



GOBAI-O₂: temporally and spatially resolved fields of ocean interior dissolved oxygen over nearly 2 decades

Jonathan D. Sharp^{1,2}, Andrea J. Fassbender², Brendan R. Carter^{1,2}, Gregory C. Johnson²,
Cristina Schultz^{3,4}, and John P. Dunne³

¹Cooperative Institute for Climate, Ocean, and Ecosystem Studies, University of Washington,
Seattle, WA 98105, USA

²Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration,
Seattle, WA 98115, USA

³Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration,
Princeton, NJ 08540, USA

⁴Department of Geosciences, Princeton University,
Princeton, NJ 08540, USA

Correspondence: Jonathan D. Sharp (jonathan.sharp@noaa.gov)

Received: 1 September 2022 – Discussion started: 20 September 2022

Revised: 7 July 2023 – Accepted: 16 August 2023 – Published: 6 October 2023

Abstract. For about 2 decades, oceanographers have been installing oxygen sensors on Argo profiling floats to be deployed throughout the world ocean, with the stated objective of better constraining trends and variability in the ocean's inventory of oxygen. Until now, measurements from these Argo-float-mounted oxygen sensors have been mainly used for localized process studies on air–sea oxygen exchange, upper-ocean primary production, biological pump efficiency, and oxygen minimum zone dynamics. Here, we present a new four-dimensional gridded product of ocean interior oxygen, derived via machine learning algorithms trained on dissolved oxygen observations from Argo-float-mounted sensors and discrete measurements from ship-based surveys and applied to temperature and salinity fields constructed from the global Argo array. The data product is called GOBAI-O₂, which stands for Gridded Ocean Biogeochemistry from Artificial Intelligence – Oxygen (Sharp et al., 2022; <https://doi.org/10.25921/z72m-yz67>); it covers 86 % of the global ocean area on a 1° × 1° (latitude × longitude) grid, spans the years 2004–2022 with a monthly resolution, and extends from the ocean surface to a depth of 2 km on 58 levels. Two types of machine learning algorithms – random forest regressions and feed-forward neural networks – are used in the development of GOBAI-O₂, and the performance of those algorithms is assessed using real observations and simulated observations from Earth system model output. Machine learning represents a relatively new method for gap filling ocean interior biogeochemical observations and should be explored along with statistical and interpolation-based techniques. GOBAI-O₂ is evaluated through comparisons to the oxygen climatology from the World Ocean Atlas, the mapped oxygen product from the Global Ocean Data Analysis Project and to direct observations from large-scale hydrographic research cruises. Finally, potential uses for GOBAI-O₂ are demonstrated by presenting average oxygen fields on isobaric and isopycnal surfaces, average oxygen fields across vertical–meridional sections, climatological seasonal cycles of oxygen averaged over different pressure layers, and globally integrated time series of oxygen. GOBAI-O₂ indicates a declining trend in the oxygen inventory in the upper 2 km of the global ocean of 0.79 ± 0.04 % per decade between 2004 and 2022.

1 Introduction

The inventory of dissolved oxygen in the global ocean has been declining over recent decades and is projected to continue declining through the current century (Keeling et al., 2010; Breitburg et al., 2018; Bindoff et al., 2019; Stramma and Schmidtko, 2019; Limburg et al., 2020), leading to detrimental consequences for aerobic marine organisms (Pörtner and Farrell, 2008; Sampaio et al., 2021) and changes in biogeochemical cycles, potentially triggering important climatological feedbacks (Gruber, 2004; Berman-Frank et al., 2008). Historical deoxygenation has been inferred from analyses of globally distributed observations (Helm et al., 2011; Schmidtko et al., 2017; Ito et al., 2017) and has been reproduced in Earth system model (ESM) reconstructions (Bopp et al., 2013; Frölicher et al., 2009; Kwiatkowski et al., 2020). Global observational studies have generally indicated a greater degree of deoxygenation than model studies over recent decades, indicating that ESMs may misrepresent the sensitivities of the physical and biological processes leading to deoxygenation, which has implications for the reliability of future projections (Oschlies et al., 2017, 2018; Stramma and Schmidtko, 2021). Model studies, however, are based on gridded output that is continuously resolved in space and time, whereas observational studies rely on the interpolation of measurements from discrete bottle samples and/or profiling sensors. These observational datasets have significant spatiotemporal gaps and may not robustly represent global deoxygenation trends.

Discrete measurements of the dissolved oxygen concentration ($[O_2]$) are typically made using Winkler titrations (Winkler, 1888; Carpenter, 1965; Langdon, 2010), which are also used to calibrate measurements from electrode (or more recently sometimes optical) dissolved oxygen sensors mounted on conductivity–temperature–depth (CTD) profilers. Globally distributed $[O_2]$ observations from discrete measurements and CTD profilers have been provided by hydrographic programs like the World Ocean Circulation Experiment (WOCE); the Climate and Ocean Variability, Predictability and Change (CLIVAR) program; and the Global Ocean Ship-Based Hydrographic Investigations Program (GO-SHIP). Data from these programs are publicly available and are conveniently compiled into databases such as the World Ocean Database (WOD; Boyer et al., 2018) and the Global Ocean Data Analysis Project (GLODAP; Lauvset et al., 2022b). Although unprecedented spatial coverage is provided by global hydrographic programs, the decadal-scale temporal resolution of WOCE, CLIVAR, and GO-SHIP data precludes robust analyses of year-to-year and/or seasonal variability in $[O_2]$.

Since the mid-2000s, approximately 1800 Argo floats equipped with oxygen sensors have been deployed. Argo floats profile the upper ~ 2000 m of the water column every ~ 10 d. Many oxygen-sensor-equipped Argo floats have been deployed as part of regional arrays such as the South-

ern Ocean Carbon and Climate Observations and Modeling (SOCCOM) project and the North Atlantic Aerosols and Marine Ecosystems Study (NAAMES). More recently, the push for a global biogeochemical Argo array has spurred the deployment of oxygen-sensor-equipped Argo floats into more sparsely sampled ocean regions (Johnson and Claustre, 2016; Claustre et al., 2020). As more floats have been deployed, improvements have been made to sensor calibration, data adjustments, and quality control. Notably, pre-deployment drift corrections (D'Asaro and McNeil, 2013; Johnson et al., 2015; Bittig and Körtzinger, 2015; Bushinsky et al., 2016; Drucker and Riser, 2016; Nicholson and Feen, 2017), climatology-based calibrations (Takeshita et al., 2013), calibrations via in-air oxygen measurements (Körtzinger et al., 2005; Fiedler et al., 2013; Bittig and Körtzinger, 2015; Johnson et al., 2015; Bushinsky et al., 2016), post-deployment drift corrections (Johnson et al., 2017; Bittig et al., 2018a), and established procedures for delayed-mode quality control (Maurer et al., 2021) have substantially reduced the uncertainty and increased the reproducibility of optode-based $[O_2]$ measurements on Argo floats.

From the time it began, the Argo-Oxygen program (now oxygen is a measured variable under the Biogeochemical Argo program) intended to document ocean deoxygenation, predict and assess anoxic and hypoxic events, and determine seasonal to interannual changes in export production (Gruber et al., 2010). To date, these goals have been achieved primarily on a regional scale. For example, $[O_2]$ measurements from biogeochemical Argo floats have been used to examine ventilation and air–sea exchange of oxygen in the Southern Ocean (Bushinsky et al., 2017) and during deep water formation in the subpolar North Atlantic (Körtzinger et al., 2004; Piron et al., 2016, 2017; Wolf et al., 2018); denitrification and the spatial extent of the oxygen minimum zone (OMZ) in the Bay of Bengal (Sarma and Udaya Bhaskar, 2018; Johnson et al., 2019; Udaya Bhaskar et al., 2021); and carbon production and export in the Pacific Ocean (Bushinsky and Emerson, 2015, 2018; Yang et al., 2017), Southern Ocean (Stukel and Ducklow, 2017; Arteaga et al., 2019), and North Atlantic Ocean (Alkire et al., 2012; Estapa et al., 2019). Recently, in an early global-scale analysis of $[O_2]$ from the Argo array, Johnson and Bif (2021) used the diel cycle of oxygen measured by the ocean-wide array of biogeochemical Argo floats to constrain net primary production in the surface ocean.

With the work presented here, we seek to capitalize on the collective efforts of global hydrographic programs, Biogeochemical Argo (BGC Argo), and Core Argo (Johnson et al., 2022) to create a novel data product: a four-dimensional monthly record of dissolved oxygen in the global ocean. We combine autonomous observations of $[O_2]$ from BGC Argo floats with discrete observations of $[O_2]$ from hydrographic cruises in the GLODAP database to create a dataset with an extensive spatial and temporal resolution. With this dataset, we train machine learning algorithms on ocean interior predictor variables co-located with $[O_2]$ observations; evaluate

those algorithms using real and simulated data; and apply the algorithms to gridded ocean interior predictor variables mapped from Core Argo to produce a gridded [O₂] data product at a monthly resolution from 2004 to 2022, on 58 pressure levels in the upper 2 km of the ocean, and on a near-global 1° × 1° (latitude × longitude) grid.

In this paper, we present the four-dimensional gridded [O₂] product, which we call GOBAI-O₂: Gridded Ocean Biogeochemistry from Artificial Intelligence – Oxygen (Sharp et al., 2022; <https://doi.org/10.25921/z72m-yz67>). Artificial intelligence (AI) is a broad term describing computerized systems that are able to recognize patterns, make decisions, and otherwise perform tasks previously reserved only for humans; GOBAI-O₂ is built using machine learning (ML), which is a subfield of AI that is focused on training, understanding, and applying algorithms that leverage data to artificially learn and reproduce patterns. We introduce GOBAI-O₂ by analyzing spatial patterns, seasonal cycles, and decadal variability. We also describe the process for creating GOBAI-O₂, show the results of evaluation exercises, assess uncertainty in the gridded [O₂] fields, and compare the data product to other gridded datasets and discrete measurements. GOBAI-O₂ represents the first step in leveraging the emerging global array of BGC Argo floats to produce spatially resolved, time-varying snapshots of global ocean biogeochemical distributions in near-real time. It also emphasizes regions that are good candidates for new observational assets; particularly where gaps in observational coverage coincide with high background variability in [O₂]. Critically, GOBAI-O₂ can be used to address the goals of the Argo-Oxygen program set by Gruber et al. (2010) over a decade ago, providing regional and global insight into ocean deoxygenation and hypoxia on timescales ranging from a few months to multiple years.

2 Methods

2.1 Data sources and processing

Hydrographic cruise data were obtained from the GLODAP version 2022 data product (GLODAPv2.2022; Key et al., 2015; Olsen et al., 2016; Lauvset et al., 2022a, b). GLODAPv2.2022 provides quality-controlled data from throughout the entire water column obtained via discrete analyses of more than 1.4 million water samples collected on 1085 research cruises. Discrete Winkler titration data were chosen rather than CTD oxygen profiles due to the issues with the quality of calibration of a subset of CTD oxygen measurements and the relatively coarse vertical resolution of the final GOBAI-O₂ product, which would not benefit from the high vertical resolution of CTD profiles. Data from GLODAP were chosen rather than data from the WOD or any other database due to the high degree of quality control applied to GLODAP data. Dissolved oxygen is the most represented biogeochemical variable in GLODAPv2.2022, with

more than 1.2 million data points from 991 research cruises. Data from GLODAPv2.2022 were filtered to retain only samples collected after 1 January 2004, from 0 to 2500 dbar (decibars), and with a quality control flag of 1 (meaning the data were manually inspected) and quality flags of 2 (good) for both salinity and [O₂]. Temperature is not assigned either flag and is assumed to be of sufficient quality if it is reported (Lauvset et al., 2022a). This filtering left 450 032 data points from 21 513 unique profiles from 393 total cruises (red points in Fig. 1).

Float data were obtained from synthetic profile (“Sprof”) files (Bittig et al., 2022) stored in the Argo Global Data Assembly Centres (GDACs) via the OneArgo-Mat toolbox (Frenzel et al., 2022) for MATLAB (MathWorks). At the time data were obtained (13 June 2023), the Argo GDACs contained data from about 1800 floats equipped with [O₂] sensors. Float data were filtered to retain only delayed-mode-adjusted data with quality flags of 1 (good), 2 (probably good), or 8 (interpolated/extrapolated) for pressure, temperature, salinity, and [O₂]. This filtering step ensured that float data had been manually reviewed by a data manager and assigned an appropriate quality flag. This filtering left 27 832 192 data points from 138 180 unique profiles from 1022 total floats (blue points in Fig. 1). Of the float profiles, 51.4 % were quality-controlled by comparisons to climatologies (World Ocean Atlas, WOA, or Commonwealth Scientific and Industrial Research Organisation Atlas of Regional Seas, CARS), 30.3 % were quality-controlled using in-air oxygen measurements, 7.0 % were quality-controlled by comparisons to subsurface measurements (WOD, OMZ assumed to have zero oxygen, or CTD profile upon deployment), 5.3 % were quality-controlled using the in situ optode calibration of Drucker and Riser (2016), 3.3 % were quality-controlled via another method, 1.9 % were not categorized, and the remaining 0.9 % were not adjusted.

The discrete temperature, salinity, and [O₂] data obtained from GLODAPv2.2022 and the Argo GDACs are archived online (see Appendix C; <https://doi.org/10.5281/zenodo.7747237>, Sharp, 2023b). To ensure that the trained machine learning algorithms were not biased toward BGC Argo float data, which (in their native format) have a higher vertical resolution than GLODAP data, each profile was interpolated to, at most, 58 standard pressure levels (the same pressure levels on which the final GOBAI-O₂ data product is provided). Interpolated temperature, salinity, and [O₂] data from each source are also archived online (see Appendix C; <https://doi.org/10.5281/zenodo.7747237>, Sharp, 2023b). After interpolation, the total number of GLODAP data points used for algorithm training increased to 1 096 324 and the total number of Argo float data points used for algorithm training decreased to 6 635 749. Co-located, interpolated GLODAP and BGC Argo profiles that fell within the same 1° × 1° monthly, depth-dependent grid cells were compared for internal consistency. The float [O₂] values were adjusted

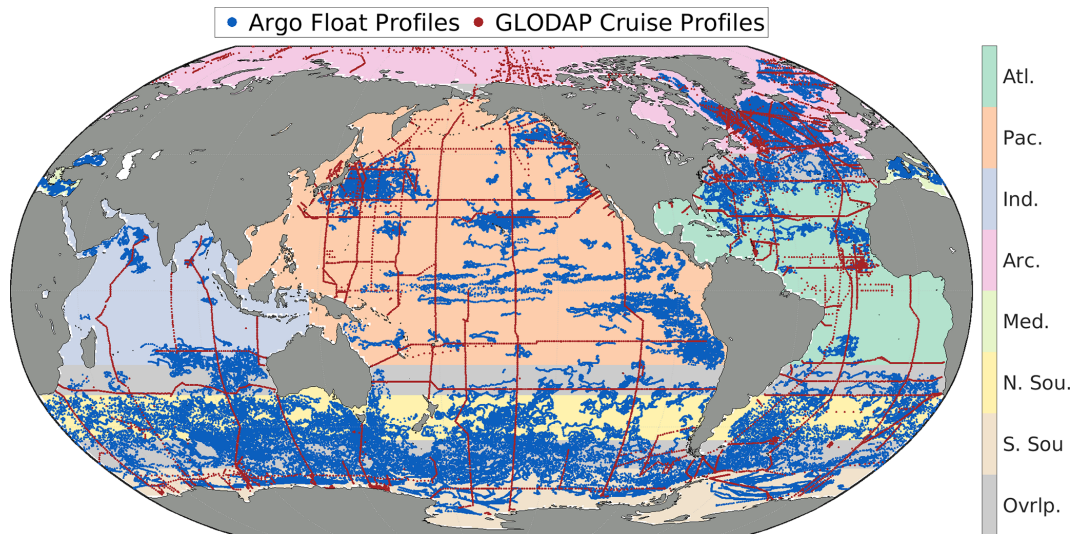


Figure 1. Discrete profile locations from oxygen-sensor-equipped Argo floats (blue) and GLODAPv2.2022 cruises (red) from 1 January 2004 to 13 June 2023. Data from these profiles were binned and used to train ML algorithms to estimate [O₂] in each of seven regions: the Atlantic Ocean (Atl.), Pacific Ocean (Pac.), Indian Ocean (Ind.), Arctic Ocean (Arc.), Mediterranean Sea (Med.), northern section of the Southern Ocean (N. Sou.), and southern section of the Southern Ocean (S. Sou.). Overlapping areas between regions are shown in grey (Ovrlp.), where [O₂] estimates are made by taking distance-weighted averages of outputs from two regional ML algorithms. The regional boundaries are presented in numerical form in Table B1.

according to the procedure in Appendix D to remove the small global discrepancy between co-located ship and float measurements in order to ensure internal consistency between the two datasets.

BGC Argo float and GLODAP cruise data were combined into a single dataset after this bias adjustment, which will be referred to as the “combined dataset” from here on. The combined dataset was grouped into seven overlapping regions (Fig. 1, Table B1). This grouping was intended to account implicitly for similar physical–biogeochemical relationships within large ocean regions and to reduce the computational burden of the machine learning (ML) algorithm fits described below. The regions were initially chosen to imitate the biomes presented by Fay and McKinley (2014), and they were then expanded to relatively large regions bound either by land masses or by overlapping boundaries along constant lines of latitude. The number of profiles made within each $1^\circ \times 1^\circ$ box by either a discrete ship cast or Argo float (Fig. A1) provides a measure of the temporal resolution of the combined dataset in addition to the spatial distribution shown in Fig. 1.

Gridded temperature and salinity data to which the trained algorithms were applied were obtained from the latest version of the Roemmich and Gilson (2009) (RG09) Argo Climatology (https://sio-argo.ucsd.edu/RG_Climatology.html, last access: 12 January 2023). The RG09 climatology is an upper-ocean (0–2000 dbar) gridded temperature and salinity product constructed exclusively from Argo observations. Long-term (2004–2018) mean fields of temperature and salinity are provided on 58 pressure levels, along with

monthly anomaly fields on each of those pressure levels from 2004 to the present day. The most recent major update of the RG09 climatology was made in 2019, and new monthly anomaly fields are provided in near-real time between major updates. Monthly gridded temperature and salinity were calculated from the RG09 long-term mean and monthly anomaly fields (Fig. A2), and they were then used for the creation of the gridded [O₂] product discussed below.

Output from the National Oceanic and Atmospheric Administration (NOAA) Geophysical Fluid Dynamics Laboratory’s Earth System Model Version 4 (GFDL-ESM4; Dunne et al., 2020) was used to assess algorithm performance. Model output was downloaded from the World Climate Research Programme database (<https://esgf-node.llnl.gov/projects/cmip6/>, last access: 8 April 2022), which hosts data from models participating in Phase 6 of the Coupled Model Intercomparison Project (CMIP6). Potential temperature, practical salinity, and [O₂] were downloaded to coincide with available ocean interior observations (Fig. A3). Historical outputs (2004–2014) and projected outputs under Shared Socioeconomic Pathway (SSP) 2-4.5 (2015–2022) were combined to cover the time period over which observations were available. A spatial mask was applied to retain only GFDL-ESM4 grid cells with corresponding temperature and salinity values in the RG09 climatology, because that is the final grid on which GOBAI-O₂ is produced.

2.2 Algorithm training

The combined dataset was used to train ML algorithms for each region to estimate [O₂] from absolute salinity, conservative temperature, potential density anomaly, hydrostatic pressure, bottom depth, and additional spatiotemporal information to allow for geographic, seasonal, and interannual variation (see Table 1). Although biology is not explicitly accounted for in the ML algorithms, Giglio et al. (2018) demonstrate that, with an appropriately distributed dataset, the inclusion of spatiotemporal variables in algorithm training can implicitly accommodate biological processes.

Absolute salinity (S_A) was calculated from practical salinity (S_p), hydrostatic pressure (P), latitude, and longitude. Conservative temperature (θ) was calculated from in situ temperature (T), S_A , and P . Potential density anomaly (σ_θ) was calculated from S_A and θ . These calculations were made using the Gibbs–SeaWater (GSW) Oceanographic Toolbox for MATLAB (McDougall and Barker, 2011). As was done by Carter et al. (2021), longitude was transformed into two separate predictors: $\cos(\text{Longitude} - 20^\circ \text{E})$ and $\cos(\text{Longitude} - 110^\circ \text{E})$. Cosine functions were applied to maintain the cyclical nature of longitude as a predictor, and offsets of 20 and 110° E were intended to shift regions where the cosine function has minimum explanatory power over landmasses. Bottom depth was determined by matching each observational location with the corresponding bathymetry from the ETOPO2v2 global relief model (NOAA National Geophysical Data Center, 2006).

Two types of ML algorithms were trained: feed-forward neural networks (FNNs; Beale et al., 2023) and random forest regressions (RFRs; Breiman, 2001). Each algorithm type was trained on the input variables given in Table 1 to produce estimates of [O₂] (Fig. A4). Three separate FNNs were trained for each of the seven basins shown in Fig. 1, with an average of the three taken to obtain one equally weighted FNN result. The FNNs were constructed using the “feedforwardnet” function and trained using the “train” function, both from Version 14.4 of the Deep Learning Toolbox for MATLAB (R2022a). Each FNN was trained using a Levenberg–Marquardt algorithm, with 15 % of the data reserved for testing the network during training steps. Each FNN had two hidden layers, with the following combinations of neurons in the first and second layer, respectively: 20 and 10, 15 and 15, and 10 and 20. One RFR was trained for each of the seven basins shown in Fig. 1. RFRs are ensembles of decision trees, each created with a bootstrapped version of the full dataset chosen randomly with replacement. Each RFR consisted of 600 trees, a minimum leaf size of 10, and 6 of the 11 predictors used for each decision split. These parameters were chosen after some trial and error to strike a balance between computational efficiency and algorithm performance. The MATLAB “treebagger” function was used to train RFRs.

In areas where two regions overlap (see Fig. 1), weighted averages of [O₂] estimates were calculated in overlapping grid cells from each regional algorithm. These averages were weighted by distance from the center latitude line of the overlapping area (e.g., a point at 33° S in the overlapping area between the N. Sou. region, whose northern border extends to 25° S, and the Atl. region, whose southern border extends to 35° S, would be calculated as $[\text{O}_2] = 0.8[\text{O}_2]_{\text{N.Sou.}} + 0.2[\text{O}_2]_{\text{Atl.}}$). Overlapping areas were used to mitigate discontinuities at the boundaries between regions in the final gridded product.

