To Reviewer 2

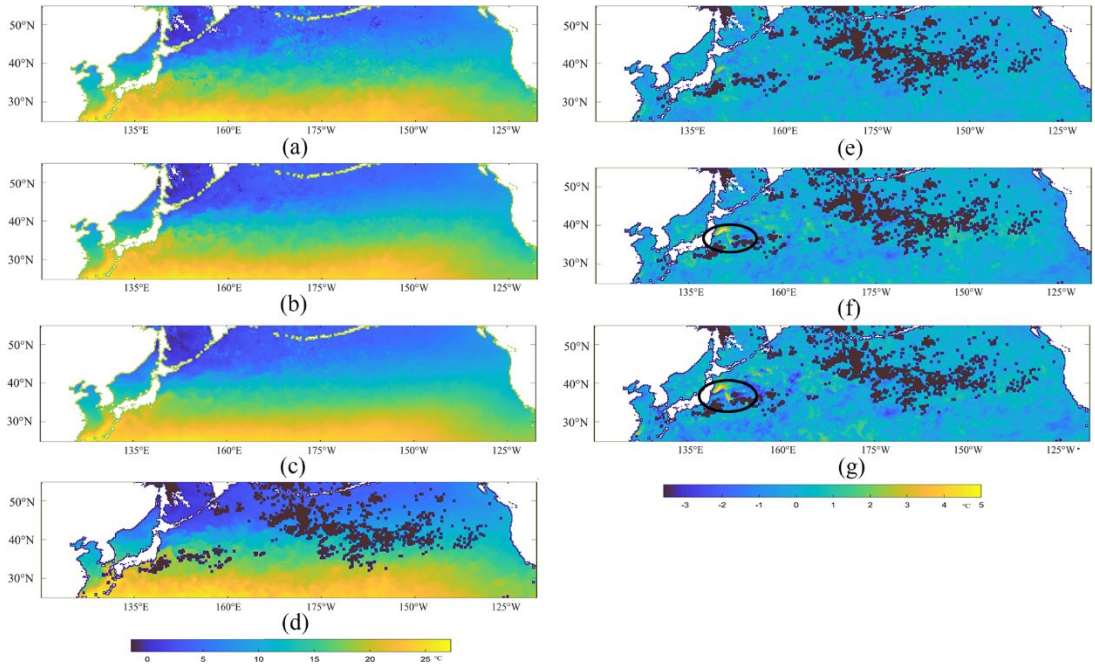
First, we appreciate the reviewer's valuable comments. For your comments, we gave our corresponding explanations and responses below:

1. Figure 7 combines data from three variables (SST, SCHL, and SSW) with different physical units into a single scatter density plot without axis labels or units. The same issue applies to Figure 13. Please either (a) plot each variable separately with proper physical units (°C for SST, mg/m³ for SCHL, and m/s for SSW) on the axes, or (b) if maintaining the combined plot format, clearly state in the figure caption that values are shown in original physical units after inverse normalization, and list the respective units for each variable.

Response: We wish to retain the combined plot format and include the following statement in the figure captions of Figs. 7 and 13: *"*Values are shown in the original physical units after inverse normalization: SST in °C, SCHL in mg/m³, and SSW in m/s.*"*

2. The longitude and latitude labels in Figures 9–11 appear compressed or distorted. Please check.

Response: In the North Pacific region (Figs. 9–11), the longitudinal range is much greater than the latitudinal range, resulting in images of size 120 × 530. To prevent stretching or distortion that could affect the interpretation of the displayed regions, these figures were generated by strictly maintaining the proportional coverage of longitude and latitude, with an aspect ratio of approximately 1:4.42. In the revised version, we regenerated the longitude and latitude labels, as shown below.



3. The validation is performed only against the original satellite-derived data. Given that the original data themselves contain uncertainties and gaps, I suggest validating the reconstructed results against other high-quality monthly products (e.g., reanalysis data) to more rigorously assess the absolute accuracy of T-DINEOF.

Response: First, we believe that reconstructing the target dataset and then evaluating the accuracy between the reconstructed data and the target data at known locations can

better reflect the reconstruction performance of the algorithm on the target dataset itself (Both Original DINEOF and Multi-DINEOF are implemented based on this method). If the reconstructed results are compared with other types of datasets, such as reanalysis data, the discrepancies between different data sources may affect the assessment of reconstruction accuracy. Taking SST as an example, reanalysis products integrate multiple SST datasets, and the retrieval algorithms used for these datasets differ from those of the target satellite observations, thereby introducing additional uncertainties. Furthermore, the interpolation procedures used in reanalysis products may generate unrealistic SST values, which can further increase the discrepancies.

