Acta Oceanol. Sin., 2017, Vol. 36, No. 3, P. 37–51
DOI: 10.1007/s13131-017-0978-2
http://www.hyxb.org.cn
E-mail: hyxbe@263.net

Assimilating OSTIA SST into regional modeling systems for the Yellow Sea using ensemble methods

JI Xuanliang[1, 2], KWON Kyung Man[3], CHOI Byoung-Ju[3, 4*], LIU Guimei[1, 2], PARK Kwang-Soon[5], WANG Hui[1, 2], BYUN Do-Seong[6], LI Yun[1, 2], JI Qiyan[7], ZHU Xueming[1, 2]

[1] National Marine Environmental Forecasting Center, State Oceanic Adminstration, Beijing 100081, China

[2] Key Laboratory of Research on Marine Hazards Forecasting, National Marine Environmental Forecasting Center, State Oceanic Adminstration, Beijing 100081, China

[3] Department of Oceanography, Kunsan National University, Gunsan 54150, Republic of Korea

[4] Department of Oceanography, Chonnam National University, Gwangju 61186, Republic of Korea

[5] Korea Institute of Ocean Science and Technology, Ansan 15627, Republic of Korea

[6] Korea Hydrographic and Oceanographic Agency, Busan 49111, Republic of Korea

[7] Marine Acoustics and Remote Sensing Laboratory, Zhejiang Ocean University, Zhoushan 316000, China

Received 13 October 2015; accepted 3 March 2016

Abstract

The effects of sea surface temperature (SST) data assimilation in two regional ocean modeling systems were examined for the Yellow Sea (YS). The SST data from the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) were assimilated. The National Marine Environmental Forecasting Center (NMEFC) modeling system uses the ensemble optimal interpolation method for ocean data assimilation and the Kunsan National University (KNU) modeling system uses the ensemble Kalman filter. Without data assimilation, the NMEFC modeling system was better in simulating the subsurface temperature while the KNU modeling system was better in simulating SST. The disparity between both modeling systems might be related to differences in calculating the surface heat flux, horizontal grid spacing, and atmospheric forcing data. The data assimilation reduced the root mean square error (RMSE) of the SST from 1.78°C (1.46°C) to 1.30°C (1.21°C) for the NMEFC (KNU) modeling system when the simulated temperature was compared to Optimum Interpolation Sea Surface Temperature (OISST) SST dataset. A comparison with the buoy SST data indicated a 41% (31%) decrease in the SST error for the NMEFC (KNU) modeling system by the data assimilation. In both data assimilative systems, the RMSE of the temperature was less than 1.5°C in the upper 20 m and approximately 3.1°C in the lower layer in October. In contrast, it was less than 1.0°C throughout the water column in February. This study suggests that assimilations of the observed temperature profiles are necessary in order to correct the lower layer temperature during the stratified season and an ocean modeling system with small grid spacing and optimal data assimilation method is preferable to ensure accurate predictions of the coastal ocean in the YS.

Key words: ensemble optimal interpolation, ensemble Kalman filter, SST, Yellow Sea, assimilation

Citation: Ji Xuanliang, Kwon Kyung Man, Choi Byoung-Ju, Liu Guimei, Park Kwang-Soon, Wang Hui, Byun Do-Seong, Li Yun, Ji Qiyan, Zhu Xueming. 2017. Assimilating OSTIA SST into regional modeling systems for the Yellow Sea using ensemble methods. Acta Oceanologica Sinica, 36(3): 37–51, doi: 10.1007/s13131-017-0978-2

1 Introduction

The Yellow Sea (YS) and the Bohai Sea (BS) are surrounded by the east coast of China and west coast of the Korean Peninsula, and the latitude ranges from 31°N to 41°N. The climate and general circulation in the YS and BS are strongly influenced by external and internal forcings, i.e., Kuroshio, the East Asia monsoon and river discharge (Wang et al., 2013). Since the YS and BS are intensely used for coastal tourism, fishery activity and marine transportation with the risk of floating sea ice, accurate predictions of the temperature and currents are essential for the YS and BS.

In winter, strong northerly winds drive a southward flow at the surface and along both the Korean and Chinese coasts (Naimie et al., 2001; Lin and Yang, 2011). The surface flow is compensated for by a deep upwind flow, the Yellow Sea Warm Current (YSWC). Satellite sea surface temperature (SST) images

Foundation item: The National Key Research and Development Program of China under contract Nos 2016YFC1401800 and 2016YFC1401605; the Cooperation on the Development of Basic Technologies for the Yellow Sea and East China Sea Operational Oceanographic System (YOOS); the project of Development of Korea Operational Oceanographic System (KOOS), Phase 2 funded by the Ministry of Oceans and Fisheries; the National Natural Science Foundation of China under contract Nos 41076011, 41206023 and 41222038; the National Basic Research Program (973 Program) of China under contract No. 2011CB403606; the Public Science and Technology Research Funds Project of Ocean under contract No. 201205018; the Strategic Priority Research Program of the Chinese Academy of Sciences under contract No. XDA1102010403; Producing map of ocean currents for the neighboring seas of Korea funded by the Ministry of Oceans and Fisheries under contract No. 2033-307-210-13.
*Corresponding author, E-mail: bjchoi@kunsan.ac.kr

show that warm water intrusion is a prominent feature on in winter and the saline and warm water are supplied by the YSWC along the western slope of the Yellow Sea Trough (YST). Lin and Yang (2011) suggested that the westward shift of the YSWC is a result of the potential vorticity input by the wind stress over the YST, and the position is strongly influenced by bottom friction through linear, barotropic and wind stress driven model simulations.

In summer, a tidal mixing front forms at approximately the 50 m isobath of the YS. The water column is highly stratified in the offshore side of the tidal mixing front and it is relatively well mixed in the shallow region. Cyclonic surface circulation develops along the front in the YS. Tidal stirring is essential to generate a bottom tidal mixing front and boundary currents along the front in the summer (Liu et al., 2003b; Xia et al., 2006; Kwon et al., 2011). The northward-flowing boundary current along the tidal mixing front is maintained through a balance between the Coriolis and pressure gradient forces, i.e., a quasi-geostrophic balance (Xia et al., 2006). Wind mixing results in a widening of the boundary current along the front and an increase of the northward transport along the eastern boundary of the YS (Kwon et al., 2011). The Yellow Sea Bottom Cold Water is a water mass that is produced by winter cooling, and it is isolated in the YST and sets up cyclonic circulation over the YS in the summer (Naimie et al., 2001).

The vertical structure of temperature has been studied using numerical models. For example, Liu et al. (2003a) simulated the density residual currents in the BS using the Blumberg and Mellor 3-dimensional (3-D) nonlinear numerical coastal circulation model that incorporates the Mellor and Yamada level 2.5 turbulent closure model, and they were able to properly reproduce the structure of the temperature. Zheng and Chen (2013) successfully simulated the 3-D structure of the summer temperature in the BS by using the Finite Volume Coastal Ocean Model (FVCOM) and reproduced the structure of the asymmetric dual-core bottom cold water, which is constituted by cold water in depressions located in the north and south of the central shallow ridge.

Since the 1980s, many data assimilation methods have been developed to improve the understanding of the ocean dynamics and the accuracy of the modeled results (Derber and Rosati, 1989). Data assimilation integrates the observed data into the model to control the model errors. Nowadays, data assimilation methods can be divided into two types according to the underlying theory and operating principle. One type is based on statistical theory, i.e., optimal interpolation (OI), Kalman filter (KF), ensemble Kalman filter (EnKF), and ensemble optimal interpolation (EnOI). Gandin (1965) introduced the OI assimilation method, which first puts forward the concept of a model background field. The KF method actually updates the results of the model through the covariance of variables, but its computational requirements are still a practical limitation. The EnKF method was introduced by Evensen (1994) and it has gained popularity because it conforms to a simple conceptual formulation and is relatively easy to implement. The EnKF method is widely used to assimilate sea level anomalies, satellite SST, and Argo profile data (Hamill et al., 2001; Evensen, 2003; Oke et al., 2007; Seo et al., 2010, 2015). The EnOI is an approximation of the EnKF and is very computationally efficient. The EnOI is widely used in both coarse ocean models (Xie and Zhu, 2010) and eddy-resolving operational systems (Lyu et al., 2014) because it has a multivariate property and a high computational efficiency. Another type is based on the variational method that includes three-dimensional variational (3-DVAR) analysis and four-dimensional variational (4-DVAR) analysis. It is based on the variational principle for which the analysis field obtains the optimal value from the forecast data and the observation data. The 3-DVAR method is widely used in assimilation research of observation data (Zhu et al., 2006; Xiao et al., 2008; Li et al., 2008).

The National Marine Environmental Forecasting Center (NMEFC) of China and the Korea Institute of Ocean Science & Technology (KIOST) of South Korea, two ocean forecasting institutions, established the Yellow Sea and East China Sea Operational Oceanographic System (YOOS) program in 2011 to ensure accurate predictions of the temperature and currents for the YS and BS by improving the 3-D circulation model. The KIOST ocean modeling system was developed in collaboration with Kunsan National University (KNU) in Korea, and the ocean circulation modeling working group was organized as one of the YOOS programs. The working group compared two operational modeling systems for the YS and collaborated to improve the efficiency and accuracy of the YS forecasting systems.

To this end the objectives of this study are to compare two regional ocean modeling systems for the YS and to examine the effects of SST assimilation using the EnOI and EnKF methods. The Regional Ocean Modeling System (ROMS) is a 3-D ocean circulation model that is used to simulate ocean circulation and temperature in both regional ocean modeling systems. The NMEFC modeling system has a smaller horizontal grid spacing of 3 km and uses a computationally efficient EnOI method to conduct data assimilation. The KNU modeling system has a larger horizontal grid spacing of 8 km and uses the EnKF method. The data-assimilative regional models can provide initial and boundary condition for high-resolution coastal models. In Section 2 of this paper, the circulation models and their configurations are introduced. Section 3 presents the main features of the SST from two experiments with one as a free-run experiment and the other as an assimilative-run experiment. The effects of the SST assimilation and the influence of the assimilation methods on the vertical structure of the temperature are discussed in Section 4. Section 5 summarizes the results obtained from this study.

2 Numerical model and assimilating methods

2.1 *Regional Ocean Modeling System*

This study uses ROMS, which is a free-surface, terrain-following, primitive equations ocean model (Haidvogel et al., 2000; Shchepetkin and McWilliams, 2005). In the vertical direction, the primitive equations are discretized over variable topography using stretched terrain-following coordinates that allow for an increase in the resolution in the areas of interest, such as the thermocline and bottom boundary layers. In the horizontal direction, primitive equations are discretized using boundary-fitted, orthogonal curvilinear coordinates on a staggered Arakawa C-grid. In ROMS, the coastal boundaries can also be specified as a finite-discretized grid via land/sea masking (http://www.myroms.org/). ROMS has various options for advection schemes and parameterizations and is also used by the scientific community for a wide range of applications (Liu and Chai, 2009; Lyu et al., 2014).