The average of FNN and RFR estimates (ENS, for ensemble average) was used as the [O₂] estimate for a given set of input data. This ensemble-averaging procedure was implemented due to insights from previous work showing that averaging the outputs of multiple ML algorithms or linear regression models often outperforms the output from just one approach on its own (Gregor et al., 2017, 2019; Bittig et al., 2018b; Carter et al., 2021; Djeutchouang et al., 2022), likely due to complementary strengths and weaknesses of each approach. For this work, any especially erroneous result from either the FNN or RFR should be mitigated by better results from the other algorithm.

2.3 Algorithm evaluation

We performed two exercises to evaluate the effectiveness of the ML algorithms used to estimate [O₂]. The first exercise involved training separate evaluation algorithms (RFR_{Data-Eval} and FNN_{Data-Eval}), as described in Sect. 2.2, using a subset of the observational dataset for training while reserving the remaining subset for assessment. For this exercise, data were split randomly into training (80 %) and assessment (20 %) groups; this split was made according to measurement platform (cruise or float; see Fig. A5) to ensure that inherent correlations among the data points from a single cruise or float did not contribute to the apparent effectiveness of each ML algorithm. Then, [O₂] values from the subset of reserved assessment data were compared to estimates of [O₂] from RFR_{Data-Eval}, FNN_{Data-Eval}, and the ensemble average of the two (ENS_{Data-Eval}). This exercise was intended to evaluate the ability of the ML algorithms to reproduce measured data that were not involved in algorithm training (Sect. 3.1.1).

The second exercise involved training evaluation algorithms (RFR_{ESM4-Eval} and FNN_{ESM4-Eval}) using synthetic “profiles” extracted from gridded GFDL-ESM4 output at the times and locations for which observational data were available and then assessing the evaluation algorithms using spatially and temporally continuous monthly GFDL-ESM4 output from 2004 through 2022. For this exercise, synthetic profiles for algorithm training were defined by matching the latitude, longitude, month, and year of each available grid cell from the binned observational dataset with the corresponding GFDL-ESM4 output. This resulted in 75 879

Table 1. Predictor variables used to train random forest regressions and feed-forward neural networks to predict [O₂]. Unitless quantities are denoted using “–”.

Predictor variable	Abbreviation	Unit	Range (approx.)
Conservative temperature	θ	°C	–2 to 32
Absolute salinity	S_A	–	14 to 40
Potential density anomaly	σ_θ	kg m ^{–3}	9.3 to 29.3
Hydrostatic pressure	P	dbar	0 to 2000
Latitude	lat	°	–78 to 90
cos(Longitude – 20°)	long _{cos20}	–	–1 to 1
cos(Longitude – 110°)	long _{cos110}	–	–1 to 1
Bottom depth	bot	m	0 to 10 000
Year	yr	years	2004 to 2023
sin(2 π · Day of Year/365.25)	doy _{sin}	–	–1 to 1
cos(2 π · Day of Year/365.25)	doy _{cos}	–	–1 to 1

synthetic profiles for algorithm training. The RFR_{ESM4-Eval} and FNN_{ESM4-Eval} algorithms were trained as described in Sect. 2.2 with the synthetic training data and then used to produce [O₂] estimates for the complete model output. These [O₂] estimates from RFR_{ESM4-Eval}, FNN_{ESM4-Eval}, and an ensemble average of the two (ENS_{ESM4-Eval}) were compared to [O₂] values from the full GFDL-ESM4 output fields at the grid-cell level. This exercise was intended to evaluate the ability of the ML algorithms to estimate [O₂] in a spatiotemporally resolved Earth system model environment when limited to training data representative of the available collection of ocean oxygen observations (Sect. 3.1.2). The four-dimensional field of [O₂] from ENS_{ESM4-Eval} that represents a reconstruction of the GFDL-ESM4 environment, which we refer to as GOBAI-O₂-ESM4, can also be used as an analogue for how well GOBAI-O₂ (trained on real observational data; Sect. 2.4) might represent [O₂] variability in the real-world environment. For this reason, the four-dimensional field of differences between GOBAI-O₂-ESM4 and GFDL-ESM4 output were used to inform the evaluation of GOBAI-O₂ uncertainty (Sects. 2.5 and 3.2.4). Additionally, we quantified global means, seasonal-cycle amplitudes, long-term trends, and interannual variabilities in [O₂] across different pressure layers of GOBAI-O₂-ESM4. To evaluate the performance of GOBAI-O₂-ESM4 on a global scale, these metrics are compared to the same metrics for the spatiotemporally resolved GFDL-ESM4 output and subsampled grid cells in GFDL-ESM4 corresponding to observational data coverage (Sect. 3.1.2). Comparisons of global means from GOBAI-O₂-ESM4 to GFDL-ESM4 are also used to approximate uncertainty in oxygen inventories for the assessment of trends (Sect. 3.2.3).

2.4 Creation of GOBAI-O₂

FNNs and RFRs for each of the seven regions shown in Fig. 1 were trained with the full combined dataset, using the predictor variables shown in Table 1, with [O₂] as a

target variable. Then, the FNNs and RFRs were applied to S_A , θ , and σ_θ calculated from RG09 temperature and salinity fields, along with spatiotemporal information from RG09 grid cells. Weighted averages were calculated where regions overlapped, and ensemble averages (ENS) were calculated from the FNN and RFR estimates. One exception occurred in the subsurface of the north Arabian Sea (above 21° N, between 900 and 1412.5 dbar), where erroneously high ENS [O₂] values caused by a clearly nonphysical feature in RFR [O₂] values were replaced by FNN [O₂] values only. This produced a monthly gridded [O₂] product in the upper 2 km of the ocean on a global grid from January 2004 to December 2022, i.e., GOBAI-O₂ (Sharp et al., 2022; <https://doi.org/10.25921/z72m-yz67>; Sect. 3.2.1–3.2.3). GOBAI-O₂ was compared to gridded climatological oxygen fields from the 2018 World Ocean Atlas (WOA18; Garcia et al., 2019; Sect. 3.2.5), the GLODAP mapped data product (Lauvset et al., 2016; Sect. 3.2.5), and discrete measurements of oxygen from select cruises between 2004 and 2022 (Sect. 3.2.6).

2.5 Uncertainty estimation

Similar to previous studies that have estimated uncertainty in observation-based biogeochemical data products (e.g., Landschützer et al., 2014; Gregor and Gruber, 2021; Keppler et al., 2020, 2023), we combine uncertainty from three separate sources – measurement, gridding, and algorithm – to estimate uncertainty in GOBAI-O₂ (Sect. 3.2.4).

Measurement uncertainty ($u([\text{O}_2])_{\text{meas.}}$) is attributable to the [O₂] observations themselves. For this quantity, gridded [O₂] from GOBAI-O₂ is multiplied by 3 %, which is an estimate based on a few factors: the consistency of the GLODAPv2.2022 dataset, the accuracy of BGC Argo observations, the relative proportion of each data source, and the recognized issue of float oxygen sensor response time. The nominal value for the consistency of the GLODAPv2.2022 cruise dataset is stated to be 1 % (Lauvset et al., 2022a). The

approximate accuracy of BGC Argo float observations is estimated to be about 3 % of surface [O₂], which is about 4 % to 5 % of average ocean [O₂], for floats quality-controlled by climatological comparisons (Takeshita et al., 2013) and about 3 μmol kg⁻¹ (Johnson et al., 2017; Maurer et al., 2021), which is about 2 % of average ocean [O₂], for floats quality-controlled by in-air measurements. We choose a value for measurement uncertainty closer to the float accuracy estimates because float observations outweigh ship observations by about six to one in the dataset used to construct GOBAI-O₂. The estimates for the accuracy of float [O₂] sensors may be somewhat optimistic, especially when the floats are crossing large vertical oxygen gradients (i.e., steep oxyclines). This is because the response time of oxygen optodes is not instantaneous (Bittig et al., 2014, 2018a; Bittig and Körtzinger, 2017). Without response-time corrections, which are not widely applied to the BGC Argo oxygen dataset, there is the potential for systematic biases where float profiles traverse steep oxyclines, such as in the eastern tropical Pacific Ocean and North Indian Ocean.

Gridding uncertainty ($u([\text{O}_2])_{\text{grid.}}$) is attributable to using a single [O₂] value to represent a four-dimensional box that is coarser in time and space than the resolution of many processes that influence [O₂]. We estimate gridding uncertainty by (1) binning the combined GLODAP and Argo observational dataset to grid cells equal in size to the RG09 grid cells; (2) calculating the standard deviation among the observations in cells with more than 10 observations (Fig. A6); (3) fitting a multivariate polynomial regression relating those standard deviations to pressure, potential density anomaly, and bottom depth; and (4) applying that regression to the RG09 grid to compute an estimated standard deviations (i.e., gridding uncertainty) in each grid cell.

Algorithm uncertainty ($u([\text{O}_2])_{\text{alg.}}$) is attributable to the ML algorithms that estimate [O₂] on the RG09 grid. We estimate algorithm uncertainty using the four-dimensional field of absolute differences between [O₂] from GFDL-ESM4 model output and GOBAI-O₂-ESM4, determined from the GFDL-ESM4 algorithm evaluation exercise described in Sect. 2.3.

The three uncertainty sources were combined in quadrature (assuming independence) to calculate a combined uncertainty estimate for each gridded [O₂] value in GOBAI-O₂ ($u([\text{O}_2])_{\text{tot.}}$):

$$u([\text{O}_2])_{\text{tot.}} = \sqrt{u([\text{O}_2])_{\text{meas.}}^2 + u([\text{O}_2])_{\text{grid.}}^2 + u([\text{O}_2])_{\text{alg.}}^2}. \quad (1)$$

3 Results and discussion

3.1 Algorithm evaluation

The evaluation exercises indicated that the ML algorithms trained on the combined GLODAP and Argo observational dataset were effective in their ability to estimate [O₂] and reconstruct seasonal to decadal variability in the global oxy-

gen inventory. Mean offsets ($\Delta[\text{O}_2] = [\text{O}_2]_{\text{obs/mod}} - [\text{O}_2]_{\text{est}}$) and root-mean-square differences (RMSDs) between [O₂] from direct measurements ([O₂]_{obs}) or GFDL-ESM4 output ([O₂]_{mod}) and [O₂] estimated from ML algorithms ([O₂]_{est}) were determined as an assessment of the ability of the algorithms to estimate [O₂] at a grid-cell level (Tables 2, B2, B3, B4; Fig. 2). Mean $\Delta[\text{O}_2]$ and RMSD determined using [O₂]_{est} from the ESPER-Mixed model (Carter et al., 2021) – an average of predictions from a neural network and moving window multiple linear regression trained on GLODAPv2.2020 data – were also determined as a point of comparison for the observational-data-based validation test (Tables 2, B4; Fig. A7). In the case of the GFDL-ESM4-based validation test, metrics to summarize means, amplitudes, trends, and variability in integrated mean [O₂] values were determined to demonstrate the ability of the GOBAI-O₂ method to capture seasonal- to decadal-scale variability in oxygen at the global scale (Table 3, Fig. 3). The results of each evaluation exercise are discussed in more detail in the following sections.

3.1.1 Test with withheld observational data

Estimates of [O₂] using the ENS_{Data-Eval} algorithms tracked closely with [O₂]_{obs} and showed no strong systematic biases with [O₂]_{est} or depth (Fig. 2a, b), although variability in $\Delta[\text{O}_2]$ was greatest from just below the surface to about 500 dbar. Mean offsets were between -1.8 and $3.2 \mu\text{mol kg}^{-1}$ for the seven regions, with a global average of $-0.2 \mu\text{mol kg}^{-1}$; RMSDs were between 7.1 and $11.0 \mu\text{mol kg}^{-1}$ for the seven regions, with a global average of $8.8 \mu\text{mol kg}^{-1}$ (Table 2). The slightly negative global average offset suggests somewhat higher estimated than measured [O₂] values, and some of the lowest RMSDs from the ENS_{Data-Eval} algorithms were found in the Southern Ocean regions (Table 2, Fig. 2c), likely because these regions have a significant number of available training data (Fig. 1). However, this evaluation exercise is influenced by the incomplete subset of data (20 %) used to test the ENS_{Data-Eval} algorithms. A cross-fold validation (e.g., repeating this exercise with five separate 20 % chunks of data withheld from algorithm training) was prohibitively computationally expensive. Therefore, the associated $\Delta[\text{O}_2]$ and RMSD values alone are not as instructive as a comparison to the $\Delta[\text{O}_2]$ and RMSD values obtained from the ESPER-Mixed model (Table 2).

Estimates of [O₂] using ESPER-Mixed (Fig. A7) showed average offsets between -3.9 and $-1.4 \mu\text{mol kg}^{-1}$ for the seven regions (with a global average of $-2.6 \mu\text{mol kg}^{-1}$) and RMSDs between 10.0 and $21.6 \mu\text{mol kg}^{-1}$ for the seven basins (with a global average of $13.1 \mu\text{mol kg}^{-1}$) (Table 2). Again, the negative global average offset suggests higher estimated than measured [O₂] values. Compared with ESPER-Mixed (Carter et al., 2021), the ENS_{Data-Eval} algorithms performed better, both in terms of $\Delta[\text{O}_2]$ and RMSD in each individual region and overall. This result is likely a reflec-

Table 2. Regional and global error statistics (mean $\Delta[\text{O}_2]$ and RMSD) for evaluation exercises using the ensemble average (ENS_{Data-Eval}) of the FNN_{Data-Eval} and RFR_{Data-Eval} algorithms trained on a subset of data from the combined GLODAP and Argo observational dataset and tested with a separate subset of withheld data (left side of the table) or using the ensemble average (ENS_{ESM4-Eval}) of the FNN_{ESM4-Eval} and RFR_{ESM4-Eval} algorithms trained on a subset of output from GFDL-ESM4 (corresponding to locations of available Argo and GLODAP data) and tested using the full field of GFDL-ESM4 output (right side of the table). Error statistics calculated using the ESPER-Mixed model are also shown for comparison to the data-based test. The numbers of data points used in the training and assessment of each algorithm are shown.

Basin	Evaluation exercise with observational data						Evaluation exercise with GFDL-ESM4 output			
			ENS _{Data-Eval}		ESPER-Mixed				ENS _{ESM4-Eval}	
	Training data points	Assessment data points	Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)	Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)	Training data points	Assessment data points	Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)
Atl.	592 099	109 134	-1.8	9.2	-3.9	10.7	184 418	28 235 064	-1.3	9.6
Pac.	1 816 367	466 788	0.6	9.9	-3.3	15.6	533 208	69 369 456	0.0	7.4
Ind.	335 768	82 491	0.9	7.1	-3.0	11.6	86 060	20 736 144	0.2	7.2
Arc.	800 328	263 873	-1.4	9.1	-2.4	12.1	293 540	11 547 744	0.0	4.1
Med.	214 540	33 899	3.2	11.0	2.3	21.6	32 110	1 096 680	1.0	5.5
N. Sou.	2 236 153	480 846	0.2	7.2	-1.4	10.0	756 444	67 626 624	-0.1	4.4
S. Sou.	1 430 492	364 133	-0.7	8.2	-2.3	11.7	519 610	31 412 472	0.0	3.3
Global	7 425 747	1 801 164	-0.2	8.8	-2.6	13.1	2 405 390	230 024 184	-0.2	6.7

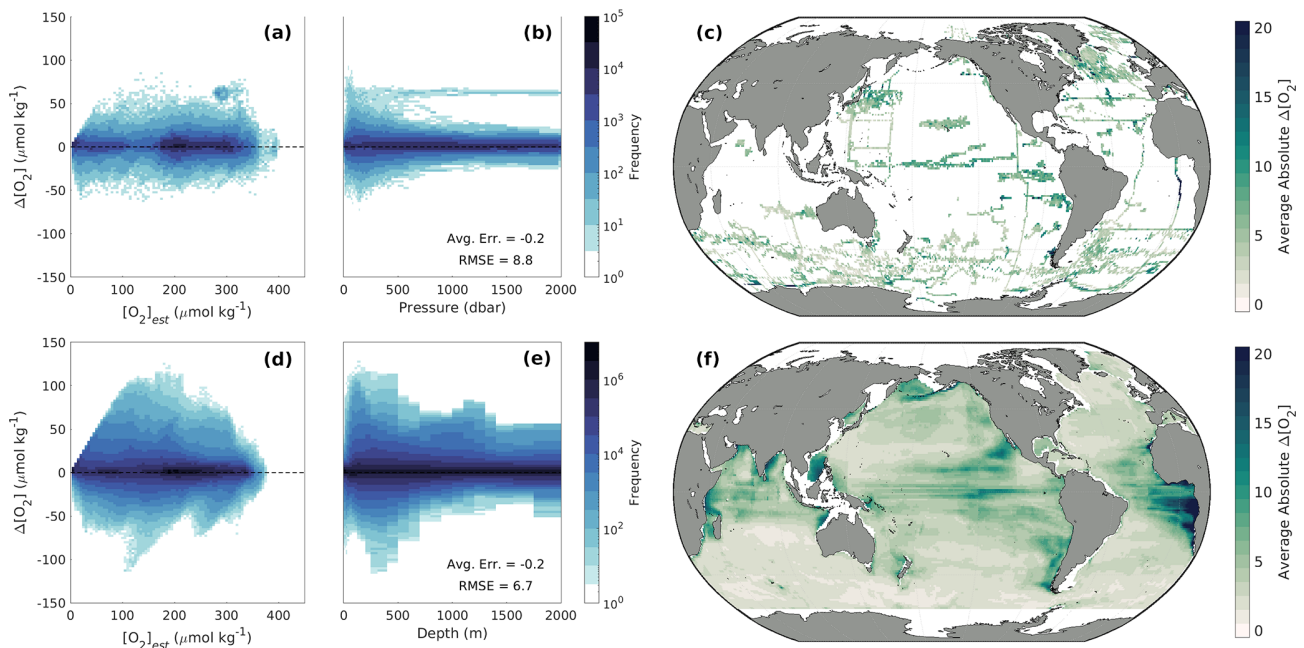


Figure 2. Two-dimensional histograms showing offsets of measured versus estimated oxygen ($\Delta[\text{O}_2] = [\text{O}_2]_{\text{obs}} - [\text{O}_2]_{\text{est}}$) for (a, b) withheld observational data and (d, e) modeled versus estimated oxygen ($\Delta[\text{O}_2] = [\text{O}_2]_{\text{mod}} - [\text{O}_2]_{\text{est}}$) for GFDL-ESM4 model output as a function of (a, d) $[\text{O}_2]_{\text{est}}$ and (b, e) depth in the water column. Offsets are binned into cells that are $2.5 \mu\text{mol kg}^{-1}$ tall in terms of $\Delta[\text{O}_2]$ and (a, d) $5 \mu\text{mol kg}^{-1}$ wide in terms of $[\text{O}_2]_{\text{est}}$ or equivalent in width to (b) the interpolated pressure levels of the data or (e) the vertical resolution of GFDL-ESM4 grid cells. The frequency of offsets that fall into a given bin is shown on a logarithmic scale, de-emphasizing the significant clustering around $\Delta[\text{O}_2] = 0$ in favor of showing the few outliers. Absolute $\Delta[\text{O}_2]$ values averaged over depth and time for $1^\circ \times 1^\circ$ (latitude \times longitude) grid cells in the global ocean for (c) withheld observational data and (f) GFDL-ESM4 model output.

tion of the fact that the ENS_{Data-Eval} algorithms were trained with more varied data than the ESPER-Mixed model (Argo and GLODAP compared with just GLODAP) and that the withheld data for which estimates were made also comprised more varied data (both Argo and GLODAP as well). Importantly, when estimates were made for just the GLODAP

dataset, the ENS_{Data-Eval} algorithms still performed better than ESPER-Mixed (Table B4), suggesting that the seasonally resolved float data supply important information to the relationships established during algorithm training.

Table 3. Statistics representing the mean values, seasonal-cycle amplitudes, long-term trends, and interannual variabilities in [O₂] from the GFDL-ESM4 model, a reconstruction of [O₂] fields from GFDL-ESM4 using the approach of GOBAI-O₂ (GOBAI-O₂-ESM4), and subsampled grid cells from GFDL-ESM4 when and where real observations are available. Global weighted means (μ) of grid-cell-level values are shown, along with differences (Δ) of the fully resolved GFDL-ESM4 means versus GOBAI-O₂-ESM4 and versus the subsampled GFDL-ESM4 grid cells.

Metric	Pressure layer (dbar)	GFDL-ESM4	GOBAI-O ₂ -ESM4		Subsampled GFDL-ESM4	
		μ	μ	Δ	μ	Δ
Mean [O ₂] ($\mu\text{mol kg}^{-1}$)	0–200	214.02	214.31	−0.29	230.12	−16.11
	200–1000	154.83	155.19	−0.36	173.54	−18.70
	0–2000	155.59	155.82	−0.23	169.63	−14.03
Seasonal-cycle amplitude ($\mu\text{mol kg}^{-1}$)	0–200	12.04	10.22	1.82	12.20	−0.16
	200–1000	3.37	2.06	1.31	5.99	−2.62
	0–2000	2.60	1.84	0.75	3.94	−1.34
Long-term trend ($\mu\text{mol kg}^{-1}$ per decade)	0–200	−0.30	−0.13	−0.17	7.10	−7.40
	200–1000	−0.48	−0.23	−0.26	5.22	−5.70
	0–2000	−0.38	−0.18	−0.20	7.28	−7.66
Interannual variability ($\mu\text{mol kg}^{-1}$)	0–200	0.22	0.20	0.03	8.95	−8.73
	200–1000	0.29	0.18	0.12	10.43	−10.13
	0–2000	0.22	0.12	0.10	10.29	−10.07

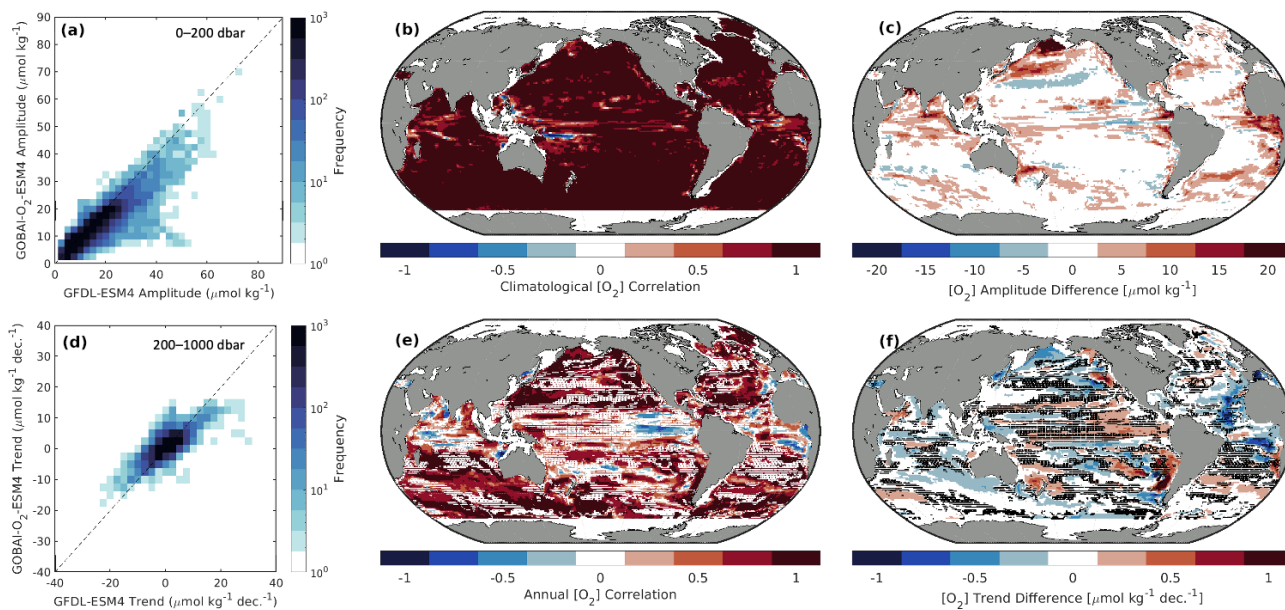


Figure 3. Two-dimensional histograms showing grid-cell-level (a) climatological seasonal amplitudes in monthly mean [O₂] (weighted means according to the size of each pressure layer) from 0 to 200 dbar and (d) trends in annual mean [O₂] from 200 to 1000 dbar between GFDL-ESM4 and GOBAI-O₂-ESM4. Pearson’s correlation coefficients between GFDL-ESM4 and GOBAI-O₂-ESM4 for (b) monthly mean [O₂] from 0 to 200 dbar, showing coherence between the surface seasonal cycles, and (e) annual mean [O₂] from 200 to 1000 dbar, showing coherence between the subsurface trends. Absolute difference between GFDL-ESM4 and GOBAI-O₂-ESM4 for (c) climatological seasonal amplitudes in monthly mean [O₂] from 0 to 200 dbar and (f) trends in annual mean [O₂] from 200 to 1000 dbar. In panels (e) and (f), stippling indicates grid cells in which the GFDL-ESM4 trend is not significantly different from zero.

