To Reviewer 3

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

1. Validation design needs to be described much more clearly.

The manuscript should provide a fully transparent description of how masks were generated, how training and test data were separated, and whether the reconstruction was evaluated in a way that prevents spatial or temporal leakage. For geophysical fields, simple random masking can be misleading because nearby points are correlated, and performance may be overestimated if contiguous structures leak into the training set. The paper should state whether cross-validation was done by space, time, or independent scenes, and whether all methods were tested under identical missingness patterns. Without this, the reported skill gains are difficult to interpret.

Response: The DINEOF, Multi-DINEOF methods and T-DINEOF method proposed in this study are both adaptive reconstruction approaches. Unlike machine learning methods, they do not require training, validation, or test datasets; that is, these methods do not build a reconstruction model, but instead operate directly on the missing matrix/tensor, reconstructing it via matrix/tensor decomposition. For accuracy assessment, the DINEOF, Multi-DINEOF, and T-DINEOF methods all evaluate reconstruction performance by comparing the reconstructed fields with satellite observations at non-missing points. Since the non-missing points in the satellite data are fixed, the accuracy comparison among the three methods is based on the same reference locations. This procedure is also consistent with the standard practice commonly adopted in DINEOF-type methods. Therefore, there is no process of generating masks or splitting data into training and test sets in this study. Currently, DINEOF and its subsequent improved methods all adopt this approach for missing data reconstruction. Furthermore, although 3% of the data were selected as cross-validation points in this study, these points were primarily used to determine the optimal EOF modes rather than to assess reconstruction accuracy. This practice follows the original DINEOF method and its improvements, and the present study adopts the same strategy to determine the optimal modes.

2. The comparison set is likely too narrow for a strong methods paper.

Comparing only against Multi-DINEOF and DINEOF may not be enough to demonstrate that the new method is broadly competitive. Ocean reconstruction and gap-filling literature includes many alternatives, such as EOF-based interpolation variants, low-rank matrix completion, tensor completion approaches, machine-learning-based infilling, and hybrid physical-statistical methods. Even if not all can be included, the paper should justify the chosen baselines and explain why they represent the relevant state of the art. A stronger benchmark suite would make the contribution more convincing.

Response: The DINEOF method is widely used for ocean data reconstruction, and many improved versions have been proposed based on it. The purpose of this study is not to introduce a state-of-the-art ocean reconstruction method, but rather to propose a tensor decomposition-based DINEOF method built upon the existing DINEOF framework, aiming to further improve the reconstruction accuracy for multivariate datasets. We acknowledge that, with the development of artificial intelligence, many machine learning and deep learning methods may achieve higher reconstruction accuracy than DINEOF-type methods. However, DINEOF-type methods still have development potential due to their adaptive nature, which does not require labeled data, the design of an explicit model, or any prior knowledge or assumptions. Therefore, the purpose of this study is to demonstrate that the tensor decomposition-based DINEOF method achieves better accuracy than the current matrix-based Multi-DINEOF reconstruction and provides a new perspective for multivariate reconstruction, rather than to claim that DINEOF-type methods are superior to emerging data-driven or compressive sensing reconstruction approaches.

3. The physical meaning of the reconstructed fields must be demonstrated.

Lower RMSE alone does not prove that the method preserves oceanographically relevant features. The authors should show whether fronts, filaments, coastal gradients, eddy structures, and seasonal patterns remain realistic after reconstruction. For a paper in ocean science, it is important to assess not only pixel-wise accuracy but also whether the reconstructed fields are dynamically plausible and consistent with known ocean variability. Maps, anomaly fields, and regional examples would help substantially.

Response: We calculated the gradient maps of the SST, SCHL, and SSW data reconstructed by T-DINEOF, Multi-DINEOF, and Single-DINEOF, respectively. Typically, oceanographic features such as fronts and eddy structures are derived from gradient fields. 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."