To verify this point, we introduced the CMEMS Global Ocean Ensemble Reanalysis product at 1/4° resolution monthly means of temperature and velocity (CMEMS GLOBAL_MULTIYEAR_PHY_ENS_001_031) and the monthly chlorophyll-a product at 1/4° resolution derived from the PISCES biogeochemical model (CMEMS GLOBAL_MULTIYEAR_BGC_001_029) to further evaluate the reconstruction accuracy of T-DINEOF and Multi-DINEOF.

The results are summarized in the following table. Overall, the reconstruction errors relative to the reanalysis and biogeochemical model datasets are larger than those obtained when directly comparing with the target satellite observations. In addition, it can be seen that the errors obtained by the T-DINEOF and Multi-DINEOF methods (first two rows of the table) are comparable to the discrepancies between the satellite observations and the reanalysis data themselves (third row). This indicates that the substantial differences between the target satellite data and the reanalysis products tend to mask the actual reconstruction errors of the algorithms. In other words, the reconstructed results are naturally closer to the target satellite observations, since the reconstruction algorithms are specifically designed to recover the target satellite data.

Therefore, we believe that the relatively large discrepancies between the target observations and the reanalysis datasets may obscure the intrinsic performance evaluation of the reconstruction algorithms. For this reason, we would prefer to retain only the accuracy assessment based on comparisons with the target satellite observations in the manuscript, and we sincerely hope for the reviewer's understanding.

	RMSE	MAE	r	R^2
T-DINEOF	4.8626	3.0245	0.9331	0.7972
Multi-DINEOF	4.8715	3.0288	0.9329	0.7965
Sat-Reanalysis	4.8695	3.0215	0.9328	0.7966

4. To better compare the detail-preserving capabilities of DINEOF, Multi-DINEOF, and T-DINEOF, I suggest adding local standard deviation maps or gradient maps.

Response: We calculated the gradient maps of the SST, SCHL, and SSW data reconstructed by T-DINEOF, Multi-DINEOF, and Single-DINEOF, respectively. The SST gradient maps have been added as Fig. 10 in the main text, while the SCHL and SSW gradient maps have been included in the supplementary file due to space limitations. The three sets of gradient maps are shown below. Overall, compared with the Multi-DINEOF and Single-DINEOF methods, the data reconstructed by the T-DINEOF method preserve richer detail information and exhibit larger gradient

magnitudes, particularly in the central and eastern North Pacific, further demonstrating the superior detail-preserving capabilities of T-DINEOF.

We have also added the following description in the main text.

"To analyze the detail-preserving capabilities of Single-DINEOF, Multi-DINEOF, and T-DINEOF, the gradients between adjacent eastward and northward pixels were calculated for the SST, SCHL, and SSW images reconstructed by the three methods. The square root of the sum of squared gradients in the two directions was then used as the pixel gradient magnitude to generate the corresponding gradient maps. The SST gradient maps are shown in Fig. 10, while the gradient maps of SCHL and SSW are presented in Figs. S4 and S5, respectively. Overall, compared with the Multi-DINEOF (a) and Single-DINEOF (b) methods, the T-DINEOF method (c) produces more gradient information with higher gradient magnitudes, indicating a better preservation of fine-scale details. The Multi-DINEOF and Single-DINEOF methods yield relatively low gradient values in the central and eastern North Pacific, failing to adequately capture the detailed structures in these regions. In contrast, although all three methods produce relatively high gradient values in the western North Pacific, the T-DINEOF method preserves substantially richer details, further demonstrating its advantage in detail-preserving capabilities."