2.2 *Two numerical modeling systems for the Yellow Sea*

As members of the Ocean Circulation Modeling Working Group for the YS, the staffs at the NMEFC of China and researchers at KNU, established separately 3-D circulation modeling systems for the YS and BS. The KNU ocean circulation modeling system for the YS and BS has been used for routine daily forecasting

of the ocean at the KIOST and Korea Hydrographic and Oceanic Agency (KHOA).

The NMEFC model domain spans from 22.2°N to 41°N, and from 114.23°E to 133°E, with a horizontal resolution of (1/30)° (Fig. 1a, dashed line). In the vertical direction, there are 30 sigma levels with an enhancement in the surface and at the bottom. The topographic data is based on the General Bathymetric Chart of the Oceans (GEBCO), with a 30 arc-second resolution. Water depths of less than 5 m were set to 5 m, and the maximum bathymetry was set to the real depth. Vertical mixing was calculated based on the parameterization of Mellor and Yamada (1982). Considering the influence of the inflow and outflow through the boundaries, the southern, eastern and northern boundaries were set as open boundaries. The boundary data was derived from Simple Ocean Data Assimilation (SODA) climatological data (http://soda.tamu.edu/assim/SODA_2.2.4/). Eight primary tides (M_2, S_2, N_2, K_2, K_1, P_1, O_1, and Q_1) and two long period tides (M_f, MS_f) were included to consider the effects of tide and tidal current, and the data was obtained from TPXO7. The river runoff from the Changjiang River (CJR) was prescribed at the upstream boundary. The monthly climatological freshwater discharge data without salinity was obtained from the Changjiang Sediment Bulletin (http://www.cjw.com.cn/zwzc/bmgb/nsgb/), and monthly climatological temperature data was obtained from SODA. To reach statistical equilibrium, the model was first spun up for 10a using climatological data obtained from the Comprehensive Oceanic and Atmosphere Data Sets (COADS) climatological data. The 6-hour data of the air-sea fluxes of momentum and buoyancy from January 2000 to February 2012 were obtained from the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR). When running the model, the internal model time step was set to 45 s, and the external model time step was 10 s.

The computational domain of the KNU ocean circulation model ranges from 21°N to 43.27°N, and from 113.53°E to 134.00°E with an average horizontal grid spacing of (1/12)° (Fig. 1a, solid line). There are 30 vertical sigma layers in the vertical direction, and topographic data was derived from Earth Topography 5 (ETOPO5) with a 5-minutes resolution. The topography data in the eastern YS and the Korea Strait were obtained from Seo (2008) and the minimum water depth was set to 5 m with a maximum bathymetry of 4 500 m. The vertical mixing parameterization was based on the Mellor and Yamada (1982) 2.5 closure scheme. The southern, eastern and western boundaries of the model domain were set as open boundaries, and the open boundary data was obtained from Seo et al. (2014).

The tides propagate into the model domain through open boundary and boundary tidal forcing contains the same ten tidal constituents as the NMEFC model. Boundary tide data were derived from TPXO6 and in contrast to the NMEFC modeling system, freshwater was discharged into the YS from seven rivers along the west coast of Korea and two rivers along the Chinese coast in the KNU modeling system. The rivers along the Korean coast are the Han River (HR), Keum River (KR), Mankyoung River (MKR), Dongjin River (DJR), and Youngsan River (YSR); the rivers along the Chinese coast are the CJR and the Huanghe River (HHR). The freshwater discharge data for the HR, KR and YSR were provided by the Flood Control Office of the Ministry of Land Infrastructure and Transport while data for the MKR and DJR were respectively assumed to be 15% and 12% of the KR discharge. The data for CJR was obtained from Senju et al. (2006), and the data for HHR was assumed to be 5% of the CJR discharge. In the KNU model, the surface heat fluxes are calculated using bulk parameterization (Fairall et al., 1996). Therefore, the surface fluxes are calculated from winds, short wave radiation, air temperature, air pressure, and relative humidity, all of which were derived from 1-hour data obtained from the Weather Research and Forecasting (WRF) Model of KISOT (Park et al., 2015).

The common area from 31°N to 41°N and from 117°E to 127°E was selected to compare the results of the simulation with two different numerical modeling systems (Fig. 1b). This area covers the region where observation data are available in the YS and the BS. Modeling period spans from September 1, 2011 to February 29, 2012.

2.3 *Data assimilation methods*

The two modeling systems used different data assimilation methods. The KNU modeling system used the EnKF method, whereas the NMEFC system used the EnOI method. Both EnKF and EnOI are sequential methods that are based on an ensemble.

2.3.1 *Ensemble Kalman filter*

The EnKF is a sequential data assimilation method with a three-step assimilation procedure to update the model state vec-

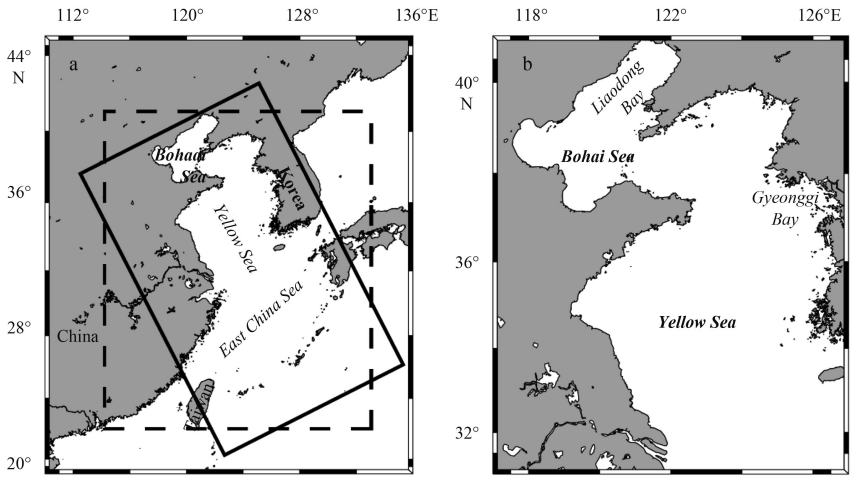

Fig. 1. KNU model domain (solid line) and NMEFC model domain (dashed line) (a), and common region of interest includes the Yellow Sea and the Bohai Sea (b).

tor. First, a numerical model integrates all ensemble members forward in time with an ensemble of initial states to obtain the forecast estimates. The Kalman gain matrix is then calculated using the model background error covariance from the model states and the observation error covariance. Finally, new ensemble members are obtained using the Kalman gain matrix, observation and model states. The analysis step for EnKF uses perturbed observation data.

The equation for the analysis step is

$$\psi_i^{\mathrm{a}} = \psi_i^{\mathrm{f}} + K\left(d_i - H\psi_i^{\mathrm{f}}\right), \tag{1}$$

where ψ_i^{a} and ψ_i^{f} are the model state vectors at a particular time. Superscripts f and a denote the forecast and analysis, respectively. Subscript i indicates the ensemble number such as $i = 1, \ldots, N_{\mathrm{e}}$, where N_{e} is the total number of ensemble members. For this study, N_{e} is 30. The observation operator H interpolates the model states on the observation points and d is the observation vector obtained by adding a random observation error to the data. The difference between the actual measurement (observation) and the predicted measurement (simulated variable), d_i-$H\psi_i^{\mathrm{f}}$, is referred to as the observational innovation.

The Kalman gain matrix K is calculated as follows:

$$K = P^{\mathrm{f}}H^{\mathrm{T}}\left(HP^{\mathrm{f}}H^{\mathrm{T}} + R\right)^{-1}, \tag{2}$$

where K represents the weighting matrix estimated from P^{f} the background error covariance from the model state vectors, and R is the observation error covariance. The background error covariance is expressed as

$$P^{\mathrm{f}} = \overline{\left(\psi^{\mathrm{f}} - \psi^{t}\right)\left(\psi^{\mathrm{f}} - \psi^{t}\right)^{\mathrm{T}}}. \tag{3}$$

The overbar represents the expected value or average. The background error covariance is localized to reduce the false correlation caused by a small number of ensemble members using a specified de-correlation length scale. The horizontal localization radius was set to 100 km, and the computational efficiency can be achieved during the analysis step due to the localized analysis. The SST from the daily composited Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) dataset was assimilated every day, and the observation error of the SST data was assumed to be 0.5°C in the KNU ocean modeling system.

2.3.2 *Ensemble optimal interpolation*

As a data assimilation method based on statistics, EnOI uses a model-based ensemble field to estimate the model errors and to update all model variables by using the covariance of the model-based ensemble. The EnOI analysis is then computed by solving the equation as follows:

$$\Psi^{\alpha} = \psi^{\mathrm{f}} + K\left(d - H\psi^{f}\right), \tag{4}$$

where Ψ^{a}, Ψ^{f}, f, a and d-$H\Psi^{\mathrm{f}}$ are the same as those in the EnKF equation. The matrix K here is as follows:

$$K = \alpha A' A'^{\mathrm{T}} H^{\mathrm{T}}\left(\alpha H A' A'^{\mathrm{T}} H^{\mathrm{T}} + \gamma\gamma^{\mathrm{T}}\right)^{-1}, \tag{5}$$

where A is defined as an ensemble matrix, A' is the ensemble of model anomalies, H is the measurement operator that transforms a model state to observation space, and γ is a matrix that stores the observation errors. Lyu et al. (2014) introduced the equations of these matrices.

α is a parameter that assigns different weights to the ensemble in relation to the observations. Evensen (2003) reported that the ensemble model state in the EnOI method could be derived from a long period of model integration and might contain the climatological variance. Hence, parameter α from 0 to 1.0 was used to reduce the climatological variance to an actual level. Parameter α was set to 0.3 considering the model properties and the configurations described in Lyu et al. (2014). In this paper, the free-running model was integrated from January 2001 to December 2010, and the results obtained from the model were stored every 5 days. Each snapshot of the model results was used as an ensemble model state vector.

This study adopted the same method of the "running" seasonal ensemble used by Xie and Zhu (2010) and Lyu et al. (2014), which means that the ensemble is composed of 180 members selected within a 90-day window centered at the assimilation date. In each year, there are 18 model state vectors within 90 days because model snapshots were stored every 5 days. The horizontal localization radius was set to 150 km, and the SST from the daily OSTIA dataset was assimilated every day. In the NMEFC ocean modeling system, the observation error of SST data was estimated from the analysis error of OSTIA dataset and the analysis error varies in space. The SST observation error ranges from 0.5°C to 2.9°C in the YS and the East China Sea for the NMEFC ocean modeling system.

