Ocean Dynamics (2020) 70:827–837
https://doi.org/10.1007/s10236-020-01366-4

On evaluating the effect of assimilating glider-observed T/S profiles with different horizontal resolutions and assimilation frequencies

Yuhang Zhu[1,5] · Yineng Li[1,2,3] · Shiqiu Peng[1,2,4,5]

Received: 20 December 2019 / Accepted: 23 March 2020 / Published online: 24 April 2020

Abstract
Underwater gliders can provide real-time and spatially flexible temperature/salinity (T/S) observations for improving the marine forecast by data assimilation. By conducting Observing System Simulation Experiments (OSSEs), this study aims to investigate the effect of assimilating glider-observed T/S profiles regarding the horizontal resolution of glider deployment and assimilation frequency, as well as the combination of assimilating satellite-derived sea level anomaly (SLA), on the forecast skill for an extreme warm eddy in the Northwestern South China Sea (SCS) in 2010. The results of OSSEs show that assimilating either glider-observed T/S profiles or satellite-derived SLA is able to improve the forecast skill, and assimilating both of them gains the largest improvement. Under the premise of a full coverage of the eddy, it is found that the higher horizontal resolution of glider deployment is, the better forecast skill will be obtained. Meanwhile, the assimilation of the glider-observed T/S profiles with a 12-h interval achieves the best forecast skill among the intervals of 4 h, 8 h, 12 h, and 24 h. These results provide valuable reference for the deployment of underwater gliders as well as the assimilation strategy of glider observations for improving the real-time marine forecast in the Northwestern SCS in the future.

Keywords Data assimilation · OSSE · Underwater glider · Marine forecast · Mesoscale eddy

1 Introduction

As the largest tropical marginal sea, the South China Sea (SCS) covers over 3.5 million square kilometers with a deep central basin surrounded by two broad shallow shelves to the north and south and two steep continental slopes to the west and east (Qu 2000; Wang et al. 2006). Located in the East Asian monsoon region, the SCS is controlled by the seasonally reversing monsoon and thus the basin-scale circulation of the SCS manifests a cyclonic circulation during winter and an anti-cyclonic circulation during summer (Fang et al. 1998; Hu et al. 2000; Liu et al. 2008). Other dynamical and physical processes, such as the strengthened western boundary current (Wang et al. 2013), the SCS throughflow (Wang et al. 2006; Song 2006; Yu et al. 2007), and the mesoscale eddies (Li et al. 2011; Chen et al. 2011, 2012), and so on, contribute to the complex marine environment of the SCS, which make the marine forecast of the SCS a great challenge.

Numerical models cooperating with data assimilation system are commonly used for marine forecast (e.g., Smedstad et al. 2003; Chassignet et al. 2007; Miyazawa et al. 2017). In the real-time marine forecast, the satellite-derived sea level anomaly (SLA) and sea surface temperature (SST) and the Argo temperature/salinity (T/S) profiles are usually assimilated to improve the forecast skill of the marine forecast system (e.g., Rhodes et al. 2002; Smedstad et al. 2003; Blockley et al. 2013). However, unlike the large amount of observations in the sea surface which are easily obtained by satellite-board instruments, the subsurface observations are very limited.

This article is part of the Topical Collection on the *11th International Workshop on Modeling the Ocean (IWMO), Wuxi, China, 17-20 June 2019*

Responsible Editor: Tal Ezer

✉ Yineng Li
lyneng@scsio.ac.cn; lyneng@scsio.ac.cn

[1] State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China

[2] Southern Marine Science and Engineering Guangdong Laboratory, Guangzhou, China

[3] Institution of South China Sea Ecology and Environmental Engineering, Chinese Academy of Sciences, Guangzhou, China

[4] Guangxi Key Laboratory of Marine Disaster in the Beibu Gulf, Bubei Gulf University, Qinzhou, China

[5] University of Chinese Academy of Sciences, Beijing 100049, China

Argo buoys can provide the subsurface T/S profiles, but they are too sparse in the SCS for the marine data assimilation and forecast with only an average of 0.92 Argo T/S profile in the Northern SCS every single day from 2006 to 2012 (Yang et al. 2013). As a relatively new type of marine observation instrument, the underwater gliders are good substitutes for the Argo buoys due to their remote controllability and flexibility. During the past decades, underwater gliders have been used widely for measuring the T/S profiles of the ocean. In China, underwater gliders have also been increasingly developed and applied in the field observation experiments during the past several years (e.g., Yu et al. 2011; Yu et al. 2013; Shu et al. 2016). However, for the real-time assimilation of the glider observations in the marine forecast system, some issues remain to be investigated, which includes how does the horizontal resolution of underwater glider deployment (hereafter denoted as horizontal deployment resolution) or the assimilation frequency influence the effect of data assimilation on the marine forecast? What is the joint effect of assimilating both the glider-observed T/S profiles and the satellite-derived SLA? This study aims to address these issues through Observing System Simulation Experiments (OSSEs).

OSSEs are commonly used to evaluate the impact of a new observing system on operational forecasts when actual observations are not available (Arnold and Dey 1986; Masutani et al. 2010). Thus, before the implement of a real field observation by underwater gliders, a set of OSSEs is conducted to investigate the issues mentioned above. Considering that mesoscale eddies are one of the important dynamic processes in the SCS, we select the extreme warm eddy occurring in the Xisha area of the Northern SCS in Aug 2010 (hereafter called Xisha warm eddy; Fig. 1) for the OSSEs, which raises the sea surface height and has a great influence on the local climate and marine ecosystem (Chu et al. 2014; McGillicuddy 2016).

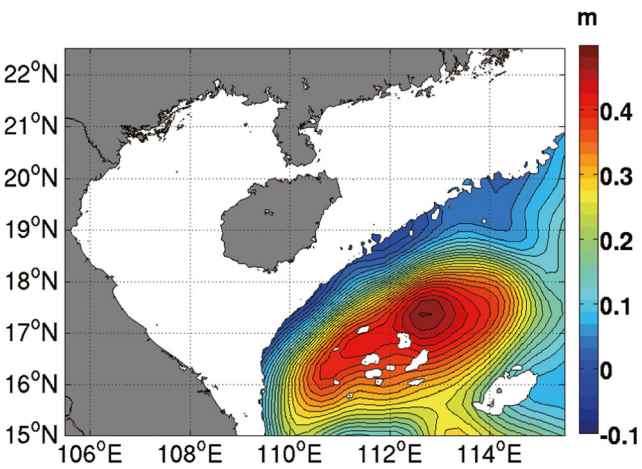

Fig. 1 The sea level anomaly (SLA, unit: m) over the regions with water depth larger than 200 m on Aug. 4, 2010; the data is provided by Copernicus Marine environment monitoring service (http://marine.copernicus.eu/)

The rest of this paper is organized as follows. The next section gives a brief introduction of the real-time marine data assimilation and forecast system for the Northwestern SCS. The designs of pseudo-field observation experiment and OSSEs are presented in Sect. 3, and Sect. 4 presents the result and discussion of OSSEs. A summary is given in the final section.

2 The multi-scale 3DVAR (MS-3DVAR) system and marine forecast system for the northwestern SCS

The 2002 version of Princeton Ocean Model (POM) is used for the marine forecast system (Blumberg and Mellor 1987; Mellor 2003). The domain of POM model covers the northwestern SCS (Fig. 2) with a horizontal resolution of 1/60° × 1/60° and sigma vertical layers of 40, which is refined in the upper and bottom layers. The atmospheric forcing is the daily NCEP Reanalysis 2 dataset (Kanamitsu et al. 2002), and the climatologically monthly mean SODA (Simple Ocean Data Assimilation, version 2.2.6; Carton et al. 2000a, b) data of temperature/salinity/current are used for the lateral boundary conditions of POM. The POM model starts free simulation from Jan. 1, 2009 to Dec. 31, 2014, from which the 6-year daily simulations from 2009 to 2014 are used for estimating the B-matrix of the assimilation system as well as providing the initial field for the OSSEs.

