

Application of the Multigrid Data Assimilation Scheme to the China Seas' Temperature Forecast

WEI LI

College of Physical and Environmental Oceanography, Ocean University of China, Qingdao, and National Marine Data and Information Service, State Oceanic Administration, Tianjin, China

YUANFU XIE

NOAA/Earth System Research Laboratory, Boulder, Colorado

ZHONGJIE HE

College of Physical and Environmental Oceanography, Ocean University of China, Qingdao, and National Marine Data and Information Service, State Oceanic Administration, Tianjin, China

GUIJUN HAN, KEXIU LIU, JIRUI MA, AND DONG LI

National Marine Data and Information Service, State Oceanic Administration, Tianjin, China

(Manuscript received 6 July 2006, in final form 2 February 2008)

ABSTRACT

Correlation scales have been used in the traditional scheme of three-dimensional variational data assimilation (3DVAR) to estimate the background (or first guess) error covariance matrix (the \mathbf{B} matrix in brief) for the numerical forecast and reanalysis of ocean for decades. However, it is challenging to implement this scheme. On the one hand, determining the correlation scales accurately can be difficult. On the other hand, the positive definite of the \mathbf{B} matrix cannot be guaranteed unless the correlation scales are sufficiently small. Xie et al. indicated that a traditional 3DVAR only corrects certain wavelength errors, and its accuracy depends on the accuracy of the \mathbf{B} matrix. Generally speaking, the shortwave error cannot be sufficiently corrected until the longwave error is corrected. An inaccurate \mathbf{B} matrix may mistake longwave errors as shortwave ones, resulting in erroneous analyses.

A new 3DVAR data assimilation scheme, called a multigrid data assimilation scheme, is proposed in this paper for quickly minimizing longwave and shortwave errors successively. By assimilating the sea surface temperature and temperature profile observations into a numerical model of the China Seas, this scheme is applied to a retroactive real-time forecast experiment and favorable results are obtained. Compared to the traditional scheme of 3DVAR, this new scheme has higher forecast accuracy and lower root-mean-square errors. Note that the new scheme demonstrates greatly improved numerical efficiency in the analysis procedure.

1. Introduction

The ocean is not observed frequently enough in space or time to allow for a direct and reasonably accurate description of the large-scale oceanic state and its variability. Such limitations can be improved by ap-

plying a data assimilation scheme that makes use of observation and background information in space and time as well as physics constraints. Many ocean data assimilation schemes have been substantially developed since the mid-1980s (Derber and Rosati 1989; Behringer et al. 1998; Gaspari and Cohn 1999; Weaver and Courtier 2001).

A three-dimensional variational data assimilation (3DVAR) technique could usually be accomplished by using correlation scales (Derber and Rosati 1989) to form the background (or first guess) error covariance

Corresponding author address: Jirui Ma, National Marine Data and Information Service, 93 Liuwei Road, Hedong District, Tianjin 300171, China.
E-mail: jrma@mail.nmdis.gov.cn

matrix (hereafter, **B** matrix), which plays an important role in data assimilation by determining the spatial spreading of observational information. However, a real field at different locations may have different correlation scales which are flow dependent and difficult to estimate. In addition, the **B** matrix cannot be guaranteed to be positive definite numerically unless the correlation scales are small enough. Another method to estimate the **B** matrix involves using a recursive filter (Hayden and Purser 1995), which can ensure that the **B** matrix is positive definite. In the context of variational analysis, the recursive filter is often interpreted as a covariance function of background errors. However, the traditional 3DVAR, using either correlation scales or recursive filters, can only correct certain wavelength errors (Xie et al. 2005). 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 in turn, a new 3DVAR data assimilation scheme, called the multigrid data assimilation scheme, is proposed in this paper. The multigrid technique is often used to solve numerical differential equation by allowing long waves to converge faster than short ones (Briggs et al. 2000). In the variational data assimilation, the multigrid technique also allows longwave errors to be corrected faster than short ones. This could prevent the longwave errors from being incorporated into shortwave analyses. For example, at observation stations in one area, observation errors contain a 1000-km wavelength error component and a 100-km wavelength error component. Before removing the 1000-km wavelength error from the observation, if an incorrect estimation of correlation scales suggests a 100-km correlation scale, the traditional 3DVAR will provide an analysis meeting the observation by taking the 1000-km wavelength error into the 100-km scale correction. However, a multigrid data assimilation scheme does not allow this to happen because the 100-km wavelength error is never corrected until the 1000-km wavelength error is removed from the observation, and it can thus provide a better or more accurate analysis. In this paper, the multigrid data assimilation scheme is applied to assimilate SST and temperature profiles of the China Seas into a numerical model in a retroactive real-time forecast experiment. A comparison of the results to those of the traditional 3DVAR using correlation scales is presented.

In the following section, the theory and verification of the multigrid data assimilation scheme are introduced. The numerical model and observational data

used in the retroactive real-time forecast experiments are described in sections 3a and 3b, respectively. In section 3c, the results of the forecast experiments are presented. A summary and conclusions are presented in section 4.

2. Theory and verification

Taking the standard 3DVAR cost functional form, which includes a background and observational term, operating on the field of temperature correction instead of the temperature itself (Derber and Rosati 1989), the cost functional to be minimized is

$$J = \frac{1}{2} \mathbf{X}^T \mathbf{B}^{-1} \mathbf{X} + \frac{1}{2} (\mathbf{H}\mathbf{X} - \mathbf{Y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{X} - \mathbf{Y}), \quad (1)$$

where \mathbf{X} is the correction of temperature referred to the background, \mathbf{B} is the background error covariance matrix, \mathbf{Y} is the difference between the observations and the interpolated background temperature at the observation locations, \mathbf{R} is the observation error covariance matrix, and \mathbf{H} is a simple bilinear interpolation operator from model space to observation space.

In a correlation-scale method, correlation scales are critically important in the estimation of a **B** matrix, which plays a major role in data assimilation methods such as optimal interpolation (OI) or 3DVAR. Derber and Rosati (1989) adopted an empirical formula to calculate the **B** matrix in which the correlation scales are dependent only on latitude. Behringer et al. (1998) modified this formula by considering the anisotropy in the zonal and meridional directions. Gaspari and Cohn (1999) discussed the construction of covariance functions for data assimilation. Weaver and Courtier (2001) described a practical algorithm that can be used to model correlation functions on the sphere. A wide range of functions can be adopted to determine the correlation with spatial distance. The commonly used formula is a Gaussian-type function in which the correlations are weighted by the separate latitudinal and longitudinal distances and/or depths. In this study, the element **B** takes the following form (Derber and Rosati 1989; Behringer et al. 1998; Zhou et al. 2004):

$$\mathbf{B}_{i,j} = a_h \exp\left(-\frac{\Delta x_{ij}^2}{L_x^2} - \frac{\Delta y_{ij}^2}{L_y^2}\right), \quad (2)$$

where L_x and L_y are characteristic of length scales that reflect the extent of spatial correlation, x and y are model coordinates, and a_h is the first-guess error variance.

