Response to Reviewer #1:

We thank the reviewer for the constructive comments, which have helped us significantly improve the manuscript. We have carefully addressed each comment and revised the manuscript accordingly. The main revisions include clarifying the open boundary conditions and model configuration, explaining how the EnOI scheme handles sparse Argo data with additional glider-based validation, updating the SLA product URL, adding OISST dataset to the SST and MHW analyses, discussing the seasonal and interannual variations of SST and SLA RMSE, confirming the consistency of the MLD estimation method, and providing a detail discussion of the high eddy counts in SCSORA during 2022–2024. Detailed point-by-point responses are given below, together with the corresponding changes in the revised manuscript.

1. L100-102: What global reanalysis data is used for open boundary conditions?

Response: The open boundary conditions are derived from SODA 3.15.2 data by spatially interpolating, as clarified in the revised manuscript (Lines 143-144). Meanwhile, we have also expanded Section 2.1 (SCSOFSv2 Model Description, lines 130-175) to describe the model configuration in more detail, including the model domain, grid, vertical coordinate, time steps, initial and boundary conditions, atmospheric forcing, advection and mixing schemes, and the EnOI data assimilation system.

Changes in the manuscript:

(1) We have clarified the open boundary data source in lines 143–144, which now read:

" The open boundary values of temperature, salinity, velocity, and SSH are obtained by spatial interpolation of SODA 3.15.2 dataset."

(2) We have expanded Section 2.1 (Lines 130-175) to provide a more complete description of the SCSOFSv2 configuration. The revised text is provided below:

" The SCSORA dataset released in this study is generated by SCSOFSv2. SCSOFSv2 is built on the Regional Ocean Modeling System (ROMS). Its domain covers the SCS and its adjacent waters (99–144° E, 4.5° S–28.3° N). The horizontal grid is curvilinear orthogonal with a resolution of 1/30° in the SCS. The vertical coordinate uses 50 terrain-following σ layers with an improved double stretching function (Shchepetkin and McWilliams, 2005), ensuring sufficient vertical resolution within the upper 300 m to resolve the mixed layer and thermocline structure; the thinnest near-surface layer is approximately 0.09 m. The barotropic and baroclinic time steps are 5 s and 150 s, respectively.

The initial temperature and salinity fields are taken from the January monthly mean of the Generalized Digital Environmental Model version 3.0 (GDEMv3; Carnes, 2009) global climatology. The surface atmospheric forcing is provided by the ECMWF ERA5 reanalysis

(Hersbach et al., 2020) at a horizontal resolution of 0.25°. The forcing variables include 10 m wind speed, mean sea level pressure, 2 m air temperature, 2 m relative humidity, downward longwave radiation flux, precipitation rate, and shortwave radiation flux. All four lateral boundaries are opened. At the open boundaries, the Chapman (1985) boundary condition is applied to the free surface, and the Flather (1976) radiation condition is applied to the depth-averaged (barotropic) velocity. The 3D velocity components, temperature, and salinity are treated using radiation–nudging boundary conditions following Marchesiello et al. (2001). The open boundary values of temperature, salinity, velocity, and SSH are obtained by spatial interpolation of SODA 3.15.2 dataset. For tracer advection, SCSOFSv2 uses the fourth-order Akima scheme (Shchepetkin and McWilliams, 2005) in both the horizontal and vertical directions. For momentum advection, a third-order upstream scheme is used in the horizontal direction and a fourth-order centred scheme in the vertical direction. Vertical turbulent mixing for both momentum and tracers is parameterised using the Mellor–Yamada level 2.5 closure scheme (Mellor and Yamada, 1982).

The system employs an EnOI data assimilation scheme (Evensen, 2003; Oke et al., 2008), assimilating three categories of observations: satellite-derived SST, along-track satellite SLA, and in-situ temperature–salinity profiles. The along-track SLA data are filtered with a Lanczos low-pass filter (cut-off wavelength 65 km, 20-day period) and are not subsampled. Observation errors are set to 3 cm for SLA (Taburet et al., 2018) and are depth-dependent for Argo profiles following Xie and Zhu (2010). To account for asynchronous observations within the assimilation time window, the First Guess at Appropriate Time (FGAT) approach is adopted to reduce assimilation biases arising from observation timing errors. The background error covariance is constructed from an ensemble of model anomalies. A Hanning low-pass filter is applied to the running mean to retain intraseasonal variability, and the anomalies are computed by subtracting the 20-day running mean from the 10-year (1991-2000) free-run daily averaged results. The assimilation cycle is 7 days. Within each cycle, all SLA and Argo observations within the 7-day time window, and SST data 1-day before the analysis date are used. It should be noted that the number of Argo profiles in the SCS was very limited in the early 2000s and remains relatively sparse in the shallow southern SCS. To mitigate this data sparsity, the assimilation system also relies on satellite SST and along-track SLA observations, which provide better spatiotemporal coverage. In addition, the ensemble-based background error covariance allows an Argo observation to influence nearby model variables within a localization radius of 150 km."

2. L125-126: The number of Argo floats should be small in the SCS in the early 2000s. Also, the number of Argo floats should be limited in the southern SCS because the depth

is shallow in the region. Please comment more about how the assimilation using Argo data works in the SCS.