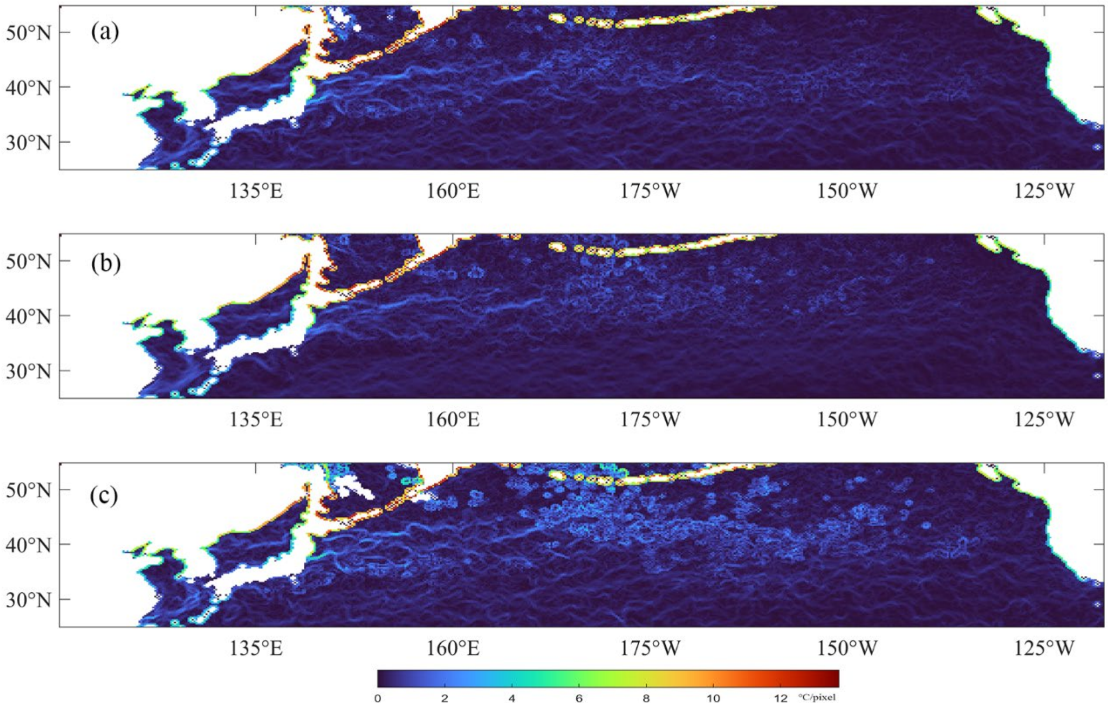

Fig. 1 Gradient maps of SST for April 2022 in the northern Pacific obtained from (a) Multi-DINEOF, (b) Single-DINEOF, and (c) T-DINEOF.