3.1.2 Test with GFDL-ESM4 output

As introduced in Sect. 2.3, we refer to the four-dimensional field of $[\text{O}_2]_{\text{est}}$ values calculated by applying $\text{ENS}_{\text{ESM4-Eval}}$ algorithms to GFDL-ESM4 output as GOBAI-O₂-ESM4. $[\text{O}_2]_{\text{est}}$ values from GOBAI-O₂-ESM4 tracked closely with $[\text{O}_2]_{\text{mod}}$ and showed no significant systematic biases with $[\text{O}_2]$ or depth (Fig. 2d, e). Similar to the data-based test, variability in $\Delta[\text{O}_2]$ was greatest from just below the surface to about 500 m. Average offsets were between -1.3 and $1.0 \mu\text{mol kg}^{-1}$ for the seven regions (with a global average of $-0.2 \mu\text{mol kg}^{-1}$) and RMSDs were between 3.3 and $9.6 \mu\text{mol kg}^{-1}$ for the seven basins (with a global average of $6.7 \mu\text{mol kg}^{-1}$) (Table 2). The near-zero global average offset suggests that $[\text{O}_2]_{\text{est}}$ values from GOBAI-O₂-ESM4 matched well with values from GFDL-ESM4 output. The lowest RMSDs were found in the Southern Ocean and Arctic regions (Table 2, Fig. 2f), again due to the high density of training data in these regions.

In addition to direct comparisons of $[\text{O}_2]$ values, GOBAI-O₂-ESM4 effectively captured local decadal-scale and seasonal variability in $[\text{O}_2]$ in the GFDL-ESM4 model environment (Figs. 3, A8, A9, A10; Table 3). The average Pearson's correlation coefficient between gridded monthly mean $[\text{O}_2]$ integrated from 0 to 200 dbar from GFDL-ESM4 output and GOBAI-O₂-ESM4 was 0.92 ± 0.17 (Fig. 3b), and the seasonal amplitudes differed in magnitude (GFDL-ESM4 minus GOBAI-O₂-ESM4) by $1.8 \pm 4.0 \mu\text{mol kg}^{-1}$ (Fig. 3c). The average Pearson's correlation coefficient between gridded annual mean $[\text{O}_2]$ integrated from 200 to 1000 dbar from GFDL-ESM4 output and GOBAI-O₂-ESM4 was 0.66 ± 0.37 (Fig. 3e), and the trends differed in magnitude (GFDL-ESM4 minus GOBAI-O₂-ESM4) by $-0.3 \pm 2.1 \mu\text{mol kg}^{-1}$ per decade (Fig. 3f).

When considered on the global scale, mean values, seasonal-cycle amplitudes, long-term trends, and interannual variabilities in $[\text{O}_2]$ matched well between GFDL-ESM4 output and GOBAI-O₂-ESM4 (Table 3). In almost every case, agreement was far better than it was when simply considering GFDL-ESM4 grid cells for which observations are available over this time period, with no spatiotemporal interpolation. For example, the trend in monthly mean $[\text{O}_2]$ integrated from 0 to 2000 dbar was $-0.38 \mu\text{mol kg}^{-1}$ per decade for GFDL-ESM4 output versus $-0.18 \mu\text{mol kg}^{-1}$ per decade for GOBAI-O₂-ESM4 (Fig. A11). On the other hand, grid cells where observations are available actually indicated an increase in monthly mean $[\text{O}_2]$ integrated from 0 to 2000 dbar of $7.3 \mu\text{mol kg}^{-1}$ per decade over this time period when no spatiotemporal interpolation is applied.

Whether the internal variability in GFDL-ESM4 is truly representative of the ocean or is biased in one or more dimensions, the success of GOBAI-O₂-ESM4 in this evaluation exercise demonstrates an ability for the ML algorithms employed here to capture that variability using the current distribution of available $[\text{O}_2]$ observations as training data.

This bodes well with respect to the ability of GOBAI-O₂, which is trained on actual observational data, to represent decadal-scale and seasonal variability in global ocean oxygen in the real world. However, the GFDL-ESM4 output has undergone substantial spatial and temporal averaging and has no observational uncertainties, and thus the assessed skill can be thought of as an upper limit of the reconstruction skill achievable with the currently available observations.

The results of the exercise with GFDL-ESM4 model output are critical for evaluating the uncertainty in gridded oxygen values in GOBAI-O₂ (Sect. 3.2.4). Further, the spatial distribution of $\Delta[\text{O}_2]$ (Fig. 2f) and the comparisons of reconstructed to modeled decadal trends and seasonal variability (Figs. 3b, c, e, and f and A8–A10) can help inform our observing efforts (e.g., future cruise planning and BGC Argo float deployments). For example, large $\Delta[\text{O}_2]$ values in the eastern tropical Pacific and eastern tropical Atlantic, coupled with some negative correlations in annual mean $[\text{O}_2]$ and differences in annual trends and seasonal amplitudes, suggest that more observations will be required for GOBAI-O₂ (or likely any observation-based gap-filled $[\text{O}_2]$ data product) to fully capture variability in that region.

3.2 GOBAI-O₂ product

3.2.1 Spatial oxygen distribution

The full GOBAI-O₂ product is available at <https://doi.org/10.25921/z72m-yz67> (Sharp et al., 2022). Vertical–meridional sections of oxygen (Figs. 4, 5) show that surface oxygen concentrations are generally high, as these waters tend to be near equilibrium with the atmosphere. This is particularly true at high latitudes, where cold, dense waters have a high capacity for dissolved oxygen. Southern Ocean surface waters, however, are generally undersaturated with respect to oxygen (Fig. A12), consistent with observations from previous studies that suggest this undersaturation is the result of O₂-depleted thermocline water upwelling into the mixed layer (Chierici et al., 2004; Reuer et al., 2007; Jonsson et al., 2013) making the Southern Ocean on average an oxygen sink (Gruber et al., 2001; Bushinsky et al., 2017). This phenomenon can also be seen in the equatorial Pacific (Fig. A12). Undersaturation in high-latitude regions that are ice-covered during parts of the year can also be the result of limited air–sea gas exchange when sea ice is present.

Isobaric maps, isopycnal maps, and vertical–meridional sections with pressure and density vertical coordinates (Figs. 4, 5) also reveal the $[\text{O}_2]$ signatures of distinct subsurface water masses. In each basin, well-ventilated Subtropical Mode Water can be identified by relatively high $[\text{O}_2]$ at mid-latitudes on the 300 dbar surface (Fig. 4a) and along dips in isopycnals plotted against pressure and latitude (Fig. 4b, c, d) or along sloping isobars plotted against density and latitude (Fig. 5b, c, d) within the upper ~ 500 dbar. Beneath mode waters in the southern portions of each basin, Antarctic In-

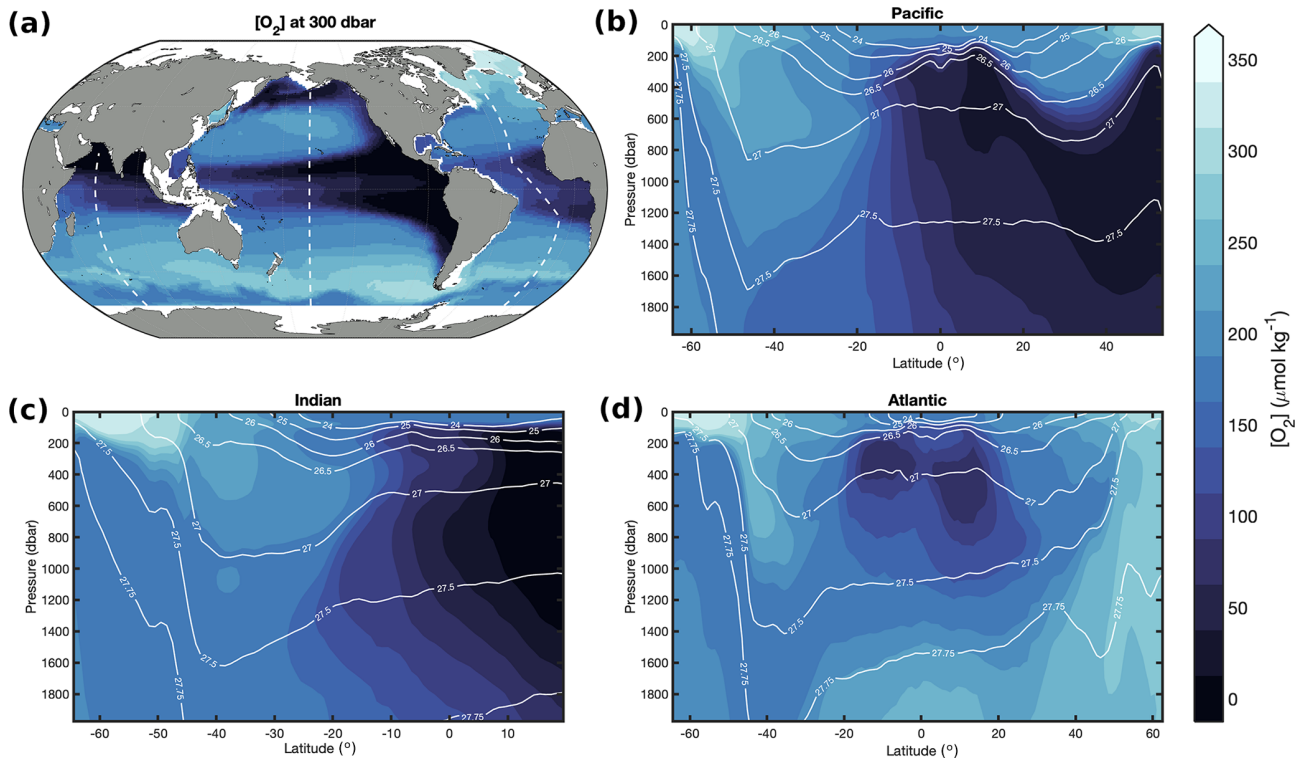


Figure 4. Long-term mean $[O_2]$ from GOBAI-O₂ at (a) 300 dbar and from the surface to 2000 dbar in the (b) Pacific, (c) Indian, and (d) Atlantic oceans. Dashed white lines in panel (a) show the locations of the sections in panels (b)–(d). White contour lines in panels (b)–(d) are potential isopycnals (kg m^{-3}).

intermediate Water that originates in the Southern Ocean with a relatively high $[O_2]$ signal is prevalent. Beneath mode waters in the northern portions of the Pacific and Indian basins, respectively, relatively old and oxygen-poor North Pacific Intermediate Water (NPIW) and Red Sea Overflow Water (RSOW) can be observed (Talley et al., 2011). Beneath mode waters in the northern portion of the Atlantic basin, intermediate waters are younger and more highly oxygenated. Near the Equator, subsurface oxygen minima are visible in each basin; this is a result of organic matter export from high production in the surface ocean that fuels strong subsurface respiration and relatively poor ventilation (old waters) in this region. Finally, the signatures of higher-oxygen deep or bottom waters can be observed near the bottom or at high latitudes in each vertical–meridional section.

Oxygen concentrations at 300 dbar (Fig. 4a) are highest in the North Atlantic and Southern oceans – where highly oxygenated, newly formed deep and intermediate waters are formed – and lowest in the northern and equatorial Pacific Ocean and the North Indian Ocean – where the oxygen content of subsurface waters has been greatly reduced by heterotrophic respiration over time. The same can be said for $[O_2]$ on the $27.0 \text{ kg m}^{-3} \sigma_\theta$ surface (Fig. 5a). Oxygen concentrations are extremely low in the deep North Pacific Ocean (Figs. 4b, 5b) and North Indian Ocean (Figs. 4c, 5c)

due to the accumulated effects of oxygen-depleting respiration over the long lifespans of those water masses (i.e., long time since gas exchange with the atmosphere).

3.2.2 Climatological seasonal oxygen cycles

Seasonal cycles in $[O_2]$ reflect a balance among physical and biological processes (Wang et al., 2022). The climatological hemispheric mean $[O_2]$ integrated over three pressure layers from GOBAI-O₂ (Fig. 6) reveals that the magnitude of the $[O_2]$ seasonal cycle is greatest near the surface and decreases with depth. The amplitude of the $[O_2]$ seasonal cycle in a near-surface layer (0–100 dbar) is about $10.7 \mu\text{mol kg}^{-1}$ in the Northern Hemisphere and $8.9 \mu\text{mol kg}^{-1}$ in the Southern Hemisphere. Maximum $[O_2]$ in this pressure layer (April/May in the Northern Hemisphere and October/November in the Southern Hemisphere) lags about 2 months behind the temperature minimum, suggesting an interaction between a thermally driven increase in oxygen solubility and biologically driven oxygen production. Minimum $[O_2]$ in the near-surface layer (October in the Northern Hemisphere and March/April in the Southern Hemisphere) is more coincident with the temperature maximum, indicating primary control by a thermally driven decrease in oxygen solubility. The amplitude of the $[O_2]$ seasonal cycle is about $2.4 \mu\text{mol kg}^{-1}$ in the Northern Hemisphere and $2.6 \mu\text{mol kg}^{-1}$ in the South-

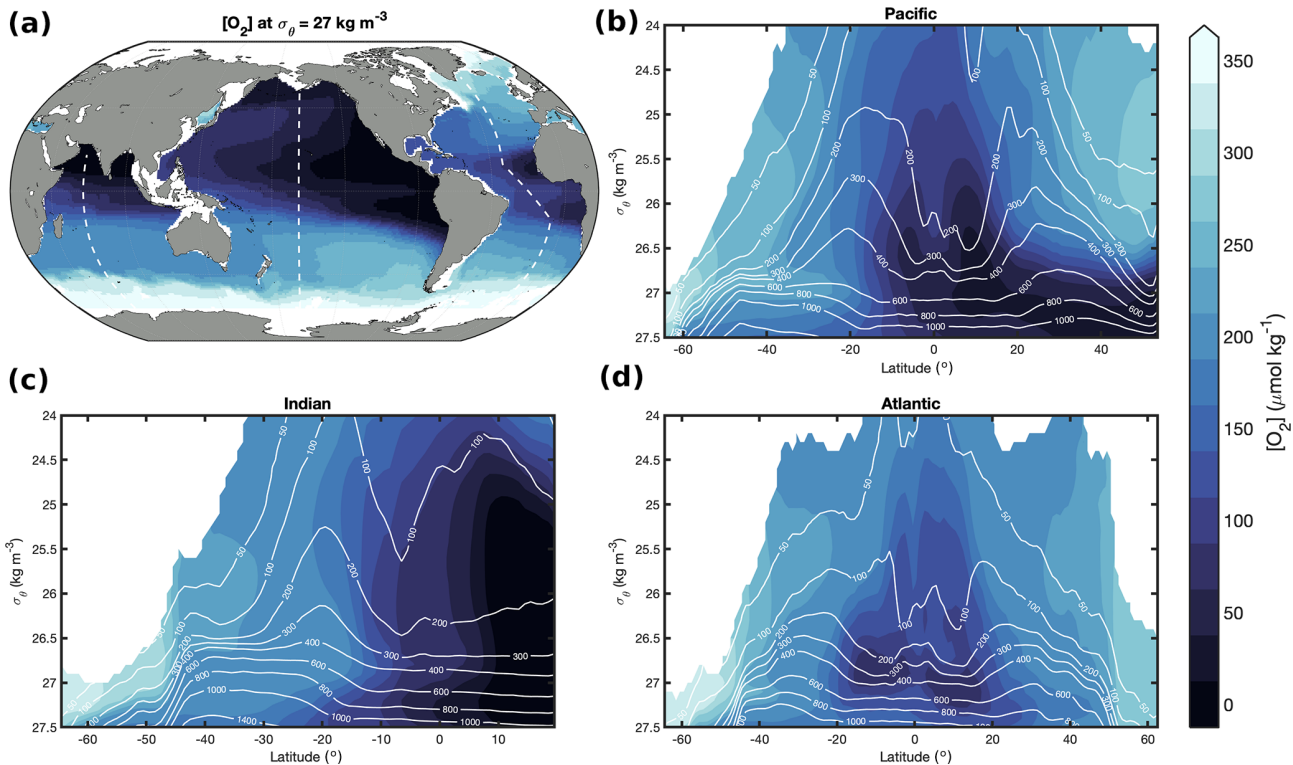


Figure 5. Long-term mean $[O_2]$ from GOBAI-O₂ at (a) $\sigma_\theta = 27 \text{ kg m}^{-3}$ and from $\sigma_\theta = 24$ to 27.5 kg m^{-3} in the (b) Pacific, (c) Indian, and (d) Atlantic oceans. Dashed white lines in panel (a) show the locations of the sections in panels (b)–(d). White contour lines in panels (b)–(d) are constant isobars (dbar).

ern Hemisphere in the intermediate layer (100–600 dbar), whereas it is about $0.2 \mu\text{mol kg}^{-1}$ in the Northern Hemisphere and $0.1 \mu\text{mol kg}^{-1}$ in the Southern Hemisphere in the deep layer (600–2000 dbar). The timing of maximum $[O_2]$ values is similar between the near-surface layer and the intermediate layer in both hemispheres, indicating the well-mixed nature of the ocean in winter and early spring when $[O_2]$ is high. On the other hand, minimum $[O_2]$ in the intermediate layer lags behind that in the near-surface layer in both hemispheres, possibly reflecting higher stratification in the upper ocean when temperatures are warmer and/or the remineralization of sinking organic matter after summer production. Further analysis of climatological $[O_2]$ cycles from GOBAI-O₂ can provide insight into the physical and biological factors that control surface and subsurface oxygen on regional and global scales.

3.2.3 Interannual oxygen trends and variability

Deoxygenation is evident in GOBAI-O₂ over the past 2 decades, coincident with ocean warming (Fig. 7, Table B5). The spatially weighted rate of deoxygenation in the upper 2 km globally (along with a 95 % confidence interval) is $-1.19 \pm 0.05 \mu\text{mol kg}^{-1}$ per decade ($-0.79 \pm 0.04 \%$ per decade). The rate of deoxygenation in GOBAI-O₂ varies over

depth, with a near-surface-pressure-layer (0–100 dbar) trend in $[O_2]$ of $-1.10 \pm 0.60 \mu\text{mol kg}^{-1}$ per decade ($-0.49 \pm 0.26 \%$ per decade), an intermediate-layer (100–600 dbar) trend in $[O_2]$ of $-1.38 \pm 0.42 \mu\text{mol kg}^{-1}$ per decade ($-0.86 \pm 0.26 \%$ per decade), and a deep-layer (600–2000 dbar) trend in $[O_2]$ of $-1.12 \pm 0.54 \mu\text{mol kg}^{-1}$ per decade ($-0.79 \pm 0.38 \%$ per decade). Interannual variability is greatest in the near-surface layer: when the multiyear trends and seasonal cycles are removed, the standard deviation of annual global mean $[O_2]$ anomalies is $0.52 \mu\text{mol kg}^{-1}$ in the near-surface layer compared with $0.26 \mu\text{mol kg}^{-1}$ in the intermediate layer and $0.12 \mu\text{mol kg}^{-1}$ in the deep layer. Trends and uncertainties were determined by fitting linear least-squares models to spatially weighted monthly mean $[O_2]$ and monthly oxygen inventories integrated over the specified pressure layers, with uncertainties in monthly values determined by comparing GOBAI-O₂-ESM4 to GFDL-ESM4; more information on this is provided in Appendix E.

Ocean warming has a direct effect on oxygen concentrations by lowering the solubility of O₂ in ocean water (Garcia and Gordon, 1992). Solubility changes explain about 56 % of deoxygenation in the near-surface ocean layer (0–100 dbar), 20 % in the intermediate ocean layer (100–600 dbar), and 15 % in the deep ocean layer (600–2000 dbar) (Fig. 7c, f).

