https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Open Access Earth System Science Data Discussions

Towards accurate daily 3D temperature and salinity reconstruction from remote sensing enhanced by explainable AI

Xihong Fang[1], Yujiao Zheng[1], Haochen Sun[1], Siming Huang[1], Jiangnan He[1], Wenfang Lu[1,2,*] Young-Heon Jo[3]

[1]Sun Yat-Sen University, Guangzhou, 519082, China
[2]Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, 519082, China
[3]Department of Oceanography and Marine Research Institute, Pusan National University, Busan, 46241, Republic of Korea.

Correspondence to: Lu Wenfang (luwf6@sysu.edu.cn)

Abstract. Using remote sensing data to reconstruct three-dimensional (3D) temperature and salinity, a field called Deep Ocean Remote Sensing (DORS), is essential for the study of ocean dynamics and climate change. However, existing DORS studies predominantly focus on monthly scales and lack explainability, leaving the mechanisms governing reconstruction errors poorly understood and hindering daily-scale operational applications. Here we report a transformer-based framework (i.e., EarthFormer) to reconstruct daily 3D temperature and salinity fields from multi-source remote sensing inputs, across 19 standard depth levels in the Northwest Pacific (NWP) (105–160°E, 0–40°N) at 0.25° resolution using reanalysis model product as the labeled data. Validated against Argo observations, the reconstruction achieves an root-mean-squared error of 0.893°C (R^2 = 0.989) for temperature and 0.141 PSU (R^2 = 0.827) for salinity, approaching reanalysis accuracy while offering near-real-time timeliness and lightweight computation. Notably, explainable AI analysis reveals that the counter-intuitively low contribution of satellite-derived sea surface salinity (SSS) stems not from weak physical relevance but from data-quality limitations; substituting high-accuracy SSS shifts the salinity error profile from a monotonic depth decrease to a V-shaped structure, with SSS contribution rising from ~10% to ~50%. Overall, this study demonstrates the feasibility of explainable-AI-enhanced daily 3D thermohaline reconstruction, providing a new technical pathway for real-time ocean monitoring and underscoring that improving satellite SSS retrieval and data quality is as critical as advancing model architectures for reliable DORS applications. The data are publicly available at https://doi.org/10.5281/zenodo.20602639 (Fang, 2026).

1. Introduction

Understanding the vertical distribution of sea temperature and salinity is increasingly critical for elucidating the complex dynamic processes and ecosystems of the ocean (Arribas et al., 2011; De Boyer Montégut et al., 2007; Helber et al., 2010; Meehl et al., 2011; Qin et al., 2015; Wilson & Coles, 2005). With the rapid development of satellite remote sensing technology, the volume of ocean observation has grown substantially (O'Carroll et al., 2019; Yan et al., 2016). However, these data are limited to the ocean surface, leaving the spatial and temporal continuous information unavailable for the ocean

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

interior (Ali et al., 2004; Bao, Zhang, et al., 2019; Hu, 2025; X. Wu et al., 2012). Previous studies have demonstrated that the subsurface feature can be empirically or dynamically related to and thus be reconstructed from remote sensing data (Z. Chen et al., 2020; Klemas & Yan, 2014). For example, as early as 1983, Khedouri et al. (1983) found that there is a strong relationship between subsurface temperature and the variability of sea surface height in the Gulf Stream region. Chu et al. (2000) used sea surface temperature to estimate subsurface temperature in the South China Sea. These findings provide a theoretical basis for inferring ocean interior structures from surface observations. In literature, this topic is often called Deep Ocean Remote Sensing (DORS), rigorously enhancing the values of remote sensing products under the impenetrable sea surface.

Recently, with the development of high-accuracy ocean reanalysis datasets such as GLORYS (Jean-Michel et al., 2021) and HYCOM (Chassignet et al., 2007), along with the widespread application of deep learning methods to nonlinear mapping problems, data-driven three-dimensional (3D) ocean reconstruction has become a research hotspot (L. Meng & Yan, 2022; Su et al., 2023), particularly at the monthly scale (Y. Chen et al., 2023). These approaches have been successfully applied to the reconstruction of variables such as temperature (Garcia-Gorriz & Garcia-Sanchez, 2007; Su, Zhang, et al., 2021; H. Wang et al., 2021; X. Wu et al., 2012), salinity (Xie et al., 2023), ocean heat content (Su et al., 2020), ocean circulation (W. Wu et al., 2023), and biogeochemical variables (Jo et al., 2012; Landschützer et al., 2013; H. Li & Ilyina, 2018; Parard et al., 2015). However, existing studies are predominantly conducted at the monthly scale, typically relying on Argo gridded data with a spatial resolution of 1° to establish mappings between surface and subsurface structures (Cheng & Zhu, 2016; Good et al., 2013; Roemmich et al., 2019; Zuo et al., 2019). Constrained by the spatiotemporal resolution and sampling density of observations, such methods exhibit clear limitations in capturing short-timescale variability and are therefore insufficient for applications requiring high timeliness, such as ocean environmental forecasting.

Practically, there exist three key differences between daily-scale and monthly-scale 3D reconstruction. Firstly, regarding application scenarios, the primary uses for daily-scale reconstruction are marine environment forecasting and operational support, which demand high data timeliness. Under high temporal resolution requirements, finer spatial resolution is also necessary (Fouchet et al., 2025; Wen et al., 2024). Thus, the 1-degree resolution commonly used for monthly scales often fails to meet practical needs (T.-T.-T. Chau et al., 2024). In such a scenario, daily and 0.25-degree resolution are the minimum requirements (Garcia-Espriu et al., 2025; B. Tang et al., 2022), some applications even achieved 1/12-degree (Cao et al., 2024). For the training data, global reanalysis products are idealized options. Finally, in terms of model training, Transformer-based architectures (self-attention networks) have become the mainstream choice for monthly-scale reconstructions. However, for daily-scale reconstruction with higher spatial resolution, the computational complexity of Transformer-based models increase quadratically, leading to very huge computation burden. Seeking a lightweight way to incorporate the nonlinear fitting capabilities is urgent. In summary, there are essential differences between 3D reconstruction of ocean structures on the daily scale and monthly scale. Therefore, we need to consider the daily scale 3D reconstruction from a new perspective.

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

One key aspect is to understand the distribution of uncertainties. Current DORS studies generally suggest that the vertical accuracy profile of temperature reconstruction exhibits a V-shaped structure. Specifically, the error is relatively small near the surface. The maximum error occurs in the thermocline, where complex dynamics associated with air sea interaction, vertical mixing, frontal processes, and mesoscale disturbances increase the difficulty of model learning. In deeper layers, temperature variability becomes weaker and the field is smoother, leading to lower reconstruction difficulty and smaller errors (Lu et al., 2019; Q. Wang et al., 2025). By contrast, salinity reconstruction studies show that, despite improvements in multiple aspects including neural network architecture, input feature combinations, and spatiotemporal resolution, accuracy often remains largest near the surface, followed by the mixed layer, and smallest in deeper layers (Liang et al., 2026; L. Meng et al., 2021; A. Wang et al., 2024; J. Zhang et al., 2023) . This shape is not consistent with either common sense or physical understanding. Since the surface layer is directly constrained by observational information, and sea surface salinity is commonly included as an important input variable, the surface reconstruction error should not theoretically remain persistently the largest.

Therefore, this study aims to bridge this gap, trying to better understand the error of temperature and salinity reconstruction with an explainable deep learning approach. We adopt the spatiotemporal attention network, EarthFormer, as the model architecture, using multi-source sea surface remote sensing products as input features and GLORYS reanalysis data as the outputs. Multiple sensitivity experiments are designed to train and test the model under different input feature combinations. The optimal feature set is selected for 3Dtemperature and salinity reconstruction in the Northwest Pacific (NWP). The reconstructed fields are then independently evaluated using Argo float observations. We then use SHAP (SHapley Additive exPlanations) analysis, on this optimized reconstruction network to quantify and explaining the influence of different SSS data sources on salinity errors, improving the understanding of the error formation mechanisms of DORS reconstruction.

2. Methods

2.1 Study region

The study focuses on the NWP (105°E–160°E, 0°N–40°N), a region characterized by complex bathymetry and diverse marginal sea systems (Cha et al., 2023; Kim et al., 2021; Seo et al., 2014). The western part of the region contains broad continental shelf system, while the eastern side faces open Pacific Ocean bordered by the Mariana Trench, the deepest trench on the Earth. This pronounced east–west bathymetric contrast forms the fundamental framework governing regional ocean circulation and the distribution of temperature and salinity, constraining the pathways of water masses and major current systems. The study region also encompasses several distinctive marginal seas, including the South China Sea, the East China Sea, and the Yellow Sea. The South China Sea is a semi-enclosed deep basin with an independent deep circulation system and active vertical exchange processes, while its upper-layer waters interact strongly with the Kuroshio through the Luzon Strait. In contrast, the East China Sea and the Yellow Sea are characterized by extensive continental shelves and are strongly

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

influenced by riverine inputs from the Yangtze and Pearl Rivers, monsoonal forcing, and Kuroshio intrusion. In addition, widespread thermohaline fronts and pronounced seasonal stratification are present throughout the region, resulting in substantial spatial variability in oceanographic characteristics among different marginal seas.

Research on ocean 3D reconstruction using deep neural networks has achieved good results, but the reconstruction remains challenging in areas with strong currents and complex dynamic processes, such as the Kuroshio Extension, as well as in marginal seas influenced by complex topography and coastal processes (Xie et al., 2022; Zhu et al., 2025). This suggests that 3D reconstruction in the northwestern Pacific must explicitly incorporate bathymetric constraints and circulation dynamics.

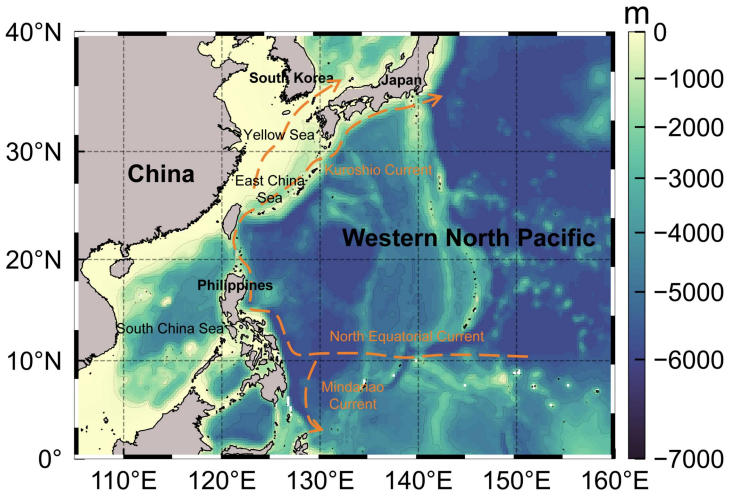

Figure 1. Bathymetric and circulation overview of the NWP (105–160°E, 0–40°N).