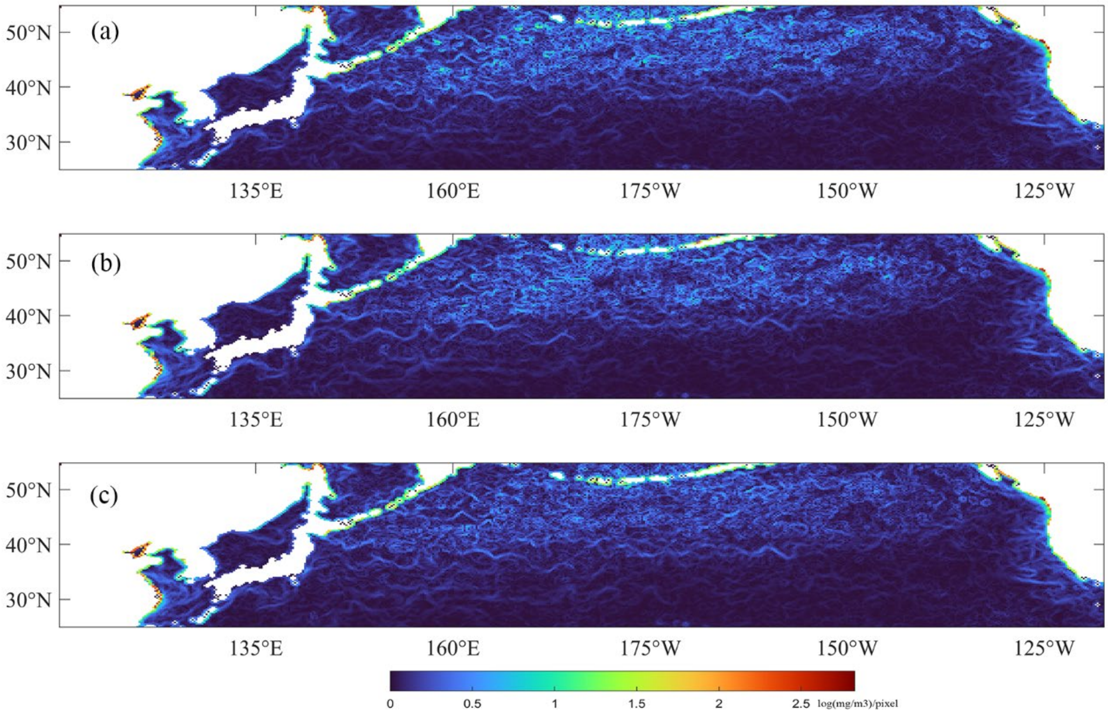

Fig. 2 Gradient maps of SCHL for July 2021 in the northern Pacific obtained from (a) Multi-DINEOF, (b) Single-DINEOF, and (c) T-DINEOF.

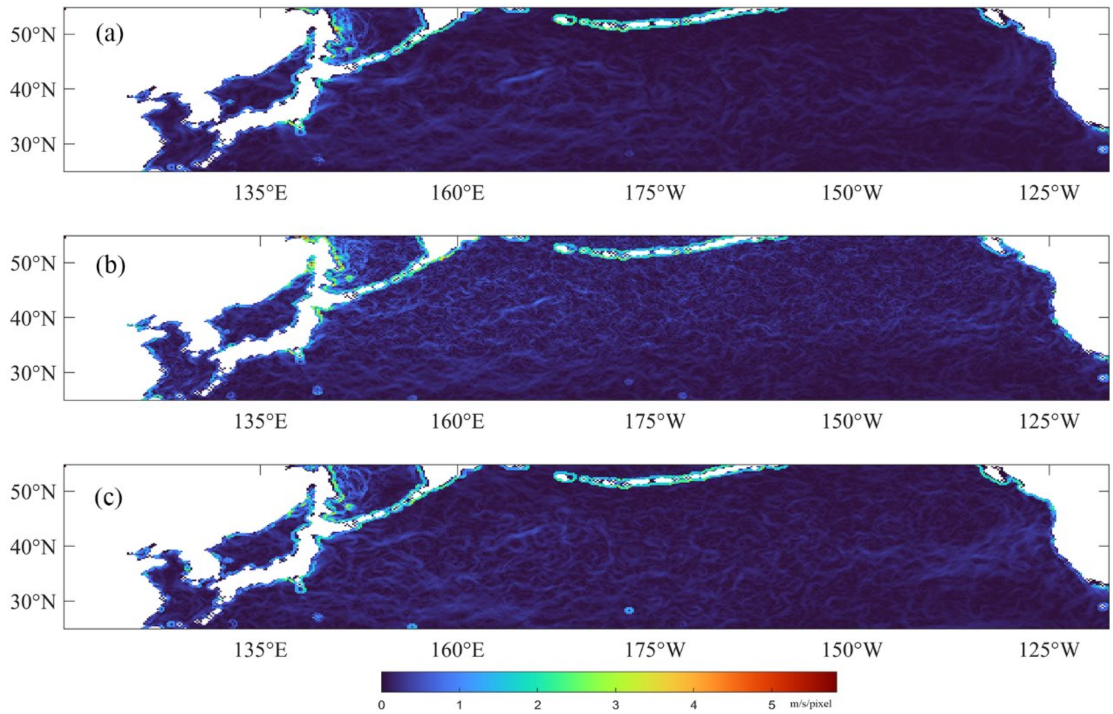

Fig. 3 Gradient maps of SSW for January 2020 in the northern Pacific obtained from (a) Multi-DINEOF, (b) Single-DINEOF, and (c) T-DINEOF.

4. The claim of improved performance in low-correlation regions needs stronger evidence.

The abstract states that T-DINEOF performs better in low-correlation scenarios, which is potentially interesting but also a demanding claim. Low-correlation regions are where multivariate methods can fail if the cross-variable relations are weak or nonstationary. The manuscript should specify how correlation was defined, how these regions were selected, and whether the improvement is statistically significant across multiple cases. If the result depends on a specific dataset or region, the authors should be explicit about that limitation.