Response: We fully agree that the number of Argo floats in the SCS was indeed very limited in the early 2000s. Below we describe the actual distribution of SCS Argo data and explain how the SCSOFSv2 assimilation system handles this data sparsity.

The global Argo array was gradually deployed in the early 2000s and only reached its target of approximately 3,000 floats around 2007 (Roemmich et al., 2009). As a marginal sea of the northwestern Pacific Ocean, the SCS began to see a gradual increase in available Argo profiles after 2006. Despite this sparse coverage, the SCSOFSv2 assimilation system still exploits the available Argo profiles through the following mechanisms:

(1) Spatially correlated ensemble background error covariance. The EnOI scheme constructs the background error covariance from model anomalies of a 10-year (1991–2000) free run. This ensemble-based covariance inherently contains spatial correlations, so an Argo observation at one location can influence nearby model variables through the covariance structure. With a localization radius of 150 km, the influence of each Argo profile is spatially limited but still allows information to propagate to neighboring regions.

(2) Complementary multi-source observations. Argo profiles are assimilated together with satellite SST and along-track SLA. In regions and periods where Argo data are sparse, satellite observations with better spatiotemporal coverage still constrain the sea surface and column-integrated ocean state.

(3) Temporal extension of observation impact through FGAT and IAU. The FGAT approach computes observation innovations within 56 three-hour slots, reducing the maximum time mismatch to 1.5 hours. IAU applies the analysis increment as a forcing term that is gradually introduced over the 7-day assimilation cycle. This means that even if Argo observations are available only on certain days, their impact is spread over the entire cycle rather than applied instantaneously, which is particularly beneficial for irregularly distributed observations (Zhu et al., 2022).

(4) Physically consistent model background. In the absence of Argo data, the model background driven by ERA5 atmospheric forcing and constrained by SODA open boundary conditions still provides a physically consistent estimate. Argo assimilation then provides corrections where and when data are available.

To evaluate the effect of assimilation in the southern SCS, we compared SCSORA and a free run without assimilation against independent glider observations from 2019. As shown in Figure R1, SCSORA reproduces the subsurface thermohaline structure substantially better than the free run, indicating that assimilation effectively improves subsurface state estimation even

in this region. In future improvements to the assimilation system, we will further assimilate additional observation types, such as gliders, subsurface moorings, and drifters, to strengthen the constraint on subsurface structure in the shallow-water region of the southern SCS.

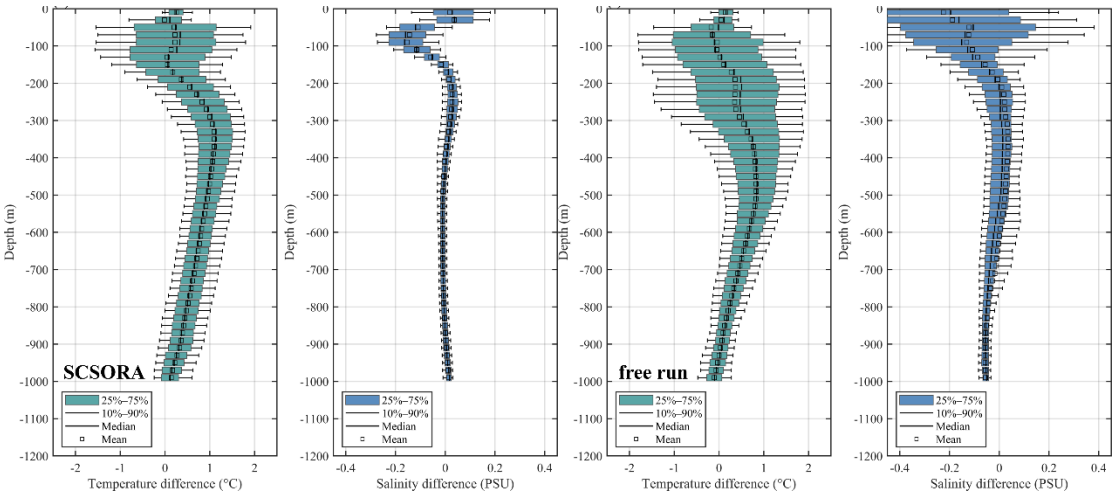


Figure R1. Vertical profiles of temperature and salinity differences for SCSORA and the free run.

Changes in the manuscript:

(1) We have added a description of the EnOI assimilation scheme and the assimilated observation types (Lines 149–151). The added text reads:

"The system employs an EnOI data assimilation scheme (Evensen, 2003; Oke et al., 2008), assimilating three categories of observations: satellite-derived SST, along-track satellite SLA, and in-situ temperature–salinity profiles. The along-track SLA data are filtered with a Lanczos low-pass filter (cut-off wavelength 65 km, 20-day period) and are not subsampled."

(2) We have added a brief description of the Argo data sparsity in the SCS and the mechanisms by which the assimilation system mitigates this sparsity in Section 2.1 (Lines 158–175). The added text reads:

" It should be noted that the number of Argo profiles in the SCS was very limited in the early 2000s and remains relatively sparse in the shallow southern SCS. To mitigate this data sparsity, the assimilation system also relies on satellite SST and along-track SLA observations, which provide better spatiotemporal coverage. In addition, the ensemble-based background error covariance allows an Argo observation to influence nearby model variables within a localization radius of 150 km."