2.2 Data

A variety of physical processes influence the 3D thermohaline structure of the ocean. Previous studies commonly adopt oceanic and atmospheric driving variables as primary input features, such as sea surface temperature (SST), sea surface salinity (SSS), absolute dynamic topography (ADT), and wind fields. In addition, mixed layer depth (MLD), ocean currents, and surface heat fluxes have also been recognized as important factors affecting ocean thermohaline structures. However, from a practical application perspective, most 3D ocean variables rely on reanalysis products, while surface heat fluxes are typically derived from atmospheric model outputs, making real-time acquisition of these variables difficult. Therefore, additional thermodynamic control variables were not incorporated in this study. This design choice considers the application scenario of daily-scale reconstruction: once the model is trained, the 3D thermohaline structure of the ocean can be rapidly and accurately inferred using a limited number of readily available and relatively accurate surface features. Rather than simply reproducing the outputs of reanalysis models such as GLORYS, the proposed approach is intended to support near-real-time to real-time applications in operational ocean nowcasting and forecasting, rather than historical-period reconstructions where reanalysis products generally have advantages.

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Based on the above considerations, and in conjunction with the discussion of environmental characteristics of the northwestern Pacific in Section 2.1, the input features used in this study include SST and its gradient which has been demonstrated important by Xie et al. (2025), ADT and its gradient, SSS, longitude, latitude, wind speed, wind stress curl (WSC), and bathymetry (See Table 1). All input variable data are at a daily scale with a spatial resolution of 0.25° consistent with the 3D data after linear bias correction. Because the input variables have different physical units and value ranges, a Z-score standardization was applied to all input features by subtracting the global mean and dividing by the global standard deviation.

$$X = \frac{x-\mu}{\sigma}, \quad (1)$$

The GLORYS dataset was used as the reference dataset for model training, as it has demonstrated superior performance compared with several global ocean forecasting systems in standardized evaluation studies. For the 3D reconstruction, only selected standard depth levels were used as output variables. Temperature and salinity values at 19 standard depth levels (5, 10, 19, 31, 41, 53, 78, 108, 120, 133, 147, 200, 245, 301, 412, 509, 697, 857, and 1046 m) were chosen as model outputs, and two independent models were trained separately to reconstruct temperature and salinity. Considering the stronger variability in the upper ocean compared with deeper layers, the output variables were not standardized. Daily data from 2010 to 2020 (11 years) were used for model training and cross-validation, while data from 2021 to 2023 were reserved as an independent test set.

Beyond comparison with GLORYS, a more critical evaluation involves comparing the results with observational data. This is essential for verifying the accuracy of both the reanalysis data and the reconstruction results, ensuring the final reconstructed fields are suitable for practical quasi-real-time applications. Approximately 28,000 Argo profiles (including those with missing values) within the study area from 2021-2023 were acquired. The raw profile data were interpolated onto the same vertical standard levels as the reconstruction results. Subsequently, the daily GLORYS data and the 3D reconstruction results were linearly interpolated to the geographic coordinates of each Argo profile for comparison.

Table 1. Dataset used in this study

Abbreviation	Full name and source	Feature	Description
GLORYS	Global Ocean Physics Reanalysis https://data.marine.copernicus.eu/products. (GLOBAL_MULTIYEAR_PHY_001_030)	3D temperature and salinity	As the ground truth for training and testing the reconstruction model. Only 19 standard levels (5, 10, 19, 31, 41, 53, 78, 108, 120, 133, 147, 200, 245, 301, 412, 509, 697, 857, 1046 m).
AVISO	Archiving, Validation and Interpretation of Satellite Oceanographic data https://www.aviso.altimetry.fr/en/data.html	ADT	As the input features. The bias between ADT and GLORYS ADT needs to be corrected.
OISST	Optimum Interpolation Sea Surface Temperature https://www.ncei.noaa.gov/products/optimum-interpolation-sst	SST	As the input features, the bias between OISST and GLORYS SST needs to be corrected.
ERA5	ECMWF Reanalysis v5 https://cds.climate.copernicus.eu/	U10, V10, WSC	As the input features. The bias between ERA5 and GLORYS U10, V10, WSC needs to be corrected.
MSSS	Multi-source product merged from multiple satellites	SSS	As the input features, the bias between MSSS and

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

	Copernicus Marine Service https://data.marine.copernicus.eu/product/MULTIOBS_GLO_PHY_S_SURFACE_MYNRT_015_013/description		GLORYS SSS needs to be corrected.
Argo	Array for the Real-time Geostrophic Oceanography https://argo.ucsd.edu/	3D temperature and salinity	Profile data with Quality Control label = 1, i.e., highest quality

2.3 Neural network

Concerning the neural network, the need for high resolution typically demands strong nonlinear learning capabilities, which in turn incur significant computational costs. Current mainstream Transformer (self-attention) based networks often struggle to handle such high computational demand scenarios.

As mentioned earlier, the self-attention (i.e., Transformer) network (Vaswani et al., 2023) has been extensively developed and applied in recent years and now serves as the foundational architecture for large language models. Mirroring the revolutionary impact of the Transformer architecture in natural language processing, the Vision Transformer has also garnered significant attention recently and is increasingly being utilized in diverse oceanographic applications, including ocean remote sensing. This mechanism was initially proposed to capture global dependencies within the input and was applied to domains like machine translation, which shares similarities with time series forecasting. By computing the covariance matrix of the input data as a similarity measure, the self-attention mechanism constructs a weighting function that assigns higher importance (hence the term "attention") to more critical input information. In self-attention, these weights are derived from the input data itself through learnable parameters, rather than using predefined parameters (the "self" attention). The resulting attention weights are then used to compute a weighted sum of the values, producing a vector that captures the relationships between each element in the input. Consequently, the self-attention mechanism can capture long-range spatiotemporal features that are crucial for the target prediction.

A notable drawback of the self-attention mechanism is its computational complexity. Consider a typical 3D reconstruction problem, temporarily ignoring that the output is a 4D spatiotemporal matrix and viewing it instead as a mapping from a spatiotemporal 3D sea surface matrix (C, H, W) to another 3D temperature/salinity matrix with an equivalent number of channels (C, H, W). For such input and output scales, global attention—where each point simultaneously considers information from all other points—has a computational complexity of $O(C^2H^2W^2)$, leading to a quadratic increase in computational cost as the problem size grows. Therefore, directly applying a point-wise spatiotemporal self-attention network is nearly infeasible for problems involving numerous features or vertical levels.

To overcome the high computational cost of global attention, a common strategy involves partitioning the image into smaller patches, such as dividing it into 16x16 sub-images in 2D (Dosovitskiy et al., 2021), or employing various strategies for spatiotemporal 3D partitioning (Bi et al., 2023), before performing attention calculations. Alternatively, the Earthformer (Gao et al., 2023) adopts a different strategy, effectively addressing spatiotemporal sequence prediction problems in Earth systems, including weather, ocean, and climate forecasting tasks. EarthFormer is based on a hierarchical Transformer

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

composed of an encoder and decoder. Its core idea involves decomposing the input matrix into smaller units called "Cuboids." Each cuboid is treated as an independent sample, and information is extracted via Cuboid Attention. The EarthFormer architecture follows three steps: Decomposition, Attention, and Aggregation. The input spatiotemporal sequence is a sea surface matrix with multiple channels (C, H, W). To reduce computational complexity, EarthFormer decomposes it into a set of non-overlapping cuboids (b_C, b_H, b_W). The computational complexity after decomposition is (1), i.e., approximately 1-2 orders of magnitude lower than that of global self-attention. This approach uses an "axial" decomposition method, meaning the spatiotemporal field data is decomposed along the three dimensions into three types of minimal cuboids: $(C, 1, 1)$, $(1, H, 1)$, and $(1, 1, W)$. The self-attention mechanism is applied to each cuboid separately, and the outputs from all cuboids are then aggregated to produce the final result.

$$O\left((b_c b_H b_W)^2 \frac{C}{b_C}\frac{H}{b_H}\frac{W}{b_W}\right) \approx O(b_c b_H b_W CHW), \quad (2)$$

Research has shown that EarthFormer outperforms other state-of-the-art spatiotemporal prediction neural networks on multiple benchmark datasets, including ConvLSTM (Shi et al., 2015) and U-Net (Ronneberger et al., 2015), which are two types of deep learning networks widely used in ocean science and remote sensing. Furthermore, EarthFormer has already been applied in our recent work such as wave forecasting (Y. Liu et al., 2024), STA forecasting (He et al., 2026) and the reconstruction of 3D ocean ecological structures (Yang et al., 2024).

The implementation of the EarthFormer-based reconstruction framework is available at https://github.com/Xihong-tomato/ESSD-Github-model, including the training, inference and evaluation scripts, configuration file, utility and model modules, and the trained temperature and salinity model weights.

2.4 Experiment design

Various sets of experiments were designed to serve two main purposes: First, starting from a relatively common feature combination in 3D reconstruction studies (longitude/latitude, SST, SSS, and ADT), we incrementally add various new features. These include water depth, which reflects static controlling factors of ocean dynamics, and dynamic controlling factors representing more dynamical processes (wind fields, SST gradients, and ADT gradients). The selection of these factors primarily considers elements less frequently addressed in past studies. After identifying the optimal feature combination through sensitivity analysis, we performed daily-scale 3D reconstruction of sea temperature and salinity based on this combination. Experiment 4-G is designed to analyze the reason for salinity error profiles, while the rationale of such design will be shown in Section 4.1.

Table 2. Sensitivity experiments

Case ID	Feature combination	Description
Case 0	SST, SSS, ADT, Lon, Lat	This combination can be routinely seen in many 3D reconstruction studies, reflecting dynamic, thermodynamic, and spatial information.
Case 1	Case 0 + H	The total depth
Case 2	Case 1 + U10, V10	/

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Open Access Earth System Science Data Discussions

Case 3	Case 2 + WSC	Wind-stress curl
Case 4	Case 3 + log(|▽SST|) + log(|▽ADT|)	The gradient calculated with corresponding feature
Case 4-G	Case 4 with GLORYS SSS replacing satellite-derived SSS	Same as Case 4, but satellite-derived SSS is replaced by GLORYS SSS. This experiment is designed to evaluate the impact of SSS data sources on salinity reconstruction and to quantify the change in SSS contribution using SHAP analysis.

To quantitatively evaluate model performance, three statistical metrics were used, including root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination (R2). These are defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}\ , \tag{3}$$

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n}\ , \tag{4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}\ , \tag{5}$$

2.5 SHAP analysis

By conducting a detailed feature contribution analysis of the model results, this study aims to further improve the predictive accuracy of the model and identify key influencing factors. To achieve this, SHAP is used in this study to evaluate the contribution of each input feature to the final prediction (Lundberg & Lee, 2017).

The core idea of SHAP is derived from the Shapley value in game theory, aiming to quantify the contribution of each feature to the model's prediction. Given a model $f(x)$ and input features $x = (x1, x2, ..., xM)$, the SHAP value ϕi represents the marginal contribution of feature xi to the current prediction. By traversing all possible feature combinations, the contribution of each feature within each combination is calculated and averaged with weights to obtain the SHAP value for that feature. The specific formula is as follows:

$$\emptyset_i = \sum_{S \subseteq \{1,\ldots,M\} \setminus \{i\}} \omega(S, i)\big(f(S \cup \{i\}) - f(S)\big)\ , \tag{6}$$

where S represents a subset of features that does not include feature i, and $S \cup \{i\}$ denotes the feature subset obtained by adding feature i to S. Here, $f(S)$ is the model prediction based on subset S, and $f(S \cup \{i\})$ is the model prediction after feature i is included. The term $w(S, i)$ is the Shapley weight, which reflects the weighted marginal contribution of feature i across different feature subsets.