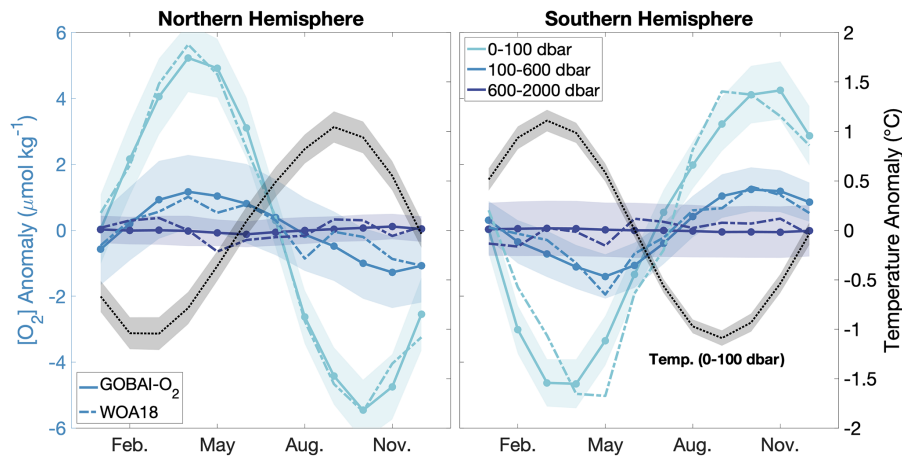


Figure 6. Climatological seasonal cycles of [O₂] anomalies (monthly [O₂] minus long-term mean [O₂]) integrated globally over three pressure layers: 0–100, 100–600, and 600–2000 dbar. The dotted black line shows the climatological temperature anomaly integrated globally over the 0–100 dbar layer. Shading indicates the standard deviation of the climatological seasonal cycle from 2004 to 2022. The dashed lines show climatological seasonal cycles of [O₂] anomalies from WOA18 over similar depth layers to GOBAI-O₂: 0–100, 100–600, and 600–1500 m.

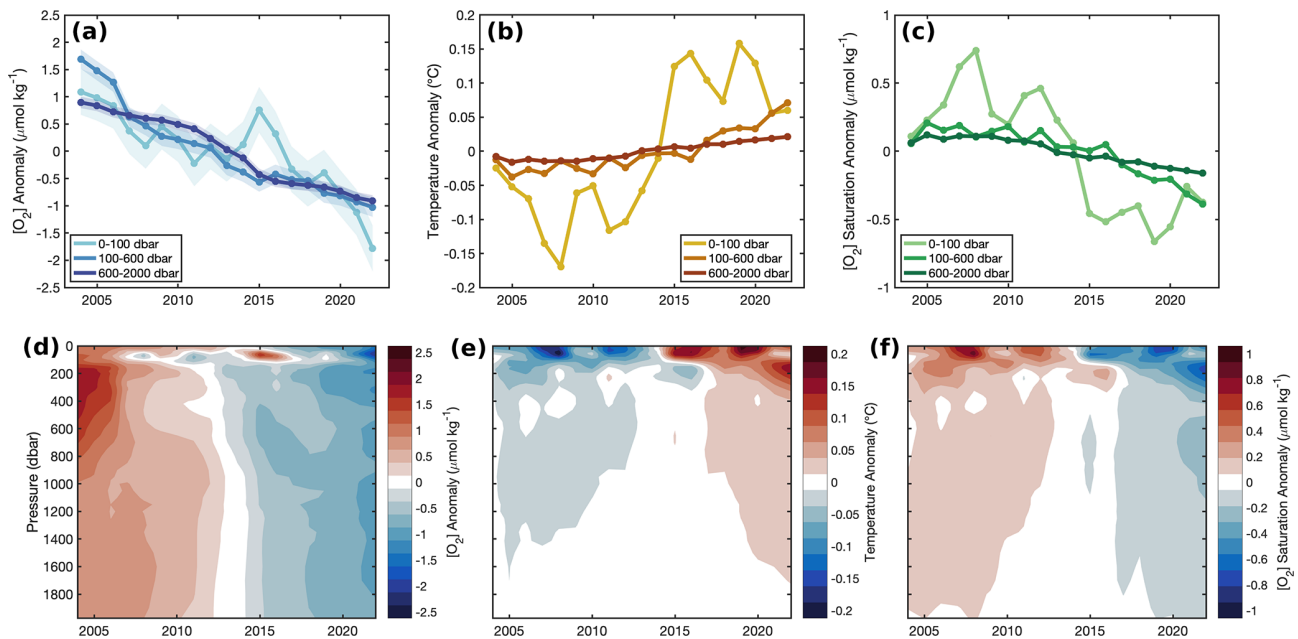


Figure 7. Annual mean (a) [O₂] anomalies from GOBAI-O₂, (b) temperature anomalies from RG09, and (c) [O₂]_{sat.} anomalies calculated from RG09 temperature and salinity fields, each integrated globally over three pressure layers: 0–100, 100–600, and 600–2000 dbar. In panel (a), shading represents uncertainty determined as the average difference of mean [O₂] from GOBAI-O₂-ESM4 versus GFDL-ESM4 in each layer. Hovmöller diagrams showing annual mean (d) [O₂] anomalies from GOBAI-O₂, (e) temperature anomalies from RG09, and (f) [O₂]_{sat.} anomalies calculated from RG09 temperature and salinity fields, each versus pressure in decibars and time from 2004 to 2022. Anomalies in each parameter are calculated as annual mean values minus the long-term mean either (a–c) integrated over a pressure layer or (d–f) on a given pressure level.

The remaining deoxygenation must then be caused by the indirect consequences of ocean warming (such as increased ocean stratification and, hence, decreased subsurface ventilation) or other processes, including changes in oxygen utilization and ocean ventilation variability (Oschlies et al., 2018),

the magnitudes of which this analysis does not attempt to deconvolve. The RG09 temperature and salinity fields are constructed such that they relax toward the climatological means during periods of low data density. For this reason, toward the beginning of the time series when fewer obser-

vations are available, temperature is biased somewhat high (Fig. 7b, e) and, therefore, O₂ solubility is biased somewhat low (Fig. 7c, f). This artifact may influence GOBAI-O₂ (Fig. 7a, d), as it was constructed using the RG09 temperature and salinity fields; however, its influence is partially mitigated because temporal information included in the training and application of the GOBAI-O₂ algorithms allows for the trend inherent in the underlying oxygen data to be retained.

GOBAI-O₂ trends can be viewed in the context of other recent analyses that have explored long-term changes in ocean oxygen. From the surface to 1000 dbar, the GOBAI-O₂ trend of -0.82 ± 0.11 % per decade from 2004 to 2022 is larger than that assessed by Bindoff et al. (2019) of -0.48 ± 0.35 % per decade from 1970 to 2010 (surface to 1000 m), which takes into account estimates from Helm et al. (2011) (-0.44 ± 0.14 % per decade), Schmidtko et al. (2017) (-0.34 ± 0.35 % per decade), and Ito et al. (2017) (-0.68 ± 0.33 % per decade). In the surface layer (0–100 dbar), the GOBAI-O₂ trend of -0.49 ± 0.26 % per decade can be compared to the Bindoff et al. (2019) assessment of -0.28 ± 0.24 % per decade; in the intermediate layer (100–600 dbar), the GOBAI-O₂ trend of -0.86 ± 0.26 % per decade can be compared to the Bindoff et al. (2019) assessment of -0.52 ± 0.36 % per decade. Considering that these comparisons represent different periods of time such that one should not expect perfect agreement, we find the results encouraging. The somewhat more negative GOBAI-O₂ trends compared with previous estimates suggest a possible acceleration of ocean deoxygenation over the last decade or so, which would be consistent with expectations (Kwiatkowski et al., 2020). Further, agreement between GOBAI-O₂ and other observation-based studies provides additional support for the notion that current ESMs, which exhibit weaker deoxygenation trends (see Sect. 3.1.2), may not fully capture the sensitivities of physical and biological processes leading to deoxygenation (Oschlies et al., 2017, 2018; Stramma and Schmidtko, 2021). This comparison not only places the GOBAI-O₂ trends in a longer-term context but suggests that the enhanced observations and analysis result in a reduced trend uncertainty despite the comparatively shorter 19-year record (± 0.04 % per decade) versus the longer but more sparse 40-year records assessed by Bindoff et al. (2019) (± 0.14 to ± 0.35 % per decade).

The trends presented here represent both natural and potentially anthropogenic variability over the interval between 2004 and 2022 as well as uncertainties in the algorithm predictions (see Sect. 3.2.4). As such, these trends should not be interpreted to be driven exclusively by ocean warming and other associated impacts of anthropogenic climate change; the period of time examined is relatively short, and the domain is not inclusive of the entire global ocean. Accordingly, decadal-scale variability in ocean ventilation, interior circulation, and biological oxygen utilization may exert significant influence over these trends. This is especially true of the regional trends. Finally, a sensitivity test indicated that in-

cluding versus withholding temporal predictors did not significantly impact the global [O₂] trend. However, a shift from the relative dominance of shipboard observations during the early portion of the GOBAI-O₂ time span to the relative dominance of float observations during the later portion cannot be ignored as a potential contributor to a deoxygenation trend, especially considering the potential for systematic biases in the float [O₂] dataset (see Sect. 2.5). This kind of shift in measurement platforms has precedent for producing spurious trends in oceanographic observational datasets (Ryckaczewski and Dunne, 2011).

3.2.4 Uncertainty

GOBAI-O₂ uncertainty fields, which were estimated as described in Sect. 2.5, can be used to assess confidence in multiyear trends and seasonal cycles of [O₂], both on a global and regional scale. Time-averaged uncertainty fields at 150 dbar (Fig. 8) suggest that the largest driver of geographic variability in uncertainty is the algorithm uncertainty. Averaged globally over space and time, $u([\text{O}_2])_{\text{meas.}}$ was equal to $4.5 \mu\text{mol kg}^{-1}$ ($5.6 \mu\text{mol kg}^{-1}$ on the 150 dbar level), $u([\text{O}_2])_{\text{grid.}}$ was equal to $3.1 \mu\text{mol kg}^{-1}$ ($5.3 \mu\text{mol kg}^{-1}$ on the 150 dbar level), and $u([\text{O}_2])_{\text{alg.}}$ was equal to $3.8 \mu\text{mol kg}^{-1}$ ($6.3 \mu\text{mol kg}^{-1}$ on the 150 dbar level). Combined, $u([\text{O}_2])_{\text{tot.}}$ (Eq. 1) was equal to $7.6 \mu\text{mol kg}^{-1}$ ($11.2 \mu\text{mol kg}^{-1}$ on the 150 dbar level), which can be compared to the global average RMSD of $8.8 \mu\text{mol kg}^{-1}$ determined independently by withholding data from algorithm training (Table 2; Fig. 2a, b, c).

Measurement uncertainty provides an estimate of confidence in an [O₂] value assigned to a water sample by direct measurement, gridding uncertainty provides an estimate of confidence that the [O₂] value provided for a four-dimensional grid cell might represent [O₂] at any point in time and space within that grid cell, and algorithm uncertainty provides an estimate of confidence that the predicted [O₂] value for a given grid cell is appropriate as the average value for that grid cell. Algorithm uncertainty, in particular, depends upon the distribution of data available to train the ML algorithms and the ability of the trained algorithms to represent underlying variability in the system. On the isobar shown in Fig. 8. (150 dbar), the underlying variability is relatively high in [O₂] minimum zones (e.g., near the Equator and on the eastern boundaries of ocean basins), hence the elevated algorithm (and total) uncertainties in those regions. Here, algorithm uncertainty was assessed via the exercise with synthetic data from GFDL-ESM4 (see Sects. 2.3 and 3.1.2).

Algorithm uncertainty should generally decrease as the spatiotemporal coverage of available training data increases. Regionally, algorithm uncertainty depends upon the degree to which the underlying variability in the system is captured by the available training observations and the ability of the ML algorithms to reconstruct that variability from concurrent

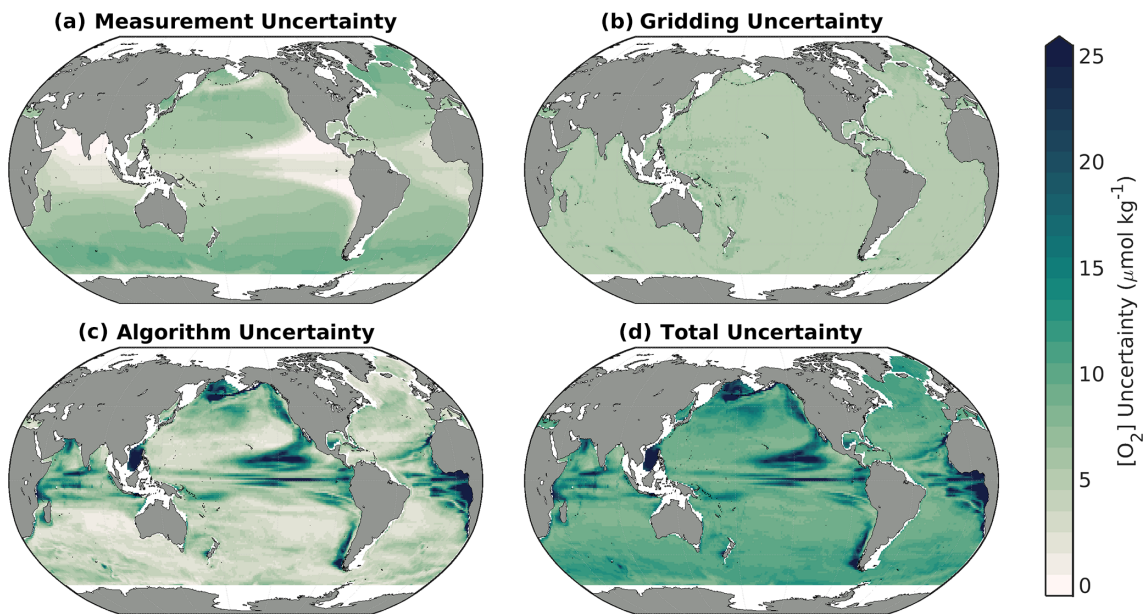


Figure 8. Long-term means of the uncertainty contributors to GOBAI-O₂ at 150 dbar, including (a) measurement uncertainty, (b) gridding uncertainty, (c) algorithm uncertainty, and (d) total uncertainty.

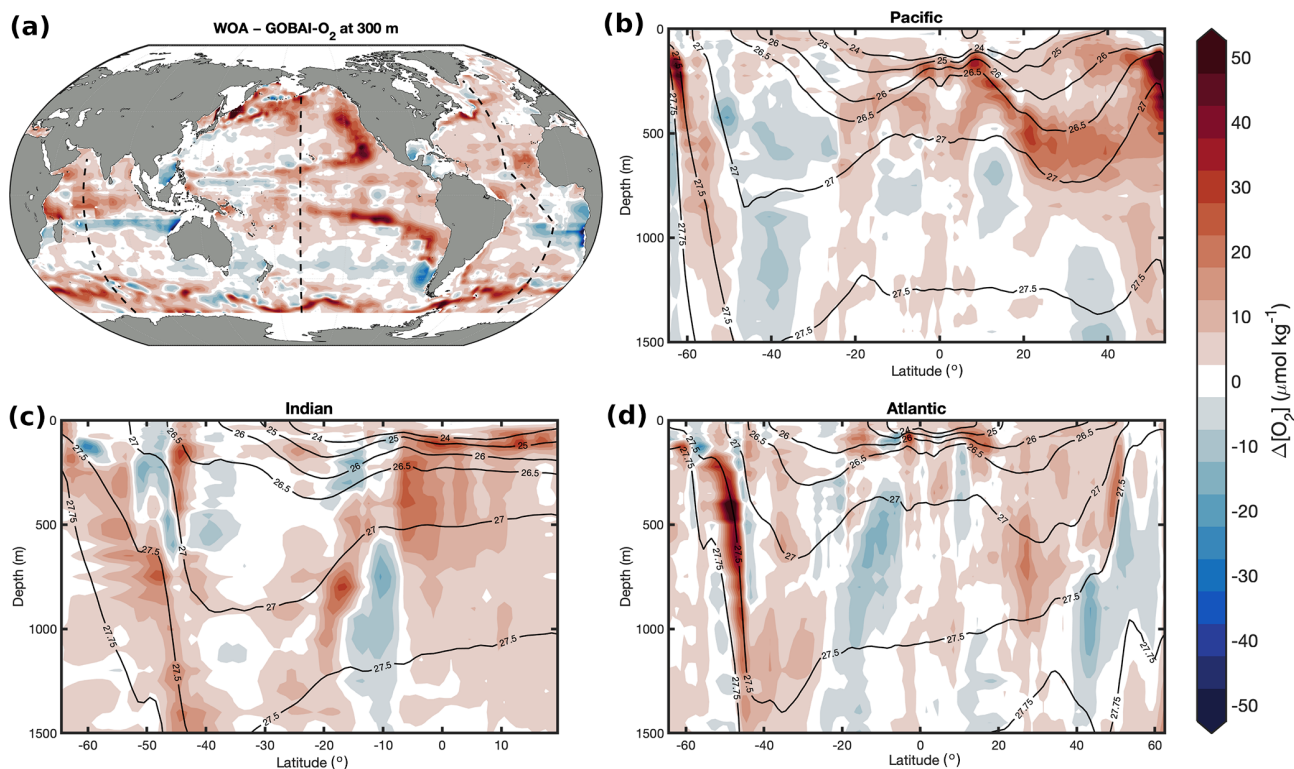


Figure 9. The difference between the climatological mean [O₂] from WOA18 and the long-term mean [O₂] from GOBAI-O₂ ($\Delta[\text{O}_2] = [\text{O}_2]_{\text{WOA}} - [\text{O}_2]_{\text{GOBAI}}$) at (a) 300 m and from the surface to 1500 dbar in the (b) Pacific, (c) Indian, and (d) Atlantic oceans.

measurements of other seawater properties. A comparison of the $\Delta[\text{O}_2]$ map in Fig. 2f or the algorithm uncertainty map in Fig. 8c to the data distribution map in Fig. 1 or Fig. A1 suggests that sparse sampling is primarily to blame for limitations related to algorithm uncertainty. Detailed analysis of GFDL-ESM4 water mass characteristics in the California Current system has also revealed that high uncertainties occur where water masses with similar physical characteristics but different oxygen signatures mix, underscoring that the measurement of additional biogeochemical parameters can supplement the $[\text{O}_2]$ estimates presented here, which are based on physical properties and spatiotemporal information. Overall, the significant influence of algorithm uncertainty is consistent with uncertainty analyses conducted for gap-filling methods applied to other ocean biogeochemical variables (e.g., Landschützer et al., 2014; Gregor and Gruber, 2021). For this reason, continued expansion of oxygen observations in undersampled regions will be critical to reduce uncertainty in our gap filling, and ultimately our understanding, of global subsurface oxygen distributions and variability. Similarly, the significant influence of measurement uncertainty underscores the importance of continued development of oxygen sensor calibration (Bittig et al., 2018a) and data quality control (Maurer et al., 2021) from the evolving BGC Argo fleet.

Global mean depth profiles of uncertainty contributors (Fig. A14) emphasize the general attenuation of uncertainty away from the surface, with subsurface maxima of algorithm uncertainty at 200 dbar and total uncertainty at 100 dbar. The algorithm uncertainty maximum corresponds to pressures at which vertical gradients in $[\text{O}_2]$ are relatively high (see Fig. 4). Here, small variations in the depths of density surfaces can influence $[\text{O}_2]$ at a given pressure level; this variability is challenging to capture, even with potential density as a predictor variable in the ML models (see Table 1). The total uncertainty maximum represents this vertical gradient effect balanced against relatively high measurement uncertainty closer to the surface, associated with higher $[\text{O}_2]$.

3.2.5 Comparison to other gridded products

The long-term mean field of $[\text{O}_2]$ from GOBAI-O₂ was compared to the corresponding mean field of $[\text{O}_2]$ from the WOA18 monthly climatology (Fig. 9) and the climatological field of $[\text{O}_2]$ from the GLODAPv2.2016 mapped product (Fig. 10). On average, the GOBAI-O₂ oxygen concentration is $1.4 \mu\text{mol kg}^{-1}$ lower than GLODAP and $9.6 \mu\text{mol kg}^{-1}$ lower than WOA18. This can be partly explained by the fact that GOBAI-O₂ is centered on the year 2012, whereas observations in GLODAPv2.2016 are centered around 2002 (Lauvset et al., 2016), WOA18 takes into account $[\text{O}_2]$ observations dating back to 1965, and global deoxygenation has occurred in recent decades (Bindoff et al., 2019). Spatially, the largest differences occur within and especially near the boundaries of oxygen minimum zones (eastern tropical Pa-

cific, eastern Atlantic coastal, and northern Indian zones), along $\sigma \approx 27.5 \text{ kg m}^{-3}$ in the Southern Ocean, and along $\sigma \approx 26.75 \text{ kg m}^{-3}$ in the North Pacific. It is difficult to determine whether these differences are functions of data availability (ship data for WOA18 and GLODAP versus ship and float data for GOBAI-O₂), representative time period, or mapping method (objective interpolation for WOA18 and GLODAP versus machine learning algorithms for GOBAI-O₂). A future intercomparison exercise between mapping methods using an identical starting dataset could be helpful in diagnosing these differences among gridded products.

3.2.6 Comparison to synoptic in situ measurements

GOBAI-O₂ was compared to direct observations from repeat hydrography cruises, including meridional transects across the Atlantic (A16 in 2013 and A20 in 2021), Pacific (P16 in 2005), and Indian (I08 and I09 in 2016) oceans, as well as a zonal transect across the Pacific Ocean (P02 in 2012). This exercise assessed how well monthly $[\text{O}_2]$ estimates from GOBAI-O₂ were able to represent high-quality $[\text{O}_2]$ measurements at distinct points in time and space. Due to fundamental differences between gridded estimates and point observations, we do not expect every matchup to be perfect. However, we would hope to see general coherence in mean values across large-scale ocean sections and to see a pattern of differences that make sense given our a priori expectations.

For the cruise datasets examined, GOBAI-O₂ estimates matched fairly well with discrete measurements in the mixed layer and below ~ 1000 dbar (Fig. 11). At intermediate depths, however, large differences occasionally occur. These large differences tended to cluster around areas with strong vertical gradients in $[\text{O}_2]$ (thin contours in Fig. 11 represent increments of $50 \mu\text{mol}$ in $[\text{O}_2]$). Comparison of Fig. 11 to Fig. A15 provides confidence in our uncertainty evaluation: larger differences between discrete measurements and GOBAI-O₂ occur where $u([\text{O}_2])_{\text{tot}}$ is large. Median biases, mean biases, and RMSDs between direct observations and GOBAI-O₂ are given in Table B6.