2.4 *Observation data for model validation*

The numerical models were validated using ocean buoy data and conductivity-temperature-depth (CTD) data (Fig. 2). Hourly buoy data were collected from September 2011 to February 2012 at three buoy stations in the western YS, three buoy stations in the BS, and four buoy stations in the eastern YS. The temperature data from six buoy stations in the western YS and the BS were provided by the North China Sea Branch of the State Oceanic Administration. The temperature data at four buoy stations in the eastern YS were supplied by the Korea Meteorological Administration, KHOA, and KIOST. The KIOST Yellow Sea Buoy was moored in the center of the Yellow Sea at 35°52′N, 124°35′E (K-K in Fig. 2a) and measures the SST every 10 minutes. The CTD data in October 2011, December 2011, and February 2012 were provided by the National Fishery Research Development Institute (NFRDI) of Korea. The NFRDI has six bi-monthly observation lines in the eastern YS (Fig. 2b). The National Oceanic and Atmospheric Administration (NOAA) (1/4)° daily Optimum Interpolation Sea Surface Temperature (OISST) was used as a supplementary satellite SST dataset for the model validation (Reynolds et al., 2007). The daily OISST is an analysis constructed using the satellite SSTs from the AVHRR and the microwave radiometers. The observational data source of OISST was only the AVHRR data from September 2011 to February 2012. However, those of OSTIA included not only the AVHRR data but also microwave radiometer and in-situ data. Although both SST datasets used the same AVHRR data, their data processing procedure, interpolation methods, and data grid spacing were different.

2.5 *Statistical metrics*

The statistical metrics that were used to compare the time series from observation data and simulated data are the Root Mean Square Error (RMSE), BIAS and correlation coefficient (R). The three statistical metrics can be described as follows:

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - X_i\right)^2\right]^{1/2}, \tag{6}$$

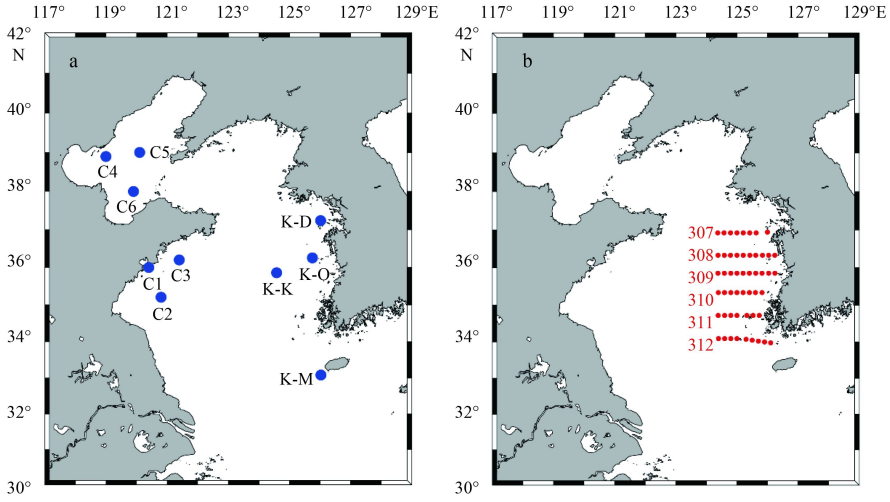

Fig. 2. Position of ocean observation buoy stations (a) and bi-monthly temperature profile observation stations (b) by the NFRDI of Korea.

$$BIAS = \frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i)\,, \tag{7}$$

$$R = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \times \sum_{i=1}^{n}(y_i - \overline{y})^2}}\,. \tag{8}$$

In Eqs (6), (7) and (8), X_i is the observation data, Y_i is the modeled data, and n is the number of data points. The RMSE is a measure of the difference between the observed and simulated time series that can also be considered as a useful measure of the accuracy of the hindcasts of the model (Kara et al., 2008). On the other hand, the BIAS in Eq. (7) can provide the systematic error, which indicates a positive deviation or negative deviation of modeled data relative to the observation data. The correlation coefficient in Eq. (8) was used to demonstrate how the variation of a modeled variable follows that of the observations.

3 Results

3.1 *Comparison of the SST dataset*

The assimilation of satellite remote sensing SST data into the operational ocean forecasting systems will have a significant impact on improving the SST and current prediction. On the other hand, the accuracy of the satellite SST products could be different in many ways, especially in shelf and coastal seas, due to the difference in the satellite data sources, integration methods and derivation techniques from the skin temperature to the foundation SST. The differences among the available SST products and their qualities in different regions were determined by Xie et al. (2008), who analyzed four global SST products and a regional product including OSTIA, Microwave only optimal interpolation SST analysis maps, High-Resolution SST/Sea Ice Analysis for GHRSST, Merged satellite and *in situ* data Global Daily SST in the global ocean, and New Generation SST. The comparison showed that the OSTIA product is considerably closer to the *in situ* observations, and the RMSE of the OSTIA dataset with ship observation data was 1.22°C. The RMSE of the OSTIA dataset with ocean buoy observation data was 0.31°C. The OSTIA data were assimilated from September 2011 to February 2012 in this study.

The OSTIA data has a resolution of (1/20)° and uses satellite SST data provided by international agencies via the Group for High Resolution SST (GHRSST) Regional/Global Task Sharing (R/GTS) framework, together with *in situ* SST data available over Global Telecommunications System (GTS) and a sea-ice concentration product from the EUMETSAT Ocean and Sea Ice Satellite Applications Facility (OSI-SAF). Donlon et al. (2012) summarized the detailed specifications of the OSTIA data.

To determine the accuracy of the OSTIA and OISST data in the BS and the YS, *in situ* observation (buoy and CTD) data were compared with the OSTIA SST and the OISST. The RMSE and bias of the OSTIA SST and OISST were calculated against the SST observed at the ocean buoys (Table 1). The warm (cold) bias of the OSTIA SST was relatively large at C1 (C3 and C4) station while the biases in other stations were less than 0.2°C. The RMSE varied from 0.29 to 0.77°C. It was lowest at the K-M station and was highest at C1. The discrepancy between C1 and other buoy stations may be due to its location; C1 is much closer to the coastline. The averaged RMSE of the OSTIA SST at the Chinese buoy stations was 0.63°C and it was 0.43°C at the Korean buoy stations. The averaged RMSE of the OISST at the Chinese buoy stations was 1.05°C and it was 0.84°C at the Korean buoy stations.

A comparison between the OSTIA SST and the SST observed from the *in situ* ship observation revealed a correlation coefficient (R) of 0.98 and a bias of -0.02°C (Fig. 3), which is similar to

Table 1. RMSE and Bias of the OSTIA SST and OISST relative to the ocean buoy SST

Satellite	Metrics	C1	C2	C3	C4	C5	C6	K-D	K-O	K-M	K-K
OSTIA	RMSE/°C	0.77	0.55	0.61	0.72	0.45	0.69	0.46	0.30	0.29	0.65
	BIAS/°C	0.60	–0.17	–0.32	–0.33	–0.19	–0.06	0.14	–0.05	–0.12	–0.03
OISST	RMSE/°C	1.84	0.60	0.78	0.84	1.50	0.74	1.37	0.76	0.59	0.63
	BIAS/°C	1.79	0.01	–0.27	0.32	0.85	0.37	0.52	0.11	–0.04	–0.05

the results reported by Xie et al. (2008). On the other hand, the RMSE is 1.07°C, which is higher than the RMSE from the buoy stations but is still lower than 1.22°C reported by Xie et al. (2008). The correlation coefficient, bias, and RMSE between the OISST and the *in situ* observation SST were 0.99, 0.04°C, and 0.84°C, respectively. The RMSE of the OISST was smaller than that reported by Xie et al. (2008). There are three groups of SST data in Fig. 3: the SST was higher than 17°C in October, ranged from 10 to 16°C in December, and was lower than 9°C in February in the southeastern YS.

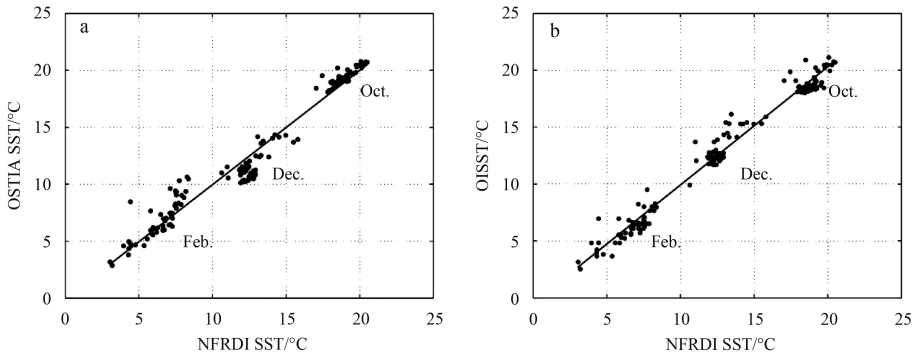

Fig. 3. Comparison of the OSTIA SST (a) and the OISST (b) with the *in situ* observed SST from ship in October 2011, December 2011 and February 2012.

3.2 *Free-running simulations*

Two systems were run without data assimilation to examine the performance of the NMEFC and KNU ocean modeling systems. The monthly mean distribution from the OSTIA SST was compared to that of the SST forecasted from the KNU and NMEFC ocean modeling systems in October 2011, December 2011, and February 2012 (Fig. 4). The spatial distribution of the SST displayed a decreasing trend from south to north. The north-south gradient of the SST increased from October to February, and the difference in the SST between the eastern and western YS was small in October with relatively cool 16°C surface water off North Korea (38.5°N, 124°E). As the air temperature decreased and vertical mixing was enhanced in December, the SST in the southern YS was approximately 3.5°C higher than that in the Liaodong Bay and the Bohai Bay. In December a warm tongue began to grow nearby the southwest of the Jeju Island (Figs 4e and f). The SST simulated from the KNU modeling system was higher than that from the OSTIA in October and December while the SST from the NMEFC modeling system was lower.

The SST of the YS ranged from 1°C to 13°C in February. Both the model simulation and the satellite observation displayed a warm tongue that formed from the west of Jeju Island and extended toward the north-northwest. The warm tongue had the largest acute angle in the southern YS, and the 8°C isotherm extended to the middle of YS (36°N, 123°E) in February (Fig. 4g). However, the 8°C isotherm from the KNU model reached 36.4°N and was extended to 35.3°N in the NMEFC model. The OSTIA SST showed that the 8°C isotherm was divided into two warm tongues between 35°N and 36°N, and these stretched out to the northwest direction and the northeast direction. Similar distribution characteristics were found for the SST isotherms in the SST distribution of the KNU and NMEFC modeling systems. In February, the spatial gradient of the SST was small in the BS and the SST ranged from 0°C to 3°C in the BS. A warm tongue appeared in the SST distribution of northern YS and warm water spread to the Bohai Gulf (Figs 4 g, h and i). The tongue axis is believed to be a trace of Yellow Sea Warm Current Water.