Applying the SHAP algorithm to the full dataset is computationally expensive. A previous study (S. L. Chau et al., 2022) as well as our experiments have both demonstrated that subsampling yields no significant differences in SHAP analysis. Therefore, sparse sampling was adopted. Specifically, one prediction point was selected every 5 degrees in space and every 5 days in time. For each prediction point, the input features within the surrounding 4×4 grid were extracted as the model input. The relative importance of each input feature was then obtained by normalizing the sum of the absolute SHAP values for each feature.

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

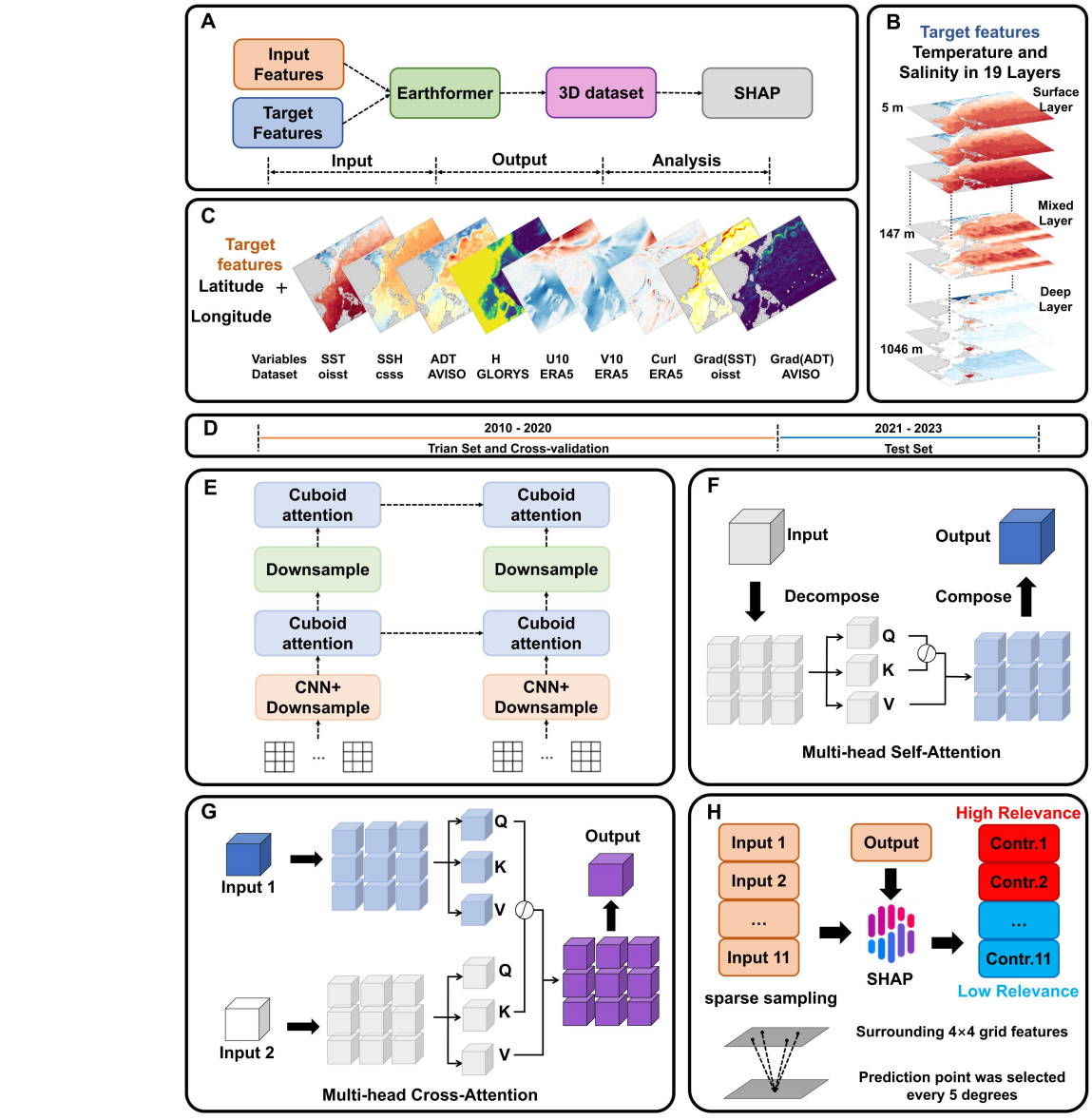

Figure 2. Framework and core mechanisms of the Earthformer-based reconstruction for 3D ocean temperature and salinity fields: (A) Overall workflow of the 3D ocean temperature and salinity field reconstruction; (B) Target features consisting of temperature and salinity profiles across 19 depth layers; (C) Input features comprising multi-source driving data, including physical variables such as SST, SSH, ADT, latitude and longitude; (D) Dataset partitioning, with 2010–2020 dedicated to model training and cross-validation, and 2021–2023 reserved for model testing and evaluation; (E) The overall architecture and mechanism of the Earthformer model; (F) Detailed decomposition and composition processes of cuboid data within the model's multi-head self-attention mechanism; (G) The multi-head cross-attention mechanism, illustrating the interaction between context and target blocks to capture spatiotemporal correlations; (H) SHAP analysis with sparse sampling. Prediction points were selected every 5° and every 5 days, using surrounding 4 × 4 grid features as inputs.

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Open Access Earth System Science Data Discussions

3. Results

3.1 Sensitivity analysis

Fig. 3 (a) and (b) present the results of the spatial sensitivity experiments validated against GLORYS, showing RMSE of temperature and salinity with depth, respectively. The temperature RMSE exhibits a pronounced V-shaped pattern: the error is relatively small near the surface, reaches a peak around the thermocline (approximately 100–150 m), and then gradually decreases with increasing depth, eventually becoming smaller than the surface error and reaching its minimum value. In contrast, the salinity RMSE is largest in the upper ocean, especially within the upper 50 m, and shows a steady decreasing trend with depth. Fig. 3 (c) and (d) show the temperature and salinity RMSE over time, respectively. Both temperature and salinity RMSE display pronounced seasonal and interannual variability, with larger errors in summer and lower errors in winter.

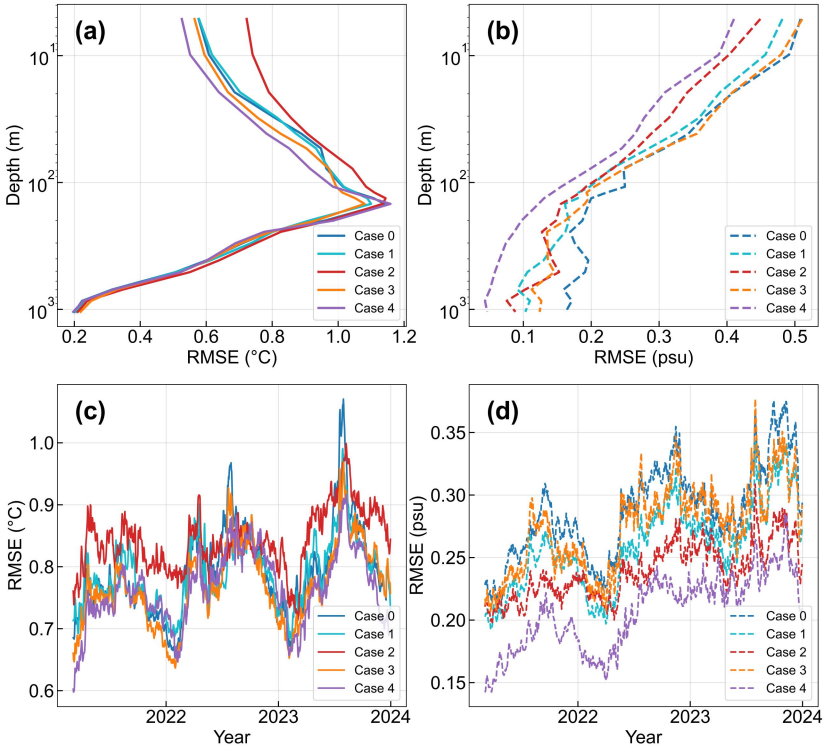

Figure 3. Sensitivity analysis of temperature and salinity field reconstruction with different feature combinations. (a) Mean RMSE of reconstructed temperature over longitude, latitude, and time as a function of depth; (b) same as (a) but for salinity; (c) Mean RMSE of reconstructed temperature over longitude and depth as a function of time; (d) same as (c), but for salinity.

Generally, including more features leads to more accurate reconstructions, which aligns with previous understanding (K. Zhang et al., 2020). Among all experiments, Case 4, which used the full feature combination, showed the most significant improvement: compared with Case 0, the salinity RMSE decreased by about 30%–50%. It is worth noting that adding more

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

features is not always effective. For example, in temperature reconstruction, after the wind speed variables were added in Case 2, the surface error increased sharply, reaching the largest value among the five experiments. However, after the wind stress curl was further introduced in Case 3, the errors in both the depth and time dimensions decreased, and the accuracy surpassed that of Case 2 and Case 1. A similar pattern was also found in salinity reconstruction. In Case 3, after the wind stress variable was added, the overall errors in both depth and time increased significantly. However, once the SST and ADT gradients were introduced in Case 4, the errors in both the depth and time dimensions decreased to their minimum values. Overall, although adding more features can improve reconstruction accuracy, the improvement does not occur uniformly across all depths and time periods. The improvement in temperature reconstruction accuracy was mainly concentrated in the upper ocean above approximately 100 m. In contrast, the error near the thermocline increased slightly, while there was almost no difference in the deep layer. For salinity reconstruction, however, the improvement extended throughout the entire water column.

Comparing cases can generate some understanding on the effects of features. For example, the errors of Case 0 are significantly higher during the summers of 2021–2023, whereas this phenomenon did not appear in the other experimental groups. By comparing the different experimental results, it can be inferred that the absence of topographic information in Case 0 was the main cause of this problem: other experiments that included topographic information showed clear improvement. These experiments suggest that the main sources of reconstruction error are either the limited ability of the neural network to learn the nonlinear relationship between the 3D field and the surface field, or the failure to effectively capture and include some key features in the input data, which will be further discussed later with SHAP results.

3.2 Spatial patterns

To further assess reconstruction performance, we compared the reconstructed fields with the GLORYS reference fields on two representative snapshots (24 January and 24 June 2023) at three depth levels: 41.18 (mixed layer), 147.41 (pycnocline), and 508.64 m (middle layer, Fig 4).