The Multi-Scale 3DVAR (MS-3DVAR) system (Li et al. 2008a, b) is applied for the POM model. The MS-3DVAR is able to assimilate different kind of observations with different sampling resolutions, e.g., the dense satellite-derived SLA and SST, the sparse T/S profile and ship-track SST. The key feature of the MS-3DVAR is that it divides the cost function into two parts technically:

$$J_L(\delta \boldsymbol{x}_L) = \frac{1}{2}\delta \boldsymbol{x}_L^T \mathbf{B}_L^{-1}\delta \boldsymbol{x}_L + \frac{1}{2}\left(\mathbf{H}\delta \boldsymbol{x}_L - \delta \boldsymbol{y}_L\right)^T \left(\mathbf{R} + \mathbf{H}\mathbf{B}_S\mathbf{H}^T\right)^{-1}\left(\mathbf{H}\delta \boldsymbol{x}_L - \delta \boldsymbol{y}_L\right) \tag{1}$$

$$J_S(\delta \boldsymbol{x}_S) = \frac{1}{2}\delta \boldsymbol{x}_S^T \mathbf{B}_S^{-1}\delta \boldsymbol{x}_S + \frac{1}{2}\left(\mathbf{H}\delta \boldsymbol{x}_S - \delta \boldsymbol{y}_S\right)^T \left(\mathbf{R} + \mathbf{H}\mathbf{B}_L\mathbf{H}^T\right)^{-1}\left(\mathbf{H}\delta \boldsymbol{x}_S - \delta \boldsymbol{y}_S\right) \tag{2}$$

In which the subscripts L and S represent large and small scale, respectively, and the superscript T represents the transpose operator. $\delta \boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}^b$ is the increment of optimal values $\boldsymbol{x}$ relative to their background values $\boldsymbol{x}^b$ of the model variables, and $\delta \boldsymbol{y} = \boldsymbol{y} - \mathbf{H}\boldsymbol{x}^b$ is the deviation of observations $\boldsymbol{y}$ from the corresponding model result $\mathbf{H}\boldsymbol{x}^b$. $\mathbf{B}$, $\mathbf{H}$, and $\mathbf{R}$ are the background error covariance matrix (B-matrix), the Jacobian

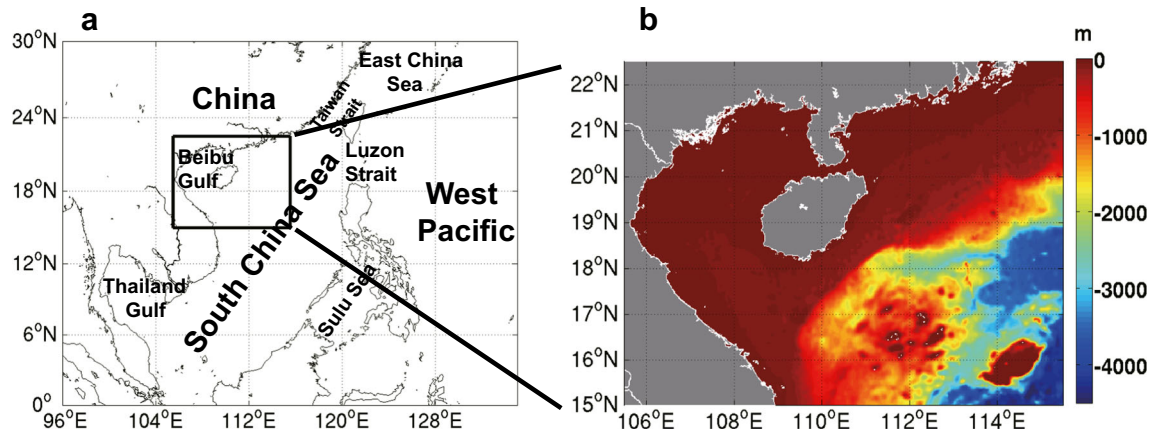

Fig. 2 **a** The geographical location of the model domain of POM, in which the black square is the model domain. **b** The model domain and topography (unit: m) of POM

matrices of the nonlinearly observational operator, and the observational error covariance matrix, respectively. In the large-scale cost function (Eq. 1), small-scale background error covariance matrix $\mathbf{B}_S$ is considered in the representativeness error ($\mathbf{HB}_S\mathbf{H}^T$) of observations, and vice versa, which could effectively reduce or eliminate the representativeness errors of observations, and thus suppress or remove the spurious increments caused by the scale-mismatch of observations. In practice, the scale separation is achieved by performing the data assimilation procedures on coarse and fine model grids sequentially. The fine grid is the same as the original grid of the ocean model, while the coarse grid is 3 times the fine grid. The large-scale increment $\delta \boldsymbol{x}_L$ after the large-scale data assimilation procedure will be interpolated and added to the small-scale background field $\boldsymbol{x}_S^b$ for the small-scale data assimilation. The dense satellite-derived SLA and SST is only assimilated in the small-scale assimilation module, and the sporadic observations, such as underwater glider-observed T/S profiles, are assimilated into both the large- and small-scale assimilation modules sequentially. Unlike the cost functions, the sporadic observations are not split into large and small scale when assimilated into two assimilation modules. After solving the cost functions, two dynamical constraints, including the hydrostatic balance and geostrophic balance, are considered to maintain the dynamical balance of the analysis field, and then a smoothing process based on the weighted spatial moving averaged method is applied to prevent the small-scale noises which may induce the shocks in the model dynamics.

Table 1 Observation errors of SLA for small scale, of temperature and salinity for large-scale and small-scale data assimilation procedures in the MS-3DVAR

	Large scale	Small scale
SLA (m)		0.03
Temperature (°C)	0.85	0.6
Salinity (psu)	0.09	0.06

The B-matrix, which represents the background error between the model results and the observations, plays a vital role in the 3DVAR. The real background errors are not applicable due to lack of adequate numbers of three-dimensional observations. Here, we generated a proxy of the background errors from the 6-year daily simulations by removing the monthly mean, which is analogous to the method provided by National Meteorological Center (the NMC method, Parrish and Derber, 1992). Following Li et al. (2008a), the control variables $\delta \boldsymbol{x}^T = \left(\delta\zeta_n^T, \delta\psi_a^T, \delta\chi_a^T, \delta T^T, \delta S^T\right)$ are adopted, which correspond to the incremental non-steric SSHs, ageostrophic stream function, ageostrophic velocity potential, temperature, and salinity, respectively. For simplicity, the background errors of the five variables are assumed uncorrelated, and then the B-matrix of each variable could be decomposed as

$$B_{\zeta_n} = \sum_{\zeta_n} C^x_{\zeta_n} C^y_{\zeta_n} \sum_{\zeta_n} \tag{3}$$

$$B_V = \sum_V C^x_V C^y_V C^z_V \sum_V \tag{4}$$

where V represents $\delta\psi_a^T$, $\delta\chi_a^T$, δT^T, or δS^T, respectively; $\sum$ is the background error standard deviation matrix that is diagonal; C is the one-dimensional correlation matrix; and x, y, and z are the east-west, south-north, and vertical directions, respectively. Readers may refer to Li et al. (2008a) or Peng et al. (2016) for details. The $\sum$ and vertical C could be calculated from the background errors directly, in which $\sum$ is estimated with respect to the coarse and fine grid separately and vertical C is the same in both grids. The horizontal C is assumed isotropic and described by a Gaussian-type function:

$$C(r_1, r_2) = exp\left[-(r_2 - r_1)^2 / \left(2L^2\right)\right] \tag{5}$$

In which r_1 and r_2 are the locations of two model grids, L is the horizontal decorrelation, which is the distance between r_1 and r_2 when the correlation coefficient decreases to $e^{-1/2}$. L of

