

Assimilation of mid-depth velocities from Argo floats in the western South China Sea

Pinqiang Wang¹, Weimin Zhang^{1,2}, Huizan Wang¹, Haijin Dai¹, Xiaohui Wang^{1,3}

1. College of Meteorology and Oceanography, National University of Defense
Technology, Changsha, China, 410073

2. Laboratory of Software Engineering for Complex Systems, Changsha, China,
410073

3. Delft Institute of Applied Mathematics, Delft University of Technology, Delft,
Netherlands

Corresponding author address: W. Zhang, College of Meteorology and Oceanography,
National University of Defense Technology, Deya Road, Changsha, China, 410073.
Email: wmzhang104@139.com

Abstract

Previous studies are mainly limited to temperature and salinity (T/S) profiling data assimilation, while data assimilation based on Argo float trajectory information has received less research focus. In this study, a new method was proposed to assimilate Argo trajectory data: the mid-depth (indicates the parking depth of Argo floats in this study, ~1200 m) velocities are estimated from Argo trajectories and subsequently assimilated into the Regional Ocean Model System (ROMS) Model using 4DVAR (4-dimensional variational) method. This method can avoid a complicated float trajectory model in direct position assimilation. The two months assimilation experiments in South China Sea (SCS) showed that this proposed method can effectively assimilate Argo trajectory information into the model and improve mid-depth velocity field by adjusting the unbalanced component in the velocity increments. The assimilation of the Argo trajectory-derived mid-depth velocity with other observations (satellite observations and T/S profiling data) together yielded the best performance, and the velocity fields at the float parking depth are more consistent with the Argo float trajectories. In addition, this method will not decrease the assimilation performance of other observations (i.e., sea level anomaly (SLA), sea surface temperature (SST), and T/S profiles), which is indicative of compatibility with other observations in the 4DVAR assimilation system.

Keywords: Argo trajectory, data assimilation, mid-depth velocity, South China Sea

58 **1 Introduction**

59 As the largest marginal sea in the west Pacific Ocean, the South China Sea (SCS)
60 is strongly influenced by the wind, which has intraseasonal variations. As a result, the
61 upper layer circulation of the SCS is characterized by anticyclonic circulation in the
62 southern SCS, cyclonic circulation in the northern SCS during summer and cyclonic
63 circulation during winter because of the wind direction reversal (Wyrтки 1961; Hu et al.
64 2000; Liu et al. 2008). The current structures below 500 m are poorly understood and
65 less studied than the upper circulation due to the lack of observations. The intermediate
66 circulation studies are mainly based on the diagnosis of hydrography data (Chu 2000),
67 numerical simulations (Chao et al. 1996; Yuan 2002) and some *in situ* observations
68 (Liao 1996). The circulation in the middle layer of the SCS appears as an overall
69 anticyclonic circulation, driven by the outflow from the SCS to the Pacific Ocean in the
70 Luzon Strait (Yuan 2002). However, the intermediate circulation still lacks accurate
71 description due to the lack of comprehensive *in situ* observations and the numerical
72 model drawbacks.

73 Strong western boundary currents (Fang et al. 2012; Wang et al. 2013; Quan et al.
74 2016) and mesoscale eddies (Wang et al. 2003; Xiu et al. 2010; Chen et al. 2011; Li et
75 al. 2011; Lin et al. 2015; Zhang et al. 2016) could also be easily detected from the
76 surface observations, i.e., many mesoscale eddy activities are observed east of Vietnam.
77 However, some mesoscale eddies appear in the subsurface and could not be observed
78 on the surface (Song et al. 2019). Two anticyclonic eddies were detected in the

subsurface of the southern SCS with *in situ* hydrography data (Zhang et al. 2014). Chen et al. (2015) found a mesoscale eddy in the deep layer of the northwestern SCS with mooring ADCPs. How the ocean circulation model captures the structure of the mesoscale eddy in the subsurface still needs to be researched.

An effective tool to improve our understanding of oceanography (i.e., intermediate circulation or mesoscale eddy) is data assimilation, which combines limited observations and numerical model. The previous ocean data assimilation in the SCS is mostly based on satellite data and *in situ* T/S profiling data and mainly improves simulation of the temperature and salinity fields in the upper 1000 m. Data assimilation for the velocity field in the middle layer is usually neglected in the SCS due to a lack of velocity observations.

The international Argo program, which has been implemented since 2000, has enriched the number of observations in the middle layer to a certain extent (Wang et al. 2012a, b). Currently, the number of active floats in the global ocean is approximately 3900. Despite monitoring of the temperature and salinity in the upper ocean, the Argo float trajectory data also provide another way to study the current structure in the middle layer of the ocean (Park et al. 2004). Over last decade, a variety of methods have been explored for Lagrangian data assimilation (Kuznetsov et al. 2003; Molcard et al. 2003, 2005; Spiller et al. 2008; Muscarella et al. 2014; Slivinski et al. 2015). However, most of the studies are using simple model and synthetic observations, i.e. from a model forecast. In fact, profile floats trajectories are influenced by vertical shear, surface drift

and topography.

In the Mediterranean Forecasting System (MFS), the position information of the Argo trajectories has been directly assimilated with other observations (i.e., sea surface height (SSH), SST, and T/S profiles). Retaining the state of the upper layer, simulation of the velocity structure at parking depth could be improved by 15% with Argo float trajectory data assimilation (Taillandier et al. 2006a, b; Nilsson et al. 2011, 2012). One big challenge on direct positions assimilation is estimation of appropriate covariances, which is inherent nonlinearity (Castellari et al. 2001; Slivinski et al. 2015). Particle filter method can deal with the nonlinearity (Slivinski et al. 2015), but it is rarely used in ocean data assimilation. Instead of assimilating the original trajectory directly, we estimate the mid-depth (~1200m) velocity field from the Argo float trajectory and then assimilate the estimated velocity observation into the model to circumvent this problem. Since the velocity is already included in the ocean model, the usage of mid-depth velocity observations allows us to forgo additional trajectory prediction models, which is a difficult task. In addition, the assimilation of Argo trajectories has not yet been used in the SCS, which is another focus of this study.

The study is organized as follows: In Section 2, the mid-depth velocity estimation method from Argo trajectories and corresponding quality control are introduced. In Section 3, the experimental setup of ocean data assimilation for the estimated velocity at the parking depth is described. In Section 4, the simulation results are analyzed. Finally, a conclusion and suggested corresponding future work are presented in Section

121 5.

122 **2 Velocity retrieval and 4DVAR**

123 The proposed Argo trajectory assimilation method is composed of the estimation
124 and assimilation of mid-depth velocity. In this method, the quality control and
125 observation error estimation of the velocity data are two important aspects. The
126 assimilation technique uses 4-dimensional variational data assimilation (4DVAR). The
127 background error covariance in 4DVAR is implicitly developed with the ocean state
128 during the assimilation time window; thus, 4DVAR is more suitable than 3DVAR for
129 data assimilation of sparse velocity observations.

130 **2.1 Mid-depth velocity retrieval**

131 Several methods have been provided to estimate the velocities at parking depth
132 (Park et al. 2005; Xie and Zhu 2008; Lebedev et al. 2007; Ollitrault et al. 2013), which
133 were then applied to estimate the regional or global ocean circulation at the basin scale
134 (Voet et al. 2010; Park and Kim 2013; Ollitrault and ColindeVerdiere 2014; Markova
135 and Bagaev 2016; Wang et al. 2018). In this paper, the method provided by Park et al.
136 (2005) is adopted due to the method's high level of accuracy.

137 The estimation of mid-depth velocity (\mathbf{v}_{deep}) is based on:

$$\mathbf{v}_{deep} = \Delta \mathbf{r} / \Delta t \quad (1)$$

138 where $\Delta \mathbf{r}$ and Δt is the drift distance and time within the subsurface. The difference
139 between different methods is how the float trajectory is treated on the surface, i.e. how

the descend start point (P_{DS}^n) and ascend end point (P_{AE}^n) is treated. In Park's method, one advantage is taking the linear velocity and the inertial velocity into consideration during the accurate estimation P_{DS}^n and P_{AE}^n . The other advantage is that instead of using linear extra interpolation, an extra interpolation method based on a least square fit of trajectory positions is adopted to estimate the P_{DS}^n and P_{AE}^n and surface velocity. According to Park et al (2005), the subsurface velocity (\mathbf{v}_{deep}^n) at the parking depth can be written as follows:

$$\mathbf{v}_{deep}^n = \frac{P_{AE}^{*n} - P_{DS}^{*n}}{\Delta t} + E_{AE}^n - E_{DS}^n - E_T^n \quad (2)$$

where P_{DS}^{*n} and P_{AE}^{*n} are the estimation of P_{DS}^n and P_{AE}^n , E_{AE}^n and E_{DS}^n are the errors due to the inaccuracy of P_{DS}^n and P_{AE}^n , respectively; E_T^n is the error related to velocity shear. Although the Argo trajectories include the tidal current information, according to the previous studies, the error introduced by tides are much smaller than the error introduced by the inaccuracy of satellite surface positioning and vertical velocity shear (Park et al. 2004; Lebedev et al. 2007; Ollitrault et al. 2013). Thus, the time averaged mid-depth velocity from Argo trajectories in this study are treated as not include tides. Park et al. (2005) estimated the total error ($E_{AE}^n - E_{DS}^n - E_T^n$), about 0.2 cm/s with over 6 satellite position points at surface, which confirms the accuracy of the estimated velocity.

2.2 Quality control procedure for velocity

Before assimilation into the model, two rounds of quality control are employed for higher velocity accuracy. The first round of quality control is applied to the original

trajectories of the Argo float before velocity estimation, which is completed in four steps (Wang et al. 2018). The first step is removing the points if the distance between two adjacent points is larger than 1.5° ; the second step is eliminating data if the time does not satisfied $T_N^{n-1} < T_{DS}^n < T_{AE}^n < T_1^n$ (T_N^{n-1} is the last satellite positioning point in the $n-1$ th cycle and T_1^n is the first satellite positioning point in the n th cycle); in the third step, data are discarded if the satellite observational position is less than one in the n th cycle; and in the fourth step, the time information of the nearest satellite position is used if T_{DS}^n or T_{AE}^n is missing.

The second round of quality control is applied to the velocity at parking depth, which is finished in two steps. The first step is discarding data with a speed of less than 1 cm/s (account for 10.1% of the total), which is comparable to the total error defined in Section 2.1 (approximately 0.2 cm/s) (Park et al. 2005). The second step is discarding velocity if the parking depth changed dramatically during the adjacent cycles (account for 7% of the total), for example, the parking depth jumped from 1200m to several hundred meters suddenly, and jumped back to 1200m.

After quality control, the mid-depth velocities have been used to estimate the current structure of the basin scale in the SCS (Wang et al. 2018), indicating a cyclonic circulation at 1200m in northern SCS. The RMSE of mid-depth velocities with YoMaHa'07 (Lebedev et al. 2007) product is 1.28 cm/s for zonal and 1.11 cm/s for meridional (Wang et al. 2018).

It should be mentioned that the mid-depth velocity obtained in section 2.1 is time

averaged velocity during the drifting period in the subsurface instead of the instantaneous velocity. Correspondingly, the background field should also use the time averaged velocity field during the same parking period, which is difficult to satisfy due to the variation in the drifting period for different Argo floats. As an alternative choice, the middle time and position during the drifting period is considered to be the observation time and position in this study.

3 Experimental setup for trajectory-derived velocity assimilation

The western SCS is characterized by a dominant western boundary current (SCSWBC) and an active mesoscale process. The mean speed is approximately 12~16 cm/s at 1000 m in the western SCS (Zhou et al. 2010). When an Argo float moved to the western SCS, the Argo float was easier to be captured by the dominant processes. Thus, the western SCS is selected as the area in which to study the impact of the Argo trajectory data assimilation on the mid-depth current structure. During the study period, No. 5903457 had just moved into the western SCS and therefore it was studied with an intense focus.

3.1 Estimation of model states

The assimilations system is SCS 4DVAR system, which combined the Regional Ocean Modeling System (ROMS, version 3.7; <https://www.myroms.org>) and observations using the primal formulation of incremental strong constraint 4D-Var (IS4DVAR) (Moore et al. 2011a, b). In ROMS-IS4DVAR method, the best estimation

201 of ocean states can be obtained by minimizing the cost function:

$$\begin{aligned}
J(\delta \mathbf{x}, \delta \mathbf{f}, \delta \mathbf{b}) = & \frac{1}{2} \delta \mathbf{x}^T \mathbf{B}_x^{-1} \delta \mathbf{x} + \frac{1}{2} (\mathbf{H} \delta \mathbf{x} - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{H} \delta \mathbf{x} - \mathbf{d}) \\
& + \frac{1}{2} \delta \mathbf{f}^T \mathbf{B}_f^{-1} \delta \mathbf{f} + \frac{1}{2} \delta \mathbf{b}^T \mathbf{B}_b^{-1} \delta \mathbf{b}
\end{aligned} \tag{3}$$

202 where \mathbf{x} denotes the state vector (T, S, ζ, u, v), \mathbf{f} denotes the atmosphere forcing (wind
203 stress, heat flux) and \mathbf{b} denotes the lateral open boundary conditions. $\mathbf{d} = \mathbf{y} - H(\mathbf{x})$ is
204 the innovation vector which represents the difference between observations and model
205 analogue in observation space. $\mathbf{B}_x, \mathbf{B}_f, \mathbf{B}_b$ and \mathbf{R} are the error covariance matrix for
206 background field, surface forcing, lateral boundary conditions and observations.