The model accurately reproduced the basin-scale seasonal contrast of the upper ocean. At 41.18 m, the domain-mean reconstructed temperature increased from 24.37°C on 24 January to 25.91°C on 24 June, corresponding to a warming of 1.54°C. The corresponding GLORYS warming was 1.44°C, with a difference of only 0.10°C. Spatially, the model captured the meridional temperature gradient in the upper layer. Along the 143–147°E band, the reconstructed temperature difference between 10–15°N and 30–35°N was 10.16°C on 24 January, close to the GLORYS value of 9.99°C, with a bias of 0.17°C. On 24 June, the reconstructed and GLORYS differences were 8.46°C and 8.09°C, respectively, with a bias of 0.37°C. For salinity, the model reproduced the broad pattern of higher salinity in the open ocean than in the marginal seas, although with noticeable smoothing. At 41.18 m, the salinity contrast between the open-ocean region 140–150°E, 15–25°N and the marginal-sea region 110–120°E, 20–30°N was 0.86 PSU in the reconstruction on 24 January, lower than the GLORYS value

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

of 1.27 PSU. On 24 June, the corresponding contrast was 0.59 PSU, compared with 0.97 PSU in GLORYS. Thus, the model underestimated the salinity contrast by approximately 0.41 and 0.38 PSU, respectively.

The main limitations were the underestimation of local gradients and salinity contrasts. Relatively large local discrepancies occurred near the Subtropical Countercurrent at approximately 19°N and the North Equatorial Countercurrent at approximately 6°N. In the Kuroshio-related region 130–138°E, 22–28°N, the model reproduced the regional mean temperature at 147.41 m reasonably well, but underestimated frontal gradients. The reconstructed mean temperature gradient was 0.56°C per 100 km on 24 January, lower than the GLORYS value of 1.00°C per 100 km. On 24 June, the reconstructed gradient increased to 0.78°C per 100 km, but remained lower than the GLORYS value of 1.06°C per 100 km. In addition, the salinity contrast between the open ocean and marginal seas was systematically underestimated.

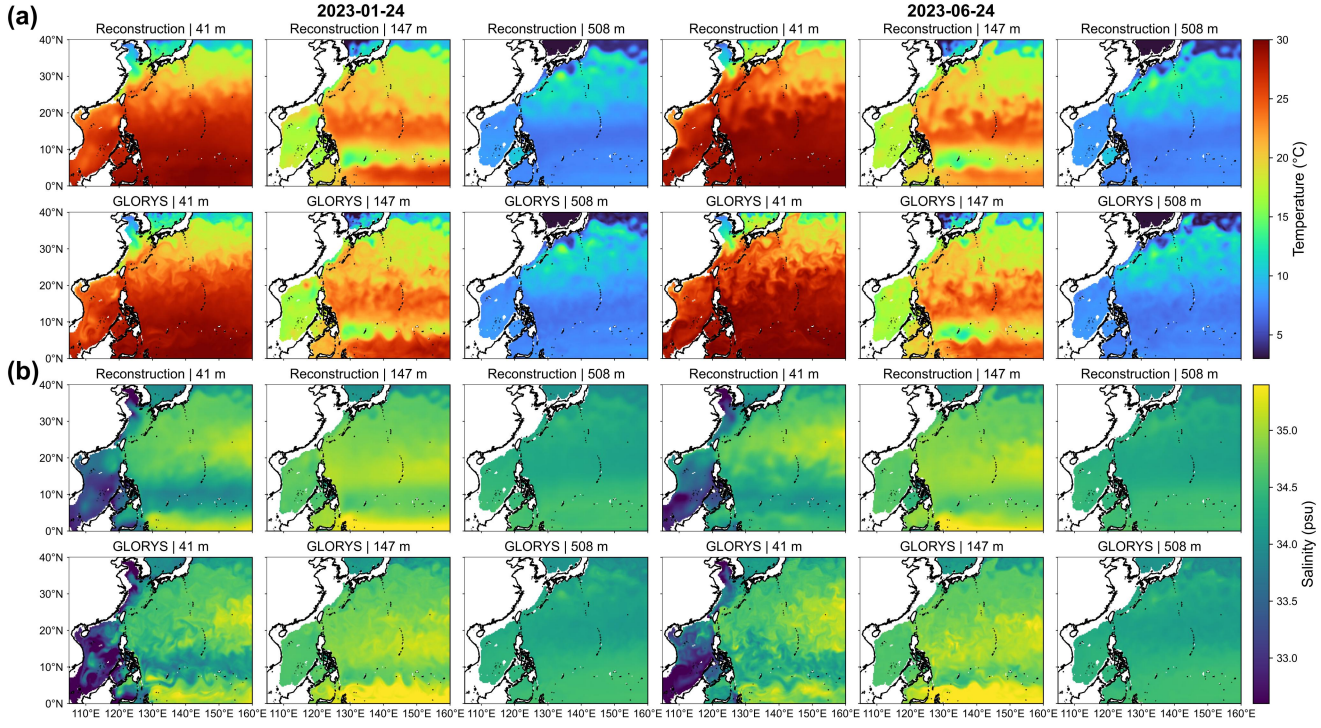

Figure 4. Spatial comparison of reconstructed and reference temperature and salinity fields at different depths in the NWP on 24 January 2023 and 24 June 2023. (a) Comparison of temperature fields, showing the spatial distributions of the reconstructed results (Reconstruction) and the reference fields (GLORYS) at depths of 41 m, 147 m, and 508 m; (b) comparison of salinity fields, showing the corresponding spatial distributions at the same depths.

3.3 Reconstruction accuracy vs. Argo

Beyond comparison with GLORYS, a more critical evaluation involves comparing the results with observational data (Fig. 5). This is essential for verifying the accuracy of both the reanalysis data and the reconstruction results, ensuring the final reconstructed fields are suitable for practical quasi-real-time applications.

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Compared with Argo observations, the GLORYS data show relatively high simulation accuracy: the RMSE for temperature is as low as 0.766°C, with a coefficient of determination (R^2) of 0.992 and an error percentage of only about 0.8%; for salinity, the RMSE is 0.111 PSU, with an R^2 of 0.887 and an error percentage of about 11.3%. The reconstruction results also achieved high accuracy: the RMSE for temperature is 0.893°C, with an R^2 of 0.989 and an error percentage of about 1.1%; the RMSE for salinity is 0.141 PSU, with an R^2 of 0.827 and an error percentage of about 17.3%, indicating slightly poorer performance for salinity compared to temperature. Salinity exhibits a noticeable underestimation in the low-salinity range. These errors are 6% larger than those of GLORYS. However, given the fact that only very limited surface information was used as the input in our reconstruction, and that GLORYS uses an equation-based numerical model to assimilate large volumes of in-situ data, this accuracy is plausible.

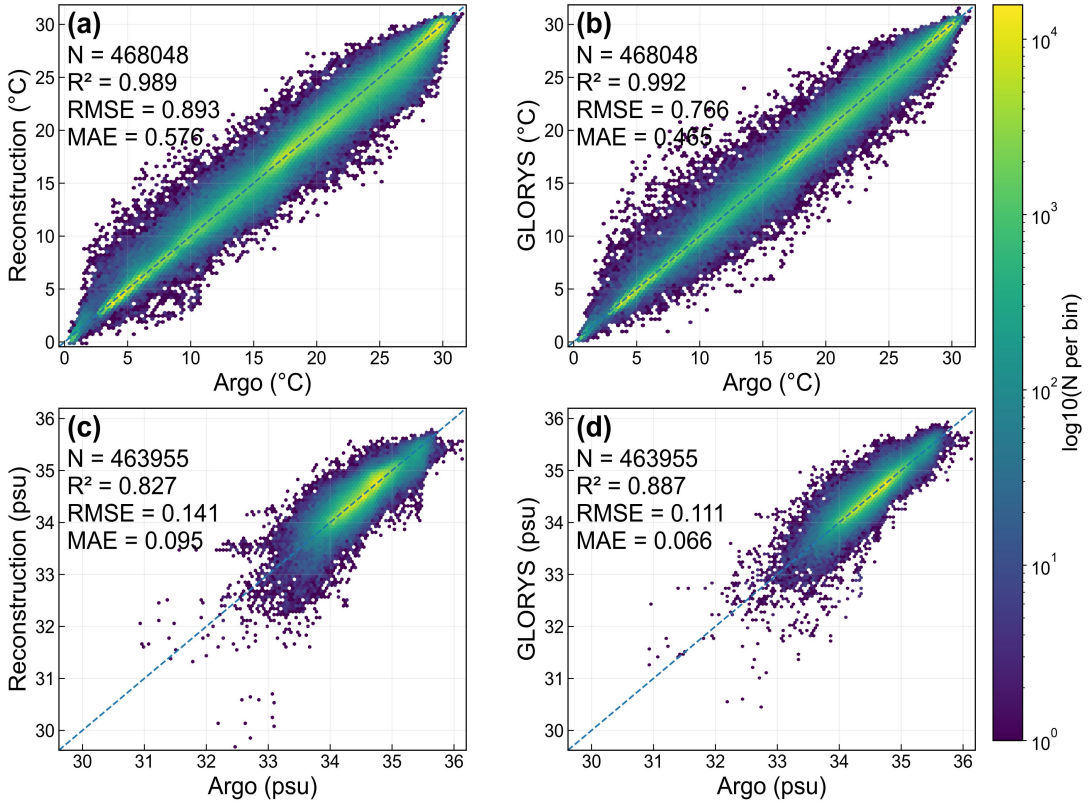

Figure 5. Scatter-density comparisons of temperature and salinity among Argo observations, reconstructed results, and GLORYS data. (a) Reconstructed temperature versus Argo temperature; (b) GLORYS temperature versus Argo temperature; (c) Reconstructed salinity versus Argo salinity; (d) GLORYS salinity versus Argo salinity. The color indicates the logarithm of the sample count in each statistical bin. Each panel also shows the sample size (N), coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE).

Further examination on the vertical accuracy pattern (Fig. 6) against Argo showed generally consistent errors of GLORYS and the reconstruction, similar with those in Fig. 3. The reconstruction had temperature and salinity errors approximately

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

0.10–0.30 °C and 0.02–0.06 PSU larger than GLORYS, respectively but with similar vertical structure. This suggests that part of the reconstruction errors may originate from the error characteristics inherent in the GLORYS training data itself, rather than being entirely attributable to limitations in the reconstruction capability of the deep learning model. This fact also provides a hint for understanding the vertical error structure: the shape does not necessarily only link to the variability of the target, but can also be associated with uncertainty sources.

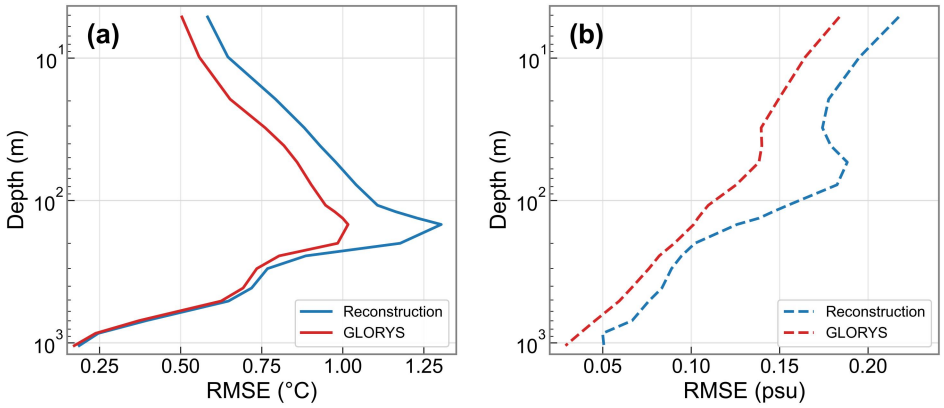

Figure 6. RMSE of the reconstruction and GLORYS compared against Argo ground truth values for (a) temperature and (b) salinity.