3. L155: Please check the URL for the SLA product.

Response: We thank the reviewer for pointing this out. We have checked and updated the URL for the SLA product (Line 251 in the revised manuscript). The data are available at

https://doi.org/10.48670/moi-00148.

Changes in the manuscript:

We have updated the URL for the SLA product in line 251 of the revised manuscript. The revised text now reads:

" The data are available at https://doi.org/10.48670/moi-00148."

4. Sec. 3.1: Could you plot or check the SST anomaly of OISST (Fig. 2top)? Do the time series mostly overlap SCSORA due to the assimilation? There are annual variations in the SST RMSE: large in winter and small in summer. Is this why spatial variations are larger in winter than summer?

Response: We thank the reviewer for the valuable comments. We address each point below.

(1) In the revised manuscript, we have added the OISST SST anomaly time series to Figure 2 (Line 338). The time series of the three datasets show broadly consistent variation trends.

(2) SCSORA assimilates OISST data. In the validation, SCSORA is compared separately with OISST and OSTIA, and OSTIA is not involved in the assimilation, so it serves as an independent validation dataset. The results show a mean RMSE of 0.340°C between SCSORA and OISST, and 0.357°C between SCSORA and OSTIA. Both values are low and comparable in magnitude, indicating that SCSORA agrees well with both satellite SST products. Since OSTIA is not assimilated, the close agreement between SCSORA and OSTIA provides an independent check of the SST field.

(3) The SST RMSE exhibits a clear seasonal cycle, and the spatial patterns of RMSE are broadly consistent between the two comparisons (SCSORA vs OISST and SCSORA vs OSTIA; Figure 3, line 381 in the revised manuscript). In winter, the RMSE increases markedly in the Taiwan Strait, forming a prominent high-error band, which corresponds to the region of strong northeasterly monsoon forcing, cold coastal water transport, and pronounced SST fronts (Kuo et al., 2018). In summer, the central SCS maintains low RMSE, while a localized high-error band appears along the Vietnam coast, which may be related to southwesterly monsoon-driven coastal upwelling and offshore cold-water expansion (Xie et al., 2003). Therefore, the spatial RMSE is generally larger in winter than in summer, with a broader high-error region. This seasonal difference may be related to the seasonal intensification of monsoon-driven processes and the increased sensitivity of the model to wind forcing, vertical mixing, horizontal advection, and COARE3.0 bulk algorithm.

Changes in the manuscript:

(1) We have added the OISST SST anomaly time series to Figure 2 (Line 338). The revised figure caption now reads:

"

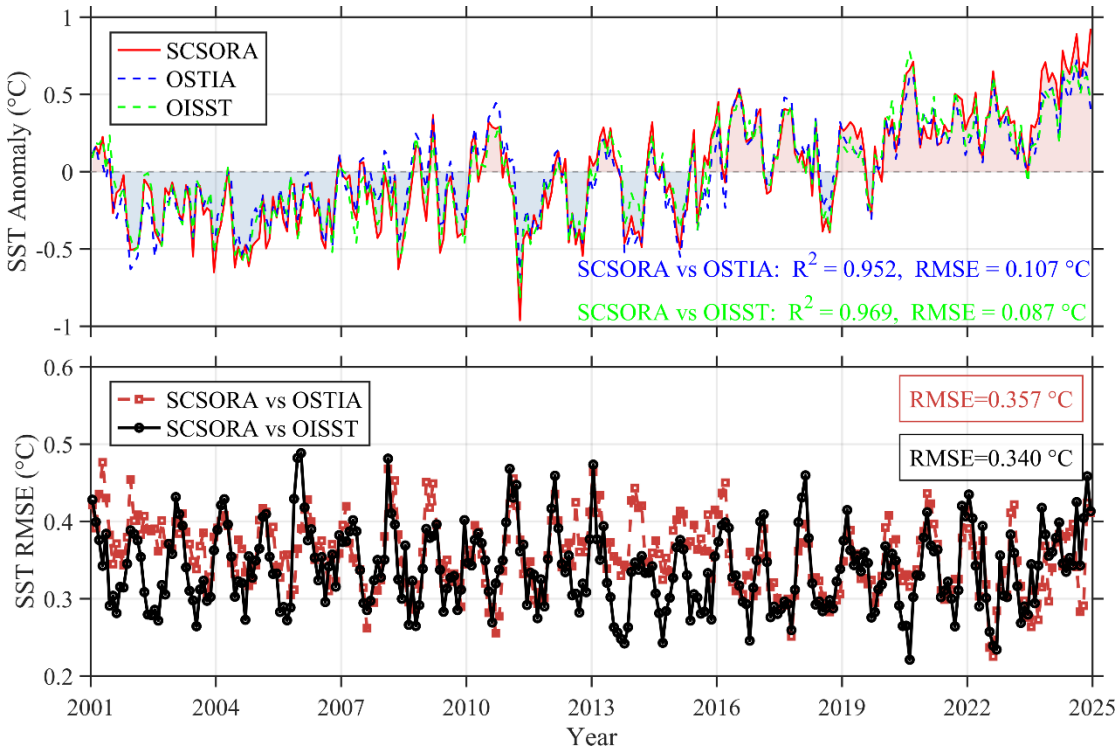


Figure 2. Time series of SST anomalies and RMSE relative to the 2001–2024 SST climatology. Comparisons are shown among SCSORA, OSTIA, and OISST. SST anomalies are computed as daily departures from the 2001–2024 climatological SST. "