207 In ROMS-IS4DVAR, the increments of model states are separated into balanced
208 and unbalanced components. All the balance components are based on the T increments
209 δT , which is estimated by minimizing the cost function (3) (Weaver et al. 2005). The
210 balanced S field is estimated through nonlinear TS relationship; the balanced density
211 field is estimated through nonlinear equation of state of δT and δS ; SSH is diagnosed
212 as a function of $\delta \rho$ (baroclinic part) and $(\delta u, \delta v)$ (barotropic part); the balanced
213 pressure increment δp is computed by hydrostatic equation; finally, the dynamically
214 consistent velocity field is obtained through geostrophic balance except near the equator.
215 Thus, a dynamically consistent mass field is obtained firstly, then a dynamically
216 consistent velocity field. The mid-depth velocity assimilation influenced the model
217 velocity field through the unbalanced components, which propagated horizontally and
218 vertically through correlation matrix (expressed as the solution of a diffusion equation),
219 then influenced other variables through model dynamics.

The assimilation procedure is shown in Fig. 1. Observations are continuously assimilated into the model in all assimilation experiments every 4 days (equal to the working cycle of Argo floats in SCS). Specifically, in one cycle, for example, the model integrates from 20130101 to 20130105 (denoted as Prior, Fig. 1 green line) to provide background fields and extract model value in observation space. Then the model states are adjusted through 4DVAR. After assimilation, the model is initialized from the analysis field at 20130101 and integrated to 20130105 again (denoted as Posterior, Fig. 1 red line). In next cycle, the model starts from the end states of previous cycle (20130101-20130105) and repeats the progress until the end date (20130228). All experiments were conducted for two months, with totally 15 data assimilation cycles.

3.2 Model configuration

The Generic Length Scale (GLS; Warner et al. 2005), $k-\omega$ vertical mixing scheme, harmonic horizontal viscosity (mixing coefficient: $4 \text{ m}^2\text{s}^{-1}$), no slip boundary conditions, quadratic bottom friction (bottom drag coefficient: 2.5×10^{-3}) and sponge layer are employed as the SCS ROMS model configuration. The model domain covers part of the northwestern Pacific Ocean ($99^\circ\text{E}\sim 134^\circ\text{E}$, $1^\circ\text{N}\sim 30^\circ\text{N}$) with a horizontal resolution of $1/10^\circ$ and 24 vertical levels. The bathymetry is derived from GEBCO08 with a minimum depth of 10 m and a maximum depth of 5500 m (Fig. 2).

The model is initialized from climatological Simple Ocean Data Assimilation (SODA) data and is forced by a climatological forcing (COADS) for 10 years, which is considered to be the model spin-up process. Then, the model is integrated for the

period of 2000-2015, which is forced by real surface forcing and boundary conditions. The heat and freshwater fluxes are extracted from ECMWF's ERA-interim dataset (<https://apps.ecmwf.int/datasets/data/interim-full-daily>) with a horizontal resolution of 0.75° and a temporal resolution of 6 hours (Dee et al. 2011). The wind stresses are calculated using the bulk formula with 10-meter U-wind and V-wind components from the CCMP 2.0 version (Atlas et al. 2011; <ftp://ftp2.remss.com>) with a horizontal resolution of 0.25° and a temporal resolution of 6 hours. The southern, eastern and northern boundaries adopt open boundary conditions from SODA version 3.3.1 (Carton et al. 2018; <https://www.atmos.umd.edu/~ocean/>), while the west boundary is closed. In detail, Radiation and Nudging boundary conditions are chosen as the 3D variables (i.e., u, v, temperature and salinity), Flather boundary conditions are chosen for the depth-integrated velocity components, and Clamped boundary conditions are chosen for the free surface.

Tides are not included in our model. It is challenging for assimilating altimetry data into a model include tide, since tides are high frequency signal compared to the mesoscale variability. Xie et al (2011) compared the model fields with or without tides, the differences are minor on monthly averaged velocity field and the influence on temperature and salinity decreased with depths below 100m. The T/S difference between with or without tide are about $0.2^\circ\text{C}/0.05\text{PSU}$ (Wang et al., 2017). For simplicity, tides are excluded in the model in this study.

3.3 Observations

The observations used in the data assimilation experiments are satellite altimeter data (SSH), satellite sea surface temperature (SST), *in situ* T/S profiling data and Argo trajectory-derived mid-depth velocity data. The SSH data are delayed time, and gridded maps of sea level anomaly (MSLA) data from AVISO (before 2016, <http://www.aviso.oceanobs.com/duacs/>). The blended SST data used in the experiments are Version 2 AVHRR-only products produced by NOAA (Reynolds et al. 2007; <ftp://eclipse.ncdc.noaa.gov>). Both SSH and SST data are available every day with a horizontal spatial resolution of 0.25° . The *in situ* data (Argo, CTD and XBT, etc.) are the quality-controlled temperature and salinity profiles from the ENSEMBLES project (EN4.2.1) provided by the UK Met Office Hadley Center (Good et al. 2013; <http://hadobs.metoffice.com/en4/download.html>, Fig.3a). Before assimilated into the model, the T/S profiles were interpolated to 23 z-levels: -5 m, -10 m, -15 m, -20 m, -25 m, -30 m, -35 m, -40 m, -50 m, -60 m, -75 m, -100 m, -125 m, -150 m, -200 m, -250 m, -300 m, -400 m, -500 m, -600 m, -800 m, -1000 m, and -1200 m. This interpolated data may cause redundancy somewhere due to the ROMS terrain-following coordinate system. However, some error data in a single profile will be eliminated simultaneously. The Argo trajectory data were provided by the Coriolis Operational Oceanography data center (<ftp://ftp.ifremer.fr/ifremer/argo>).

The observation errors are assumed to be uncorrelated at any time or at any point as defined in most ocean data assimilation systems. The variances along the main

diagonal of observation error covariance are a combination of measurement error and representative error. Measurement errors from different data sources are obtained with the following standard deviations (Moore et al. 2011b): 2 cm for SSH; 0.4 °C for SST; 0.1 °C for *in situ* temperature (T); and 0.01 PSU for *in situ* salinity (S).

For the mid-depth velocity, since the Argo positions are not perfect, a total error is given without distinguishing different error sources which mentioned in Section 2.1. This total error is estimated through HYbrid Coordinate Ocean Model - Navy Coupled Ocean Data Assimilation product (HYCOM-NCODA; http://tds.hycom.org/thredds/dodsC/GLBu0.08/exp_90.9) (as shown in Fig. 4). Most of the misfits are distributed between ± 5 cm/s (accounting for approximately 83% of the total number), and the standard deviation of these misfits (between ± 5 cm/s) is about 2 cm/s. Thus, the observation error has been tested with standard deviation 2 cm/s and 5 cm/s for the 2-month assimilation. The total root mean square errors (RMSEs) and correlation coefficients (CCs) of the mid-depth (u , v) with Argo trajectory-derived (u , v) are shown in Fig. 5 (since there are no direct velocity observations, the Argo trajectory-derived velocities are used as the true velocities). As the 2 cm/s velocity error performed better than the 5 cm/s velocity error, with smaller RMSEs (2.6 cm/s for u and 2.8 cm/s for v) and higher CCs (0.66 for u and 0.69 for v). The observation error of the mid-depth is set to 2 cm/s in this study.

3.4 Experimental design

Four experiments were designed to evaluate the Lagrangian velocity assimilation

influence on the model forecast (Table 1). The reference experiment does not assimilate any data (Ctrl). In Exp1, only satellite data (SST and SLA) are assimilated. In Exp2 both satellite data and *in situ* T/S profile data are assimilated. In Exp3, additional Argo trajectory-derived mid-depth velocities are added to the Exp2 dataset. In Exp1 to Exp3, the assimilation window is 4 day. All experiments begin from the same initial conditions (model's the real forecast on 1 January, 2013) and are driven by the same surface forcing and boundary conditions. To evaluate the effects of assimilating different data more accurately, the climatological nudging has been closed in all assimilation experiments.

4 Results analysis

During the first two months in 2013, there were 88 active Argo floats in the whole model region, among which 23 Argo floats were distributed in SCS. After quality control, these floats provided 382 mid-depth velocity observations in Jan and Feb 2013 (Fig. 3a). The mean parking depth is 1179 m, and 358 observations deeper than 1100 m (accounting for ~94% of the total number). The mean speed is 3.4 cm/s with a maximum speed of 13.1 cm/s. The Argo floats were mainly irregularly distributed in the northern SCS (Fig. 3a). Two floats are in western SCS, No. 5903457 float completed a cycle near the east coast of Vietnam in Jan-Feb 2013 and No.5902163 drifted southwestward. The following analysis is mainly based on the No. 5903457 float (Fig. 3b) and with less focus on No.5902163.

4.1 The impacts of different assimilation experiments in the western SCS

The impacts of assimilating different observations were compared based on the 2-month averaged current structure at 1200 m depth, as shown in Fig. 6. An anticyclonic eddy appeared at $\sim 14^\circ\text{N}$ in all experiments, while apparent differences are identified, such as eddy radius, position of eddy center (EC) and speed. These observation influences on mid-depth velocity field can be expressed as:

$$\delta \mathbf{v}^k = K_{vT} \delta T^k + K_{vS} \delta S^k + K_{v\zeta} \delta \zeta^k + \delta \mathbf{v}_U^k \quad (4)$$

where K_{vT} , K_{vS} , $K_{v\zeta}$ are balance operator based on geostrophic balance. δT^k , δS^k and $\delta \zeta^k$ are balanced increments of T , S and SSH , $\delta \mathbf{v}_U^k$ is the unbalanced increment of velocities (u , v).

In Exp1, the assimilation of SLA and SST introduced barotropic and baroclinic adjustment into velocity field, among which the barotropic adjustment is dominant, and yielded the largest radius of an eddy in Exp1 (Fig. 6b) at approximately 104 km (defined as the distance of model EC to the western boundary). The barotropic and baroclinic geostrophic adjustment together narrowed the radius of the eddy to 74 km in Exp2 (Fig. 6c). However, the increments introduced by balance operator (i.e., $K_{vT} \delta T^k + K_{vS} \delta S^k + K_{v\zeta} \delta \zeta^k$) are part of the geostrophic currents (i.e., $\delta \mathbf{v}^k$). In Exp3, the radius further decreased to 61 km due to the introduction of mid-depth velocities (Fig. 6d). Except the balanced components through geostrophic adjustment, the unbalanced components (i.e., $\delta \mathbf{v}_U^k$) are also adjusted in Exp3.

Along with the changes in radius, the position of the EC changed. The EC of Exp3

is obviously closest to the EC of the No. 5903457 float trajectory. The EC in Exp2 is closer than in Exp1, proving that assimilation of *in situ* T/S profile observations is necessary for improving subsurface states. The EC misfits between the model and trajectories are listed in Table 2. The misfit between Exp3 and the observations is the smallest (approximately 0.9 km), followed by Exp2 (approximately 25.6 km). When the balanced components in the velocity increment are overly strong, an opposite correction is made through the introduction of unbalanced components in the velocity increment (Fig. 6e).

The impact on trajectory prediction had also been compared in Fig.6. Due to the topography, the particles mainly went northward in Jan, 2013, which could be seen in all assimilation experiments. However, obvious differences occurred in Feb, 2013. The particles released in Exp3 moved almost along the realistic trajectories, while in Exp1 and Exp2 the particles drifted away from realistic trajectories. The movements of particles, approximately two-dimensional in 4 days, were in agreement with the Eulerian velocity field. Thus, an improvement of Lagrangian velocities could be expected with the improvement of Eulerian velocity field. This improvement was quantified in terms of relative differences (%) between Exp3 and Exp2 float position misfits (16.4 and 18.2 km, respectively, in Table 2) and found to be about 7% (take the Ctrl as a reference).