3.4 Neural network interpretability

After well validation against multi-source data, the next question is to understand the impact of different inputs on the reconstruction of temperature and salinity field across different depth layers using the SHAP method, as shown in Fig. 7. Overall, both temperature and salinity reconstructions mainly rely on three fundamental sea surface variables (SST, ADT, and SSS) as well as location parameters (longitude, latitude, and bathymetry), while wind- and front-related factors provide only complementary information. This is consistent with the results of the sensitivity experiments presented above: Case 1, which uses only SST, ADT, SSS, longitude, latitude, and bathymetry as input variables, already shows good reconstruction capability, whereas Case 4, which further incorporates wind- and front-related factors, yields an additional improvement in reconstruction accuracy of approximately 10%–20%.

In temperature reconstruction, the relative importance of SST and ADT varies clearly with depth. SST makes the largest contribution in the surface and with its contribution exceeding 20%, because it provides direct thermal information. In the mixed layer, the contribution of ADT is comparable to that of SST, and in deeper layers it becomes larger than that of SST. A similar complementary relationship between SST and ADT is also evident in the SHAP time series (Fig. 8) for both temperature and salinity reconstruction, where the two variables show alternating or jointly enhanced contributions over time. This highlights the physical relation of heaving processes leading to close co-variance of ADT and temperature (L. Zhang et al., 2012).

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Open Access Earth System Science Data Discussions

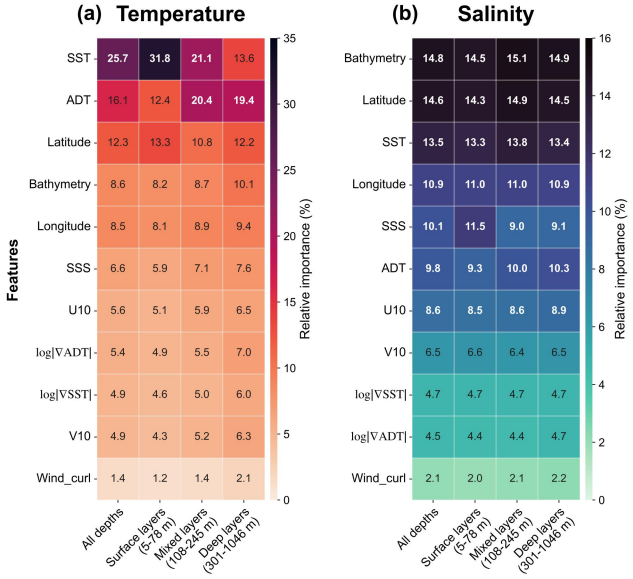

Figure 7. Relative importance of input features for (a) temperature and (b) salinity across different depth layers based on SHAP values.

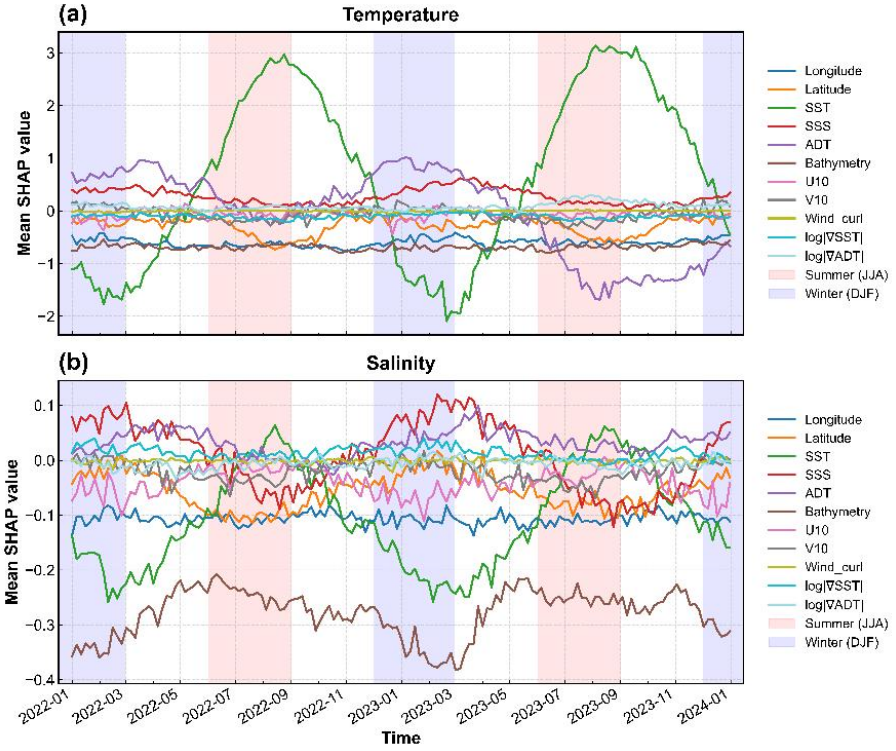

Figure 8. Time series of relative feature importance for (a) temperature and (b) salinity. Shaded areas indicate summer (JJA) and winter (DJF) seasons.

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

An apparently counterintuitive pattern is observed in the SHAP results for salinity. Unlike SST contribution for temperature reconstruction (reaching predominate 31.8% at surface), SSS plays a merely minor role in salinity reconstruction. Its overall contribution is only about 10%, ranking fifth among all input features, showing a more vertically uniform distribution. The contribution is even lower than static input features, i.e., bathymetry, latitude, and longitude. The low contribution is also contradictory with that in Tian et al. (2022), in which surface salinity is a vital predictor for 3D salinity. A candidate explanation for such low contribution is the different accuracies of the SST and SSS products used in this study. Compared with Argo observations during 2022, SST shows much higher accuracy than SSS, with an R^2 of 0.991 versus 0.732. The relatively low accuracy of remote-derived MSSS product may thus constrain its role in salinity reconstruction, which will be further examined below.

4. Discussion

4.1 Revisiting SHAP for Salinity Reconstruction

To further examine this phenomenon, we designed Case4-G, in which the remote-derived SSS was replaced by GLORYS SSS, to examine the reconstruction accuracy in the case of 100% SSS accuracy (vertical RMSE profile in Fig. 9 and SHAP results in Fig. 10). Apparently, this substitution substantially changed the vertical RMSE profile of salinity reconstruction, markedly reducing the RMSE in the upper ocean. More importantly, the error profile no longer shows a monotonic decrease with depth, but instead exhibits a V-shaped structure, very similar with that for temperature reconstruction. Further substitution SSS with the average of GLORYS SSS and remote-derived SSS ("Mean" case in Fig. 9), provides further evidence for this interpretation. Its RMSE profile lies between those of the two experiments, indicating a gradual transition as the quality of the SSS input improves. This result suggests that the vertical structure of salinity reconstruction errors is strongly controlled by the information content and reliability of the surface salinity field.

SHAP analysis supports this interpretation. After replacement with GLORYS SSS, the contribution increased to 49%, dominantly contributing to salinity. The increase was strongest in the surface layer, where the contribution rose from 11.5% to 54%. In the deep layer, the SSS contribution also increased from 9.1% to 34.2%, despite of no visual improvements in reconstruction accuracy. From a snapshot SSS contrasting GLORYS SSS with MSSS (Fig. 11), it is clear that MSSS has substantial pepper-like noises with vague meandering signals. When training a reconstruction model, these noises would inevitably contaminate the training process, degrading the important role of SSS. In summary, SSS is not intrinsically unimportant for 3D salinity reconstruction. Instead, the limited accuracy of current satellite products restricts SSS effectiveness.

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Open Access Earth System Science Data Discussions

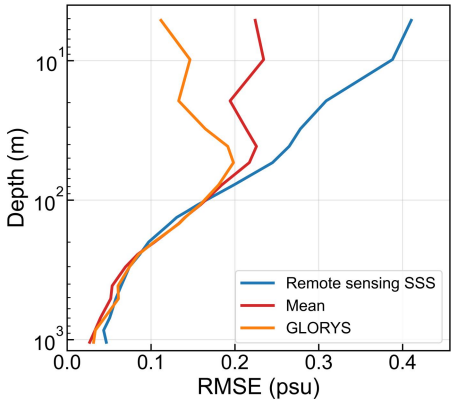

Figure 9. Vertical variations of RMSE in 3D salinity reconstruction under different SSS input conditions.

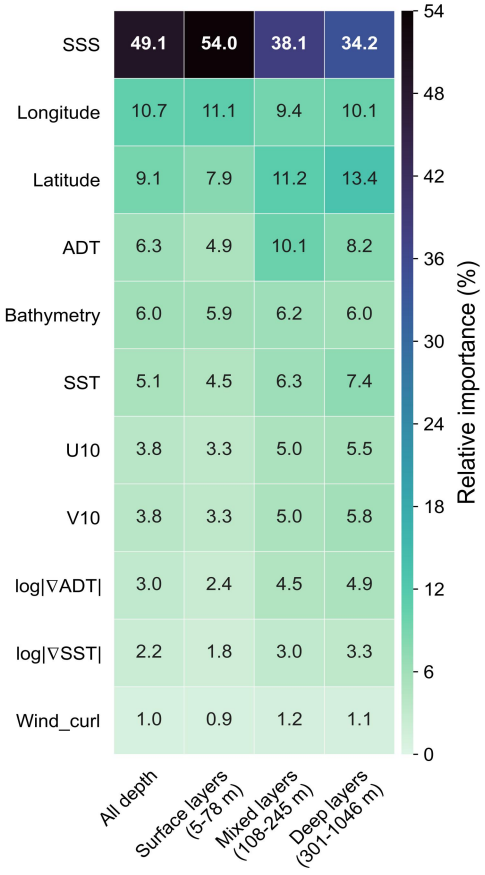

Figure 10. Relative importance of input features for salinity across different depth layers based on SHAP values after replacing satellite-derived SSS with GLORYS SSS.

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

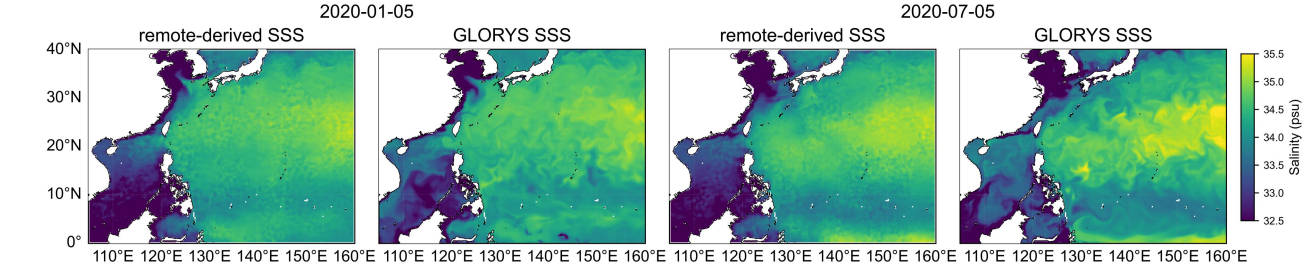

Figure 11. Spatial distributions of surface salinity from remote-derived SSS and GLORYS SSS on 5 January 2020 and 5 July 2020.

