An EnKF-based ocean data assimilation system improved by adaptive observation error inflation (AOEI)

Shun Ohishi[1, 2, 3, 4], Takemasa Miyoshi[1, 2, 3], and Misako Kachi[5]

[1]RIKEN Center for Computational Science, Kobe, 6500047, Japan
5   [2]RIKEN Cluster for Pioneering Research, Kobe, 6500047, Japan
[3]RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), Kobe, 6500047, Japan
[4]Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, 4648601, Japan
[5]Earth Observation Research Center, Japan Aerospace Exploration Agency, Tsukuba, 3058505, Japan

10   *Correspondence to*: Shun Ohishi (shun.ohishi@riken.jp) and Takemasa Miyoshi (takemasa.miyoshi@riken.jp)

Abstract.

A previous study proposed an adaptive observation error inflation (AOEI) method for an ensemble Kalman filter-based atmospheric data assimilation system to assimilate all-sky infrared brightness temperatures. Brightness temperature differences between clear- and cloudy-sky radiances are large, and observation-minus-forecast differences or innovations are therefore likely to be large around boundaries between clear- and cloudy-sky regions. The AOEI method mitigates these discrepancies by adaptively inflating observation errors. Ocean frontal regions have similar characteristics to the borders between clear- and cloudy-sky regions with large innovations. Consequently, we have implemented the AOEI with an EnKF-based regional ocean data assimilation system, in which the assimilation interval is set to one day to utilize frequent satellite observations. We conducted sensitivity experiments to investigate the impacts of the AOEI on salinity structure, geostrophic balance, and accuracy. A control run, in which the AOEI is not applied, shows the degradation of low-salinity North Pacific Intermediate Water around the Kuroshio Extension region, where the innovation amplitude and forecast ensemble spread are large in association with the fronts and eddies. The resulting large temperature and salinity increments weaken the density stratification, leading to large vertical diffusivity. As a result, the low salinity water in the intermediate layer is lost through strong vertical diffusion. When the AOEI is used, the salinity structure in the ocean interior is preserved because the AOEI suppresses the salinity degradation by reducing the temperature and salinity increments. We also demonstrate that the AOEI provides significant improvement of the geostrophic balance and the analysis accuracy of temperature, salinity, and surface flow fields.

Short summary (487/500 character)

An adaptive observation error inflation (AOEI) method was proposed for atmospheric data assimilation to mitigate erroneous analysis updates caused by large observation-minus-forecast differences for satellite brightness temperature around clear- and cloudy-sky boundaries. This study implemented the AOEI with an ocean data assimilation system, leading to an improvement of analysis accuracy and dynamical balance around the frontal regions with large meridional temperature differences.

1. Introduction

The ensemble Kalman filter (EnKF) estimates flow-dependent forecast errors from an ensemble of model forecasts, and calculates the best estimates (i.e., analyses) by combining forecasts and observations with their error covariances (Evensen, 1994, 2003). The EnKF has the advantage of being easy to implement for various models (See table 1 of Ohishi et al., 2022), but it has been used in only two ocean reanalysis datasets thus far (Balmaseda et al., 2015; Martin et al., 2015): the Predictive Ocean Atmosphere Model for Australia (PAOMA) Ensemble Ocean Data Assimilation System (PEODAS; Yin et al., 2011) and TOPAZ4 (Sakov et al., 2012). In contrast, the three-dimensional variational method (3D-VAR) is the most widely used in ocean analysis datasets (e.g. Miyazawa et al., 2017; Zuo et al., 2019).

With the enhancement of in-situ and satellite observations, the number of observations has increased dramatically. Argo profiling float observations since the 2000s provide a large number of in-situ temperature and salinity data in the ocean interior. Although satellite sea surface salinity (SSS) data since 2010 are relatively inaccurate, particularly in coastal and high-latitude regions (Abe and Ebuchi, 2014), previous studies have demonstrated the positive effects of SSS assimilation on the analyses of ocean interior structure such as mixed and barrier layers (Chakraborty et al., 2015), low salinity water caused by river discharge (Toyoda et al., 2015), and El Niño-Southern Oscillation prediction (Hackert et al., 2011). A Japanese geostationary satellite, Himawari-8 (Bessho et al., 2016; Kurihara et al., 2016), has observed sea surface temperatures (SSTs) in the Pacific region at high spatiotemporal resolutions of 2 km and 10 minutes since July 2015. The Surface Water and Ocean Topography (SWOT) satellite is scheduled to be launched in 2022 and will provide high-resolution and two-dimensional sea surface height (SSH) anomalies (SSHAs).

For effective use of dense and frequent satellite observations, Ohishi et al. (2022) performed sensitivity experiments using an EnKF-based ocean data assimilation system with an assimilation interval of one day, which is more frequent than the five- and seven-day intervals in the existing EnKF-based systems (PEODAS and TOPAZ4, respectively). They demonstrated that the combination of incremental analysis updates (IAU; Bloom et al., 1996) and relaxation-to-prior perturbations (RTPP; Kotsuki et al., 2017; Zhang et al., 2004) to restore the forecast ensemble perturbations toward the analysis by 80%–90% produced optimal results in terms of both dynamic balance and accuracy. However, their system contained several tuning parameters such as observation errors, ensemble size, and localization scale. Previous studies have prescribed observation errors in various ways, e.g., using spatiotemporally fixed constants (Miyazawa et al., 2012; Xu and Oey, 2014); assuming them to be standard deviations calculated from historical observations (Miyazawa et al., 2009; Usui et al., 2006); estimating them from other assimilation datasets (Penny et al., 2013); and assuming that observation error covariance matrices are proportional to the forecast error covariance matrices (Carton et al., 2018; Yin et al., 2011). A technique to adaptively inflate the observation errors based on the innovation statistics (Desroziers et al., 2005), which is called as an adaptive observation error inflation (AOEI) method, was recently proposed for assimilating all-sky infrared satellite brightness temperatures in an atmospheric data assimilation system (Minamide and Zhang, 2017; Zhang et al., 2016). Since the brightness temperature differences between clear- and cloudy-sky radiances are large, there are large observation-minus-forecast differences or innovations around boundaries between clear- and cloudy-sky regions even for the tiny boundary differences between forecasts and

observations. This results in erroneous analysis increments and degrades the analysis. The AOEI mitigates the large discrepancies between forecasts and observations by adaptively inflating the observation errors. Ocean frontal regions such Kuroshio and Kuroshio Extension (KE) regions have large spatiotemporal variations, and the innovations around the frontal regions also tend to be large even for small differences of frontal positions between the forecasts and observations. Therefore, ocean fronts have similar characteristics to the borders between clear- and cloudy-sky regions with large innovations, and the AOEI method is therefore expected to be useful for improving EnKF-based ocean data assimilation systems.

This study aims to investigate the causes of the salinity degradation around the KE region, and to evaluate the impacts of the AOEI on the salinity structure, dynamical balance, and accuracy. The remainder of this paper is organized as follows: details of the AOEI method, the experiment design, the temperature and salinity budget equations, and the methods to evaluate geostrophic balance and accuracy in sensitivity experiments are presented in Sect. 2; Section 3 describes the causes of the salinity degradation in the intermediate layer and the positive impacts of the AOEI on the geostrophic balance and accuracy for temperature, salinity, and surface flow. A summary is provided in Sect. 4.

2. Methods

2.1. Adaptive observation error inflation (AOEI)

Manual tuning of observation errors is computationally expensive, and several studies have proposed adaptive estimation methods using the innovation statistics of Desroziers et al. (2005):

$$\langle \boldsymbol{d}_b^o (\boldsymbol{d}_b^o)^{\mathrm{T}} \rangle \approx \mathbf{H}\mathbf{P}^{\mathbf{b}}\mathbf{H}^{\mathrm{T}} + \mathbf{R}. \tag{1}$$

Here, $\langle \cdot \rangle$ denotes the statistical expectation, $\boldsymbol{d}_b^o$ $(= \boldsymbol{y} - \mathbf{H}\overline{\boldsymbol{x}^b})$ is an innovation vector, where $\boldsymbol{y}$, $\mathbf{H}$, and $\overline{\boldsymbol{x}^b}$ denote an observation vector, linear observation operator, and forecast ensemble mean state vector, respectively. $\mathbf{P}^{\mathbf{b}}$ and $\mathbf{R}$ are the forecast and observation error covariance matrices, respectively. Expressing Eq. (1) in a scalar form, the observation error σ_{est-o} may be estimated by

$$\sigma_{est-o}^2 = (d_b^o)^2 - \sigma_{H(x^b)}^2, \tag{2}$$

where $\sigma_{H(x^b)}$ is the forecast ensemble spread in observation space. Here, the forecast ensemble spreads are assumed to be accurate, and $(d_b^o)^2$ is assumed to be equivalent to $\langle (d_b^o)^2 \rangle$. To avoid underestimation of the observation errors, larger observation errors σ_o between the estimated and prescribed errors are used in the AOEI method (Minamide and Zhang, 2017; Zhang et al., 2016):

$$\sigma_o^2 = \max\{\sigma_{pre-o}^2, \sigma_{est-o}^2\}, \tag{3}$$

100

where σ_{pre-o} is the prescribed observation error. As described in Sect. 1, the AOEI suppresses erroneous analysis increments associated with systematic errors, biases, and representation errors by adaptively inflating the observation errors when the squared innovation is larger than the sum of the prescribed observation and ensemble-based forecast error variances.

2.2. Experimental design

This study uses an EnKF-based regional ocean data assimilation system known as sbPOM-LETKF (Ohishi et al., 2022), comprising a sigma-coordinate regional ocean model, the Stony Brook Parallel Ocean Model version 1.0 (sbPOM; Jordi and Wang, 2012; Ohishi et al., 2022), and a three-dimensional local ensemble transform Kalman filter (3D-LETKF; Hunt et al., 2007; Miyoshi and Yamane, 2007). The sbPOM is configured for the northwestern Pacific region (117°–180°E, 15°–50°N) with horizontal resolution of 0.25° and 50 sigma layers. The bottom topography is taken from a 1 arc-minute global relief model of Earth's surface (ETOPO1; Amante and Eakins, 2009) and is smoothed by a Gaussian filter with a 200 km e-folding scale to reduce pressure gradient errors at steep bottom slopes (Mellor et al., 1994). Monthly (seasonal) temperature and salinity climatologies from the World Ocean Atlas 2018 (WOA18; Locarnini et al., 2019; Zweng et al., 2019) with horizontal resolution of 1° and 57 (103) layers are used for the initial conditions over depths shallower (deeper) than 1500 m. Lateral boundary conditions for temperature, salinity, and horizontal velocity are derived from the Simpler Ocean Data Assimilation (SODA) version 3.7.2 (Carton et al., 2018) with horizontal resolution of 0.5° and 50 layers. The Japanese 55-year Reanalysis (JRA55; Kobayashi et al., 2015) with horizontal and temporal resolution of 1.25° and 6 hours, respectively, is adopted for the atmospheric boundary conditions including air temperature and specific humidity at 2 m, wind velocity at 10 m, shortwave radiation, total cloud fraction, sea level pressure, and precipitation. River discharge is obtained from the Japan Aerospace Exploration Agency (JAXA)'s land surface and river simulation system, Today's Earth (TE)-Global (https://www.eorc.jaxa.jp/water/, last access: 24 November 2022), with horizontal and temporal resolution of 0.25° and 3 hours, respectively. To avoid filter divergence, the atmospheric and lateral boundary conditions other than rainfall and river discharge are perturbed in the same way as in Ohishi et al. (2022). The model with 100 ensemble members is spun up from 1 January 2011 to 6 July 2015, using the initial conditions with no motion. During the spin-up period, simulated temperature and salinity are nudged towards the monthly climatology from the WOA18 with a 90-day timescale to prevent northward overshoot of the Kuroshio along the east coast of Japan.

The LETKF with 100 ensemble members and on a one-day assimilation cycle is used to assimilate observations as below: satellite SSTs from Himawari-8 (Bessho et al., 2016; Kurihara et al., 2016) and the Global Change Observation Mission–Water (GCOM-W: https://gportal.jaxa.jp/gpr/?lang=en, last access: 24 November 2022); satellite SSS from Soil Moisture and Ocean Salinity (SMOS: http://www.esa.int/Applications/Observing_the_Earth/SMOS, last access: 24 November 2022) and Soil Moisture Active Passive (SMAP) version 4.3 (Meissner et al., 2018); SSH estimated by summing satellite SSH anomalies from the Copernicus Marine Environment Monitoring Service (CMEMS: https://marine.copernicus.eu/, last access: 24 November 2022) and mean dynamic ocean topography obtained by averaging the simulated SSH over 2012–14; and in-situ

temperatures and salinity from the Global Temperature and Salinity Profile Programme (GTSPP; Sun et al., 2010) and Advance

135 automatic QC (AQC) Argo Data version 1.2a (AQC: https://www.jamstec.go.jp/argo_research/dataset/aqc/index_dataset.html, last access: 24 November 2022).

Covariance localization in observation space is applied using the Gaussian function with horizontal and vertical localization scales $L = 300$ km and 100 m, respectively, following Miyazawa et al. (2012) and Penny et al. (2013). We assume that the localization function becomes zero beyond $2\sqrt{10/3}\,L \approx 1100$ km (370 m) in the horizontal (vertical)

140 direction (Miyoshi et al., 2007). Following Miyazawa et al. (2012), the prescribed observation errors for temperature, SSH, and salinity are set to 1.0 °C, 0.2 m, and 0.3, respectively. Here, we set the larger salinity observation errors than those from Miyazawa et al. (2012), because the SSS satellite observations are relatively noisy and the measurement errors would be large (Meissner et al., 2018). We adopt the combination of the IAU (Bloom et al., 1996; Ohishi et al., 2022) and RTPP (Kotsuki et al., 2017; Zhang et al., 2004), in which the analysis ensemble perturbations are relaxed toward the forecast ensemble

145 perturbations by 90 % while maintaining the analysis ensemble mean, because the sensitivity experiments in Ohishi et al. (2022) demonstrated that it results in the best dynamical balance and accuracy. Although this may not be optimal, our computational resources are limited, and thus the RTPP relaxation parameter is fixed at 90%.

To highlight the impacts of the AOEI on the ocean salinity structure and dynamical balance, we conduct AOEI and control (CTL) runs with and without applying the AOEI, respectively, from the start date of the Himawari-8 observation (7

150 July 2015) to 31 December 2015. Furthermore, we perform the 1.5Terr run with the same setting as the CTL run but with the temperature observation error of 1.5 °C, and then compare the accuracy between the CTL, AOEI, and 1.5Terr runs. During the assimilation period, the SSS nudging with a 90-day timescale is applied to prevent a surface freshening drift as in the spin-up period.