4 Data availability

GOBAI-O₂ is available as a NetCDF file at <https://doi.org/10.25921/z72m-yz67> (Sharp et al., 2022); additional information and animations can be found at <https://www.pmel.noaa.gov/gobai/> (last access: 28 September 2023). GLODAPv2.2022 is available at <https://doi.org/10.25921/1f4w-0t92> (Lauvset et al., 2022b); the GLODAP database is updated annually and is made available at <https://www.glodap.info> (last access: 28 September 2023). GFDL-ESM4 model output is available at <https://doi.org/10.22033/ESGF/CMIP6.1407> (Krasting et al., 2018). Data from the 2018 World Ocean Atlas are available at <https://www.ncei.noaa>

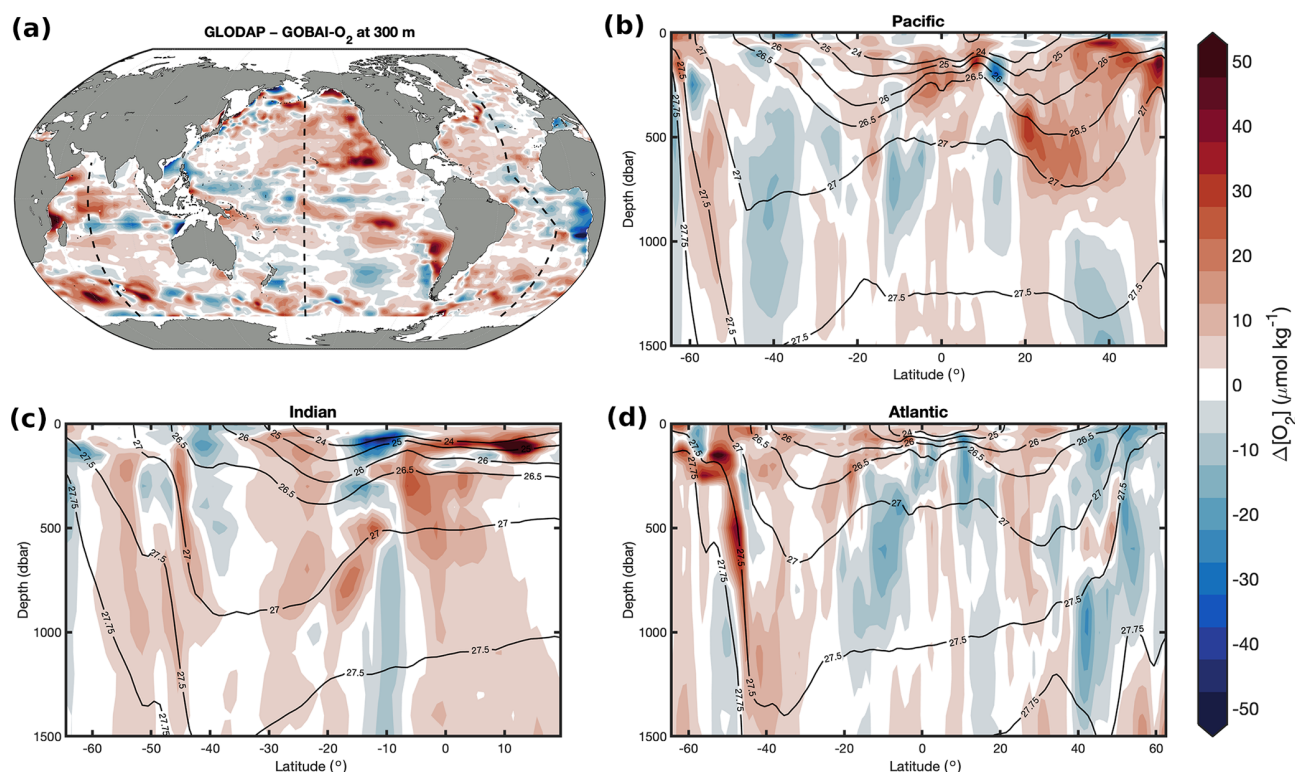


Figure 10. The difference between the 2002-centered mean $[O_2]$ from the GLODAPv2.2016 mapped product and the long-term mean $[O_2]$ from GOBAI-O₂ ($\Delta[O_2] = [O_2]_{\text{GLODAP}} - [O_2]_{\text{GOBAI}}$) at (a) 300 m and from the surface to 1500 dbar in the (b) Pacific, (c) Indian, and (d) Atlantic oceans.

gov/archive/accession/NCEI-WOA18 (Garcia et al., 2019; last access: 01 Oct. 2021). The OneArgo-Mat toolbox used to download Argo float data is available at <https://doi.org/10.5281/zenodo.6588041> (Frenzel et al., 2022); the toolbox acquires data from two global data assembly centers: Coriolis and US GODAE (<https://doi.org/10.17882/42182>, Argo, 2000). The Roemmich and Gilson (2009) Argo-based temperature and salinity product is available at https://sio-argo.ucsd.edu/RG_Climatology.html (last access: 12 January 2023).

5 Conclusions

GOBAI-O₂ is a major step toward the fulfillment of the primary goal set out by Gruber et al. (2010): “to determine, on a global-scale, seasonal to decadal time-scale variations in dissolved oxygen concentrations throughout the upper ocean”. Quantifying these variations is important for documenting ocean deoxygenation, determining global net primary productivity and carbon export, and facilitating studies of the oceanic uptake of anthropogenic CO₂. In addition, insights into ocean biogeochemical dynamics, when observations are unavailable, often come from ocean models, and GOBAI-O₂ can bring value to modeling studies by providing fields of

$[O_2]$ to be used for boundary conditions and model initialization. GOBAI-O₂ can also be useful as a dynamic reference check for new, sensor-based $[O_2]$ measurements that would otherwise be compared to a static monthly climatology like WOA18. Still, users should carefully consider the associated spatial uncertainty fields, especially when conducting regional analyses. Spatial and temporal errors and discontinuities may be significant when GOBAI-O₂ is analyzed over small areas, but they are mitigated when looking at broader scales.

The uncertainty analysis conducted here confirms that GOBAI-O₂ remains limited, largely by sparse sampling and inadequate representation of $[O_2]$ across strong gradients. The most consequential actions to improve GOBAI-O₂ fields over the next decade will be the continued deployment of Argo floats with oxygen optodes, emphasizing the importance of bolstering the biogeochemical Argo array and expanding the international OneArgo network into high latitudes, the deep ocean, and marginal seas (Roemmich et al., 2019, 2021; Schofield et al., 2022); continued work toward ensuring reliable measurements from those optodes; and the continued collection of discrete dissolved oxygen observations, primarily through the international GO-SHIP program, both for use in $[O_2]$ mapping and for calibration/validation of the Argo oxygen data.

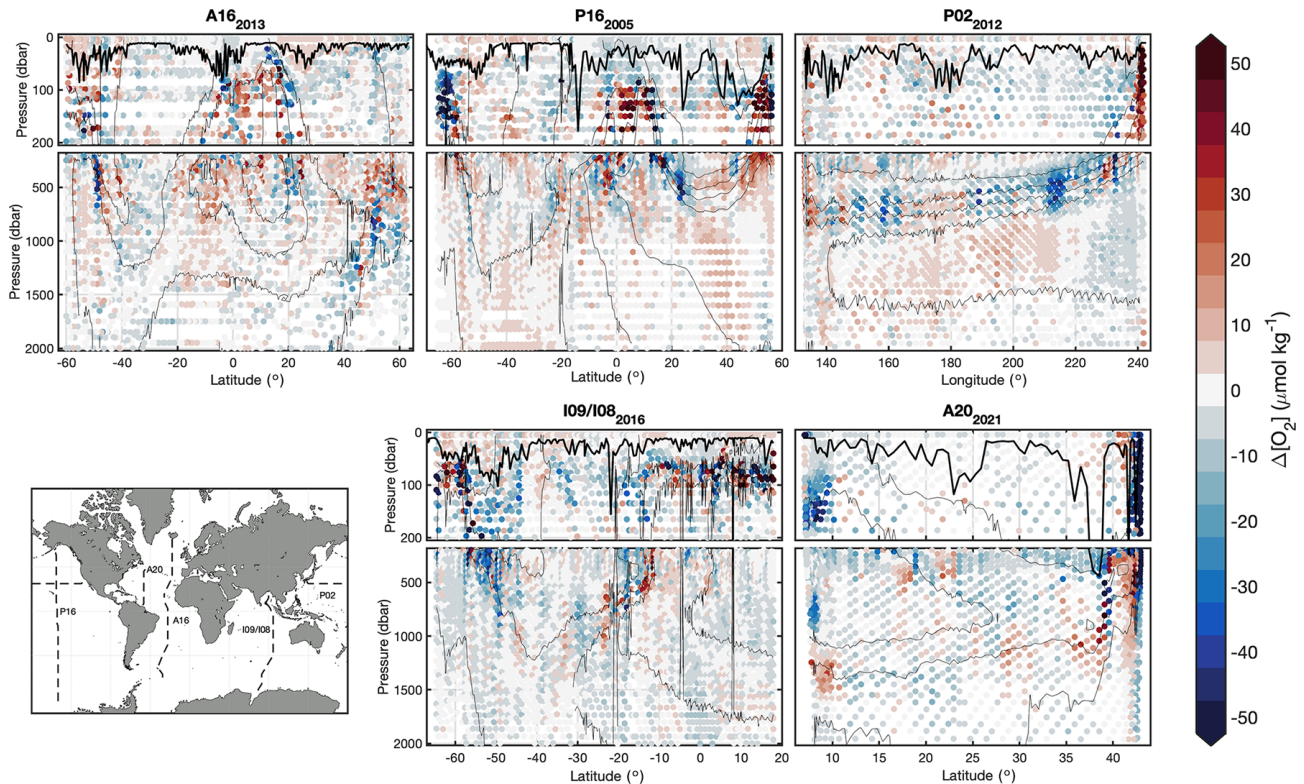


Figure 11. Section plots displaying comparisons between discrete observations of $[O_2]$ from repeat hydrography cruises and $[O_2]$ extracted from corresponding grid cells in GOBAI-O₂. The thick lines in each panel represent the mixed-layer depth calculated as the depth at which the potential density anomaly increased to 0.03 kg m^{-3} greater than the potential density anomaly at 10 dbar. The thin lines are contours representing increments of $50 \mu\text{mol kg}^{-1}$ in $[O_2]$.

Besides these actions, additional steps can be taken to improve GOBAI-O₂ fields. For one, more predictor variables and ML algorithms can be tested. Different processes dominate $[O_2]$ variability in different regions (Keeling et al., 2010; Oschlies et al., 2018; Garcia-Soto et al., 2021), and certain predictor variables will be better suited for capturing these processes. Also, ML algorithms adapt to data sparseness and modes of variability in different ways (Ritter et al., 2017; Gregor et al., 2019), so estimates in a given region that are worse using one algorithm may be better using another. Therefore, regionally tuned predictors and more diverse ensembles of ML algorithms should lead to increased confidence in estimates of ocean interior $[O_2]$. Another action that could result in improved fidelity of GOBAI-O₂ fields is the use of predictor variable fields with higher spatial and temporal resolution across sharp biogeochemical gradients. Ocean profiles of temperature and salinity tend to be relatively smooth, so a depth resolution on the order of tens of meters in the upper ocean increasing to hundreds of meters at depth is sufficient for gridded products. Biogeochemical parameters like oxygen, on the other hand, tend to be characterized by profiles with sharp gradients and with distinct minima and maxima in the water column (Sarmiento and Gruber, 2006). These minima and maxima can occur very near the

surface or hundreds of meters below it. For this reason, comparisons of GOBAI-O₂ to direct measurements of $[O_2]$ can be uniquely problematic in the $\sim 100\text{--}1000$ dbar range when sharp gradients are present (Fig. 11). A complicating factor is the lack of response-time corrections applied to float sensor data (Sect. 2.5), which contributes to uncertainty in observation-based $[O_2]$ products like GOBAI-O₂. Biogeochemical gradients over horizontal space and time can also be sharp, especially in highly dynamic coastal zones and in the surface ocean where the residence time of oxygen is often less than a month (Luz and Barkan, 2000). Recent work from Lyman and Johnson (2023) uses Argo observations coupled with machine learning to provide well-resolved ($7 \text{ d} \times 1/4^\circ$ grid) ocean heat content maps, and continued development toward maps of temperature and salinity could be helpful for overcoming the issue of resolving sharp biogeochemical gradients. Alternatively, $[O_2]$ estimates could be made using temperature and salinity observations at their original resolution and then mapped onto four-dimensional grids that are uniquely suited in their spatial resolution for biogeochemical parameters. A necessary consideration of the latter option would be computing resources: applying complex ML algorithms to temperature and salinity measurements from Argo floats at their original resolution may prove to be impractic-

cal. Finally, observations from additional platforms could be incorporated into approaches like this one to map [O₂] in the global ocean. Ocean gliders and moored profilers have long been equipped with oxygen optodes. These platforms collect data at unique spatiotemporal scales and could add predictive information for [O₂] that is not provided by Argo float observations nor discrete shipboard measurements. To facilitate the incorporation of new data streams into the development of gridded data products, accessible databases should be created and maintained (Testor et al., 2019; Grégoire et al., 2021).

The method used to develop GOBAI-O₂ can be applied in a similar way to other ocean chemical parameters. In addition to dissolved oxygen, the BGC Argo program has deployed floats with sensors for measuring dissolved nitrate, pH, chlorophyll-*a*, particle backscatter, and downwelling irradiance. Machine learning methods have been used to develop four-dimensional fields of optical properties, i.e., chlorophyll-*a* and particle backscatter (Sauzède et al., 2015, 2016), and continued refinement of those fields is ongoing (Sauzède et al., 2021). Chemical properties, i.e., nitrate and pH, that exhibit distributions more similar to [O₂] are good candidates for adoption into the GOBAI mapping approach. Together with property estimation algorithms for total alkalinity (Bittig et al., 2018b; Carter et al., 2021), a mapped ocean interior pH product could be used to resolve the entire ocean carbonate system in four dimensions in near-real time.

Ultimately, global changes in the amount of dissolved oxygen in ocean waters will have profound effects on the metabolism of marine organisms (Pörtner and Farrell, 2008; Sampaio et al., 2021) and the cycling of biogeochemically important elements (Gruber, 2004; Berman-Frank et al., 2008). Whereas ocean models agree that the ocean's oxygen inventory has been declining and will continue to decline, disagreement remains as to regional patterns of this deoxygenation. Direct observations are critical for the confirmation or contradiction of model trends. With this work, we have turned to autonomous and discrete observations, with the assistance of machine learning algorithms, to bridge the model–observational gap. We produce and analyze a multiyear gridded product of ocean dissolved oxygen called GOBAI-O₂, independently confirming a phenomenon that has been demonstrated previously: the ocean is losing dissolved oxygen at a rapid rate (0.79 ± 0.04 % per decade in the upper 2 km according to GOBAI-O₂). In addition, we provide this valuable observation-based product for community use. GOBAI-O₂ can be turned to as a reference for [O₂] observations and model boundary conditions, compared with new and existing observational and model-based reconstructions of ocean deoxygenation, and used for critical analyses of seasonal to decadal and regional to global oxygen variability.

Appendix A

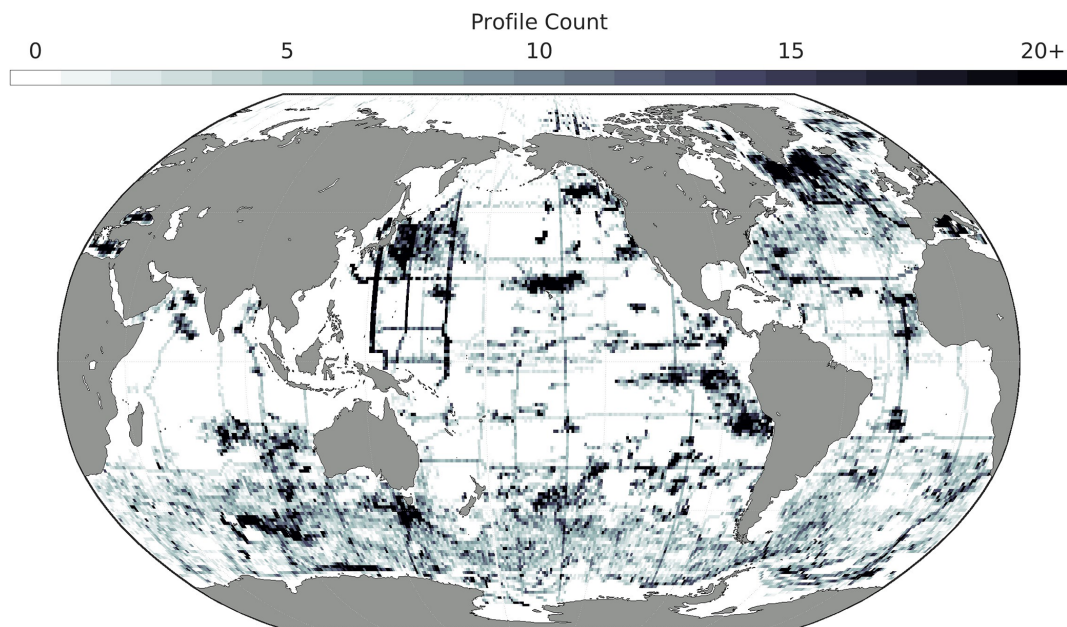


Figure A1. The number of profiles (either ship-based or Argo-float-based) from the combined dataset used to train machine learning algorithms to produce GOBAI-O₂ that are contained within each 1° × 1° box in the global ocean.

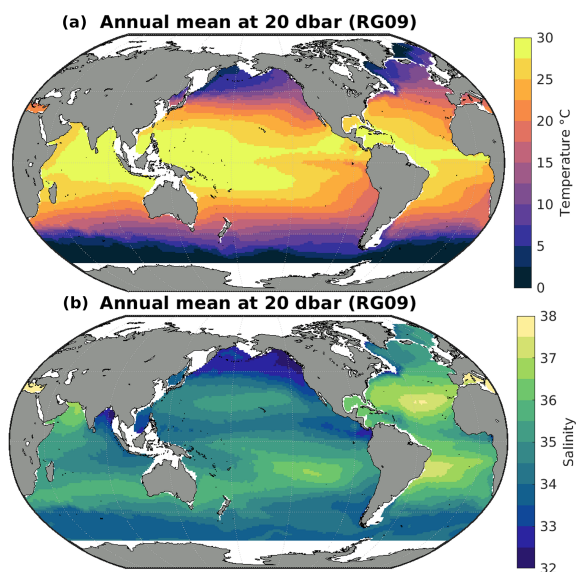


Figure A2. Annual mean in situ (a) temperature and (b) salinity from RG09 (2004–2022) at 20 dbar.

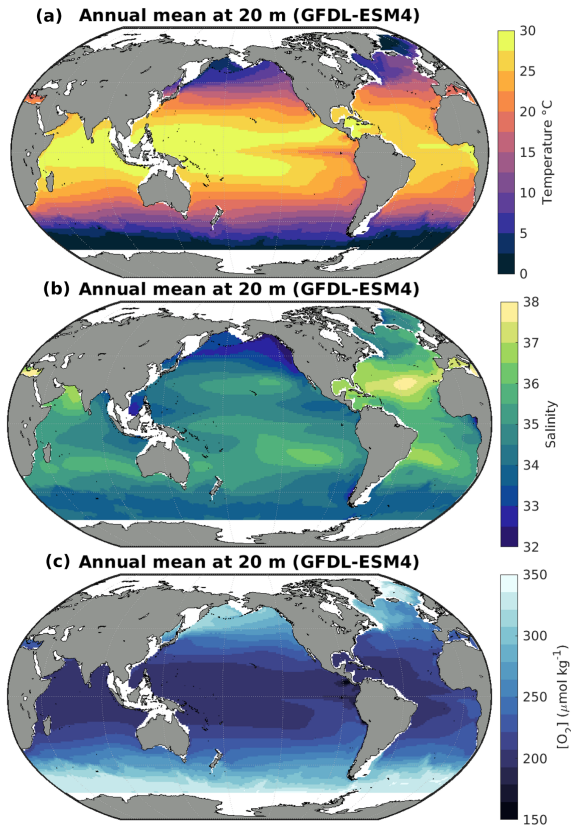


Figure A3. Annual mean in situ (a) temperature, (b) salinity, and (c) dissolved oxygen concentration from GFDL-ESM4 (2004–2021) at 20 m.

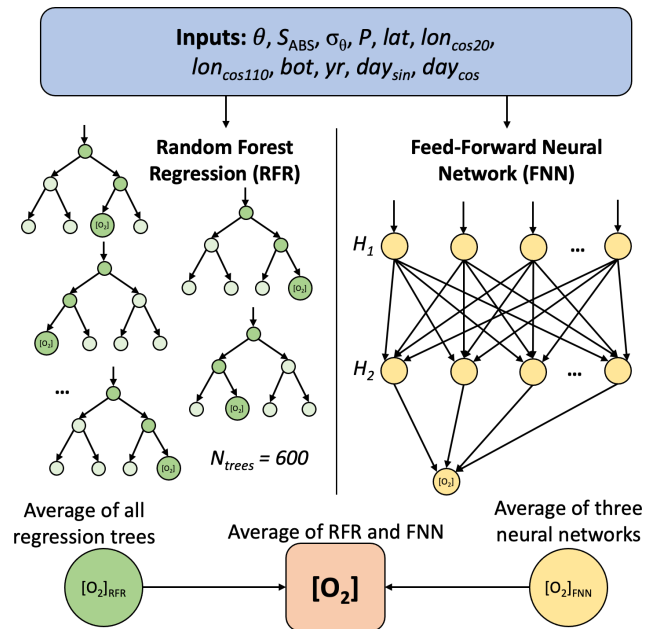


Figure A4. A schematic for the random forest regressions (RFRs) and feed-forward neural networks (FNNs). A random subset of the predictors is used for each tree in the RFR, and a randomly chosen predictor is used for each node split. The two hidden layers (H_1 and H_2) in each of the three FNNs have 10 and 20, 15 and 15, and 20 and 10 nodes. Each machine learning algorithm is trained with input data and $[O_2]$ observations and then used to predict $[O_2]$ from new input data.