This finding has important implications for future 3D salinity reconstruction, calling for better remote-derived SSS products. When more accurate SSS is provided, the model relies strongly on SSS, with improved reconstruction accuracy up to 10%. Its apparent low contribution in experiments using satellite-derived SSS mainly reflects data-quality limitations rather than weak physical relevance. Improving satellite SSS retrievals (Dinnat et al., 2019; Z. Li et al., 2024; Yu et al., 2017), multi-source fusion (Song et al., 2024; J. Zhang, 2010), and bias correction (Bao et al., 2023; Pasula & Subramani, 2026) may therefore be as important as optimizing model architecture. Currently, SSS retrieval relies on NASA's Soil Moisture Active Passive (SMAP) and Soil Moisture and Ocean Salinity (SMOS) observatory, which has only an 8-day revisit cycle and a weekly RMSE of 0.26 PSU leading to even higher daily errors in the MSSS product (Buongiorno Nardelli et al., 2016). In addition, it is also documented that the phase difference between SMAP and buoy observations at some salinity peaks can lead to positive biases in certain periods or regions (Bao, Wang, et al., 2019; W. Tang et al., 2017). In this case, to achieve high accuracy daily SSS dataset, the community could seek for indirect data-driven methods. In this regard, deep learning aided by explainability approaches could be helpful. These understandings are directly yielded from SHAP analysis: unlike previous studies using SHAP only to reported the relative contributions of individual input features (F. Liu et al., 2026), we further use SHAP to identify future directions for improving the accuracy of daily 3D temperature and salinity reconstruction.

4.2 Practical Advantages of the Proposed Reconstruction Framework

Previous results demonstrate that Earthformer is capable of reconstructing daily 3D temperature and salinity structures in the NWP using surface-only observations. For an operational purpose, we further highlight three advantages in our method: accuracy, timeliness and lightweight.

Firstly, using the IV-TT criterion (Ryan et al., 2015), we showed that the performance is lower than yet close to that of GLORYS. The accuracy of the reconstructed temperature is only 0.3% lower than that of GLORYS, while the accuracy of the reconstructed salinity is 6% lower than that of GLORYS. In addition, the accuracy of our method surpasses that of other intelligent reconstruction approaches, including convolutional neural network (CNN), Light gradient boosting Machine (LightGBM) (Su et al., 2021) and physics-guided generative adversarial networks ((Y. Meng et al., 2023). A directly

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

comparable study by Su et al. (2021b) showed that the RMSE of the CNN model was close to 2.5 °C in the mixed layer, while the RMSE of the LightGBM model was close to 1.5 °C. In contrast, our model achieved an RMSE of less than 1.2 °C in the mixed layer. Moreover, our model also showed lower RMSEs in both the surface and deep layers, indicating better reconstruction performance across the full depth range.

In terms of timeliness, our method also has advantages over GLORYS, which is a reanalysis product and is therefore subject to an inherent production latency of about one to two months. For example, as of May 9, 2026, the latest available GLORYS reanalysis are only up to March 26, 2026. Our framework relies only on surface observations and is thus more suitable for near-real-time applications. Moreover, under identical input-output configuration and spatial-temporal resolutions, our model requires substantially less computational time. Specifically, the reconstruction of daily temperature fields for the entire year of 2023 across 19 depth levels was completed in approximately 60 s using an NVIDIA GeForce RTX 4080 SUPER graphic card. This indicates that the proposed model can achieve efficient 3D temperature reconstruction with relatively low computational cost, promisingly applied in local operational centers or for governance purposes.

4.3 Study limitations

Nevertheless, two inherent limitations of this approach remain. First, the model estimates still exhibit some roughness in the spatial patterns, suggesting that the existing input information may be insufficient to fully reduce estimation errors. Therefore, future work should incorporate additional information, such as higher-quality sea surface observations and more relevant input variables (L. Meng et al., 2022; Z. Wu et al., 2025). Second, the current model reconstructs temperature and salinity separately, without explicitly constraining their physical consistency (Dong et al., 2022). Future work could incorporate the seawater equation of state into AI, for example through Physics-Informed Neural Networks (PINNs) (Raissi et al., 2019), to better preserve the physical linkage between temperature and salinity.

5. Data availability

The reconstructed daily three-dimensional temperature and salinity dataset for the Northwest Pacific is available from Zenodo at https://doi.org/10.5281/zenodo.20602639 (Fang, 2026).The dataset includes six NetCDF files corresponding to Case 0 to Case 4 and Case 4-G. The dataset is distributed under the Creative Commons Attribution 4.0 International license.

The source datasets used to generate and validate the reconstructed product are publicly available from their original providers. The GLORYS Global Ocean Physics Reanalysis product, GLOBAL_MULTIYEAR_PHY_001_030, used as the reference data for model training and testing, is available from the Copernicus Marine Service at https://data.marine.copernicus.eu/products. The AVISO absolute dynamic topography data used as input features are available from https://www.aviso.altimetry.fr/en/data.html. The NOAA Optimum Interpolation Sea Surface Temperature data are available from the National Centers for Environmental Information at

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

https://www.ncei.noaa.gov/products/optimum-interpolation-sst. The ERA5 reanalysis data used to derive U10, V10, and wind stress curl are available from the Copernicus Climate Data Store at https://cds.climate.copernicus.eu/. The multi-source sea surface salinity product, MULTIOBS_GLO_PHY_S_SURFACE_MYNRT_015_013, is available from the Copernicus Marine Service at https://data.marine.copernicus.eu/product/MULTIOBS_GLO_PHY_S_SURFACE_MYNRT_015_013/description. The Argo profile observations used for independent validation are available from the International Argo Program at https://argo.ucsd.edu/. Only Argo profiles with the highest-quality control flag were used in the validation.

6. Conclusion

In this study, an EarthFormer-based deep learning framework was developed to reconstruct daily 3D temperature and salinity fields in the NWP at a spatial resolution of 0.25° using multi-source satellite remote sensing products. Sensitivity experiments demonstrated that incorporating bathymetry, wind forcing, and frontal information significantly improved reconstruction performance. Independent validation against Argo observations showed that the reconstructed temperature achieved an RMSE of 0.893 °C, while the reconstructed salinity achieved an RMSE of 0.141 PSU. The overall reconstruction accuracy was close to that of the GLORYS reanalysis data, which had RMSE values of 0.766 °C for temperature and 0.111 PSU for salinity. Moreover, using explainability analysis, we identified a counter-intuitively low contribution of SSS in salinity reconstruction unlike SST which dominates temperature reconstruction. Such low contribution is largely due to low accuracy of remote-sensing-derived SSS. When replaced with noise-free GLORYS SSS, SSS actually contributes similarly with temperature, showing a V-shaped RMSE structure. The analysis also highlights the need for improved SSS products to better reconstruct the 3D thermohaline structure. Improving satellite SSS retrievals, multi-source fusion, and bias correction may therefore be as important as optimizing model architecture.

Overall, this study demonstrates the feasibility of lightweight daily scale 3D temperature and salinity reconstruction using advanced deep learning and multi-source remote sensing data. More importantly, the results highlight the critical role of remote-derived product quality in shaping the vertical error structures of ocean reconstruction, providing new insight into the physical interpretation of 3D thermohaline reconstruction errors and offering a potential pathway toward future real-time ocean monitoring applications.

Author contribution

Xihong Fang and Wenfang Lu conceptualized the study. Xihong Fang developed the model code, performed the formal analysis and validation, and prepared the original draft. Yujiao Zheng and Jiangnan He contributed to data curation. Haochen Sun, Siming Huang and Young-Heon Jo contributed to the investigation. Wenfang Lu contributed to data curation and formal analysis, supervised the study, administered the project, acquired funding, and reviewed and edited the manuscript.

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This study is supported by Guangdong Province Natural Science Foundation (2025A1515012024). X. F. H.S. and S.H. were also supported by College Students Innovative Entrepreneurial Training Plan Program (No. 20260826). The authors gratefully acknowledge the Copernicus Marine Service for providing the GLORYS reanalysis and multi-source satellite sea surface salinity products, AVISO for providing absolute dynamic topography data, NOAA for providing the Optimum Interpolation Sea Surface Temperature data, the European Centre for Medium-Range Weather Forecasts for providing the ERA5 reanalysis dataset, and the International Argo Program and its national contributors for providing Argo profile observations. The authors used ChatGPT by OpenAI for language polishing and assistance with manuscript submission preparation

Code availability

The supplementary code, configuration file, utility modules, and trained model weights used in this study are available at https://github.com/Xihong-tomato/ESSD-Github-model. The repository contains the training script, inference and evaluation scripts, model architecture modules, utility modules, configuration file, and trained temperature and salinity model weights.

References

Ali, M. M., Swain, D., & Weller, R. A. (2004). Estimation of ocean subsurface thermal structure from surface parameters: A neural network approach. Geophysical Research Letters, 31(20), 2004GL021192. https://doi.org/10.1029/2004GL021192

Arribas, A., Glover, M., Maidens, A., Peterson, K., Gordon, M., MacLachlan, C., Graham, R., Fereday, D., Camp, J., Scaife, A. A., Xavier, P., McLean, P., Colman, A., & Cusack, S. (2011). The GloSea4 Ensemble Prediction System for Seasonal Forecasting. Monthly Weather Review, 139(6), 1891–1910. https://doi.org/10.1175/2010MWR3615.1

Bao, S., Wang, H., Zhang, R., Yan, H., & Chen, J. (2019). Comparison of Satellite-Derived Sea Surface Salinity Products from SMOS, Aquarius, and SMAP. Journal of Geophysical Research: Oceans, 124(3), 1932–1944. https://doi.org/10.1029/2019JC014937

Bao, S., Zhang, R., Wang, H., Yan, H., Chen, J., & Wang, Y. (2023). Correction of Satellite Sea Surface Salinity Products Using Ensemble Learning Method. IEEE Access, 11, 17870–17881. https://doi.org/10.1109/ACCESS.2021.3057886

Bao, S., Zhang, R., Wang, H., Yan, H., Yu, Y., & Chen, J. (2019). Salinity Profile Estimation in the Pacific Ocean from Satellite Surface Salinity Observations. Journal of Atmospheric and Oceanic Technology, 36(1), 53–68. https://doi.org/10.1175/JTECH-D-17-0226.1

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2023). Accurate medium-range global weather forecasting with 3D neural networks. Nature, 619(7970), 533–538. https://doi.org/10.1038/s41586-023-06185-3

Buongiorno Nardelli, B., Droghei, R., & Santoleri, R. (2016). Multi-dimensional interpolation of SMOS sea surface salinity with surface temperature and in situ salinity data. Remote Sensing of Environment, 180, 392–402. https://doi.org/10.1016/j.rse.2015.12.052

Cao, X., Liu, C., Zhang, S., & Gao, F. (2024). A Method for Predicting High-Resolution 3D Variations in Temperature and

