https://doi.org/10.5194/essd-2025-699
Preprint. Discussion started: 1 December 2025

GEOXYGEN: a global long-term dissolved oxygen dataset based on biogeochemistry-aware machine learning framework and multi-source observations

Zhenguo Wang[1], Weiwei Fu[1,2], Cunjin Xue[3,4], Guihua Wang[1]

[1]Department of Atmospheric and Oceanic Sciences, Fudan University, Shanghai, 200438, China

[2] Institute of Eco-Chongming (IEC), 1050 Baozhen, Lühua Town, Chongming District, Shanghai 202151, China

[3]International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China;

[4]Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

Correspondence to: Weiwei Fu (wwfu@fudan.edu.cn)

Abstract. Dissolved oxygen (DO) serves as an essential indicator of marine ecosystem health. However, sparse and uneven observations have limited our ability to characterize its full spatiotemporal variability, underscoring the continued need for long-term, high-resolution, and physically consistent global DO datasets. Here, we present GEOXYGEN, a global dataset of monthly DO fields at 0.5° × 0.5° resolution spanning 1960–2024 and depths from the surface to 5500 m (Wang et al., 2025, https://doi.org/10.5281/zenodo.17615657). GEOXYGEN is generated with a hierarchical modeling framework that accounts for regional and vertical heterogeneity. By integrating physical and biogeochemical predictors with an adaptive feature-selection strategy, GEOXYGEN achieves high predictive accuracy across all depth layers on an independent out-of-time test ($R^2 > 0.92$). The reconstructed spatial patterns align closely with the World Ocean Atlas 2023 climatology, and in subsurface and deep waters, GEOXYGEN demonstrates superior generalization relative to existing data-driven products. A sensitivity analysis further reveals that including coastal data in model training increases basin-wide uncertainty by approximately 7.5%, underscoring that current observing systems remain insufficient to reliably resolve nearshore DO dynamics. GEOXYGEN provides a consistent, physically informed baseline for analyzing global and regional variability of DO. It also offers a valuable benchmark for evaluating and improving the representation of DO in climate and Earth system models and can support future studies on long-term deoxygenation trends and regional hotspots.

Keywords: Dissolved oxygen, Machine learning, Hierarchical modeling, Long-term dataset, Coastal uncertainty

1 Introduction

Ocean dissolved oxygen (DO) concentration serves as an essential indicator of marine ecosystem health and biogeochemical status (Robinson, 2019; Grégoire et al., 2023). Beyond its ecological significance, DO plays a critical role in modulating climate-relevant biogeochemical feedbacks in the global carbon cycle (Gregoire et al., 2021; Oschlies, 2021; Yamaguchi et al., 2024). Observations over recent decades reveal marked spatiotemporal variability in DO, accompanied by a clear trend toward deoxygenation (Ito et al., 2017), particularly within tropical oxygen minimum zones (OMZs) and in subsurface waters at mid to high latitudes (Bopp et al., 2013; Li et al., 2020). This loss of oxygen is projected to persist under continued global warming (Gong et al., 2021; Zhou et al., 2022), with growing consequences for marine habitats, fisheries, and ecosystem services (Breitburg et al., 2018; Kim et al., 2023; Chen et al., 2024; Humphries et al., 2024).

Sparse and heterogeneous observational coverage hampers an accurate estimate of the global oxygen inventory and the quantification of its long-term changes. A seminal study by Schmidtko et al. (2017) estimates a 2% decline (4.8 ± 2.1 Pmol) in the global ocean oxygen inventory from 1960 to 2009. Yet, the accuracy of such an estimate depends heavily on the observations used. Historically, DO measurements have been sourced from ship-based campaigns compiled in global databases such as the World Ocean Database (WOD) and the Global Ocean Data Analysis Project (GLODAP), which exhibit strong spatial and temporal sampling biases (Garcia et al., 1998). The resulting unevenness in data availability across time, space, and quality standards, especially in coastal waters, complicates robust quantification of deoxygenation rates, particularly in dynamic and vulnerable systems such as coastal shelves and polar oceans. These limitations underscore the pressing need for a spatially continuous, long-term, and accurate global DO reconstruction.

Multiple approaches have been developed to address these observational gaps. Earth system models (ESMs) simulate four-dimensional DO fields continuously but often suffer from systematic biases and incomplete representation of multi-scale processes (Cocco et al., 2013; Oschlies et al., 2018). Limited observational constraints further compound uncertainties in model evaluation and in projections. Traditional statistical interpolation methods can reproduce mean climatologies but frequently underestimate trends in data-sparse regions and fail to capture seasonal to interannual variability (Ito et al., 2024b; Gouretski et al., 2024b). In recent years, data-driven machine learning (ML) has emerged as a promising alternative, leveraging relationships between DO and physical or biogeochemical covariates to reconstruct continuous four-dimensional fields from sparse in situ measurements (Sharp et al., 2023; Garabaghi et al., 2023; Huang et al., 2023; Wang et al., 2024; Lu et al., 2024). In principle, ML can recover local variability and identify deoxygenation risk without relying on computationally expensive coupled simulations.

Despite this potential, several methodological challenges remain. First, many existing ML reconstructions employ a single model trained on global-ocean data, which struggles to represent regional variations in the dominant physical–biogeochemical controls on DO as well as the spatial heterogeneity of processes across biogeochemical provinces (Garabaghi et al., 2023).

This global approach tends to blur contrasts between water-mass regimes and degrade skill in dynamically distinct regions such as oxygen minimum zones and boundary currents. Second, a common workflow is to first reconstruct DO at scattered profile locations and then interpolate these point estimates onto a regular grid, often using a limited set of predictors such as temperature and salinity (Sharp et al., 2023; Wang et al., 2024; Liu et al., 2025). In data-sparse regions, this two-step procedure encourages extrapolation, propagates local errors, and can generate spurious fine-scale structure that is not supported by the underlying observations, particularly near sharp DO gradients and in historically undersampled basins. Third, training models directly on raw profiles amplifies sampling biases: autonomous platforms such as Argo repeatedly sample specific regions and depth ranges, whereas historical ship-based surveys are concentrated along cruise tracks (Huang et al., 2023; Lu et al., 2024). Without explicit rebalancing or weighting, ML models place disproportionate emphasis on well-observed areas and generalize poorly elsewhere, leading to reconstructions that systematically underrepresent variability and trends in data-poor regions.

At the dataset level, available ML-based global DO products, such as GOBAI-O_2 (Sharp et al., 2023), G4D-DOC (Xue et al., 2024), and ML4O_2 (Ito et al., 2024a), represent important advances, providing monthly gridded DO fields at 1° resolution over multi-year to multi-decadal periods and resolving much of the upper and intermediate ocean. However, they are generally limited either to the Argo era or to depths shallower than about 1000–2000 m. To our knowledge, there is currently no single observation-based product that combines pre-Argo coverage from the 1960s, full-depth global fields, and sub-degree horizontal resolution.

To address these methodological and dataset-level limitations, we generated GEOXYGEN, a monthly global DO dataset at 0.5° × 0.5° resolution on 187 depth levels from 1960 to 2024 (Wang et al., 2025, https://doi.org/10.5281/zenodo.17615657). The dataset is generated by combining a global compilation of in situ DO profiles with objectively analyzed temperature–salinity fields and related sea-surface environmental variables, and by learning their relationships with DO through a regionally structured, depth-aware, and adaptively constrained machine-learning framework. Our approach explicitly accounts for strong spatiotemporal heterogeneity by partitioning the ocean into ecogeographical macro-regions and training separate submodels for different depth layers. Within each region–depth unit, we adaptively select predictive features from a suite of variables including temperature, salinity, oxygen saturation, physical indicators, carbonate-system parameters, and bio-optical properties, thereby ensuring physical interpretability while minimizing redundancy. To mitigate sampling bias and discontinuities, we implement inverse-density weighting, year-grouped cross-validation, and cross-boundary fusion techniques. The resulting GEOXYGEN product provides a consistent, long-term, and spatially complete representation of global DO suitable for quantifying global and regional deoxygenation, diagnosing underlying drivers, and evaluating Earth system and biogeochemical models.

2 Data

2.1 In situ oxygen observations

To support a data-driven reconstruction of dissolved oxygen (DO), we compiled a global collection of in situ measurements by integrating six major international data sources: the CLIVAR and Carbon Hydrographic Data Office (CCHDO), the Global Ocean Data Analysis Project (GLODAP), the GEOTRACES Intermediate Data Product 2021 (IDP2021), Biogeochemical Argo (BGC-Argo), the World Ocean Database 2023 (WOD23; Mishonov et al., 2024), and the OceanSITES fixed-point observatory network. This compilation spans the period 1950–2024 and includes millions of DO profiles collected via bottle casts (OSD), Conductivity–Temperature–Depth (CTD) profilers, and Argo floats.

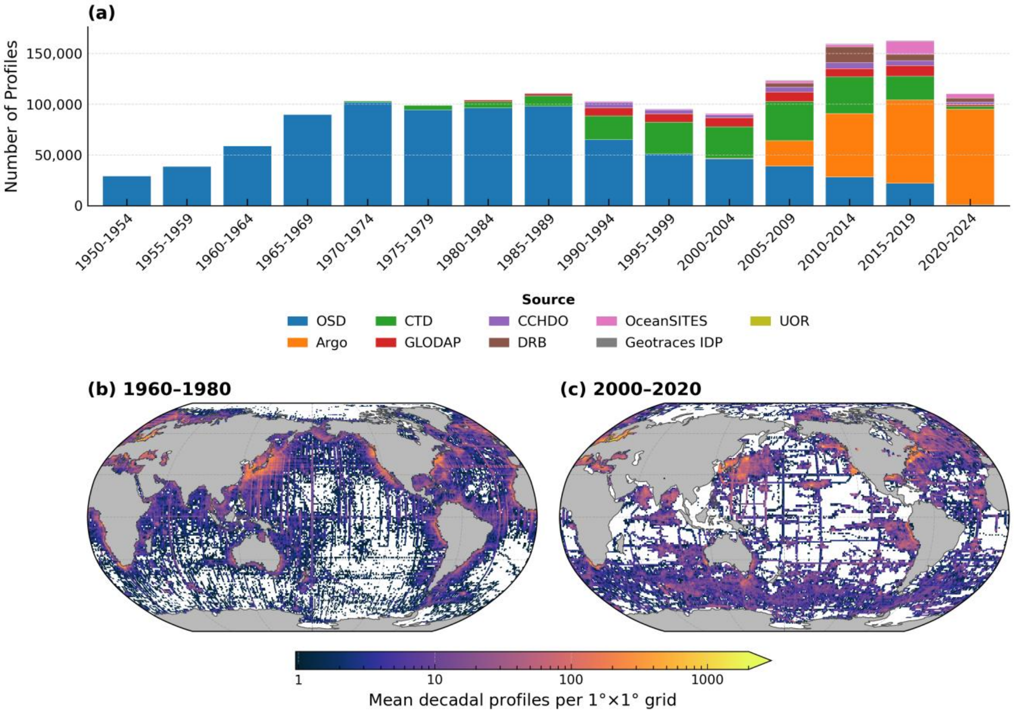


Figure 1: Global distribution and temporal coverage of DO profiles. (a) Changes in the number of profiles from major data sources during 1950–2024. (b–c) Spatial distribution of decadal-mean profile counts for 1960–1980 and 2000–2020, computed on a 1° × 1° grid, showing a transition from predominantly Northern Hemisphere sampling to much denser coverage in the Southern Hemisphere. The color bar indicates the decadal-mean number of profiles per grid cell (log scale).