The distribution of observations in the ocean is highly inhomogeneous, and inaccurate correlation

scales could result in more errors for regions with sparse observations. For a given observation system, data-sparse regions could provide longwave information and data-dense regions could provide both longwave and relative shortwave information. An ideal data assimilation scheme would retrieve longwave information over the whole domain and shortwave information over data-dense regions. One way to realize this idea is to obtain an accurate flow-dependent \mathbf{B} matrix, which is practically impossible. The other way is to retrieve these waves by a sequence of 3DVAR and combine these waves together for the final analysis. This has to be done from longwave to shortwave. Otherwise, the shortwave information will meet the observation perfectly and mistakenly destroy the longwave information.

For instance, a temperature field can be separated into longwave information and shortwave information using the Fourier theory

$$\mathbf{X}^b = \mathbf{X}_L^b + \mathbf{X}_S^b \quad \text{and} \quad (3)$$

$$\mathbf{X}_L = \mathbf{X}_L^a - \mathbf{X}_L^b, \quad (4)$$

where the subscript L represents the longwave information and S the shortwave information. The superscript b represents the background field and a the analysis field. The longwave and shortwave errors can be minimized in turn and the cost functional for retrieving longwave information can be modified as follows:

$$J = \frac{1}{2} \mathbf{X}_L^T \mathbf{B}_L^{-1} \mathbf{X}_L + \frac{1}{2} (\mathbf{H} \mathbf{X}_L - \mathbf{Y}_L)^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{X}_L - \mathbf{Y}_L). \quad (5)$$

The analysis of temperature becomes

$$\mathbf{X}^a = \mathbf{X}^b + \mathbf{X}_L = \mathbf{X}_L^b + \mathbf{X}_S^b + \mathbf{X}_L^a - \mathbf{X}_L^b = \mathbf{X}_L^a + \mathbf{X}_S^b, \quad (6)$$

which means that the final analysis is a combination of the analyzed longwave information obtained from the observations and the shortwave information obtained from the background field. A functional for relative shortwave information can be iteratively formed by keeping the relative longwave information \mathbf{X}_L . The multigrid data assimilation scheme will be applied to correct the longwave and shortwave errors successively.

A multigrid technique is initially developed for differential equations (Briggs et al. 2000). To solve differential equations numerically, a finite difference scheme is usually used. However, the oscillatory modes of the error are eliminated effectively, whereas smooth modes are corrected very slowly for many numerical iterative methods. To overcome such limitations, the multigrid

technique is developed because coarse grids can be used to speed up the convergence of smooth modes (i.e., longer waves over a coarser grid). Minimizing a functional over a coarse grid to obtain longwave information and a functional over a relative fine grid for shortwave information could fit the ideal data assimilation scheme discussed above. By introducing the multigrid technique into the 3DVAR data assimilation, called the multigrid data assimilation scheme hereafter, the cost functional should take the following form:

$$J^{(n)} = \frac{1}{2} \mathbf{X}^{(n)T} \mathbf{X}^{(n)} + \frac{1}{2} [\mathbf{H}^{(n)} \mathbf{X}^{(n)} - \mathbf{Y}^{(n)}]^T \mathbf{R}^{(n-1)} [\mathbf{H}^{(n)} \mathbf{X}^{(n)} - \mathbf{Y}^{(n)}], \quad (7)$$

$(n = 1, 2, 3, \dots, N),$

where the superscript n represents the n th level grid and N is the final level (which depends on the observations' distribution).

Let $\mathbf{Y}^{(1)} = \mathbf{Y}^{\text{obs}} - \mathbf{H} \mathbf{X}^b$ in the first-level grid be the difference between the observations and the interpolated background temperature at the observation locations, and in the other grid levels it is defined as

$$\mathbf{Y}^{(n)} = \mathbf{Y}^{(n-1)} - \mathbf{H}^{(n-1)} \mathbf{X}^{(n-1)} \quad (n = 2, 3, \dots, N). \quad (8)$$

Here, coarse grids correspond to smooth modes (longwave information) and fine grids correspond to oscillatory modes (shortwave information). The \mathbf{B} matrix is simplified to an identity matrix because the correlation scales are reflected by coarse or fine grids. During the procedure of sequential multiple-scale analysis, the half V cycle (Briggs et al. 2000) property of the multigrid technique is employed as it continuously refines the resolutions. The final analysis is

$$\mathbf{X}^a = \mathbf{X}^b + \mathbf{X}_L = \mathbf{X}^b + \sum_{n=1}^N \mathbf{X}^{(n)}. \quad (9)$$

An idealized experiment is performed to verify the validity of this new data assimilation scheme. The model domain covers a square region, extending over 30° – 40° N in latitude and 100° – 110° E in longitude (Fig. 1). Figure 1 shows the distribution of observations (Fig. 1a), which includes 600 random points, and the profile of the true temperature field (Fig. 1b). The true temperature field, simulating a warm front, is depicted in Fig. 2f.

In the traditional 3DVAR using correlation scales, the cost functional is minimized by the preconditioned conjugate gradient algorithm via an iterative procedure

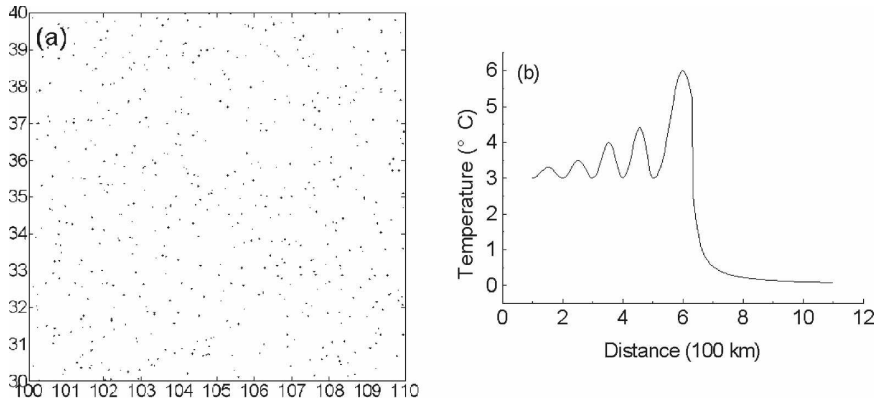


FIG. 1. The distribution of (a) 600 random observations and (b) the true profile.