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Salinity Fields Using Multi-Source Ocean Data. Journal of Marine Science and Engineering, 12(8), 1396. https://doi.org/10.3390/jmse12081396
Cha, H., Moon, J.-H., Kim, T., & Song, Y. T. (2023). A process-based assessment of the sea-level rise in the northwestern Pacific marginal seas. Communications Earth & Environment, 4(1), 300. https://doi.org/10.1038/s43247-023-00965-5
Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Hogan, P. J., Wallcraft, A. J., Baraille, R., & Bleck, R. (2007). The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. Journal of Marine Systems, 65(1–4), 60–83. https://doi.org/10.1016/j.jmarsys.2005.09.016
Chau, S. L., Hu, R., Gonzalez, J., & Sejdinovic, D. (2022). RKHS-SHAP: Shapley Values for Kernel Methods (arXiv:2110.09167). arXiv. https://doi.org/10.48550/arXiv.2110.09167
Chau, T.-T.-T., Gehlen, M., Metzl, N., & Chevallier, F. (2024). CMEMS-LSCE: A global, 0.25°, monthly reconstruction of the surface ocean carbonate system. Earth System Science Data, 16(1), 121–160. https://doi.org/10.5194/essd-16-121-2024
Chen, Y., Liu, L., Chen, X., Wei, Z., Sun, X., Yuan, C., & Gao, Z. (2023). Data driven three-dimensional temperature and salinity anomaly reconstruction of the northwest Pacific Ocean. Frontiers in Marine Science, 10, 1121334. https://doi.org/10.3389/fmars.2023.1121334
Chen, Z., Wang, X., & Liu, L. (2020). Reconstruction of Three-Dimensional Ocean Structure From Sea Surface Data: An Application of isQG Method in the Southwest Indian Ocean. Journal of Geophysical Research: Oceans, 125(6), e2020JC016351. https://doi.org/10.1029/2020JC016351
Cheng, L., & Zhu, J. (2016). Benefits of CMIP5 Multimodel Ensemble in Reconstructing Historical Ocean Subsurface Temperature Variations. Journal of Climate, 29(15), 5393–5416. https://doi.org/10.1175/JCLI-D-15-0730.1
Chu, P. C., Fan, C., & Liu, W. T. (2000). Determination of Vertical Thermal Structure from Sea Surface Temperature. Journal of Atmospheric and Oceanic Technology, 17(7), 971–979. https://doi.org/10.1175/1520-0426(2000)017%3C0971:DOVTSF%3E2.0.CO;2
De Boyer Montégut, C., Mignot, J., Lazar, A., & Cravatte, S. (2007). Control of salinity on the mixed layer depth in the world ocean: 1. General description. Journal of Geophysical Research: Oceans, 112(C6), 2006JC003953. https://doi.org/10.1029/2006JC003953
Dinnat, E. P., Le Vine, D. M., Boutin, J., Meissner, T., & Lagerloef, G. (2019). Remote Sensing of Sea Surface Salinity: Comparison of Satellite and In Situ Observations and Impact of Retrieval Parameters. Remote Sensing, 11(7), 750. https://doi.org/10.3390/rs11070750
Dong, C., Xu, G., Han, G., Bethel, B. J., Xie, W., & Zhou, S. (2022). Recent Developments in Artificial Intelligence in Oceanography. Ocean-Land-Atmosphere Research, 2022, 2022/9870950. https://doi.org/10.34133/2022/9870950
Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale (arXiv:2010.11929). arXiv. https://doi.org/10.48550/arXiv.2010.11929
Fang, X. (2026). Daily 3D Temperature and Salinity Reconstruction Dataset for the Northwest Pacific from Remote Sensing with Explainable AI [Dataset]. Zenodo. https://doi.org/10.5281/ZENODO.20602639
Fouchet, E., Benkiran, M., Le Traon, P.-Y., & Remy, E. (2025). Comparison of a global high-resolution ocean data assimilation system with SWOT observations. Frontiers in Marine Science, 12, 1563934. https://doi.org/10.3389/fmars.2025.1563934
Gao, Z., Shi, X., Wang, H., Zhu, Y., Wang, Y., Li, M., & Yeung, D.-Y. (2023). Earthformer: Exploring Space-Time Transformers for Earth System Forecasting (arXiv:2207.05833). arXiv. https://doi.org/10.48550/arXiv.2207.05833
Garcia-Espriu, A., González-Haro, C., & Aguilar-Gómez, F. (2025). On the global reconstruction of ocean interior variables: A feasibility data-driven study with simulated surface and water column observations. Ocean Science, 21(5), 2579–2603. https://doi.org/10.5194/os-21-2579-2025
Garcia-Gorriz, E., & Garcia-Sanchez, J. (2007). Prediction of sea surface temperatures in the western Mediterranean Sea by neural networks using satellite observations. Geophysical Research Letters, 34(11), 2007GL029888. https://doi.org/10.1029/2007GL029888
Good, S. A., Martin, M. J., & Rayner, N. A. (2013). EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. Journal of Geophysical Research: Oceans, 118(12), 6704–6716. https://doi.org/10.1002/2013JC009067
He, J., Lu, W., Liu, Y., Yang, G., Jo, Y.-H., & Lai, Z. (2026). Accurate Sea Level Anomaly Forecast over North Pacific with

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Improved AI Training Strategies. Ocean-Land-Atmosphere Research, 5, 0128. https://doi.org/10.34133/olar.0128

Helber, R. W., Richman, J. G., & Barron, C. N. (2010). The influence of temperature and salinity variability on the upper ocean density and mixed layer. Mixed Layer/In situ Observations/All Geographic Regions/Temperature, Salinity and Density Fields. https://doi.org/10.5194/osd-7-1469-2010

Hu, S. (2025). Subsurface marine heatwaves: Challenging to detect at the surface yet critically important. Journal of Oceanology and Limnology, s00343-025-5225–5227. https://doi.org/10.1007/s00343-025-5225-7

Jean-Michel, L., Eric, G., Romain, B.-B., Gilles, G., Angélique, M., Marie, D., Clément, B., Mathieu, H., Olivier, L. G., Charly, R., Tony, C., Charles-Emmanuel, T., Florent, G., Giovanni, R., Mounir, B., Yann, D., & Pierre-Yves, L. T. (2021). The Copernicus Global 1/12° Oceanic and Sea Ice GLORYS12 Reanalysis. Frontiers in Earth Science, 9, 698876. https://doi.org/10.3389/feart.2021.698876

Jo, Y., Dai, M., Zhai, W., Yan, X., & Shang, S. (2012). On the variations of sea surface p CO2 in the northern South China Sea: A remote sensing based neural network approach. Journal of Geophysical Research: Oceans, 117(C8), 2011JC007745. https://doi.org/10.1029/2011JC007745

Khedouri, E., Szczechowski, C., & Cheney, R. (1983). Potential Oceanographic Applications Of Satellite Altimetry For Inferring Subsurface Thermal Structure. Proceedings OCEANS '83, 274–280. https://doi.org/10.1109/OCEANS.1983.1152138

Kim, Y.-Y., Kim, B.-G., Jeong, K. Y., Lee, E., Byun, D.-S., & Cho, Y.-K. (2021). Local Sea-Level Rise Caused by Climate Change in the Northwest Pacific Marginal Seas Using Dynamical Downscaling. Frontiers in Marine Science, 8, 620570. https://doi.org/10.3389/fmars.2021.620570

Klemas, V., & Yan, X.-H. (2014). Subsurface and deeper ocean remote sensing from satellites: An overview and new results. Progress in Oceanography, 122, 1–9. https://doi.org/10.1016/j.pocean.2013.11.010

Landschützer, P., Gruber, N., Bakker, D. C. E., Schuster, U., Nakaoka, S., Payne, M. R., Sasse, T. P., & Zeng, J. (2013). A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. Biogeosciences, 10(11), 7793–7815. https://doi.org/10.5194/bg-10-7793-2013

Li, H., & Ilyina, T. (2018). Current and Future Decadal Trends in the Oceanic Carbon Uptake Are Dominated by Internal Variability. Geophysical Research Letters, 45(2), 916–925. https://doi.org/10.1002/2017GL075370

Li, Z., Guo, F., Zhang, X., Zhang, Z., Zhu, Y., Yang, W., Wu, Z., & Yue, L. (2024). Integrating spaceborne GNSS-R and SMOS for sea surface salinity retrieval using artificial neural network. GPS Solutions, 28(4), 162. https://doi.org/10.1007/s10291-024-01709-4

Liang, Z., Bao, S., Su, H., Zhang, W., Wang, H., & Yan, H. (2026). Daily Subsurface Salinity Reconstruction From Multisource Satellite Observations Using Wavelet-Enhanced 3-D Mamba. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 19, 3371–3388. https://doi.org/10.1109/JSTARS.2025.3649568

Liu, F., Wei, L., & Guan, L. (2026). Ocean temperature reconstruction from satellite observations in the North Atlantic using an explainable deep learning framework. International Journal of Digital Earth, 19(1), 2632430. https://doi.org/10.1080/17538947.2026.2632430

Liu, Y., Lu, W., Wang, D., Lai, Z., Ying, C., Li, X., Han, Y., Wang, Z., & Dong, C. (2024). Spatiotemporal wave forecast with transformer-based network: A case study for the northwestern Pacific Ocean. Ocean Modelling, 188, 102323.

Lu, W., Su, H., Yang, X., & Yan, X.-H. (2019). Subsurface temperature estimation from remote sensing data using a clustering-neural network method. Remote Sensing of Environment, 229, 213–222. https://doi.org/10.1016/j.rse.2019.04.009

Lundberg, S., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions (arXiv:1705.07874). arXiv. https://doi.org/10.48550/arXiv.1705.07874

Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A., & Trenberth, K. E. (2011). Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. Nature Climate Change, 1(7), 360–364. https://doi.org/10.1038/nclimate1229

Meng, L., Yan, C., Zhuang, W., Zhang, W., Geng, X., & Yan, X.-H. (2022). Reconstructing High-Resolution Ocean Subsurface and Interior Temperature and Salinity Anomalies From Satellite Observations. IEEE Transactions on Geoscience and Remote Sensing, 60, 1–14. https://doi.org/10.1109/TGRS.2021.3109979

Meng, L., Yan, C., Zhuang, W., Zhang, W., & Yan, X. (2021). Reconstruction of Three-Dimensional Temperature and Salinity Fields From Satellite Observations. Journal of Geophysical Research: Oceans, 126(11), e2021JC017605. https://doi.org/10.1029/2021JC017605

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Meng, L., & Yan, X.-H. (2022). Remote Sensing for Subsurface and Deeper Oceans: An overview and a future outlook. IEEE Geoscience and Remote Sensing Magazine, 10(3), 72–92. https://doi.org/10.1109/MGRS.2022.3184951

Meng, Y., Rigall, E., Chen, X., Gao, F., Dong, J., & Chen, S. (2023). Physics-Guided Generative Adversarial Networks for Sea Subsurface Temperature Prediction. IEEE Transactions on Neural Networks and Learning Systems, 34(7), 3357–3370. https://doi.org/10.1109/TNNLS.2021.3123968