155 **2.3. Temperature and salinity budget equations in the ocean interior**

To quantitatively investigate ocean interior temperature and salinity differences between the AOEI and CTL runs, respectively, we use the temperature T and salinity S budget equations:

$$\frac{\partial T}{\partial t} = \nabla \cdot (\boldsymbol{\kappa} \circ \nabla T) - \boldsymbol{v} \cdot \nabla T + \frac{q_{sw}}{\rho_0 c_p} + (T\ increment)\ \text{and} \qquad (4)$$

160
$$\frac{\partial S}{\partial t} = \nabla \cdot (\boldsymbol{\kappa} \circ \nabla S) - \boldsymbol{v} \cdot \nabla S + (S\ increment). \qquad (5)$$

Here, $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ denotes the three-dimensional gradient operator, $\boldsymbol{\kappa} = (\kappa_x, \kappa_y, \kappa_z)$ is a diffusivity vector, $\circ$ indicates a Schur product, $\boldsymbol{v} = (u, v, w)$ is three-dimensional velocity, $\rho_0 = 1025\ kg\ m^{-3}$ is the reference density, $c_p = 4190\ J\ kg^{-1}\ °C^{-1}$ is the specific heat of the seawater, and q_{sw} is downward shortwave radiation parameterized by

165

$$q_{sw} = Q_{sw} \left\{ R_{sw} exp\left(-\frac{|z|}{\gamma_1}\right) + (1 - R_{sw})exp\left(-\frac{|z|}{\gamma_2}\right) \right\} \qquad (6)$$

(Paulson and Simpson, 1977), where Q_{sw} is shortwave radiation at the sea surface, $R_{sw} = 0.62$ is a separation constant, and $\gamma_1 = 0.60\ m$ and $\gamma_2 = 20.0\ m$ are attenuation length scales. These values are set to the case of Type IA from Jerlov (1976).

170 ($T\ increment$) and ($S\ increment$) indicate the temperature and salinity analysis increments, respectively. Equations (4) and (5) do not include a residual term because each term of the temperature and salinity budget equations is accumulated at each model timestep and each grid, respectively, and the daily-mean outputs are saved in this system.

2.4. Evaluation method

175 As in Ohishi et al. (2022), this study evaluates geostrophic balance and accuracy using the nonlinear balance equation (NBE) and root mean square deviations (RMSDs) relative to observations, respectively (see subsections 2.4.1 and 2.4.2).

2.4.1. Nonlinear balance equation

For the analysis fields, the geostrophic balance equation is represented as

180

$$f\mathbf{k} \times \delta\mathbf{u} = -g\mathbf{\nabla_h}\delta\eta, \qquad (7)$$

where f is the vertical component of the Coriolis parameter, $\mathbf{k}$ is a unit vector in the vertical direction, δ is the analysis increment, $\mathbf{u} = (u, v)$ denotes horizontal velocity at the sea surface, $g = 9.8\ m\ s^{-2}$ is gravitational acceleration, $\mathbf{\nabla_h} =$

185 $(\partial/\partial x, \partial/\partial y)$ is the horizontal gradient operator, and η denotes SSH. By taking $\partial/\partial x$ of the x-component plus $\partial/\partial y$ of the y-component of Eq. (7), the geostrophic equation can be reduced to the nonlinear balance equation (NBE; Shibuya et al., 2015; Zhang et al., 2001):

$$-f\delta\zeta + \beta\delta u + g\mathbf{\nabla_h^2}\delta\eta = 0, \qquad (8)$$

190

where $\zeta = \partial v/\partial x - \partial u/\partial y$ is the relative vorticity at the sea surface, and $\beta = \partial f/\partial y$ is the planetary vorticity gradient. If the analysis fields do not satisfy geostrophic balance, there is an absolute NBE residual:

$$\Delta NBE \equiv \left| -f\delta\zeta + \beta\delta u + g\mathbf{\nabla_h^2}\delta\eta \right|, \qquad (9)$$

195

where $|\cdot|$ indicates taking the absolute value. A smaller (larger) ΔNBE indicates more (less) geostrophic balance, and smaller (larger) initial shocks tend to occur.

2.4.2. RMSD

200 We evaluate the analysis accuracy of temperature, salinity, horizontal velocities, and SSH using the RMSDs calculated relative to the following observations: in-situ temperature and salinity over 1–525 m depth and in-situ horizontal velocity over 8–36 m depth at (144.6°E, 32.3°N) south of the KE from the Kuroshio Extension Observatory (KEO) buoy (https://www.pmel.noaa.gov/ocs/, last access: 24 November 2022; See Fig. 11), SSH and SSHA gridded datasets with horizontal resolution of 0.25° from Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO; Ducet

205 et al., 2000), in-situ surface horizontal velocity from surface drifter buoys of the Global Drifter Program (Elipot et al., 2016), and Himawari-8 SSTs. We note that the AVISO and Himawari-8 are not independent because the satellite SSH anomalies and SSTs are used in this system, respectively, whereas the KEO and surface drifter buoys are independent observations. The validation in the ocean interior in this study is limited due to the paucity of available independent observations.

In this study, we calculate the ΔNBE (RMSDs) using daily outputs from the CTL and AOEI runs (the CTL, AOEI,

210 and 1.5Terr runs). To compare the AOEI run (AOEI and 1.5Terr runs) with the CTL run, we also calculate improvement ratios IR for ΔNBE (RMSD):

$$IR_{NBE} = \frac{(\Delta NBE)_{CTL} - (\Delta NBE)_{AOEI}}{(\Delta NBE)_{CTL}} \times 100 \text{ and} \tag{10}$$

$$IR_{RMSD} = \frac{(RMSD)_{CTL} - (RMSD)_{AOEI/1.5Terr}}{(RMSD)_{CTL}} \times 100. \tag{11}$$

215

Here, the subscripts CTL, $AOEI$, $1.5Terr$ indicate the CTL, AOEI, and 1.5Terr runs, respectively. Using the bootstrap method with 10,000 cycles, we detect significant improvement and degradation in the AOEI and 1.5Terr runs relative to the CTL run at the 99% confidence level.

220 **3. Results**

In subsection 3.1, the degradation of the low-salinity structure in the CTL run is described. Subsection 3.2 presents how the AOEI is applied to SST, SSS, and SSH fields. The detail of the improvement of the low-salinity structure by the AOEI is provided in subsection 3.3, and the results of the geostrophic balance and accuracy are described in subsection 3.4.

225 **3.1. Salinity degradation of the NPIW around the KE region in the CTL run**

As shown in Fig. 1, the SST field in the CTL run agrees well with the satellite observations. Although the satellite-derived SSS has large errors, especially in coastal and high-latitude regions (Abe and Ebuchi, 2014), the SSS spatial pattern appears to be reproduced well in the analysis field (Fig. 2). However, the CTL run has noisier signals in the latter half of the experiment period, particularly in the SSS analysis fields. We also assess the monthly-mean temperature, salinity, and potential

230 density σ_θ along 150°E and 35°N sections across the KE. During the initial stages of the experiment period, the North Pacific Intermediate Water (NPIW), characterized by low minimum salinity, is distributed within $\sigma_\theta = 26.5$–27.25 kg m^{-3} (Talley,

1993; Yasuda, 1997) (Fig. 3a, b). However, as the assimilation period progresses, the low salinity structure in the intermediate layer around the KE region is lost along with the noisy signals (Fig. 3d, e, g, h). In contrast, the temperature stratification dominates for the density stratification in this region, and therefore the density and temperature structure persists with higher density and lower temperatures at deeper depths, respectively. The noisy signals and degradation of the low-salinity structure do not appear during the spin-up period. To quantitatively investigate the cause of the salinity degradation in the CTL run, we calculate the salinity budget equation [Eq. (5)] in the intermediate layer around the KE region (white boxes in Fig. 3g–i; 140°–160°E, 30°–40°N, 500–1000 m depth) (See the detail in Appendix A). The result shows that the vertical diffusion is the main cause of the salinity degradation.

3.2. Spatiotemporal characteristics of the AOEI application in the surface fields

To investigate how much the AOEI applies to the SST, SSS, and SSH fields, we calculate the monthly-mean ratio of the area where the AOEI is applied to the entire system domain (Fig. 4). Application of the AOEI to the SSS field is the highest at around 35%–40% of the domain because the instantaneous satellite observations are noisy (cf. Fig.2; https://smos-diss.eo.esa.int/socat/SMOS_Open, last access: 24 November 2022). The AOEI is also applied to the SST field at a relatively high ratio of 5%–10%, whereas the ratio in the SSH field is exceedingly small (less than 0.1%). This indicates that the AOEI method is applied substantially to the SST and SSS fields, and rarely to the SSH field.

We also examine the spatial characteristics of where the AOEI is applied to the SST and SSS fields by calculating the ratio of the period when the AOEI method is applied compared with the total experiment period (Figs. 5, 6). High SST ratios are distributed in the coastal and frontal regions, including the Kuroshio, the KE, and a subpolar front along J1 around 150°E, 40°N (Isoguchi et al., 2006; Kida et al., 2015; Fig.5a). The SSS ratios are high in the East China Sea, Japan Sea, and high-latitude regions (Fig. 6a). The spatial pattern of the positive and negative innovation phases is asymmetric in both the SST and SSS fields (Figs. 5b, c, 6b, c). In the positive innovation phase, the high SST ratios are distributed only along the northeastern coast of Japan at 140°–150°E, 40°–50°N (Fig. 5b), whereas in the negative innovation phase, high SST ratios are more widely distributed, covering coastal and frontal regions (Fig. 5c). In the negative innovation phase, the SSS ratios are higher in the East China Sea, Japan Sea, and high-latitude regions (Fig. 6b, c). In the SST and SSS fields, the spatial patterns of the positive forecast biases correspond closely to the high ratios in the negative innovation phase (Figs. 5c, 6c, 7). Therefore, the forecast SST and SSS biases lead to the asymmetry in which the AOEI is applied more during negative innovation phases than during positive phases, as seen in Fig. 4a, b.

In the SST field, large innovation amplitude and forecast ensemble spread are distributed along the KE and the J1 (Fig. 5d, e). In the SSS field, the ensemble spread is large along the KE and J1, where the salinity innovation amplitude is large and exceeds 1.0 (Fig. 6d, e). This demonstrates that large temperature and salinity analysis increments are likely to be generated in the KE and J1 regions if the AOEI is not applied, as in the CTL run.

3.3. Improvements of the salinity structure by the AOEI

We compare monthly temperature and salinity fields between the CTL and AOEI runs at the sea surface and along the 150°E and 35°N sections. In the AOEI run, noisy signals are reduced in the temperature and salinity fields especially in the ocean interior (right panels in Fig. 3), and the low salinity water persists in the intermediate layer. The salinity budget analysis indicates that this salinity improvement results from the reduction of the vertical salinity diffusion in the AOEI run relative to the CTL run (See the detail in Appendix B). Figure 8 shows the vertical profile of the vertical diffusivity κ_z averaged over the KE region (140°–160°E, 30°–40°N) for the whole experiment period and the maximum of the averaged diffusivity over the whole experiment period within 300–1000 m depth. As is consistent with the results of the salinity budget analysis, there is exceedingly large vertical diffusivity at 300–800 m depth around the KE region, which results in salinity degradation induced by strong vertical diffusion in the CTL run. In contrast, the low salinity water in the intermediate layer persists in the AOEI run because the vertical diffusivity is smaller.

Weak density stratification and strong vertical shear are favorable conditions for the generation of large vertical diffusivity (Davis et al., 2016; Pacanowski and Philander, 1981). To gain dynamical insight into the vertical diffusivity difference between the AOEI and CTL runs, the temporal tendency of the vertical diffusivity κ_z, the squared buoyancy frequency $N^2 = - g/\sigma_\theta \, \partial\sigma_\theta/\partial z$, and the squared vertical shear $\boldsymbol{u}_z^2 = |\partial\boldsymbol{u}/\partial z|^2$ ($\partial\kappa_z/\partial t$, $\partial N^2/\partial t$, and $\partial\boldsymbol{u}_z^2/\partial t$, respectively) are summed during the positive vertical diffusivity tendency ($\partial\kappa_z/\partial t > 0$), and are then averaged in the KE region (140°–160°E, 30°–40°N). The buoyancy frequency can be represented as the sum of the contributions from the temperature and salinity vertical gradients (N_T^2 and N_S^2, respectively):

$$N^2 = -\frac{g}{\sigma_\theta}\frac{\partial\sigma_\theta}{\partial z} = g\left(\alpha_T\frac{\partial T}{\partial z} - \beta_S\frac{\partial S}{\partial z}\right) \equiv N_T^2 + N_S^2. \tag{12}$$

Here, α_T (β_S) is the thermal (salinity) expansion coefficient.

Figure 9a, b shows that the total vertical diffusivity tendency is smaller in the AOEI run than the CTL run, which agrees qualitatively with the diffusivity averaged over the whole period (Fig. 8a, b). As is clear from Fig. 9c, d, the total shear tendency is almost zero in both the CTL and AOEI runs. The total buoyancy frequency tendency makes substantial contributions in the CTL and AOEI runs, and its amplitude is smaller in the AOEI run than the CTL run. Since the negative values indicate weakening of the density stratification, the density stratification is less weakened in the AOEI run than the CTL run. The difference in the total buoyancy frequency tendency between the AOEI and CTL runs is caused by the differences of both the total $\partial N_T^2/\partial t$ and $\partial N_S^2/\partial t$ (Fig. 9e). Although $\partial N_T^2/\partial t$ ($\partial N_S^2/\partial t$) can be decomposed into the temporal tendency terms of the vertical temperature (salinity) gradient and the thermal (salinity) expansion coefficient, we confirmed that the latter terms are almost zero and have almost no impact on $\partial N_T^2/\partial t$ and $\partial N_S^2/\partial t$. Therefore, the differences in the temperature and salinity vertical gradient tendencies result in less weakening of the density stratification in the AOEI run than the CTL run.

To investigate the causes of the differences in the temperature and salinity vertical gradient tendencies between the AOEI and CTL runs, we derive the temperature and salinity stratification tendency equations, respectively, by taking the vertical derivatives of the temperature and salinity budget equations [Eqs. (4) and (5)] in the ocean interior:

300

$$\frac{\partial}{\partial t}\left(\frac{\partial T}{\partial z}\right) = \frac{\partial}{\partial z}\{\boldsymbol{\nabla} \cdot (\boldsymbol{\kappa} \circ \boldsymbol{\nabla}T)\} - \frac{\partial}{\partial z}(\boldsymbol{v} \cdot \boldsymbol{\nabla}T) + \frac{1}{\rho_0 c_p}\frac{\partial q_{sw}}{\partial z} + \frac{\partial}{\partial z}(T\ increment)\ \text{and} \tag{13}$$

$$\frac{\partial}{\partial t}\left(\frac{\partial S}{\partial z}\right) = \frac{\partial}{\partial z}\{\boldsymbol{\nabla} \cdot (\boldsymbol{\kappa} \circ \boldsymbol{\nabla}S)\} - \frac{\partial}{\partial z}(\boldsymbol{v} \cdot \boldsymbol{\nabla}S) + \frac{\partial}{\partial z}(S\ increment). \tag{14}$$

As in the total vertical diffusivity tendency calculated the above, each term in Eqs. (13) and (14) is summed when $\partial \kappa_z / \partial t > 0$, and then averaged in the KE region (140°–160°E, 30°–40°N) (Fig. 10). We note that positive values in Eqs. (13) and (14) indicate opposite effects on the density stratification: a positive temperature (salinity) vertical gradient tendency strengthens (weakens) the density stratification.