(Navon and Legler 1987; Derber and Rosati 1989). The analyses by 3DVAR using different correlation scales—500, 200, 100, and 50 km, respectively—are shown in Fig. 2. The results demonstrate that the larger the correlation scale is, the longer the corrected error

wavelength becomes. In the multigrid data assimilation scheme, five-level grids with a grid ratio of 0.5 [i.e., the grid spacing in the n th level grid is half of that in the $(n - 1)$ th-level grid] are employed, ranging from $5^\circ \times 5^\circ$ to $0.3125^\circ \times 0.3125^\circ$. Figure 2e gives the result of the

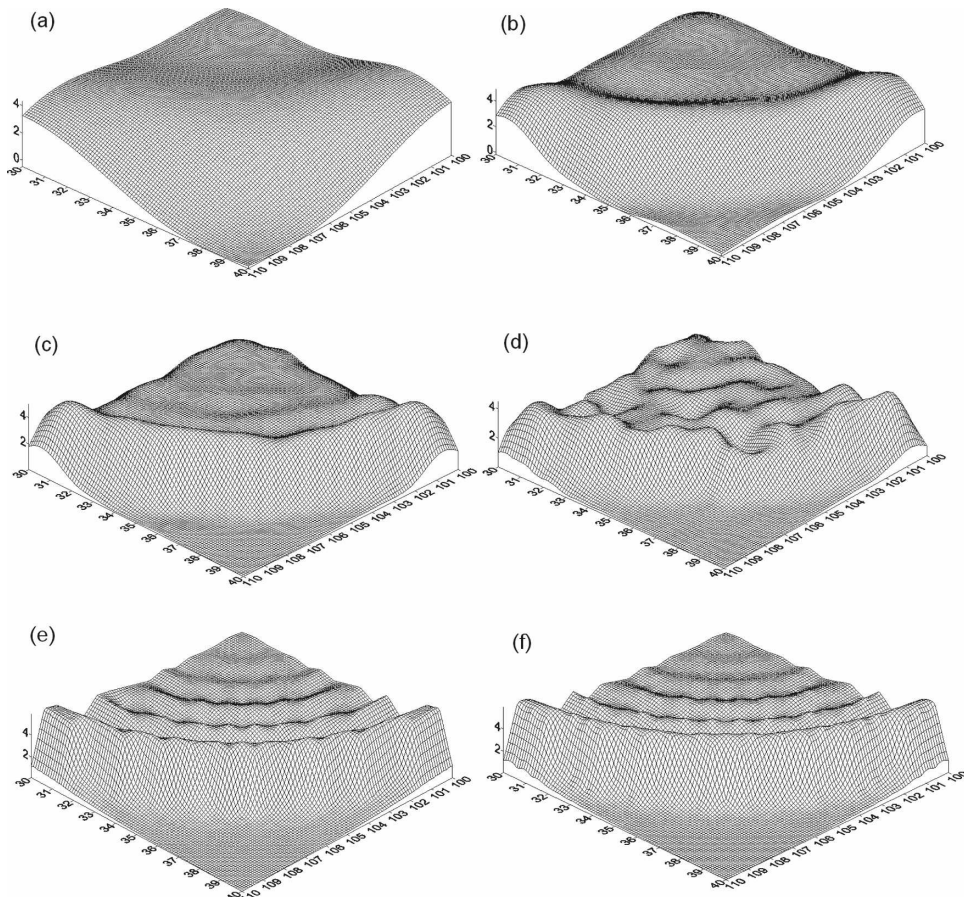


FIG. 2. Analysis results by using 3DVAR of different correlation scales: (a) 500, (b) 200, (c) 100, and (d) 50 km. (e) The analysis result of the multigrid data assimilation scheme; (f) the true temperature field.

TABLE 1. The RMS difference between the analysis results of the multigrid data assimilation scheme (short for multigrid) and the traditional 3DVAR using correlation scales (short for correlation scales) and the true temperature field.

Scheme		RMS ($^{\circ}\text{C}$)
Multigrid		9.28×10^{-6}
Correlation scales	500 km	0.845
	200 km	0.570
	100 km	0.474
	50 km	0.957

multigrid data assimilation scheme, which is better than the former ones, compared to the true field in Fig. 2f. Table 1 shows the RMS difference of different methods.

3. Retroactive real-time forecast experiment

Section 2 has demonstrated that the proposed multigrid data assimilation scheme can give much higher accuracy compared to the traditional 3DVAR using correlation scales in the idealized experiment. To examine the performance of the new scheme in numerical forecasts of temperature, a retroactive real-time forecast experiment is carried out as described in this section.

a. The numerical model

The numerical model used in the following retroactive real-time forecast experiment is a coastal ocean circulation model based on the Princeton Ocean Model with generalized coordinate system (POM; POMgcs) in which sigma- and/or z -level vertical grids can be chosen (Ezer and Mellor 2004). This is a fully nonlinear, prognostic model incorporating the free surface and the

Mellor and Yamada (1974, 1982) level-2.5 turbulent closure scheme for vertical mixing. The z -level coordinate system of POMgcs is employed in this paper.

The model domain covers the China Seas, including the Pohai (or Bohai) Sea, the Yellow Sea, the East China Sea, and the South China Sea (Fig. 3c), and an adjacent sea area extending 10°S – 41°N in latitude and 99° – 142°E in longitude. The bathymetry employed in the simulations is based on 5-min gridded elevations/bathymetry for the world (ETOPO5). The bathymetry values provided by ETOPO5 in shallow regions (e.g., the Pohai Sea) are found to be highly questionable. Therefore, the bathymetry of these regions is modified by local nautical charts to obtain a more realistic coastline. The model grid spacing is varied from $1/12^{\circ}$ to $1/2^{\circ}$, which produces a moderate number of computational points in the model domain, thereby reducing the computation time. There are 20 vertical level of which 13 are above 450 m. In the retroactive real-time forecast experiment, the model is forced by wind stress and air temperature reanalysis products from the National Centers for Environmental Prediction (NCEP), and heat flux is calculated using bulk formulas. The open boundary conditions of currents are provided by a global model. Orlanski radiation conditions are employed in open boundary conditions for temperature, salinity, and sea surface height, which are relaxed to those of the global model.

b. Observational data processing

The data type, accuracy, and distribution in space and time govern the quality of the results produced by the data assimilation system. If all predicted variables