(2) We have added the seasonal spatial distributions of SST RMSE between SCSORA and OISST (top row) and between SCSORA and OSTIA (bottom row) as Figure 3 (Line 381). The revised figure caption now reads:

"

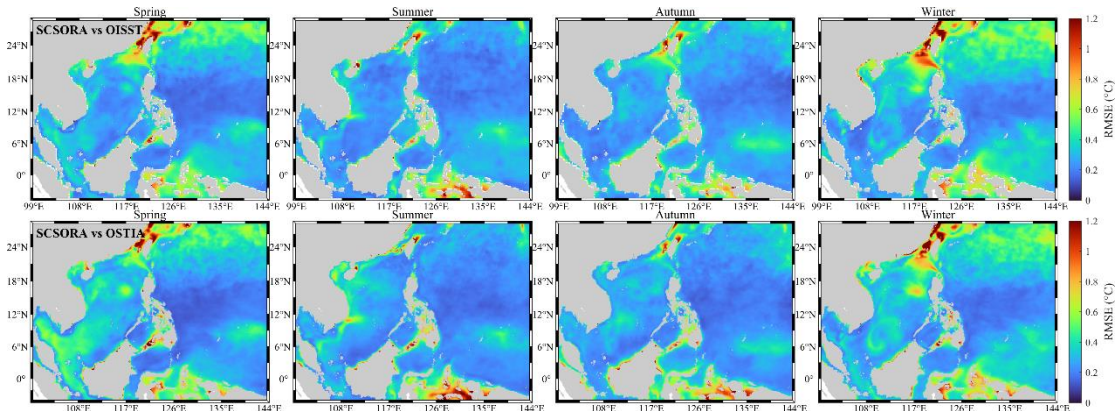


Figure 3. Seasonal spatial distributions of SST RMSE between SCSORA and OISST (top row) and between SCSORA and OSTIA (bottom row) for spring, summer, autumn, and winter (from left to right). "

(3) We have added the corresponding discussion of the seasonal RMSE patterns in Section 3.1 (Lines 354–380). The added text reads:

"The RMSE of SST between SCSORA and both OISST and OSTIA exhibits pronounced regional disparities and seasonal evolution (Figure 3). Overall, the RMSE in the central open waters of the SCS is generally below 0.5°C, whereas larger errors are predominantly distributed in the Taiwan Strait, the northern SCS continental shelf, and certain coastal regions, with localized values exceeding 0.8°C. Such spatial discrepancies may be associated with complex ocean dynamical processes in nearshore and shelf seas, such as monsoon-driven coastal circulations, tidal mixing, and local upwelling.

In terms of seasonal variability, the RMSE increases markedly during winter in the northern SCS and adjacent to the Taiwan Strait, forming a relatively prominent high-error region. Previous studies have demonstrated that the northeasterly monsoon can drive the southward transport of cold coastal water and generate a distinct SST front in the Taiwan Strait; concurrently, intensified wintertime sea surface cooling in the northern SCS further enhances the regional temperature gradient (Kuo et al., 2018). Therefore, the relatively large RMSE in these regions may be attributed to model biases in simulating monsoon-driven sea surface heat fluxes, coastal water mass transport, and the positioning of temperature fronts. As the northeasterly monsoon gradually weakens, the high-error extent in the northern SCS diminishes accordingly. During summer, the central open waters of the SCS maintain relatively low RMSE, yet a localized high-value band emerges along the Vietnam coast. Xie et al. (2003) noted that the southwesterly monsoon can drive coastal upwelling off Vietnam and promote the offshore expansion of cold water through local wind stress curl and anticyclonic circulation. These processes may intensify the local SST spatial gradient, rendering the simulated results in this region more sensitive to the accuracy of wind fields, vertical mixing, and horizontal advection. Nevertheless, SCSORA is capable of reproducing SST over most of the SCS reasonably well across all seasons, particularly exhibiting low RMSE in open waters, which indicates its favorable capability for SST reconstruction."

5. Sec. 3.2. If possible, please comment on why the SLA RMSE is relatively large in 2017-2022 (Fig. 3bottom).

Response: We further examined the temporal variations of SLA and its RMSE over 2001–2024. Figure 4 (Line 398 in the revised manuscript) shows that AVISO SLA exhibits pronounced increases around 2017 and 2021, and SCSORA reproduces these increases reasonably well, although the amplitude is generally weaker, leading to relatively larger SLA RMSE during these periods. Such increases have been reported to be associated with multiple factors, including the long-term sea level rise trend, La Niña, the PDO transition from a warm to a cool phase, typhoon-induced storm surges, and wind stress and Ekman transport during the monsoon transition (Wang et al., 2026). In SCSOFSv2, the assimilation of along-track SLA

involves a Lanczos low-pass filter, which removes part of the signal and weakens the model's ability to reproduce the full amplitude of the SLA increase.