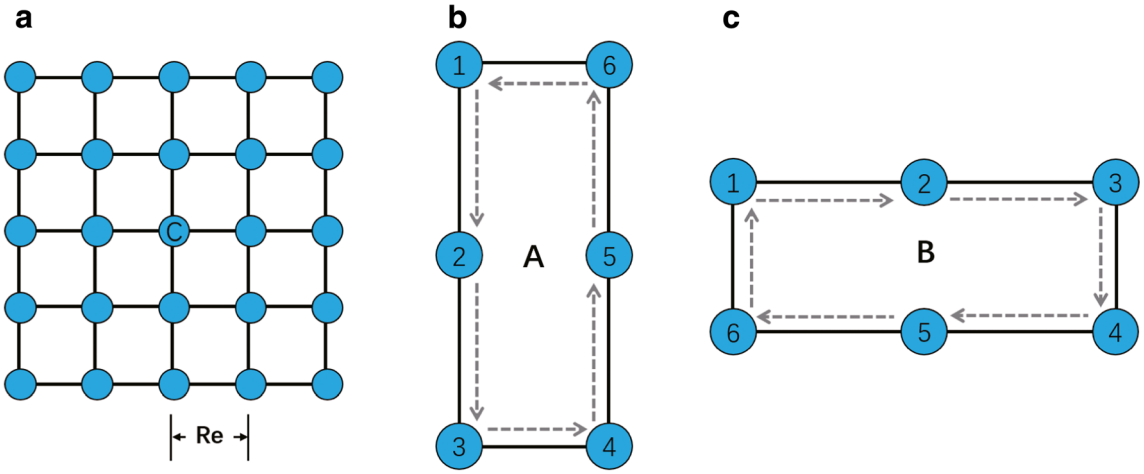

Fig. 3 **a** The equal-spaced observational network of the 25 pseudo-underwater gliders; the blue circles represent the deployment locations of the pseudo underwater gliders, "C" is the center of the network, and "Re" is the "horizontal deployment resolution." **b** The movement strategy A of pseudo-underwater glider; the numbers in the blue circles represent the location of each T/S profile with temporal/spatial interval of 4 h/0.04° which follow the movement direction represented by the arrows (upward direction is north); the number "1" is the first observations obtained at 0400 UTC Aug. 4, 2010; all underwater gliders share the same movement strategy. **c** The same as **b** except for movement strategy B

all control variables is set to the same value and is estimated from the background errors. L of the large- (small-)scale B-matrix for the POM model are 82.5 km and 66 km (27.5 km and 22 km) for x and y directions, respectively. For the observational errors, the error of SLA varies from 0.01 to 0.04 m (Chambers et al. 2003), and here, we set it to be 0.03 m. According to Guan and Kawamura (2004), the errors of T/S profile are set to be 0.6 °C and 0.06 psu in the small-scale assimilation module, respectively. In the large-scale assimilation module, extra representative errors are generated and assumed to depend on the horizontal resolution of model grid linearly. By using the 6-year daily simulations and the small-scale observational errors, the large-scale observational errors of T/S profile are estimated to be 0.85 °C and 0.09 psu, respectively (Table 1).

3 Experimental design

3.1 Pseudo-field observation experiment

In order to investigate the effect of assimilating glider-observed T/S profiles on the forecast skill of the Xisha warm eddy, 25 pseudo-underwater gliders are adopted in the OSSEs.

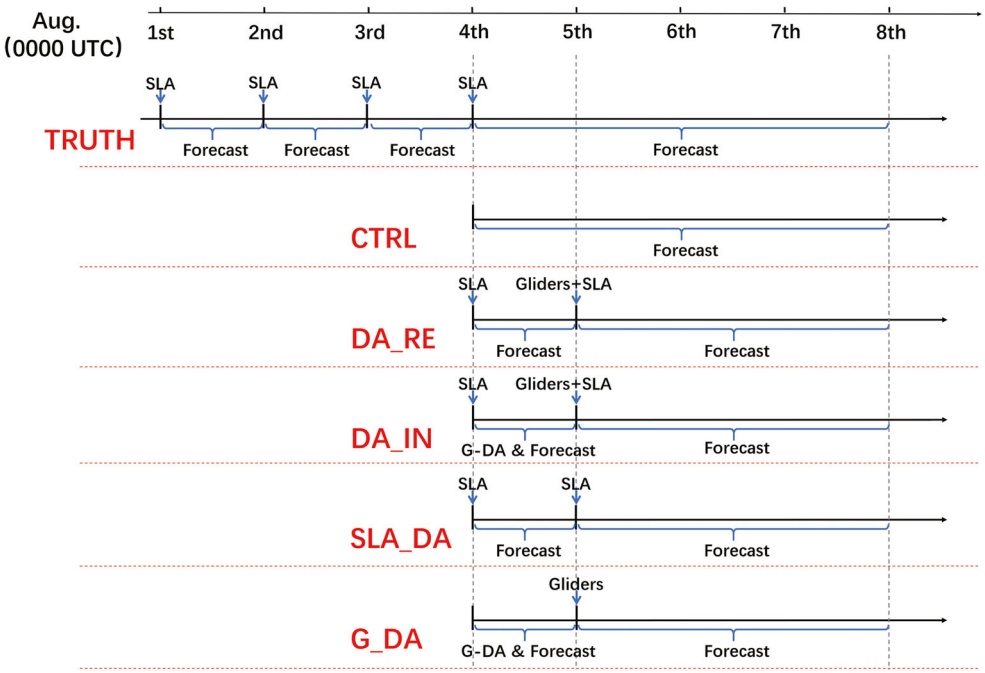

Fig. 4 The flowcharts of the TRUTH, CTRL, DA_RE, DA_IN, SLA_DA, and G_DA. The "Forecast" means the free fun without any data assimilation, and "DA and Forecast" means the successive data assimilation of glider-observed T/S profiles and forecast with a fixed assimilation interval. The blue arrow means the glider-observed T/S profiles or the SLA is assimilated at the pointed moment

Table 2 Designs of the experiments, "DA" means data assimilation and "DA of SLA" means data assimilation of satellite-derived SLA in the TRUTH or pseudo SLA in the OSSEs

DA strategies Experiments			Initial field	Movement strategies of underwater gliders	Horizontal deployment resolutions of underwater gliders	DA interval of underwater gliders	DA of SLA
TRUTH			6-year daily simulation result on Aug. 1, 2010	×	×	×	√
CTRL			6-year daily simulation result on Aug. 4, 2009	×	×	×	×
OSSEs	DA_RE	ReA025		A	0.25° × 0.25°	×	√
		ReB025		B	0.25° × 0.25°		
		ReA05		A	0.5° × 0.5°		
		ReB05		B	0.5° × 0.5°		
		ReA075		A	0.75° × 0.75°		
		ReB075		B	0.75° × 0.75°		
	DA_IN	In4		A	0.5° × 0.5°	4 h	√
		In8		A		8 h	
		In12		A		12 h	
		In24		A		24 h	
	SLA_DA			×	×	×	√
	G_DA			A	0.5° × 0.5°	12 h	×

They are deployed around the center of Xisha warm eddy with a 5 × 5 equally spaced observational network. The deployment of each pseudo-underwater glider locates in the node of the network (such as "C" in Fig. 3a) and varies with the horizontal deployment resolution Re × Re (Fig. 3a). Referred to configuration of real underwater gliders (Peng et al. 2019), the pseudo-underwater gliders are set to dive and rise from the surface to the 1000-m depth with horizontal distance of 4 km (we take 0.04° in this study). The diving-rising process takes about 4 h and thus, each glider can provide 6 T/S profiles per