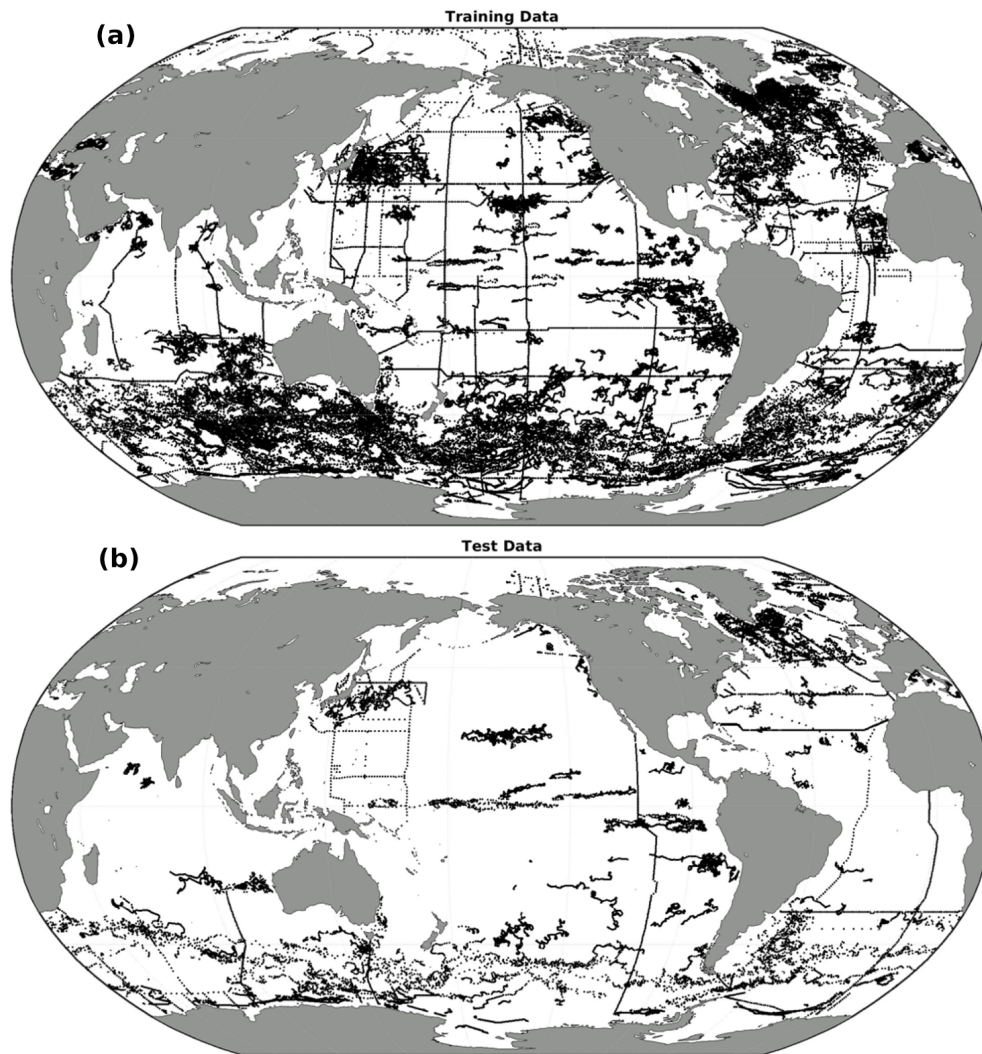


Figure A5. The spatial distribution of profile data used to (a) train and (b) test the $RFR_{Data-Eval}$ and $FNN_{Data-Eval}$ algorithms.

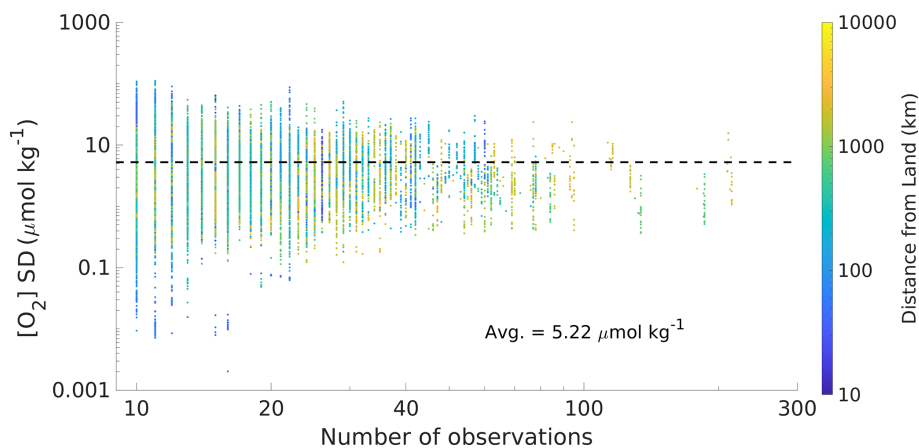


Figure A6. A comparison between the number of observations binned within a four-dimensional grid cell and the standard deviation in $[O_2]$ among those observations. The horizontal black line shows the mean standard deviation ($5.21 \mu\text{mol kg}^{-1}$).

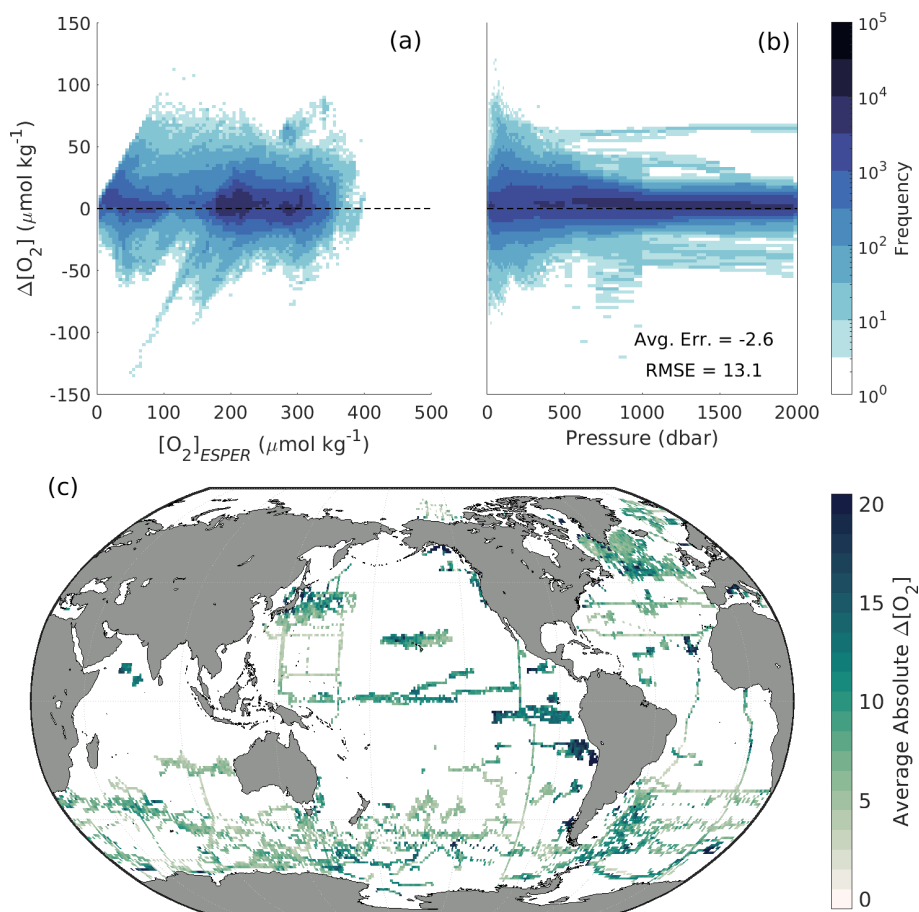


Figure A7. For withheld Argo and GLODAP data, two-dimensional histograms showing offsets between measured and ESPER-Mixed-calculated oxygen ($\Delta[O_2] = [O_2]_{\text{meas}} - [O_2]_{\text{ESPER}}$) as a function of (a) $[O_2]_{\text{ESPER}}$ and (b) pressure in the water column. Offsets are binned into cells that are $2.5 \mu\text{mol kg}^{-1}$ tall in terms of $\Delta[O_2]$ and $5 \mu\text{mol kg}^{-1}$ wide in terms of (a) $[O_2]_{\text{ESPER}}$ or (b) equivalent in width to the interpolated pressure levels of the data. (c) Absolute $\Delta[O_2]$ values averaged over depth and time for $1^\circ \times 1^\circ$ (latitude \times longitude) grid cells in the global ocean.

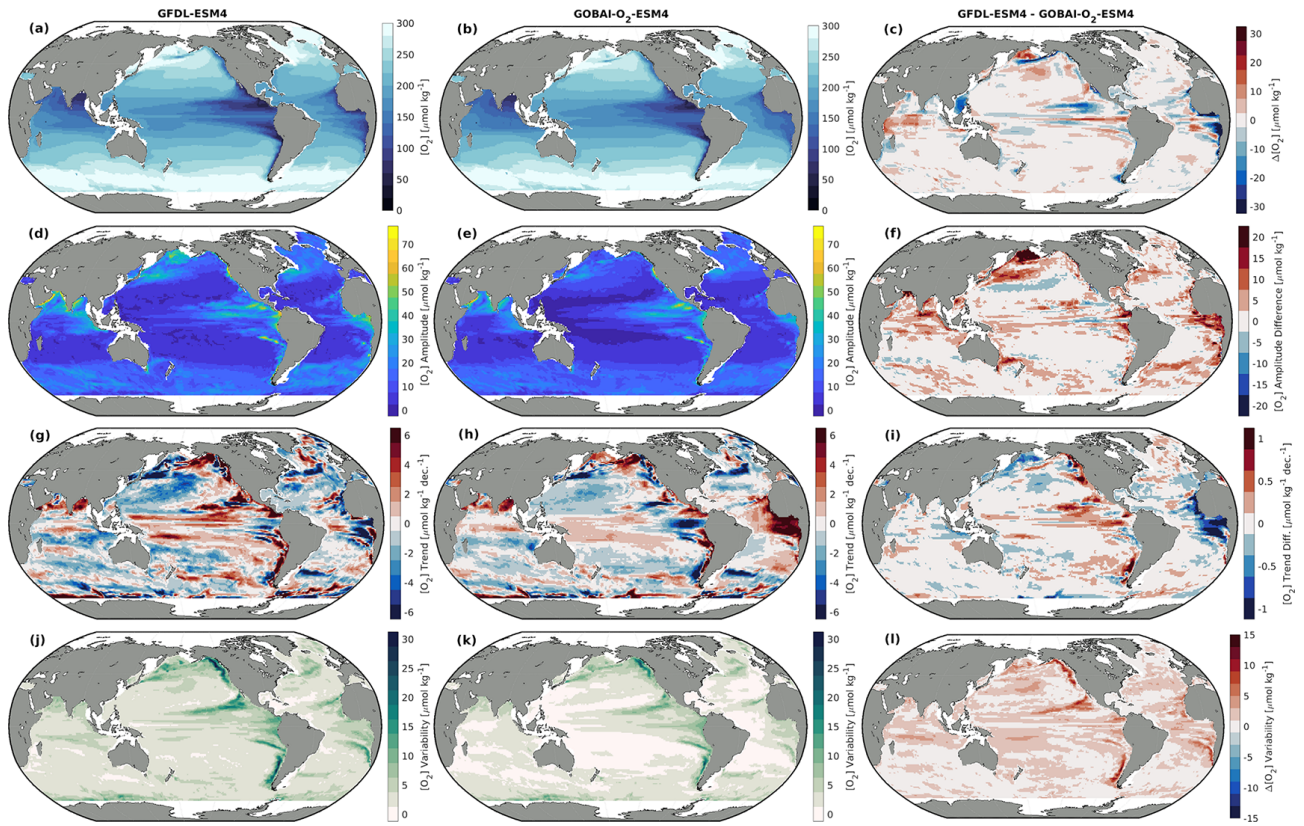


Figure A8. Integrated from 0 to 200 dbar: (a, b) long-term mean [O₂], (d, e) seasonal [O₂] amplitudes, (g, h) trends in [O₂], and (j, k) interannual variability in [O₂] for (a, d, g, j) GFDL-ESM4 and (b, e, h, k) GOBAI-O₂-ESM4, along with (c, f, i, l) the difference between the two.

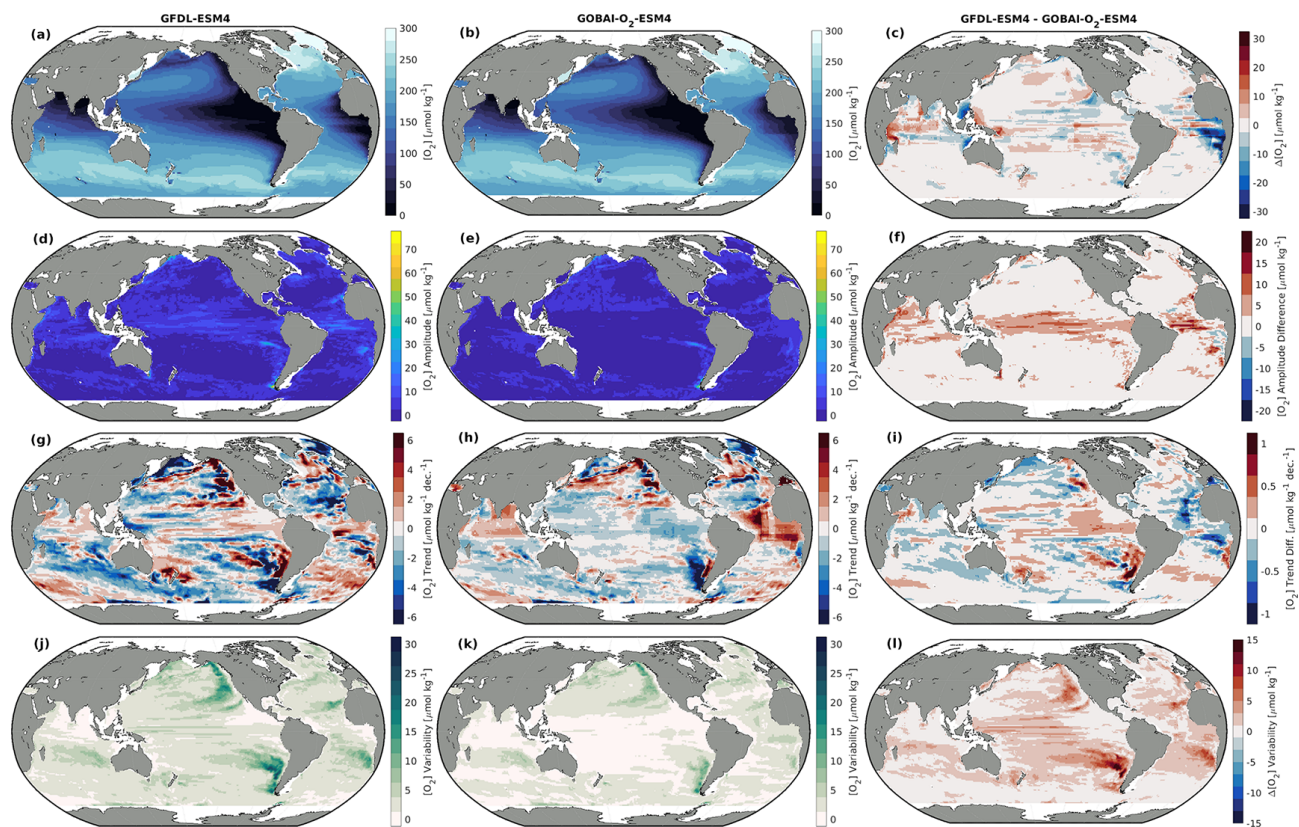


Figure A9. Integrated from 200 to 1000 dbar: (a, b) long-term mean [O₂], (d, e) seasonal [O₂] amplitudes, (g, h) trends in [O₂], and (j, k) interannual variability in [O₂] for (a, d, g, j) GFDL-ESM4 and (b, e, h, k) GOBAI-O₂-ESM4, along with (c, f, i, l) the difference between the two.

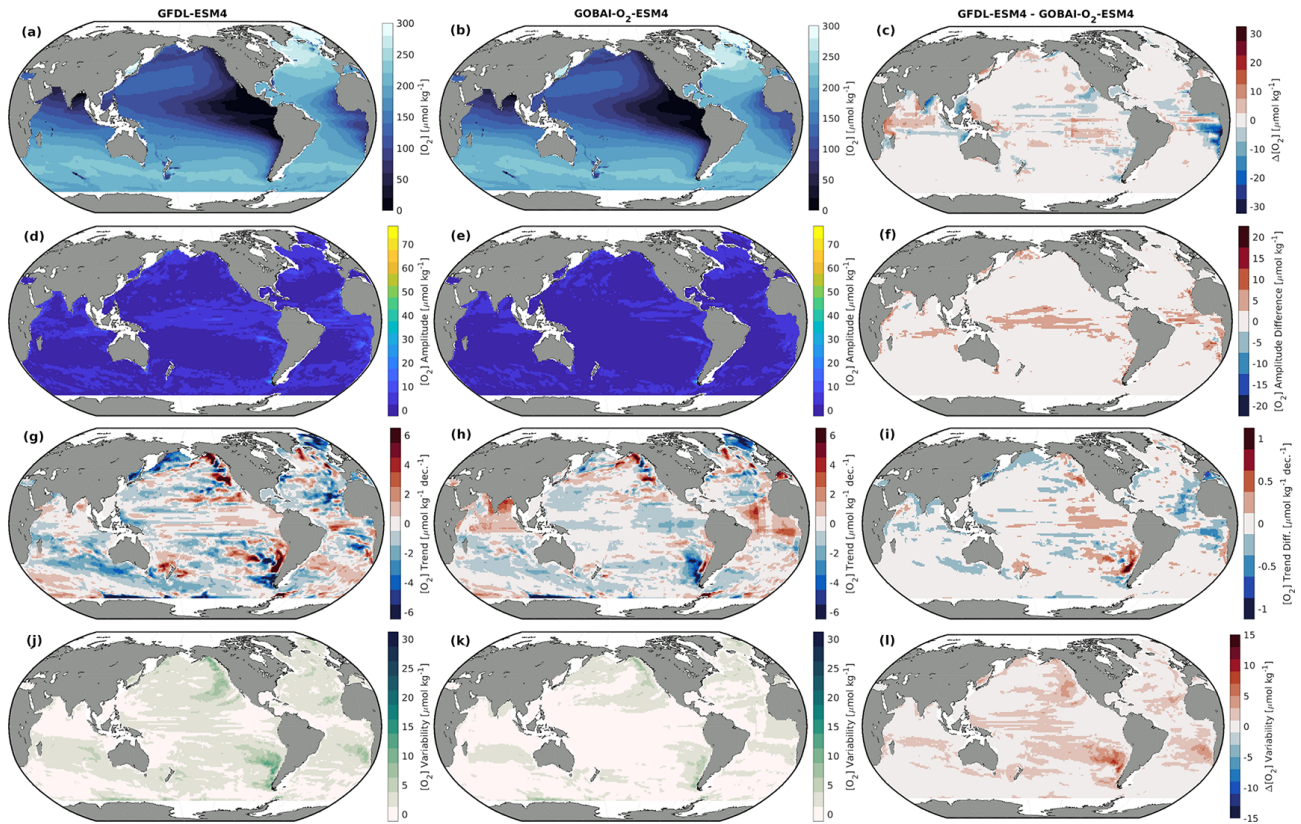


Figure A10. Integrated from 0 to 2000 dbar: (a, b) long-term mean [O₂], (d, e) seasonal [O₂] amplitudes, (g, h) trends in [O₂], and (j, k) interannual variability in [O₂] for (a, d, g, j) GFDL-ESM4 and (b, e, h, k) GOBAI-O₂-ESM4, along with (c, f, i, l) the difference between the two.

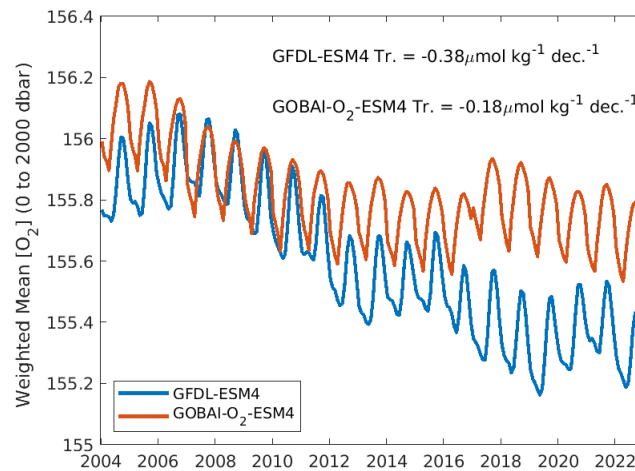


Figure A11. Monthly area-weighted mean [O₂] integrated globally from 0 to 2000 dbar from GFDL-ESM4 (blue) and GOBAI-O₂-ESM4 (orange).

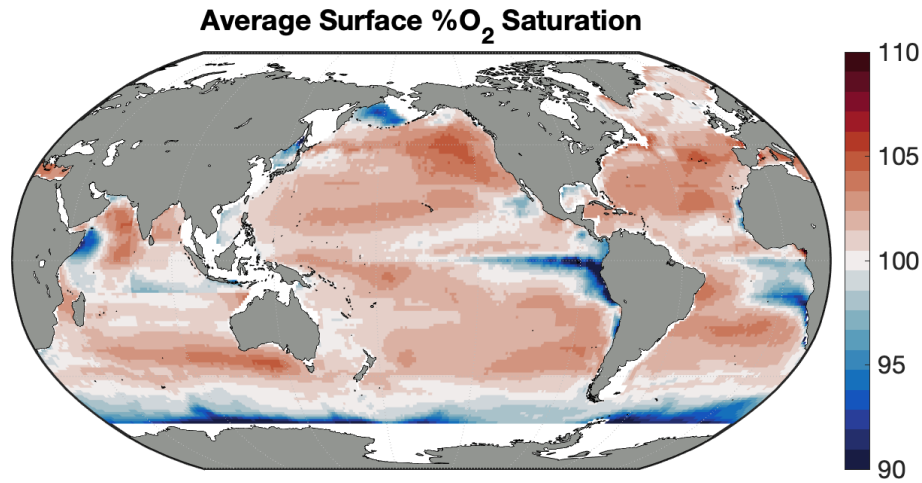


Figure A12. Long-term mean oxygen saturation percentage on the uppermost pressure level in GOBAI-O₂.