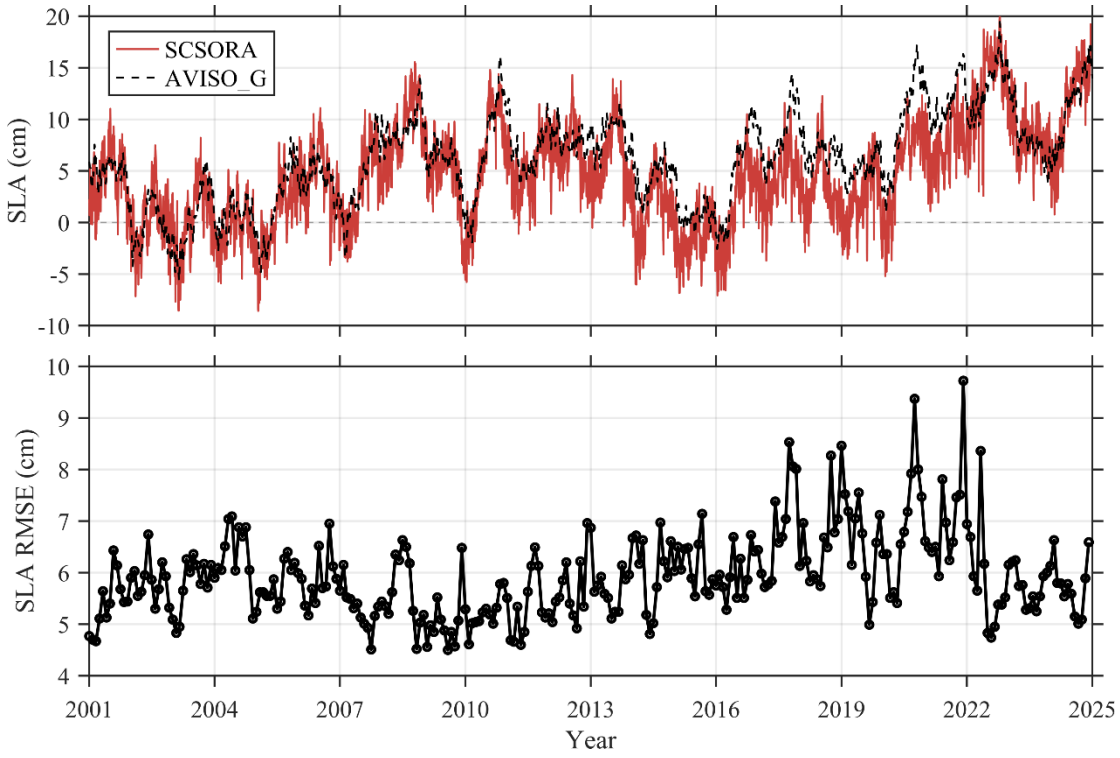


Figure 4. Time series of SLA and RMSE between the SCSORA and AVISO observations. RMSE represents the temporal deviation of the SCSORA from the observations over the same period.

6. Fig. 6: Near the coast and southern SCS south to the equator, the MLD in GLORYS is not estimated. Did you use the strictly same method to estimate the MLD in both dataset?

Response: We thank the reviewer for this comment. Upon re-confirmation, MLD was estimated using the same method for both SCSORA and GLORYS: a 0.2°C threshold criterion with a 10 m reference depth, and the temperature baseline was taken from the climatological temperature data of each dataset. No special treatment or exception was applied to either dataset. Regarding the missing MLD in GLORYS along the coast and in the southern SCS south of the equator (Figure 9, line 518), we checked the code output and confirmed that no conditions satisfying the 0.2°C threshold criterion were found in these regions, and thus no valid MLD was produced. This is most likely due to the coarser vertical resolution of GLORYS.