O'Carroll, A. G., Armstrong, E. M., Beggs, H. M., Bouali, M., Casey, K. S., Corlett, G. K., Dash, P., Donlon, C. J., Gentemann, C. L., Høyer, J. L., Ignatov, A., Kabobah, K., Kachi, M., Kurihara, Y., Karagali, I., Maturi, E., Merchant, C. J., Marullo, S., Minnett, P. J., … Wimmer, W. (2019). Observational Needs of Sea Surface Temperature. Frontiers in Marine Science, 6, 420. https://doi.org/10.3389/fmars.2019.00420

Parard, G., Charantonis, A. A., & Rutgerson, A. (2015). Remote sensing the sea surface CO2 of the Baltic Sea using the SOMLO methodology. Biogeosciences, 12(11), 3369–3384. https://doi.org/10.5194/bg-12-3369-2015

Pasula, A., & Subramani, D. N. (2026). Bias correction with a deep neural model to improve Global Climate Model salinity projections in the Bay of Bengal. PLOS Climate, 5(3), e0000749. https://doi.org/10.1371/journal.pclm.0000749

Qin, S., Zhang, Q., & Yin, B. (2015). Seasonal variability in the thermohaline structure of the Western Pacific Warm Pool. Acta Oceanologica Sinica, 34(7), 44–53. https://doi.org/10.1007/s13131-015-0696-6

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

Roemmich, D., Alford, M. H., Claustre, H., Johnson, K., King, B., Moum, J., Oke, P., Owens, W. B., Pouliquen, S., Purkey, S., Scanderbeg, M., Suga, T., Wijffels, S., Zilberman, N., Bakker, D., Baringer, M., Belbeoch, M., Bittig, H. C., Boss, E., … Yasuda, I. (2019). On the Future of Argo: A Global, Full-Depth, Multi-Disciplinary Array. Frontiers in Marine Science, 6, 439. https://doi.org/10.3389/fmars.2019.00439

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation (arXiv:1505.04597). arXiv. https://doi.org/10.48550/arXiv.1505.04597

Ryan, A. G., Regnier, C., Divakaran, P., Spindler, T., Mehra, A., Smith, G. C., Davidson, F., Hernandez, F., Maksymczuk, J., & Liu, Y. (2015). GODAE OceanView Class 4 forecast verification framework: Global ocean inter-comparison. Journal of Operational Oceanography, 8(sup1), s98–s111. https://doi.org/10.1080/1755876X.2015.1022330

Seo, G., Cho, Y., Choi, B., Kim, K., Kim, B., & Tak, Y. (2014). Climate change projection in the Northwest Pacific marginal seas through dynamic downscaling. Journal of Geophysical Research: Oceans, 119(6), 3497–3516. https://doi.org/10.1002/2013JC009646

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W., & Woo, W. (2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting (arXiv:1506.04214). arXiv. https://doi.org/10.48550/arXiv.1506.04214

Song, T., Xu, G., Yang, K., Li, X., & Peng, S. (2024). Convformer: A Model for Reconstructing Ocean Subsurface Temperature and Salinity Fields Based on Multi-Source Remote Sensing Observations. Remote Sensing, 16(13), 2422. https://doi.org/10.3390/rs16132422

Su, H., Lu, W., Wang, A., & Zhang, T. (2023). AI-Based Subsurface Thermohaline Structure Retrieval from Remote Sensing Observations. In X. Li & F. Wang (Eds.), Artificial Intelligence Oceanography (pp. 105–123). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-6375-9_5

Su, H., Wang, A., Zhang, T., Qin, T., Du, X., & Yan, X.-H. (2021a). Super-resolution of subsurface temperature field from remote sensing observations based on machine learning. International Journal of Applied Earth Observation and Geoinformation, 102, 102440. https://doi.org/10.1016/j.jag.2021.102440

Su, H., Wang, A., Zhang, T., Qin, T., Du, X., & Yan, X.-H. (2021b). Super-resolution of subsurface temperature field from remote sensing observations based on machine learning. International Journal of Applied Earth Observation and Geoinformation, 102, 102440. https://doi.org/10.1016/j.jag.2021.102440

Su, H., Zhang, H., Geng, X., Qin, T., Lu, W., & Yan, X.-H. (2020). OPEN: A new estimation of global ocean heat content for upper 2000 meters from remote sensing data. Remote Sensing, 12(14), 2294.

Su, H., Zhang, T., Lin, M., Lu, W., & Yan, X.-H. (2021). Predicting subsurface thermohaline structure from remote sensing data based on long short-term memory neural networks. Remote Sensing of Environment, 260, 112465. https://doi.org/10.1016/j.rse.2021.112465

Tang, B., Zhao, D., Cui, C., & Zhao, X. (2022). Reconstruction of ocean temperature and salinity profiles in the Northern

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

South China Sea using satellite observations. Frontiers in Marine Science, 9, 945835. https://doi.org/10.3389/fmars.2022.945835

Tang, W., Fore, A., Yueh, S., Lee, T., Hayashi, A., Sanchez-Franks, A., Martinez, J., King, B., & Baranowski, D. (2017). Validating SMAP SSS with in situ measurements. Remote Sensing of Environment, 200, 326–340. https://doi.org/10.1016/j.rse.2017.08.021

Tian, T., Cheng, L., Wang, G., Abraham, J., Wei, W., Ren, S., Zhu, J., Song, J., & Leng, H. (2022). Reconstructing ocean subsurface salinity at high resolution using a machine learning approach. Earth System Science Data, 14(11), 5037–5060. https://doi.org/10.5194/essd-14-5037-2022

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). Attention Is All You Need (arXiv:1706.03762). arXiv. https://doi.org/10.48550/arXiv.1706.03762

Wang, A., Su, H., Huang, Z., & Yan, X.-H. (2024). Knowledge-Informed Deep Learning Model for Subsurface Thermohaline Reconstruction From Satellite Observations. IEEE Transactions on Geoscience and Remote Sensing, 62, 1–16. https://doi.org/10.1109/TGRS.2024.3509616

Wang, H., Song, T., Zhu, S., Yang, S., & Feng, L. (2021). Subsurface Temperature Estimation from Sea Surface Data Using Neural Network Models in the Western Pacific Ocean. Mathematics, 9(8), 852. https://doi.org/10.3390/math9080852

Wang, Q., Zhang, X., Wu, X., Zhang, D., Qi, J., Ning, P., & Qiao, X. (2025). Deep Learning–based Eddy-resolving Reconstruction of Subsurface Temperature and Salinity in the South China Sea. Advances in Atmospheric Sciences, 42(8), 1675–1692. https://doi.org/10.1007/s00376-024-4038-5

Wen, Y., Chen, P., Zhang, Z., & Li, Y. (2024). Cross-Attention-Based High Spatial-Temporal Resolution Fusion of Sentinel-2 and Sentinel-3 Data for Ocean Water Quality Assessment. Remote Sensing, 16(24), 4781. https://doi.org/10.3390/rs16244781

Wilson, C., & Coles, V. J. (2005). Global climatological relationships between satellite biological and physical observations and upper ocean properties. Journal of Geophysical Research: Oceans, 110(C10), 2004JC002724. https://doi.org/10.1029/2004JC002724

Wu, W., Zhai, F., Liu, C., Gu, Y., & Li, P. (2023). Three-dimensional structure of summer circulation in the Bohai Sea and its intraseasonal variability. Ocean Dynamics, 73(11), 679–698. https://doi.org/10.1007/s10236-023-01576-6

Wu, X., Yan, X.-H., Jo, Y.-H., & Liu, W. T. (2012). Estimation of Subsurface Temperature Anomaly in the North Atlantic Using a Self-Organizing Map Neural Network. Journal of Atmospheric and Oceanic Technology, 29(11), 1675–1688. https://doi.org/10.1175/JTECH-D-12-00013.1

Wu, Z., Lu, W., Roobaert, A., Song, L., Yan, X.-H., & Cai, W.-J. (2025). A machine-learning reconstruction of sea surface p CO2 in the North American Atlantic Coastal Ocean Margin from 1993 to 2021. Earth System Science Data, 17(1), 43–63. https://doi.org/10.5194/essd-17-43-2025

Xie, H., Dong, C., & Xu, Q. (2025). Dual U–Vision–Transformer for reconstructing the three-dimensional eddy-resolving oceanic physical parameters from satellite observations. International Journal of Applied Earth Observation and Geoinformation.

Xie, H., Xu, Q., Cheng, Y., Yin, X., & Fan, K. (2023). Reconstructing three-dimensional salinity field of the South China Sea from satellite observations. Frontiers in Marine Science, 10, 1168486. https://doi.org/10.3389/fmars.2023.1168486

Xie, H., Xu, Q., Cheng, Y., Yin, X., & Jia, Y. (2022). Reconstruction of Subsurface Temperature Field in the South China Sea From Satellite Observations Based on an Attention U-Net Model. IEEE Transactions on Geoscience and Remote Sensing, 60, 1–19. https://doi.org/10.1109/TGRS.2022.3200545

Yan, X., Boyer, T., Trenberth, K., Karl, T. R., Xie, S., Nieves, V., Tung, K., & Roemmich, D. (2016). The global warming hiatus: Slowdown or redistribution? Earth's Future, 4(11), 472–482. https://doi.org/10.1002/2016EF000417

Yang, G. G., Wang, Q., Feng, J., He, L., Li, R., Lu, W., Liao, E., & Lai, Z. (2024). Can three-dimensional nitrate structure be reconstructed from surface information with artificial intelligence? — A proof-of-concept study. Science of The Total Environment, 924, 171365. https://doi.org/10.1016/j.scitotenv.2024.171365

Yu, X., Xiao, B., Liu, X., Wang, Y., Cui, B., & Liu, X. (2017). Retrieval of remotely sensed sea surface salinity using MODIS data in the Chinese Bohai Sea. International Journal of Remote Sensing, 38(23), 7357–7373. https://doi.org/10.1080/01431161.2017.1375570

Zhang, J. (2010). Multi-source remote sensing data fusion: Status and trends. International Journal of Image and Data Fusion, 1(1), 5–24. https://doi.org/10.1080/19479830903561035

https://doi.org/10.5194/essd-2026-461
Preprint. Discussion started: 29 June 2026

Zhang, J., Zhang, X., Wang, X., Ning, P., & Zhang, A. (2023). Reconstructing 3D ocean subsurface salinity (OSS) from T–S mapping via a data-driven deep learning model. Ocean Modelling, 184, 102232. https://doi.org/10.1016/j.ocemod.2023.102232

Zhang, K., Geng, X., & Yan, X.-H. (2020). Prediction of 3-D Ocean Temperature by Multilayer Convolutional LSTM. IEEE Geoscience and Remote Sensing Letters, 17(8), 1303–1307. https://doi.org/10.1109/LGRS.2019.2947170

Zhang, L., Sun, C., & Hu, D. (2012). Relationship between oceanic heat content and sea surface height on interannual time scale. Chinese Journal of Oceanology and Limnology, 30(6), 1026–1032. https://doi.org/10.1007/s00343-012-1247-z

Zhu, Q., Li, H., Sun, H., Xia, T., Wang, X., & Han, Z. (2025). 3DV-Unet: Eddy-Resolving Reconstruction of Three-Dimensional Upper-Ocean Physical Fields from Satellite Observations. Remote Sensing, 17(19), 3394. https://doi.org/10.3390/rs17193394

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