To ensure dataset consistency, we implemented a multi-stage quality-control (QC) procedure. First, we standardized formats and units across sources and retained only measurements flagged as "good/probably good." Second, we removed duplicates from different sources on the same date and at the same location with vertical separation less than 1 m. Third, we applied a physiologically plausible DO threshold of 0–600 μmol kg^{-1} to exclude outliers. Although differences in QC protocols and metadata conventions across datasets may introduce cross-source biases (Gregoire et al., 2021), prior assessments have indicated no significant systematic offsets between Winkler-calibrated high-accuracy data and sensor measurements at the global scale (Schmidtko et al., 2017). Therefore, we did not apply cross-source corrections. Instead, we focused on local outlier detection at each standard depth level, whereby within each 0.5° × 0.5° grid cell and a ±10-day time window, observations deviating by more than 3σ from the local mean were excluded. Across depth levels, removal rates range from 0 to 0.105%, with most levels below 0.08%, indicating that only a very small fraction of the quality-controlled observations was flagged as outliers.

Following QC, the final observational archive for 1950–2024 comprises approximately 1.4 million DO profiles (Fig. 1a). This full 1950–2024 record is used for model training and evaluation, whereas the reconstructed DO fields are produced over 1960–2024. The OSD, CTD, and Argo collectively account for ~91.15% of all profiles. Vertical sampling distribution has evolved with technological advances: OSD dominates in surface waters (~80% above 100 m) but declines to less than 10% by 500 m, whereas Argo represents over half of all profiles below 500 m and is the primary source in the mesopelagic layer. This shift reflects broader observational trends: discrete bottle sampling prevailed through the 1960s, CTD profiling expanded in the 1970s–1990s through programs such as GLODAP and CCHDO, and autonomous float observations increased by an order of magnitude after 2005. By 2010, Argo became the dominant platform in the open ocean.

Spatial coverage has shifted from a strongly Northern Hemisphere bias to a more Southern Hemisphere focus, but remains highly uneven (Fig. 1b–c). From 1960 to 1980, sampling was concentrated along ship tracks in the Northern Hemisphere, with limited data in the Southern Hemisphere and basin interiors. Quantitatively, analysis of our profile dataset shows that the number of Southern Hemisphere profiles increased from roughly one quarter of the Northern Hemisphere count in 1960–1980 to about 2.4 times the Northern Hemisphere count in 2000–2020, reflecting a pronounced shift of observing effort into the Southern Hemisphere. Since 2000, autonomous platforms have substantially improved coverage across the Southern Ocean and the open ocean, partly alleviating the historical Southern Hemisphere undersampling. Nevertheless, marginal seas and high-latitude ice-covered regions remain undersampled. In our reconstruction approach, we mitigate these biases through regionalized modeling, gridding to standard depth levels, and inverse-density weighting.

2.2 Physical–biogeochemical variables

Accurately predicting ocean deoxygenation requires denser, longer DO observations, and drivers that characterize physical transport and cross-scale biogeochemical processes (Oschlies et al., 2018). To meet this need, we complement objectively analyzed temperature–salinity fields with a curated suite of sea-surface environmental variables (SSEVs), providing physically consistent and process-informed predictors for DO reconstruction (Table 1).

We derive three-dimensional temperature and salinity fields from the Coriolis Ocean Dataset for Reanalysis (CORA) (Szekely et al., 2025), a CMEMS objective analysis that compiles in situ temperature and salinity using the ISAS objective-mapping system, integrating observations from ships, Argo floats, and other in situ platforms. This dataset undergoes delayed-mode quality control to ensure long-term stability and global coherence. Using the TEOS-10 standard, we further calculate oxygen saturation (O_2_sat) from CORA temperature and salinity data (IOC, 2010); its deviation from observed DO reflects biological respiration and physical mixing.

In parallel, we assemble SSEVs spanning thermodynamic, dynamical, bio-optical, and carbon-chemistry processes, supporting data-driven reconstruction (Shao et al., 2024; Ma et al., 2025). Wind vectors (zonal and meridional components, U and V) are taken from NASA's Cross-Calibrated Multi-Platform (CCMP) product (Mears et al., 2022). Mixed-layer depth (MLD) is obtained from the CMEMS Multi-Observation Global Ocean 3D product (Guinehut et al., 2012). Dynamical variables include sea surface height (SSH) and eddy kinetic energy (EKE), both derived from AVISO satellite altimetry (Hauser et al., 2020). Bio-optical variables comprise photosynthetically active radiation (PAR) and chlorophyll a (Chl-a) from NASA Level-3/Level-4 ocean-color products (NASA Ocean Biology Processing Group, 2018). Carbon-chemistry variables include dissolved inorganic carbon (DIC), total alkalinity, pH, sea surface partial pressure of CO_2 (pCO_2), and CO_2 flux, all obtained from the CMEMS Surface Ocean Carbon Fields product (Chau et al., 2022; Chau et al., 2024). All variables were last accessed in March 2025.

For consistency across sources, we regridded each variable to a uniform 0.5° × 0.5° grid, aggregated them to a monthly resolution, and aligned them with both the DO observation locations and the reconstruction grid. This standardization ensures coherent feature availability during model training and prediction while retaining large-scale physical and biogeochemical signals essential for reconstructing the deoxygenation trend. To preserve historical information and reduce imputation uncertainty, we retain all QC-passed DO data, even if some covariates (e.g., SSEVs) are missing. This sampling library provides standardized inputs for subsequent model training.

Table 1. Details of the Physical–biogeochemical variables

Variable	Description	Spatial Resolution	Temporal Resolution	Temporal Coverage	Data Source
Temperature (°C)	Seawater temperature	0.25° × 0.25°; 187 standard depth levels (surface–5500 m)	Monthly	1960/01 – 2024/06	(Szekely et al., 2025)
Salinity	Seawater salinity	0.25° × 0.25°; 187 standard depth levels (surface–5500 m)	Monthly	1960/01 – 2024/06	
O_2_sat	Seawater oxygen saturation (calculated)	0.25° × 0.25°; 187 standard depth levels (surface–5500 m)	Monthly	1960/01 – 2024/06	

U (m s^{-1})	U-wind vector component at 10 meters	0.25° × 0.25°	Monthly	1993/01 – 2023/08	(Mears et al., 2022)
V (m s^{-1})	V-wind vector component at 10 meters				
MLD (m)	Ocean mixed layer depth	0.25° × 0.25°	Monthly	1993/01 – 2022/12	(Guinehut et al., 2012)
DIC (µmol kg^{-1})	Surface ocean dissolved inorganic carbon	0.25° × 0.25°	Monthly	1985/01 – 2023/12	(Chau et al., 2022; Chau et al., 2024)
pH	Surface pH on total scale				
pCO_2 (µatm)	Surface aqueous partial pressure of CO_2				
CO_2 flux (mol m^{-2} yr^{-1})	Surface downward flux of total CO_2				
Alkalinity (µmol kg^{-1})	Total alkalinity in surface seawater				
PAR (mol m^{-2} d^{-1})	Photosynthetically available radiation	4 km / 9 km	Monthly	1997/10 – 2025/2	(NASA Ocean Biology Processing Group, 2018)
Chl-a (mg m^{-3})	Mass concentration of chlorophyll in surface water				
SSH (m)	Sea surface height above geoid	0.25° × 0.25°	Monthly	1993/01 – 2023/08	(Hauser et al., 2020)
EKE (cm^2 s^{-2})	Surface averaged eddy kinetic energy				

2.3 Depth-dependent relationships between oxygen and drivers

We characterized the depth-dependent relationships between dissolved oxygen (DO) and a suite of environmental variables

(Fig. 2). Near the surface, DO exhibits a strong negative correlation with temperature, consistent with solubility control, while salinity contributes via the salting-out effect. With increasing depth, the direct influence of solubility declines, yet temperature and salinity remain valuable as proxies for water-mass identity and ventilation history. For reference, we also show correlations with TEOS-10 derived oxygen saturation (O_2_sat), which serves as a physicochemical baseline reflecting equilibrium conditions.

Sea-surface environmental variables (SSEVs) display distinct vertical patterns in their coupling with DO. Sea surface height (SSH) maintains a strong and stable correlation through much of the water column, while surface wind components (U, V) contribute dynamical context related to wind-driven advection and upwelling (Hollitzer et al., 2024). Carbonate system variables retain explanatory power below the euphotic zone, indicative of remineralization signals and association with specific water masses. In contrast, Chl-a is mainly informative within the upper tens of metres, reflecting its biogeochemical role in the sunlit layer. Several SSEVs also display notable nonlinear relations with subsurface DO, which supports their value for reconstructing underwater DO (Ping et al., 2024; Cao et al., 2024).

Overall, these correlation patterns illustrate how surface forcings and water-mass structure jointly shape subsurface oxygen variability. They provide a theoretical and empirical foundation for subsequent modeling of the often nonlinear relationships between these variables and DO.