The daily OSTIA SST field was compared with the daily mean SST field simulated by the NMEFC and KNU modeling systems (Table 2). The RMSE, bias, and correlation coefficients were calculated for each month. The SST biases of the NMEFC modeling system were lower than –0.20°C, which means that the SST from the NMEFC system was colder than that from the OSTIA with the average RMSE of 1.53°C. The mean of the correlation coefficients was 0.88 between the observed and simulated SST. The SST bias of the KNU modeling system was higher than 0.05°C except in February, which indicated that the SST from the KNU modeling system had a warm bias. The average RMSE and the correlation coefficient were 1.38°C and 0.88, which shows that the SST from the KNU modeling system was relatively closer to the OSTIA SST.

The daily OISST field was compared to the daily mean SST field simulated by the NMEFC and the KNU modeling systems (Table 3). The average SST bias of the NMEFC modeling system was –0.95°C, and the maximum bias was –2.32°C in December. The averaged RMSE was 1.78°C, and the maximum RMSE was 2.95°C in December. The lowest correlation coefficient was 0.72 for the NMEFC modeling system in September 2011 with correlation coefficients greater than 0.85 in the other five months. The biases of the monthly mean from the KNU modeling system indicate that the SST from the KNU system was warmer than the OISST from September to November, but colder than the OISST from December to February. The mean RMSE and the correlation coefficient of the SST were 1.46°C and 0.82, respectively, for the KNU modeling system. The results of this comparison show that the SST from the KNU modeling system has a smaller RMSE, but the SST from the NMEFC modeling system has a better correlation with the OISST.

The performance of the simulated products in the coastal and shelf seas of the YS and BS was examined by calculating the RMSE and bias of the SST simulated by the modeling systems relative to the SST observed at the ocean buoy stations (Table 4). The SST RMSE of the KNU modeling system ranged from 0.72°C to 1.53°C. The SST RMSE was smallest at the C3 station and largest at the K-M station for the KNU modeling system. The SST simulated by the NMEFC modeling system was colder than that from the buoy stations except at the C1, K-O and K-K stations. The RMSE varied from 0.79°C to 2.34°C and the highest and lowest RMSE were observed at the K-D station and at the C1 station, respectively. The average SST RMSEs for the KNU and NMEFC modeling systems were 1.17 and 1.39°C, respectively. The correl-

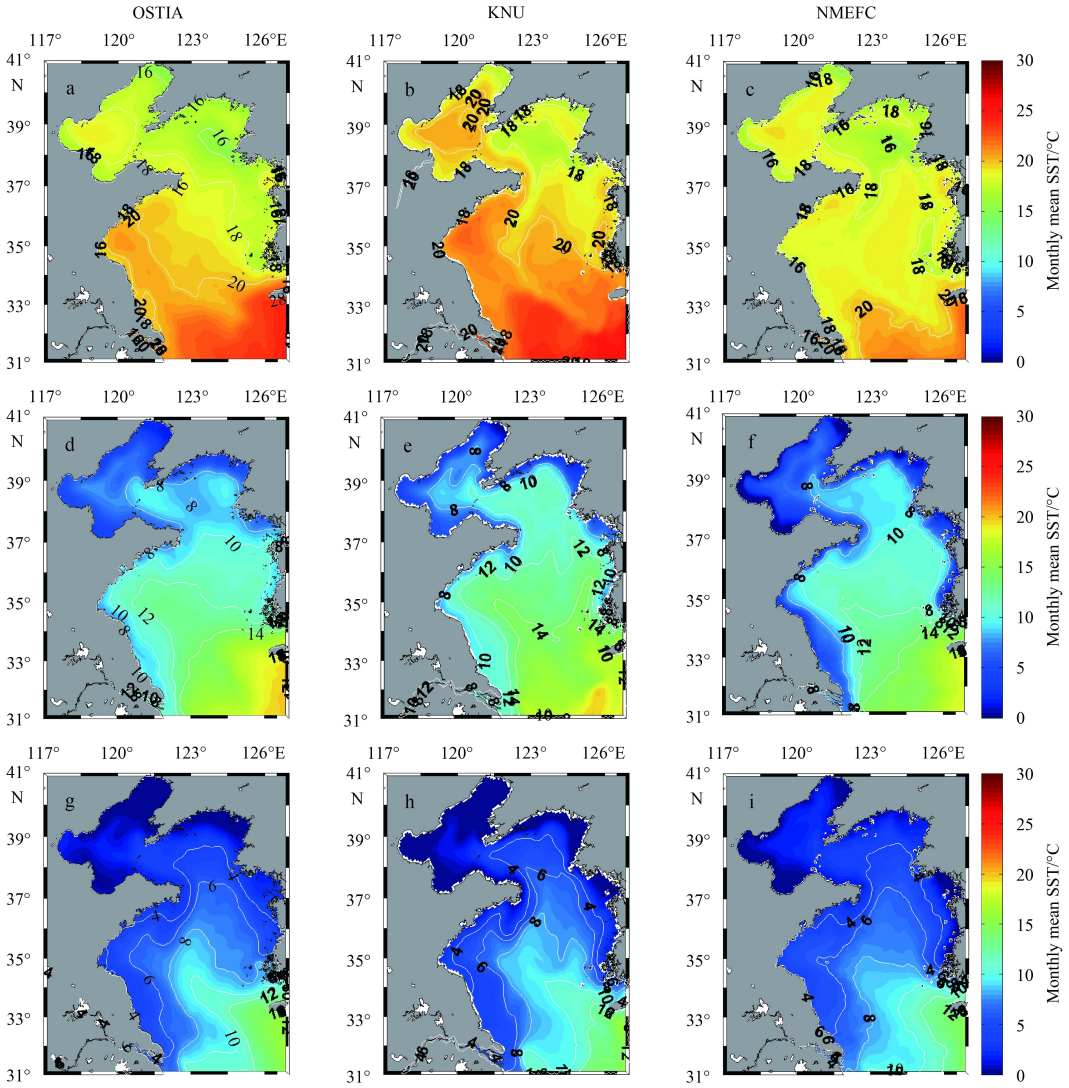

Fig. 4. Monthly mean SST from the OSTIA and the non-assimilative free running KNU and NMEFC modeling systems in October 2011 (a, b, c), December 2011 (d, e, f) and February 2012 (h, i, j).

Table 2. Comparison of the daily OSTIA SST and the modeled SST from the NMEFC and KNU modeling systems without data assimilation

Modeling system	Metrics	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mean
NMEFC	RMSE/°C	1.38	1.45	1.30	**2.09**	1.61	1.29	1.53
	BIAS/°C	–0.20	–0.62	–0.58	**–1.38**	–1.05	–0.30	–0.69
	R	**0.70**	0.85	0.91	0.91	0.94	0.95	0.88
KNU	RMSE/°C	**1.50**	1.39	1.19	1.41	1.40	1.40	1.38
	BIAS/°C	1.09	**1.10**	0.78	0.55	0.05	–0.50	0.51
	R	**0.80**	0.89	0.90	0.89	0.90	0.92	0.88

Note: The largest RMSE, the largest bias and the lowest correlation coefficient in each row are in bold face.

ation coefficients were 0.98 and 0.97 for the KNU and NMEFC modeling systems, respectively (Fig. 5). The results of the comparison suggest that the surface temperature simulated from the KNU modeling system was relatively closer to the surface temperature from the buoy stations.

The accuracy of the simulated subsurface temperature in the coastal and shelf seas of the YS was estimated by comparing the temperature simulated by the non-assimilative models and the temperature observed from the research vessels (Fig. 5). When the simulated temperature profiles were compared to the ship observations, the RMSEs of the temperature for the KNU modeling system were 1.76°C, the temperature biases were 1.21°C, and the correlation coefficients were 0.96 and 0.98, respectively. When the temperature simulated by the NMEFC modeling system was compared to the ship and buoy observations, the RMSEs were 1.72°C, the biases were 0.19°C, and the correlation coefficients were 0.94. The vertical temperature profiles from the NMEFC system were relatively closer to the *in situ* temperature

Table 3. Comparison of the daily OISST and the modeled SST from NMEFC and KNU modeling systems without data assimilation

Modeling system	Metrics	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mean
MEFC	RMSE/°C	1.57	1.54	1.41	**2.95**	2.01	1.20	1.78
	BIAS/°C	0.17	–0.92	–0.80	**–2.32**	–1.46	–0.35	–0.95
	R	**0.72**	0.85	0.90	0.90	0.93	0.94	0.87
KNU	RMSE/°C	**1.76**	1.21	1.14	1.53	1.63	1.46	1.46
	BIAS/°C	**1.45**	0.81	0.60	–0.34	–0.29	–0.50	0.29
	R	**0.76**	0.81	0.84	0.82	0.83	0.85	0.82

Note: The largest RMSE, the largest bias and the lowest correlation coefficient in each row are in bold face.

Table 4. Comparison between the buoy SST data and the modeled SST without data assimilation

Modeling system	Metric	C1	C2	C3	C4	C5	C6	K-D	K-O	K-M	K-K
KNU	RMSE/°C	1.32	0.80	0.72	1.33	1.30	1.05	1.39	1.25	1.53	1.03
	BIAS/°C	–1.32	0.50	0.43	–0.09	1.23	0.87	–0.15	0.92	–1.27	0.97
NMEFC	RMSE/°C	0.79	1.71	1.22	1.59	0.89	1.50	2.34	0.84	2.15	0.86
	BIAS/°C	0.12	–1.51	–1.01	–0.69	–0.33	–0.66	–0.97	0.05	–1.72	0.23

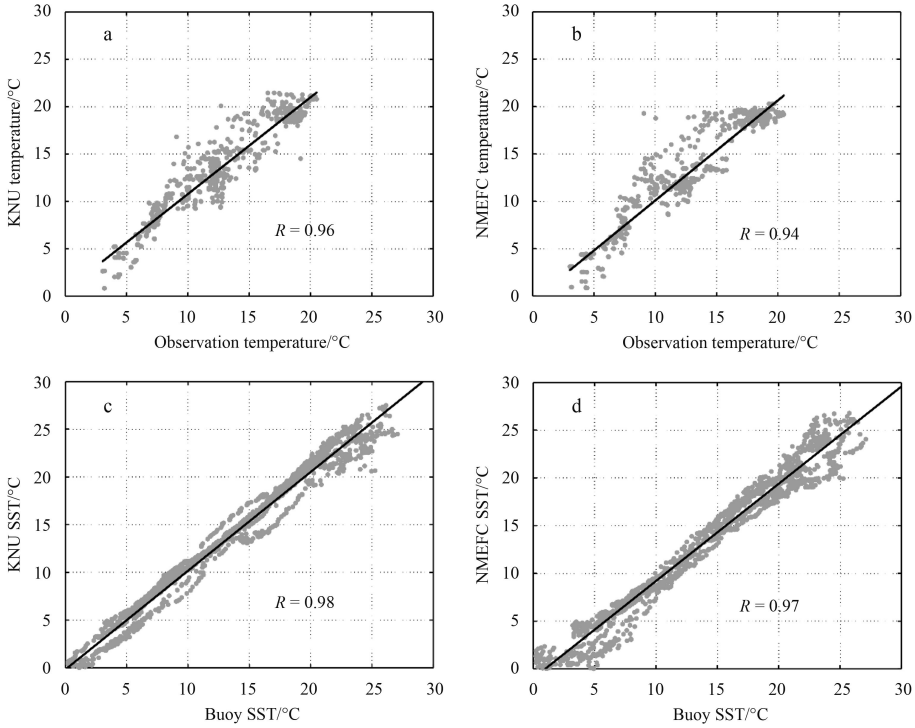

Fig. 5. Comparison of the simulated temperature from non-assimilative models and temperature observed from research vessels (upper panels) and ocean surface buoys (lower panels). *R* is the correlation coefficient obtained from linear regression analysis.

profiles obtained from research vessels.