Response: In this study, correlations were calculated directly based on the entire input three-dimensional tensor. The correlation coefficients between the input variables SST, SCHL, and SSW were -0.6 for SST–SCHL, 0.03 for SCHL–SSW, and -0.04 for SST–SSW, all of which were statistically significant ($p<0.01$), indicating that SSW has a relatively low correlation with both SST and SCHL. Previous studies have also shown that SST and SCHL are negatively correlated, while both SST and SCHL exhibit low correlation with SSW. Because DINEOF-type methods operate directly on the missing-value matrix/tensor, no specific local regions were selected for analysis, and the correlation calculation was performed over the entire study area. As a result, there are no multiple cases. Reconstruction results demonstrate that, compared with the conventional Multi-DINEOF method, the T-DINEOF method significantly improves the reconstruction accuracy for SSW. Therefore, it can be concluded that T-DINEOF achieves better reconstruction performance for input variables with low correlations.

5. Uncertainty and robustness are not sufficiently addressed.

For a reconstruction method, the paper should discuss sensitivity to missing-data fraction, noise, variable scaling, and tensor rank or hyperparameter choices. The reported improvements may depend strongly on the chosen decomposition rank, normalization strategy, or stopping criterion. A sensitivity analysis would make the method easier to trust and reproduce. At minimum, the paper should include parameter selection details and a robustness check over a range of conditions.

Response: DINEOF is an adaptive reconstruction approach, for which the only key parameter is the determination of the optimal number of modes. In this study, the Method Section provides a detailed description of how the optimal modes are determined, following a procedure similar to that of the original DINEOF and Multi-DINEOF methods. Furthermore, in the Discussion section, we report the specific optimal mode numbers and the corresponding RMSE values for both T-DINEOF and Multi-DINEOF. The main conclusion is that the two methods have similar optimal modes, indicating comparable convergence rates, while T-DINEOF achieves better reconstruction accuracy.

Regarding missing-data fraction and noise, we applied T-DINEOF to daily datasets with a high proportion of missing values (see Supplementary File S1), where the missing fraction exceeded 78%. Despite the high level of missing data, T-DINEOF still outperformed Multi-DINEOF in reconstruction accuracy. This conclusion further confirms the robustness of the T-DINEOF method and its effectiveness in reconstructing datasets with a high proportion of missing values.

As stated in the manuscript (Section 2, Materials), since the ranges of SST, SCHL, and SSW differ, a max–min normalization method was applied to scale each ocean variable (SCHL data were first log-transformed) in each subregion to the range [0, 1]. The maximum and minimum values of each ocean variable in each subregion are provided in Table 1.

During each decomposition, reconstruction is performed using a fixed number of modes, and the reconstruction accuracy is evaluated. If the accuracy does not meet the requirements, the reconstructed tensor is decomposed again using the same number of modes, and the accuracy is recalculated. This process is repeated until the desired accuracy is achieved (RMSE < 0.00001) or a stopping criterion is met (100 iterations, Section 3, Methodology), thereby completing the reconstruction for a given number of modes. The number of modes is then increased, and the process is repeated. The number of modes that yields the highest accuracy is selected as the optimal mode for reconstruction. This procedure is similar to that of the original DINEOF and Multi-DINEOF methods. In the Discussion section, we also compare the optimal mode numbers between T-DINEOF and Multi-DINEOF. Therefore, the tensor decomposition process does not involve a specific design or choice of tensor rank.

In addition, this study also analyzed the impact of different cross-validation fractions on reconstruction accuracy (see Fig. 18 in the Discussion). The results confirm that varying the fraction of cross-validation points does not significantly affect the reconstruction accuracy, and the optimal number of modes remains similar. The 3% fraction chosen in this study—also used in the original DINEOF and its improved versions—yields slightly higher accuracy compared to other fractions. Therefore, we retained the 3% cross-validation fraction in this study.

Table 1 Maximum and minimum values of SST, SCHL and SSW in the three subregions.

	SST (°C)		SCHL (log(mg/m^3))		SSW(m/s)	
	Max	Min	Max	Min	Max	Min
Subregion 1	35.65	-1.50	4.05	-4.16	27.05	0
Subregion 2	33.57	0	4.18	-4.56	15.26	0
Subregion 3	31.83	-1.81	3.32	-5.16	24.91	0

6. The treatment of variables with different units and distributions requires careful justification. SST, chlorophyll-a, and wind speed have very different magnitude ranges, statistical distributions, and physical drivers. The manuscript should explain how these variables were normalized and whether the reconstruction operates on standardized anomalies, log-transformed values, or raw

fields. If the method jointly reconstructs variables on different scales without careful preprocessing, one variable may dominate the tensor decomposition. This is an important methodological detail that should be clarified.

