To Reviewer 2

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

1. The authors argue that large satellite-reanalysis discrepancies justify omitting independent validation. However, the manuscript mentions daily data experiments (Jan–Mar 2022, Sect. 4.2). Were daily reconstructed results validated against daily reanalysis or buoy data? Monthly averaging may amplify discrepancies; daily comparison would better assess whether the large RMSE reflects true reconstruction error or temporal resolution mismatch. I suggest including daily validation if available.

Response: We still believe that biases among different data sources can affect the accuracy assessment of the T-DINEOF and Multi-DINEOF methods. Therefore, comparing the reconstruction results with their corresponding original data sources can better reflect the actual performance of the algorithms. As discussed in the previous response, compared with the monthly reanalysis data, the reconstruction accuracy of both T-DINEOF and Multi-DINEOF decreased. Because the discrepancies between the original satellite observations and the reanalysis data are relatively large, the resulting errors reflect not only reconstruction uncertainty but also the biases between different datasets.

Similarly, we further evaluated the reconstruction accuracy of T-DINEOF and Multi-DINEOF using the daily average SST at 20 cm depth derived from the C3S global Sea Surface and Sea Ice Temperature Reprocessed product at 0.05° × 0.05° resolution (SST_GLO_SST_L4_REP_OBSERVATIONS_010_024), the 4-km daily chlorophyll-a product derived from the Copernicus-GlobColour processor (OCEANCOLOUR_GLO_BGC_L4_MY_009_104), and the CMEMS Global Ocean Ensemble Reanalysis daily velocity product at 0.25° resolution (GLOBAL_MULTIYEAR_PHY_ENS_001_031). Among these datasets, the SST and SCHL products were resampled to 0.25° resolution to ensure consistency with the SSW data and the reconstructed datasets.

Similar to the results obtained using the monthly datasets, the reconstruction accuracies evaluated with the daily reanalysis data were lower than those evaluated using the satellite observations, and were even lower than the accuracies obtained with the monthly reanalysis datasets. On the one hand, consistent with the conclusions derived from the monthly datasets, the discrepancies between the input satellite observations and the reanalysis products are inherently large. Consequently, the reconstruction errors evaluated against the reanalysis data mainly reflect the differences between data sources rather than the actual reconstruction capability of the methods. Since the reconstruction results of both T-DINEOF and Multi-DINEOF are more strongly constrained by the input satellite observations, the comparison with reanalysis data may not adequately represent the relative reconstruction performance of the two methods.

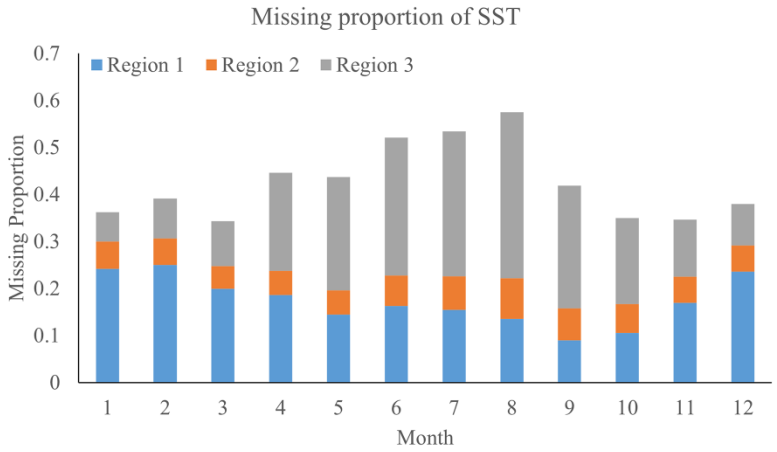
On the other hand, the daily datasets contain substantially higher missing-data

ratios than the monthly datasets. As a result, the reanalysis products are required to estimate a larger number of unknown grid values, which may introduce additional uncertainties and generate more erroneous estimations of oceanic variables. This is likely another reason why the daily datasets exhibit larger errors than the monthly datasets. 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	6.9053	5.2532	0.9002	0.5482
Multi-DINEOF	6.9124	5.2620	0.8994	0.5389
Sat-Reanalysis	6.9058	5.2512	0.9001	0.3248

2. I think there may be a misunderstanding in the explanation. The response states that summer SST exhibits "higher homogeneity" leading to "lower reconstruction errors," while winter has "stronger variability" causing "larger reconstruction errors." However, Figure 15a-c clearly shows higher SST RMSE in summer and lower RMSE in winter—the opposite pattern. I would suggest the authors verify whether the seasonal RMSE pattern aligns with missing data proportions.

Response: We found that in subregion 1, summer is generally not the period with the largest amount of missing data, as shown in the figure below. Therefore, similar to the case of SCHL, we believe that variations in the regional parameter itself are the primary cause of the RMSE fluctuations. In our previous response, we stated that the variability is smaller in summer and larger in winter; however, this interpretation appears to be opposite to the actual situation. We would like to apologize to the reviewer for this mistake.



We calculated the monthly standard deviation (std) distribution in subregion 1, as shown in the figure below. The results indicate that the std values are generally higher in summer and lower in winter, suggesting that SST variability is stronger in summer than in winter. In addition, we analyzed the correlation between RMSE and std values, which reached 0.39 in subregion 1. This further supports the conclusion that regional SST variability is a major factor driving the periodic variations in RMSE.

As mentioned in our previous response, in the SCHL section we also investigated the relationship between the periodic variations of RMSE and std. Therefore, to avoid

repetitive discussion in the revised manuscript, we did not include a similar analysis for the periodic SST variations. We hope the reviewer can understand this consideration.