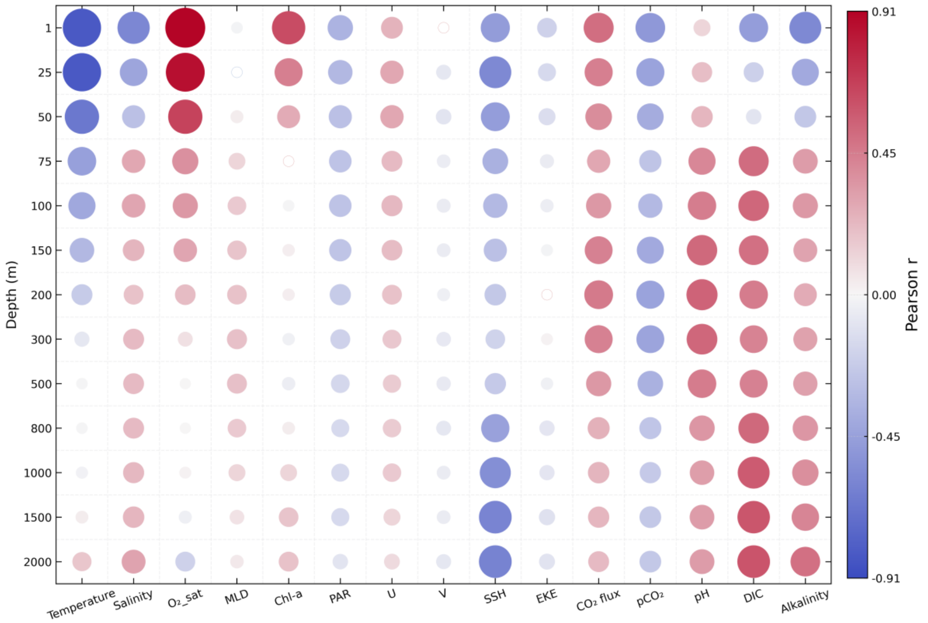


Figure 2: Vertical correlations between DO and physical–biogeochemical variables. Each bubble represents the Pearson correlation coefficient ((r)) between DO and an environmental variable at a standard depth level. Bubble color encodes the sign and magnitude (red = positive; blue = negative), and bubble area scales with (|r|). Filled bubbles denote correlations significant at (q<0.05) after Benjamini–Hochberg false-discovery-rate control; hollow bubbles indicate non-significant results.

3 Method

Here, we develop a hierarchical modeling framework that resolves regional and vertical heterogeneity by partitioning the ocean into ecogeographic regions and training depth-stratified, region-specific CatBoost models with adaptive feature selection and year-grouped cross-validation.

3.1 Partitioning based on heterogeneity

Dissolved oxygen (DO) trends exhibit substantial variability across ocean basins, latitude bands, and depth layers, with dominant controlling mechanisms shifting both regionally and vertically (Ma et al., 2025). To capture regional physical–biogeochemical coupling while mitigating sampling biases, we implement a modeling framework that is stratified in both the vertical and horizontal dimensions.

Vertical stratification. Given the depth dependence of processes governing DO distribution, we divide the 0–5500 m water column into 187 standard depth levels, following the vertical grid of the CORA objective analysis and ISAS17 climatology (Szekely et al., 2019; Kolodziejczyk et al., 2023). An independent model is trained at each depth, allowing the framework to resolve depth-varying controls while preserving realistic vertical structure.

Horizontal regionalization. To account for spatial heterogeneity in the physical–biogeochemical controls on DO, we follow the global ocean biogeochemical province classification of Fay and Mckinley (2014) and refine it using basin boundaries and the statistical robustness of sample distributions, allowing region-specific submodels to better represent distinct DO–environment relationships. Gaps in the original map are filled by nearest-neighbor interpolation constrained by ocean connectivity, combined with a KD-tree projection. To avoid local complexities, we exclude enclosed or exchange-limited seas (e.g., the Mediterranean Sea and the Red Sea) and mask coastal waters shallower than the 200 m isobath. A sensitivity analysis of this coastal threshold is presented in Sect. 4.3.

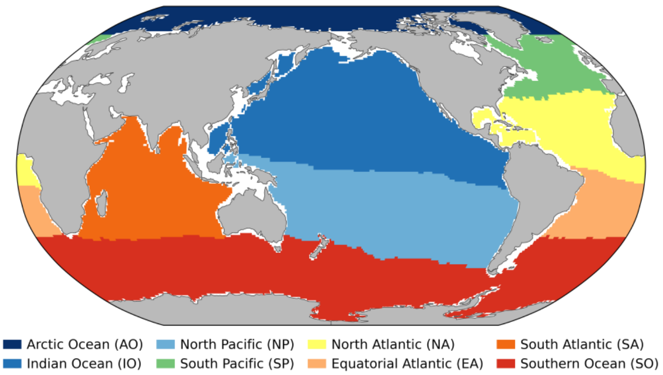


Figure 3: Partitioning of the global open ocean into eight macro biogeochemical provinces.

The open ocean is partitioned into eight macro biogeochemical provinces (Fig. 3): the Arctic Ocean (AO), North Pacific (NP), South Pacific (SP), North Atlantic (NA), Equatorial Atlantic (EA), South Atlantic (SA), Indian Ocean (IO), and Southern Ocean (SO). Each province serves as an independent modeling unit, enabling regionally tailored representation of key processes, reducing spurious extrapolation across biogeochemical domains, and introducing structured constraints that enhance the physical consistency of the reconstructed DO fields.

3.2 Adaptive modeling

Within each modeling unit, we employ the CatBoost gradient-boosting framework—an interpretable decision-tree method—to learn the functional mapping between sparse dissolved oxygen (DO) observations and their environmental predictors. This method builds ensembles of oblivious trees with ordered boosting, a training scheme that reduces variance and limits target leakage. The algorithm offers several advantages for our application: it directly accommodates missing covariates without imputation, incorporates sample weighting to address spatial biases, and employs multiple regularization strategies—including L2 regularization on leaf values, subsampling, and early stopping—to prevent overfitting.

For historical periods (e.g., 1960–1997) where satellite-derived predictors are unavailable, the model treats missing covariate values as a distinct input state, allowing it to rely primarily on temperature, salinity, oxygen saturation, and spatiotemporal coordinates. This strategy avoids introducing systematic biases that can arise from statistical imputation and maintains consistency across the reconstruction period. The CatBoost model uses 19 predictor variables as inputs. They include: temperature, salinity, O_2_sat, year_norm, month_sin, month_cos, Latitude, U, V, SSH, EKE, MLD, PAR, Chl-a, DIC, pCO_2, pH, Alkalinity, and CO_2 flux. The temporal encodings are defined in Eq. (1)-(3), where days_since_start denotes the number of days elapsed since 1950-01-01 00:00:00 UTC.

$$\text{year\_norm} = \frac{\text{days\_since\_start}}{365.25}, \quad (1)$$

$$\text{month\_sin} = \sin\left(\frac{2\pi\,(m-1)}{12}\right), \quad (2)$$

$$\text{month\_cos} = \cos\left(\frac{2\pi\,(m-1)}{12}\right), m \in \{1,2,\ldots,12\}, \quad (3)$$

To enhance model interpretability and generalization, we implemented a two-stage feature selection procedure for each region–depth submodel, using only training data and a year-grouped cross-validation scheme consistent with our model evaluation framework. Tree-based models tend to perform better with compact, informative feature sets, as redundant predictors can dilute predictive accuracy (Garabaghi et al., 2023). First, we estimated permutation importance under five-fold cross-validation grouped by year and retained an initial subset of features using an adaptive rule $K = \max(10, 2\sqrt{p})$, where $p = 19$ is the number of candidate features, discarding predictors with negligible contribution. Second, we perform recursive feature elimination with cross-validation (RFECV) using the same year-grouped folds, iteratively removing the least important feature and selecting the combination that minimizes validation RMSE. The five independent test years are excluded from both stages.

This design allows feature sets to adapt to regional and vertical regimes—surface models emphasise rapidly varying surface and biogeochemical terms, mid-depth models highlight meridional gradients and water-mass transition metrics, and deep models rely primarily on temperature–salinity structure, while avoiding temporal information leakage and reducing overfitting.

To mitigate biases arising from heterogeneous spatiotemporal sampling (Fig. 1), we applied inverse-density weighting within a fixed binning scheme. The sample domain is partitioned on a 5° × 5° latitude–longitude grid and non-overlapping 10-year time windows in each partition. Sample weights were computed as the inverse of the observational density within each spatiotemporal bin and standardized at each depth level, thereby reducing the influence of over-sampled regions and periods during model training.

Let $b(i)$ denote the spatiotemporal bin containing sample i, and let $n_{b(i)}$ be the number of samples in that bin. The initial per-sample weight is the inverse square root of this count (Eq. (4)):

$$\tilde{w}_i = \frac{1}{\sqrt{n_{b(i)}}}, \tag{4}$$

Specifically, each sample weight is set proportional to the inverse square root of the local sample density. To preserve the aggregate information content, we then normalize the weights to unit mean (Eq. (5)):

$$w_i = \frac{\tilde{w}_i}{\frac{1}{N}\sum_{k=1}^{N}\tilde{w}_k}, \tag{5}$$

This strategy increases the influence of observations from sparse regions and earlier periods without altering the aggregate sample distribution.

3.3 Hyperparameter optimization and validation

To efficiently explore the hyperparameter space and avoid ad hoc manual tuning, hyperparameters for each regional modeling unit were optimized independently using Bayesian optimization (Optuna; Table 2) with the objective of minimizing validation RMSE. We employed a five-fold cross-validation scheme grouped by calendar year, wherein the observational record was divided into five non-overlapping temporal blocks. Each block serves as a validation fold (~20%), ensuring that data from the same year never appears in both training and validation splits. This design prevents temporal information leakage and avoids inflated performance estimates that can arise from autocorrelated observations within narrow time windows (Salazar et al., 2022). During each fold, an early stopping halts training if validation RMSE fails to improve for 50 consecutive iterations and determines the optimal boosting rounds. The final model used the median of the optimal iteration counts across folds. Early stopping was then disabled, and the model was retrained on the full training set prior to reconstruction.

To provide a final, unbiased assessment of model performance, we construct an independent out-of-time test set by randomly selecting five calendar years—1964, 1972, 1985, 2012, and 2019 (without replacement)—and withholding them entirely from training and hyperparameter optimization. Generalization capability was evaluated on this holdout set using RMSE, mean absolute error (MAE), and the coefficient of determination (R^2), providing robust and interpretable metrics of predictive accuracy and strengthening confidence in the reconstruction. These five withheld years are used in Sect. 4.1, 4.2 and 4.5 for independent evaluation and product intercomparison.