Response: As mentioned in response to the previous comment, the normalization methods for different variables are described in Section 2, Materials. Specifically, a max–min normalization was applied to each ocean variable in each subregion. For SCHL, the data were first log-transformed and then normalized using the max–min method, while SST and SSW were directly normalized without transformation. The maximum and minimum values for each variable in each subregion are provided in Table 1. The reconstruction process is based on these normalized data. The original manuscript states: “Since the ranges of SST, SCHL, and SSW differ, a max–min normalization method was applied to scale each ocean variable in each subregion to the range [0, 1]. The maximum and minimum values of each ocean variable in each subregion are provided in Table 1.”

7. Computational cost and scalability should be reported.

Tensor methods can be more expensive than matrix-based approaches, especially for large spatial grids or long time series. The paper should provide runtime, memory demand, convergence behavior, and the practical limits of the method. This matters for operational oceanography and large satellite archives. If the method is meant as a practical alternative, the computational overhead must be discussed honestly.

Response: In this study, the T-DINEOF method was developed based on Matlab R2023b. During the algorithm development, we focused primarily on improving reconstruction accuracy rather than computational efficiency. Therefore, it should be acknowledged that there is room to optimize the execution speed of the code. Furthermore, tensor operations involved in T-DINEOF, such as T-SVD decomposition and tensor transposition, were independently implemented based on the tensor definitions. Steps such as data preprocessing, land masking, and selection of cross-validation pixels were also integrated into the execution code. As a result, it is currently difficult to provide absolute computation times for T-DINEOF.

In our experiments, we confirmed that T-DINEOF requires longer computation times compared with matrix-based methods, even though they have similar computational complexity, likely because tensor operations are inherently more computationally demanding than matrix operations. As shown in Figures 17 and 18 in the Discussion, the optimal number of modes is typically between 30 and 40, indicating that both the T-DINEOF and Multi-DINEOF methods can reach convergence relatively quickly. Since the code was independently developed without extensive optimization for execution efficiency, we are unable to provide absolute runtime comparisons, and can only offer relative time comparisons.

The code was executed on an Intel Xeon W-2223 CPU with 128 GB of memory, capable of performing computations for the three subregions ($119{,}939\times91\times3$; $174{,}091\times91\times3$; $207{,}779\times91\times3$) as well as the combined full-domain reconstruction.

This issue has already been addressed in the manuscript.

“However, due to the tensor operations involved, T-DINEOF requires longer computation times to reconstruct the same region. This may limit its application to larger-scale tensors, such as longer time-series images (91 scenes were used in this study) or tensors with more dimensions (three variables in this study).”

8. The manuscript should better explain novelty relative to Multi-DINEOF.

The paper should clearly state what is genuinely new in T-DINEOF beyond “tensorizing” a known matrix method. If the novelty is mainly mathematical, the authors should articulate why the tensor formulation changes the reconstruction behavior. If the novelty is algorithmic, they should describe the specific improvement in a way that readers can reproduce. At present, the contribution may read as a natural extension rather than a fully established methodological advance.

Response: We consider the main innovation of this study to be the extension of the widely used matrix-based DINEOF method to a tensor-based formulation. In current oceanographic studies, DINEOF-type methods are commonly employed to reconstruct missing data, making reconstruction accuracy crucial for subsequent data applications. Nearly 20 years since its introduction, the DINEOF method remains fundamentally matrix-based. Whether this framework can be improved or enhanced is the key question that our study aims to address. By extending DINEOF to the tensor domain, this work not only improves reconstruction accuracy but also provides a more reliable data foundation for oceanographic research.

As illustrated in Fig. 1, the Multi-DINEOF framework essentially remains a matrix-based decomposition method. It flattens multiple variables into a concatenated two-dimensional matrix, which inevitably obscures the intrinsic coupling relationships among variables, since different parameters (e.g., SST, SCHL, and SSW) are forced to be represented along the same dimension. As a result, the extracted principal components may mix heterogeneous signals, making it difficult to identify the true inter-variable relationships and thus reducing the physical interpretability of the results. Moreover, matrix-based approaches can only characterize pairwise (i.e., second-order) correlations, making them insufficient for capturing higher-order interactions across spatial, temporal, and variable dimensions.