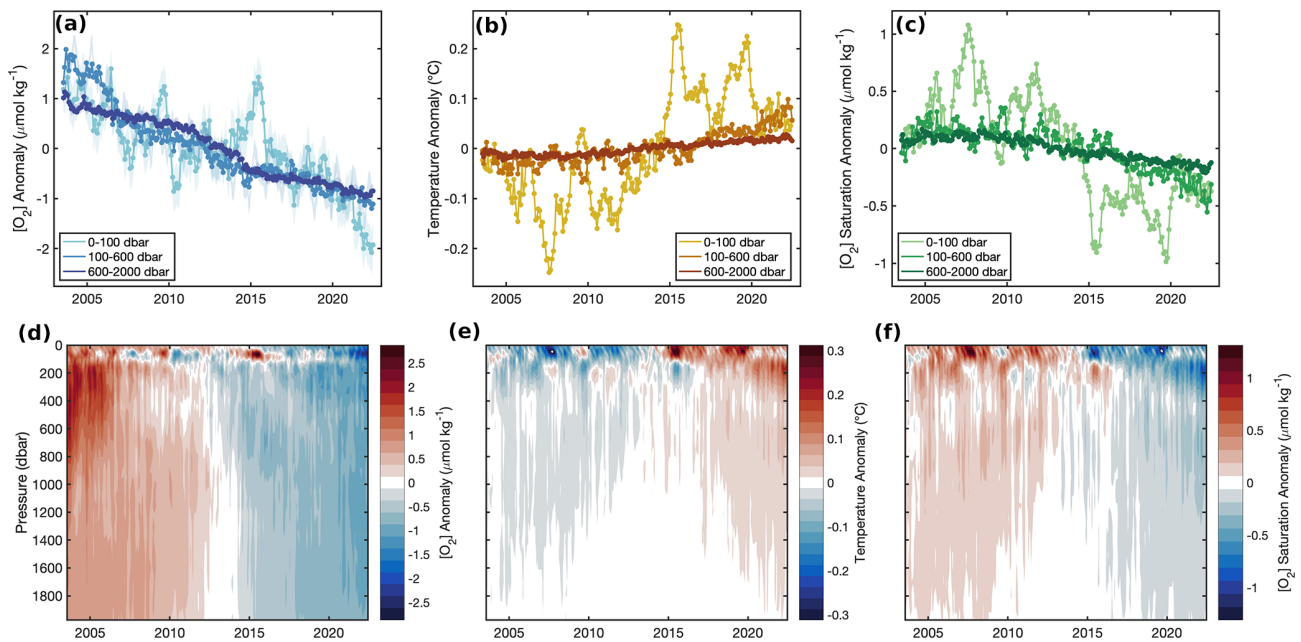


Figure A13. Monthly mean de-seasonalized (a) [O₂] anomalies from GOBAI-O₂, (b) temperature anomalies from RG09, and (c) [O₂]_{sat} anomalies calculated from RG09 temperature and salinity fields, each integrated globally over three pressure layers: 0–100, 100–600, and 600–2000 dbar. In panel (a), shading represents uncertainty determined as the average difference of mean [O₂] from GOBAI-O₂-ESM4 versus GFDL-ESM4 in each layer. Hovmöller diagrams showing monthly mean de-seasonalized (d) [O₂] anomalies from GOBAI-O₂, (e) temperature anomalies from RG09, and (f) [O₂]_{sat} anomalies calculated from RG09 temperature and salinity fields, each over depth in decibars from 2004 to 2022. Anomalies in each parameter are calculated as monthly mean values with a seasonal cycle removed and minus the long-term mean either (a–c) integrated over a pressure layer or (d–f) on a given pressure level.

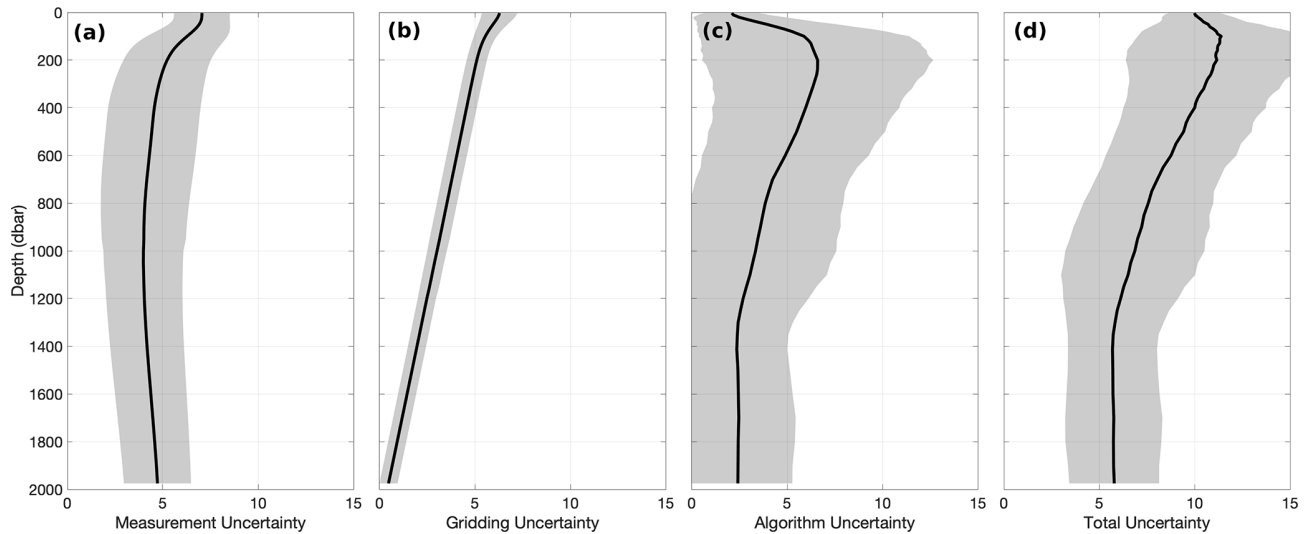


Figure A14. Global mean depth profiles of uncertainty contributors to GOBAI-O₂, including (a) measurement uncertainty, (b) gridding uncertainty, (c) algorithm uncertainty, and (d) total uncertainty. The shaded region represents variability in space and is calculated as the standard deviation on each depth level of the mean uncertainties over time for each grid cell.

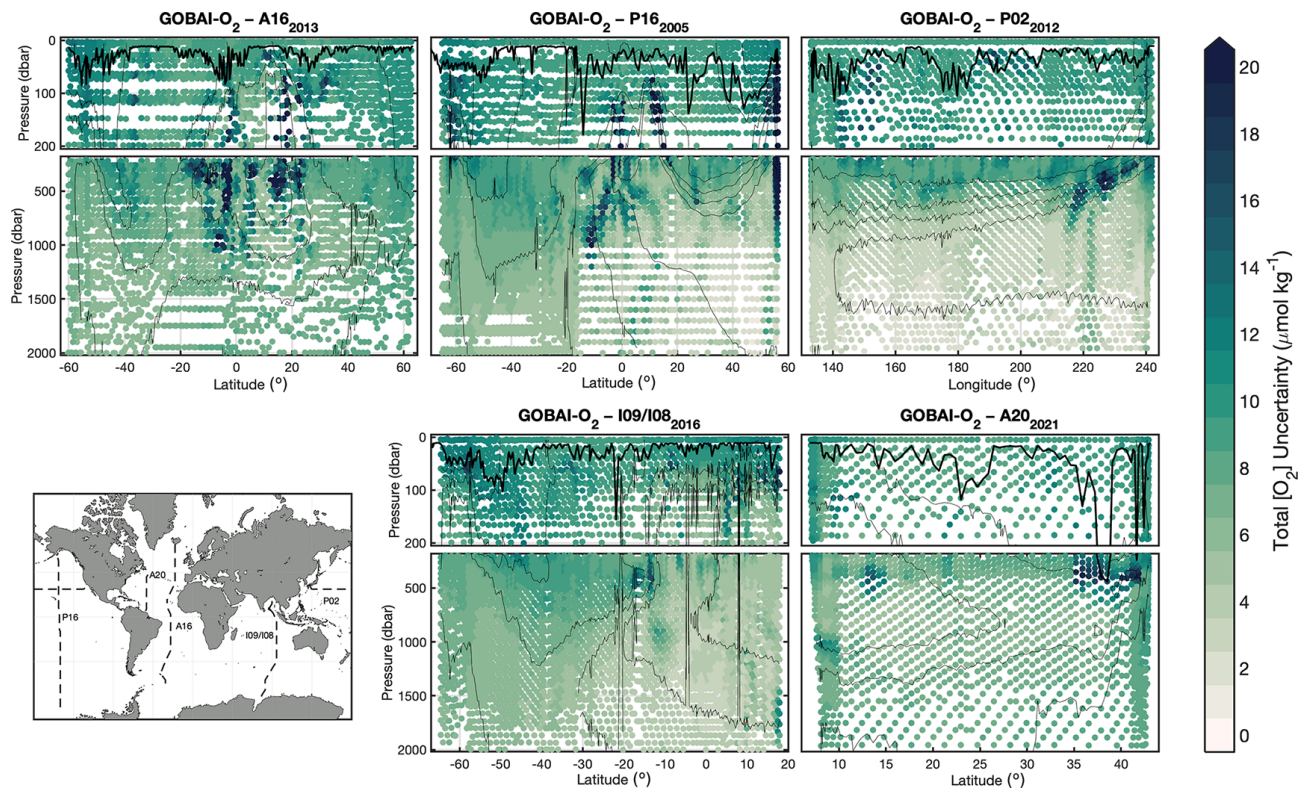


Figure A15. Section plots displaying total uncertainty estimates from GOBAI-O₂ that correspond to discrete measurements of [O₂] from repeat hydrography cruises, to be compared to Δ [O₂] values in Fig. 11.

Appendix B

Table B1. Boundaries for the seven large ocean regions used to fit machine learning algorithms.

Basin	Polygon vertices: [longitude, latitude; ...]
Atl.	[-60, 0; -79, 9.4; -81, 8.4; -100, 22; -100, 45; -6, 45; -6, 35; 4, 15; 25, 0; 22, -35; -68, -35; -60, 0]
Pac.*	[104, 0; 104, 70; 181, 70; 181, 0; 181, -35; 145, -35; 131, -30; 131, 0; 104, 0] [-180, 0; -180, 70; -150, 70; -150, 67; -120, 67; -100, 22; -81, 8.4; -79, 9.4; -60, 0; -68, -35; -180, -35; -180, 0]
Ind.	[22, -35; 25, 10; 38, 35; 104, 35; 104, 0; 131, 0; 131, -30; 116, -35; 22, -35]
Arc.	[-180, 64; -180, 90; 181, 90; 181, 67; 90, 67; 0, 50; 0, 40; -6, 40; -6, 35; -90, 35; -120, 64; -180, 64]
Med.	[-6.5, 40; 0, 40; 0, 45; 20, 47; 38, 35; 34, 30; -5, 30; -6.5, 40]
N. Sou.	[-180, -60; -180, -25; 181, -25; 181, -60; -180, -60]
S. Sou.	[-180, -90; -180, -50; 181, -50; 181, -90; -180, -90]

* Two sets of boundaries are given for the Pacific to accommodate crossing the International Date Line.

Table B2. Error statistics (mean $\Delta[\text{O}_2]$ and RMSD) for tests using RFR and FNN algorithms trained on a subset of Argo and GLODAP data and tested with a separate subset of withheld data. Also shown are error statistics corresponding to the ensemble average (ENS) of the estimates from both algorithms.

Evaluation exercise with observational data								
Basin	Training data points	Assessment data points	RFR _{Data-Eval}		FNN _{Data-Eval}		ENS _{Data-Eval}	
			Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)	Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)	Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)
Atl.	592 099	109 134	-2.0	10.9	-1.5	8.9	-1.8	9.2
Pac.	1 816 367	466 788	0.5	10.0	0.8	10.9	0.6	9.9
Ind.	335 768	82 491	0.3	7.1	1.5	7.9	0.9	7.1
Arc.	800 328	263 873	-1.1	9.1	-1.7	9.6	-1.4	9.1
Med.	214 540	33 899	3.8	10.8	2.6	12.0	3.2	11.0
N. Sou.	2 236 153	480 846	0.1	7.3	0.4	7.6	0.2	7.2
S. Sou.	1 430 492	364 133	-0.6	8.2	-0.8	8.6	-0.7	8.2
All	7 425 747	1 801 164	-0.2	9.0	-0.2	9.5	-0.2	8.8

Table B3. Error statistics (mean $\Delta[\text{O}_2]$ and RMSD) for tests using RFR and FNN algorithms trained on a subset of output from GFDL-ESM4 (corresponding to locations of available Argo and GLODAP data) and tested using a separate subset of withheld output from GFDL-ESM4. Also shown are error statistics corresponding to the ensemble average (ENS) of the estimates from both algorithms.

Evaluation exercise with GFDL-ESM4 output								
Basin	Training data points	Assessment data points	RFR _{ESM4-Eval}		FNN _{ESM4-Eval}		ENS _{ESM4-Eval}	
			Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)	Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)	Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)
Atl.	184 418	28 235 064	−2.1	12.8	−0.6	9.5	−1.3	9.6
Pac.	533 208	69 369 456	0.1	7.5	−0.1	8.5	0.0	7.4
Ind.	86 060	20 736 144	0.7	8.8	−0.3	7.3	0.2	7.2
Arc.	293 540	11 547 744	0.1	4.1	−0.1	4.6	0.0	4.1
Med.	32 110	1 096 680	0.8	4.8	1.3	7.5	1.0	5.5
N. Sou.	756 444	67 626 624	−0.1	4.4	−0.1	4.9	−0.1	4.4
S. Sou.	519 610	31 412 472	0.1	3.4	−0.1	3.6	0.0	3.3
All	2 405 390	230 024 184	−0.2	7.7	−0.2	7.3	−0.2	6.7

Table B4. Error statistics (mean $\Delta[\text{O}_2]$ and RMSD) for tests using RFR and FNN algorithms trained on a subset of Argo and GLODAP data and tested with all available GLODAP data. Also shown are error statistics corresponding to the ensemble average (ENS) of the estimates from both algorithms and corresponding to the ESPER-Mixed model (Carter et al., 2021).

Evaluation exercise with observational data (tested with GLODAP data only)										
Basin	Training data points	Assessment data points	RFR _{Data-Eval}		FNN _{Data-Eval}		ENS _{Data-Eval}		ESPER-Mixed	
			Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)	Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)	Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)	Mean $\Delta[\text{O}_2]$ ($\mu\text{mol kg}^{-1}$)	RMSD ($\mu\text{mol kg}^{-1}$)
Atl.	592 099	180 374	0.0	7.9	0.5	9.4	0.2	7.9	0.1	11.3
Pac.	1 816 367	495 035	0.6	6.1	1.2	9.3	0.9	7.2	−0.3	11.0
Ind.	335 768	42 460	0.6	4.2	1.1	7.2	0.9	5.4	−1.3	9.2
Arc.	800 328	227 905	0.0	5.5	−0.1	8.7	0.0	6.7	1.2	11.0
Med.	214 540	60	−3.6	8.0	0.0	4.3	−1.8	5.3	−5.5	7.7
N. Sou.	2 236 153	174 368	0.7	4.9	0.9	7.2	0.8	5.7	−0.7	8.4
S. Sou.	1 430 492	141 065	0.4	5.3	−0.2	7.9	0.1	6.2	−0.3	9.5
All	7 425 747	1 261 267	0.4	6.1	0.7	8.9	0.6	6.9	0.0	10.7

Table B5. Estimated decadal trends and uncertainties in [O₂] ($\mu\text{mol kg}^{-1}$ per decade) and oxygen inventory (% per decade) in different pressure layers of GOBAI-O₂. Uncertainties are determined according to the procedure in Appendix E, both using the autocorrelation of residuals to the linear least-squares model (Autocov.) and by incorporating estimated uncertainty in global mean GOBAI-O₂ fields ($u(\text{ESM4})$). The value used to represent uncertainty in each trend (larger value) is in bold.

Pressure layer	[O ₂] trend ($\mu\text{mol kg}^{-1}$ per decade)	Trend uncertainty ($\mu\text{mol kg}^{-1}$ per decade)		O ₂ inventory trend (% per decade)	Trend uncertainty (% per decade)	
		Autocov.	$u(\text{ESM4})$		Autocov.	$u(\text{ESM4})$
		0–100 dbar	–1.10		0.60	0.53
100–600 dbar	–1.38	0.42	0.36	–0.86	0.26	0.22
600–2000 dbar	–1.12	0.45	0.54	–0.79	0.32	0.38
0–1000 dbar	–1.28	0.14	0.18	–0.82	0.09	0.11
0–2000 dbar	–1.19	0.03	0.05	–0.79	0.02	0.04

Table B6. Summary error statistics between direct observations from repeat hydrography cruises and GOBAI-O₂ and WOA18.

Cruise	GOBAI-O ₂		WOA18	
	Mean $\Delta[\text{O}_2]$	RMSD	Mean $\Delta[\text{O}_2]$	RMSD
A16 (2013)	–0.2	9.2	0.2	12.0
P16 (2005)	0.1	14.9	0.2	14.5
P02 (2012)	–1.3	9.9	–0.4	12.9
I08/I09 (2016)	–2.3	10.9	–1.1	13.0
A20 (2021)	–7.8	23.4	–2.2	21.4

Appendix C

As stated above, the original and vertically interpolated observational datasets from the BGC Argo and GLO-DAP databases that are used to develop GOBAI-O₂ are archived online: <https://doi.org/10.5281/zenodo.7747237> (Sharp, 2023b). The algorithms trained on vertically interpolated observational data that were applied to predictor variables to produce GOBAI-O₂ are also archived online: <https://doi.org/10.5281/zenodo.7747308> (Sharp, 2023a).

Appendix D

A negative median bias ($-1.18 \mu\text{mol kg}^{-1}$) in float [O₂] measurements compared with co-located ship [O₂] measurements (below 300 dbar, to avoid the impact of high-frequency variability near the surface) was adjusted by fitting the differences ($\Delta[\text{O}_2]$) to a linear least-squares model as a function of float [O₂] and then adding that [O₂]-dependent adjustment back on to the float [O₂] measurements. The $\Delta[\text{O}_2]$ values as a function of float [O₂] before (Fig. D1a) and after (Fig. D1b) this adjustment are shown. This resulted in a reduced median $\Delta[\text{O}_2]$ of $0.33 \mu\text{mol kg}^{-1}$.

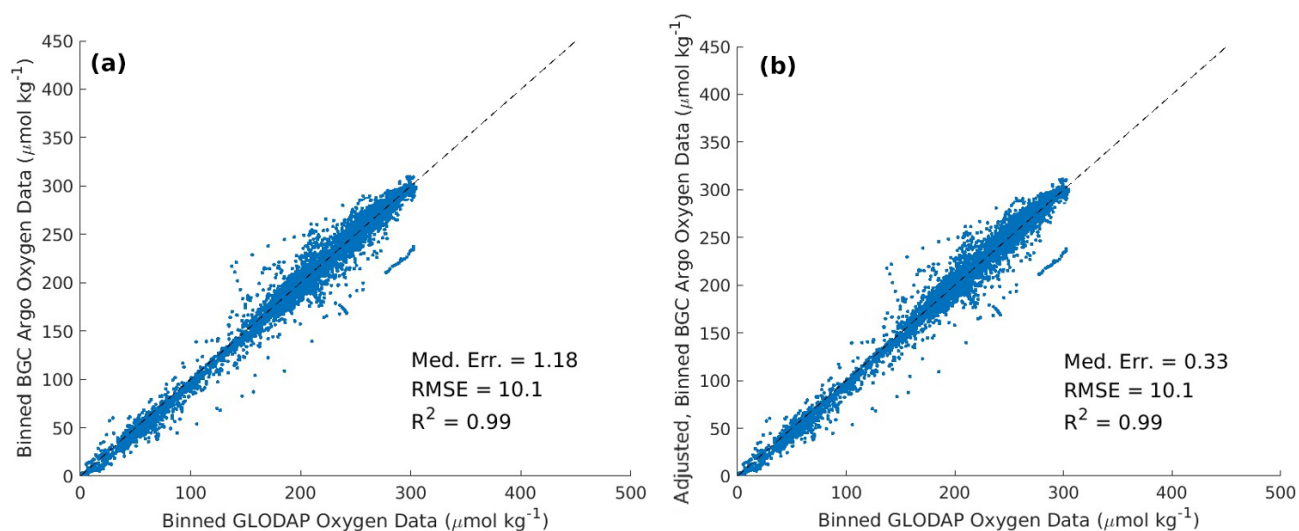


Figure D1. Unadjusted (a) and adjusted (b) matchups between BGC Argo [O₂] measurements (y axis) and GLODAP [O₂] measurements (x axis). The adjustment procedure does not mitigate the scatter between the matchups, but it does reduce the median error.

Appendix E

Trends and associated uncertainties in GOBAI-O₂ were determined via the four-step procedure outlined below.

1. Spatially weighted monthly mean [O₂] values for the entire GOBAI-O₂ domain or within specified pressure layers were calculated from gridded [O₂], using relative grid-cell volumes as weights.
2. A linear least-squares model with a trend, intercept, and annual (12-month) and semiannual (6-month) harmonics was fit to monthly mean [O₂] values. The monthly trend from the least-squares model was multiplied by 120 to obtain a decadal trend of weighted mean [O₂].
3. Uncertainty in the decadal trends was assessed in two different ways, and the largest of the two uncertainty estimates taken for each analyzed pressure layer, indicating that either (a) uncertainty in the linear least-squares model or (b) uncertainty in the GOBAI-O₂ fields was driving uncertainty in the trend. In all cases, the uncertainty estimate from the second method was chosen, indicating that decadal trend uncertainty was controlled primarily by uncertainty in the GOBAI-O₂ fields. The two different methods used to assess uncertainty are outlined below.
 - a. Uncertainty was assessed using the autocovariance of residuals from the linear least-squares model via the following procedure:
 - i. the standard error on the trend was calculated from the covariance matrix of the linear least-squares model;
 - ii. the autocovariance of the residuals from the least-squares model was examined to compute the *e*-folding timescale, and the effective degrees of freedom were obtained by dividing the number of monthly mean [O₂] values by the *e*-folding timescale and subtracting the number of least-squares model parameters; and
 - iii. the standard error on the trend was scaled by the effective degrees of freedom, multiplied by 2 to obtain a 95 % confidence interval, and multiplied by 120 to obtain an uncertainty on the decadal trend of weighted mean [O₂].
 - b. Uncertainty was assessed by incorporating estimated uncertainty in global mean GOBAI-O₂ fields via the following procedure:
 - i. uncertainties in monthly mean [O₂] values were determined as the standard deviations of monthly differences between GOBAI-O₂-ESM4 and GFDL-ESM4 (Sect. 3.1.2);
 - ii. these uncertainties were used to compute a weight matrix for the linear least-squares fit,

and the effective degrees of freedom were obtained as previously described; and

- iii. the standard error on the trend was scaled by the effective degrees of freedom, multiplied by 2 to obtain a 95 % confidence interval, and multiplied by 120 to obtain an uncertainty on the decadal trend of weighted mean [O₂].
4. The process was repeated for oxygen inventories for the entire GOBAI-O₂ domain or within each specified pressure layer; inventories were determined from gridded [O₂], volumes of each grid cell, and densities of each grid cell.

Author contributions. JDS, AJF, BRC, and GCJ conceptualized and planned the project. JDS produced the data product, conducted the analysis, created the data visualizations, and wrote the original draft of the manuscript. JDS, AJF, BRC, GCJ, JPD, and CS reviewed and edited the manuscript.