Except for No.5903457 float, the float No.5902163 is also in the western SCS. All experiments showed that No.5902163 float affected by a cyclonic eddy and move

southward in Jan, 2013 (Fig. 7). However, the velocity field in Exp3 was closer to the No.5902163 trajectory after the float left land boundary at the end of Jan (20130123 ~ 20130131) with introducing of mid-depth velocities. In Feb, the velocity field was affected by both floats (No.5903457 and No.5902163). As a result, the core of the anticyclonic eddy in Exp3 was distorted and closer to the north land boundary than in Exp2. Thus, this float can be captured by the south edge of eddy core in Exp3 and go westward.

4.2 Vertical propagation of corrections

The velocity corrections at ~1200 m are likely to propagate vertically due to the cross-correlation in the background error covariance, and propagate forward through the model dynamics. This vertical propagation has been studied above the parking depth of the No. 5903457 float (Fig. 9). In Exp2, the eddy structure at different depths is similar, such as the radius and position of EC. The EC distance between 500 m and 1000 m is only 20.9 km, which means that the eddy is almost vertically distributed without tilt. When the mid-depth velocities are assimilated, the radius (especially in the zonal direction) of the eddy decreases with depth, and the EC moves southwest, approaching the trajectory structure. The EC distance between 500 m and 1000 m increases to 75.5 km in Exp3, which means the EC tilts towards the southwest.

The vertical influence could also be seen from predicted trajectories at different depth. The predicted trajectories in Exp2 and Exp3 were consistent with Eulerian velocity field at different depth, as at parking depth. In Exp3, the position misfits

amplified with the increase of distance to parking depth, indicating that the influence of mid-depth velocity was weakened with the increase of distance. This could also be proved from Fig. 11 and Fig. 12, in which the difference between Exp2 and Exp3 are little at surface. This probably one reason for the EC tilt with depth.

The assimilation of the mid-depth velocity also yielded changes in the mass field through model dynamics, i.e. advection process. A zonal transect through trajectory EC (111.1°E, 13.8°N) is shown in Fig. 10. In Exp2, a cold core is distributed mainly between 110.5°E and 112°E, which is broader than the diameter of the trajectory (Fig. 6). In Exp3, the cold core diameter is narrowed (between 110.5°E and 111.5°E) and more consistent with the float trajectory. The largest change in the temperature fields ($\sim 0.2^{\circ}\text{C}$) due to velocity correction is approximately 600 m \sim 1000 m (Fig. 10e), although the float parking depth is approximately 1200 m. The contribution of velocity corrections to the salinity field is on the order of 0.005 PSU (Fig. 10f).

Although Argo floats are important for the subsurface states, SST and SLA are more important at surface. The assimilation of Argo T/S profile and mid-depth introduced some errors in the velocity field at the region near floats (Fig. 11 and Fig. 12), but the main pattern is similar to Exp2 and observations. It should be mentioned that the movement of No.5903457 was contrast to the surface velocity (Fig. 11a) field, indicating that the eddy in 1200m cannot be detected from surface. Differ to No.5903457 float, the trajectory of No.5903163 float drifting from northeast to southeast at surface (Fig. 11) and 1200m (Fig. 7). An anticyclonic eddy appeared near

No. 5902163 float in Jan and disappeared in Feb at surface (Fig. 11a and Fig. 12a). All assimilation experiments captured the pattern at surface due to the assimilation of satellite data. However, influenced by the topography, the anticyclonic eddy near No. 5902163 float cannot extend to 1200 m (Fig. 7 and Fig. 8). The slow movement ($\sim 3\text{cm/s}$) of float indicated that this float was also influenced the mid-depth velocity field.

4.3 System compatibility

When adding a new dataset, it is important that the forecast quality of the SST, SLA, T profile and S profile not decrease. The compatibility of the trajectory with other observations in the 4DVAR system was checked through statistical analysis. The RMSEs between the model field and observations are defined as follows:

$$\text{RMSE}(x) = \sqrt{\sum_{i=1}^N (x_i^{\text{model}} - x_i^{\text{observation}})^2 / N} \quad (3)$$

where x denotes the SST, SLA, T (temperature profile) or S (salinity profile), and N is the total number of x for 2 months. In terms of SST and SLA, only water depths greater than 200 m are counted. The 2-month RMSEs in whole region and western SCS in our concern ($110\text{--}114^\circ\text{E}$, $13\text{--}16^\circ\text{N}$) are shown in Table 3. The assimilation experiments show better performances overall than the Ctrl experiment. Exp1 performed best in SST and SLA but increased the error in salinity forecast. While *in situ* T/S profile data are assimilated, the error in salinity forecast can be reduced but the RMSEs in SST and SLA increased slightly at the same time. When the additional Argo trajectory-derived mid-depth velocities are introduced, the forecast level of other variables is maintained at the previous levels as in Exp2 except slight improvement in

SLA. This demonstrates that the method we used can effectively assimilate the Argo trajectory into the model without introducing other problems.

The biases of T/S are shown in Fig. 13 (whole model region) and Fig. 14 (western SCS). The assimilation of mid-depth velocities increased the T/S biases in upper 300m slightly but reduced the T/S biases below 300m. The largest biases and RMSEs occurred about 150m due to the drawback of model.

Generally, the assimilation of satellite and T/S profile data are not enough to yield accurate modeling of the mid-depth velocity field in the western SCS. Except for balanced information, assimilation of the Argo trajectory can aid in providing additional unbalanced information, which further improves the forecast quality of the velocity fields at the float parking depth.

5 Summary and discussion

Argo trajectory data are useful in studying basin scale current structures and detecting subsurface mesoscale eddy in the middle layer of the ocean but are rarely used in ocean data assimilation. In this paper, the Argo trajectory data are assimilated into a regional model of the SCS. Instead of directly assimilating the trajectory data, which requires a complicated trajectory prediction model, the mid-depth velocities are first retrieved from the float trajectories and then assimilated into the ocean model. As pointed out by Molcard et al. (2005), pseudo-Lagrangian method is efficient provided that position sampling period is smaller than Lagrangian time scale. In SCS, the sampling period of Argo position is 4 days, while the Lagrangian time scale is about 7-

15 days in ocean interior. Thus, this method can be used in SCS to assimilate Argo position into model. Before assimilated into the model, the trajectories were subjected to two rounds of quality control, and the observation error of the mid-depth velocities was estimated from the distribution of velocity misfits between the Argo and HYCOM-NCODA product and finally set to 2 cm/s.

Four experiments were designed to evaluate the influence of the Argo trajectory assimilation on the mid-depth velocity field. A comparison of 2-months of results from four experiments showed that the assimilation of mid-depth velocities is useful in simulation of the progresses which are not detected from surface, i.e., anti-cyclonic eddy east off Vietnam at ~1200m. With the introducing of the mid-depth velocities, the Eulerian velocity are improved for ~11% (zonal) and ~72% (meridional) compared with traditional data assimilation experiment (Exp2); the radius and EC misfit of eddy reduced to 61km and 0.9km, respectively. Along with the radius decreased, a cold core diameter is narrowed. Except the Eulerian aspect, the Lagrangian velocity is also improved, the predicted trajectory matched well with the realistic trajectory, with an improvement of 7% compared with Exp2. A statistical analysis of the whole region and western SCS showed that this pseudo-Lagrangian assimilation method can reduce the T/S biases and RMSEs below 300 m, and maintain the forecast level of other observations without introducing other problems, such as spurious values in the model states.

The influence of mid-depth velocity can be insight from the balance relationship.

Although the main component of velocity at 1200 m is geostrophic current, accurately modeling the complicated velocity field in the SCS subsurface cannot be adequately accomplished by only adjusting balanced components through satellite and *in situ* T/S profile data assimilation. Mid-depth velocity data assimilation can adjust the unbalanced component of velocity field, which influences the properties of the mesoscale process in the subsurface, such as position and strength. Velocity field corrections are not limited to the observation position at the float parking depth but also propagate horizontally and vertically. The vertical influence is weakened with the shallow of depth. As a result, the eddy east off Vietnam tilted southwest with the increase of depth. A comparison of Exp2 and Exp3 showed that the velocity field corrections also impact the mass fields. However, this process is achieved through model dynamics but not balance relationship, since temperature is the base variable in ROMS IS4DVAR. This probably one reason for the slightly increase of T/S biases and RMSEs in upper 300 m. The above conclusion is mainly drawn from the No. 5903457 float and No.5902163 in the western SCS near Vietnam but can also be applied to other regions with Argo float trajectories in the SCS (not shown). This finding will be useful for more accurately describing the mid-depth velocity in the SCS and deployment of Argo floats that short cycle period (smaller than Lagrangian time) should be considered for better effect.

Some shortcomings remain in this paper. One is our model does not include tides, which are important for progress in deeper than 2400m in SCS (Wang et al., 2016;

2017). The major problem of including tides in model is the changing for assimilating SLA due to its high frequency variability. Add tides to the model will be considered in future work. The data assimilation system made some simple assumptions as in most ocean data assimilation, including the correlation matrix is isotropic and uniform and observation errors between u and v is uncorrelated, which are not true. The horizontal and vertical correlation length scales are uniform (50 km horizontally and 30 m vertically), which is one reason why the assimilation of satellite data imposed a larger radius in Exp1 than in the other experiments. The assimilation window is 4 day in the assimilation experiments following Moore et al. (2011b), which may not be the optimal assimilation window in the SCS. These drawbacks will be addressed further in future studies.

Acknowledgements

This study benefitted from the AVISO SLA data (<ftp://ftp-access.aviso.altimetry.fr>), AVHRR SST data (<ftp://eclipse.ncdc.noaa.gov>), EN4.2.1 *in situ* data (<http://hadobs.metoffice.com/en4/download.html>), Argo trajectory data from Coriolis (<ftp://ftp.ifremer.fr/ifremer/argo>), ERA-interim dataset from ECMWF (<https://apps.ecmwf.int/datasets/data/interim-full-daily>), CCMP wind data (<ftp://ftp2.remss.com>), SODA3.3.1 data (<https://www.atmos.umd.edu/~ocean/>) and HYCOM-NCODA product (http://tds.hycom.org/thredds/dodsC/GLBu0.08/exp_90.9). This study was jointly supported by the National Key R&D Program of China

511 (2018YFC1406202) and National Natural Science Foundation of China (41675097,
512 41830964).
513

514 **Reference**

- 515 Atlas, R., R. N. Hoffman, J. Ardizzone, S. M. Leidner, J. C. Jusem, D. K. Smith, and
 516 D. Gombos, 2011: A cross-calibrated, multiplatform ocean surface wind velocity
 517 product for meteorological and oceanographic applications. *B. Am. Meteorol. S.*,
 518 **92**, 157-174, <https://doi.org/10.1175/2010BAMS2946.1>.
- 519 Carton, J. A., G. A. Chepurin, and L. Chen, 2018: SODA3: a new ocean climate
 520 reanalysis. *J. Climate*, **31**, 6967-6983, <https://doi.org/10.1175/JCLI-D-18-0149.1>.
- 521 Castellari, S., A. Griffa, T. M. Özgökmen, and P. M. Poulain, 2001: Prediction of
 522 particle trajectories in the adriatic sea using lagrangian data assimilation. *J. Marine*
 523 *Syst.*, **29**, 33-50, [https://doi.org/10.1016/S0924-7963\(01\)00008-2](https://doi.org/10.1016/S0924-7963(01)00008-2).
- 524 Chao, S. Y., P. T. Shaw, and S. Y. Wu, 1996: Deep water ventilation in the South China
 525 Sea. *Deep-Sea Res. Pt. I*, **43**, 445-466, [https://doi.org/10.1016/0967-](https://doi.org/10.1016/0967-0637(96)00025-8)
 526 [0637\(96\)00025-8](https://doi.org/10.1016/0967-0637(96)00025-8).
- 527 Chen, G., Y. Hou, and X. Chu, 2011: Mesoscale eddies in the South China Sea: Mean
 528 properties, spatiotemporal variability, and impact on thermohaline structure. *J.*
 529 *Geophys. Res. Ocean.*, **116**, <https://doi.org/10.1029/2010JC006716>.
- 530 Chen, G., D. Wang, C. Dong, T. Zu, H. Xue, Y. Shu, X. Chu, Y. Qi, and H. Chen, 2015:
 531 Observed deep energetic eddies by seamount wake. *Sci. Rep.*, **5**, 17416,
 532 <https://doi.org/10.1038/srep17416>.
- 533 Chu, P. C., and R. Li, 2000: South China Sea isopycnal-surface circulation. *J. Phys.*

534 *Oceanogr.*, **30**, 2419-2438, <https://doi.org/10.1175/1520->
535 0485(2000)030<2419:SCSISC>2.0.CO;2.

536 Courtier, P., J. N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational
537 implementation of 4D-Var, using an incremental approach. *Q. J. Roy. Meteor. Soc.*,
538 **120**, 1367-1387, <https://doi.org/10.1002/qj.49712051912>.