In the CTL and AOEI runs, the temperature gradient tendency term [the left-hand side (LHS) term of Eq. (13)] is negative and indicates that the temperature and density stratification are weakened at all depths (Fig. 10a, b). The amplitude of this term is smaller in the AOEI run than the CTL run, and thus the temperature and density stratification is less weakened. As shown in Fig. 10c, the difference in the temperature gradient tendency between the AOEI and CTL runs is due mainly to the temperature analysis increment gradient term [the last term on the right-hand side (RHS) of Eq. (13)] and in part to the advection gradient term [the second term on the RHS of Eq. (13)], whereas the diffusion and shortwave penetration gradient terms [the first and third terms on the RHS of Eq. (13), respectively] make almost no contribution.

In the CTL run, the salinity gradient tendency term [the LHS term of Eq. (14)] indicates that the salinity (density) stratification is strengthened (weakened) at all depths (Fig. 10d). In the AOEI run, the salinity (density) stratification is weakened (strengthened) at 200–400 m depth and slightly strengthened (weakened) at 400–1000 m depth (Fig. 10e). The salinity gradient tendency term is smaller in the AOEI run than the CTL run at all depths, and thus the salinity (density) stratification is less strengthened (weakened) in the AOEI run relative to the CTL run (Fig. 10f). The difference in the salinity analysis increment gradient terms [the last term on the RHS of Eq. (14)] between the AOEI and CTL runs dominates that in the salinity gradient tendency term, whereas the differences between the diffusion and advection gradient terms [the second and third terms on the RHS of Eq. (14), respectively] have little influence. This indicates that less strengthening (weakening) of the salinity (density) stratification in the AOEI run relative to the CTL run is due to the smaller salinity analysis increment. The impacts of the SST, SSS, and SSH assimilation are limited to between the surface and about 370 m depth because of the prescribed vertical localization scale of 100 m described in subsection 2.2, and consequently only in-situ temperature and salinity assimilation generates the analysis increments in the intermediate layer.

The AOEI contributes to maintaining the density stratification by reducing the temperature and salinity analysis increments and preventing the occurrence of large vertical diffusivity that degrades low salinity water in the intermediate layer around the KE region. In the CTL run, the salinity analysis increments restore the degraded low salinity water (See Appendix

330   A) but lead to degradation through the formation of large vertical diffusivity at the same time. Thus, it seems that positive feedback exists that may degrade the salinity structure.

3.4 Improvement of geostrophic balance and accuracy by the AOEI

In this subsection, we investigate the impacts of the AOEI on the geostrophic balance and accuracy. Figure 11 shows ΔNBE

335   averaged over the whole period in the CTL and AOEI runs. In the CTL run, ΔNBE is large in the mid-latitude regions, especially along the KE (Fig. 11a). In the AOEI run, ΔNBE is smaller than the CTL run for the entire domain (Fig. 11b). The spatiotemporal averaged ΔNBE over the whole experiment period and domain is 0.57×10^{-10} s^{-2} and 0.35×10^{-10} s^{-2} for the CTL and AOEI runs, respectively, and the balance is significantly improved in the AOEI run relative to the CTL run. This is probably because the analysis increments are smaller in the AOEI run than the CTL run.

340   To investigate the analysis accuracy in the ocean interior, we calculate the RMSDs of the CTL, AOEI, and 1.5Terr runs relative to in-situ temperature, salinity, and horizontal velocity observations from the KEO buoy south of the KE (Figs. 11a, 12). Results are only presented for the temperature and salinity because no significant results are obtained for the horizontal velocities. The RMSDs for both temperature and salinity are smaller in the AOEI run than the CTL run, and the AOEI run provides significant temperature (salinity) improvements at 0–150 m (50–400 m) depth relative to the CTL run.

345   This is probably because the AOEI suppresses the development of the strong vertical diffusion that leads to the salinity degradation and because of the improvement in the geostrophic balance. We have confirmed that the AOEI run also has smaller temperature (salinity) RMSDs than the 1.5Terr run throughout the depth except for two observation points at 225 and 275 m (150 and 525 m) depth. Therefore, among the experiments, the AOEI run is the best for the accuracy of temperature and salinity south of the KE.

350   We also investigate the analysis accuracy of the SSH, surface flow, and SST fields, respectively, calculating the spatiotemporally averaged RMSDs relative to the SSH and SSHA datasets from the AVISO, to in-situ surface horizontal velocity observations from the drifter buoys, and to Himawari-8 SSTs (Fig. 13). Although the AOEI run slightly degrades the SST accuracy relative to the CTL run (Fig. 13e), the RMSDs in the AOEI run are smaller for all other variables, and indicate significant improvements except for surface meridional velocity (Fig. 13a–d). The improvement of the surface flow field in

355   the AOEI run would result from the better geostrophic balance and accuracy of the density structure in the ocean interior. Relative to the 1.5Terr run, the AOEI run improves the accuracy for all variables.

       Kurihara et al. (2016), for example, show that the RMSDs of the Himawari-8 SSTs relative to the buoys are about 0.5°C and are larger in the higher latitude regions with larger zenith angle, and observation error variances have substantial contributions to the RMSDs. However, the ensemble spreads are much smaller than the RMSDs for all variables (Fig. 13). We

360   have found that the ensemble spreads in the subtropical region appear to be under-dispersive even if the perturbed atmospheric and lateral boundary conditions are applied (cf. Figs. 5e, 6e). Methods to inflate the ensemble spread more would be required for further improvements of the accuracy, but this will be a future issue.

4. Summary

365   We have implemented the AOEI with the sbPOM-LETKF ocean data assimilation system and conducted sensitivity experiments to investigate the impacts on the low-salinity NPIW around the KE region, geostrophic balance, and analysis accuracy. In the CTL run, the large analysis increments by in-situ temperature and salinity assimilation weaken the density stratification. The resulting exceedingly large vertical diffusivity induces the strong vertical diffusion that breaks the low salinity structure in the NPIW around the KE region. The salinity analysis increment contributes to restoring the low salinity

370   water, but at the same time causes the salinity degradation by generating strong vertical diffusion. Therefore, the positive feedback appears to occur, degrading the salinity structure.

        The AOEI decreases the temperature and salinity analysis increments around the KE region by adaptively inflating the temperature and salinity observation errors, respectively. As a result, the AOEI mitigates the salinity degradation seen in the CTL run, and therefore the low salinity water is maintained in the AOEI run. In addition, the AOEI significantly improves

375   the geostrophic balance probably because of the reduction of the analysis increments. Moreover, the AOEI prevents the development of strong vertical diffusion and improves the accuracy of temperature and salinity in the ocean interior. Furthermore, the improvements of the geostrophic balance and density structure in the ocean interior contribute to more accurate SSH and surface flow fields. In summary, this study demonstrates the positive impacts of the AOEI on the balance and accuracy of the temperature, salinity, and surface flow fields.

380   Since our available computational resources were limited, we fixed the tuning parameter of the RTPP, perturbed atmospheric forcing, ensemble size, localization scale, and prescribed observation errors. Further experiments to explore more optimal settings are required, and this will be investigated in the future. Coastal data assimilation systems with high horizontal resolution might reach the stage where they capture sub-mesoscale phenomena such as filaments with strong temperature and salinity gradients. For such systems, the position errors of fronts, eddies, and filaments might cause degradation as seen in the

385   CTL run. Furthermore, low salinity water is distributed in the intermediate layer in western boundary current regions in all ocean basins. Consequently, we would expect that this study will be helpful for improving and developing EnKF-based ocean data assimilation systems. Minamide and Zhang (2017) noted that the AOEI has the advantage of being easily implemented with various EnKF-based systems, and this study serves as a good example for the usefulness of the AOEI. We are currently constructing high-resolution reanalysis datasets in the western North Pacific and Maritime Continent regions based on this

390   system, and plan to develop (near) real-time ensemble forecast systems.

Appendices

Appendix A: Cause of salinity degradation in the CTL run

        To investigate the cause of the salinity degradation in the CTL run, we calculate the salinity budget equation [Eq. (5)]

395   in the intermediate layer around the KE region (140°–160°E, 30°–40°N, 500–1000 m depth; white boxes in Fig. 3g, h). Figure A1a indicates that the salinity tendency term [the LHS term of Eq. (5)] is positive and corresponds to the salinity increase shown in left and middle panels of Fig. 3. The positive salinity tendency term is caused mainly by the diffusion term [the first

term on the RHS of Eq. (5)] and partly by the advection term [the second term on the RHS of Eq. (5)]. The diffusion term is dominated only by the vertical diffusion, and the horizontal diffusion makes almost no contribution (Fig. A1b). The advection
400 term consists of different components in different months during the experiment period (Fig. A1c): meridional advection in July, zonal and meridional advection in August, and zonal and vertical advection in September–December 2015. In contrast, the salinity analysis increment term [the last term on the RHS of Eq. (5)] has only a minor impact but plays a role in restoring the low salinity water. Therefore, the vertical diffusion is the main cause of the salinity degradation in the intermediate layer around the KE region.

405

Appendix B: Cause of salinity improvement in the AOEI run

To investigate the cause of the salinity improvement in the AOEI run relative to the CTL run, we calculate the salinity budget equation [Eq. (5)] difference between the AOEI and CTL runs (Fig. B1):

410 $$\Delta\left(\frac{\partial S}{\partial t}\right) = \Delta\{\nabla \cdot (\kappa \circ \nabla S)\} - \Delta\{v \cdot \nabla S\} + \Delta(S \ increment), \qquad (B1)$$

where Δ indicates the AOEI run minus the CTL run. The salinity tendency difference term [the LHS term of Eq. (B1)] indicates that the salinity structure is maintained in the AOEI run by suppressing the salinity increase throughout the experiment period (Fig. B1a). The diffusion and advection difference terms [the first and second terms on the RHS of Eq. (B1), respectively]
415 contribute almost equally to the salinity tendency difference term. The diffusion difference term is dominated by only the vertical diffusion difference, whereas the advection difference term is dominated by different components in different months: by the meridional advection difference in July; by all advection differences in August–September; and by vertical and partly zonal advection differences in October–December 2015. The reduction in the vertical diffusion is therefore the main cause of the improvement for low salinity water in the AOEI run relative to the CTL run.

420

Code and data availability

The source code for sbPOM version 1.0 and LETKF are available from https://zenodo.org/record/6482744 (last access: 24 November 2022, Jordi and Wang, 2012; Miyoshi and Yamane, 2007; Ohishi et al. 2022). The source code for COARE version 3.5 (Brodeau et al., 2017; Edson et al., 2013) was downloaded from https://github.com/brodeau/aerobulk (last
425 access: 24 November 2022).

We thank Dr. Kenshi Hibino for providing us with an earlier version of TE-Global, before the official release of the latest version (https://www.eorc.jaxa.jp/water/, last access: 24 November 2022). The observation datasets are: the surface drifter buoy data (https://www.aoml.noaa.gov/phod/gdp/hourly_data.php, last access: 24 November 2022, Elipot et al., 2016); the KEO buoy data (https://www.pmel.noaa.gov/ocs/, last access: 24 November 2022); ETOPO1
430 (https://www.ngdc.noaa.gov/mgg/global/, last access: 24 November 2022, Amante and Eakins, 2009); WOA18

(https://www.ncei.noaa.gov/access/world-ocean-atlas-2018/, last access: 24 November 2022; Locarnini et al., 2019; Zweng et al., 2019); the satellite SSTs from Himawari-8 (https://www.eorc.jaxa.jp/ptree/index.html, last access: 24 November 2022; Bessho et al., 2016; Kurihara et al., 2016) and GCOM-W (https://gportal.jaxa.jp/gpr/?lang=en, last access: 24 November 2022); the satellite-derived SSS from SMOS (http://www.esa.int/Applications/Observing_the_Earth/SMOS, last access: 24 November 2022) and SMAP version 4.3 (https://podaac.jpl.nasa.gov/, last access: 24 November 2022, Meissner et al., 2018); the satellite-derived SSHA and AVISO (Ducet et al., 2000) from CMEMS (https://marine.copernicus.eu/, last access: 24 November 2022); and in-situ temperatures and salinity from GTSPP (https://www.ncei.noaa.gov/products/global-temperature-and-salinity-profile-programme, last access: 24 November 2022, Sun et al., 2010) and AQC Argo version 1.2a (https://www.jamstec.go.jp/argo_research/dataset/aqc/index_dataset.html, last access: 24 November 2022). The global JRA55 atmosphere and SODA 3.7.2 ocean reanalysis datasets are from http://search.diasjp.net/en/dataset/JRA55 (last access: 24 November 2022, Kobayashi et al., 2015) and https://www.soda.umd.edu/soda3_readme.htm (last access: 24 November 2022, Carton et al., 2018), respectively.

Author contribution

SO developed the code for the ocean data assimilation system, conducted the sensitivity experiments, and analyzed the outputs. SO and TM prepared the paper with contributions from MK.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank two anonymous reviewers for constructive comments. Numerous comments from Drs. Nariaki Hirose, Takahiro Toyoda, Yosuke Fujii, and Norihisa Usui at the Meteorological Research Institute, Yoichi Ishikawa at JAMSTEC, Katsumi Takayama at IDEA Consultants, Inc., Naoki Hirose at Kyushu University, and participants in the ocean data assimilation summer school helped us to develop the presented system. This work used computational resources of the JAXA Supercomputer System Generation 2 and 3 (JSS2 and JSS3, respectively) and the supercomputer Fugaku provided by RIKEN through the HPCI Research Project (Project ID: hp210166, hp220167, ra000007).

Financial support

This work was supported by JST AIP Grant Number JPMJCR19U2, Japan; MEXT (JPMXP1020200305) as "Program for Promoting Researches on the Supercomputer Fugaku" (Large Ensemble Atmospheric and Environmental Prediction for Disaster Prevention and Mitigation); the COE research grant in computational science from Hyogo Prefecture and Kobe City through Foundation for Computational Science; JST, SICORP Grant Number JPMJSC1804, Japan; JSPS KAKENHI Grant Number JP19H05605; the Japan Aerospace Exploration Agency (JX-PSPC-452680, -500973, -509736, -513414, -519799, and

465 -527843); JST, CREST Grant Number JPMJCR20F2, Japan; Cabinet Office, Government of Japan, Moonshot R&D Program for Agriculture, Forestry and Fisheries (funding agency: Bio-oriented Technology Research Advancement Institution) No. JPJ009237; RIKEN Pioneering Project "Prediction for Science"; JST, CREST Grant Number : JPMJSA2109.