3.3 *Data assimilation impact on surface temperature*

The OSTIA SST data were assimilated into the KNU and the NMEFC ocean models by applying the EnKF and EnOI methods, respectively. Figure 6 provides a comparison of the monthly mean SST fields simulated from the KNU and the NMEFC modeling systems with data assimilation. The SST simulated by the KNU system became colder toward the OSTIA SST, while the SST simulated by the NMEFC system became warmer toward the observation (Fig. 4). For the KNU modeling system in particular, surface water cooler than 16°C was enhanced off North Korea in October. For the NMEFC modeling system, the SST in the southern YS became warmer.

The bias and RMSE for the SST simulated by the data assimilative modeling systems were calculated relative to the OSTIA SST (Table 5). After assimilation, the monthly mean RMSEs of the SST simulated by the NMEFC modeling system decreased relative to SST RMSEs of the non-assimilative model by 29.52%, 37.36%, 38.97%, 46.31%, 31.07%, and 8.27%, respectively, from September to February (Table 2). For the KNU modeling system, the RMSEs decreased by 26.96%, 39.01%, 35.36%, 36.84%, 14.04% and 9.31%, respectively, from September to February. The correlation coefficients for the KNU and NMEFC modeling systems in September 2011 increased to 0.87 and 0.89, respectively. The mean correlation coefficients increased from 0.88 to 0.94 for the NMEFC modeling system and from 0.88 to 0.92 for the KNU modeling system. The SST assimilation using both EnKF and EnOI methods improved the simulation of SST in terms of the SST RMSE and bias.

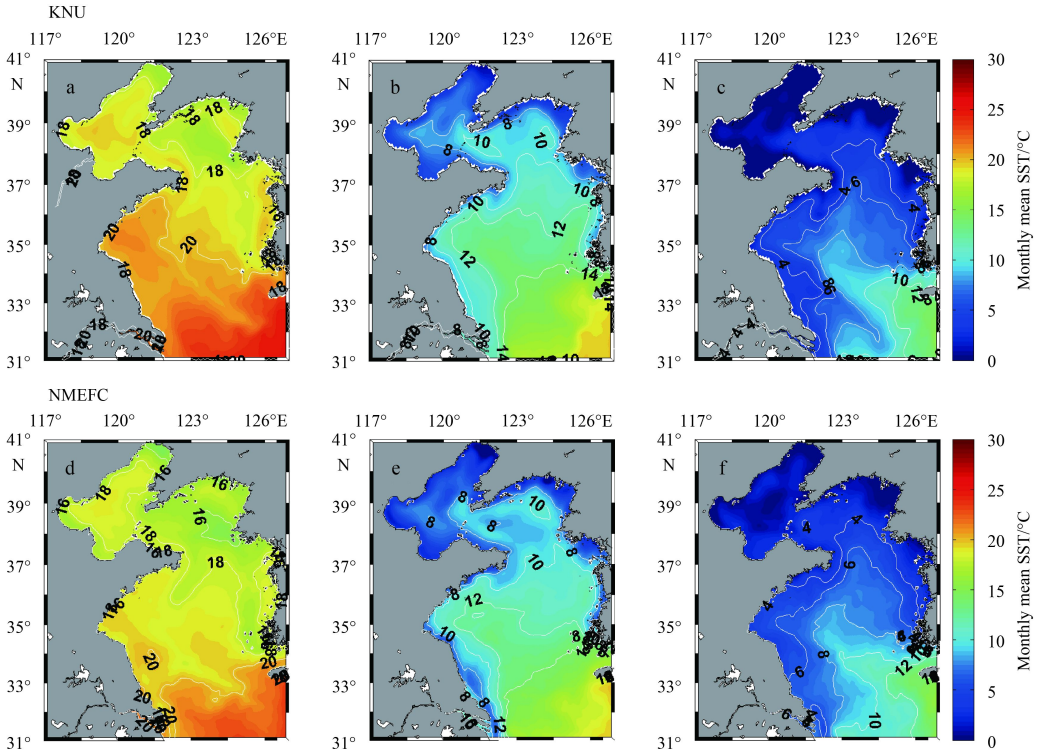

Fig. 6. Monthly mean SST from the assimilative KNU and NMEFC modeling systems in October 2011 (a, d), December 2011(b, e) and February 2012 (c, f).

Table 5. Comparison of the daily OSTIA SST and the modeled SST from the NMEFC and KNU modeling systems with data assimilation

Modeling system	Metrics	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mean
NMEFC	RMSE/°C	0.97	0.91	0.80	1.12	1.11	**1.19**	1.02
	BIAS/°C	–0.45	**–0.49**	–0.23	–0.32	–0.31	–0.08	–0.31
	R	**0.89**	0.95	0.94	0.94	0.95	0.96	0.94
KNU	RMSE/°C	1.10	0.85	0.77	0.89	1.21	**1.27**	1.02
	BIAS/°C	0.82	0.60	0.45	0.38	–0.66	**–0.89**	0.12
	R	**0.87**	0.93	0.92	0.92	0.92	0.93	0.92

Note: The largest RMSE, the largest bias and the lowest correlation coefficient in each row are in bold face.

The bias and RMSE of the simulated SST relative to the OISST were calculated (Table 6). After assimilation, the monthly RMSEs of the SST simulated from the NMEFC modeling system decreased by 35.67%, 29.59%, 29.74%, 34.24%, 22.18% and 0.01%, respectively, from September to February. For the KNU modeling system, the monthly RMSEs decreased by 19.66%, 39.87%, 29.59%, 16.01%, 0.14% and 5.31%, respectively, from September to February. However, the RMSEs in some months were still larger than 1.0°C after assimilation. There are several sources of error which could result in these deviations. One of possible sources of error is atmospheric forcing fields, especially the surface heat flux, which might contain errors, and the incorrect surface boundary condition can induce errors in SST. Model biases due to deficiency in bottom topography data, bottom friction coefficient, and mixing parameterization can be another source of error. The satellite data also have observation errors, which was assumed to be larger than 0.5°C in the data assimilation systems, and the observation error changes in season due to cloud cover variations in the YS. The mean correlation coefficients increased from 0.87 to 0.93 for the NMEFC modeling system and from 0.82 to 0.85 for the KNU modeling system by the assimilation of the OSTIA SST data. This confirmed that data assimilation

Table 6. Comparison of the OISST and the modeled SST from the NMEFC and KNU modeling systems with data assimilation

Modeling system	Metrics	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mean
NMEFC	RMSE/°C	1.01	1.08	0.99	**1.94**	1.56	1.20	1.30
	BIAS/°C	–0.08	–0.79	–0.44	**–1.23**	–0.72	–0.13	–0.57
	R	**0.88**	0.94	0.94	0.93	0.93	0.94	0.93
KNU	RMSE/°C	1.42	0.73	0.80	1.29	**1.63**	1.38	1.21
	BIAS/°C	**1.18**	0.32	0.27	–0.51	–1.02	–0.91	–0.11
	R	**0.81**	0.85	0.87	0.85	0.86	0.87	0.85

Note: The largest RMSE, the largest bias and the lowest correlation coefficient in each row are in bold face.

had a positive influence on the simulation of the SST.

The bias and RMSE of the SST simulated from the data assimilative numerical modeling systems were calculated relative to the ocean buoy observations (Fig. 7). The SST bias was relatively large (>0.5°C) at the C1 and K-M ocean buoy stations for both modeling systems. The C1 station is close to the coast and the K-M station is located near the frontal region of the Jeju Warm Current. The SST bias was relatively small (< 0.5°C) at the C4, C5, C6, K-D, K-O, and K-K ocean buoy stations, which were distributed over the BS and the southeastern YS. After assimilation using the EnOI method in the NMEFC modeling system, the SST RMSEs at the buoy stations decreased to 0.75, 1.11, 1.01, 0.61, 0.63, 0.65, 1.08, 0.63, 1.47, and 0.42°C from the C1 to K-K stations, respectively, and the average RMSE was 0.83°C (Fig. 7). Compared to the buoy observations, the SST RMSEs for the KNU modeling system based on the EnKF method, decreased to 1.18, 0.64, 0.51, 0.71, 0.50, 0.52, 1.28, 0.98, 1.07, and 0.49°C from the C1 to K-K stations, respectively, and the average RMSE was 0.79°C.

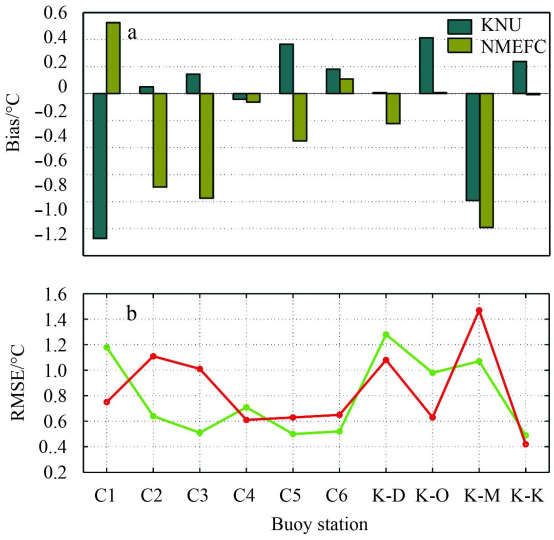

Fig. 7. Bias (a) and RMSE (b) of the SST modeled from the NMEFC (red) and KNU (green) modeling systems relative to the buoy observations (see Fig. 2 for the buoy locations).

The SST data from the assimilative modeling systems were compared to those observed from ocean buoys (Figs 8c and d). Correlation coefficients were 0.99 for both KNU and NMEFC modeling systems. The RMSEs of the SST simulated from the KNU and the NMEFC modeling systems decreased to 0.79 and 0.83°C, respectively. The KNU modeling system was relatively better to simulate the SST in terms of the temperature RMSE.

3.4 *SST assimilation impact on vertical temperature profile*

In this section, independent temperature profile data were used to validate the assimilative modeling systems using the EnKF and EnOI assimilation methods. The vertical profiles of the temperature were observed in the southeastern YS (Fig. 2b) in October 2011, December 2011 and February 2012. The temperature data from the assimilative modeling systems were compared to the temperature profile data observed from the research vessels. Correlation coefficients were 0.96 for the KNU modeling system and 0.95 for the NMEFC modeling system, respectively (Figs 8a and b). The RMSEs of the temperature for the KNU and the NMEFC modeling systems decreased to 1.56°C and 1.58°C, respectively.

