *A Multigrid Tutorial*, Society for Industrial and Applied Mathematics, 2nd edition, 2000.
- [19] G. J. Han, W. Li, X. F. Zhang et al., "A new version of regional ocean reanalysis for coastal waters of China and adjacent seas," *Advances in Atmospheric Sciences*, vol. 30, no. 4, pp. 974–982, 2013.
- [20] H. Fu, P. C. Chu, G. Han, Z. He, W. Li, and X. Zhang, "Improvement of short-term forecasting in the northwest Pacific through assimilating Argo data into initial fields," *Acta Oceanologica Sinica*, vol. 32, no. 7, pp. 57–65, 2013.
- [21] G. J. Han, W. Li, X. F. Zhang et al., "A regional ocean reanalysis system for coastal waters of China and adjacent seas," *Advances in Atmospheric Sciences*, vol. 28, no. 3, pp. 682–690, 2011.
- [22] X. Wang, G. Han, Y. Qi, and W. Li, "Impact of barrier layer on typhoon-induced sea surface cooling," *Dynamics of Atmospheres and Oceans*, vol. 52, no. 3, pp. 367–385, 2011.
- [23] X. F. Zhang, G. J. Han, D. X. Wang, Z. Deng, and W. Li, "Summer surface layer thermal response to surface gravity waves in the Yellow Sea," *Ocean Dynamics*, vol. 62, no. 7, pp. 983–1000, 2012.
- [24] T. Yu, Z. Deng, G. Han, X. Wu, H. Fu, and K. Wu, "The Reanalysis of Currents and Throughflow Volume Transport in the Taiwan Strait," *Marine Geodesy*, vol. 35, no. 1, pp. 16–31, 2012.