Table 2. CatBoost hyperparameters and their Optuna prior search spaces

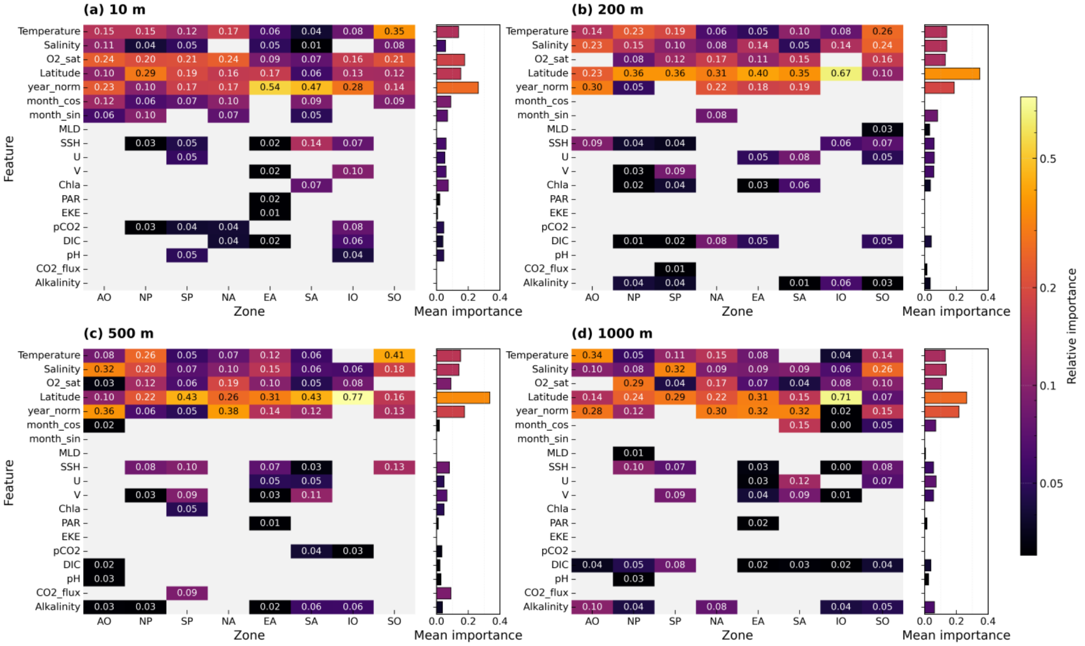
Hyperparameter	Explanation	Search range
iterations	Maximum boosting rounds	100–2000
learning_rate	Learning rate (shrinkage)	0.02–0.10
depth	Tree depth	4–8
l2_leaf_reg	L2 regularization on leaf values	2–10
bagging_temperature	Temperature for Bayesian bootstrap	0.05–1.0

4 Results and Discussion

4.1 Feature Importance

Under the adaptive feature-selection framework, each biogeographic region develops depth-specific optimal predictor sets (Fig. 4). Across all region–depth submodels, the final feature sets typically contain 5-10 predictors. At 10 m, year_norm is the dominant control. Among sea-surface environmental variables (SSEVs), sea surface height (SSH) and partial pressure of CO_2 (pCO_2) consistently enhance model performance in multiple regions. In the Indian Ocean, for instance, SSH and meridional wind velocity (V) account for substantial DO variance—a pattern consistent with wind-driven transport and upwelling processes that modulate surface oxygen concentrations. Similarly, surface pCO_2 patterns have been shown to closely reflect the upwelling of carbon-rich, oxygen-poor waters (Franco et al., 2014), underscoring its utility as a predictor of DO variability in upwelling and OMZ regions.

At 200 and 1000 m, latitude replaces year_norm as the most influential variable, particularly in the Indian Ocean. This result suggests that latitude serves as an effective spatial proxy, capturing broad-scale oxygen gradients and the influence of unobserved physical or biogeochemical drivers. Under this dominant latitudinal effect, the marginal contributions of most SSEVs are generally subdued, though not negligible (Milà et al., 2024). In summary, although SSEVs contribute little on average across depth levels, they remain important predictors in specific regions at certain depths. These results demonstrate that region and depth jointly shape the effective input space. Adaptive regional modeling preserves key controls and removes redundancy, improving both accuracy and physical interpretability. Feature subsets are not universal across regions, supporting the regionalized approach's advantage and smaller bias on the independent test. Our feature selection provides mechanistic evidence for these findings.



285 **Figure 4: Heatmap of relative feature importance across depths and provinces. Colors are on a logarithmic scale. The bar chart on the right shows each feature's mean importance computed over the provinces in which that feature is available.**

4.2 Model evaluation

The reconstructed dissolved oxygen (DO) fields demonstrate high accuracy across depth layers and biogeographic regions. Overall model performance was evaluated by aggregating predictions from all regions and comparing them with observed 290 values (Fig. 5). The model demonstrates high accuracy in estimating DO concentrations, with the scatter plot closely following the 1:1 line. At representative depths, the global RMSE/MAE on the independent test set is on the order of 14.6/8.7 μmol kg^{-1} at 10 m, 17.1/11.3 μmol kg^{-1} at 200 m, and 8.7/5.2 μmol kg^{-1} at 1000 m. Most R^2 values are above 0.92, and the R^2 for deep-layer DO reconstruction reaches 0.99. A slight positive bias is noted under severely hypoxic conditions, where extremely low DO values are marginally overestimated—a known behavior of regularized tree ensembles that tend to smooth outliers toward 295 local means. These cases are rare, confined to strongly hypoxic conditions, and have negligible impact on large-scale statistics. It is also worth noting that some of the lowest recorded DO values may reflect measurement uncertainty. Residuals are generally symmetric and centered near zero, indicating that the model produces unbiased estimates with only occasional negative outliers.

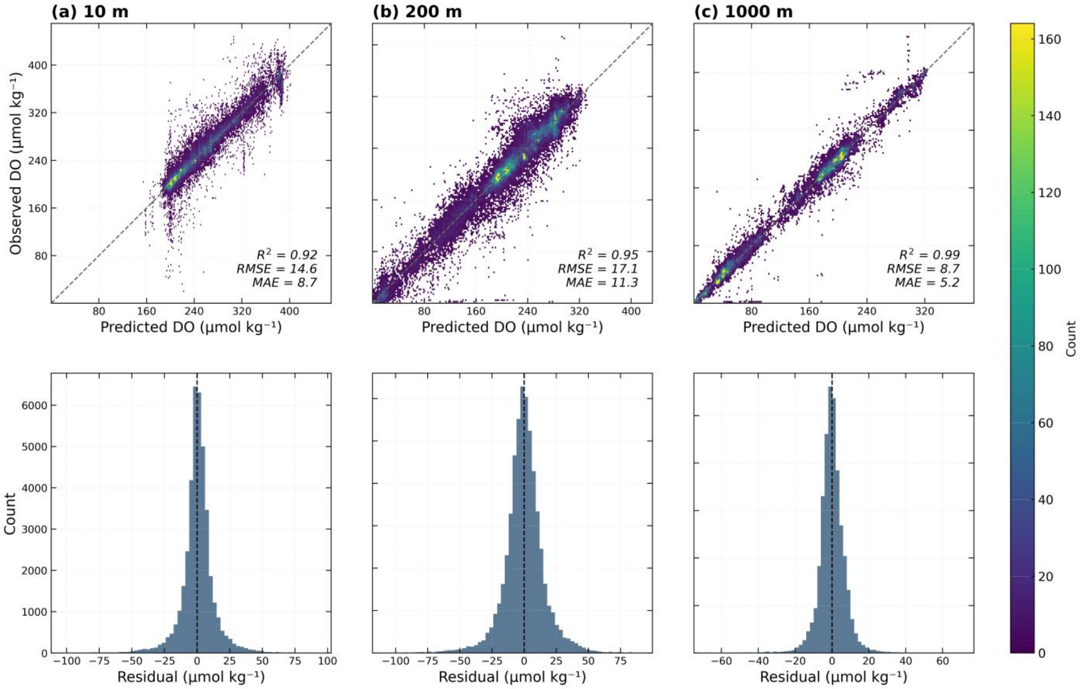


Figure 5: Model performance for DO predictions across depth layers (independent test set). Top row: hexagon-binned scatterplots of predicted vs. observed values; the gray dashed line denotes the 1:1 reference. The color bar indicates sample counts per hexbin. Each panel reports the R², RMSE, and MAE. Bottom row: corresponding histograms of residuals (observed − predicted).

Vertical error structure exhibits consistent stratification across regions (Fig. 6): RMSE is highest in the surface and thermocline layers, decreasing with depth, consistent with a strong vertical gradient in the thermocline and the relative stability of intermediate to deep water masses. This pattern underscores the role of water mass stratification in shaping the estimation uncertainty of DO. Regional differences are primarily reflected in the position and amplitude of mid-layer peaks. Regarding error profile patterns, stable regions (e.g., SP–IO, NA–SA) show consistently high R², except at the surface, with RMSE monotonically or nearly monotonically decreasing with depth. In mid-layer-sensitive regions (e.g., NP), an RMSE peak is observed around 100–600 m, although R² remains relatively high. Lower surface R² values in the South Pacific, Equatorial Atlantic, and Indian Ocean arise primarily from limited DO variance in those layers rather than increased absolute error.

Deep layers, constrained by water masses, show a consistent pattern of low errors and high correlations. In the thermocline and OMZs, water properties and boundary positions are more sensitive to rapid seasonal and mesoscale–submesoscale fluctuations (Bettencourt et al., 2015). Together with limitations in historical coverage and resolution, this increases reconstruction uncertainty in these regions. Performance differences among regional models further confirm the pronounced spatial heterogeneity of DO. In summary, the regionalized modeling framework generalizes effectively, capturing the large-scale DO distribution with high overall skill. Errors are primarily concentrated in regions and depths where oxygen gradients are strongest, aligning with expected patterns of physical and biogeochemical complexity.