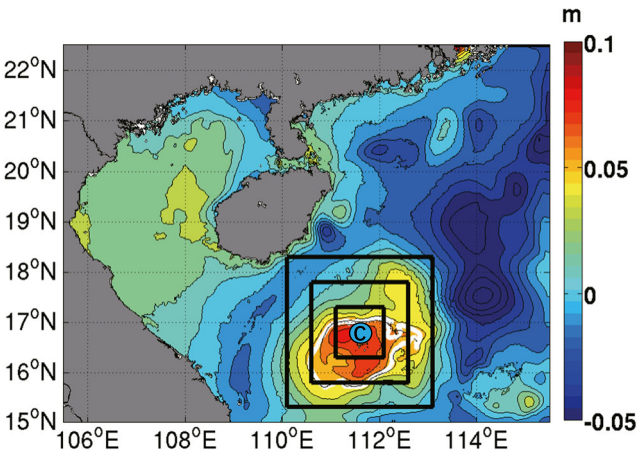

Fig. 5 The mean sea level anomaly (SLA) of TRUTH during the 72-h forecast. The blue circle with C inside represents the center-located underwater glider C as introduced in Fig. 3; the black squares from outside to inside represent the observation area of the underwater glider network with the horizontal resolutions of 0.75°, 0.5°, and 0.25°, respectively; and the region surrounded by the white curve line is the edge of Xisha warm eddy 2010 from the TRUTH

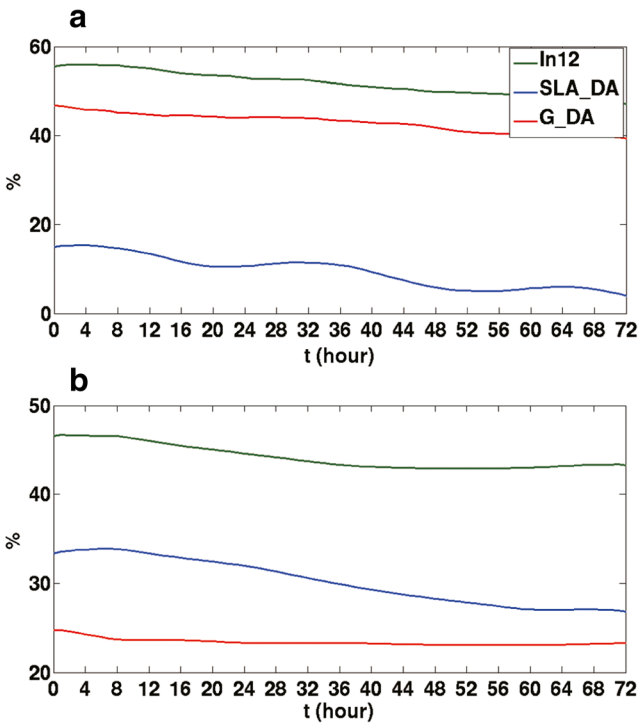

Fig. 6 The evolution of vertical-averaged RMSEs reductions of the forecasting temperature (unit: °C) (**a**) and salinity (unit: psu) (**b**) from In12, SLA_DA, and G_DA validated against the TRUTH in the target region (In12, SLA_DA, and G_DA are indicated in Table 2)

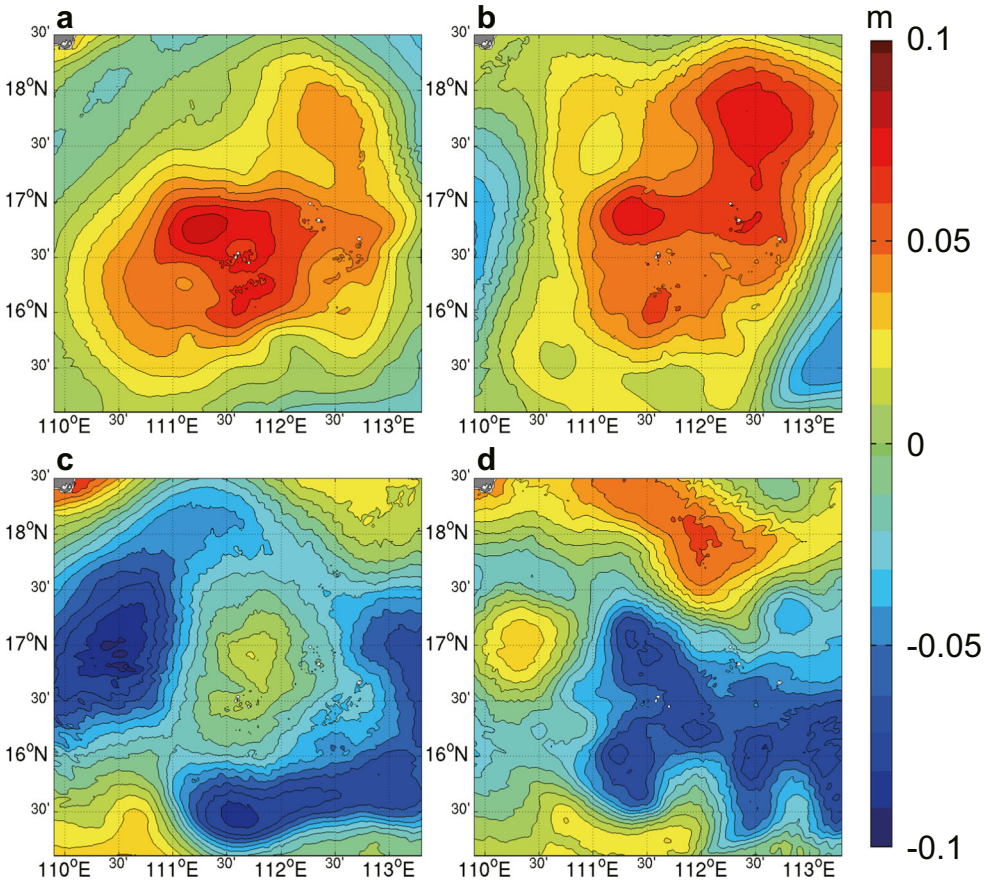

Fig. 7 The 72-h mean SLA forecasts (unit: m) from the **a** TRUTH, **b** In12, **c** SLA_DA, and **d** G_DA around the target region (In12, SLA_DA, and G_DA are indicated in Table 2)

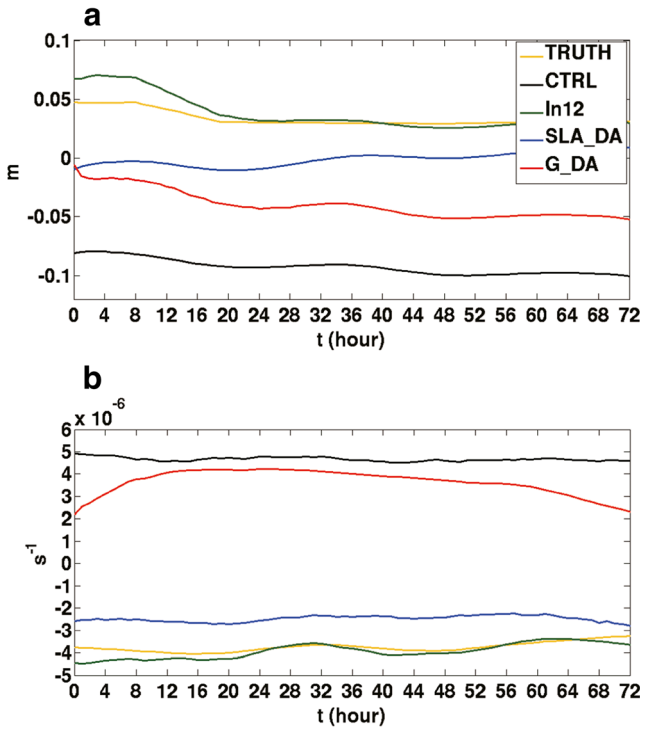

Fig. 8 The evolution of the spatial mean SLA (**a**, unit: m) and 0–200-m spatial mean relative vorticity (**b**, unit: s^{-1}) forecasts in the target region from the TRUTH, CTRL, In12, SLA_DA, and G_DA (CTRL, In12, SLA_DA, and G_DA are indicated in Table 2)

day. For simplicity, the following assumptions are made: (1) All 25 pseudo-underwater gliders are strictly synchronous in observing time and movement. (2) The observing times (locations) of T/S profiles are the times when (where) the pseudo-underwater gliders reach the sea surface. (3) The transmission and pre-processing time of the pseudo glider-observed T/S profiles are ignored. Two different movement strategies A and B of each pseudo-underwater glider are designed to investigate the impact of horizontal movement of underwater glider on the data assimilation (Fig. 3b, c). The pseudo-field observation experiment is conducted from 0000 UTC Aug. 4, 2010 to 0000 UTC Aug. 5, 2010, and the first T/S profile is obtained at 0004 UTC Aug. 4 2010 which is referred as number "1" in Fig. 3b, c.