In October, the water column was stratified with a warm upper layer (0–30 m) and a cold lower layer (below 30 m) in the central YS (Fig. 9a). A cold water mass with a temperature of less than 10°C, known as Yellow Sea Bottom Cold Water, sitting in the deep trough of the YS. The tidal mixing front formed along the bottom slope, and the water column was relatively mixed over the shallow coastal region. In October, the simulated temperature was slightly warmer than the observed temperature over the slope region. The water column was well mixed vertically due to strong wind mixing in December and February (Figs 9d–i). In December, a band of warm water formed along the lower part of the slope region and the temperature was warmer than 12°C in the observations and in numerical simulations (Figs 9d, e and f). In February, the water temperature was about 1°C cooler in the shallow coastal region of the models than that in the observation (Figs 9g, h and i).

The differences between the observed and simulated temperature (RMSE) were calculated to quantitatively measure the effects of the SST assimilation on the vertical structure of the water column (Fig. 10). In October, the RMSE of the temperature was smaller than 1.5°C in the upper layer from the surface to 10 m and was larger than 2.7°C in the lower layer below 30 m in both the non-assimilative free runs and assimilative runs (Fig. 10a). The large RMSE of the temperature was attributed to the horizontal misplacement of the simulated tidal front over the sloping bottom. As the air temperature decreased and wind speed increased in the winter, vertical mixing was enhanced and the water column was well mixed (Fig. 9). Without data assimilation, the RMSE of temperature decreased to less than 2.2°C in December and February. The SST assimilation reduced the RMSE of the temperature and the RMSE was about 1°C in both data assimilative modeling systems in February.

The RMSEs of the temperature were averaged horizontally at each depth from the surface to a depth of 80 m (Fig. 11). For the KNU system, the horizontally averaged RMSE of the temperature decreased by 0.30°C in the upper layer above 20 m, and by 0.15°C for the lower layer below 30 m. For the NMEFC system, the horizontally averaged RMSE of the temperature is reduced by 0.25°C in the upper layer above 20 m, and by 0.07°C for the lower layer below 30 m. Averages of the temperature RMSEs were 1.56 and 1.58°C for the data assimilative KNU and NMEFC modeling systems, respectively (Fig. 8; Fig. 11 solid lines). Weighted averages of temperature RMSEs over the water column were 1.91 and 1.87°C for the free-running KNU and NMEFC modeling systems, respectively (Fig. 11 dotted lines). The weighted averages of the temperature RMSEs over the water column were reduced to 1.73 and 1.77°C for the data assimilative KNU and NMEFC modeling systems, respectively (Fig. 11 solid lines). Before the SST data were assimilated, the NMEFC modeling system was relatively better in the simulation of temperature profiles. After the SST data were assimilated, the KNU modeling system was relatively better in the simulation of temperature profiles.

Both the EnKF and EnOI methods improved the vertical temperature distribution by assimilating the SST. The multivariate covariance in the EnKF and EnOI methods led to an improvement in the temperature profiles, but when the water column was highly stratified, the innovation (corrected information) in the surface by the assimilation of SST was not transferred to the lower layer. During the winter, vertical mixing was enhanced and the water column in the YS became vertically homogeneous,

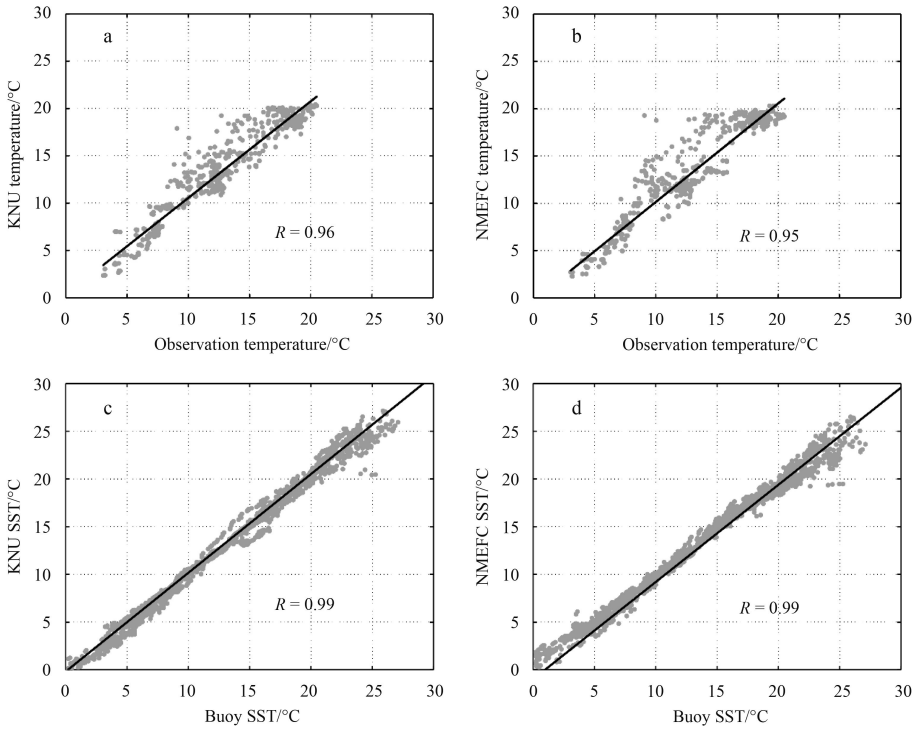

Fig. 8. Comparison of the temperature simulated by data assimilative models and the temperature observed from research vessels (a, b) and ocean buoys (c, d). *R* represents the correlation coefficients obtained from the linear regression analysis.

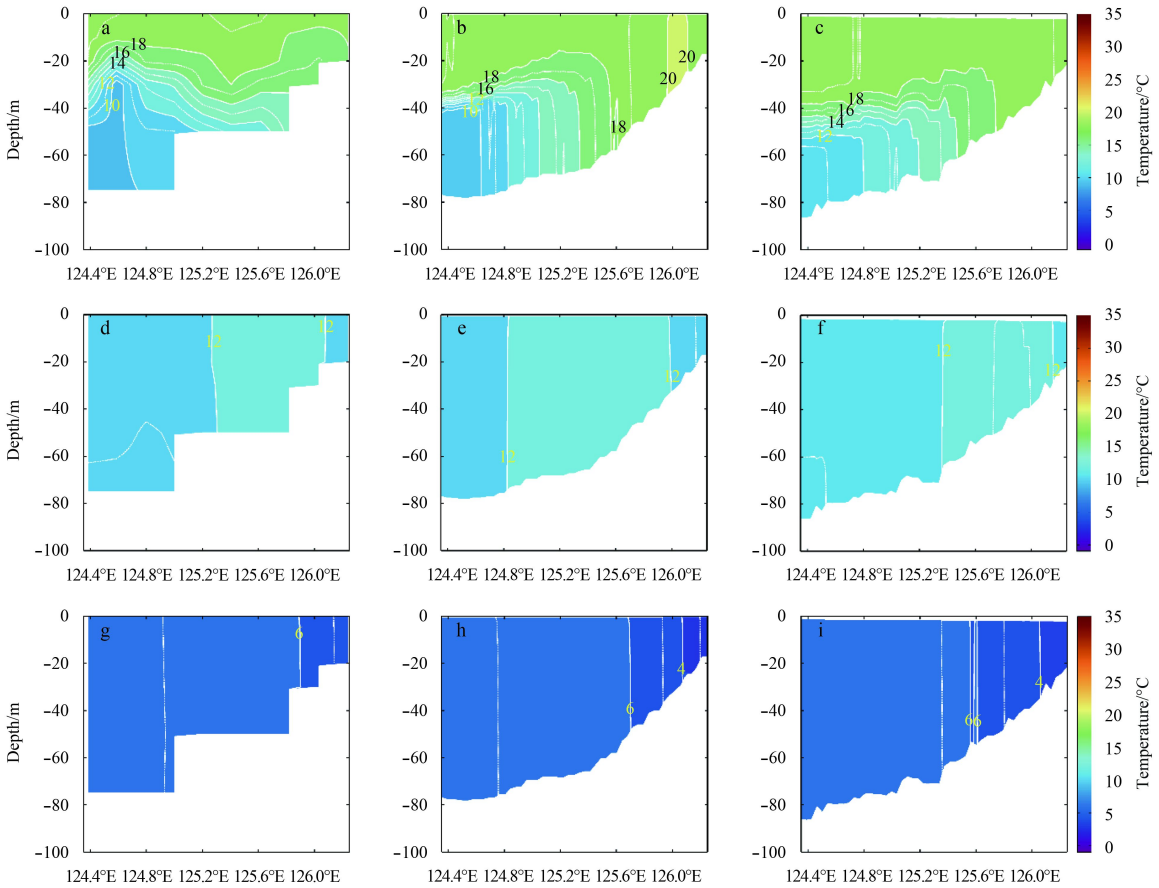

Fig. 9. Vertical temperature section along NFRDI-Korea line 309 from observation (a, d, g), KNU system (b, e, h) and NMEFC system (c, f, i) in October 2011, December 2011, and February 2012, respectively.

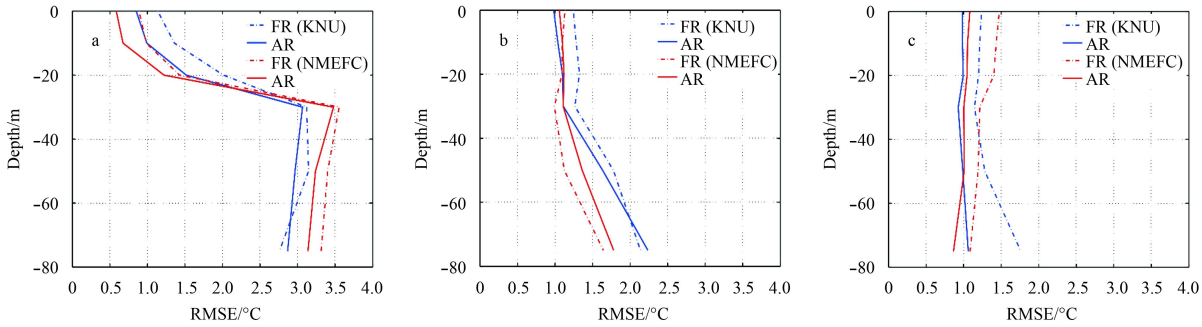

Fig. 10. Vertical structure of the RMSE between temperatures obtained from the data assimilative models and the observed temperature for the KNU and NMEFC systems in October 2011(a), December 2011 (b), and February 2012 (c) from the free-running simulation FR and assimilative simulation AR.