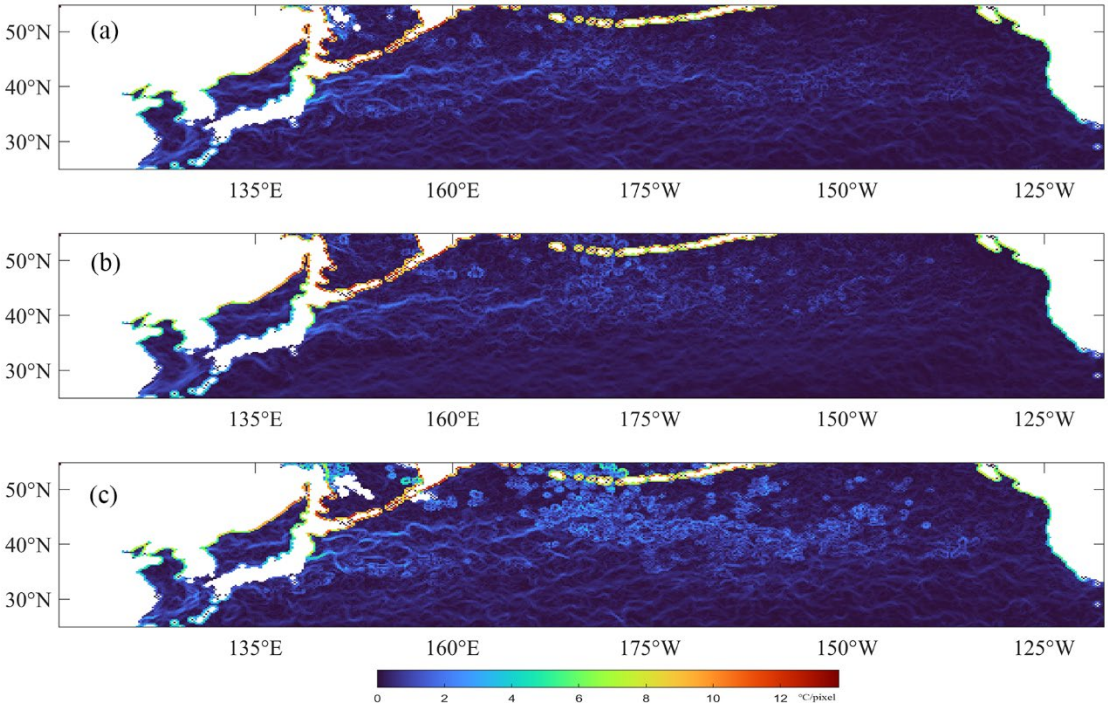


Fig. 1 Gradient maps of SST for April 2022 in the northern Pacific obtained from (a) Multi-DINEOF, (b) Single-DINEOF, and (c) T-DINEOF.

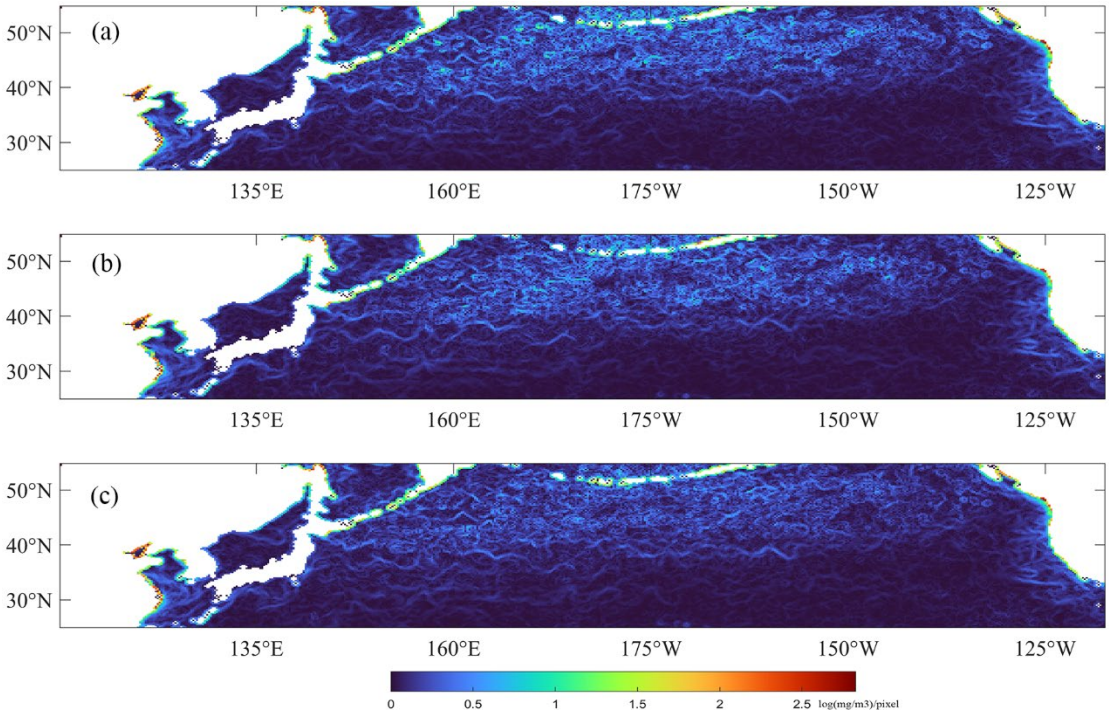


Fig. 2 Gradient maps of SCHL for July 2021 in the northern Pacific obtained from (a) Multi-DINEOF, (b) Single-DINEOF, and (c) T-DINEOF.