3.2 OSSEs

The results from the 6-year daily simulations at 0000 UTC Aug. 1, 2010 are used as the initial conditions of the model to generate the true state (TRUTH) of OSSEs. To better reproduce the Xisha warm eddy in TRUTH, the daily satellite-derived and gridded SLA data with a horizontal resolution of 1/4° × 1/4° provided by Copernicus Marine environment monitoring service are assimilated into the model using MS-3DVAR at 0000 UTC from Aug. 1, 2010 to Aug. 4, 2010.

Table 3 The 72-h mean vertical-averaged RMSEs of the temperature, salinity forecasts, and the 72-h mean biases of SLA and 0–200-m mean relative vorticity forecasts averaged in the target region of the eddy validated against TRUTH for CTRL, the experiments in DA_RE

	CTRL	ReA025	ReB025	ReA05	ReB05	ReA075	ReB075
Temp (°C)	1.77	1.258	1.259	1.167	1.167	1.218	1.220
Salt (psu)	0.293	0.168	0.169	0.166	0.167	0.163	0.164
Mean SLA (m)		0.038		0.020		0.046	
Relative vorticity (10^{-6} s^{-1})		2.952		0.196		0.375	

Only the SLA data over the regions with water depth larger than 200 m are used in the data assimilation due to the poor quality of SLA data nearshore (Cipollini et al. 2010). After the assimilation of SLA, a free run starts from 0000 UTC Aug. 4, 2010 to 0000 UTC Aug. 8, 2010 and the results are used as the TRUTH (Fig. 4; Table 2). The pseudo-glider-observed T/S profiles and SLA observations are extracted from the TRUTH during 0000 UTC Aug. 4, 2010 to 0000 UTC Aug. 5, 2010 without any perturbation, among which the pseudo-SLA observations have the same resolution as the satellite-derived SLA assimilated in the TRUTH. According to the location of Xisha warm eddy, the center of pseudo-underwater glider observation network C is fixed at the point of 111.6°E, 16.8°N (Fig. 5).

The results from the 6-year daily simulations at 0000 UTC Aug. 4, 2009 when no significant warm eddy occurred in the region are used as the initial conditions for both the control experiment (CTRL) and the OSSEs. The CTRL starts from 0000 UTC Aug. 4, 2010 to 0000 UTC Aug. 8, 2010 without any data assimilation (Fig. 4; Table 2). In the OSSEs, the pseudo-glider-observed T/S profiles or/and SLA are assimilated through MS-3DVAR to optimize the initial conditions. Two sets of sensitivity experiments are designed in OSSEs, i.e., experiments with different horizontal deployment resolutions (denoted as DA_RE) and those with different assimilation frequencies (denoted as DA_IN). In DA_RE, the horizontal deployment resolutions are set to be 0.25° × 0.25°, 0.5° × 0.5° or 0.75° × 0.75°, respectively, corresponding to three different underwater glider deployment regions (Fig. 5). Combining the different horizontal deployment resolutions and different movement strategies A and B, we designed six experiments in DA_RE, which are denoted as ReA025, ReB025, ReA5, ReB05, ReA075, and ReB075 (Table 2). Each experiment in DA_RE starts from 0000 UTC Aug. 4, 2010. One hundred fifty (25 × 6) underwater glider-observed T/S profiles obtained during 0004 UTC Aug. 4, 2010 to 0000 UTC Aug. 5, 2010 are assimilated at 0000 UTC Aug. 5, 2010, followed by a 72-h model forecast (Fig. 4).

In DA_IN, the glider-observed T/S profiles with the same horizontal deployment resolution and movement strategy as those of ReA05 are continually assimilated into the model with different assimilation intervals of 4 h, 8 h, 12 h, and 24 h (denoted as In4, In8, In12, and In24) from 0000 UTC Aug. 4, 2010 to 0000 UTC Aug. 5, 2010, followed by a 72-h

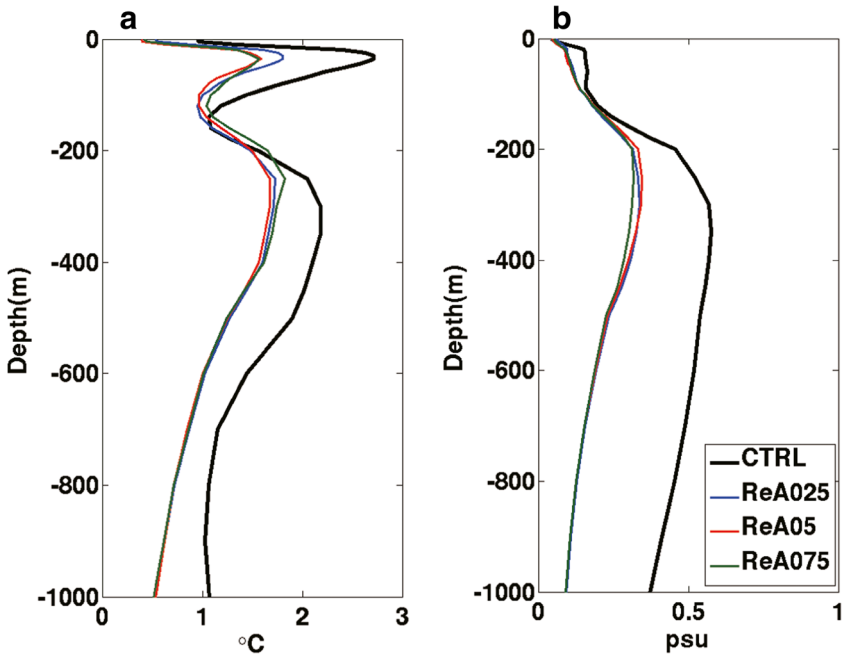

Fig. 9 The vertical profiles of 72-h mean RMSEs of temperature (**a**, unit: °C) and salinity (**b**, unit: psu) forecasts from CTRL (black line) and DA_RE (colorful lines) validated against the TRUTH in the target region (ReA025, ReA05, and ReA075 are indicated in Table 2)

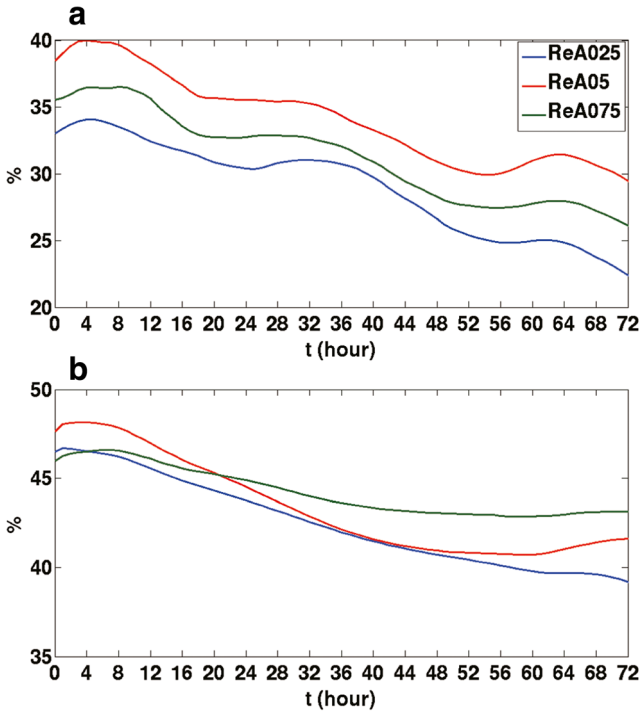

Fig. 10 The evolution of vertical-averaged RMSEs reductions of the forecasting temperature (unit: °C) (**a**) and salinity (unit: psu) (**b**) from experiments in DA_RE validated against the TRUTH in the target region (ReA025, ReA05, and ReA075 are indicated in Table 2)