Review statement: This paper was edited by Christopher Horvat and reviewed by two anonymous reviewers.

470

References

Abe, H. and Ebuchi, N.: Evaluation of sea-surface salinity observed by Aquarius, J. Geophys. Res. Ocean., 119, 8109–8121, doi:10.1002/2014JC010094, 2014.

Amante, C. and Eakins, B. W.: ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis., 2009.

475 Balmaseda, M. A., Hernandez, F., Storto, A., Palmer, M. D., Alves, O., Shi, L., Smith, G. C., Toyoda, T., Valdivieso, M., Barnier, B., Behringer, D., Boyer, T., Chang, Y. S., Chepurin, G. A., Ferry, N., Forget, G., Fujii, Y., Good, S., Guinehut, S., Haines, K., Ishikawa, Y., Keeley, S., Köhl, A., Lee, T., Martin, M. J., Masina, S., Masuda, S., Meyssignac, B., Mogensen, K., Parent, L., Peterson, K. A., Tang, Y. M., Yin, Y., Vernieres, G., Wang, X., Waters, J., Wedd, R., Wang, O., Xue, Y., Chevallier, M., Lemieux, J. F., Dupont, F., Kuragano, T., Kamachi, M., Awaji, T., Caltabiano, A., Wilmer-Becker, K. and

480 Gaillard, F.: The ocean reanalyses intercomparison project (ORA-IP), J. Oper. Oceanogr., 8(0), s80–s97, doi:10.1080/1755876X.2015.1022329, 2015.

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., Kumagai, Y., Miyakawa, T., Murata, H., Ohno, T., Okuyama, A., Oyama, R., Sasaki, Y., Shimazu, Y., Shimoji, K., Sumida, Y., Suzuki, M., Taniguchi, H., Tsuchiyama, H., Uesawa, D., Yokota, H. and Yoshida, R.: An introduction to Himawari-8/9 – Japan's new-generation geostationary

485 meteorological satellites, J. Meteorol. Soc. Japan, 94(2), 151–183, doi:10.2151/jmsj.2016-009, 2016.

Bloom, S. C., Takacs, L. L., da Silva, A. M. and Ledvina, D.: Data assimilation using incremental analysis updates, Mon. Weather Rev., 124(6), 1256–1271, doi:10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2, 1996.

Brodeau, L., Barnier, B., Gulev, S. K. and Woods, C.: Climatologically significant effects of some approximations in the bulk parameterizations of turbulent air–sea fluxes, J. Phys. Oceanogr., 47(1), 5–28, doi:10.1175/JPO-D-16-0169.1, 2017.

490 Carton, J. A., Chepurin, G. A. and Chen, L.: SODA3: A new ocean climate reanalysis, J. Clim., 31(17), 6967–6983, doi:10.1175/JCLI-D-17-0149.1, 2018.

Chakraborty, A., Sharma, R., Kumar, R. and Basu, S.: Joint assimilation of Aquarius-derived sea surface salinity and AVHRR-derived sea surface temperature in an ocean general circulation model using SEEK filter: Implication for mixed layer depth and barrier layer thickness, J. Geophys. Res. Ocean., 120(10), 6927–6942, doi:10.1002/2015JC010934, 2015.

495 Davis, P. E. D., Lique, C., Johnson, H. L. and Guthrie, J. D.: Competing effects of elevated vertical mixing and increased freshwater input on the stratification and sea ice cover in a changing Arctic Ocean, J. Phys. Oceanogr., 46(5), 1531–1553, doi:10.1175/JPO-D-15-0174.1, 2016.

Desroziers, G., Berre, L., Chapnik, B. and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Q. J. R. Meteorol. Soc., 131(613), 3385–3396, doi:10.1256/qj.05.108, 2005.

500 Ducet, N., Le Traon, P. Y. and Reverdin, G.: Global high-resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and -2, J. Geophys. Res., 105(C8), 19477–19498, doi:10.1029/2000JC900063, 2000.

Edson, J. B., Jampana, V., Weller, R. A., Bigorre, S. P., Plueddemann, A. J., Fairall, C. W., Miller, S. D., Mahrt, L., Vickers, D. and Hersbach, H.: On the exchange of momentum over the open ocean, J. Phys. Oceanogr., 43(8), 1589–1610, doi:10.1175/JPO-D-12-0173.1, 2013.

505 Elipot, S., Lumpkin, R., Perez, R. C., Lilly, J. M., Early, J. J. and Sykulski, A. M.: A global surface drifter data set at hourly resolution, J. Geophys. Res. Ocean., 121(5), 2937–2966, doi:10.1002/2016JC011716, 2016.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10143–10162, 1994.

Evensen, G.: The Ensemble Kalman Filter: Theoretical formulation and practical implementation, Ocean Dyn., 53, 343–367,
510 doi:10.1007/s10236-003-0036-9, 2003.

Hackert, E., Ballabrera-Poy, J., Busalacchi, A. J., Zhang, R. H. and Murtugudde, R.: Impact of sea surface salinity assimilation on coupled forecasts in the tropical Pacific, J. Geophys. Res. Ocean., 116, C05009, doi:10.1029/2010JC006708, 2011.

Hunt, B. R., Kostelich, E. J. and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble
515 transform Kalman filter, Phys. D, 230, 112–126, doi:10.1016/j.physd.2006.11.008, 2007.

Isoguchi, O., Kawamura, H. and Oka, E.: Quasi-stationary jets transporting surface warm waters across the transition zone between the subtropical and the subarctic gyres in the North Pacific, J. Geophys. Res. Ocean., 111(10), C10003, doi:10.1029/2005JC003402, 2006.

Jerlov, N. G.: Marine Optics., 1976.

520 Jordi, A. and Wang, D. P.: sbPOM: A parallel implementation of Princenton Ocean Model, Environ. Model. Softw., 38, 59–61, doi:10.1016/j.envsoft.2012.05.013, 2012.

Kida, S., Mitsudera, H., Aoki, S., Guo, X., Ito, S., Kobashi, F., Komori, N., Kubokawa, A., Miyama, T., Morie, R., Nakamura, H., Nakamura, T., Nakano, H., Nishigaki, H., Nonaka, M., Sasaki, H., Sasaki, Y. N., Suga, T., Sugimoto, S., Taguchi, B., Takaya, K., Tozuka, T., Tsujino, H. and Usui, N.: Oceanic fronts and jets around Japan: a review, J. Oceanogr.,
525 71, 469–497, doi:10.1007/s10872-015-0283-7, 2015.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K. and Takahashi, K.: The JRA-55 reanalysis: General specifications and basic characteristics, J. Meteorol. Soc. Japan, 93(1), 5–48, doi:10.2151/jmsj.2015-001, 2015.

Kotsuki, S., Ota, Y. and Miyoshi, T.: Adaptive covariance relaxation methods for ensemble data assimilation: experiments in
530 the real atmosphere, Q. J. R. Meteorol. Soc., 143(705), 2001–2015, doi:10.1002/qj.3060, 2017.

Kurihara, Y., Murakami, H. and Kachi, M.: Sea surface temperature from the new Japanese geostationary meteorological Himawari-8 satellite, Geophys. Res. Lett., 43(3), 1234–1240, doi:10.1002/2015GL067159, 2016.

Locarnini, R. A., Mishonov, A. V., Baranova, O. K., Boyer, T. P., Zweng, M. M., Garcia, H. E., Reagan, J. R., Seidov, D., Weathers, K. W., Paver, C. R. and Smolyar, I. V.: World Ocean Atlas 2018, Volume 1: Temperature. A. Mishonov,

535   Technical Editor., NOAA Atlas NESDIS, 81(July), 52, 2019.

Martin, M. J., Balmaseda, M., Bertino, L., Brasseur, P., Brassington, G., Cummings, J., Fujii, Y., Lea, D. J., Lellouche, J. M., Mogensen, K., Oke, P. R., Smith, G. C., Testut, C. E., Waagbø, G. A., Waters, J. and Weaver, A. T.: Status and future of data assimilation in operational oceanography, J. Oper. Oceanogr., 8(0), s28–s48, doi:10.1080/1755876X.2015.1022055, 2015.

540   Meissner, T., Wentz, F. J. and Vine, D. M. Le: The salinity retrieval algorithms for the NASA Aquarius version 5 and SMAP version 3 releases, Remote Sens., 10(7), 1121, doi:10.3390/rs10071121, 2018.

Mellor, G. L., Ezer, T. and Oey, L.-Y.: The pressure gradient conundrum of sigma coordinate ocean models, J. Atmos. Ocean. Technol., 11(4), 1126–1134, doi:10.1175/1520-0426(1994)011<1126:TPGCOS>2.0.CO;2, 1994.

Minamide, M. and Zhang, F.: Adaptive observation error inflation for assimilating all-sky satellite radiance, Mon. Weather

545   Rev., 145(3), 1063–1081, doi:10.1175/MWR-D-16-0257.1, 2017.

Miyazawa, Y., Zhang, R., Guo, X., Tamura, H., Ambe, D., Lee, J. S., Okuno, A., Yoshinari, H., Setou, T. and Komatsu, K.: Water mass variability in the western North Pacific detected in a 15-year eddy resolving ocean reanalysis, J. Oceanogr., 65(6), 737–756, doi:10.1007/s10872-009-0063-3, 2009.

Miyazawa, Y., Miyama, T., Varlamov, S. M., Guo, X. and Waseda, T.: Open and coastal seas interactions south of Japan

550   represented by an ensemble Kalman filter, Ocean Dyn., 62(4), 645–659, doi:10.1007/s10236-011-0516-2, 2012.

Miyazawa, Y., Varlamov, S. M., Miyama, T., Guo, X., Hihara, T., Kiyomatsu, K., Kachi, M., Kurihara, Y. and Murakami, H.: 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(6), 713–728, doi:10.1007/s10236-017-1056-1, 2017.

555   Miyoshi, T. and Yamane, S.: Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution, Mon. Weather Rev., 135(11), 3841–3861, doi:10.1175/2007MWR1873.1, 2007.

Miyoshi, T., Yamane, S. and Enomoto, T.: Localizing the error covariance by physical distances within a local ensemble transform Kalman filter (LETKF), SOLA, 3, 89–92, doi:10.2151/sola.2007-023, 2007.

Ohishi, S., Hihara, T., Aiki, H., Ishizaka, J., Miyazawa, Y., Kachi, M. and Miyoshi, T.: An ensemble Kalman filter system

560   with the Stony Brook Parallel Ocean Model v1.0, Geosci. Model Dev., 15(22), 8395–8410, doi:10.5194/gmd-15-8395-2022, 2022.

Pacanowski, R. C. and Philander, S. G. H.: Parameterization of vertical mixing in numerical models of tropical oceans, J. Phys. Oceanogr., 11, 1443–1451, doi:10.1175/1520-0485(1981)011<1443:POVMIN>2.0.CO;2, 1981.

Paulson, C. A. and Simpson, J. J.: Irradiance measurements in the upper ocean, J. Phys. Oceanogr., 7, 952–956,

565   doi:10.1175/1520-0485(1977)007<0952:IMITUO>2.0.CO;2, 1977.

Penny, S. G., Kalnay, E., Carton, J. A., Hunt, B. R., Ide, K., Miyoshi, T. and Chepurin, G. A.: The local ensemble transform
Kalman filter and the running-in-place algorithm applied to a global ocean general circulation model, Nonlinear Process.
Geophys., 20(6), 1031–1046, doi:10.5194/npg-20-1031-2013, 2013.

Sakov, P., Counillon, F., Bertino, L., Lisæter, K. A., Oke, P. R. and Korablev, A.: TOPAZ4: an ocean-sea ice data

570   assimilation system for the North Atlantic and Arctic, Ocean Sci., 8(4), 633–656, doi:10.5194/os-8-633-2012, 2012.

Shibuya, R., Sato, K., Tomikawa, Y., Tsutsumi, M. and Sato, T.: A study of multiple tropopause structures caused by inertia-
gravity waves in the antarctic, J. Atmos. Sci., 72(5), 2109–2130, doi:10.1175/JAS-D-14-0228.1, 2015.

Sun, C., Thresher, A., Keeley, R., Hall, N., Hamilton, M., Chinn, P., A.Tran, Goni, G., Villeon, L. P. de la, Carval, T.,
Cowen, L., Manzella, G., Gopalakrishna, V., Guerrero, R., Reseghetti, F., Kanno, Y., Klein, B., Rickard, L., Baldoni, A.,

575   Lin, S., Ji, F. and Nagaya, Y.: The data management system for the global temperature and salinity profile programme, in
Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society, pp. 931–938, European Space
Agency., 2010.

Talley, L. D.: Distribution and formation of North Pacific intermediate water, J. Phys. Oceanogr., 23(3), 517–537,
doi:10.1175/1520-0485(1993)023<0517:DAFONP>2.0.CO;2, 1993.

580   Toyoda, T., Fujii, Y., Kuragano, T., Matthews, J. P., Abe, H., Ebuchi, N., Usui, N., Ogawa, K. and Kamachi, M.:
Improvements to a global ocean data assimilation system through the incorporation of Aquarius surface salinity data, Q. J. R.
Meteorol. Soc., 141(692), 2750–2759, doi:10.1002/qj.2561, 2015.

Usui, N., Tsujino, H., Fujii, Y. and Kamachi, M.: Short-range prediction experiments of the Kuroshio path variabilities south
of Japan, Ocean Dyn., 56(5–6), 607–623, doi:10.1007/s10236-006-0084-z, 2006.

585   Xu, F. H. and Oey, L. Y.: State analysis using the Local Ensemble Transform Kalman Filter (LETKF) and the three-layer
circulation structure of the Luzon Strait and the South China Sea, Ocean Dyn., 64(6), 905–923, doi:10.1007/s10236-014-
0720-y, 2014.

Yasuda, I.: The origin of the North Pacific intermediate water, J. Geophys. Res. Ocean., 102(C1), 893–909,
doi:10.1029/96JC02938, 1997.