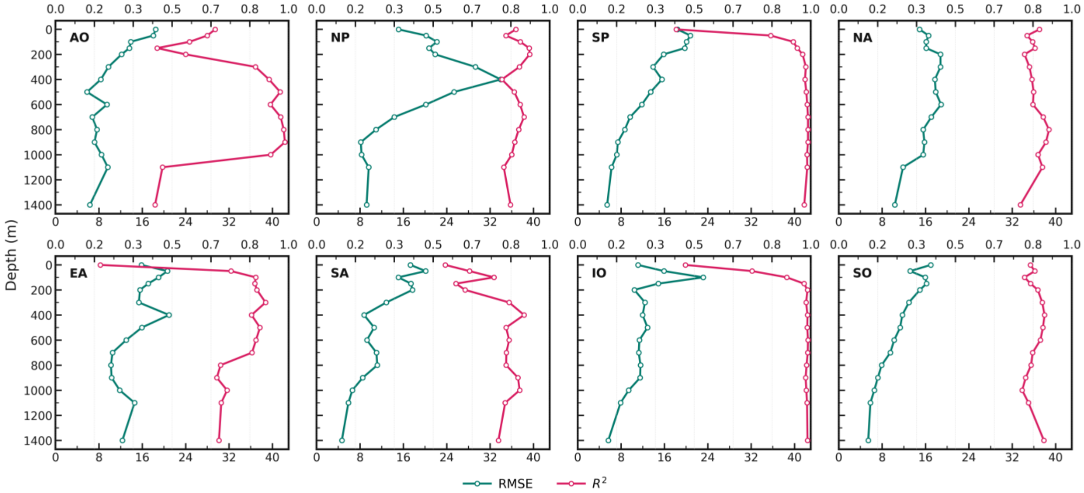


Figure 6: Depth profiles of model performance across provinces.

4.3 Coastal uncertainty

To assess the influence of nearshore regions on basin-wide reconstruction accuracy, we performed a depth-threshold sensitivity analysis. This involved progressively excluding coastal and marginal seas by increasing the bathymetric cutoff from 0 to 400 m, then evaluating model performance over the remaining open-ocean basin (Fig. 7). The results indicate that nearshore shallow waters are a major source of uncertainty. Increasing the threshold from 0 to 200 m yields a marked reduction in domain-mean error (**7.5%** relative to no exclusion; Fig. 7d–e). Further increasing the exclusion threshold to 250–300 m yields only marginal improvements, while a 400 m cutoff increases error—likely due to the loss of representative slope waters and a substantial decline in training sample size. In practice, the 200 m isobath offers a pragmatic coastal-exclusion standard when nearshore processes such as terrestrial inputs, eutrophication, fine-scale stratification, and benthic-pelagic coupling are not fully captured by our predictor set.

These findings are consistent with the physical and ecological characteristics of nearshore environments. Nearshore oxygen variability is driven by high-frequency, localized processes including phytoplankton bloom pulses, riverine discharge, tidal mixing, and anthropogenic effects, leading to strong diel and seasonal fluctuations that are poorly represented in existing open-ocean focused predictors (Gilbert et al., 2010; Regier et al., 2023; Giomi et al., 2023; Liu et al., 2024). Such nonlinear, rapidly varying processes produce error statistics that differ systematically from the open ocean. As a result, the model has limited ability to generalize features across regions (Valera et al., 2020). Statistically, shallow shelf systems are not representative of the large-scale oceanic DO background; their inclusion during model training and evaluation artificially elevates domain-integrated error and obscures broad-scale oxygen patterns. Consequently, GEOXYGEN provides DO fields only for open-ocean grid cells deeper than 200 m, with shallower coastal grid cells flagged as missing values.

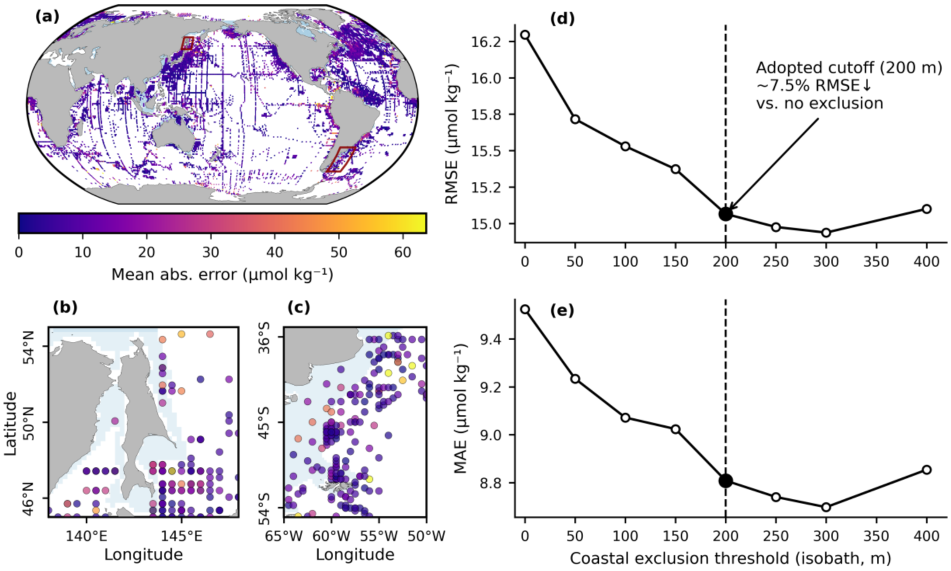


Figure 7: Mean absolute error of the test set at the ocean surface (panel a). Blue shading indicates shallow coastal and marginal sea regions shallower than 200 m, and red boxes mark two representative high-error coastal areas. Panels (b) and (c) show zoomed views of the Sea of Okhotsk shelf and the Patagonian Shelf, respectively. Panels (d) and (e) show RMSE and MAE as functions of the coastal exclusion threshold, quantifying how model performance changes with the chosen depth cutoff.

4.4 Long-term dataset

The trained models, forced by external physical analysis fields, generate spatially and temporally consistent monthly dissolved oxygen (DO) reconstructions. We use the CORA ocean analysis from CMEMS to provide temperature and salinity, together with the SSEVs, as large-scale physical and biogeochemical predictors on the target grid. Using these objectively analyzed gridded fields instead of interpolating from sparse in situ profiles avoids an extra interpolation step and the associated secondary uncertainties. At each standard depth, the corresponding feature fields are evaluated by the matching regional model to generate monthly DO predictions, yielding 0.5° × 0.5° fields from the surface to 5500 m that are aligned with the CMEMS grid and calendar, but with coastal waters shallower than the 200 m isobath masked out. Because CORA analyses undergo delayed-mode objective mapping and rigorous quality control, they provide long-term consistency, global continuity and full traceability. Compared to methods that interpolate sparse in situ temperature and salinity observations, this approach reduces uncertainties associated with secondary statistical interpolation.

To mitigate potential discontinuities at the boundaries between biogeochemical provinces—often termed a "step-effect"—we implement boundary fusion within the transition zones (Wagstaff and Bean, 2022). Let $\mathbf{x}$ be the prediction location, $\hat{y}_i(\mathbf{x})$ the prediction of the i regional model at $\mathbf{x}$, and d_i the minimum great-circle distance from $\mathbf{x}$ to the boundary of that region. Let the smoothing bandwidth be defined as $S = 300$ km. The fused estimate is then defined as:

$$\hat{y}'(\boldsymbol{x}) = \frac{\sum_i w(d_i|S)\hat{y}_i(x)}{\sum_i w(d_i|S)}, \qquad w(d_i \mid S) = \begin{cases} \left(\frac{(S-d_i)}{S}\right)^2, & d_i \le S \\ 0, & d_i > S \end{cases} \tag{6}$$

Thus, far from a boundary ($d > S$), only the local provincial model contributes; within S of a boundary, predictions from adjacent provinces are blended smoothly in proportion to their distance from the edge. This preserves continuity across boundaries without excessive smoothing.

Finally, we obtained GEOXYGEN, a global DO dataset that provides monthly fields on a 0.5° × 0.5° latitude–longitude grid at 187 standard depth levels from the surface to 5500 m, spanning 1960–2024. GEOXYGEN is distributed as CF-compliant NetCDF files, with one file per month (GEOXYGEN_DO_YYYYMM_0p5deg_v1.nc) containing a four-dimensional DO variable (time × depth × lat × lon) and the corresponding coordinate variables (time, depth, latitude, longitude). Time is encoded as days since 1950-01-01 00:00:00 UTC using a Gregorian calendar and represents monthly means, and missing values are flagged with a large sentinel value. An accompanying biogeochemical province mask, which also serves as a valid-ocean mask (open-ocean grid cells deeper than 200 m), is provided in a separate NetCDF file. All data files are openly available at https://doi.org/10.5281/zenodo.17615657 (Wang et al., 2025).

4.5 Comparison with other products

We compare GEOXYGEN against four existing dissolved oxygen (DO) products (Table 3) with the independent test dataset. Performance is evaluated using spatiotemporally co-located samples from the five withheld years, analyzed separately for the early (1960–1980) and recent (2000–2020) periods (Table 4). Only grid-month-depth points with concurrent observations across all products are included.

Table 3. Summary of our product and other DO products

Product	Time coverage	Vertical levels	Temporal resolution	Horizontal resolution
Our product (GEOXYGEN)	1960-01 – 2024-06	0–5500 m (187 levels)	Monthly	0.5° × 0.5°
ML_4O_2	1965-01 – 2020-12	6–1000 m (20 levels)	Monthly	1° × 1°
GOBAI-O_2	2004-01 – 2023-12	2.5–1975 m (58 levels)	Monthly	1° × 1°
G4D-DOC	2005-01 – 2022-12	10–1995 m (26 levels)	Monthly	1° × 1°
IAP Oxygen (Gouretski et al., 2024a)	1960-01 – 2022-12	0–6000 m (119 levels)	Monthly	1° × 1°