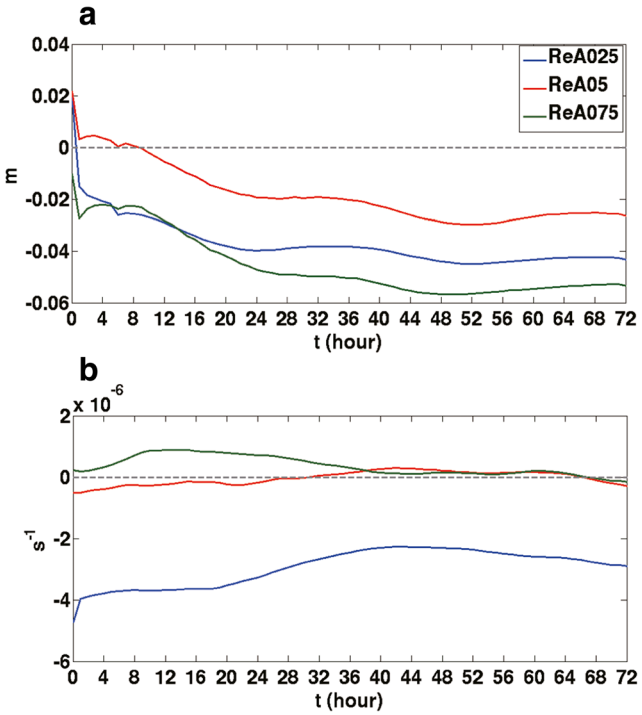

Fig. 12 The evolution of the spatial mean biases of SLA (**a**, unit: m) and 0–200-m mean relative vorticity (**b**, unit: s^{-1}) forecasts in the target region for ReA025, ReA05, and ReA075 (ReA025, ReA05, and ReA075 are indicated in Table 2)

forecast (Table 2; Fig. 4). The glider-observed T/S profiles assimilated in each assimilation cycle of DA_IN are those obtained during the cycle; as such, In24 is identical to ReA05. For example, in the first assimilation cycle of In8, the total 50 pseudo-glider-observed T/S profiles obtained at 0004 UTC and 0008 UTC will be assimilated at 0008 UTC Aug. 4, 2010. The pseudo-SLA is assimilated simultaneously along with pseudo-glider-observed T/S profiles at 0000 UTC Aug. 4 and 0000 UTC Aug. 5 in DA_RE and DA_IN. Besides, two experiments which assimilate only the pseudo-SLA (denoted as SLA_DA) or only the pseudo-glider-observed T/S profiles (denoted as G_DA) are conducted for evaluating the combined effect of assimilating the glider-observed T/S profiles and satellite-derived SLA (Fig. 4; Table 2). The configuration of G_DA is the same as that of In12 in DA_IN except for the absence of pseudo-SLA assimilation.

4 Result and discussion

To define the main area of Xisha warm eddy, the Okubo-Weiss criterion method (Isern-Fontanet et al., 2003) is used. The eddy center is surrounded by the closed contour line of

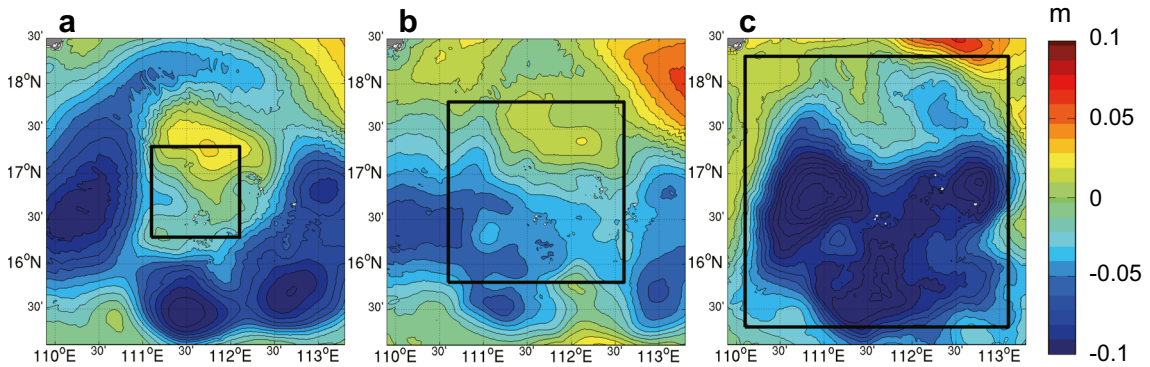

Fig. 11 The biases of 72-h mean SLA (unit: m) forecasts for ReA025 (**a**), ReA05 (**b**), and ReA075 (**c**); the black squares represent the observation area of the underwater glider network (CTRL, ReA025, ReA05, and ReA075 are indicated in Table 2)

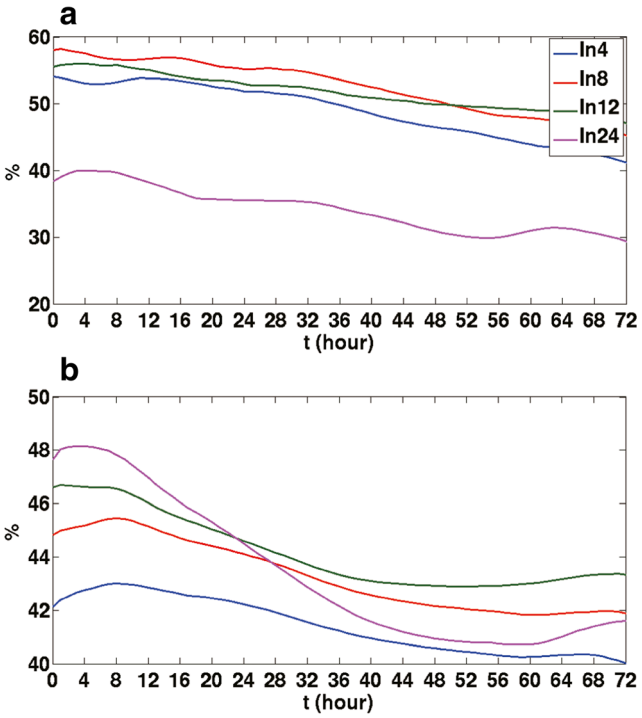

Fig. 13 The same as Fig. 10 except for the experiments in DA_IN (In4, In8, In12, and In24 are indicated in Table 2)