539 Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U.
540 Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars,
541 L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J.
542 Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P.
543 Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J.
544 Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavalato, J.-N. Thépaut, and
545 F. Vitart, 2011: The ERA-Interim reanalysis: Configuration and performance of
546 the data assimilation system. *Q. J. Roy. Meteor. Soc.*, **137**, 553-597,
547 <https://doi.org/10.1002/qj.828>.

548 Fang, G., G. Wang, Y. Fang, and W. Fang, 2012: A review on the South China Sea
549 western boundary current. *Acta Oceanol. Sin.*, **31**, 1-10,
550 <https://doi.org/10.1007/s13131-012-0231-y>.

551 Good, S. A., M. J. Martin, and N. A. Rayner, 2013: EN4: Quality controlled ocean
552 temperature and salinity profiles and monthly objective analyses with uncertainty
553 estimates. *J. Geophys. Res. Ocean.*, **118**, 6704-6716,
554 <https://doi.org/10.1002/2013JC009067>.

555 Hu, J., H. Kawamura, H. Hong, and Y. Qi, 2000: A review on the currents in the South
 556 China Sea: seasonal circulation, South China Sea warm current and Kuroshio
 557 intrusion. *J. Oceanogr.*, **56**, 607-624, <https://doi.org/10.1023/A:1011117531252>.
 558 Kuznetsov, L. V., K. Ide, and C. K. Jones, 2003: A Method for Assimilation of
 559 Lagrangian Data. *Mon. Wea. Rev.*, **131**, 2247-2260, [https://doi.org/10.1175/1520-0493\(2003\)131<2247:AMFAOL>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<2247:AMFAOL>2.0.CO;2).
 560
 561 Lebedev, K. V., H. Yoshinari, N. A. Maximenko, and P. W. Hacker, 2007: Velocity
 562 data assessed from trajectories of Argo floats at parking level and at the sea surface.
 563 *IPRC Tech. Note*, **4**, <https://doi.org/10.13140/RG.2.2.12820.71041>.
 564 Li, J., R. Zhang, and B. Jin, 2011: Eddy characteristics in the northern South China Sea
 565 as inferred from Lagrangian drifter data. *Ocean Sci.*, **7**, 661,
 566 <https://doi.org/10.5194/os-7-661-2011>.
 567 Liao, G., Y. Yuan, and X. Xu, 2005: The three dimensional structure of the circulation
 568 in the South China Sea during the winter of 1998. *Acta Oceanol. Sin.*, **27**, 8-17.
 569 (In Chinese).
 570 Lin, X., C. Dong, D. Chen, Y. Liu, J. Yang, B. Zou, and Y. Guan, 2015: Three-
 571 dimensional properties of mesoscale eddies in the South China Sea based on eddy-
 572 resolving model output. *Deep-Sea Res. Pt. I*, 2015, **99**, 46-64,
 573 <https://doi.org/10.1016/j.dsr.2015.01.007>.
 574 Liu, Q., A. Kaneko, and J. Su, 2008: Recent progress in studies of the South China Sea
 575 circulation. *J. Oceanogr.*, **64**, 753-762, <https://doi.org/10.1007/s10872-008-0063->

576 8.

577 Markova, N. V., and A. V. Bagaev, 2016: The Black Sea Deep Current Velocities

578 Estimated from the Data of Argo Profiling Floats. *Phys. Oceanogr.*, **3**,

579 <https://doi.org/10.22449/1573-160X-2016-3-23-35>.

580 Molcard, A., A. Griffa, Özgökmen, and M. Tamay, 2005: Lagrangian Data

581 Assimilation in Multilayer Primitive Equation Ocean Models. *J. Atmos. Oceanic*

582 *Technol.*, **22**, 70-83. <https://doi.org/10.1175/jtech-1686.1>.

583 Molcard, A., L. I. Piterbarg, A. Griffa, T. M. Özgökmen, and A. J. Mariano, 2003:

584 Assimilation of drifter observations for the reconstruction of the Eulerian

585 circulation field, *J. Geophys. Res.*, **108**, 3056,

586 <https://doi.org/10.1029/2001JC001240>, C3.

587 Moore, A. M., H. G. Arango, G. Broquet, B. S. Powell, A. T. Weaver, and J. Zavala-

588 Garay, 2011a: The Regional Ocean Modeling System (ROMS) 4-dimensional

589 variational data assimilation systems: Part I–System overview and formulation.

590 *Prog. Oceanogr.*, **91**, 34-49, <https://doi.org/10.1016/j.pocean.2011.05.004>.

591 Moore, A. M., H. G. Arango, G. Broquet, C. Edwards, M. Veneziani, B. Powell, D.

592 Foley, J. D. Doyle, D. Costa and P. Robinson, 2011b: The Regional Ocean

593 Modeling System (ROMS) 4-dimensional variational data assimilation systems:

594 part II–performance and application to the California Current System. *Prog.*

595 *Oceanogr.*, **91**, 50-73, <https://doi.org/10.1016/j.pocean.2011.05.003>.

596 Muscarella, P., M. Carrier, H. Ngodock, S. Smith, B. L. Lipphardt, A. D. Kirwan, and

597 H. S. Huntley, 2015: Do Assimilated Drifter Velocities Improve Lagrangian
 598 Predictability in an Operational Ocean Model. *Mon. Wea. Rev.*, **143**, 1822-1832,
 599 <https://doi.org/10.1175/MWR-D-14-00164.1>.

600 Nilsson, J. A. U., S. Dobricic, N. Pinardi, P.-M. Poulain, and D. Pettenuzzo, 2012:
 601 Variational assimilation of Lagrangian trajectories in the Mediterranean ocean
 602 Forecasting System. *Ocean Sci*, **8**, 249-259, [https://doi.org/10.5194/os-8-249-](https://doi.org/10.5194/os-8-249-2012)
 603 2012.

604 Nilsson, J. A. U., S. Dobricic, N. Pinardi, V. Taillandier, and P.-M. Poulain, 2011: On
 605 the assessment of Argo float trajectory assimilation in the Mediterranean
 606 Forecasting System. *Ocean Dynam.*, **61**, 1475-1490,
 607 <https://doi.org/10.1007/s10236-011-0437-0>.

608 Ollitrault, M., and J. P. Rannou, 2013: ANDRO: An Argo-based deep displacement
 609 dataset. *J. Atmos. Oceanic Technol.*, **30**, 759-788, [https://doi.org/10.1175/JTECH-](https://doi.org/10.1175/JTECH-D-12-00073.1)
 610 D-12-00073.1

611 Ollitrault, M., and Colin de Verdière A., 2014: The ocean general circulation near 1000-
 612 m depth. *J. Phys. Oceanogr.*, **44**: 384-409, [https://doi.org/10.1175/JPO-D-13-](https://doi.org/10.1175/JPO-D-13-030.1)
 613 030.1.

614 Park, J. J., K. Kim, and W. R. Crawford, 2004: Inertial currents estimated from surface
 615 trajectories of ARGO floats. *Geophys. Res. Lett.*, **31**,
 616 <https://doi.org/10.1029/2004GL020191>.

617 Park, J. J., K. Kim, B. A. King, and S. C. Riser, 2005: An advanced method to estimate

618 deep currents from profiling floats. *J. Atmos. Oceanic Technol.*, **22**: 1294-1304,
619 <https://doi.org/10.1175/JTECH1748.1>.

620 Park, J. J. and K. Kim, 2013: Deep currents obtained from Argo float trajectories in the
621 Japan/East Sea. *Deep-Sea Res. Pt II*, 2013, **85**, 169-181,
622 <https://doi.org/10.1016/j.dsr2.2012.07.032>.

623 Quan, Q., H. Xue, H. Qin, X. Zeng, S. Peng, 2016: Features and variability of the South
624 China Sea western boundary current from 1992 to 2011. *Ocean Dynam.*, **66**, 795-
625 810, <https://doi.org/10.1007/s10236-016-0951-1>.

626 Slivinski, L., E. T. Spiller, A. Apte, and B. Sandstede, 2015: A Hybrid Particle–
627 Ensemble Kalman Filter for Lagrangian Data Assimilation. *Mon. Wea. Rev.*, **143**,
628 195-211, <https://doi.org/10.1175/MWR-D-14-00051.1>.

629 Song B., H. Wang, C. Chen, R. Zhang, and S. Bao, 2019: Observed subsurface eddies
630 near the Vietnam coast of the South China Sea. *Acta Oceanol. Sin.*, **38**, 39 - 46,
631 <https://doi.org/10.1007/s13131-019-1412-8>.

632 Spiller, E., A. Budhiraja, K. Ide, and C. Jones, 2008: Modified particlefilter methods
633 for assimilating Lagrangian data into a point-vortexmodel. *Physica D*, **237**, 1498-
634 1506, <https://doi.org/10.1016/j.physd.2008.03.023>.

635 Reynolds, R. W., T. M. Smith, C. Liu, C., D. B. Chelton, K. S. Casey, and M. G. Schlax,
636 2007: Daily high-resolution-blended analyses for sea surface temperature. *J.*
637 *Climate*, **20**, 5473-5496, <https://doi.org/10.1175/2007JCLI1824.1>.

638 Taillandier, V., A. Griffa, P. M. Poulain, and K. Béranger, 2006a: Assimilation of Argo

float positions in the north western Mediterranean Sea and impact on ocean
circulation simulations. *Geophys. Res. Lett.*, **33**,
<https://doi.org/10.1029/2005GL025552>.

Taillandier, V., and A. Griffa, 2006b: Implementation of position assimilation for
ARGO floats in a realistic Mediterranean Sea OPA model and twin experiment
testing. *Ocean Sci.*, **3**, 255-289, <https://doi.org/10.5194/os-2-223-2006>.

Voet, G., D. Quadfasel, K. A. Mork, and H. S  iland, 2010: The mid-depth circulation
of the Nordic Seas derived from profiling float observations. *Tellus A*, **62**, 516-
529, <https://doi.org/10.1111/j.1600-0870.2009.00444.x>.

Wang, D., Q. Liu, Q. Xie, Z. He, W. Zhuang, Y. Shu, X. Xiao, B. Hong, X. Wu, and D.
Sui, 2013: Progress of regional oceanography study associated with western
boundary current in the South China Sea. *Chin. Sci. Bull.*, **58**, 1205-1215,
<https://doi.org/10.1007/s11434-012-5663-4>.

Wang, G., J. Su, and P. C. Chu, 2003: Mesoscale eddies in the South China Sea
observed with altimeter data. *Geophys. Res. Lett.*, 2003, **30**,
<https://doi.org/10.1029/2003GL018532>.

Wang, H., G. Wang, D. Chen, and R. Zhang, 2012a: Reconstruction of Three-
Dimensional Pacific Temperature with Argo and Satellite Observations. *Atmos.*
Ocean, **50**, 116-128, <https://doi.org/10.1080/07055900.2012.742421>.

Wang, H., R. Zhang, Wang G, Y. An, and B. Jin, 2012b: Quality control of Argo
temperature and salinity observation profiles. *Chinese J. Geophys.*, **55**, 577-588,

660 <https://doi.org/10.6038/j.issn.0001-5733.2012.02.020>.(in Chinese)

661 Wang, X., S. Peng, Z. Liu, R. Huang, Y. Qian, and Y. Li, 2016: Tidal Mixing in the
662 South China Sea: An Estimate Based on the Internal Tide Energetics. *J. Phys.*
663 *Oceanogr.*, **46**, 107-124, <https://doi.org/10.1175/JPO-D-15-0082.1>.

664 Wang, X., Z. Liu, and S. Peng, 2017: Impact of Tidal Mixing on Water Mass
665 Transformation and Circulation in the South China Sea. *J. Phys. Oceanogr.*, **47**,
666 419-432, <https://doi.org/10.1175/JPO-D-16-0171.1>.

667 Wang, X., W. Zhang, P. Wang, J. Yang, and H. Wang, 2018: Research on mid-depth
668 current of basin scale in the South China Sea based on historical Argo observations.
669 *Haiyang Xuebao*, **40**, 1-14, <https://doi.org/10.3969/j.issn.0253-4193.2018.06.001>.
670 (in Chinese).

671 Warner, J. C., C. R. Sherwood, H. G. Arango, and R. P. Signell, 2005: Performance of
672 four turbulence closure models implemented using a generic length scale method.
673 *Ocean Model.*, **8**, 81-113, <https://doi.org/10.1016/j.ocemod.2003.12.003>.

674 Weaver, A. T., C. Deltel, Machu É, S. Ricci, and N. Daget, 2005: A multivariate balance
675 operator for variational ocean data assimilation. *Q.J.R. Meteorol. Soc.*, **131**, 3605-
676 3625, <https://doi.org/10.1256/qj.05.119>.

677 Weaver, A. T., J. Vialard, and D. L. T. Anderson, 2003: Three-and four-dimensional
678 variational assimilation with a general circulation model of the tropical Pacific
679 Ocean. Part I: Formulation, internal diagnostics, and consistency checks. *Mon.*
680 *Wea. Rev.*, **131**, 1360-1378, <https://doi.org/10.1175/1520->

0493(2003)131<1360:TAFVAW>2.0.CO;2.