Table 4. Accuracy by depth for each product relative to observations

Product	Depth	1960-1980			2000-2020		
		RMSE	Bias	R^2	RMSE	Bias	R^2
Our product (GEOXYGEN)	10	**15**	**0.1**	**0.92**	**8.2**	**-0.5**	**0.96**
IAP Oxygen		17.1	1.1	0.9	9.5	1.3	0.94
ML$_4$O$_2$		17.3	0.7	0.89	10.5	1.7	0.93
GOBAI-O$_2$		NA	NA	NA	10	1.1	0.93
G4D-DOC		NA	NA	NA	7.4	0.2	0.96
Our product (GEOXYGEN)	50	**20**	**-0.3**	**0.9**	**13.5**	**-0.1**	**0.94**
IAP Oxygen		21.1	1.4	0.89	17.2	2.5	0.9
ML$_4$O$_2$		21.3	0	0.89	16.1	0.6	0.91
GOBAI-O$_2$		NA	NA	NA	16.9	2.1	0.9
G4D-DOC		NA	NA	NA	13.6	0	0.93
Our product (GEOXYGEN)	100	**20.3**	**-0.2**	**0.93**	**15.3**	**0**	**0.95**
IAP Oxygen		20.6	1.5	0.93	19.4	1.8	0.93
ML$_4$O$_2$		20.8	1	0.93	18.6	-0.1	0.93
GOBAI-O$_2$		NA	NA	NA	18	0.7	0.94
G4D-DOC		NA	NA	NA	18	0.7	0.94
Our product (GEOXYGEN)	200	**19.3**	**-0.1**	**0.93**	**12.7**	**0**	**0.97**
IAP Oxygen		20.2	-0.3	0.92	16.9	1.4	0.95
ML$_4$O$_2$		19.8	-0.1	0.93	20	-1.7	0.94
GOBAI-O$_2$		NA	NA	NA	21	-1.6	0.93
G4D-DOC		NA	NA	NA	16.4	1	0.96
Our product (GEOXYGEN)	500	**21.2**	**-3.6**	**0.93**	**14.2**	**-0.3**	**0.97**
IAP Oxygen		18.6	-2.9	0.94	23.5	-2.2	0.92
ML$_4$O$_2$		19.6	-6.1	0.94	26	-4	0.91
GOBAI-O$_2$		NA	NA	NA	28.4	-5.3	0.9
G4D-DOC		NA	NA	NA	16.1	0.1	0.96
Our product (GEOXYGEN)	1000	**13.1**	**-0.9**	**0.97**	**6.5**	**0.2**	**0.99**
IAP Oxygen		14.5	1.6	0.96	8.2	3	0.99
ML$_4$O$_2$		14.6	-0.3	0.96	8.1	0.5	0.99
GOBAI-O$_2$		NA	NA	NA	7.1	1.4	0.99
G4D-DOC		NA	NA	NA	7.7	-0.5	0.99

The comparison highlights three main features. First, during the data-sparse early period (1960–1980), GEOXYGEN matches or exceeds the performance of IAP and ML4O$_2$ at 10–200 m depth, exhibiting lower RMSE and near-zero mean bias, and achieves the best performance at 1000 m. Although errors are marginally higher at 500 m, the high R^2 (>0.9) confirms that large-scale oxygen structure is well reproduced even under limited observational constraints. Second, in the recent, data-rich period (2000–2020), GEOXYGEN systematically reduces RMSE and bias at 50–200 m relative to all reference products, while maintaining $R^2 \geq 0.94$. In this depth range, several existing products exhibit persistent positive or negative biases, whereas GEOXYGEN remains close to unbiased, consistent with the benefits of regionalized modeling, inverse-density weighting, and the inclusion of physically informed predictors. Third, at 1000 m, all products perform well, reflecting the stability of deep water masses, yet GEOXYGEN ranks among the top performers, with RMSE values that are slightly lower or comparable to the best alternatives. Overall, the out-of-time, co-located comparison indicates that GEOXYGEN provides reconstruction skill that exceeds existing products, particularly at mid-depths where deoxygenation signals and model spread are most pronounced. These outcomes align with our methodological strategy: the use of a long observational archive (1960–2024) enables training across diverse climate states; CORA temperature–salinity fields, process-relevant SSEVs, regionally adapted models, adaptive feature selection, inverse-density weighting, and boundary fusion collectively reduce bias, most notably between 50–500 m. GEOXYGEN agrees with existing products at the surface and in the deep ocean and achieves systematically lower or comparable RMSE and bias at key mid-depths. Taken together, it provides a high-resolution, broad-coverage, and accurate DO dataset that supports more comprehensive analyses of deoxygenation trends and their driving mechanisms.

4.6 Comparison with WOA23

To evaluate the large-scale credibility and multi-year stability of our reconstructed dissolved oxygen (DO) fields, we compare our product's annual-mean DO climatology with the World Ocean Atlas 2023 (WOA23; Garcia et al., 2024) and examine vertical structure using representative profiles from three major basins (Fig. 8).

In the upper ocean (0–300 m, depth-averaged), both products capture consistent basin-scale spatial patterns. The subtropical gyres exhibit relatively high DO concentrations, while the equatorial region and eastern boundary upwelling systems form distinct oxygen-deficient belts. The spatial extent and location of these structures are in close agreement between the two climatologies. GEOXYGEN reproduces these banded structures continuously, with cross-frontal gradients in transition zones closely matching those in WOA23 (Fig. 8, top row).

In the vertical, our product accurately depicts the largest hypoxic zone in climatology, which is mainly located in the intermediate depth, ranging from 300 to 1,300 m. Meridional sections across the Pacific, Atlantic, and Indian Oceans show broad consistency between GEOXYGEN and WOA23 in the alignment and closure of key oxygen isopleths, as well as in the positioning of high-gradient transition layers. This structural agreement underscores the physical consistency of our product across depth and region.

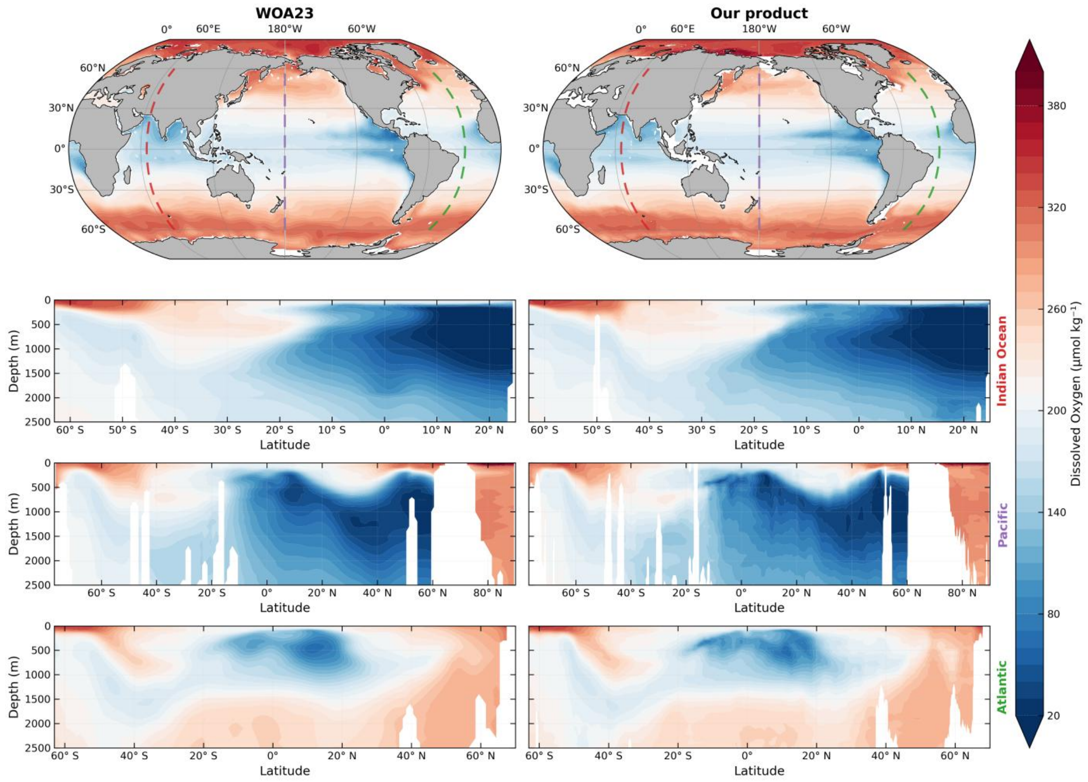


Figure 8: Climatological comparison between WOA23 and our product (GEOXYGEN). Row 1 shows the global-mean DO distribution averaged over 0–300 m. Colored dashed lines mark the locations of three sections: 65°E (red), 180° (purple), and 30°W (green). Rows 2–4 show cross-sectional DO along these three sections.

Most global DO products are distributed on a 1° grid; in contrast, GEOXYGEN adopts a 0.5° horizontal grid, chosen to match the resolution of the external analysis fields used as predictors. Within each ecogeographic region, the model learns relationships between DO and objectively analyzed physical and biogeochemical fields and then applies these relationships across the grid. Consequently, the reconstruction is not merely a reflection of where DO observations are dense; it also resolves regional gradients and delineates the geometry of OMZs more clearly. Although performance is ultimately bounded by the quality of the underlying observations and external analysis products, comparisons with 1° products show that the 0.5° configuration yields consistently lower—or comparable—errors at key depths and can accurately resolve regions of strong DO gradients. We therefore regard 0.5° as an appropriate operating resolution that exploits available information without over-interpreting sparsely sampled areas.

5 Conclusion

We introduce GEOXYGEN, a monthly, four-dimensional global ocean dissolved oxygen (DO) product spanning 1960–2024 at 0.5° × 0.5° resolution, developed to address long-standing challenges of data sparsity and spatiotemporal heterogeneity in historical DO observations. Evaluated on an independent out-of-time test composed of withheld years, the reconstruction demonstrates consistently high skill across all depth layers (typically $R^2 > 0.92$), confirming its robustness under conservative validation protocols.

In comparison with other existing global DO products, GEOXYGEN reproduces observed spatial variability and attains superior overall performance at mid-depths (50–200 m) and in the deep ocean while offering a higher spatial resolution (0.5° vs. 1°). Our approach is distinguished by three principal innovations:

- **Heterogeneity-aware hierarchical modeling.** By combining vertical stratification with biogeographic provincialization, we train CatBoost regressors within each region–depth unit. This design directly addresses the limitations of single global ML models, which struggle to represent the spatially varying physical–biogeochemical controls on DO. In combination with adaptive feature selection, inverse-density weighting, year-grouped cross-validation, and cross-boundary fusion, the framework enhances robustness in undersampled regions, minimizes temporal leakage and boundary artefacts, and yields parsimonious, interpretable submodels. The same hierarchical strategy can be readily applied to other biogeochemical tracers or observing systems that exhibit strong regional and vertical heterogeneity.
- **Adaptive multi-source feature selection.** Starting from a rich set of physical, biogeochemical, and spatiotemporal predictors, we employ a two-stage feature-selection procedure within each region–depth unit to retain only variables that add independent skill. This adaptive, region- and depth-aware integration of multi-source environmental predictors strengthens the representation of upper-ocean processes and water-mass transitions while suppressing noise and redundancy, providing physically interpretable feature sets.
- **A physically consistent, long-record, high-resolution product.** GEOXYGEN delivers global monthly DO fields from 1960 to 2024 on a consistent 0.5° × 0.5° horizontal grid, spanning depths from the surface to 5500 m. Its high skill and climatological consistency support robust estimates of deoxygenation trends and decadal variability, providing a stringent benchmark for assessing and constraining Earth system models.