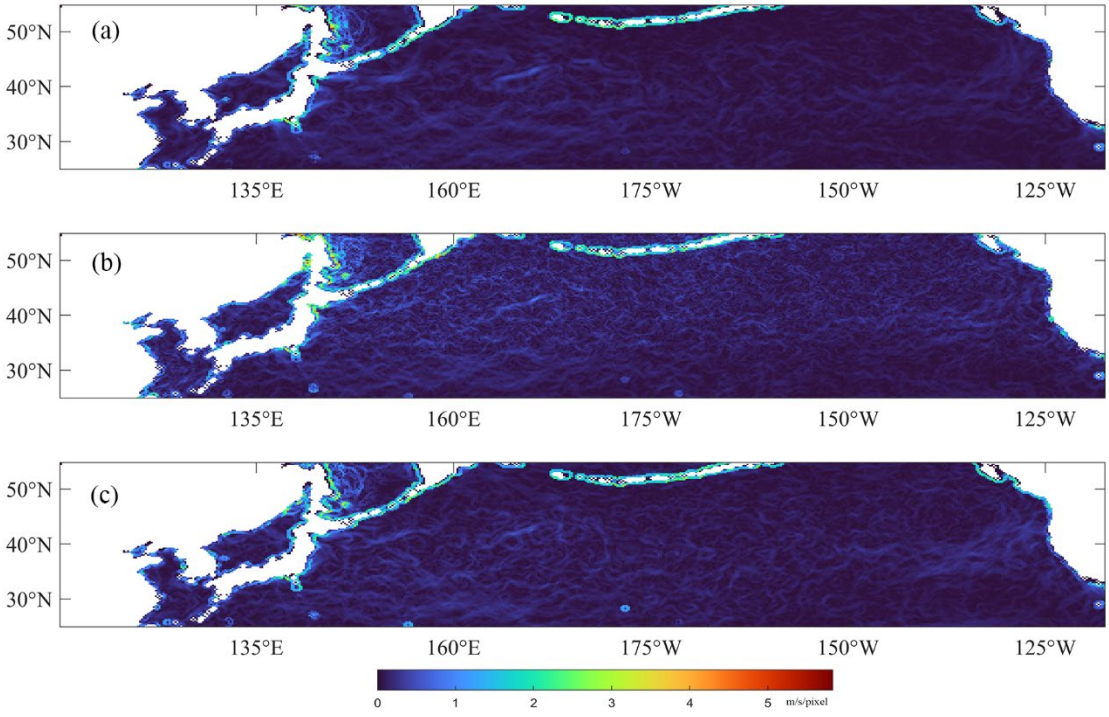


Fig. 3 Gradient maps of SSW for January 2020 in the northern Pacific obtained from (a) Multi-DINEOF, (b) Single-DINEOF, and (c) T-DINEOF.

5. In Figures 7 and 13, a small number of points show underestimation relative to the 1:1 line. What causes this? Please analyze whether these underestimated points correspond to specific geographic regions (e.g., coastal upwelling zones, frontal regions), specific seasons (e.g., summer stratification periods), or specific variables

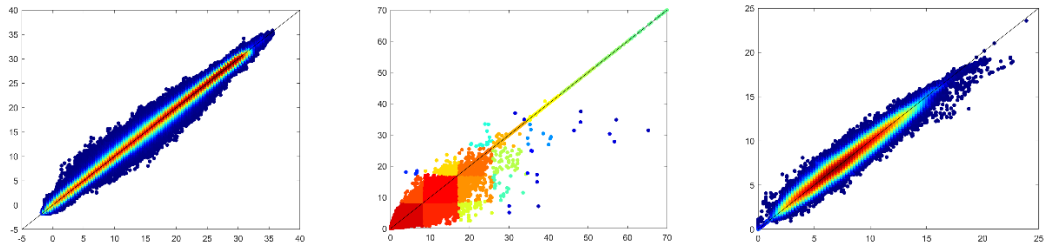
(e.g., high chlorophyll values). This would help clarify the physical mechanisms behind the reconstruction bias.

Response: First, based on the T-DINEOF reconstruction results, we generated separate scatter plots for SST, SCHL, and SSW, as shown in the first row of the figure below. It can be observed that the underestimated points mainly originate from the SCHL data. Based on this observation, we further analyzed the SCHL scatter plots for the three study subregions (second row). The results show that underestimated values mainly occur in the equatorial region (subregion 2) and the Northern Hemisphere mid-latitude region (subregion 1), whereas the SCHL scatter distribution in the Southern Hemisphere mid-latitude region (subregion 3) is comparatively well distributed. This indicates that the underestimated points in Fig. 7 mainly originate from the SCHL data in the Northern Hemisphere mid-latitude and equatorial regions.