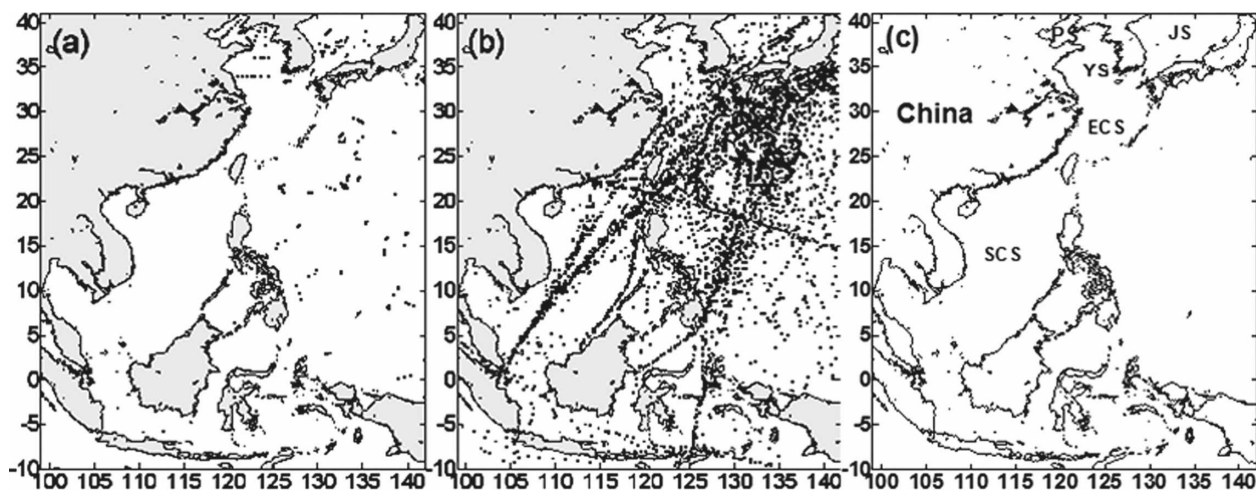


FIG. 3. The distribution of (a) temperature profiles and (b) shipboard SST in August 2004. (c) Place names used in this paper. PS: Pohai Sea; YS: Yellow Sea; ECS: East China Sea; SCS: South China Sea; JS: Japan Sea.

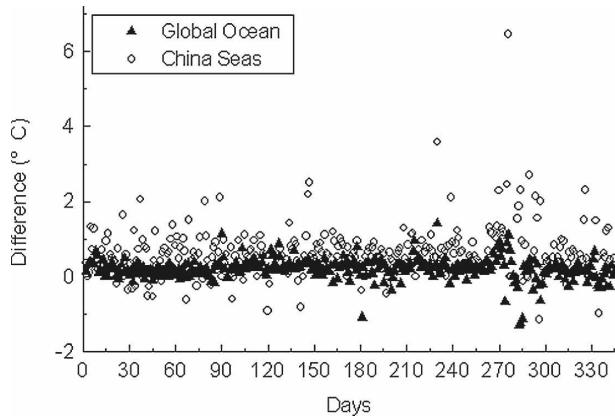


FIG. 4. The deviation of AVHRR SST from shipboard SST in the period of the experiment. The triangle represents the global ocean; the circle is for the China Seas.

were observed perfectly and continuously in space and time, there would be little need for a data assimilation system as long as the numerical model could handle the shortest waves. Unfortunately, this is not the case. The observational system, which is highly inhomogeneous, typically covers large regions either without observations or without measuring all of the forecast variables. Thus, in this study we would like to use as many types of observations as possible. The conventional temperature observations used in the following experiments consist of shipboard SST and temperature profiles including expendable bathythermograph (XBT) ocean station data and conductivity–temperature–depth (CTD) data, obtained from the Global Temperature and Salinity Profile Project (GTSP). A typical

monthly distribution (August 2004) of temperature profiles and shipboard SST is shown in Fig. 3. The unconventional temperature observations used consist of satellite remote sensing SST of Advanced Very High Resolution Radiometer (AVHRR) Pathfinder Version 5.0 (hereafter, AVHRR SST in brief) and Argo profiles.

The shipboard SST and temperature profiles contain observations with errors, making it necessary to perform quality control on them. For example, the shipboard SST observations are grossly checked by omitting the data with deviations larger than 3.5σ . Here, σ is the standard deviation statistically obtained from several decades' observations. A similar gross check is also done for temperature profiles. After that, the temperature profiles are vertically interpolated to the top 13 model levels. No information is inserted below level 13 because of the lack of data and our focus being primarily in the upper ocean.

As for AVHRR SST, the quality control procedure consists of four steps:

Step 1: The dataset undergoes a gross check by omitting the SST data with quality flags less than 3. (The overall quality flag is a relative assignment of SST quality based on a hierarchical suite of tests. The quality flag varies from 0 to 7, with 0 being the lowest quality and 7 the highest. See <http://www.nodc.noaa.gov/SatelliteData/pathfinder4km/userguide.html>). Then the SST data are averaged in each $\frac{1}{2}^\circ \times \frac{1}{2}^\circ$ bin.

Step 2: Compared with the monthly averaged shipboard SST, the AVHRR SSTs which have a devia-

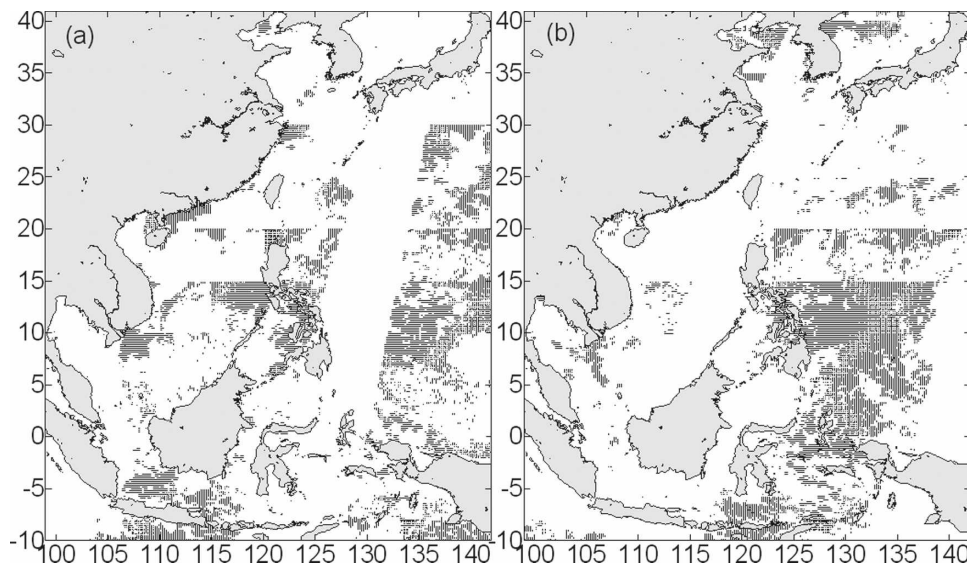


FIG. 5. The distribution of AVHRR SST data on (a) 19 Apr and (b) 24 Apr 2004.