Nonetheless, GEOXYGEN is subject to certain limitations inherent in observational coverage and methodological assumptions. It is also noted that uncertainty is elevated in early decades, ice-covered high latitudes, data-sparse deep basins, and unresolved nearshore regions. Extremely low DO values in severely hypoxic environments may be conservatively biased upward, and the product's sensitivity to uncertainties in external physical drivers merits further assessment. Future efforts will focus on cross-platform calibration, dynamically refined regionalization, and inclusion of additional subsurface constraints. Despite these limitations, it provides a long-term, internally consistent, and process-informed foundation for detecting, attributing, and interpreting global ocean deoxygenation. The GEOXYGEN, together with basic usage examples, is available at Zenodo (Wang et al., 2025, https://doi.org/10.5281/zenodo.17615657). Future updates will be versioned at the same DOI.

Code and data availability

The dissolved oxygen (DO) dataset generated in this study, together with the biogeochemical province mask used for regionalization and coastal exclusion, **is** openly available at https://doi.org/10.5281/zenodo.17615657 (Wang et al., 2025), where details of the data files and metadata are documented. The codes used to train models and generate the data product in this paper are openly available at https://github.com/layne1202/GEOXYGEN-code.

Author contribution

Z. W. and W. F. conceived the project. Z. W. collected and processed the data with contributions from C. X. Z. W. carried out the study and generated the GEOXYGEN product with contributions from W. F. Z. W. and W. F. wrote the manuscript with contributions from C. X. and G. W. All authors discussed the results and commented on the manuscript.

Acknowledgments

This research is sponsored by Natural Science Foundation of Shanghai under the grant number 24ZR1404500. The Argo data used here were collected and made freely available by the International Argo Program and by the national programs that contribute to it (https://doi.org/10.17882/42182). The OceanSITES data were collected and made freely available by the international OceanSITES project and the national programs that contribute to it.

Competing interests

The contact author has declared that none of the authors has any competing interests.

References

Bettencourt, J. H., Lopez, C., Hernandez-Garcia, E., Montes, I., Sudre, J., Dewitte, B., Paulmier, A., and Garcon, V.: Boundaries of the Peruvian oxygen minimum zone shaped by coherent mesoscale dynamics, Nat. Geosci., 8, 937–940, https://doi.org/10.1038/ngeo2570, 2015.

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., and Seferian, R.: Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models, Biogeosciences, 10, 6225–6245, https://doi.org/10.5194/bg-10-6225-2013, 2013.

https://doi.org/10.5194/essd-2025-699
Preprint. Discussion started: 1 December 2025

Breitburg, D., Levin, L. A., Oschlies, A., Gregoire, M., Chavez, F. P., Conley, D. J., Garcon, V., Gilbert, D., Gutierrez, D., Isensee, K., Jacinto, G. S., Limburg, K. E., Montes, I., Naqvi, S. W. A., Pitcher, G. C., Rabalais, N. N., Roman, M. R., Rose, K. A., Seibel, B. A., Telszewski, M., Yasuhara, M., and Zhang, J.: Declining oxygen in the global ocean and coastal waters, Science, 359, eaam7240, https://doi.org/10.1126/science.aam7240, 2018.

Cao, R., Wang, S., Bao, S., Li, X., Tan, J., and Shao, C.: SE-LeNet: A data reconstruction method for dissolved oxygen in tropical Pacific with deep learning, in: Proc. 2024 IEEE Int. Conf. Parallel Distrib. Process. Appl. (ISPA), https://doi.org/10.1109/ISPA63168.2024.00031, 2024.

Chau, T.-T.-T., Gehlen, M., Metzl, N., and Chevallier, F.: CMEMS-LSCE: a global, 0.25∘, monthly reconstruction of the surface ocean carbonate system, Earth Syst. Sci. Data, 16, 121–160, https://doi.org/10.5194/essd-16-121-2024, 2024.

Chau, T. T. T., Gehlen, M., and Chevallier, F.: A seamless ensemble-based reconstruction of surface ocean pCO_2 and air–sea CO_2 fluxes over the global coastal and open oceans, Biogeosciences, 19, 1087–1109, https://doi.org/10.5194/bg-19-1087-2022, 2022.

Chen, Z., Siedlecki, S., Long, M., Petrik, C. M., Stock, C. A., and Deutsch, C. A.: Skillful multiyear prediction of marine habitat shifts jointly constrained by ocean temperature and dissolved oxygen, Nat. Commun., 15, 900, https://doi.org/10.1038/s41467-024-45016-5, 2024.

Cocco, V., Joos, F., Steinacher, M., Frölicher, T. L., Bopp, L., Dunne, J., Gehlen, M., Heinze, C., Orr, J., and Oschlies, A.: Oxygen and indicators of stress for marine life in multi-model global warming projections, Biogeosciences, 10, 1849–1868, https://doi.org/10.5194/bg-10-1849-2013, 2013.

Fay, A. R. and McKinley, G. A.: Global open-ocean biomes: mean and temporal variability, Earth Syst. Sci. Data, 6, 273–284, https://doi.org/10.5194/essd-6-273-2014, 2014.

Franco, A. C., Hernández-Ayón, J. M., Beier, E., Garçon, V., Maske, H., Paulmier, A., Färber-Lorda, J., Castro, R., and Sosa-Ávalos, R.: Air-sea CO_2 fluxes above the stratified oxygen minimum zone in the coastal region off Mexico, J. Geophys. Res.-Oceans, 119, 2923–2937, https://doi.org/10.1002/2013JC009337, 2014.

Garabaghi, F. H., Benzer, S., and Benzer, R.: Modeling dissolved oxygen concentration using machine learning techniques with dimensionality reduction approach, Environ. Monit. Assess., 195, 879, https://doi.org/10.1007/s10661-023-11492-3, 2023.

Garcia, H., Cruzado, A., Gordon, L., and Escanez, J.: Decadal-scale chemical variability in the subtropical North Atlantic deduced from nutrient and oxygen data, J. Geophys. Res.-Oceans, 103, 2817–2830, https://doi.org/10.1029/97JC03037, 1998.

Garcia, H. E., Wang, Z., Bouchard, C., Cross, S. L., Paver, C. R., Reagan, J. R., Boyer, T. P., Locarnini, R. A., Mishonov, A. V., Baranova, O., Seidov, D., and Dukhovskoy, D.: World Ocean Atlas 2023, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation, A. Mishonov (Ed.), NOAA Atlas NESDIS 91, 109 pp., https://doi.org/10.25923/rb67-ns53, 2024.

Gilbert, D., Rabalais, N. N., Díaz, R. J., and Zhang, J.: Evidence for greater oxygen decline rates in the coastal ocean than in the open ocean, Biogeosciences, 7, 2283–2296, https://doi.org/10.5194/bg-7-2283-2010, 2010.

Giomi, F., Barausse, A., Steckbauer, A., Daffonchio, D., Duarte, C. M., and Fusi, M.: Oxygen dynamics in marine productive ecosystems at ecologically relevant scales, Nat. Geosci., 16, 560–566, https://doi.org/10.1038/s41561-023-01217-z, 2023.

Gong, H., Li, C., and Zhou, Y.: Emerging global ocean deoxygenation across the 21st century, Geophys. Res. Lett., 48, e2021GL095370, https://doi.org/10.1029/2021GL095370, 2021.

Gouretski, V., Cheng, L., Du, J., Xing, X., and Chai, F.: A quality-controlled and bias-adjusted global ocean oxygen profile dataset [data set], https://doi.org/10.12157/IOCAS, 2024a.

Gouretski, V., Cheng, L., Du, J., Xing, X., Chai, F., and Tan, Z.: A consistent ocean oxygen profile dataset with new quality control and bias assessment, Earth Syst. Sci. Data, 16, 5503–5530, https://doi.org/10.5194/essd-16-5503-2024, 2024b.

Gregoire, M., Garcon, V., Garcia, H., Breitburg, D., Isensee, K., Oschlies, A., Telszewski, M., Barth, A., Bittig, H. C., Carstensen, J., Carval, T., Chai, F., Chavez, F., Conley, D., Coppola, L., Crowe, S., Currie, K., Dai, M., Deflandre, B., Dewitte, B., Diaz, R., Garcia-Robledo, E., Gilbert, D., Giorgetti, A., Glud, R., Gutierrez, D., Hosoda, S., Ishii, M., Jacinto, G., Langdon, C., Lauvset, S. K., Levin, L. A., Limburg, K. E., Mehrtens, H., Montes, I., Naqvi, W., Paulmier, A., Pfeil, B., Pitcher, G., Pouliquen, S., Rabalais, N., Rabouille, C., Recape, V., Roman, M., Rose, K., Rudnick, D., Rummer, J., Schmechtig, C., Schmidtko, S., Seibel, B., Slomp, C., Sumalia, U. R., Tanhua, T., Thierry, V., Uchida, H., Wanninkhof, R., and Yasuhara, M.: A global ocean oxygen database and atlas for assessing and predicting deoxygenation and ocean health in the open and coastal ocean, Front. Mar. Sci., 8, https://doi.org/10.3389/fmars.2021.724913, 2021.

Grégoire, M., Oschlies, A., Canfield, D., Castro, C., Ciglenecki, I., Croot, P., Salin, K., Schneider, B., Serret, P., and Slomp, C.: Ocean Oxygen: the role of the Ocean in the oxygen we breathe and the threat of deoxygenation, European Marine Board, Ostend, Belgium, https://doi.org/10.5281/zenodo.7941157, 2023.

Guinehut, S., Dhomps, A.-L., Larnicol, G., and Le Traon, P.-Y.: High resolution 3-D temperature and salinity fields derived from in situ and satellite observations, Ocean Sci., 8, 845–857, https://doi.org/10.5194/os-8-845-2012, 2012.

Hauser, D., Tourain, C., Hermozo, L., Alraddawi, D., Aouf, L., Chapron, B., Dalphinet, A., Delaye, L., Dalila, M., and Dormy, E.: New observations from the SWIM radar on-board CFOSAT: Instrument validation and ocean wave measurement assessment, IEEE Trans. Geosci. Remote Sens., 59, 5–26, https://doi.org/10.1109/TGRS.2020.2994372, 2020.