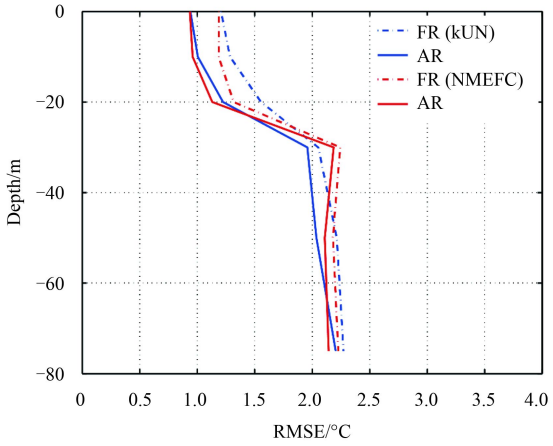

Fig. 11. Vertical structure of horizontally averaged RMSE between the temperature obtained from the data assimilative models and the observed temperature for the KNU (blue line) and NMEFC (red line) systems from September 2010 to February 2012. Temperature RMSEs for the free-running simulation FR (dotted line) and assimilative simulation AR (solid line) are compared.

which allowed the assimilation of the SST to effectively correct the subsurface temperature.

4 Discussion

Before assimilating the SST data, the KNU modeling system produced a better SST and upper layer above the 20 m temperature field with respect to the observation while the NMEFC modeling system produced a better subsurface temperature below 20 m. This difference might be related to the difference in the methods used to calculate the surface heat flux at the sea surface, horizontal grid spacing, boundary data and forcing fields. The KNU modeling system uses bulk parameterization to calculate the surface heat flux and the horizontal grid spacing is approximately 8 km. The NMEFC modeling system enforces a surface net heat flux on the ocean surface, and the nominal horizontal grid spacing is about 3 km. Meanwhile, the boundary data and the forcing fields of the two systems were also different. For the KNU modeling system, the boundary data were daily data derived from Seo et al. (2014). For the NMEFC modeling system, the boundary data were monthly data derived from SODA climatological data. 1-hour forcing field data from the WRF model were used for the KNU ocean modeling system whereas 6-hour forcing data obtained from NCEP CFSR were used for the NMEFC ocean modeling system.

The assimilation of the SST data with the EnOI and EnKF methods significantly reduced the temperature errors. The spatial and temporal distributions of the modeled SST were consistent with those of the observed SST, particularly for the low temperature patch off North Korea in October 2011 and the warm tongue of the SST in the southwestern YS in February 2012 (Figs 4 and 6). The RMSE in the SST was reduced from 1.46–1.78°C to 1.21–1.30°C compared to supplementary SST observations (NOAA OISST dataset). The errors in the subsurface temperature also decreased by the SST assimilation in December and February when the water column was well mixed.

A comparison of the simulated SST and the OSTIA SST revealed that the SST RMSEs for the modeling systems were relatively large in December and January compared to the other months (Tables 3 and 6). The difference between the simulated SST and the OSTIA SST were calculated on 2011 December 15 in order to determine the causes of the large RMSE in December (Fig. 12). A relatively large warm bias exists for the SST along the eastern boundary of the YS, and a cold SST bias exists in the northeastern East China Sea (33.5°N, 124.5°E,) in the KNU free run (Fig. 12b). The warm bias along the eastern boundary might be related to anomalously strong northward along-shore currents. The warm and cold SST bias of the KNU modeling system was reduced over the entire modeling domain by data assimilation with the EnKF method (Fig. 12d). For the NMEFC free run, there was a cold SST bias over the western and central YS except along the eastern boundary (Fig. 12a). In December, the monthly mean bias was reduced from –2.23 to –1.23°C by data assimilation. Although the SST bias was reduced in most of the study area by the data assimilation in the NMEFC system with the EnOI method, the results of the assimilation became worse or had a relatively large warm SST bias in some areas, e.g., the northeastern YS and the southeastern YS (star markers in Figs 12 a and c).

The ensemble SST spread (e-spread) relative to the ensemble average at each grid was calculated to quantify the discrepancy of the results simulated by the model and the observations and to determine the reason for which the SST bias increased in the eastern coastal region for the NMEFC modeling system as follows:

$$\text{e-spread} = \sqrt{\frac{1}{(N_\text{e}-1)}\sum_{i=1}^{N_\text{e}}(|\psi_i - \bar{\psi}|)^2}, \tag{9}$$

where N_e is the number of ensemble members, ψ_i is the SST at a

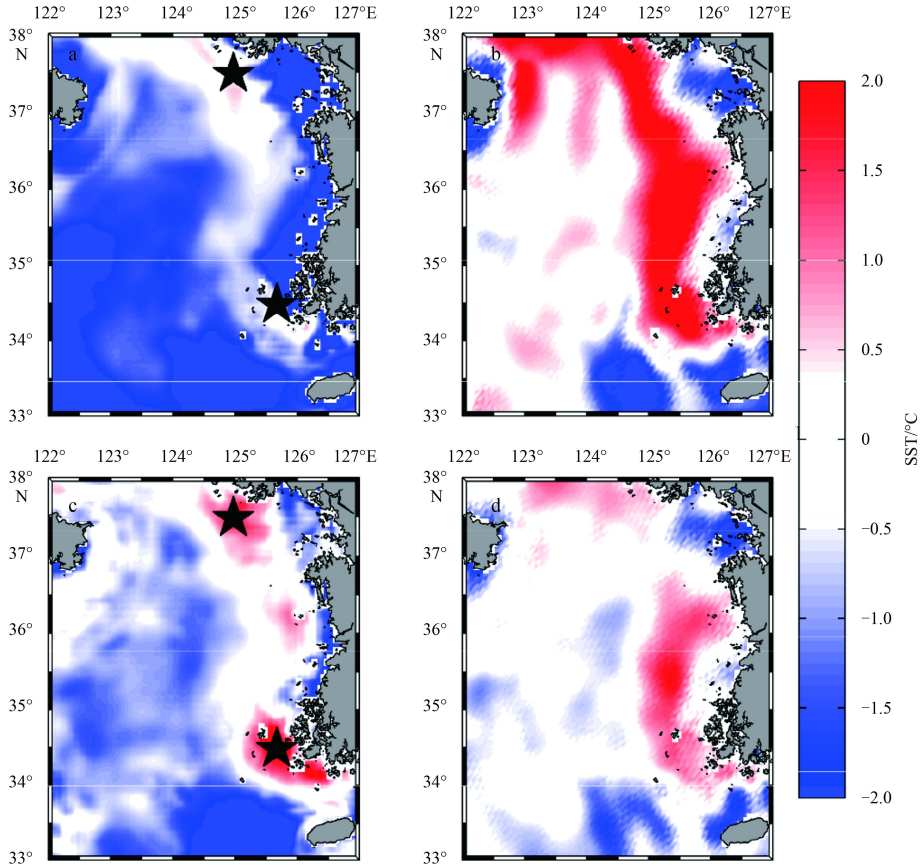

Fig. 12. Difference in the SST between the simulated SST and the observed SST on 15 December 2011. SST difference in NMEFC free run (a), KNU free run (b), NMEFC assimilative run (c) with the EnOI method, and KNU assimilative run (d) with the EnKF method.

grid point from the *i*th member, and $\bar{\psi}$ is the ensemble mean of SST at the grid point. The number of ensemble members (N_e) was 180 for the NMEFC system and that was 30 for the KNU system. The SST spread and the correlation coefficient of the model SST anomalies were calculated for the NMEFC and KNU modeling systems (Fig. 13). The model SST anomaly is a deviation of the SST relative to the ensemble mean SST at a given grid point. The SST spread for the NMEFC system with the EnOI assimilation method was larger than that for the KNU system with the EnKF method (Figs 13a and d). For the NMEFC system, the SST spread was larger than 0.25°C in the coastal ocean around China and Korea, and it decreased to 0.10–0.15°C offshore. Large SST spreads were obtained with the EnOI method because the ensemble members of the NMEFC system were selected from multiple year simulation results with a 90-day window. However, the SST spread for the KNU system was larger than 0.15°C near Gyeonggi Bay and it was lower than 0.15°C offshore. In general, the SST spread was smaller in the KNU modeling system than that in the NMEFC modeling system. When there is a large (small) SST difference among the ensemble members at a given grid point, the analyzed SST become closer to (away from) the observed SST (modeled SST) during the data assimilation process according to Eqs (4) and (5).

The spatial patterns were examined for the correlation coefficients between the SST anomalies at a reference point (star marker in Fig. 13) and the SST anomalies at other grid points. The SST anomalies are the deviation of the SST relative to the ensemble mean. The pattern of the correlation coefficients was similar to that of the background error covariance. In the NMEFC system, the correlation coefficients on December 15 between the SST anomalies at a reference point to the SST anomalies at other surrounding grid points were larger than 0.85 for all regions because of seasonal variation signals in the ensemble of the NMEFC modeling system. However, the correlation coefficients for the KNU modeling system quickly decreased from 1.0 to small values around the reference point. Therefore, the disparity in the SST bias after the SST assimilation in the modeling systems shown in Fig. 12 might be caused by differences in the distributions of the ensemble spread and background error covariance as well as differences in the priors, e.g., a large negative SST bias in the NMEFC modeling system. Since the temperature was lower at areas around the two reference points (Fig. 4d) and the correlation coefficients were relatively high between grid points in the NMEFC modeling system, the SSTs at the reference points probably became excessively warm through the assimilation of the SST data away from the reference points.