In contrast, tensor decomposition represents data as a higher-order array (e.g.,space×time×variable), thereby preserving the native multi-dimensional structure of the dataset. This framework allows simultaneous extraction of spatial, temporal, and inter-variable modes, maintaining the inherent coupling among dimensions. Tensor-based models, such as T-SVD decompositions, can reconstruct missing data more accurately by leveraging multi-dimensional constraints, leading to more robust and physically consistent results. Furthermore, each decomposed mode corresponds to a specific physical meaning (e.g., spatial pattern, temporal variation, or variable contribution), enhancing the interpretability of the analysis. Therefore, tensor decomposition provides a more natural and efficient framework for representing and reconstructing complex, multi-variable oceanic or environmental systems.

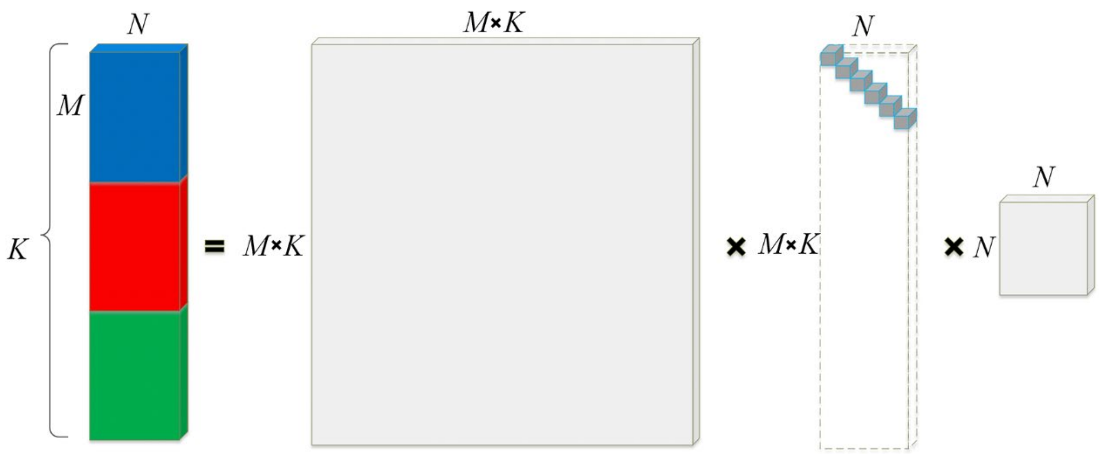

Fig. 1 Matrix decomposition in the Multi-DINEOF method.

Minor comments

9. The abstract is concise, but the manuscript should ensure that the main novelty is stated in the introduction in a way that is understandable to readers outside the immediate DINEOF community.

Response: In the Introduction, we have described the limitations of the DINEOF-type methods and presented the main ideas and innovations of the proposed T-DINEOF approach using both figures and text. Considering the page limit (currently 31 pages, 10,704 words), we did not provide further details on DINEOF-type methods, as this is not intended to be a review article. Nevertheless, we believe that in the current version, the main innovations of the study have been clearly presented in the Introduction.

10. The terminology for the variables should be defined consistently throughout the paper, especially if abbreviations are used for SST, SCHL, and SSW.

Response: We have carefully checked the terminology used throughout the manuscript and did not find any inconsistencies. If any discrepancies exist, they may have been an oversight on our part. We would appreciate it if the reviewer could kindly point out any specific instances of inconsistent terminology so that we can make the necessary corrections.

11. The paper should include more visual comparison figures, ideally with original, corrupted, reconstructed, and error maps shown side by side.

Response: We selected SST data from April 2022, SCHL data from July 2021, and SSW data from January 2020 for visual comparison (Figures 9, 11, and 12). Each of these three figures includes the original, reconstructed, and error maps, presented side by side. As mentioned previously, since DINEOF-type methods do not require generating a mask, there is no corresponding corrupted map; in this case, the original map itself serves as the "corrupted" map containing missing values. Due to space limitations, it is not feasible to show all 91 scenes. In the revised version, we have also included the reconstructed gradient maps to further enhance the visual comparison figures.