Wyrski, K., 1961: Physical oceanography of the Southeast Asian waters. *UC San Diego: Scripps Institution of Oceanography*. Retrieved from <https://escholarship.org/uc/item/49n9x3t4>.

Xie, J., F. Counillon, J. Zhu, and L. Bertino, 2011: An eddy resolving tidal-driven model of the South China Sea assimilating along-track SLA data using the EnOI. *Ocean Sci.*, **7**, 609-627, <https://doi.org/10.5194/os-7-609-2011>.

Xie, J., and J. Zhu, 2008: Estimation of the surface and mid-depth currents from Argo floats in the Pacific and error analysis. *J. Marine Sys.*, **73**, 61-75, <https://doi.org/10.1016/j.jmarsys.2007.09.001>.

Xiu, P., F. Chai, L. Shi, H. Xue, and Y. Chao, 2010: A census of eddy activities in the South China Sea during 1993–2007. *J. Geophys. Res.*, 2010, **115**, <https://doi.org/10.1029/2009JC005657>.

Yuan, D., 2012: A numerical study of the South China Sea deep circulation and its relation to the Luzon Strait transport. *Acta Oceanol. Sin.*, **21**, 187-202.

Zhang, Z., F. Qiao, and J. Guo, 2014: Subsurface eddies in the southern South China Sea detected from in-situ observation in October 2011. *Deep-Sea Res. Pt. I*, **87**, 30-34, <https://doi.org/10.1016/j.dsr.2014.02.004>.

Zhang, Z., J. Tian, B. Qiu, W. Zhao, P. Chang, D. Wu, and X. Wan, 2016: Observed 3D structure, generation, and dissipation of oceanic mesoscale eddies in the South China Sea. *Sci. Rep.*, **6**, 24349, <https://doi.org/10.1038/srep24349>.

702 Zhou, H., D. Yuan, R. Li, and L. He, 2010: The western South China Sea currents from
703 measurements by Argo profiling floats during October to December 2007. *Chin.*
704 *J. Ocean. Limnol.*, **28**, 398-406, <https://doi.org/10.1007/s00343-010-9052-z>.

705 **Tables**

706 **Table 1.** Table 1 Experimental design

Exp. Name	SST	SLA	T/S	VEC
Ctrl	N	N	N	N
Exp1	Y	Y	N	N
Exp2	Y	Y	Y	N
Exp3	Y	Y	Y	Y

707 Note: The assimilated observations are marked with Y; otherwise, the observations
708 are marked with N. SST (satellite SST), SLA (sea level anomaly observation), T (*in*
709 *situ* temperature profile), S (*in situ* salinity profile), and VEC (velocity derived from
710 Argo trajectory).

711
712 **Table 2.** RMS misfits between model values and observations at 1200 m for EC,
713 u, v and position misfit. The CCs are displaced within the brackets

	Ctrl	Exp1	Exp2	Exp3
EC (km)	87.7	54.1	25.6	0.9
U (cm/s)	6.1 (-0.15)	6.1 (0.22)	3.9 (0.81)	3.2 (0.87)
V (cm/s)	7.2 (0.06)	10.9 (-0.51)	7.7 (0.58)	2.5 (0.94)
Distance (km)	25.5	31.2	18.2	16.4

714 Note: The eddy center from the model output (EC) is the location of the minimum
715 speed in the eddy. The EC (111.1°E, 13.8°N) of the trajectory observations is identified
716 by 8 nearly symmetrical points (3, 4, 5, 6, 8, 9, 10 and 11 in Fig. 3b) of the trajectory,
717 i.e., the position where the sum of the distance to these points is the shortest.

718

719 **Table 3.** Two-month averaged RMSEs between model fields.

	Ctrl		Exp1		Exp2		Exp3	
	Whole	WSCS	Whole	WSCS	Whole	WSCS	Whole	WSCS
SST (°C)	0.88	1.0	0.26	0.15	0.27	0.19	0.27	0.2
SLA (cm)	8.6	6.6	4.1	2.3	5.1	2.9	4.8	2.7
T(°C)	0.86	0.86	0.81	0.92	0.65	0.38	0.65	0.37
S(PSU)	0.23	0.06	0.32	0.11	0.17	0.03	0.17	0.04

720 Note: Whole denotes the whole model region (99°E~134°E, 1°N~30°N) where
721 deeper than 200 m; WSCS the denotes the western SCS near eastern coast of Vietnam
722 (110-113°E, 13-16°N).

723 **Figure Caption List**

724 **Fig. 1.** The process of continuous data assimilation in 4DVAR. $X(t)$ denotes the model
725 states at time t .

726 **Fig. 2.** Bathymetry (unit: m) and model domain ($99^{\circ}\text{E}\sim 134^{\circ}\text{E}$, $1^{\circ}\text{N}\sim 30^{\circ}\text{N}$). The black
727 contour represents the 200 m isobath. TW represents Taiwan Island.

728 **Fig. 3.** Distribution of mid-depth velocities (a) and No. 5903457 Argo float T/S profile
729 position (b) during 1 Jan-28 Feb, 2013. The range in (b) is marked by a black box in
730 (a); the number in (b) represents the observational sequence in the first 2 months of
731 2013.

732 **Fig. 4.** Distribution of velocity deviation between HYCOM-NCODA product and Argo
733 data at 1200 m during 2006~2016. (Blue is the distribution of zonal ocean current u ,
734 and green is the distribution of meridional ocean current v).

735 **Fig. 5.** RMSEs and CCs of u and v for different observation error standard deviations
736 in Jan and Feb, 2013. The red line represents the result of experiment with observation
737 error 2 cm/s, and the blue line represents the result of experiment with observation error
738 5 cm/s.

739 **Fig. 6.** Two-month averaged velocity field from (a) Ctrl, (b) Exp1, (c) Exp2, (d) Exp3
740 and (e) the difference between Exp3 and Exp2. The No. 5903457 float trajectory is
741 represented by the magenta line, the starting position of the float in January is
742 represented by the magenta square, the final position is represented by the triangle, and

the center is represented by the magenta pentagram. The EC in the model fields is represented by a black pentagram. The predicted No.5903457 float trajectories (released at realistic position) of 4 days are marked with blue line, the last positions of predicted trajectories are marked with blue dots.

Fig. 7. Month averaged velocity field at 1200m in Jan, 2013 for (a) Ctrl, (b) Exp1, (c) Exp2, (d) Exp3. The No.5903457 float and No.5902163 trajectory is represented by the magenta line, the starting position of the float in January is represented by the square, the final position is represented by the triangle, and the pentagram denotes the location of floats at 20130131.

Fig. 8. Month averaged velocity field at 1200m in Feb, 2013 for (a) Ctrl, (b) Exp1, (c) Exp2, (d) Exp3. The No.5903457 float and No.5902163 trajectory is represented by the magenta line, the starting position of the float in January is represented by the square, the final position is represented by the triangle, the pentagram denotes the location of floats at 20130131.

Fig. 9. Two-month average velocity field of Exp2 (left), Exp3 (middle) and the difference between them (right, Exp3-Exp2). Top is 500 m, middle is 750 m and bottom is 1000 m. Contours represent stream function (interval 2000, unit: m³/s). No. 5903457 float trajectory is represented by the magenta line. The predicted No.5903457 float trajectories (released at realistic position) of 4 days are marked with blue line, the last positions of predicted trajectories are marked with blue dots.

Fig. 10. Temperature (left) and salinity (right) zonal transects (along 13.8°N) of the

difference between Exp2 and Ctrl (top panel), the difference between Exp3 and Ctrl (middle panel) and the difference between Exp3 and Exp2 (bottom panel).

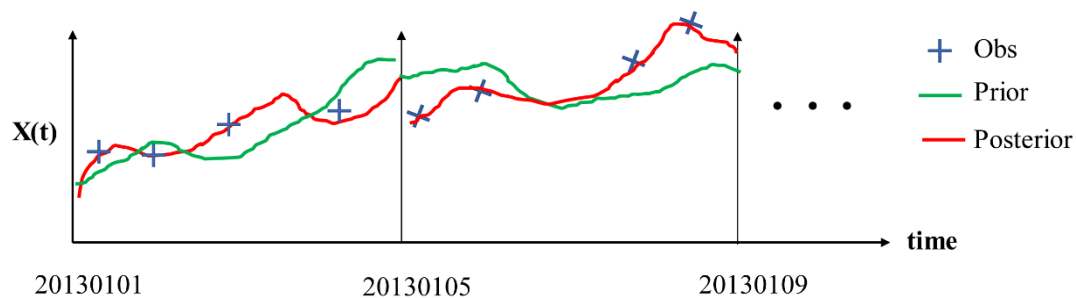
Fig. 11. The SST in 20130115 for a) AVHRR; b) Exp1; c) Exp2; d) Exp3. The vectors are surface geostrophic currents. The surface geostrophic currents in a) are from AVISO, others are from model results. The No.5903457 float and No.5902163 trajectory is represented by the magenta line, the starting position of the floats in January is represented by the square, the final position is represented by the triangle, and the pentagram denotes the location of floats at 20130115.

Fig. 12. The SST in 20130214 for a) AVHRR; b) Exp1; c) Exp2; d) Exp3. The vectors are surface geostrophic currents. The surface geostrophic currents in a) are from AVISO, others are from model results. The No.5903457 float and No.5902163 trajectory is represented by the magenta line, the starting position of the floats in January is represented by the square, the final position is represented by the triangle, and the pentagram denotes the location of floats at 20130214.

Fig. 13. Mean bias (Top panel) and RMSE (Bottom panel) of temperature and salinity over 2 months in the whole model region ($99^{\circ}\text{E}\sim 134^{\circ}\text{E}$, $1^{\circ}\text{N}\sim 30^{\circ}\text{N}$) where deeper than 200 m.

Fig. 14. Mean bias and RMSE of temperature and salinity over 2 months in the western SCS ($110\text{-}114^{\circ}\text{E}$, $13\text{ - }16^{\circ}\text{N}$).

785 **Figure**

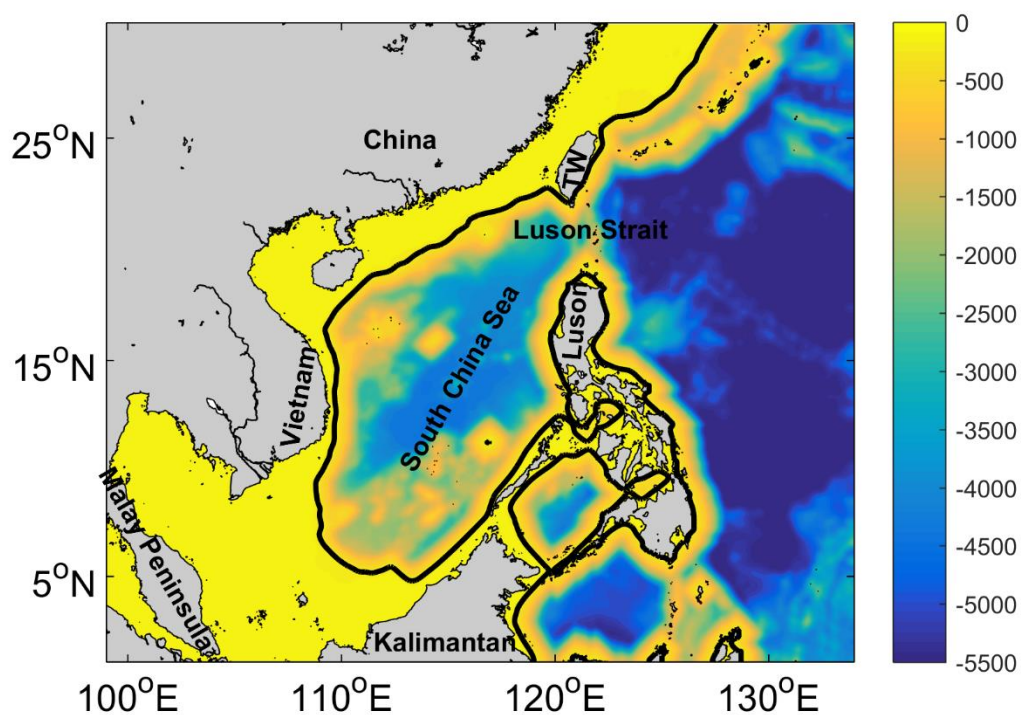


786

787 Fig. 1 The process of continuous data assimilation in 4DVAR. $X(t)$ denotes the model

788

states at time t .



789

790 **Fig. 2.** Bathymetry (unit: m) and model domain ($99^{\circ}\text{E}\sim 134^{\circ}\text{E}$, $1^{\circ}\text{N}\sim 30^{\circ}\text{N}$). The black
791 contour represents the 200 m isobath. TW represents Taiwan Island.