5 Summary

Two ocean circulation modeling systems were introduced and compared for the BS and the YS. The ocean circulation modeling systems have been used at the NMEFC and at KNU (KIOST and KHOA). These modeling systems assimilate SST observation data using EnOI and EnKF assimilation methods. A “running” seasonal ensemble, which was composed of 180 members selected within a 90-day window, was adopted for the EnOI method. The ocean modeling system of KNU with the EnKF method used 30 ensemble numbers with a one-day assimilation window. The horizontal localization radius was set to 150 km for the NMEFC

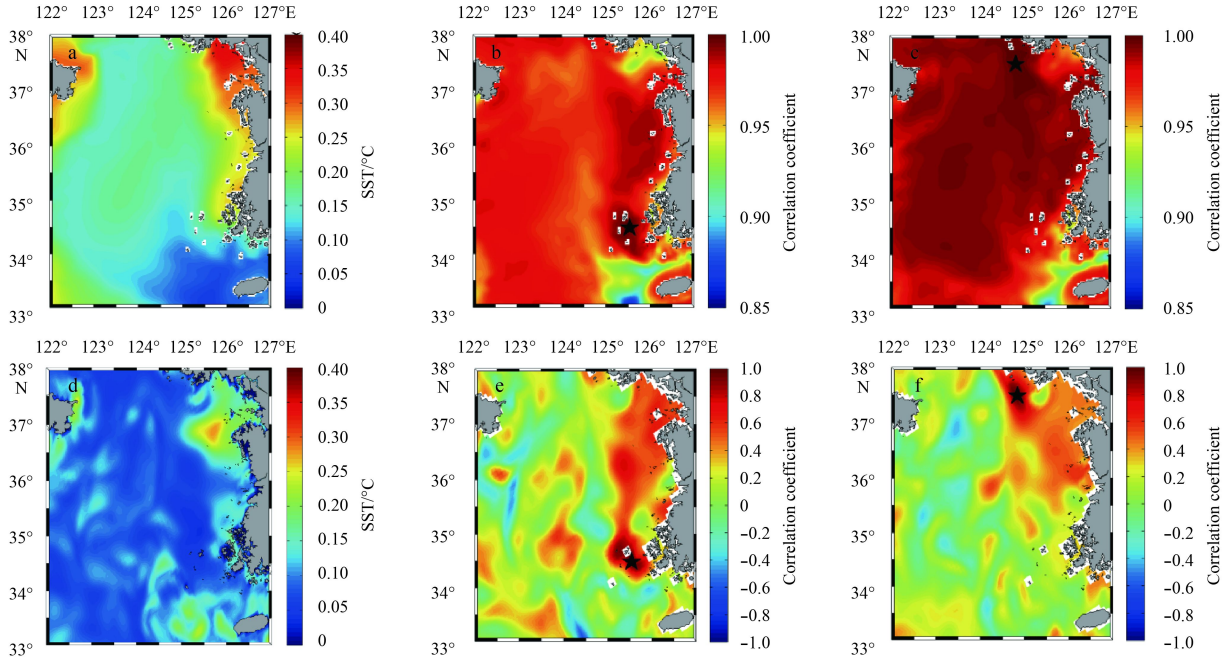

Fig. 13. SST spread for the NMEFC (a) and KNU (d) modeling systems on 15 December 2011. Correlation coefficients between the SST anomalies at a reference point (star marker) to the SST anomalies at the other grid points in the NMEFC (b, c) and KNU (e, f) modeling systems without localization.

system and to 100 km for the KNU system. The OSTIA data had a higher horizontal resolution (smaller grid spacing) compared to the other gridded SST products, and it uses both satellite-borne SST data and the other *in situ* ocean SST data to produce a combined SST product. Since the OSTIA dataset has relatively small biases and RMSEs against other *in situ* observed data, the OSTIA SST from September 2011 to February 2012 was assimilated in both ocean modeling systems.

After assimilating the OSTIA SST, the assimilation products from the ocean circulation modeling systems were compared to the OSTIA and OISST dataset. Compared with OSTIA SST dataset, the RMSE of the simulated SST was reduced from 1.56°C (1.38°C) to 1.02°C (1.02°C) for the NMEFC (KNU) modeling system. When the SST was compared to the NOAA OISST dataset, a supplementary satellite-borne SST product, the RMSE of the SST was reduced from 1.78°C (1.46°C) to 1.30°C (1.21°C) for the NMEFC (KNU) modeling system. Ocean buoy data were collected in order to validate the assimilative numerical modeling systems. The comparisons with ocean buoy SST observation data indicated that the RMSE of SST was reduced from 1.55°C (1.18°C) to 0.91°C (0.82°C) for the assimilative NMEFC (KNU) modeling system. A comparison with independent temperature profiles from ship observations in the southeastern YS from October 2011 to February 2012 showed that the SST assimilation also led to improvements in the subsurface temperature. The RMSE of the temperature was reduced by 0.25°C (0.45°C) in the upper ocean from the surface to a depth of 20 m and was reduced by 0.15°C (0.17°C) in the lower part of the ocean by the SST assimilation for the NMEFC (KNU) modeling system. The RMSE of the temperature was about 1.0°C in the upper 20 m and was approximately 2.1°C in the lower layer despite the SST assimilation. This highlights the importance of assimilating the temperature profiles to improve the prediction of the subsurface temperature when the water column is highly stratified.

The ensemble spread and the correlation coefficients of the background errors were calculated to investigate the relatively large error for the SST in December 2011 off the west coast of Korea. The results showed that the high spread value and the seasonal variation characteristics of the ensembles from the NMEFC modeling system caused an inferior distribution of the SST under the EnOI method. The EnOI method has its limitation as a suboptimal solution to the EnKF method. However, considering the advantage of computational cost, the EnOI method is acceptable as an assimilation technology for the regional ocean forecast. Overall, the low-resolution model with the EnKF method (KNU modeling system) and the high-resolution model with the EnOI method (NMEFC modeling system) showed comparable performance in the marginal seas. This study suggests that the horizontal grid spacing of the numerical models should be reduced and that a more elegant assimilation method (e.g., EnKF) is required for future operational modeling systems in the YS. In future research, the observed temperature and salinity profiles will be assimilated into both ocean modeling systems to examine their effect and to improve the predictability of the ocean state in the YS.

References

Derber J, Rosati A. 1989. A global oceanic data assimilation system. Journal of Physical Oceanography, 19(9): 1333–1347

Donlon C J, Martin M, Stark J, et al. 2012. The operational sea surface temperature and sea ice analysis (OSTIA) system. Remote Sensing of Environment, 116: 140–158

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research, 99(C5): 10143–10162

Evensen G. 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dynamics, 53(4): 343–367

Fairall C W, Bradley E F, Rogers D P, et al. 1996. Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere

Coupled-Ocean Atmosphere Response Experiment. Journal of Geophysical Research, 101(C2): 3747–3764
Gandin L S. 1965. Objective analysis of meteorological fields. Jerusalem: Israel Program for Scientific Translations
Haidvogel D B, Arango H G, Hedstrom K, et al. 2000. Model evaluation experiments in the North Atlantic Basin: simulations in nonlinear terrain-following coordinates. Dynamics of Atmospheres and Oceans, 32(3–4): 239–281
Hamill T M, Whitaker J S, Snyder C. 2001. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. Monthly Weather Review, 129(11): 2776–2790
Kara A B, Wallcraft A J, Martin P J, et al. 2008. Performance of mixed layer models in simulating SST in the equatorial Pacific Ocean. Journal of Geophysical Research, 113(C2): C02020
Kwon K M, Choi B J, Lee S H, et al. 2011. Coastal current along the eastern boundary of the Yellow Sea in summer: numerical simulations. The Sea, 16(4): 155–168
Li Wei, Xie Yuanfu, He Zhongjie, et al. 2008. Application of the multigrid data assimilation scheme to the china seas' temperature forecast. Journal of Atmospheric and Oceanic Technology, 25(11): 2106–2116
Lin Xiaopei, Yang Jiayan. 2011. An asymmetric upwind flow, Yellow Sea Warm Current: 2. Arrested topographic waves in response to the northwesterly wind. Journal of Geophysical Research, 116(C4): C04027
Liu Guimei, Chai Fei. 2009. Seasonal and interannual variability of primary and export production in the South China Sea: a three-dimensional physical-biogeochemical model study. ICES Journal of Marine Science, 66(2): 420–431
Liu Guimei, Wang Hui, Sun Song, et al. 2003a. Numerical study on density residual currents of the Bohai Sea in summer. Chinese Journal of Oceanology and Limnology, 21(2): 106–113
Liu Guimei, Wang Hui, Sun Song, et al. 2003b. Numerical study on the velocity structure around tidal fronts in the Yellow Sea. Advances in Atmospheric Sciences, 20(3): 453–460
Lyu Guokun, Wang Hui, Zhu Jiang, et al. 2014. Assimilating the along-track sea level anomaly into the regional ocean modeling system using the ensemble optimal interpolation. Acta Oceanologica Sinica, 33(7): 72–82
Mellor G L, Yamada T. 1982. Development of a turbulence closure model for geophysical fluid problems. Reviews of Geophysics, 20(4): 851–875
Naimie C E, Blain C A, Lynch D R. 2001. Seasonal mean circulation in the Yellow Sea-a model-generated climatology. Continental Shelf Research, 21(6–7): 667–695
Oke P R, Sakov P, Corney S P. 2007. Impacts of localisation in the EnKF and EnOI: experiments with a small model. Ocean Dynamics, 57(1): 32–45
Park K S, Heo K Y, Jun K, et al. 2015. Development of the operational oceanographic system of Korea. Ocean Science Journal, 50(2): 353–369
Reynolds R W, Smith T M, Liu Chunying, et al. 2007. Daily high-resolution-blended analyses for sea surface temperature. Journal of Climate, 20(22): 5473–5496
Senju T, Enomoto H, Matsuno T, et al. 2006. Interannual salinity variations in the Tsushima Strait and its relation to the Changjiang discharge. Journal of Oceanography, 62(5): 681–692
Seo S N. 2008. Digital 30sec gridded bathymetric data of Korea marginal Seas-Korbathy30s. Journal of Korean Society of Coastal and Ocean Engineers, 20(1): 110–120
Seo G H, Cho Y K, Choi B J. 2014. Variations of heat transport in the northwestern Pacific marginal seas inferred from high-resolution reanalysis. Progress in Oceanography, 121: 98–108
Seo G H, Choi B J, Cho Y K, et al. 2010. Assimilation of sea surface temperature in the Northwest Pacific Ocean and its marginal seas using the ensemble Kalman filter. Ocean Science Journal, 45(4): 225–242
Seo G H, Choi B J, Cho Y K, et al. 2015. Evaluation of a regional ocean reanalysis system for the East Asian marginal seas based on the ensemble Kalman filter. Ocean Science Journal, 50(1): 29–48
Shchepetkin A F, McWilliams J C. 2005. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Modelling, 9(4): 347–404
Wang Fan, Meng Qingjia, Tang Xiaohui, et al. 2013. The long-term variability of sea surface temperature in the seas east of China in the past 40 a. Acta Oceanologica Sinica, 32(3): 48–53
Xia Changshui, Qiao Fangli, Yang Yongzeng, et al. 2006. Three-dimensional structure of the summertime circulation in the Yellow Sea from a wave-tide-circulation coupled model. Journal of Geophysical Research, 111(C11): C11S03, doi: 10.1029/2005JC003218
Xiao Xianjun, Wang Dongxiao, Yan Changxiang, et al. 2008. Evaluation of a 3dVAR system for the South China Sea. Progress in Natural Science, 18(5): 547–554
Xie Jiping, Zhu Jiang. 2010. Ensemble optimal interpolation schemes for assimilating Argo profiles into a hybrid coordinate ocean model. Ocean Modelling, 33(3–4): 283–298
Xie Jiping, Zhu Jiang, Li Yan. 2008. Assessment and inter-comparison of five high-resolution sea surface temperature products in the shelf and coastal seas around China. Continental Shelf Research, 28(10–11): 1286–1293
Zheng Peng, Chen Xueen. 2013. Numerical simulation to 3-D structure of water temperature of Bohai Sea in summer. Periodical of Ocean University of China (in Chinese), 43(11): 9–16
Zhu Jiang, Zhou Guangqing, Yan Changxiang, et al. 2006. A three-dimensional variational ocean data assimilation system: scheme and preliminary results. Science in China Series D: Earth Sciences, 49(11): 1212–1222