Hollitzer, H. A. L., Patara, L., Terhaar, J., and Oschlies, A.: Competing effects of wind and buoyancy forcing on ocean oxygen trends in recent decades, Nat. Commun., 15, https://doi.org/10.1038/s41467-024-53557-y, 2024.

Huang, S., Shao, J., Chen, Y., Qi, J., Wu, S., Zhang, F., He, X., and Du, Z.: Reconstruction of dissolved oxygen in the Indian Ocean from 1980 to 2019 based on machine learning techniques, Front. Mar. Sci., 10, 1291232, https://doi.org/10.3389/fmars.2023.1291232, 2023.

Humphries, N. E., Fuller, D. W., Schaefer, K. M., and Sims, D. W.: Highly active fish in low oxygen environments: Vertical movements and behavioural responses of bigeye and yellowfin tunas to oxygen minimum zones in the eastern Pacific Ocean, Mar. Biol., 171, 55, https://doi.org/10.1007/s00227-023-04366-2, 2024.

IOC, SCOR, and IAPSO: The international thermodynamic equation of seawater – 2010: calculation and use of thermodynamic properties, Intergovernmental Oceanographic Commission, Manuals and Guides No. 56, UNESCO, 196 pp., 2010.

Ito, T., Minobe, S., Long, M. C., and Deutsch, C.: Upper ocean O_2 trends: 1958–2015, Geophys. Res. Lett., 44, 4214–4223, https://doi.org/10.1002/2017GL073613, 2017.

Ito, T., Cervania, A., Cross, K., Ainchwar, S., and Delawalla, S.: Mapping dissolved oxygen concentrations by combining shipboard and argo observations using machine learning algorithms, J. Geophys. Res.: Mach. Learn. Comput., 1, e2024JH000272, https://doi.org/10.1029/2024JH000272, 2024a.

Ito, T., Garcia, H. E., Wang, Z. K., Minobe, S., Long, M. C., Cebrian, J., Reagan, J., Boyer, T., Paver, C., Bouchard, C., Takano, Y., Bushinsky, S., Cervania, A., and Deutsch, C. A.: Underestimation of multi-decadal global O_2 loss due to an optimal interpolation method, Biogeosciences, 21, 747–759, https://doi.org/10.5194/bg-21-747-2024, 2024b.

Kim, H., Franco, A. C., and Sumaila, U. R.: A selected review of impacts of ocean deoxygenation on fish and fisheries, Fishes, 8, https://doi.org/10.3390/fishes8060316, 2023.

Kolodziejczyk, N., Prigent-Mazella, A., and Gaillard, F.: ISAS temperature, salinity, dissolved oxygen gridded fields, SEANOE [data set], https://doi.org/10.17882/52367, 2023.

Li, C., Huang, J., Ding, L., Liu, X., Yu, H., and Huang, J.: Increasing escape of oxygen from oceans under climate change, Geophys. Res. Lett., 47, e2019GL086345, https://doi.org/10.1029/2019GL086345, 2020.

Liu, G., Yu, X., Zhang, J., Wang, X., Xu, N., and Ali, S.: Reconstruction of the three-dimensional dissolved oxygen and its spatio-temporal variations in the Mediterranean Sea using machine learning, J. Environ. Sci., 157, 710–728, https://doi.org/10.1016/j.jes.2025.01.010, 2025.

Liu, Q., Liu, C., Meng, Q., Su, B., Ye, H., Chen, B., Li, W., Cao, X., Nie, W., and Ma, N.: Machine learning reveals biological activities as the dominant factor in controlling deoxygenation in the South Yellow Sea, Cont. Shelf Res., 283, 105348, https://doi.org/10.1016/j.csr.2024.105348, 2024.

Lu, B., Zhao, Z., Han, L., Gan, X., Zhou, Y., Zhou, L., Fu, L., Wang, X., Zhou, C., and Zhang, J.: OxyGenerator: reconstructing global ocean deoxygenation over a century with deep learning, in: Proc. 41st Int. Conf. Mach. Learn. (ICML 2024), Vienna, Austria, 2024.

Ma, D., Zhao, F., Zhu, L., Li, X., Wei, J., Chen, X., Hou, L., Li, Y., and Liu, M.: Deep learning reveals hotspots of global oceanic oxygen changes from 2003 to 2020, Int. J. Appl. Earth Obs. Geoinf., 136, 104363, https://doi.org/10.1016/j.jag.2025.104363, 2025.

Mears, C., Lee, T., Ricciardulli, L., Wang, X., and Wentz, F.: Improving the Accuracy of the Cross-Calibrated Multi-Platform (CCMP) Ocean Vector Winds, Remote Sens., 14, 4230, https://doi.org/10.3390/rs14174230, 2022.

Milà, C., Ludwig, M., Pebesma, E., Tonne, C., and Meyer, H.: Random forests with spatial proxies for environmental modelling: opportunities and pitfalls, Geosci. Model Dev., 17, 6007–6033, https://doi.org/10.5194/gmd-17-6007-2024, 2024.

NASA Ocean Biology Processing Group: Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Level-2 Ocean Color Data, version R2018.8, NASA Ocean Biology Distributed Active Archive Center [data set], https://doi.org/10.5067/ORBVIEW-2/SEAWIFS/L2/OC/2018, 2018.

Mishonov, A. V., Boyer, T. P., Baranova, O. K., Bouchard, C. N., Cross, S. L., Garcia, H. E., Locarnini, R. A., Paver, C. R., Reagan, J. R., Wang, Z., Seidov, D., Grodsky, A. I., and Beauchamp, J. G.: World Ocean Database 2023, NOAA Atlas NESDIS, 97, NOAA National Environmental Satellite, Data, and Information Service and NOAA National Centers for Environmental Information, Silver Spring, MD, USA, https://doi.org/10.25923/z885-h264, 2024.

Oschlies, A.: A committed fourfold increase in ocean oxygen loss, Nat. Commun., 12, https://doi.org/10.1038/s41467-021-22584-4, 2021.

Oschlies, A., Brandt, P., Stramma, L., and Schmidtko, S.: Drivers and mechanisms of ocean deoxygenation, Nat. Geosci., 11, 467–473, https://doi.org/10.1038/s41561-018-0152-2, 2018.

Ping, B., Meng, Y., Su, F., Xue, C., and Li, Z.: Retrieval of subsurface dissolved oxygen from surface oceanic parameters based on machine learning, Mar. Environ. Res., 199, 106578, https://doi.org/10.1016/j.marenvres.2024.106578, 2024.

Regier, P. J., Ward, N. D., Myers-Pigg, A. N., Grate, J., Freeman, M. J., and Ghosh, R. N.: Seasonal drivers of dissolved oxygen across a tidal creek – marsh interface revealed by machine learning, Limnol. Oceanogr., 68, 2359–2374, https://doi.org/10.1002/lno.12426, 2023.

Robinson, C.: Microbial respiration, the engine of ocean deoxygenation, Front. Mar. Sci., 5, https://doi.org/10.3389/fmars.2018.00533, 2019.

Salazar, J. J., Garland, L., Ochoa, J., and Pyrcz, M. J.: Fair train-test split in machine learning: Mitigating spatial autocorrelation for improved prediction accuracy, J. Pet. Sci. Eng., 209, 109885, https://doi.org/10.1016/j.petrol.2021.109885, 2022.

Schmidtko, S., Stramma, L., and Visbeck, M.: Decline in global oceanic oxygen content during the past five decades, Nature, 542, 335–339, https://doi.org/10.1038/nature21399, 2017.

Shao, J., Huang, S., Chen, Y., Qi, J., Wang, Y., Wu, S., Liu, R., and Du, Z.: Satellite-based global sea surface oxygen mapping and interpretation with spatiotemporal machine learning, Environ. Sci. Technol., 58, 498–509, https://doi.org/10.1021/acs.est.3c08833, 2024.

Sharp, J. D., Fassbender, A. J., Carter, B. R., Johnson, G. C., Schultz, C., and Dunne, J. P.: GOBAI-O_2: Temporally and spatially resolved fields of ocean interior dissolved oxygen over nearly two decades, Earth Syst. Sci. Data, 15, 4481–4518, https://doi.org/10.5194/essd-15-4481-2023, 2023.

Szekely, T., Gourrion, J., Pouliquen, S., and Reverdin, G.: The CORA 5.2 dataset for global in situ temperature and salinity measurements: data description and validation, Ocean Sci., 15, 1601–1614, https://doi.org/10.5194/os-15-1601-2019, 2019.

Szekely, T., Gourrion, J., Pouliquen, S., Reverdin, G., and Merceur, F.: CORA, Coriolis Ocean Dataset for Reanalysis, SEANOE [data set], https://doi.org/10.17882/46219, 2025.

615 Valera, M., Walter, R. K., Bailey, B. A., and Castillo, J. E.: Machine learning based predictions of dissolved oxygen in a small coastal embayment, J. Mar. Sci. Eng., 8, 1007, https://doi.org/10.3390/jmse8121007, 2020.

Wagstaff, J. and Bean, B.: remap: Regionalized models with spatially smooth predictions, R J., 14, 160–178, https://doi.org/10.32614/RJ-2023-004, 2022.

Wang, Z., Fu, W., Xue, C., and Wang, G.: GEOXYGEN: a global long-term dissolved oxygen dataset (v1.0) [data set], https://doi.org/10.5281/zenodo.17615657, 2025.

Wang, Z., Xue, C., and Ping, B.: A reconstructing model based on time–space–depth partitioning for global ocean dissolved oxygen concentration, Remote Sens., 16, 228, https://doi.org/10.3390/rs16020228, 2024.

Xue, C., Wang, Z., Yue, L., and Niu, C.: A global four-dimensional gridded dataset of ocean dissolved oxygen concentration retrieval from Argo profiles, Geosci. Data J., 11, 775–789, https://doi.org/10.1002/gdj3.251, 2024.

625 Yamaguchi, R., Kouketsu, S., Kosugi, N., and Ishii, M.: Global upper ocean dissolved oxygen budget for constraining the biological carbon pump, Commun. Earth Environ., 5, https://doi.org/10.1038/s43247-024-01886-7, 2024.

Zhou, Y., Gong, H., and Zhou, F.: Responses of Horizontally Expanding Oceanic Oxygen Minimum Zones to Climate Change Based on Observations, Geophys. Res. Lett., 49, https://doi.org/10.1029/2022gl097724, 2022.