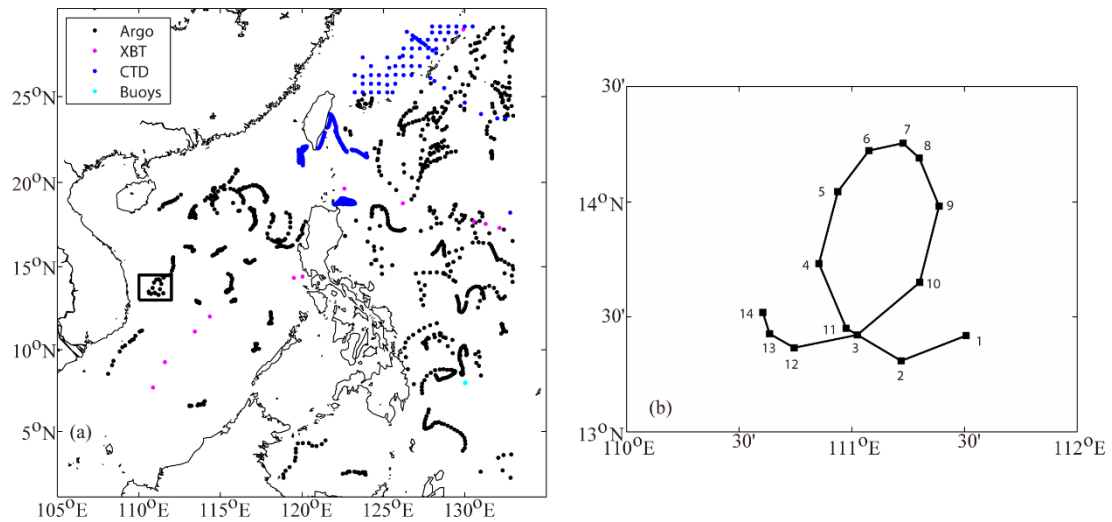


Fig. 3. Distribution of in-situ observations (a) and No. 5903457 Argo float T/S profile position (b) during 1 Jan-28 Feb, 2013. The range in (b) is marked by a black box in (a); the number in (b) represents the observational sequence in the first 2 months of 2013.

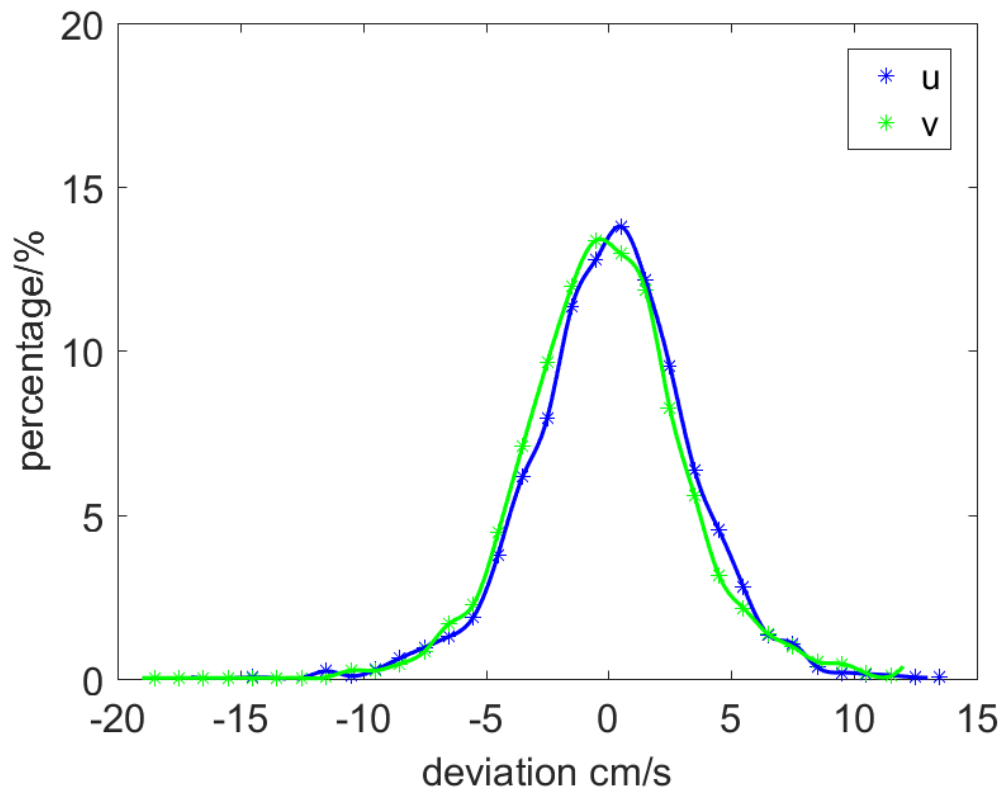
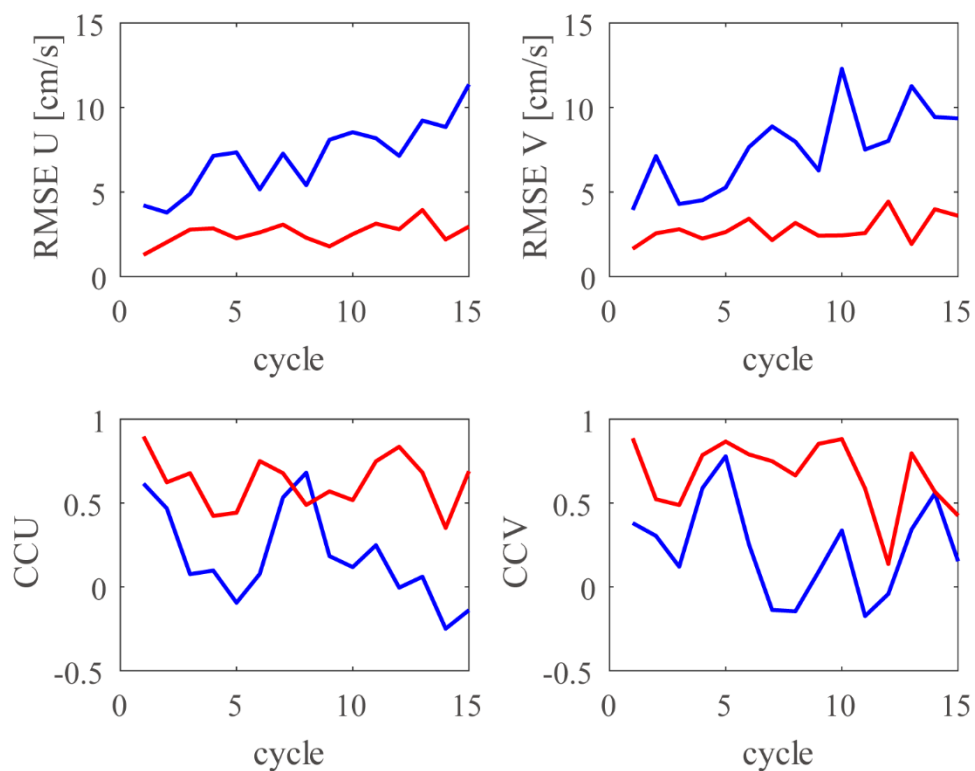


Fig. 4. Distribution of velocity deviation between HYCOM-NCODA product and Argo data at 1200 m during 2006~2016. (Blue is the distribution of zonal ocean current u , and green is the distribution of meridional ocean current v).

802



803

804 **Fig. 5.** RMSEs and CCs of u and v for different observation error standard deviations
805 in Jan and Feb, 2013. The red line represents the result of experiment with observation
806 error 2 cm/s, and the blue line represents the result of experiment with observation error
807 5 cm/s.

808

809

810

811

812

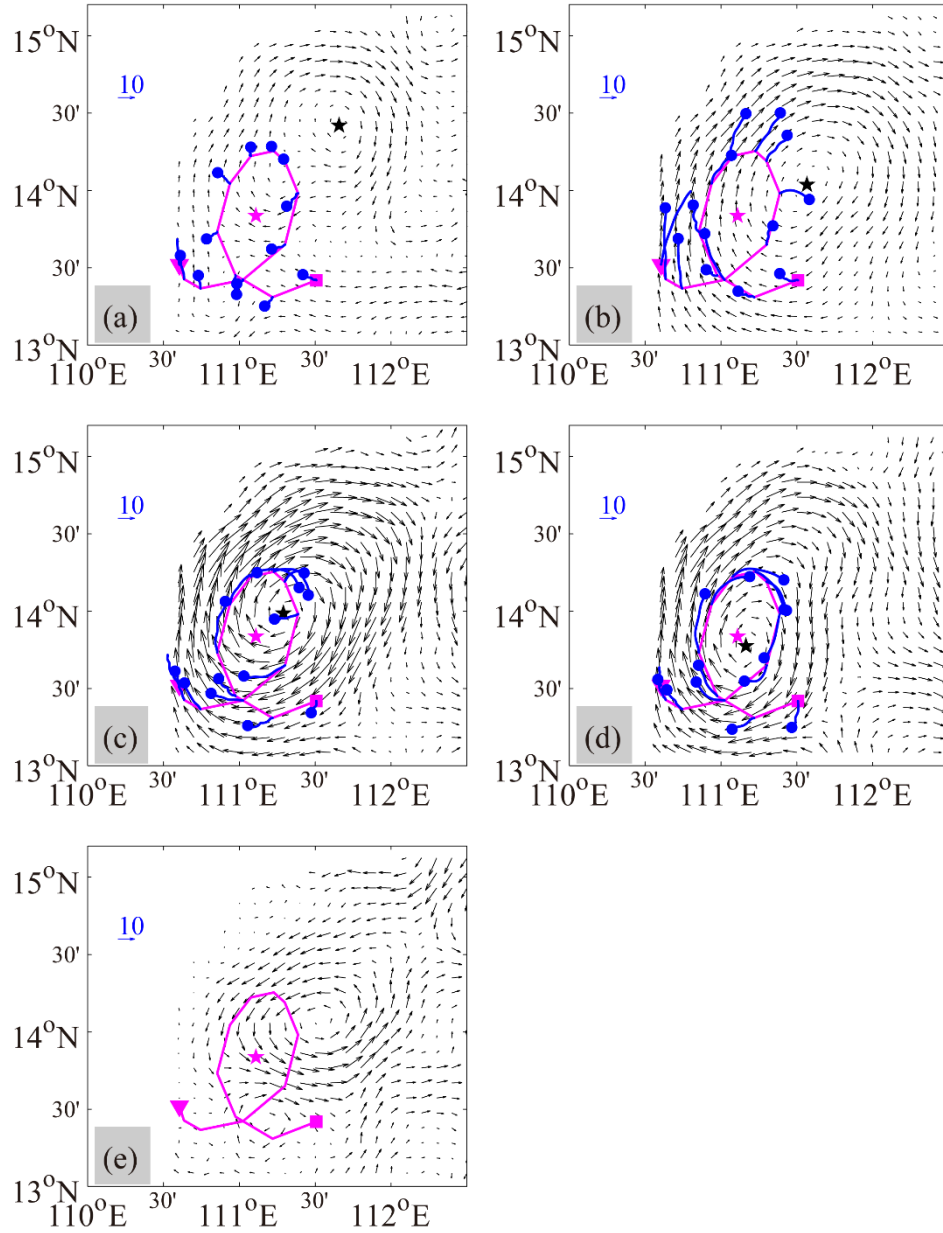


Fig. 6. Two-month averaged velocity field from (a) Ctrl, (b) Exp1, (c) Exp2, (d) Exp3 and (e) the difference between Exp3 and Exp2. The No. 5903457 float trajectory is represented by the magenta line, the starting position of the float in January is represented by the magenta square, the final position is represented by the triangle, and the center is represented by the magenta pentagram. The EC in the model fields is represented by a black pentagram. The predicted No.5903457 float trajectories (released at realistic position) of 4 days are marked with blue line, the last positions of predicted trajectories are marked with blue dots.