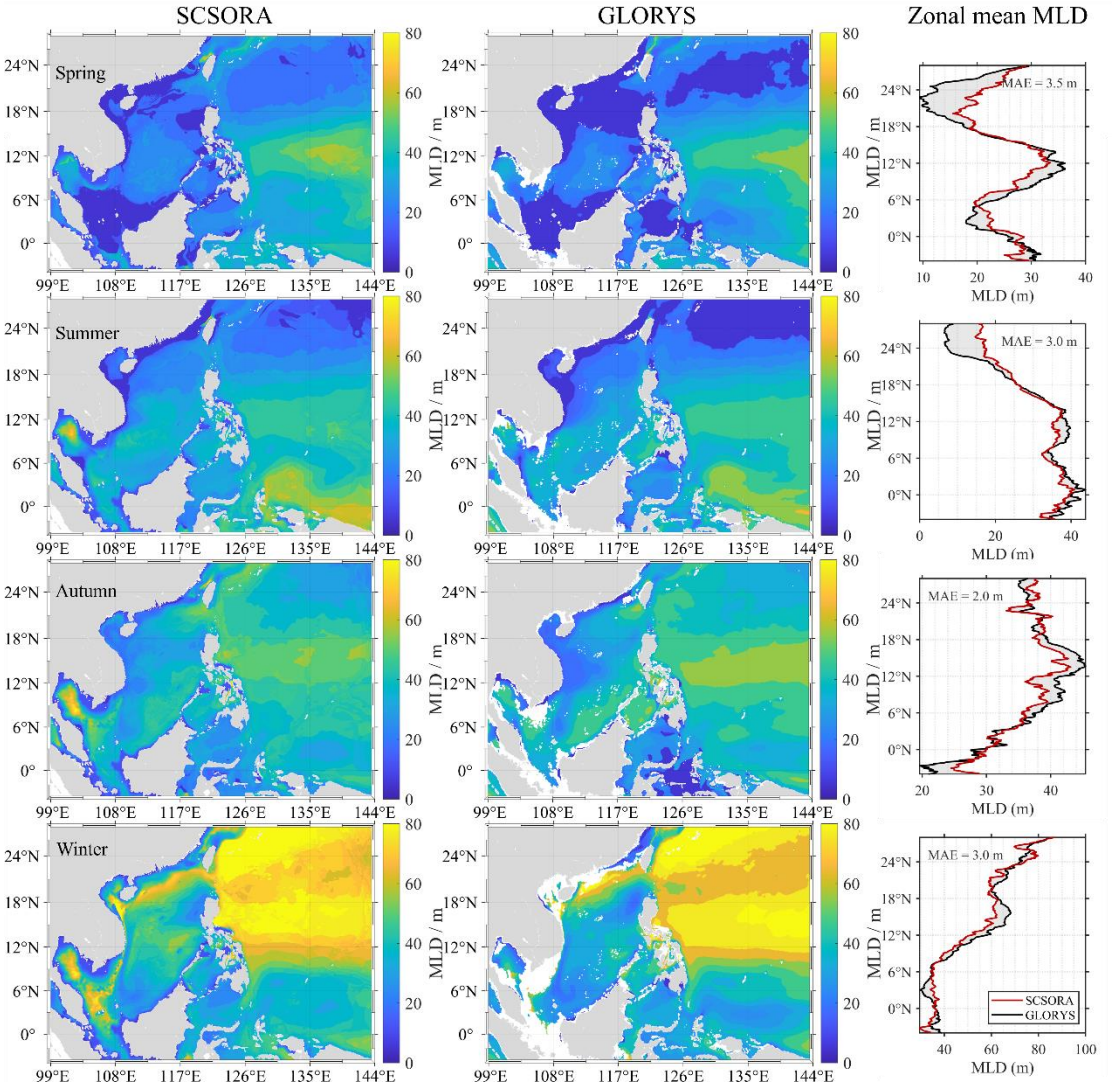


Figure 9. Seasonal climatological MLD from SCSORA (left column) and GLORYS (middle column) for spring, summer, autumn, and winter, along with the corresponding zonal mean MLD profiles (right column; red: SCSORA, black: GLORYS). The MAE of zonal mean MLD between the two products is annotated in each right panel.

7. Sec. 4.2: Could you plot or check the time series of the MHW using OISST in Fig. 10 and Fig. 11? Do these time series overlap with those using SCSORA? How the assimilated data constrains the SST in SCSORA should be checked.

Response: In the revised manuscript (Lines 712 and 736), we have added the OISST time series. In the revised Figure 15 (formerly Figure 10), the interannual time series of MHW frequency, duration, and cumulative intensity now include SCSORA, OSTIA, and OISST. In the revised Figure 16 (formerly Figure 11), the daily SST, climatology, and MHW threshold are also presented for all three products. OISST and SCSORA show broadly consistent interannual variations in MHW frequency, duration, and cumulative intensity. SCSORA reproduces the main peaks and the increasing trend evident in OISST, although slight differences exist in

amplitude at the event scale. For example, in the NWPO in 2020, OISST shows the highest peaks in MHW frequency and cumulative intensity, followed by SCSORA, while OSTIA is relatively lower. SCSORA assimilates satellite-blended OISST through the EnOI scheme of SCSOFSv2, which directly constrains the model SST. In addition, along-track SLA and Argo temperature–salinity profiles are assimilated simultaneously, constraining the upper-ocean thermal structure and further influencing SST through model dynamics and coherence relationships among them. To examine whether the agreement with OISST arises solely from assimilation, we also compared SCSORA with OSTIA, which is not assimilated. The mean RMSE between SCSORA and OSTIA is 0.357°C, comparable to the 0.340°C between SCSORA and OISST, indicating that SCSORA agrees better with OISST than that of OSTIA. In addition, the subsurface thermohaline structure is validated against independent glider and Argo observations, providing further evidence.

Changes in the manuscript:

(1) We have added the OISST time series to the revised Figure 15 (formerly Figure 10) and Figure 16 (formerly Figure 11) (Lines 712 and 736). The revised figure captions now read:

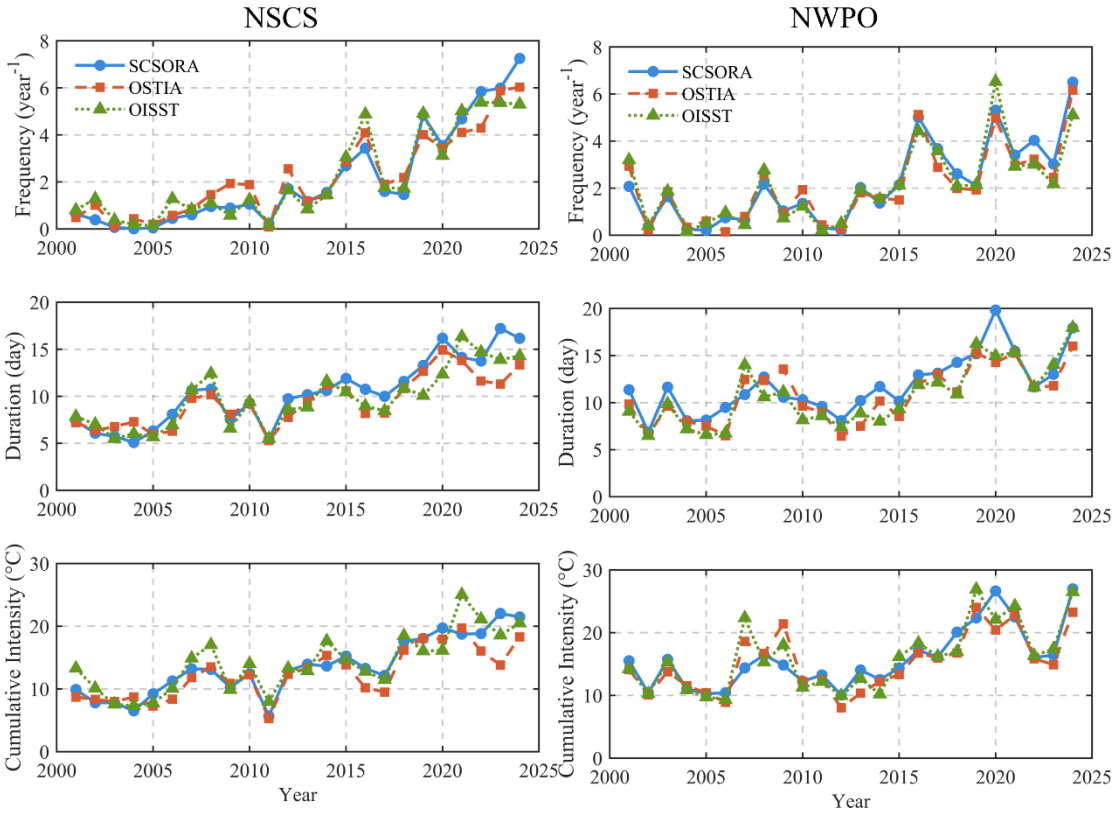


Figure 15. Interannual variability of marine heatwave (MHW) characteristics in the NSCS (left column) and the NWPO (right column) during 2001–2024, including annual frequency (top row, units: year^{-1}), mean duration (middle row, units: days), and cumulative intensity (bottom row, units: °C).

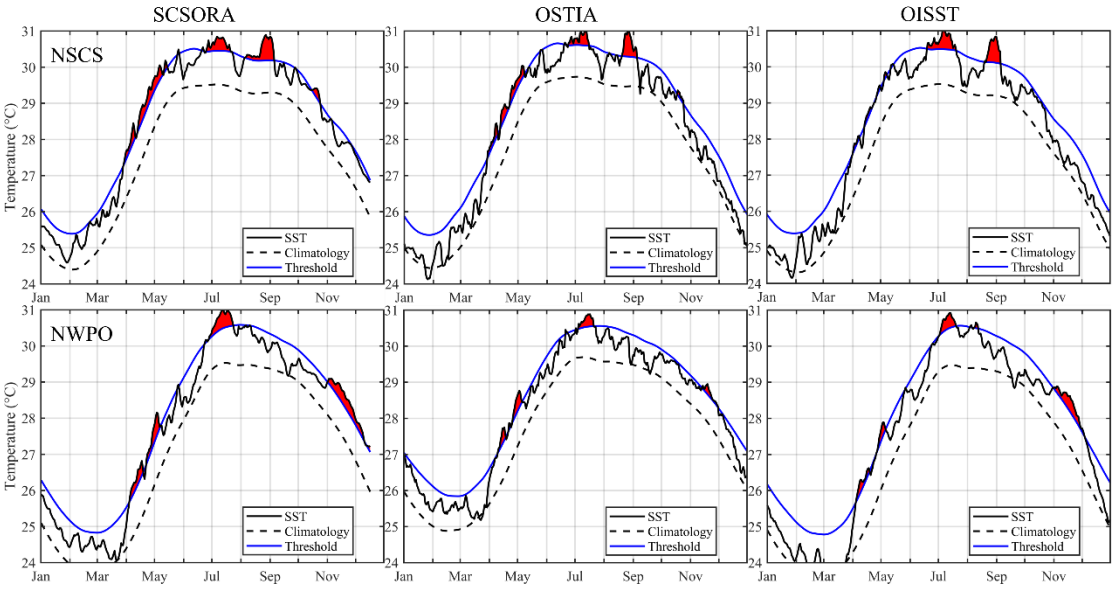


Figure 16. Seasonal evolution of area-averaged SST from SCSORA (left column), OSTIA (middle column), and OISST (right column) in the NSCS (top row) and the NWPO (bottom row) for the year 2024. Black solid lines indicate daily mean SST in 2024, black dashed lines indicate climatological mean SST, and blue solid lines indicate the MHW threshold (90th percentile of climatology). Red shading denotes marine heatwave (MHW) events where daily SST exceeds the threshold.