TABLE 2. The RMS difference of temperature forecasts using the traditional 3DVAR with different correlation scales.

Correlation scales		SST (°C)	Temp profiles (°C)
L_x (km)	L_y (km)		
1000	500	1.58	1.72
500	250	1.53	1.68
240	70	1.51	1.46
100	50	1.54	1.61
10	5	1.93	1.88

tion larger than the standard deviation of the shipboard SST are removed. The deviation of the remaining AVHRR SST from shipboard SST is shown in Fig. 4.

Step 3: The AVHRR SST is further controlled by omitting AVHRR SSTs that have deviations from the daily averaged shipboard SST larger than 1.5°C. About 1/3 of the data that have passed step 1 are omitted in steps 2 and 3. The remaining AVHRR SST data can still cover the whole study domain and are shown in Fig. 5.

Step 4: The remaining AVHRR SSTs are adjusted by making the daily averaged AVHRR SST equal to that of the shipboard SST in the whole domain. For the global ocean, the RMS differences between shipboard SST and AVHRR SST are 1.05° and 0.83°C, respectively, before and after such adjustment, whereas for the China Seas, the RMS differences are 1.20° and 1.16°C, respectively.

c. *Retroactive real-time forecast experiment and results*

The retroactive real-time forecast experiment includes three subexperiments, denoted Emodel, ETV,

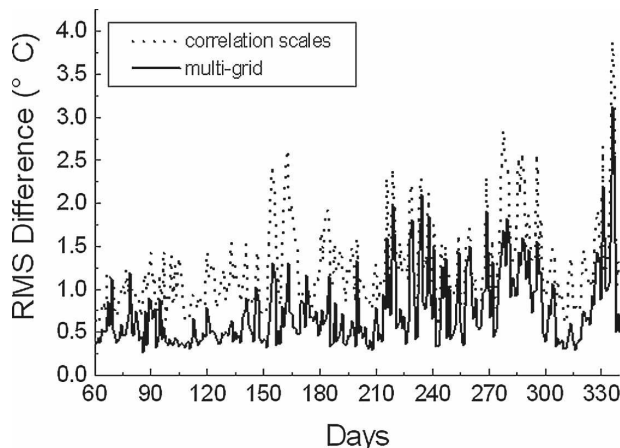


FIG. 6. Time series of RMS difference between analyzed and observed SST.

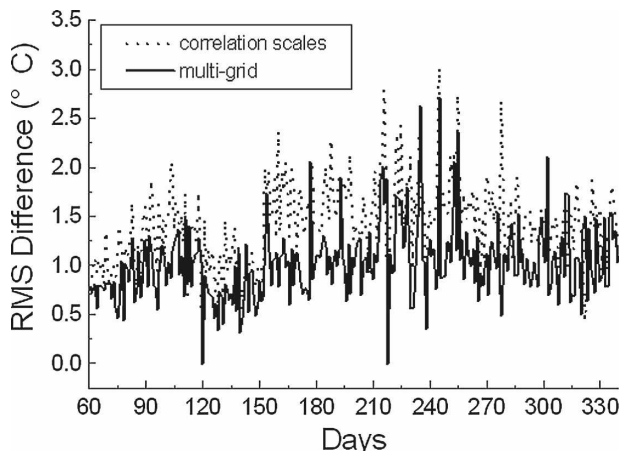


FIG. 7. Time series of RMS difference between the analyzed profiles and observations.

and EMG: EMG is the proposed model, Emodel is the model run without assimilation, and ETV employs the traditional 3DVAR data assimilation using correlation scales.

Similar to the traditional 3DVAR scheme implemented by Derber and Rosati (1989), data assimilation is performed on each model level with the vertical correlations ignored for both ETV and EMG. In ETV, the **B** matrix takes the form in Eq. (2). The correlation scales $L_x = 240$ km and $L_y = 70$ km in the China Seas are selected from tuning experiments to optimize the results (see Table 2 for temperature forecasts results using the traditional 3DVAR with some selected different correlation scales for demonstration). In EMG, five-level grids are employed, ranging from about $10^\circ \times 10^\circ$ to the model horizontal resolution. A grid ratio of 0.5 is employed in the first four levels, and the horizontal resolution of last grid is the same as that of model

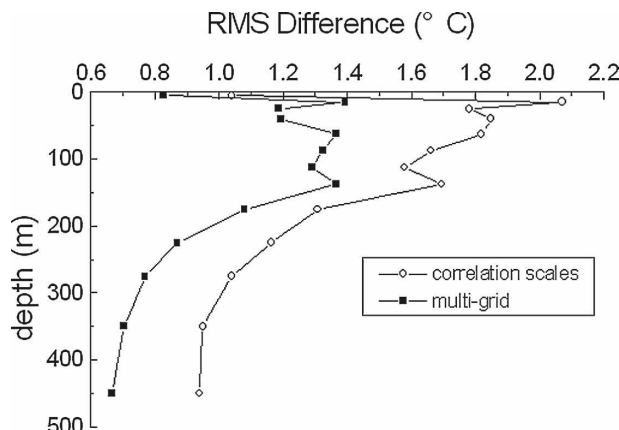


FIG. 8. The vertical structure of the RMS difference between the analyzed profiles and observations.