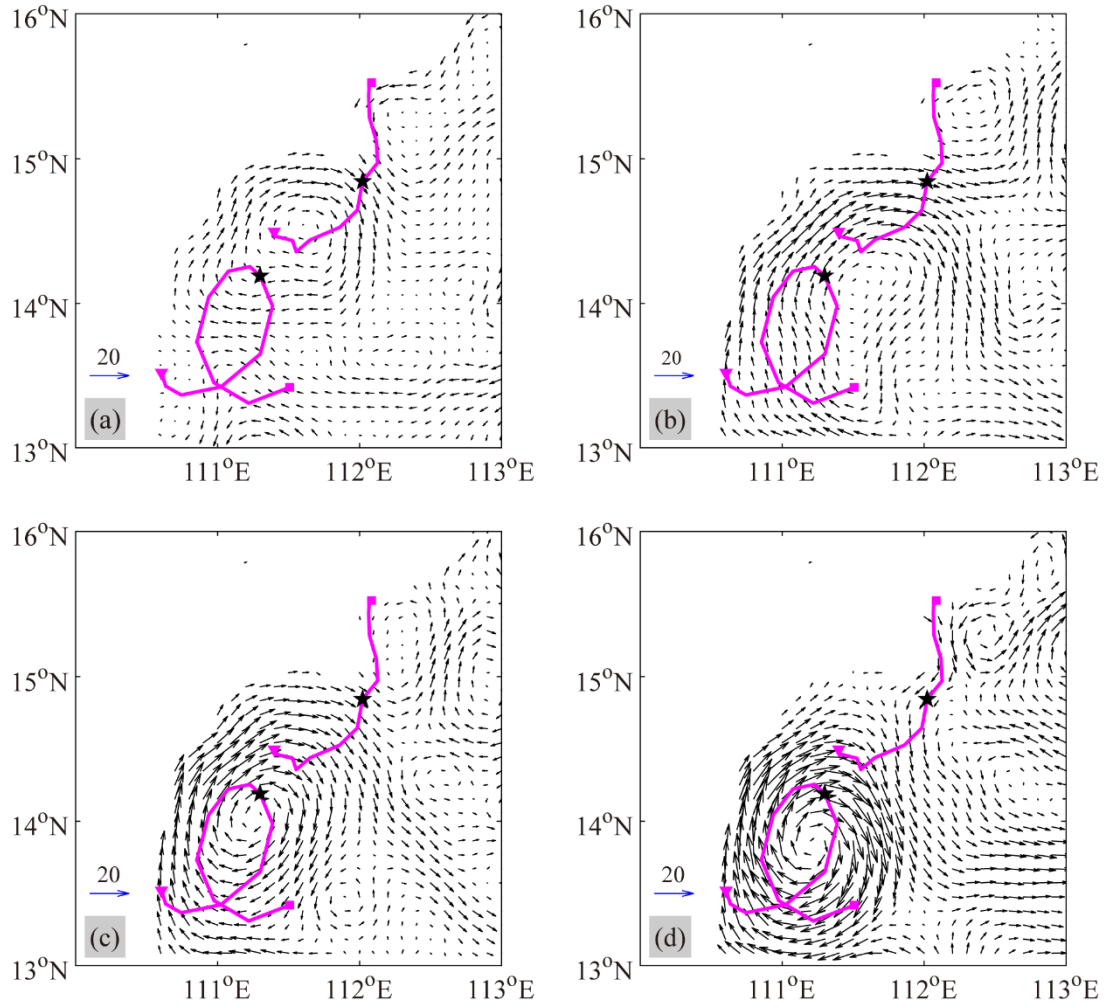


Fig. 7. Month averaged velocity field at 1200m in Jan, 2013 for (a) Ctrl, (b) Exp1, (c) Exp2, (d) Exp3. The No.5903457 float and No.5902163 trajectory is represented by the magenta line, the starting position of the float in January is represented by the square, the final position is represented by the triangle; the pentagram denotes the location of floats at 20130131.

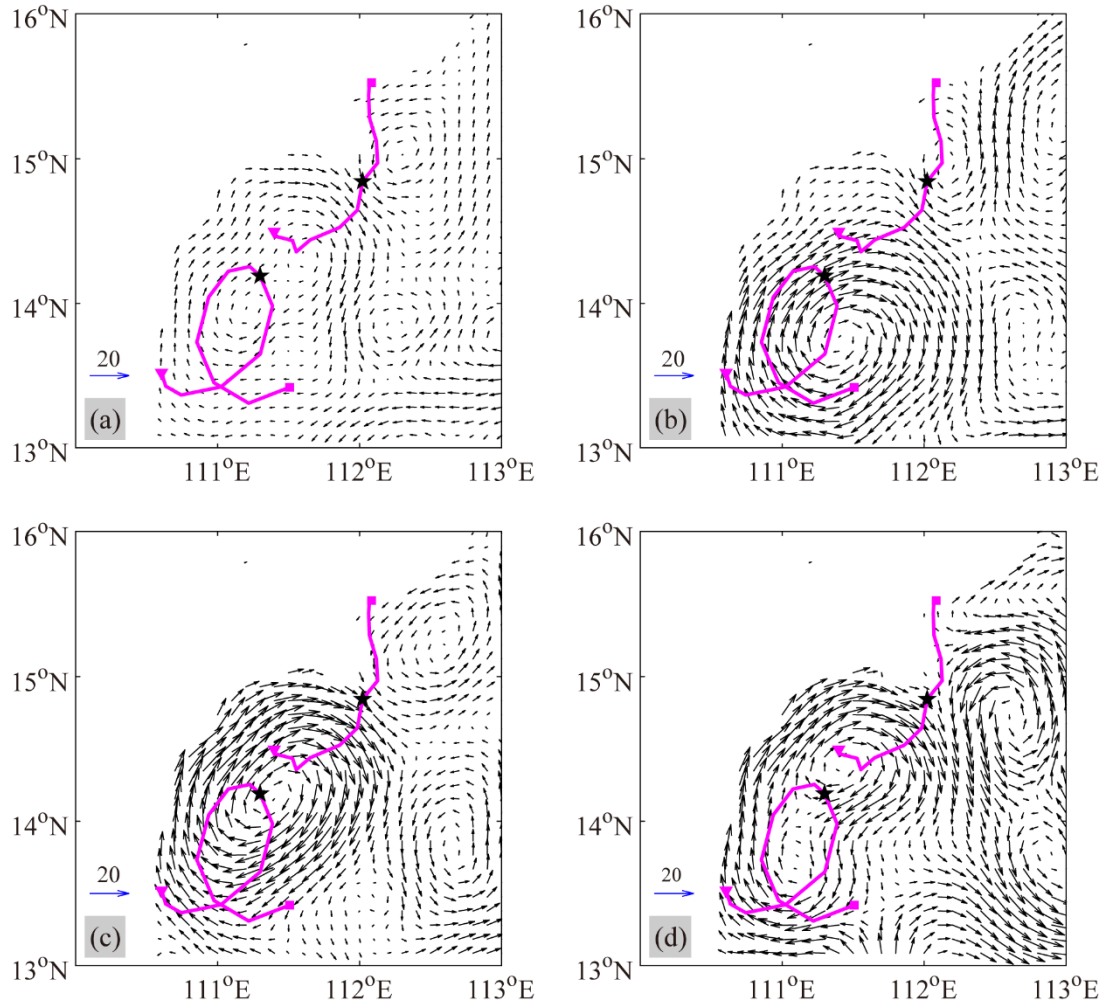


Fig. 8. Month averaged velocity field at 1200m in Feb, 2013 for (a) Ctrl, (b) Exp1, (c) Exp2, (d) Exp3. The No.5903457 float and No.5902163 trajectory is represented by the magenta line, the starting position of the float in January is represented by the square, the final position is represented by the triangle, the pentagram denotes the location of floats at 20130131.

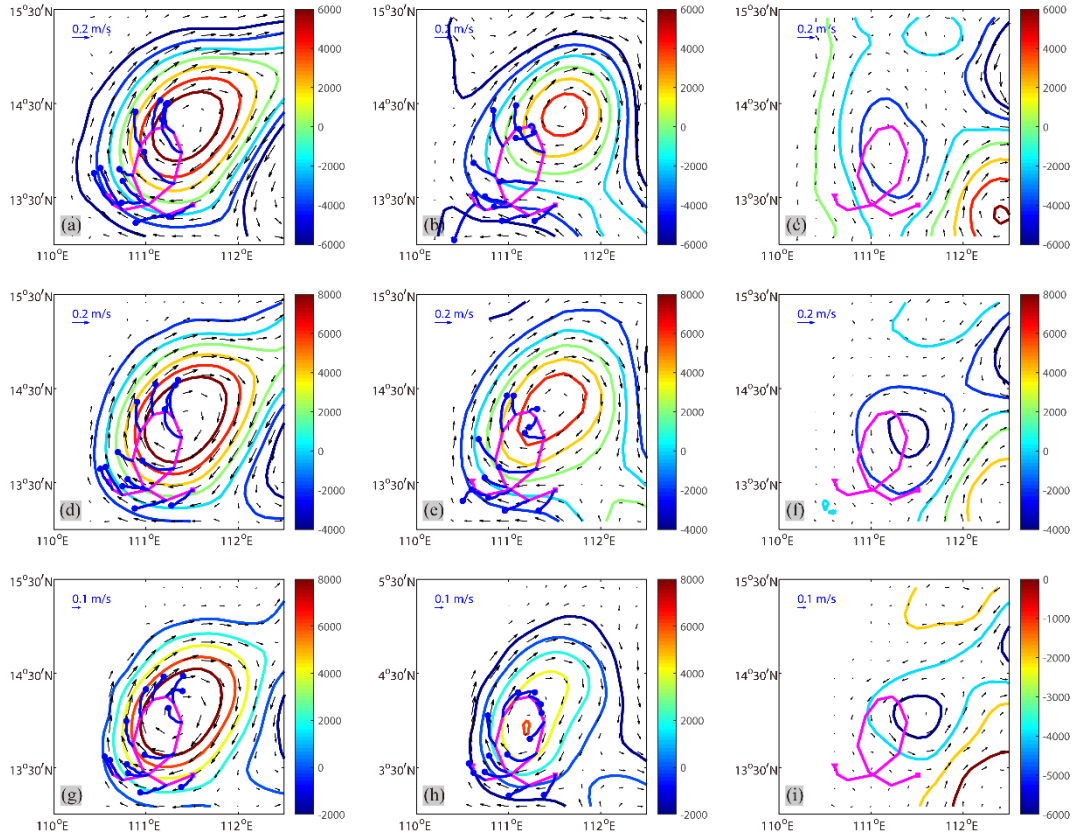


Fig. 9. Two-month average velocity field of Exp2 (left), Exp3 (middle) and the difference between them (right, Exp3-Exp2). Top is 500 m, middle is 750 m and bottom is 1000 m. Contours represent stream function (interval 2000, unit: m^3/s). No. 5903457 float trajectory is represented by the magenta line. The predicted No.5903457 float trajectories (released at realistic position) of 4 days are marked with blue line, the last positions of predicted trajectories are marked with blue dots.

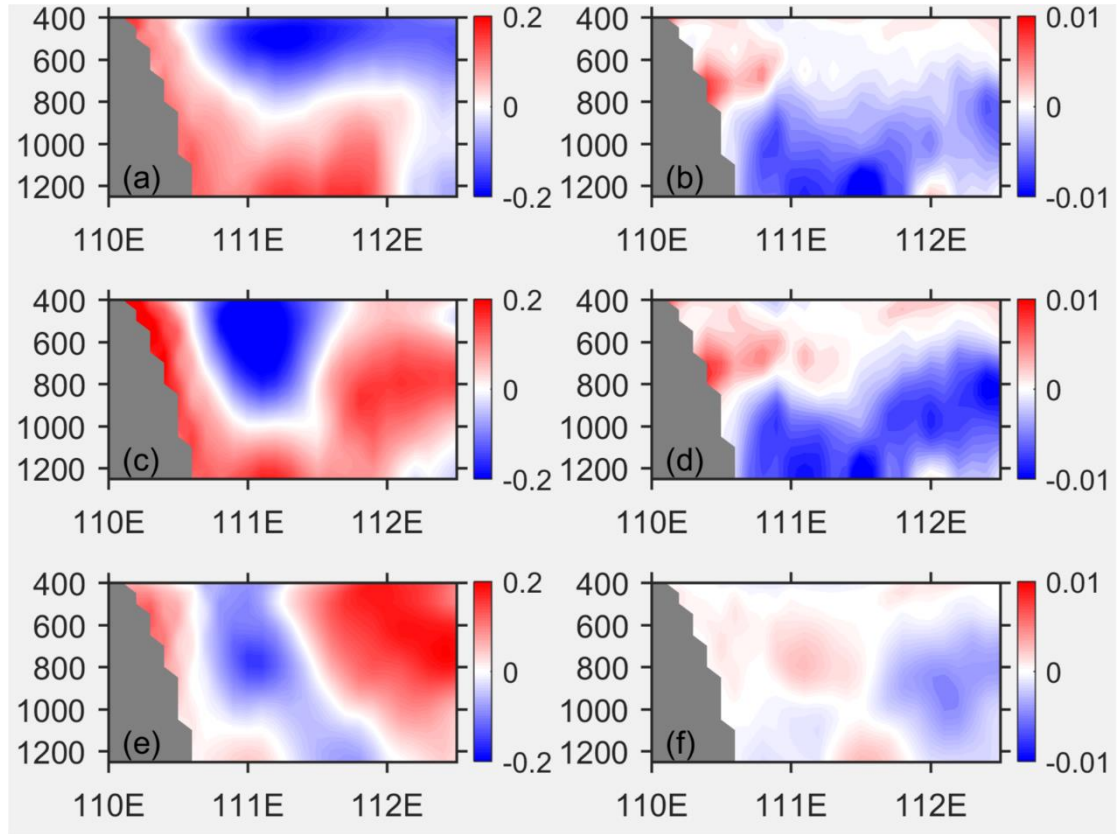


Fig. 10. Temperature (left) and salinity (right) zonal transects (along 13.8°N) of the difference between Exp2 and Ctrl (top panel), the difference between Exp3 and Ctrl (middle panel) and the difference between Exp3 and Exp2 (bottom panel).