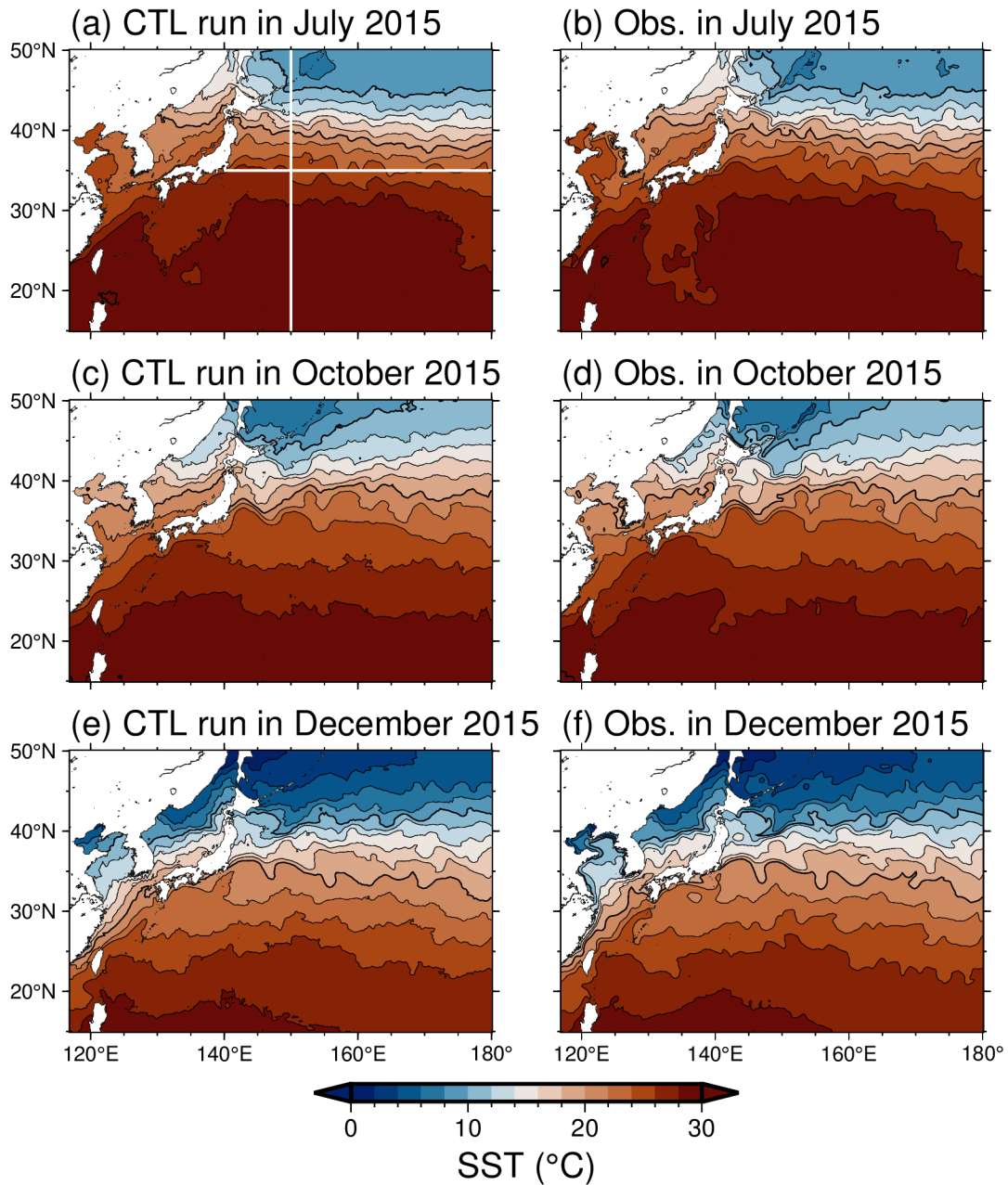
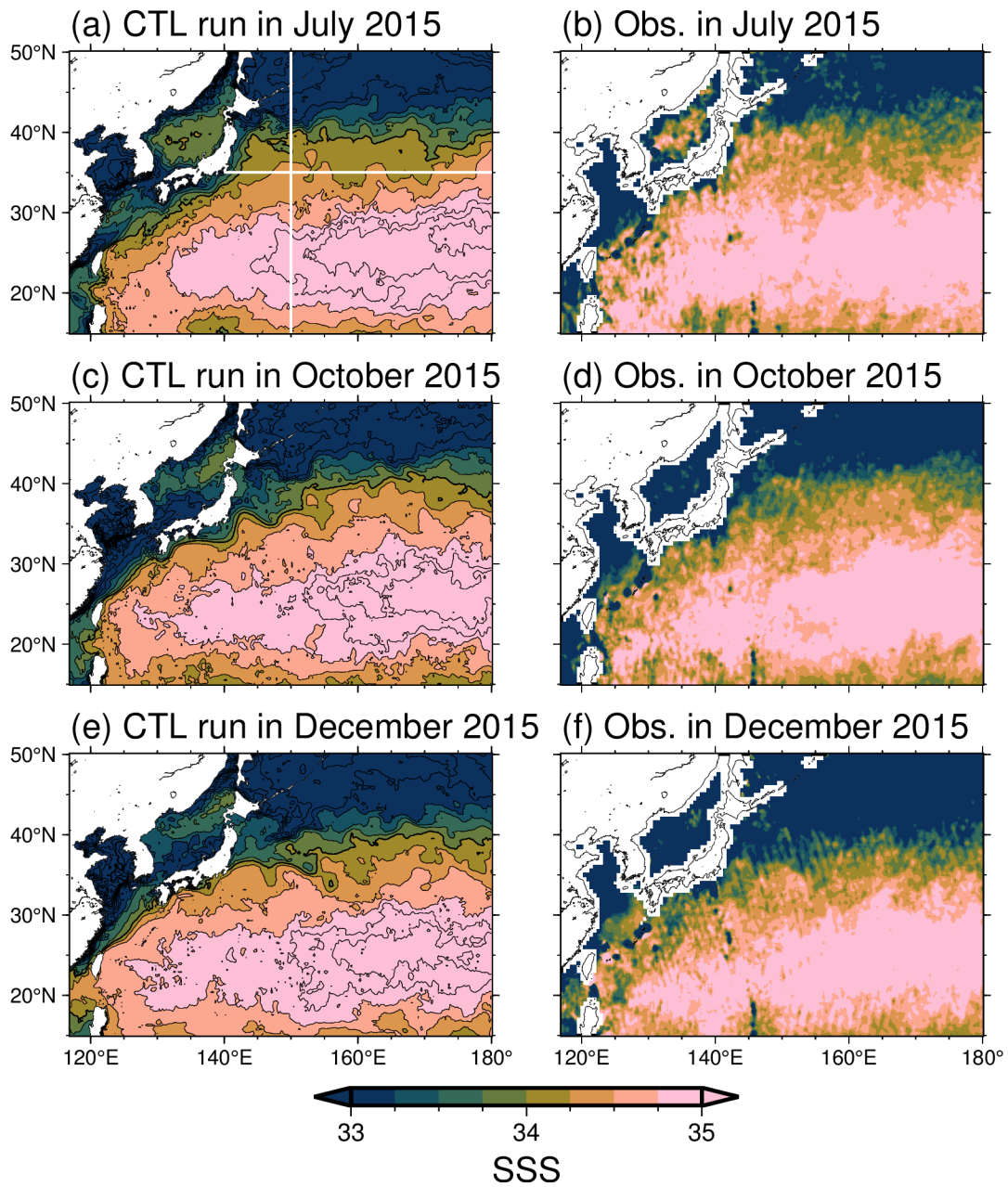
590   Yin, Y., Alves, O. and Oke, P. R.: An ensemble ocean data assimilation system for seasonal prediction, Mon. Weather Rev.,
139(3), 786–808, doi:10.1175/2010MWR3419.1, 2011.

Zhang, F., Davis, C. A., Kaplan, M. L. and Koch, S. E.: Wavelet analysis and the governing dynamics of a large-amplitude
mesoscale gravity-wave event along the east coast of the United States, Q. J. R. Meteorol. Soc., 127(577), 2209–2245,
doi:10.1002/qj.49712757702, 2001.

595   Zhang, F., Snyder, C. and Sun, J.: Impacts of initial estimate and observation availability on convective-scale data
assimilation with an ensemble Kalman filter, Mon. Weather Rev., 132(5), 1238–1253, doi:10.1175/1520-
0493(2004)132<1238:IOIEAO>2.0.CO;2, 2004.

Zhang, F., Minamide, M. and Clothiaux, E. E.: Potential impacts of assimilating all-sky infrared satellite radiances from GOES-R on convection-permitting analysis and prediction of tropical cyclones, Geophys. Res. Lett., 43(6), 2954–2963, doi:10.1002/2016GL068468, 2016.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K. and Mayer, M.: The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment, Ocean Sci., 15(3), 779–808, doi:10.5194/os-15-779-2019, 2019.

Zweng, M. M., Reagan, J. R., Seidov, D., Boyer, T. P., Antonov, J. I., Locarnini, R. A., Garcia, H. E., Mishonov, A. V., Baranova, O. K., Weathers, K. W., Paver, C. R. and Smolyar, I. V.: World Ocean Atlas 2018, Volume 2: Salinity. A. Mishonov, Technical Editor, NOAA Atlas NESDIS, 82(July), 50, 2019.

610 **Figure 1: Monthly-mean SSTs for (a) July, (c) October, and (e) December 2015 in the CTL run. (b), (d), and (f) Same as (a), (c), and (e), but for assimilated satellite SSTs. Thin (thick) black contour intervals are 2 (10) °C. White lines in (a) indicate 150°E (35°N) used for the zonal (meridional) sections in Fig. 3.**

615 **Figure 2: Same as Fig.1 but for SSS. Thin (thick) black contour intervals are 0.25 (2). In (b), (d), and (f), contour intervals are not shown because the satellite observations are noisy.**

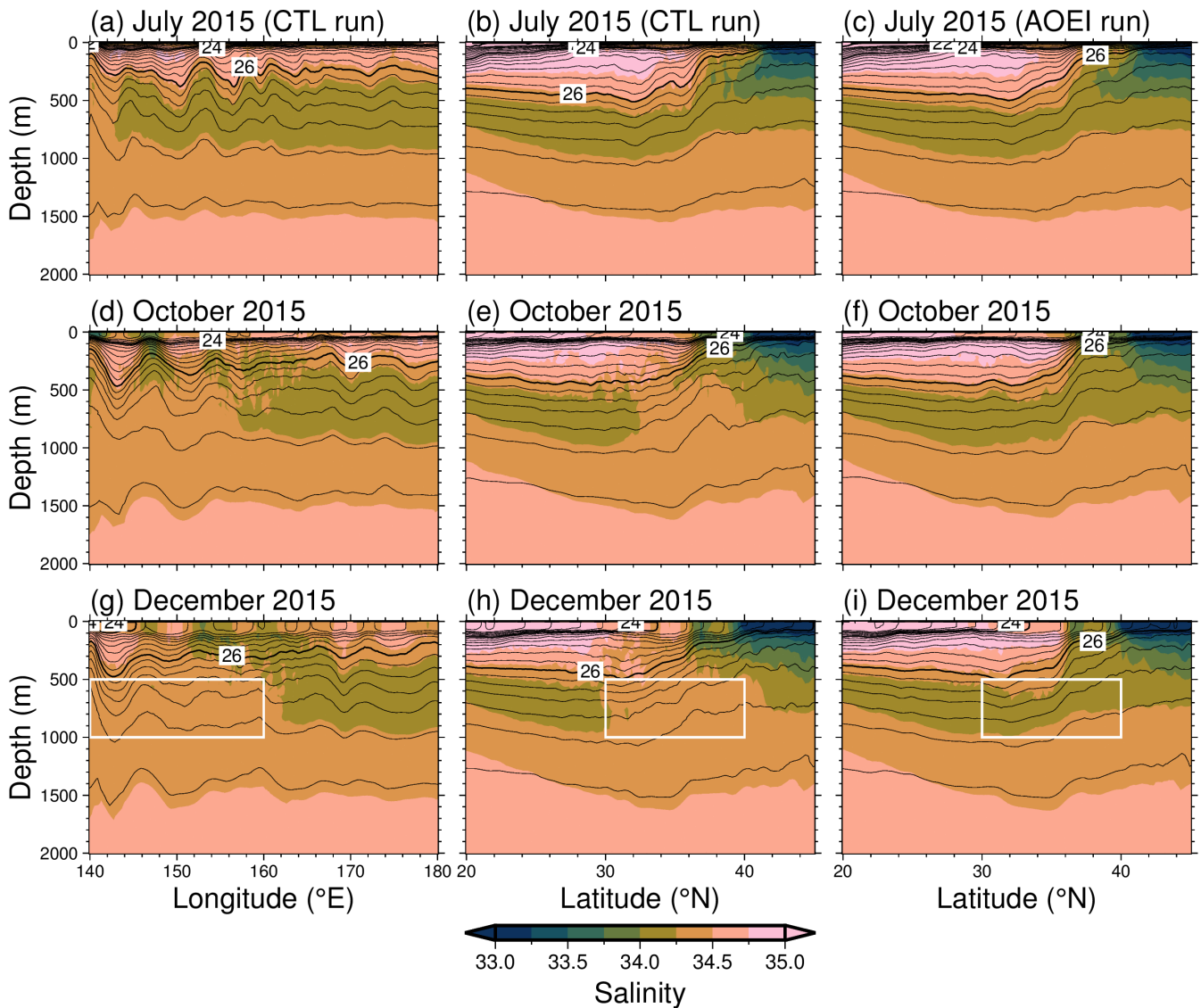



Figure 3: The zonal section of monthly-mean salinity (color) and potential density (contour) along 35°N in (a) July, (d) October, and (g) December 2015 in the CTL run. (b), (e), and (h) Same as (a), (d), and (g), but for the meridional section along 150°E. (c), (f), (i) Same as (b), (e), (h) but for AOEI run. Thin (thick) contour intervals are 0.25 (2) kg m⁻³. The white box in (g) [(h) and (i)] encloses a longitude (latitude)-depth section of 140°–160°E (30°–40°N) and 500–1000 m depth.