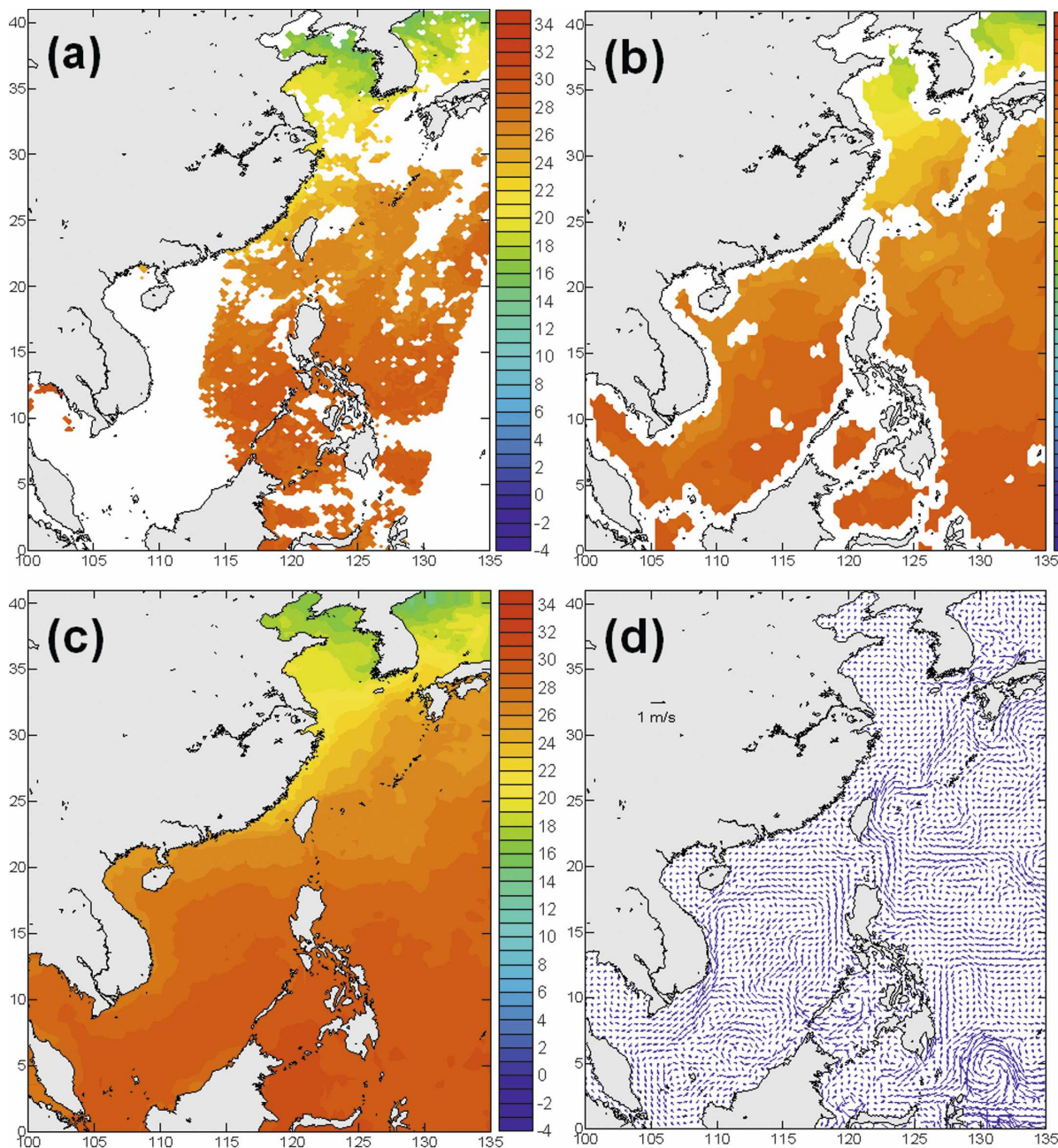


FIG. 9. (a) AVHRR SST observations, (b) AMSR-E SST observations, (c) SST forecast, and (d) surface circulation forecast using the multigrid data assimilation scheme on 31 Oct 2004. Note that AMSR-E was not assimilated in this experiment.

grid. At each stage of the sequential multiple-scale analyses in EMG, the procedure is similar to ETV, except that the **B** matrix is implemented as an identity matrix. The observation error covariance matrices used in the two schemes, which are assumed to be diagonal, are the same.

The model is initialized by using the monthly average (January) temperatures and salinities from the National Marine Data and Information Service (NMDIS) of the

State Oceanic Administration of the People’s Republic of China and is spun up over 1 yr in the diagnostic mode with temperature and salinity held to reach a steady state. The observational data of January and February 2004 processed in section 3b are then assimilated into the model. After the 2-month run, the initial field of the forecast is obtained. Three sets of 72-h forecasts, each of which spans 10 months from March to December 2004, are carried out afterward.

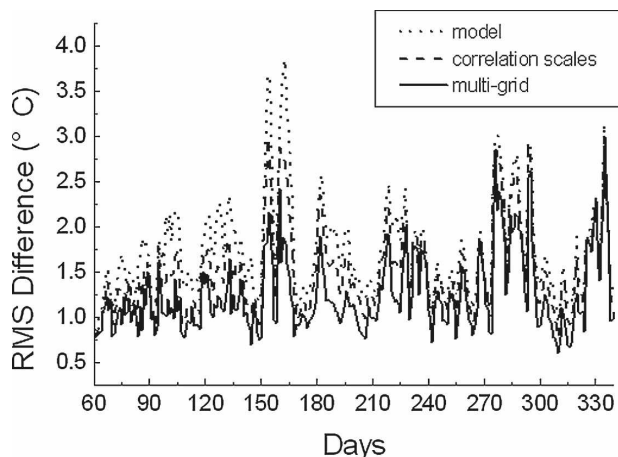


FIG. 10. Time series of RMS difference between SST forecasts and observed SST.

The analyses from the assimilation are compared to the observations. Because the observations are used in the assimilation, this comparison is a measure of the assimilation technique’s ability to force the model solution toward the data. Figure 6 depicts the RMS differences between the analyzed SST and observations; Fig. 7 depicts the RMS difference between the analyzed temperature profiles and observations. Figure 8 is the vertical structure of the RMS differences between the analyzed temperature profiles and observations. In Figs. 6–8, the RMS difference in EMG is the smallest, which indicates that the multigrid data assimilation scheme can provide better analysis.