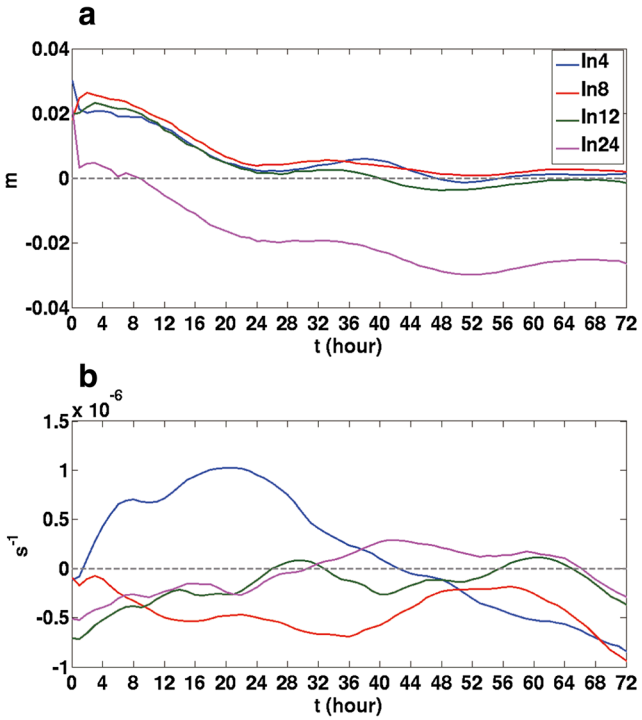

Fig. 14 The same as Fig. 12 except for the experiments in DA_IN (In4, In8, In12, and In24 are indicated in Table 2)

the Okubo-Weiss parameter $W = -2 \times 10^{-12}\ s^{-2}$, and we define the contour line as the edge of the eddy. By this method, the region surrounded by the mean edge of Xisha warm eddy during the TRUTH is chosen as the target region (Fig. 5). In the following evaluation, the temperature, salinity, SLA, and 0–200-m mean relative vorticity forecasts in the target region are considered, of which the relative vorticity Ω is defined by

$$\Omega = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \tag{6}$$

where u and v represent the perturbations of the sea current u- and v-component in the target region.

Firstly, the combined effect of assimilating the glider-observed T/S profiles and satellite-derived SLA is evaluated. The temperature and salinity forecasts from the In12 in DA_IN, SLA_DA, and G_DA are validated against the TRUTH in the target region (Fig. 6). The RMSEs of both the temperature and salinity forecasts from In12, SLA_DA, and G_DA are all reduced, of which the RMSEs reductions from In12 are significantly larger than those from SLA_DA and G_DA. In addition, the SLA forecasts around the target region show that the assimilation of pseudo-glider-observed T/S profiles help in maintaining the positive SLA during the 72-h forecast, which is close to the TRUTH (Fig. 7). Similar results are found in the evolution of the spatial mean SLA and 0–200-m spatial mean relative vorticity forecasts (Fig. 8). It indicates that assimilating either the glider-observed T/S profiles or the satellite-derived SLA can improve the forecast skill for the Xisha warm eddy, and assimilating both of them achieves the largest improvement.

Table 4 The same as Table 3 except for the experiments in DA_IN

	CTRL	In4	In8	In12	In24
Temp (°C)	1.777	0.903	0.837	0.854	1.167
Salt (psu)	0.293	0.171	0.166	0.164	0.166
Mean SLA (m)		0.006	0.008	0.006	0.020
Relative vorticity ($10^{-7}\ s^{-1}$)		5.203	4.389	2.026	1.959

As seen in Table 3, the RMSEs of temperature and salinity forecasts from the experiments with the two movement strategies are very close, though those from the experiments with movement strategy A are slightly smaller. It indicates that the meridional movement is a little more important than the zonal movement in this case. Therefore, the results with movement strategy B are not shown in the following analysis. The 72-h mean RMSEs of salinity forecasts reduce at all layers, while those of temperature forecasts increase at some layers for experiments in DA_RE (Fig. 9). Figure 10 shows the time series of vertical-averaged RMSEs reductions for temperature and salinity forecasts from experiments in DA_RE, from which the RMSEs reductions of both temperature salinity forecasts decreased with forecast time. The largest RMSEs reduction for temperature forecast is from the experiment ReA05 with a 72-h mean vertical-averaged RMSE of 1.167 °C (Table 3), while those for the salinity forecasts are close among all experiments. When looking at

the biases of the 72-h mean SLA around the target region which represents the characteristic of Xisha warm eddy, significant differences are found between the experiments with different horizontal deployment resolutions and ReA05 performs the best while ReA075 the worst (Fig. 11). Similar results can be found in the time series of spatial mean biases of SLA and 0–200 m spatial mean relative vorticities (Fig. 12), in which the smallest biases are from ReA05 with values of 0.020 m and 1.96×10^{-7} s^{-1}, respectively (Table 3). Given that the deployment region of ReA05 is very close to the target region, we conclude that the effect of assimilating glider-observed T/S profiles on the forecast of an extreme eddy is influenced by both the horizontal deployment resolution and deployment coverage. The best deployment strategy for the forecast of an extreme eddy is to cover the main area of the eddy with higher horizontal deployment resolution.

As shown in Fig. 13, assimilating the glider-observed T/S profiles in DA_IN can reduce over 30% and 40% RMSEs for temperature and salinity forecasts, respectively. The 72-h mean vertical-averaged RMSE of the temperature from In8 is the smallest and that of In24 is the largest among the experiments in DA_IN (Table 4). The RMSEs of temperature firstly decrease and then increase with the decrease of the assimilation frequency (Table 4). For the salinity forecasts, the vertical-averaged RMSE from In4 is the largest, while those from In8, In12, and In24 are close to each other (Table 4). The spatial mean bias evolutions of SLA from In24 are the largest and those from In8, In12, and In24 are close to each other (Table 4; Fig. 14a). The spatial mean biases of the 0–200 m mean relative vorticities from In4 and In8 are the smallest at the initial moment but fluctuate greatly after that (Fig. 14b), and the mean biases from In12 and In24 are close although that from In24 is the smallest (Table 4). It should be noted that although the numbers of glider-observed T/S profiles assimilated in each assimilation cycle are different for different experiments in DA_IN, the total amount of T/S profiles is the same and the coverage of the profiles is similar during the 24 h. These results imply that assimilation with very high frequency will break the dynamical balance between the variables and weaken the forecast skill of the mesoscale eddy (In4 and In8 in Fig. 14b), while assimilation with very low frequency results in less temperature improvement (In24 in Fig. 13a) and more SLA biases with integration (Fig. 14a). Considering the forecast results of the temperature, salinity, SLA, and relative vorticity, the optimal assimilation interval in DA_IN is 12 h.

5 Summary

In this study, the OSSEs are employed to investigate the effect of assimilating glider-observed T/S profiles regarding the horizontal resolutions of underwater glider deployment and assimilation frequencies, as well as the combination of assimilating satellite-derived SLA, on the forecast skill for the extreme warm eddy in the Northwestern SCS. The results can be summarized as follows:

1. Although assimilating either the glider-observed T/S profiles or the satellite-derived SLA can improve the forecast skill for the Xisha warm eddy, assimilating both of them achieves the largest improvement
2. The horizontal resolution of $0.5° \times 0.5°$ for underwater glider deployment gains the best forecast skill for the Xisha warm eddy. This implies that, under the premise of a full coverage of the eddy, the higher horizontal resolution of underwater glider deployment is, the better forecast skill will be obtained
3. The assimilation of the glider-observed T/S profiles with a 12-h interval achieves the best forecast skill for the Xisha warm eddy accounting for the improvements of temperature, salinity, SLA, and relative vorticity, instead of a higher assimilation frequency (shorter assimilation interval). The reason could be that the assimilation with a higher frequency will break the dynamical balance between the variables and thus degrade the forecast skill of the eddy.
4. The difference of data assimilation effects on the forecast skill for the Xisha warm eddy regarding different strategies of the underwater glider horizontal movement is minor, although that with meridional movement is slightly better.

The results of OSSEs obtained in this study can provide valuable reference for the deployment of underwater gliders as well as the assimilation strategy of glider observations. Based on these results, using the real glider observations to improve the forecast skill of the real-time forecast system for the northwestern SCS will be our future work.

Acknowledgments The authors gratefully acknowledge the use of the HPCC at the South China Sea Institute of Oceanology, Chinese Academy of Sciences.