(2) We have added the corresponding discussion in Section 4.2 (Lines 716–735). The added text reads:

"Figure 16 presents the daily SST, climatology, and MHW thresholds based on SCSORA, OSTIA, and OISST for the NSCS and NWPO subregions in 2024, with MHW events exceeding the threshold marked by red shading. Overall, the three products exhibit a consistent seasonal cycle: SST reaches its minimum value during January–March and its peak value during June–September, while the threshold curve lies distinctly above the climatology, defining the identification criterion for MHWs (Hobday et al., 2016). In the NSCS, all three products identify significant summer MHW events, but differences exist at the event scale. The SCSORA SST time series is relatively smooth with weaker high-frequency fluctuations, resulting in longer event durations (up to 34 days) and slightly lower peak intensities (0.71°C). OSTIA is more sensitive to high-frequency variability, capable of capturing shorter events (up to 23 days), with slightly higher peak intensities (0.75°C). In the NWPO, the MHW signal is generally weaker, with lower duration and intensity; the strongest events recorded by the three products differ, with SCSORA at 0.64°C, OSTIA at 0.46°C, and OISST at 0.53°C. OISST exhibits a variation trend broadly consistent with SCSORA, but with MHW intensity and MHW duration slightly lower than those of SCSORA."

8. Sec. 4.3: Why is the number of eddies extremely large in 2022, 2023, and 2024 in SCSORA, which is not observed in AVISO (Fig 12 top panels)? Do the small-scale (submesoscale) eddies increase, which cannot be detected in AVISO?

Response: We thank the reviewer for this question. SCSORA and AVISO are generally consistent in the interannual variability of eddy counts, with both showing an increasing trend in eddy number during 2020–2023; however, the increase in SCSORA is more pronounced, and SCSORA identifies noticeably more eddies than AVISO, especially during 2022–2024. This increase may be related to the following factors. In one side, eddy activities in the SCS are modulated by the monsoon, ENSO, and PDO through local wind stress curl (Wang et al., 2003; Cheng and Qi, 2010; He et al., 2016), and the negative PDO and consecutive La Niña conditions in the 2020s may favor cyclonic eddy generation. Meanwhile, the Kuroshio intrusion through the Luzon Strait can enhance frontal and shear instability in the northern SCS, also favoring the generation of new eddies (Wu and Chiang, 2007).

In the other side, the eddy identification method itself may introduce differences of eddy counts, since the differences in effective resolution between those two datasets. Although SCSORA and AVISO employ the same velocity-vector-based identification method, their spatial resolutions differ, and the parameter choices in the method (e.g., grid-scale-related thresholds, minimum area, and maximum velocity criteria) need to be specified separately for each dataset. In this study, both datasets use A = 3 and B = 2 in common, but this parameter combination was designed for the resolution of AVISO. When applied directly to the higher-resolution SCSORA, it may lead to the identification of more small-scale eddies, thereby causing differences in eddy counts. The above phenomenon and its possible causes are also discussed in lines 742–812 of the revised manuscript.

Therefore, in subsequent research, sensitivity experiments will be designed based on SCSORA and SCSOFSv2 to investigate the mechanisms underlying the increase in eddy number during 2020–2023, including large-scale background forcing (monsoon, ENSO, PDO, and Kuroshio intrusion) and the response of eddies at different scales to this increase. In addition, identification parameters adapted to SCSORA will be developed to systematically examine the effects of identification parameters (e.g., grid-scale-related thresholds, minimum area, and maximum velocity criteria) on eddy identification results, and to analyze the differences in eddy number, size distribution, and eddy core amplitude identified under different parameter settings, so as to distinguish whether the higher eddy counts in SCSORA during 2022–2024 mainly arise from a real increase in small-scale eddies or from the capture of more small-scale eddies by the identification parameters under higher-resolution data.

Changes in the manuscript:

We have added a discussion of the possible causes of the higher eddy counts in SCSORA

during 2022–2024 in Section 4.3 (Lines 742–812). The revised text now reads:

" The statistical results indicate that in the eddy-active region of the SCS (102–122°E, 5–24°N), cyclonic eddies outnumber anticyclonic eddies (Figure. 17). SCSORA and AVISO are generally consistent in interannual variability, but SCSORA identifies a slightly higher number of eddies, especially during 2022–2024. Notably, both AVISO and SCSORA exhibit an increasing trend during 2020–2023, with a more pronounced increase in SCSORA, especially during 2022–2024. Eddy activity in the SCS is modulated by the monsoon, ENSO, and PDO through local wind stress curl (Wang et al., 2003; Cheng and Qi, 2010; He et al., 2016). The negative PDO and consecutive La Niña conditions in the 2020s may favor cyclonic eddy generation; meanwhile, the Kuroshio intrusion through the Luzon Strait can enhance frontal and shear instability in the NSCS, also favoring the generation of new eddies (Wu and Chiang, 2007). In addition, differences in effective resolution among datasets may also affect the comparison of eddy counts. Such method- and resolution-dependent discrepancies are common in eddy detection and eddy data products (Dong et al., 2025). However, the specific mechanisms can only be clarified through dedicated sensitivity experiments and validation against higher-resolution observations. In terms of seasonal variability (bottom row of Figure. 17), cyclonic eddies exhibit a bimodal distribution, with two peaks occurring in March–May and August–October, respectively. In contrast, the seasonal variability of anticyclonic eddies is weaker, manifesting only as a single broad peak in May–August."