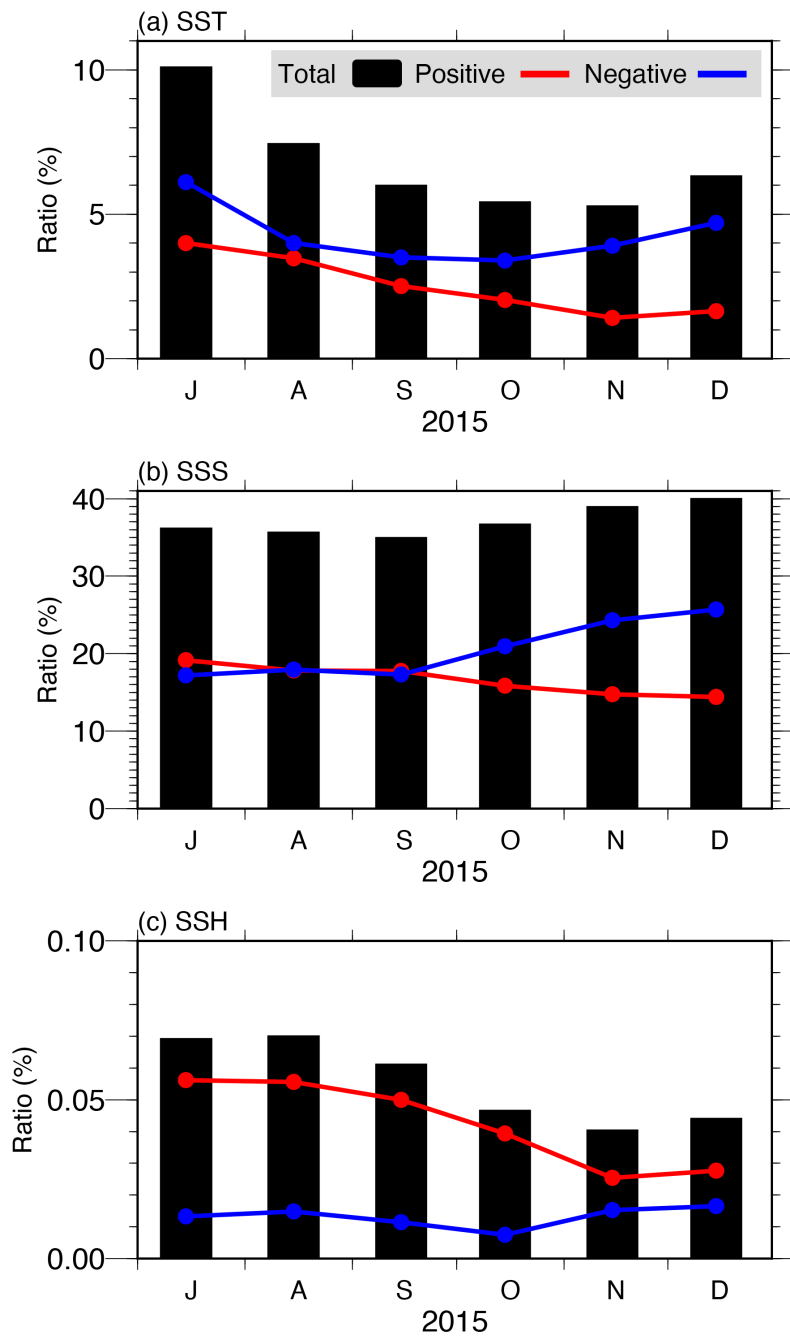



Figure 4: Monthly-mean ratio of the area where the AOEI is applied to the system domain in the (a) SST, (b) SSS, and (c) SSH fields in the AOEI run (black bars). Red (blue) lines indicate when the innovation is positive (negative).

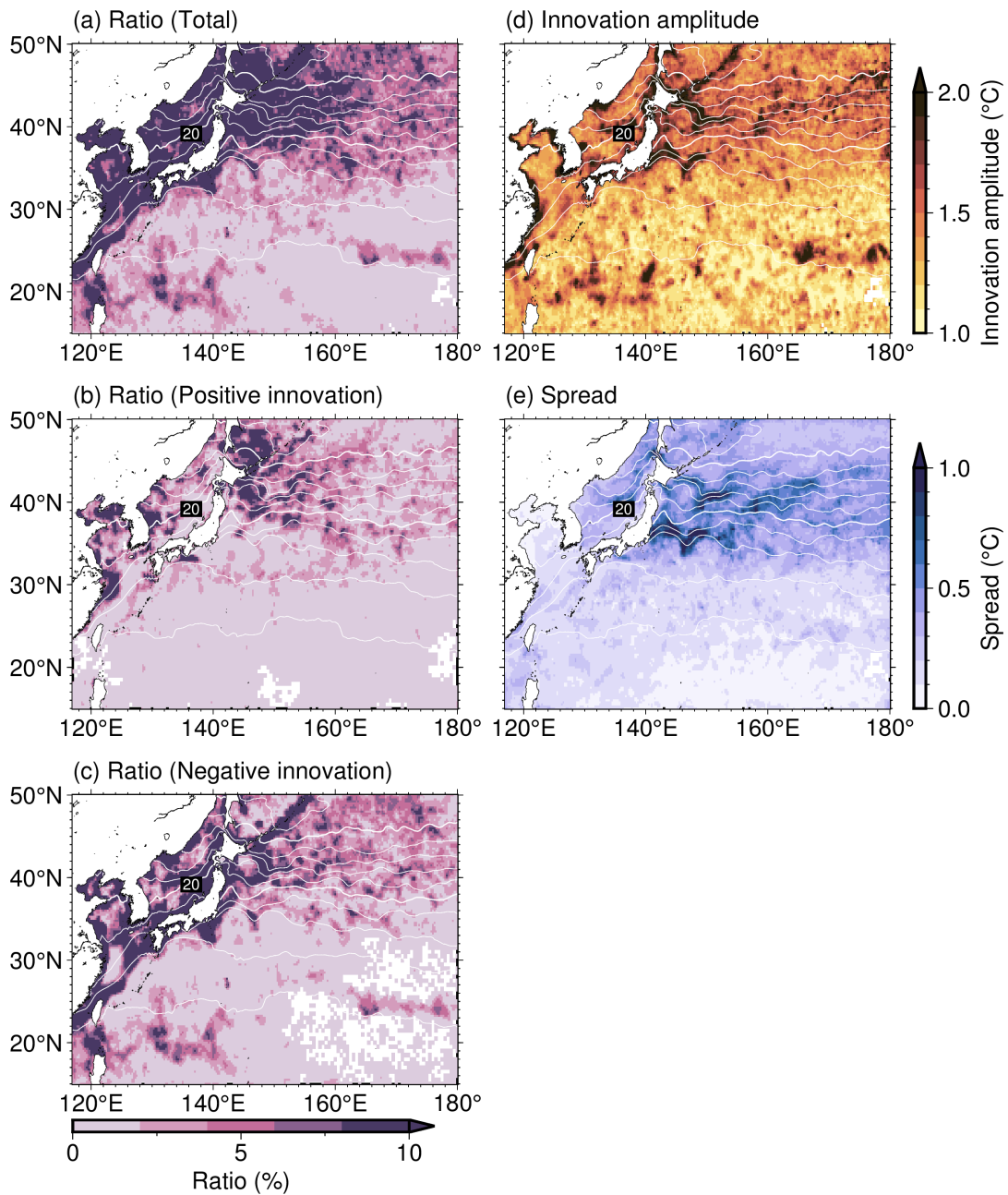



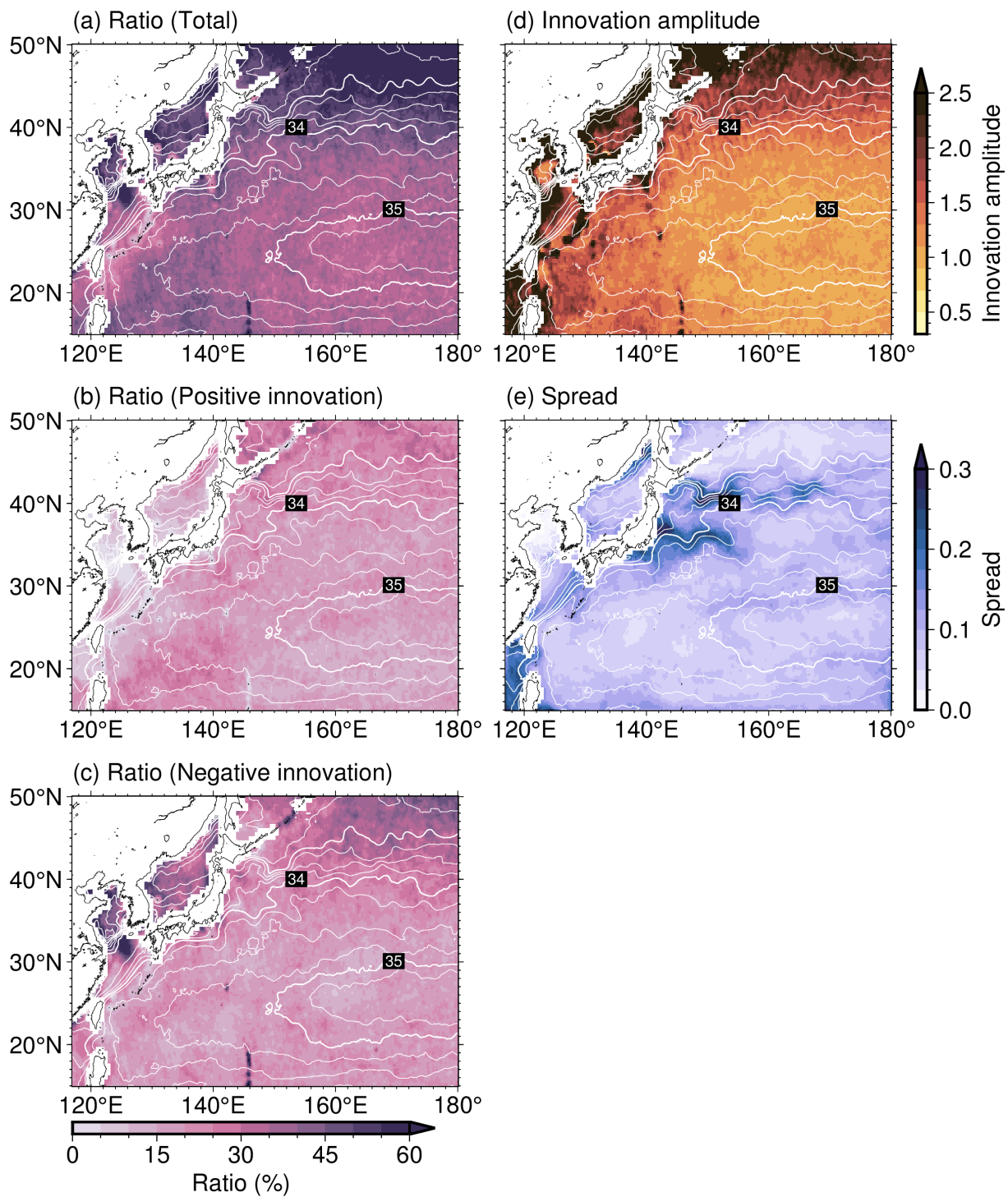
Figure 5: (a) Ratio of the period when the AOEI is applied to the SST compared with the whole experiment period in the AOEI run. (b) and (c) Same as (a) but for when the innovation is positive and negative, respectively. (d) and (e) innovation amplitude and ensemble spread averaged during the period when the AOEI is applied, respectively. White contours indicate SST averaged over the whole period. Thin (thick) contour intervals are 2 (10) °C. White areas indicate no AOEI application.

635

Figure 6: Same as Fig. 5 but for SSS. Thin (thick) contour intervals are 0.25 (1).

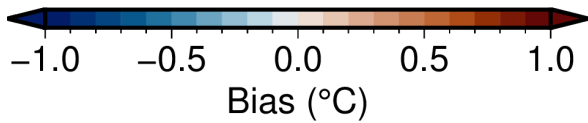
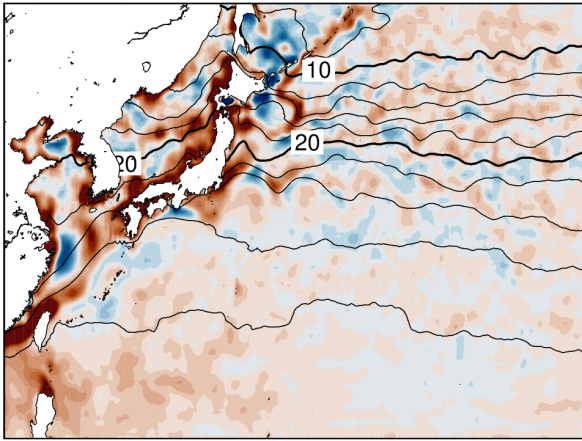
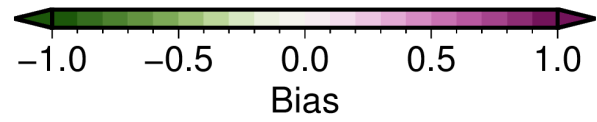
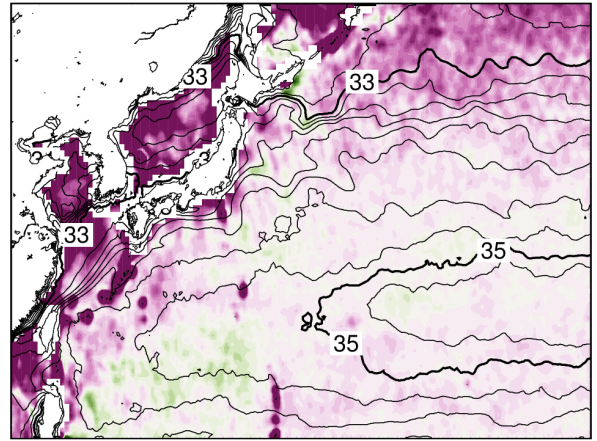



Figure 7: (a) SST and (b) SSS forecast biases (color) and averages (contour) over the whole experiment period in the
AOEI run. Thin (thick) contour intervals are 2 (10) °C in (a) and 0.25 (2) in (b).