Figure 9 shows AVHRR SST observations, AMSR-E SST observations, the forecast SST, and the forecast surface circulation derived by using the multigrid data assimilation scheme on 10 November 2004. The forecasts of SST coincide well with AVHRR SST (assimilated) and the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E SST), which is not assimilated in this experiment. In Fig. 9d, the main current systems are well presented, for example, Kuroshio and its extension, a southward coastal jet off the west coast of Kyushu as part of a Kuroshio meander, the Tsushima Warm Current, the Ryukyu Current, the China Coastal Current, the Taiwan Warm Current entering the East China Sea from Taiwan Strait while turning anticyclonically, the Mindanao Current, and the circulation in South China Sea. Figure 10 shows the time series of RMS differences between forecasts and SST observations. The RMS differences curve for the multigrid data assimilation scheme displays obvious improvement. For SST observations, the forecast RMS error is 1.51°C in ETV and 1.21°C in EMG, which indicates that the new scheme has higher

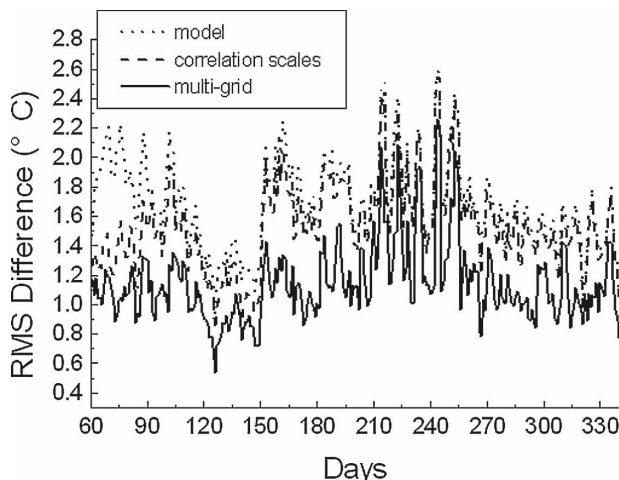


FIG. 11. Time series of RMS difference between the profile forecasts and observations.

forecast accuracy and a smaller forecast RMS error by about 0.30°C, and the rate of improvement is about 19.9%.

Figure 11 shows the time series of RMS differences between the forecasts and profile observations. The RMS differences between the forecasts of the multigrid data assimilation scheme and profile observations are less than those of the other two forecasts, indicating the advantage of the new method. Figure 12 shows the vertical structure of the RMS differences. For profile observations, the forecast RMS error is 1.46°C in ETV and 1.06°C in EMG, which indicates that the new scheme has higher forecast accuracy and a lower forecast RMS error by about 0.40°C, and the improvement is about 27.4%.

Figure 13 displays the spatial distribution of RMS

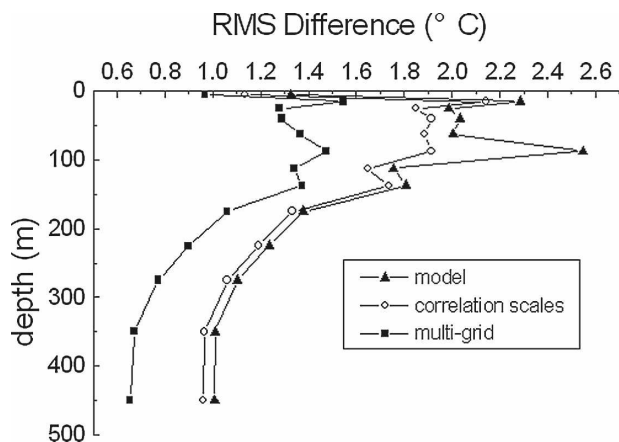


FIG. 12. The vertical structure of the RMS difference between the profile forecasts and observations.

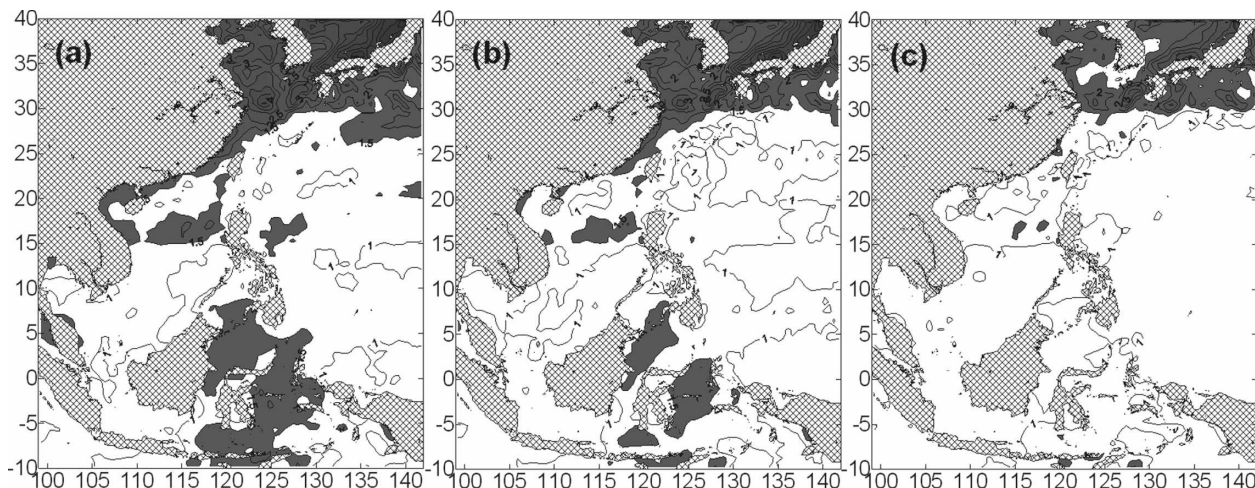


FIG. 13. Spatial distribution of RMS difference of SST forecasts between (a) Emodel, (b) ETV, and (c) EMG and the observed SST. The contour interval is 0.5°C and the black shading is for values greater than 1.5°C.

differences of SST in these retroactive real-time forecasts. In Emodel, large RMS differences appear almost everywhere. RMS differences are reduced only in some regions in ETV; however, most of the RMS differences are reduced in EMG, especially in the South China Sea and the northwestern Pacific Ocean. The possible reasons for regions where RMS differences are large in both EMG and ETV are varied. For instance, in the area near the southern and northern boundaries, including the southern part of the Japan Sea, the RMS differences could be caused by the poor open boundary conditions that should be improved in further studies; inhomogeneous distribution of observations might be another reason. As for the Pohai Sea, the Yellow Sea, and part of the East China Sea, there are some impor-

tant near-shore fronts, such as the Changjiang River plume and tidal fronts. The grid resolution in this area is about $\frac{1}{2}^\circ \times \frac{1}{4}^\circ$, which cannot simulate these fronts very well. The large RMS differences may partly be the result of coarse model grid resolution. Another reason could be that the observations are relatively sparse in the above region. In the Kuroshio extension, the large RMS differences for SST forecasts also can perhaps be attributed to poor distribution of SST observations. The vertical average RMS differences of the three forecast results are drawn in Fig. 14. It can be observed in Fig. 14 that another area with large RMS difference is the Mindanao Current, where there is an eddy; the position cannot be well simulated by the model. Although all the above-mentioned large RMS difference areas