Funding information This work was jointly supported by Major Projects of the National Natural Science Foundation of China (grant number 41890851 and 41931182), Innovation Research Group of National Natural Science Foundation of China (grant number 41521005), Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou, grant number GML2019ZD0303), National Natural Science Foundation of China (grant number 41676016 and 41776028), Guangdong Key Project (grant number 2019BT2H594), Strategic Priority Research Program of the Chinese Academy of Sciences (grant number XDA13030103 and XDA19060503), CAS/SAFEA International Partnership Program for Creative Research Teams, Chinese Academy of Sciences (grant number ZDRW-XH-2019-2 and ISEE2018PY05), and the Independent Research Project Program of State Key Laboratory of Tropical Oceanography (LTOZZ1902).

References

Arnold CP, Dey CH (1986) Observing-systems simulation experiments: past, present, and future. Bull Am Meteorol Soc 67:687–695

Blockley EW, Martin MJ, McLaren AJ et al (2013) Recent development of the Met Office operational ocean forecasting system:an overview and assessment of the newGlobal FOAM forecasts. Geosci Model Dev Discuss 6:6219–6278. https://doi.org/10.5194/gmdd-6-6219-2013

Blumberg AF, Mellor GL (1987) A description of a three-dimensional coastal ocean circulation model. Three-dimensional coastal ocean models. In: Heaps N (ed) coastal and estuarine studies series, vol 4. Amer, Geophys, Union, pp 1–16

Carton JA, Chepurin G, Cao XH et al (2000a) A simple ocean data assimilation analysis of the global upper ocean 1950-95. Part I: Methodology J Phys Oceanogr 30:294–309

Carton JA, Chepurin G, Cao XH (2000b) A simple ocean data assimilation analysis of the global upper ocean 1950-95. Part II: Results J Phys Oceanogr 30:311–326

Chambers DP, Ries JC, Urban TJ (2003) Calibration and verification ofJason-1 using global along-track residuals with TOPEX. Mar Geod 26:305–317. https://doi.org/10.1080/01490410390256691

Chassignet EP, Hurlburt HE, Smedstad OM et al (2007) The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. J Mar Syst 65:60–83

Chen GX, Gan JP, Xie Q et al (2012) Eddy heat and salt transports in the South China Sea and their seasonal modulations. J Geophys Res-Oceans 117:C05021. https://doi.org/10.1029/2011jc007724

Chen GX, Hou YJ, Chu XQ (2011) Mesoscale eddies in the South China Sea: mean properties, spatiotemporal variability, and impact on thermohaline structure. J Geophys Res-Oceans 116, C06018. do1: https://doi.org/10.1029/2010jc006716

Chu XQ, Xue HJ, Qi YQ (2014) An exceptional anticyclonic eddy in the South China Sea in 2010. J Geophys Res-Oceans:119. https://doi.org/10.1002/2013JC009314

Cipollini P et al (2010) The role of altimetry in coastal observing systems, in Proceeding s of OceanObs'09: sustained ocean observations and information for society, vol. 2, In: Hall J, Harrison DE, and Stammer D (ed). Eur Space Agency Publ 2:WPP-306

Fang G, Fang W, Fang Y et al (1998) A survey of studies on the South China Sea upper ocean circulation. Acta Oceanogr Taiwan 37:1–16

Guan L, Kawamura H (2004) Merging satellite infrared and microwave SSTs: methodology and evaluation of the new SST. J Oceanogr 60: 905–912. https://doi.org/10.1007/s10872-004-5782

Hu J, Kawamura H, Hong H et al (2000) A review on the currents in the South China Sea: seasonal circulation, South China Sea warm current and Korishio intrusion. J Oceanogr 56:607–624

Isern-Fontanet J, Arcía-Ladona E, Font J (2003) Identification of marine eddies from altimetric maps. J Atmos Ocean Technol 20:772–778

Kanamitsu M, Ebisuzaki W, Woollen J et al (2002) NCEP-DOE AMIP-II reanalysis (R-2). Bull Amer Meteor Soc 83:1631–1643

Li JX, Zhang R, Jin BG (2011) Eddy characteristics in the northern South China Sea was inferred from lagrangian drifter data. Ocean Sci 7: 661–669

Liu Q, Kaneko A, Jilan S (2008) Recent progress in studies of the South China Sea circulation. J Oceanogr 64:753–762

Li ZJ, Chao Y, Mcwilliams JC et al (2008a) A three-dimensional variational data assimilation scheme for the regional ocean modeling system. J Atmos Ocean Technol 25:2074–2090

Li Z, Chao Y, Mcwilliams JC et al (2008b) A three-dimensional variational data assimilation scheme for the Regional Ocean Modeling System: implementation and basic experiments. J Geophys Res 113: C05002

Masutani M, Woollen JS, Lord SJ et al (2010) Observing system simulation experiments at the National Centers for Environmental Prediction. J Geophys Res-Atmos 115

McGillicuddy DJ (2016) Mechanisms of physical-biological-biogeochemical interaction at the oceanic mesoscale. Annual Review of Marine Science 8, annurev-marine-010814-015606

Mellor GL (2003) Users guide for a three-dimensional, primitive equation, numerical ocean model. Program in Atmospheric and Oceanic Sciences, Princeton University 53 pp.

Miyazawa Y, Varlamov SM, Miyama T et al (2017) Assimilation of high-resolution sea surface temperature data into an operational nowcast/forecast system around Japan using a multi-scale three-dimensional variational scheme. Ocean Dyn 67:713–728

Parrish DF, Derber JC (1992) The National Meteorological Center's spectral statistical-interpolation analysis system. Mon Weather Rev 120: 1747–1763. https://doi.org/10.1175/1520-0493(1992)120<1747:tnmcss>2.0.co;2

Peng SQ, Zeng XZ, Li ZJ (2016) A three-dimensional variational data assimilation system for the South China Sea: preliminary results from observing system simulation experiments. Ocean Dyn 66: 737–750

Peng SQ, Zhu YH, Li ZJ, Li Y, Xie Q, Liu S, Luo Y, Tian Y, Yu J (2019) Improving the real-time marine forecasting of the northern South China Sea by assimilation of glider-observed T/S profiles. Sci Rep 9:17845. https://doi.org/10.1038/s41598-019-54241-8

Qu TD (2000) Upper-layer circulation in the South China Sea. J Phys Oceanogr 30:1450–1460

Rhodes RC, Hurlburt HE, Wallcraft AJ et al (2002) Navy real-time global modeling systems. Oceanography 15:16. https://doi.org/10.5670/oceanog.2002.34

Shu YQ, Xiu P, Xue HJ et al (2016) Glider-observed anticyclonic eddy in northern South China Sea. Aquat Ecosyst Health 19:233–241

Smedstad OM, Hurlburt HE, Metzger EJ et al (2003) An operational real-time eddy-resolving 1/16° global ocean nowcast/forecast system. J Mar Syst 40:341–361

Song YT (2006) Estimation of interbasin transport using ocean bottom pressure: theory and model for Asian marginal seas. J Geophys Res 111:C11S19. https://doi.org/10.1029/2005JC003189

Wang D, Liu Q, Huang RX et al (2006) Interannual variability of the South China Sea throughflow inferred from wind data and an ocean data assimilation product. Geophys Res Lett 33:L14605

Wang DX, Liu QY, Xie Q et al (2013) Progress of regional oceanography study associated with western boundary current in the South China Sea. Chin Sci Bull 58:1205–1215

Yang QX, Zhou L, Tian JW et al (2013) The roles of Kuroshio intrusion and mesoscale eddy in upper mixing in the northern South China Sea. J Coast Res 30:192–198. https://doi.org/10.2112/jcoastres-d-13-00012.1

Yu JC, Zhang AQ, Jin WM et al (2011) Development and experiments of the sea wing underwater glider. China Ocean Eng 25:721–736

Yu JC, Zhang FM, Zhang AQ et al (2013) Motion parameter optimization and sensor scheduling for the sea-wing underwater glider. IEEE J Ocean Eng 38:243–254

Yu Z, Shen S, McCreary JP et al (2007) South China Sea throughflow as evidenced by satellite images and numerical experiments. Geophys Res Lett 34:L01601. https://doi.org/10.1029/2006GL028103.s