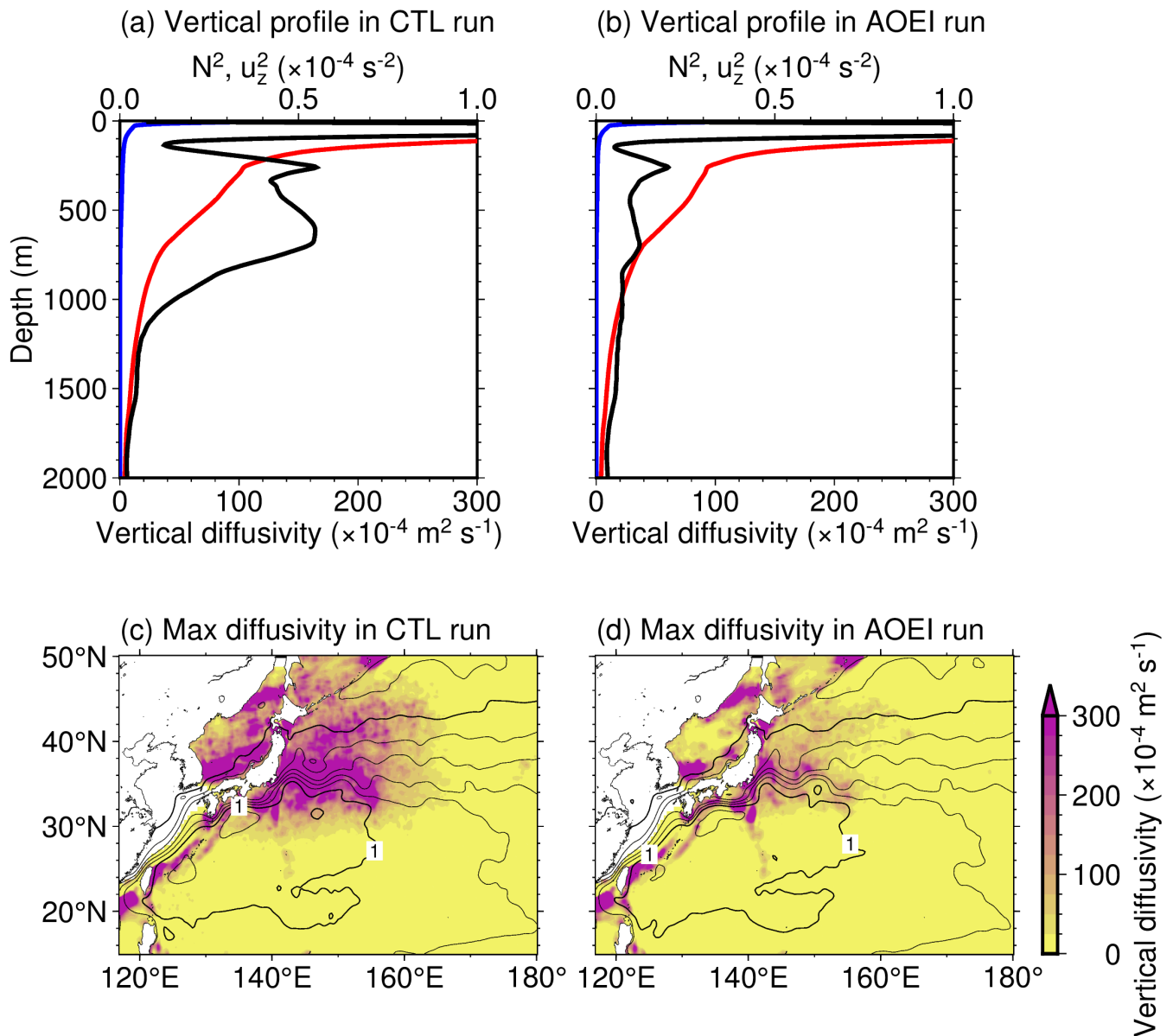



Figure 8: Vertical diffusivity (black), and squared buoyancy frequency (red) and shear (blue) averaged over the KE region (140°–160°E, 30°–40°N) for the whole experiment period in the (a) CTL and (b) AOEI runs. Maxima of the averaged vertical diffusivity for the whole experiment period within 300–1000 m depth in the (c) CTL and (d) AOEI runs. Black contours show SSH averaged over the whole period. Thin (thick) contour intervals are 0.2 (1) m.

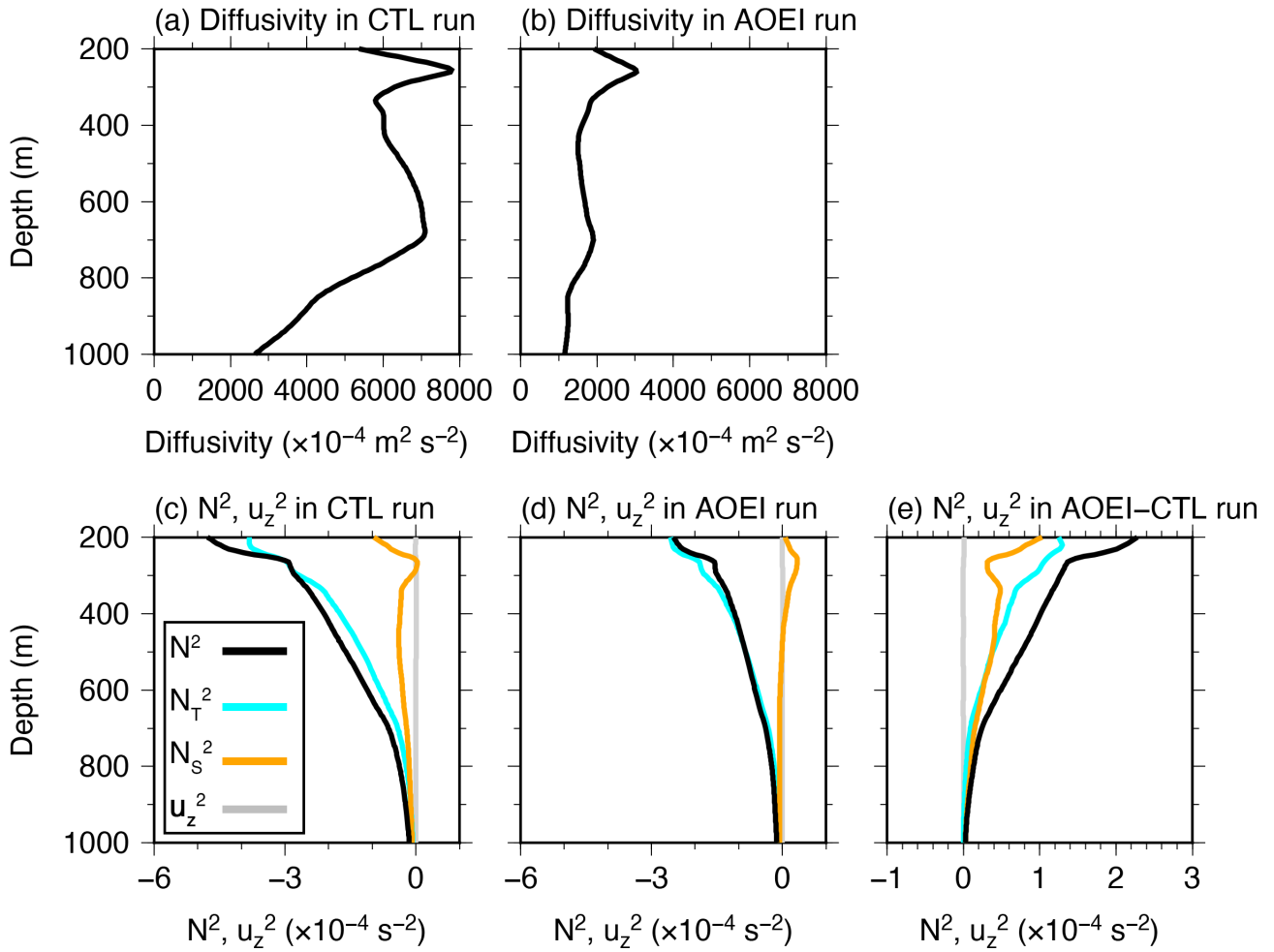
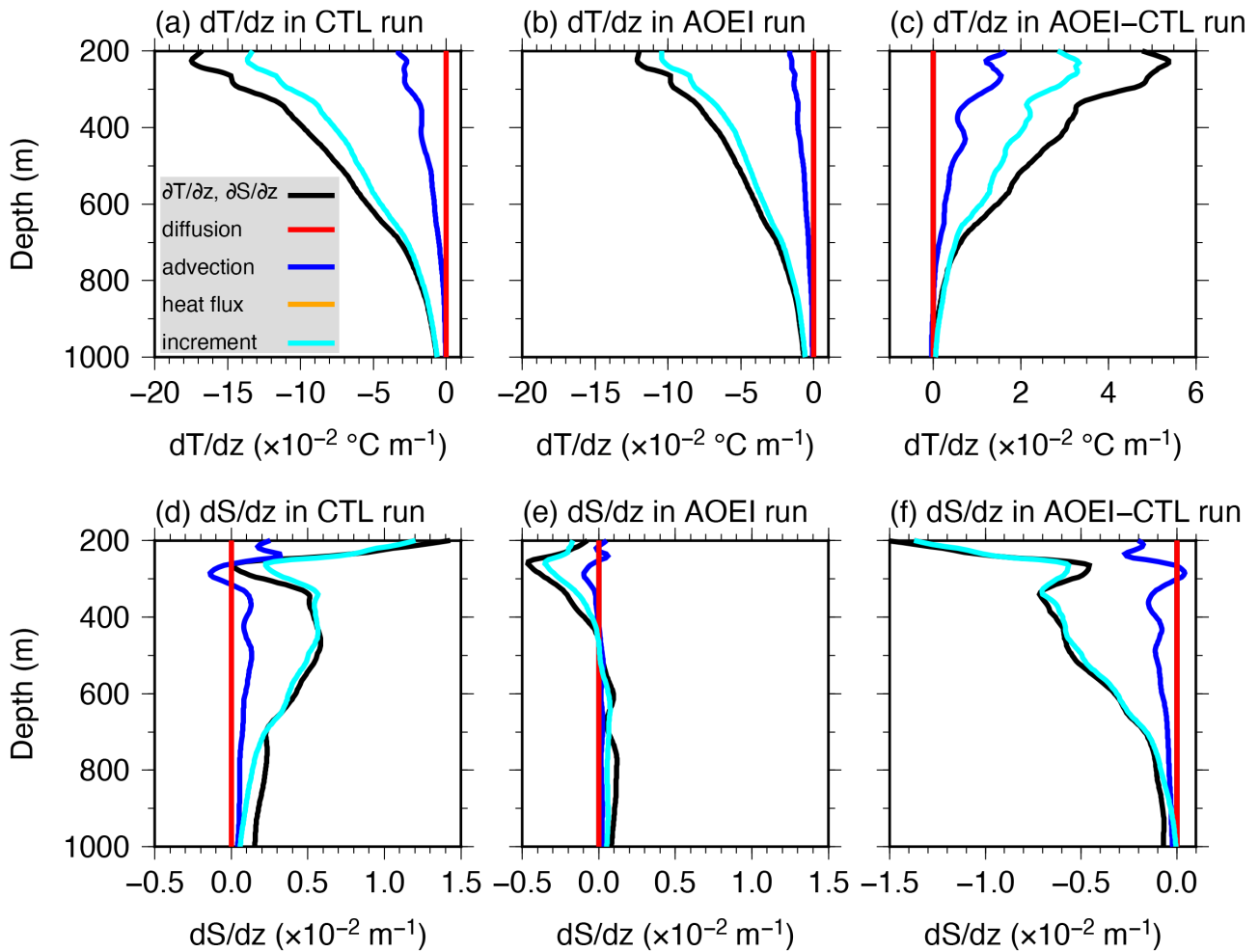



Figure 9: Total vertical diffusivity tendency during the positive vertical diffusivity tendency period averaged over the KE region (140°–160°E, 30°–40°N) in the (a) CTL and (b) AOEI runs. (c) and (d) Same as (a) but for the squared buoyancy frequency (black) and shear (gray) tendency. (e) Same as (c) and (d), but for the AOEI minus CTL run. In (c)–(e), cyan (orange) lines indicate contributions from $\partial N_T^2/\partial t$ ($\partial N_S^2/\partial t$).

650

655 **Figure 10: (a)–(c) Same as Fig. 9 but for each term of the temperature stratification tendency equation [Eq. (13)]: Temperature gradient tendency term (the LHS term; black), temperature diffusion gradient term (the first term on the RHS; red), temperature advection gradient term (the second term on the RHS; blue), shortwave penetration gradient term (the third term on the RHS; orange), and temperature increment gradient term (the last term on the RHS; cyan). (d)–(f) Same as (a)–(c) but for the salinity stratification tendency equation [Eq. (14)]. In (a)-(c), we note**
660 **that the shortwave penetration gradient term is almost zero and overlaps with the temperature diffusion gradient term.**

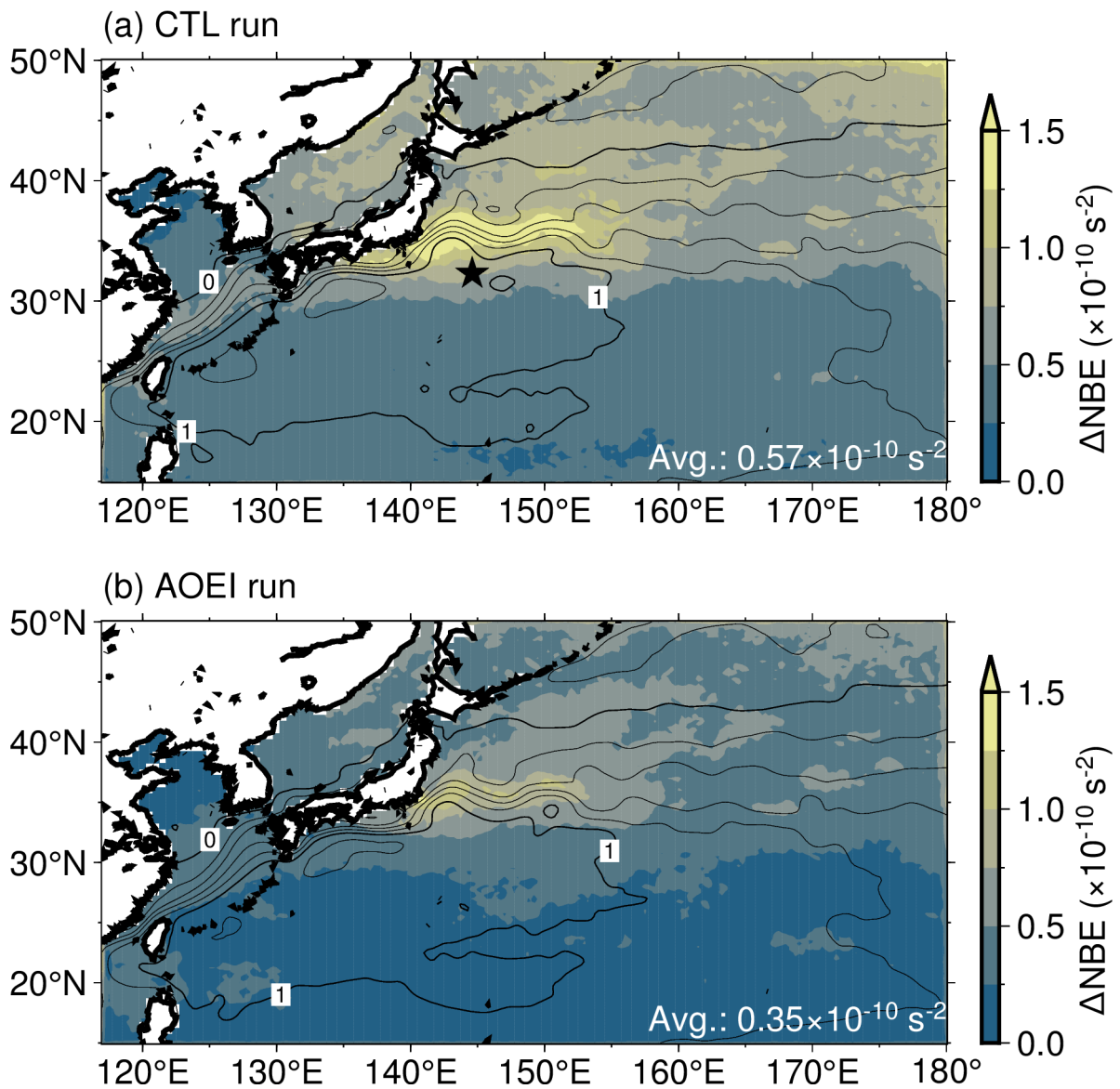



Figure 11: Δ*NBE* (color) and SSH (contour) averaged over the whole experiment period for the (a) CTL and (b) AOEI runs. Thin (thick) contour intervals are 0.2 (1) m. Spatiotemporally averaged Δ*NBE* over the whole period and domain is shown in the lower right corners. The black star in (a) denotes the position of the KEO buoy.