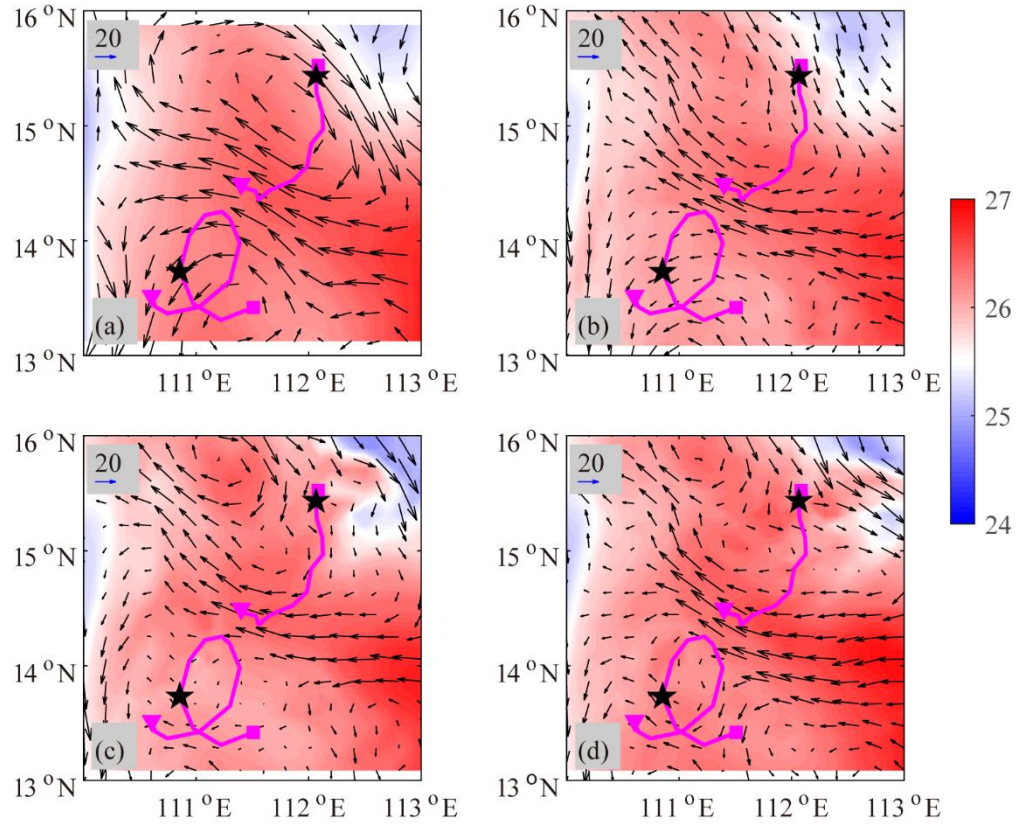


Fig. 11. The SST in 20130115 for a) AVHRR; b) Exp1; c) Exp2; d) Exp3. The vectors are surface geostrophic currents. The surface geostrophic currents in a) are from AVISO, others are from model results. The No.5903457 float and No.5902163 trajectory is represented by the magenta line, the starting position of the floats in January is represented by the square, the final position is represented by the triangle, and the pentagram denotes the location of floats at 20130115.

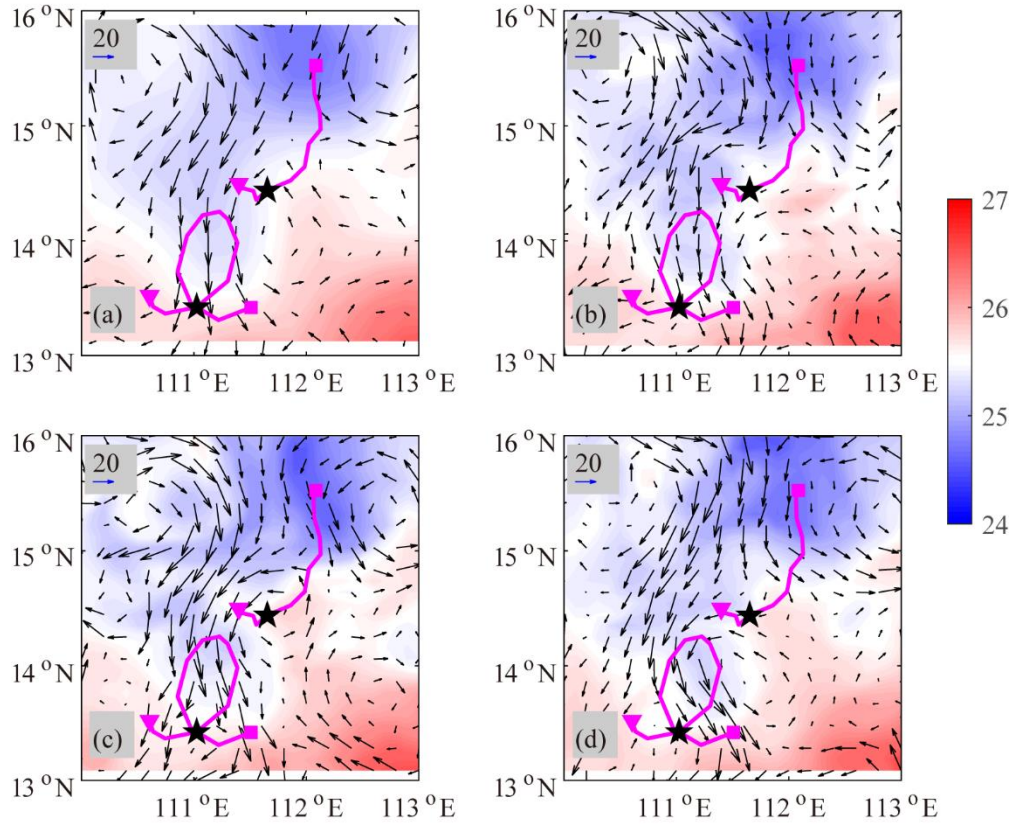


Fig. 12. The SST in 20130214 for a) AVHRR; b) Exp1; c) Exp2; d) Exp3. The vectors are surface geostrophic currents. The surface geostrophic currents in a) are from AVISO, others are from model results. The No.5903457 float and No.5902163 trajectory is represented by the magenta line, the starting position of the floats in January is represented by the square, the final position is represented by the triangle, and the pentagram denotes the location of floats at 20130214.

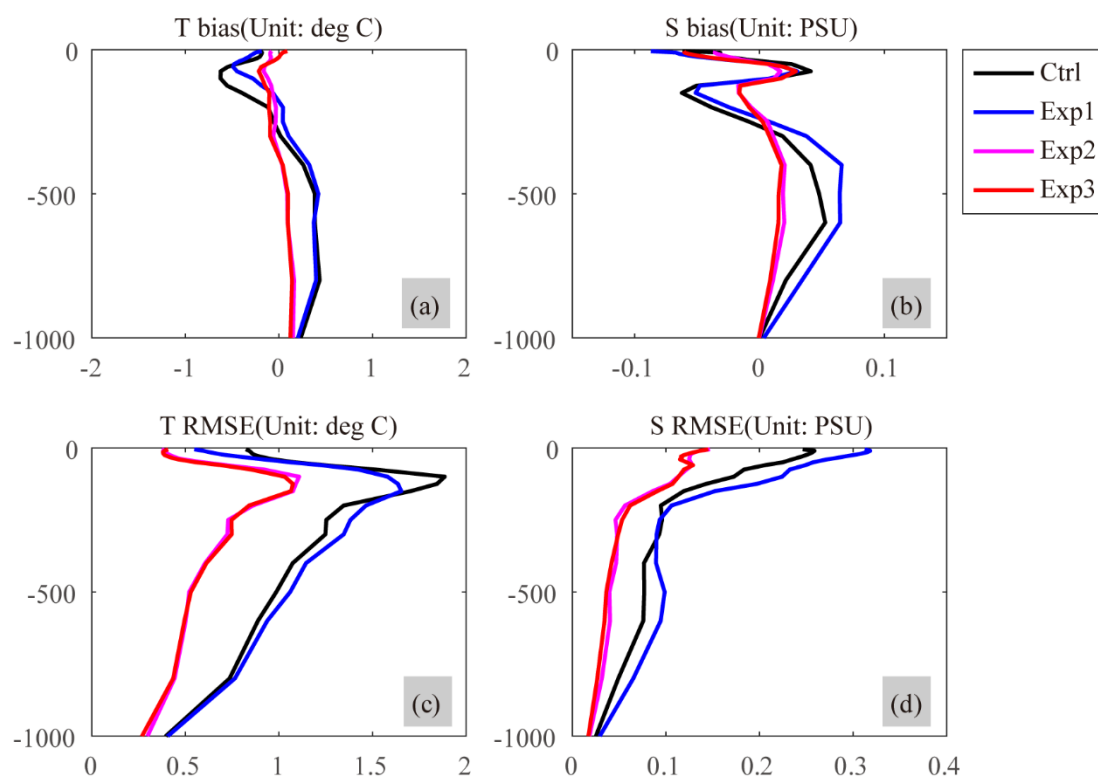


Fig. 13. Mean bias (Top panel) and RMSE (Bottom panel) of temperature and salinity over 2 months in the whole model region ($99^{\circ}\text{E}\sim 134^{\circ}\text{E}$, $1^{\circ}\text{N}\sim 30^{\circ}\text{N}$) where deeper than 200 m.

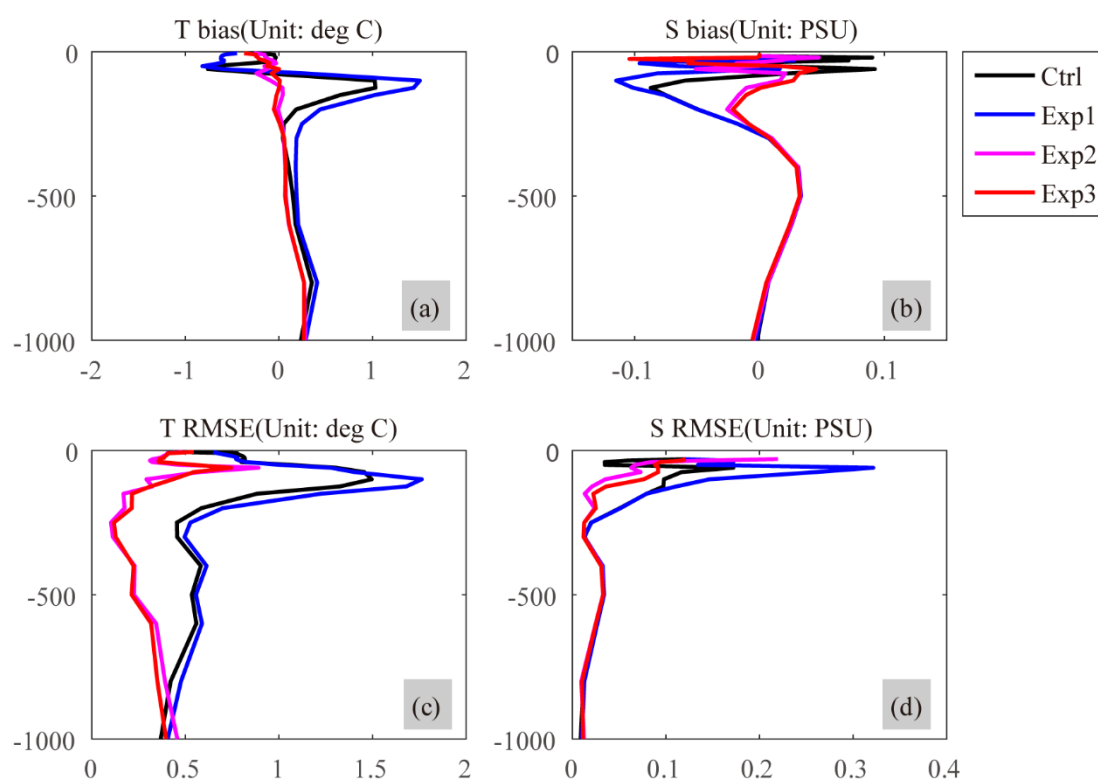


Fig. 14. Mean bias and RMSE of temperature and salinity over 2 months in the western

875 SCS (110-114°E, 13 - 16°N).