12. Metric definitions should be stated explicitly, including any averaging over space, time, or multiple scenes.

Response: We have added the relevant explanations in the Metric Definitions section of Section 3.3 (Methodology). The specific content is as follows:

"where R_{exist} and O_{exist} signify the reconstructed and original values at the existing pixels within all images in the study area during the target period, respectively, and l_{exist} indicates the number of existing pixels."

13. If the method is applied to satellite products, the preprocessing chain, cloud masking, and regridding procedure should be documented carefully.

Response: We used remote sensing data products distributed by Ocean Color and Remote Sensing Systems, rather than direct satellite observations. The received data products have already undergone preprocessing procedures such as cloud masking and resampling. On one hand, these procedures are well documented and can be directly accessed on the respective websites; on the other hand, they are not the focus of this study and are only marginally related to our research. Considering the page limit, we did not include further details on these preprocessing steps. However, the principles behind the generation of the SST, SCHL, and SSW data relevant to this study have been described in the main text.

14. The discussion should note the main limitations of the approach rather than only emphasizing improvement.

Response: In the second paragraph of the Discussion, we have provided a detailed description of the limitations of the proposed T-DINEOF method, including the use of a single optimal mode, the lack of consideration for temporal correlations, and the computational cost. The specific content is as follows:

"Although T-DINEOF outperforms the original DINEOF and Multi-DINEOF methods, it still relies on a single global optimal mode, which may not be ideal for capturing heterogeneous local dynamics within the study area or achieving the best reconstruction in specific regions and during intermediate iterations. Therefore, integrating improvements from original DINEOF methods and their variants into third-order tensor reconstruction is necessary. Furthermore, this study uses monthly oceanic variables, so temporal correlations are not considered. If daily or weekly data were used, incorporating temporal correlations among various oceanic variables would be essential for enhancing reconstruction accuracy. Additionally, compared to monthly data, daily or weekly datasets typically exhibit higher proportions of missing values. In this study, we demonstrated that T-DINEOF outperforms both Multi-DINEOF and Single-DINEOF in regions with high missing data proportions, suggesting that it may be more advantageous for reconstructing daily or weekly data (Fig. S2). However, due to the tensor operations involved, T-DINEOF requires longer computation times to reconstruct the same region. This may limit its application to larger-scale tensors, such as longer time-series images (91 scenes were used in this study) or tensors with more dimensions (three variables in this study). It is evident that, over the same time span, the volume of daily or weekly data is significantly greater than that of monthly data. Given the substantial data volume inherent in a third-order tensor, optimizing computational processes and accelerating processing speeds are crucial for the effective application of T-DINEOF to high-temporal-resolution datasets. Moreover, to unify MODIS SST and SCHL (high resolution) with AMSR2 SSW (low resolution), the nearest-neighbor method was used to downscale MODIS data to the AMSR2 resolution. Upscaling low-resolution data is an ill-posed (NS) problem that may introduce larger errors, whereas downscaling helps reduce error propagation. Therefore, MODIS SST and SCHL data were downscaled to the AMSR2 resolution rather than upscaling AMSR2 data to match the MODIS resolution. However, although the nearest-neighbor method is widely used for downscaling high-resolution imagery, it may still introduce artifacts that affect the reconstruction process. Consequently, adopting more advanced resampling techniques or using multi-source datasets with consistent spatial resolution could further improve reconstruction accuracy. Finally, it should be noted that if the source datasets contain systematic biases, such errors may be propagated through the reconstruction process, potentially affecting the accuracy of the final results. In particular, tensor-based methods may be more susceptible to such bias propagation, as errors in one variable can influence others through the coupled decomposition."

Main questions

1. How were the missing-data masks generated, and do they preserve realistic cloud-gap or sampling-gap structure?

2. Were the test regions or time periods completely independent from the data used to train the decomposition?

3. How sensitive is the method to tensor rank, normalization, and initialization?

4. Does the method remain advantageous when variables are weakly correlated or when one variable is much noisier than the others?

5. Can the authors demonstrate that the method preserves mesoscale and coastal features, not just global error statistics?

6. How does T-DINEOF compare in runtime and memory use to Multi-DINEOF and DINEOF?

7. Would the method generalize to other ocean variables or other basins beyond the example shown?

Response: For responses to the above issues, please refer to the explanations provided above.