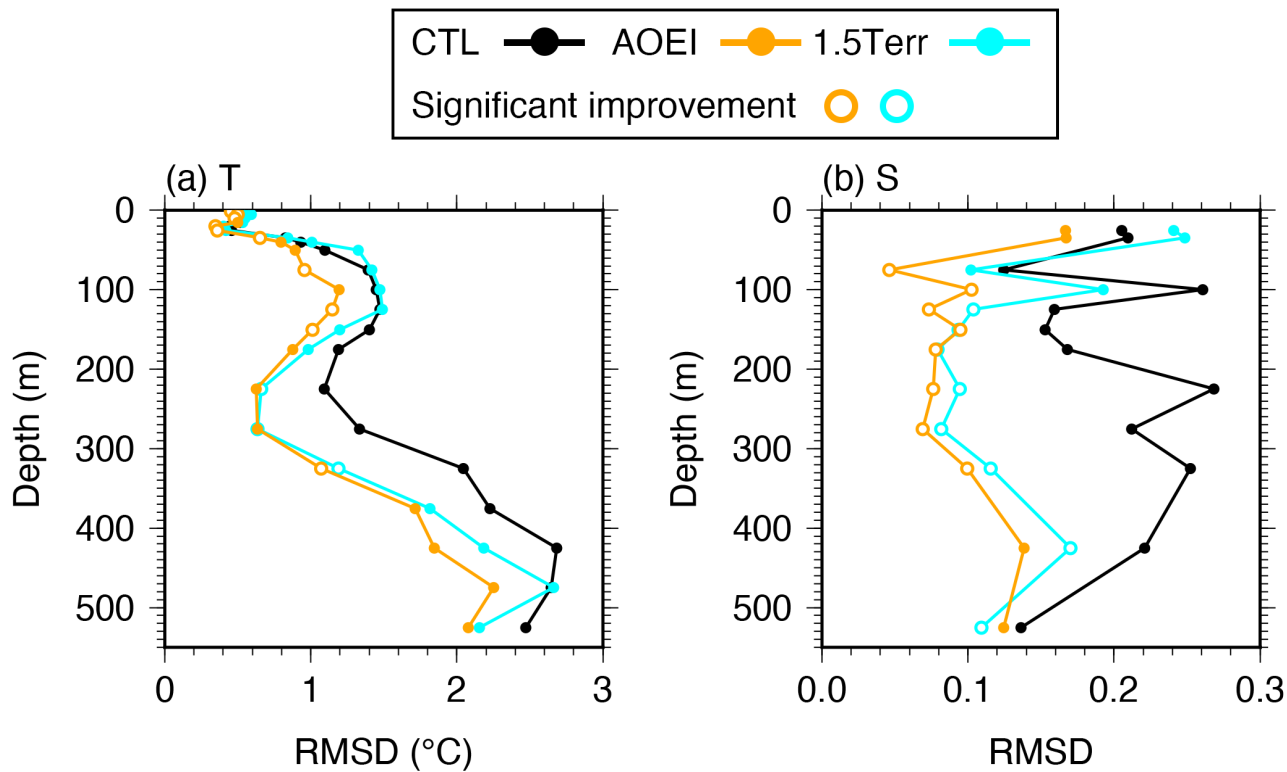



Figure 12: (a) Temperature and (b) salinity RMSDs averaged over the whole experiment period at the KEO buoy in the CTL (black), AOEI (orange), and 1.5Terr (cyan) runs. Open circles indicate significant improvement in the AOEI and 1.5Terr runs relative to the CTL run.

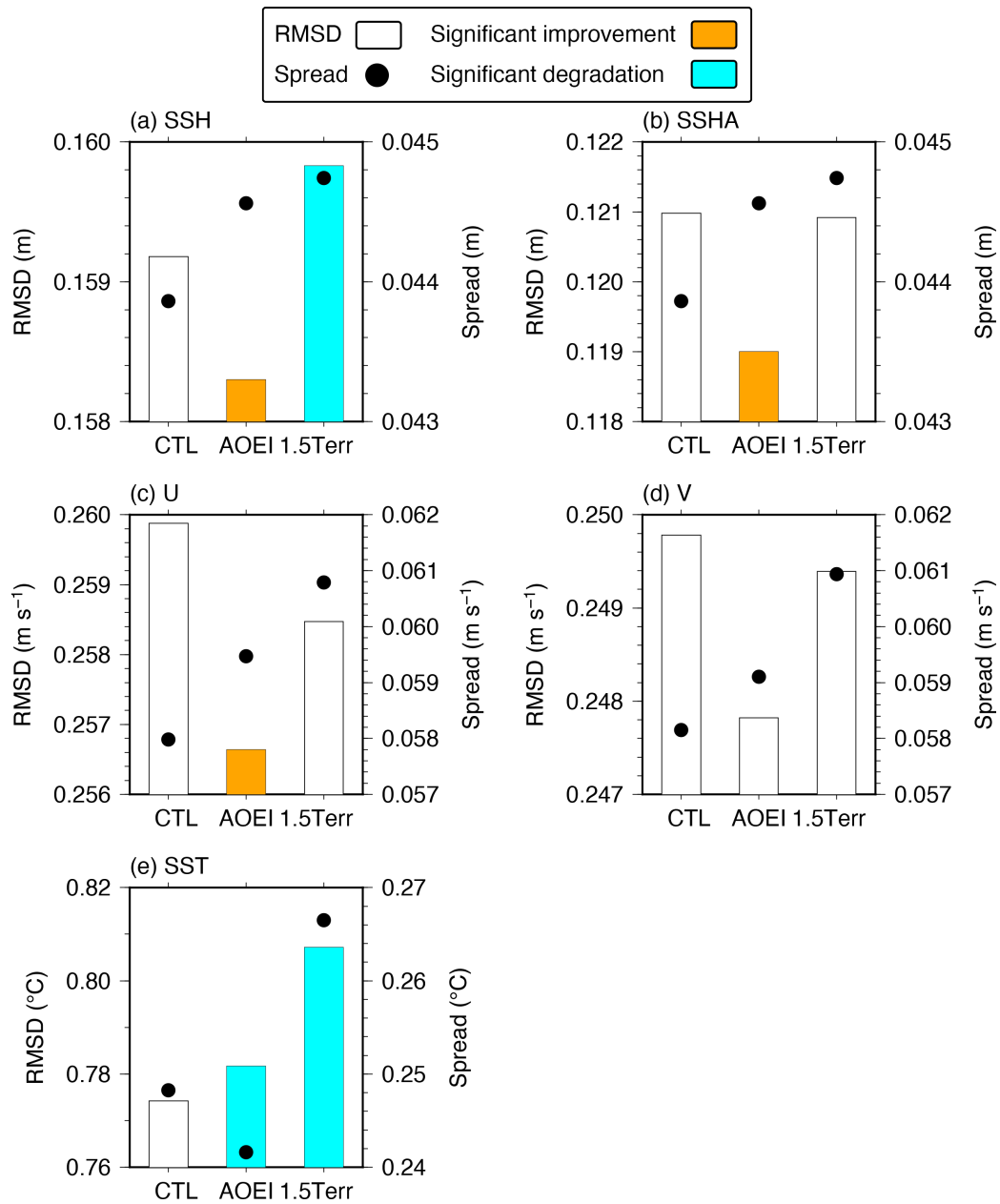



Figure 13: RMSDs for (a) SSH and (b) SSHA relative to the AVISO, surface (c) zonal and (d) meridional velocity relative to the drifter buoys, and (e) SSTs relative to the Himawari-8 averaged over the whole domain and period. Black dots indicate the ensemble spread in the observation space. We note that the ranges of the vertical axis are different between the RMSDs and ensemble spreads.

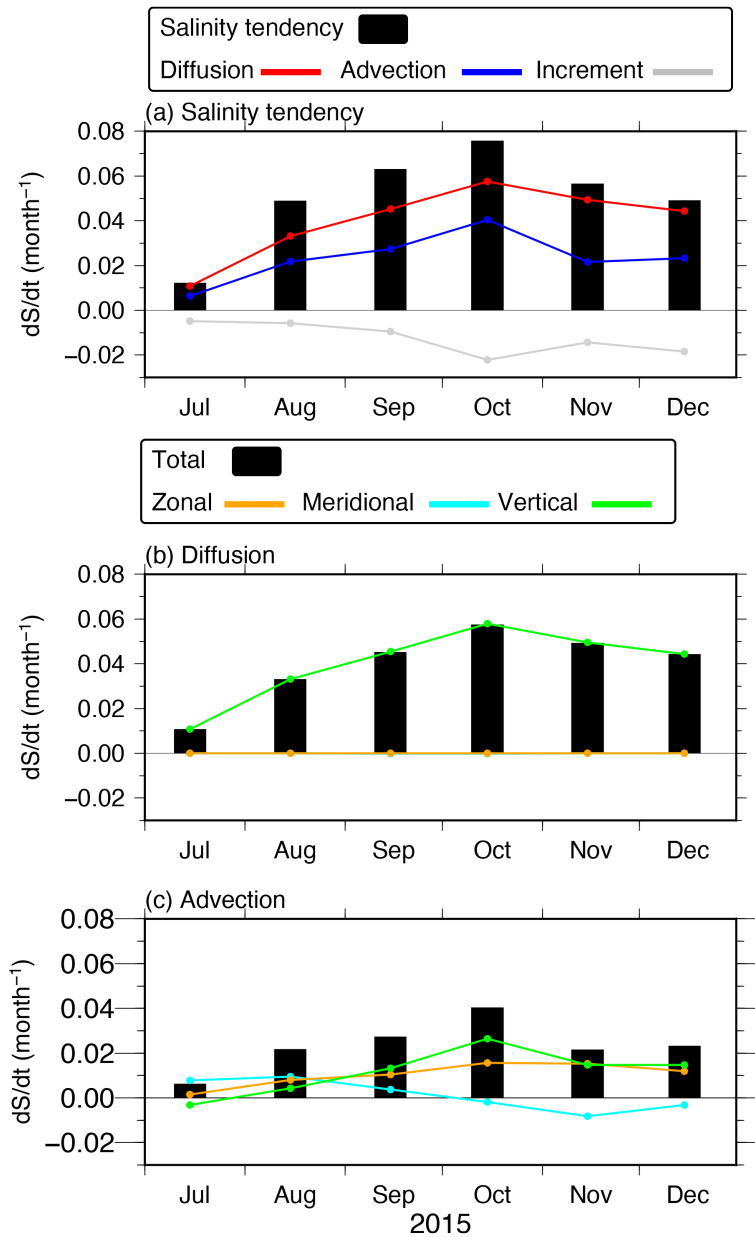



Figure A1: (a) Monthly mean for each term in the salinity budget equation [Eq. (5)] averaged over the KE region in the intermediate layer (140°–160°E, 30°–40°N, and 500–1000 m depth) in the CTL run: the salinity tendency term (the LHS term; black bars), salinity diffusion term (the first term on the RHS; red line), salinity advection term (the second term on the RHS, blue line), and salinity analysis increment term (the last term on the RHS; gray line). (b) and (c) Same as (a), but for salinity diffusion and advection terms (black bars), and zonal (orange lines), meridional (cyan lines), and vertical (green lines) components.

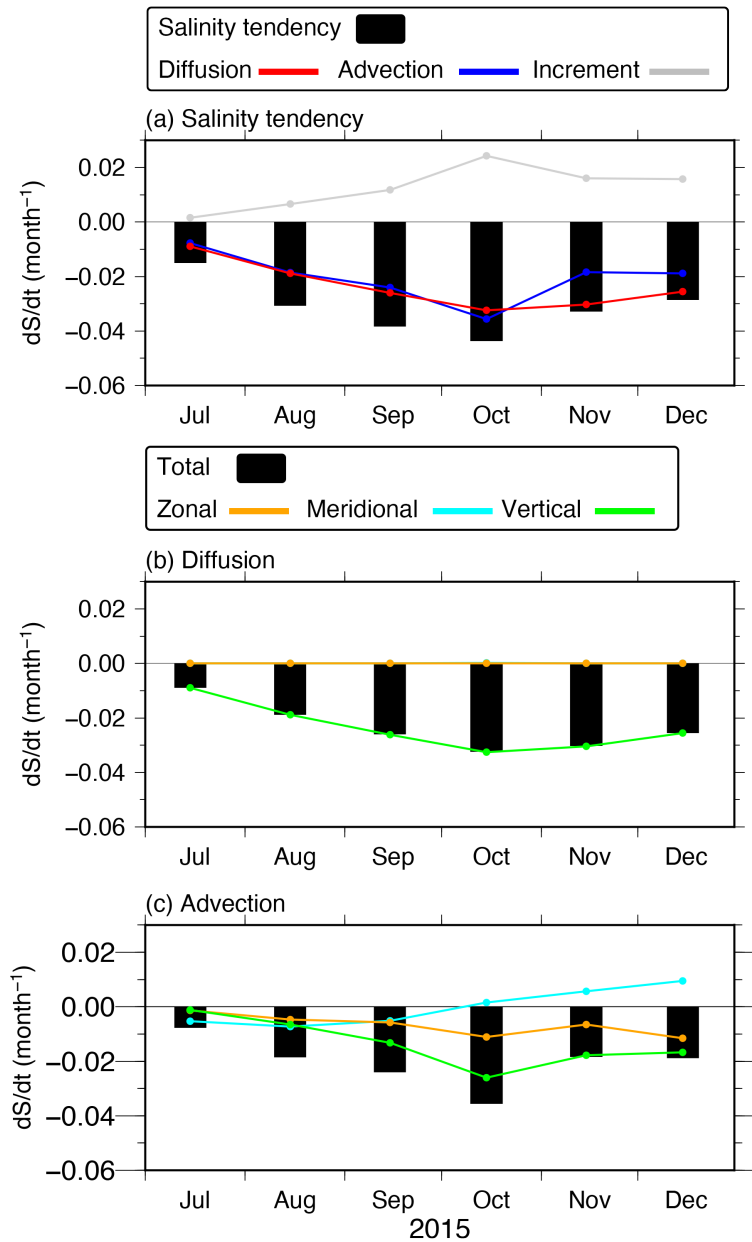



Figure B1: (a) Monthly mean for each term for the difference between the AOEI and CTL runs in the salinity budget equation [Eq. (B1)]: Salinity tendency difference term (the LHS term; black bars), salinity diffusion difference term (the first term on the RHS; red line), salinity advection difference term (the second term on the RHS; blue line), and salinity analysis increment difference term (the last term on the RHS; gray line). (b) and (c) Same as (a) but for the salinity diffusion and advection difference terms (black bars), respectively, and zonal (orange lines), meridional (cyan lines), and vertical (green lines) components.