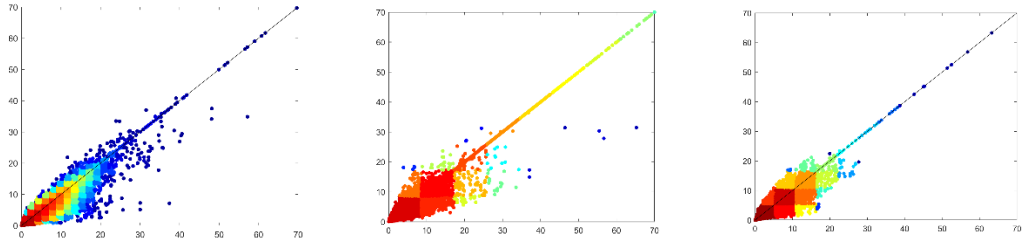
However, due to the relatively limited number of underestimated points, it is difficult for us to further determine whether they are associated with specific geographic regions or particular seasons. Moreover, we believe that a small number of underestimated points alone is insufficient to demonstrate a systematic bias of the reconstruction method for a certain region or season, since most reconstructed values do not exhibit underestimation and only a small fraction show underestimated behavior. We hope for the reviewer's understanding.

In the revised manuscript, we added the following statement.

"By analyzing the scatterplot distributions of SST, SCHL, and SSW, it was found that these underestimated values mainly originate from the SCHL data in subregion 1 and subregion 2, indicating that SCHL values may be underestimated in these two subregions."



SST SCHL SSW



Subregion 1 Subregion 2 Subregion 3

6. Figure 15 shows obvious periodic fluctuations in SST RMSE over time, particularly in subregion 1. Is this due to lower data availability in summer?

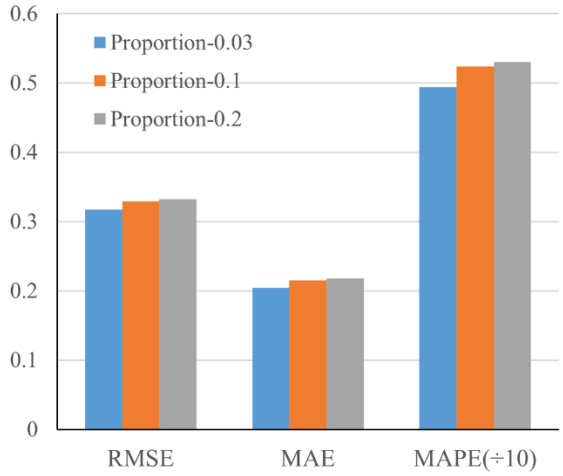
Response: We believe that the periodic variation in SST RMSE is caused by a mechanism similar to that of SCHL, namely the spatial homogeneity of the region.

During summer, SST generally exhibits higher homogeneity, and regions with higher homogeneity tend to produce lower reconstruction errors. In contrast, SST variability is stronger in winter, leading to larger reconstruction errors. Since this explanation has already been discussed in the SCHL section and would be largely repetitive, we chose to retain the discussion only in the SCHL section. We hope for the reviewer's understanding.

7. In Figure 18, the three colored lines (blue, red, grey) are difficult to distinguish. Please revise the figure using more distinct colors, line styles, or symbols to ensure clear visual separation of the three cross-validation proportions.

Response: In the revised version, we replaced the radar chart with a bar chart to better illustrate the differences in reconstruction accuracy. Accordingly, the corresponding text in the main manuscript has been revised as follows.

"Overall, the accuracies obtained under the three proportions are quite similar, while the 3% proportion yields relatively lower error values."



8. On page 10 (line 1), the manuscript states that 3% of existing pixels are selected as cross-validation pixels, while page 23 tests three proportions (3%, 10%, and 20%). Please clarify the rationale for selecting 3% in Section 3.2 (Methodology) rather than deferring this justification to the Discussion section. Readers should understand the basis for this choice when first encountering the method, not later in the paper.

Response: In the revised version, we replaced the corresponding content in Section 3.2 (Methodology) with the following text.

"The same cross-validation pixel selection strategy as in the original DINEOF method was adopted, i.e., 3% of the existing data from the corresponding spatiotemporal matrix of each oceanic variable were randomly selected as cross-validation pixels. In addition, we analyzed the impact of different cross-validation pixel proportions (see Discussion). The results indicate that the choice of proportion has only a minor effect on the reconstruction accuracy, and the 3% setting performs slightly better overall."