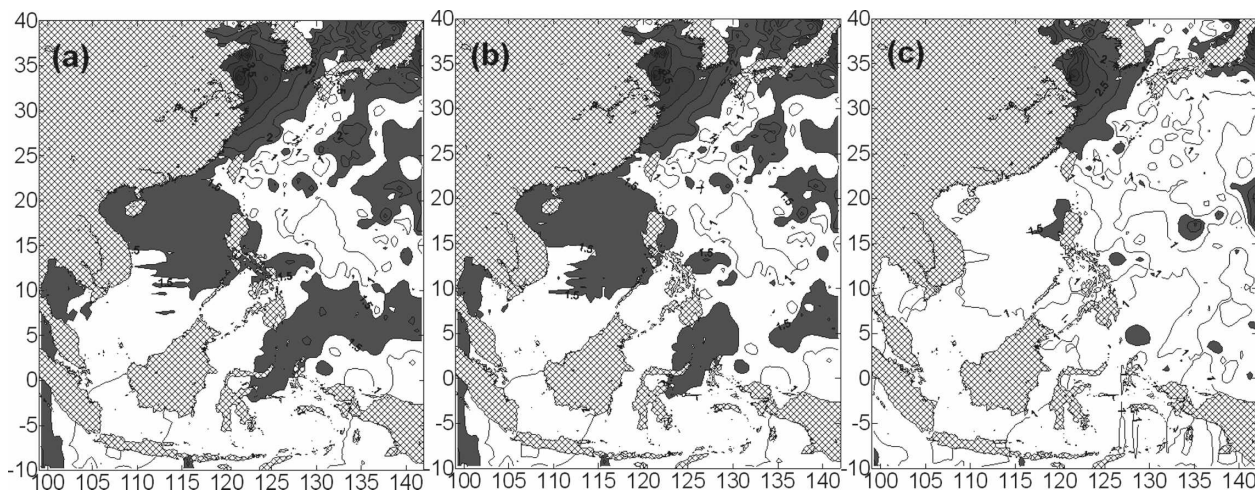


FIG. 14. Spatial distribution of vertical averaged RMS difference of the profile forecasts between observations and (a) Emodel, (b) ETV, and (c) EMG. The contour interval is 0.5°C and the black shading is for values greater than 1.5°C.

exist not only in EMG but also in ETV, EMG has a relatively small RMS difference area and smaller RMS differences, which indicates that the multigrid data assimilation scheme can produce more accurate forecasts and can generate a better forecast initial field under an identical forecast environment.

Forecasts using the multigrid data assimilation scheme are better than those using the traditional 3DVAR on nearly every day during the experimental period. With the exception of the assimilation scheme, the common influence factors (such as open boundary conditions, atmospheric forcing conditions, and datasets used) are the same in these subexperiments. It can be concluded that the new method proposed in this paper generates substantially better initial conditions for temperature forecasts.

4. Summary and conclusions

In this paper, a new data assimilation scheme, called the multigrid data assimilation scheme, was developed. The idealized experiment to verify the validity of this new data assimilation scheme was made by comparing it to the traditional 3DVAR using correlation scales. The new scheme was applied to assimilate the shipboard and AVHRR SST and a variety of temperature profile observations in a retroactive real-time forecast experiment in the China Seas. The main conclusions can be summarized as follows:

- (i) For the traditional 3DVAR using correlation scales, only the errors matching the specified length scales can be corrected, and thus its analysis critically depends on how accurate the correlation scales are. In general, in a complex multiscale SST field, for example, it is almost impossible to obtain accurate correlation scales, as demonstrated in the idealized experiment. In addition, shortwave errors should not be corrected until longwave errors are corrected; otherwise, corrections in the short correlation scales make the analysis fit observations too closely, so that longwave information is mistakenly treated as shortwave information. However, the multigrid data assimilation scheme always performs well, producing an analysis with much higher accuracy than the traditional 3DVAR, because it can minimize the longwave and shortwave errors in turn.
- (ii) In the retroactive real-time sea temperature forecast experiment, the forecast accuracy of profiles and SST by using the multigrid data assimilation

scheme is much higher than that of the traditional 3DVAR, so it can be concluded that the multigrid data assimilation scheme can generate a substantially better initial field for numerical sea temperature forecasts.

Acknowledgments. The authors thank two reviewers for their thorough and helpful comments and suggestions, which contributed to greatly improving the original manuscript. The research for this paper was jointly supported by grants of the National Basic Research Program of China (2007CB816001), the National Natural Science Foundation of China (40776016, 40476006, and 40231014), the National High-Tech R&D Program of China (2006AA09Z138), and CAS Key Laboratory of Tropical Marine Environmental Dynamics (LED).

REFERENCES

- Behringer, D. W., M. Ji, and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The ocean data assimilation system. *Mon. Wea. Rev.*, **126**, 1013–1021.
- Briggs, W. L., V. E. Henson, and S. F. McCormick, 2000: *A Multigrid Tutorial*. 2nd ed. Society for Industrial and Applied Mathematics, 193 pp.
- Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347.
- Ezer, T., and G. L. Mellor, 2004: A generalized coordinate ocean model and a comparison of the bottom boundary layer dynamics in terrain-following and in z -level grids. *Ocean Modell.*, **6**, 379–403.
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.
- Hayden, C. M., and R. J. Purser, 1995: Recursive filter objective analysis of meteorological fields: Applications to NESDIS operational processing. *J. Appl. Meteor.*, **34**, 3–15.
- Mellor, G., and T. Yamada, 1974: A hierarchy of turbulence closure models for planetary boundary layers. *J. Atmos. Sci.*, **31**, 1791–1806.
- , and —, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.*, **20**, 851–875.
- Navon, I. M., and D. M. Legler, 1987: Conjugate-gradient methods for large-scale minimization in meteorology. *Mon. Wea. Rev.*, **115**, 1479–1502.
- Weaver, A., and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion equation. *Quart. J. Roy. Meteor. Soc.*, **127**, 1815–1846.
- Xie, Y., S. E. Koch, J. A. McGinley, S. Albers, and N. Wang, 2005: A sequential variational analysis approach for mesoscale data assimilation. Preprints, *21st Conf. on Weather Analysis and Forecasting/17th Conf. on Numerical Weather Prediction*, Washington, DC, Amer. Meteor. Soc., 15B.7. [Available online at <http://ams.confex.com/ams/pdfpapers/93468.pdf>.]
- Zhou, G., W. Fu, J. Zhu, and H. Wang, 2004: The impact of location-dependent correlation scales in ocean data assimilation. *Geophys. Res. Lett.*, **31**, L21306, doi:10.1029/2004GL020579.