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

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. The authors thank the international Argo program and the national programs that contribute to it (<https://argo.ucsd.edu>, last access: 22 September 2023, <https://www.ocean-ops.org>, last access: 22 September 2023); the Argo program is part of the Global Ocean Observing System. The authors also thank the many scientists who have contributed to the GLODAPv2 database (<https://www.glodap.info>, last access: 22 September 2023) by securing funding, dedicating their time to collecting and sharing data, and assembling and quality-controlling those data. Jonathan D. Sharp and Brendan R. Carter were funded by the National Oceanic and Atmospheric Administration (NOAA) Climate Program Office's Climate Observations and Monitoring (COM) and Climate Variability and Predictability (CVP) programs, in partnership with NOAA's Global Ocean Monitoring and Observing (GOMO) program, through award no. NA21OAR4310251. Cristina Schultz and Gregory C. Johnson were funded, and Brendan R. Carter and Jonathan D. Sharp were partially funded, by the Ocean Observing and Monitoring Division (now GOMO) for the establishment and use of a biogeochemical Argo array within the California Current system. Jonathan D. Sharp and Brendan R. Carter were funded through the Cooperative Institute for Climate, Ocean, & Ecosystem Studies (CIOCES) under NOAA cooperative agreement no. NA20OAR4320271. Jonathan D. Sharp, Andrea J. Fassbender, and Gregory C. Johnson were supported by NOAA's Pacific Marine Environmental Laboratory (PMEL). The authors are grateful to Hartmut Frenzel (CICOES) for computational assistance. Hernan Garcia and two anonymous reviewers are also acknowledged

for providing comments that prompted significant improvements to this manuscript. This is CICOES contribution no. 2022-1226 and PMEL contribution no. 5416.

Financial support. This research has been supported by the Global Ocean Monitoring and Observing Program (grant no. NA21OAR4310251), the Global Ocean Monitoring and Observing Program (grant no. 100007298), and the Cooperative Institute for Climate, Ocean, and Ecosystem Studies (grant no. NA20OAR4320271).

Review statement. This paper was edited by Anton Velo and reviewed by Hernan Garcia and two anonymous referees.

References

- Alkire, M. B., D'Asaro, E., Lee, C., Perry, M. J., Gray, A., Cetinić, I., Briggs, N., Rehm, E., Kallin, E., Kaiser, J., and González-Posada, A.: Estimates of net community production and export using high-resolution, Lagrangian measurements of O₂, NO₃⁻, and POC through the evolution of a spring diatom bloom in the North Atlantic, *Deep-Sea Res. Pt. I*, 64, 157–174, <https://doi.org/10.1016/j.dsr.2012.01.012>, 2012.
- Argo: Argo float data and metadata from Global Data Assembly Centre (Argo GDAC), SEANOE [data set], <https://doi.org/10.17882/42182>, 2000.
- Arteaga, L. A., Pahlow, M., Bushinsky, S. M., and Sarmiento, J. L.: Nutrient controls on export production in the Southern Ocean, *Global Biogeochem. Cy.*, 33, 942–956, <https://doi.org/10.1029/2019GB006236>, 2019.
- Berman-Frank, I., Chen, Y. B., Gao, Y., Fennel, K., Follows, M. J., Milligan, A. J., and Falkowski, P. G.: Feedbacks between the nitrogen, carbon and oxygen cycles, in: *Nitrogen in the Marine Environment*, Elsevier Inc., 1539–1563, ISBN 978-0-12-372522-6, 2008.
- Bindoff, N. L., Cheung, W. W. L., Kairo, J. G., Arístegui, J., Guinder, V. A., Hallberg, R., Hilmi, N., Jiao, N., Karim, M. S., Levin, L., O'Donoghue, S., Purca Cuicapusa, S. R., Rinkevich, B., Suga, T., Tagliabue, A., and Williamson, P.: Changing ocean, marine ecosystems, and dependent communities, in: *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, edited by: Pörtner, H.-O., Roberts, D. C., Masson-Delmotte, V., Zhai, P., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegría, A., Nicolai, M., Okem, A., Petzold, J., Rama, B., and Weyer, N. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, 447–588, <https://doi.org/10.1017/9781009157964>, 2019.
- Bittig, H., Wong, A., and Plant, J.: BGC-Argo synthetic profile file processing and format on Coriolis GDAC, Ifremer, Plouzane, France, <https://doi.org/10.13155/55637>, 2022.
- Bittig, H. C. and Körtzinger, A.: Tackling oxygen optode drift: Near-surface and in-air oxygen optode measurements on a float provide an accurate in situ reference, *J. Atmos. Ocean. Tech.*, 32, 1536–1543, <https://doi.org/10.1175/JTECH-D-14-00162.1>, 2015.
- Bittig, H. C. and Körtzinger, A.: Technical note: Update on response times, in-air measurements, and in situ drift for oxygen optodes on profiling platforms, *Ocean Sci.*, 13, 1–11, <https://doi.org/10.5194/os-13-1-2017>, 2017.
- Bittig, H. C., Fiedler, B., Scholz, R., Krahnemann, G., and Körtzinger, A.: Time response of oxygen optodes on profiling platforms and its dependence on flow speed and temperature, *Limnol. Oceanogr.-Meth.*, 12, 617–636, <https://doi.org/10.4319/lom.2014.12.617>, 2014.
- Bittig, H. C., Körtzinger, A., Neill, C., van Ooijen, E., Plant, J. N., Hahn, J., Johnson, K. S., Yang, B., and Emerson, S. R.: Oxygen optode sensors: Principle, characterization, calibration, and application in the ocean, *Front. Mar. Sci.*, 4, 429, <https://doi.org/10.3389/fmars.2017.00429>, 2018a.
- Bittig, H. C., Steinhoff, T., Claustre, H., Fiedler, B., Williams, N. L., Sauzède, R., Körtzinger, A., and Gattuso, J. P.: An alternative to static climatologies: Robust estimation of open ocean CO₂ variables and nutrient concentrations from T, S, and O₂ data using Bayesian neural networks, *Front. Mar. Sci.*, 5, 328, <https://doi.org/10.3389/fmars.2018.00328>, 2018b.
- Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J., and Vichi, M.: 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.
- Boyer, T. P., Baranova, O. K., Coleman, C., Garcia, H. E., Grodsky, A., Locarnini, R. A., Mishonov, A. V., Paver, C. R., Reagan, J. R., Seidov, D., Smolyar, I. V., Weathers, K., and Zweng, M. M.: *World Ocean Database 2018*, edited by: Mishonov, A. V. (technical ed.), NOAA Atlas NESDIS 87, 2018.
- Breiman, L.: Random forests, *Mach. Learn.*, 45, 5–32, <https://doi.org/10.1023/A:1010933404324>, 2001.
- Breitbart, D., Levin, L. A., Oschlies, A., Grégoire, M., Chavez, F. P., Conley, D. J., Garçon, V., Gilbert, D., Gutiérrez, 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., and Zhang, J.: Declining oxygen in the global ocean and coastal waters, *Science*, 359, eaam7240, <https://doi.org/10.1126/science.aam7240>, 2018.
- Bushinsky, S. M. and Emerson, S.: Marine biological production from in situ oxygen measurements on a profiling float in the sub-arctic Pacific Ocean, *Global Biogeochem. Cy.*, 29, 2050–2060, <https://doi.org/10.1002/2015GB005251>, 2015.
- Bushinsky, S. M. and Emerson, S. R.: Biological and physical controls on the oxygen cycle in the Kuroshio Extension from an array of profiling floats, *Deep-Sea Res. Pt. I*, 141, 51–70, <https://doi.org/10.1016/j.dsr.2018.09.005>, 2018.
- Bushinsky, S. M., Emerson, S. R., Riser, S. C., and Swift, D. D.: Accurate oxygen measurements on modified Argo floats using in situ air calibrations, *Limnol. Oceanogr.-Meth.*, 14, 491–505, <https://doi.org/10.1002/lom3.10107>, 2016.
- Bushinsky, S. M., Gray, A. R., Johnson, K. S., and Sarmiento, J. L.: Oxygen in the Southern Ocean from Argo floats: Determination of processes driving air-sea fluxes, *J. Geophys. Res.-Oceans*, 122, 8661–8682, <https://doi.org/10.1002/2017JC012923>, 2017.
- Carpenter, J. H.: The Chesapeake Bay Institute technique for the Winkler dissolved oxygen method, *Limnol. Oceanogr.*, 10, 141–143, <https://doi.org/10.4319/lo.1965.10.1.0141>, 1965.
- Carter, B. R., Bittig, H. C., Fassbender, A. J., Sharp, J. D., Takeshita, Y., Xu, Y. Y., Álvarez, M., Wanninkhof, R., Feely, R. A., and

- Barbero, L.: New and updated global empirical seawater property estimation routines, *Limnol. Oceanogr.-Meth.*, 19, 785–809, <https://doi.org/10.1002/lom3.10461>, 2021.
- Chierici, M., Fransson, A., Turner, D. R., Pakhomov, E. A., and Froneman, P. W.: Variability in pH, $f\text{CO}_2$, oxygen and flux of CO_2 in the surface water along a transect in the Atlantic sector of the Southern Ocean, *Deep-Sea Res. Pt. II*, 51, 2773–2787, <https://doi.org/10.1016/j.dsr2.2001.03.002>, 2004.
- Claustre, H., Johnson, K. S., and Takeshita, Y.: Observing the global ocean with biogeochemical-Argo, *Annu. Rev. Mar. Sci.*, 12, 23–48, <https://doi.org/10.1146/annurev-marine-010419-010956>, 2020.
- D'Asaro, E. A. and McNeil, C.: Calibration and stability of oxygen sensors on autonomous floats, *J. Atmos. Ocean. Tech.*, 30, 1896–1906, <https://doi.org/10.1175/JTECH-D-12-00222.1>, 2013.
- Demuth, H., Beale, M., and Hagan, M.: *Neural Network Toolbox 6 User's Guide*, The MathWorks, Inc., Natick, MA, https://www.mathworks.com/help/pdf_doc/deeplearning/nnet Ug.pdf (last access: 28 September 2023), 2008.
- Djeutchouang, L. M., Chang, N., Gregor, L., Vichi, M., and Monteiro, P. M. S.: The sensitivity of $p\text{CO}_2$ reconstructions to sampling scales across a Southern Ocean sub-domain: a semi-idealized ocean sampling simulation approach, *Biogeosciences*, 19, 4171–4195, <https://doi.org/10.5194/bg-19-4171-2022>, 2022.
- Drucker, R. and Riser, S. C.: In situ phase-domain calibration of oxygen Optodes on profiling floats, *Methods in Oceanography*, 17, 296–318, <https://doi.org/10.1016/j.mio.2016.09.007>, 2016.
- Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. A., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M.: The GFDL Earth System Model version 4.1 (GFDL-ESM4.1): Model description and simulation characteristics, *J. Adv. Model. Earth Sy.*, 12, e2019MS002015, <https://doi.org/10.1029/2019MS002015>, 2020.
- Estapa, M. L., Feen, M. L., and Breves, E.: Direct observations of biological carbon export from profiling floats in the subtropical North Atlantic, *Global Biogeochem. Cy.*, 33, 282–300, <https://doi.org/10.1029/2018GB006098>, 2019.
- 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.
- Fiedler, B., Fietzek, P., Vieira, N., Silva, P., Bittig, H. C., and Körtzinger, A.: In Situ CO_2 and O_2 Measurements on a Profiling Float, *J. Atmos. Ocean. Tech.*, 30, 112–126, <https://doi.org/10.1175/JTECH-D-12-00043.1>, 2013.
- Frenzel, H., Sharp, J. D., Fassbender, A. J., and Buzby, N.: OneArgo-Mat: A MATLAB toolbox for accessing and visualizing Argo data (v1.0.2), Zenodo [data set], <https://doi.org/10.5281/zenodo.6588041>, 2022.
- Frölicher, T. L., Joos, F., Plattner, G. K., Steinacher, M., and Doney, S. C.: Natural variability and anthropogenic trends in oceanic oxygen in a coupled carbon cycle–climate model ensemble, *Global Biogeochem. Cy.*, 23, GB1003, <https://doi.org/10.1029/2008GB003316>, 2009.
- Garcia, H. E. and Gordon, L. I.: Oxygen solubility in seawater: Better fitting equations, *Limnol. Oceanogr.*, 37, 1307–1312, <https://doi.org/10.4319/lo.1992.37.6.1307>, 1992.
- Garcia, H. E., Weathers, K., Paver, C. R., Smolyar, I., Boyer, T. P., Locarnini, R. A., Zweng, M. M., Mishonov, A. V., Baranova, O. K., Seidov, D., Reagan, J. R.: *World Ocean Atlas 2018*, vol. 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Dissolved Oxygen Saturation, NOAA Atlas NESDIS 83, <https://archimer.ifremer.fr/doc/00651/76337> (last access: 1 October 2021), 2019 (data available at: <https://www.ncei.noaa.gov/archive/accession/NCEI-WOA18>, last access: 1 October 2021).
- Garcia-Soto, C., Cheng, L., Caesar, L., Schmidtko, S., Jewett, E. B., Cheripka, A., Rigor, I., Caballero, A., Chiba, S., Báez, J. C., Zielinski, T., and Abraham, J. P.: An Overview of Ocean Climate Change Indicators: Sea Surface Temperature, Ocean Heat Content, Ocean pH, Dissolved Oxygen Concentration, Arctic Sea Ice Extent, Thickness and Volume, Sea Level and Strength of the AMOC (Atlantic Meridional Overturning Circulation), *Front. Mar. Sci.*, 8, 642372, <https://doi.org/10.3389/fmars.2021.642372>, 2021.
- Giglio, D., Lyubchich, V., and Mazloff, M. R.: Estimating oxygen in the Southern Ocean using Argo temperature and salinity, *J. Geophys. Res.-Oceans*, 123, 4280–4297, <https://doi.org/10.1029/2017JC013404>, 2018.
- Gregor, L. and Gruber, N.: OceanSODA-ETHZ: a global gridded data set of the surface ocean carbonate system for seasonal to decadal studies of ocean acidification, *Earth Syst. Sci. Data*, 13, 777–808, <https://doi.org/10.5194/essd-13-777-2021>, 2021.
- Gregor, L., Kok, S., and Monteiro, P. M. S.: Empirical methods for the estimation of Southern Ocean CO_2 : support vector and random forest regression, *Biogeosciences*, 14, 5551–5569, <https://doi.org/10.5194/bg-14-5551-2017>, 2017.
- Gregor, L., Lebehot, A. D., Kok, S., and Scheel Monteiro, P. M.: A comparative assessment of the uncertainties of global surface ocean CO_2 estimates using a machine-learning ensemble (CSIR-ML6 version 2019a) – have we hit the wall?, *Geosci. Model Dev.*, 12, 5113–5136, <https://doi.org/10.5194/gmd-12-5113-2019>, 2019.
- Grégoire, M., Garçon, 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., Mehrrens, 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, 724913, <https://doi.org/10.3389/fmars.2021.724913>, 2021.
- Gruber, N.: The dynamics of the marine nitrogen cycle and its influence on atmospheric CO_2 variations, in: *The ocean*

- carbon cycle and climate, Springer, Dordrecht, 97–148, <https://doi.org/10.1007/978-1-4020-2087-2>, 2004.
- Gruber, N., Gloor, M., Fan, S.-M., and Sarmiento, J. L.: Air-sea flux of oxygen estimated from bulk data: Implications For the marine and atmospheric oxygen cycles, *Global Biogeochem. Cy.*, 15, 783–803, <https://doi.org/10.1029/2000GB001302>, 2001.
- Gruber, N., Doney, S., Emerson, S., Gilbert, D., Kobayashi, T., Körtzinger, A., Johnson, G., Johnson, K., Riser, S., and Ulloa, O.: Adding oxygen to Argo: Developing a global in-situ observatory for ocean deoxygenation and biogeochemistry, in: Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society, Venice, Italy, 21–25 September 2009, ESA Publication WPP-306, <https://doi.org/10.5270/OceanObs09.cwp.39>, 2010.
- Helm, K. P., Bindoff, N. L., and Church, J. A.: Observed decreases in oxygen content of the global ocean, *Geophys. Res. Lett.*, 38, 1–6, <https://doi.org/10.1029/2011GL049513>, 2011.
- Ito, T., Minobe, S., Long, M. C., and Deutsch, C.: Upper ocean O₂ trends: 1958–2015, *Geophys. Res. Lett.*, 44, 4214–4223, <https://doi.org/10.1002/2017GL073613>, 2017.
- Johnson, G. C., Hosoda, S., Jayne, S. R., Oke, P. R., Riser, S. C., Roemmich, D., Suga, T., Thierry, V., Wijffels, S., and Xu, J.: Argo – Two decades: Global oceanography, revolutionized, *Annu. Rev. Mar. Sci.*, 14, 379–403, <https://doi.org/10.1146/annurev-marine-022521-102008>, 2022.
- Johnson, K. and Claustre, H.: Bringing biogeochemistry into the Argo age, *Eos*, 97, <https://doi.org/10.1029/2016EO062427>, 2016.
- Johnson, K. S. and Bif, M. B.: Constraint on net primary productivity of the global ocean by Argo oxygen measurements, *Nat. Geosci.*, 14, 769–774, <https://doi.org/10.1038/s41561-021-00807-z>, 2021.
- Johnson, K. S., Plant, J. N., Riser, S. C., and Gilbert, D.: Air oxygen calibration of oxygen optodes on a profiling float array, *J. Atmos. Ocean. Tech.*, 32, 2160–2172, <https://doi.org/10.1175/JTECH-D-15-0101.1>, 2015.
- Johnson, K. S., Plant, J. N., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Riser, S. C., Swift, D. D., Williams, N. L., Boss, E., Haëntjens, N., Talley, L. D., and Sarmiento, J. L.: Biogeochemical sensor performance in the SOCCOM profiling float array, *J. Geophys. Res.-Oceans*, 122, 6416–6436, <https://doi.org/10.1002/2017JC012838>, 2017.
- Johnson, K. S., Riser, S. C., and Ravichandran, M.: Oxygen variability controls denitrification in the Bay of Bengal oxygen minimum zone, *Geophys. Res. Lett.*, 46, 804–811, <https://doi.org/10.1029/2018GL079881>, 2019.
- Jonsson, B. F., Doney, S. C., Dunne, J., and Bender, M.: Evaluation of the Southern Ocean O₂/Ar-based NCP estimates in a model framework, *J. Geophys. Res.-Biogeo.*, 118, 385–399, <https://doi.org/10.1002/jgrg.20032>, 2013.
- Keeling, R. F., Arne, K., and Gruber, N.: Ocean Deoxygenation in a Warming World, *Annu. Rev. Mar. Sci.*, 2, 199–229, <https://doi.org/10.1146/annurev.marine.010908.163855>, 2010.
- Keppler, L., Landschützer, P., Gruber, N., Lauvset, S. K., and Stemmler, I.: Seasonal carbon dynamics in the near-global ocean, *Global Biogeochem. Cy.*, 34, e2020GB006571, <https://doi.org/10.1029/2020GB006571>, 2020.
- Keppler, L., Landschützer, P., Lauvset, S. K., and Gruber, N.: Recent trends and variability in the oceanic storage of dissolved inorganic carbon, *Global Biogeochem. Cy.*, 37, e2022GB007677, <https://doi.org/10.1029/2022GB007677>, 2023.
- Key, R. M., Olsen, A., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F. F., and Suzuki, T.: Global Ocean Data Analysis Project, Version 2 (GLODAPv2), ORNL/CDIAC-162, NDP-093, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee, <https://epic.awi.de/id/eprint/39306/> (last access: 28 September 2023), 2015.
- Körtzinger, A., Schimanski, J., Send, U., and Wallace, D.: The ocean takes a deep breath, *Science*, 306, 1337, <https://doi.org/10.1126/science.1102557>, 2004.
- Körtzinger, A., Schimanski, J., and Send, U.: High Quality Oxygen Measurements from Profiling Floats: A Promising New Technique, *J. Atmos. Ocean. Tech.*, 22, 302–308, <https://doi.org/10.1175/JTECH1701.1>, 2005.
- Krasting, J. P., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., Rand, K., Zadeh, N.T., Balaji, V., Durachta, J., Dupuis, C., Menzel, R., Robinson, T., Underwood, S., Vahlenkamp, H., Dunne, K. A., Gauthier, P. G., Ginoux, P., Griffies, S. M., Hallberg, R., Harrison, M., Hurlin, W., Malyshev, S., Naik, V., Paulot, F., Paynter, D. J., Ploshay, J., Reichl, B. G., Schwarzkopf, D. M., Seman, C. J., Silvers, L., Wyman, B., Zeng, Y., Adcroft, A., Dunne, J. P., Dussin, R., Guo, H., He, J., Held, I. H., Horowitz, L. W., Lin, P., Milly, P. C. D., Shevliakova, E., Stock, C., Winton, M., Wittenberg, A. T., Xie, Y., and Zhao, M.: NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP, Version 20190726, Earth System Grid Federation [data set], <https://doi.org/10.22033/ESGF/CMIP6.1407>, 2018.
- Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., Dunne, J. P., Gehlen, M., Ilyina, T., John, J. G., Lenton, A., Li, H., Lovenduski, N. S., Orr, J. C., Palmieri, J., Santana-Falcón, Y., Schwinger, J., Séférian, R., Stock, C. A., Tagliabue, A., Takano, Y., Tjiputra, J., Toyama, K., Tsujino, H., Watanabe, M., Yamamoto, A., Yool, A., and Ziehn, T.: Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections, *Biogeosciences*, 17, 3439–3470, <https://doi.org/10.5194/bg-17-3439-2020>, 2020.
- Landschützer, P., Gruber, N., Bakker, D. C. E., and Schuster, U.: Recent variability of the global ocean carbon sink, *Global Biogeochem. Cy.*, 28, 927–949, <https://doi.org/10.1002/2014GB004853>, 2014.
- Langdon, C.: Determination of Dissolved Oxygen in Seawater by Winkler Titration using Amperometric Technique, in: The GO-SHIP Repeat Hydrography Manual: A Collection of Expert Reports and Guidelines. Version 1, edited by: Hood, E. M., Sabine, C. L., and Sloyan, B. M., IOCCP Report no. 14, ICPO Publication Series no. 134, 18 pp., <https://doi.org/10.25607/OBP-1350>, 2010.
- Lauvset, S. K., Key, R. M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F. F., Suzuki, T., and Watelet, S.: A new global interior ocean mapped climatology: the 1° × 1° GLODAP version 2, *Earth Syst. Sci. Data*, 8, 325–340, <https://doi.org/10.5194/essd-8-325-2016>, 2016.

- Lauvset, S. K., Lange, N., Tanhua, T., Bittig, H. C., Olsen, A., Kozyr, A., Alin, S., Álvarez, M., Azetsu-Scott, K., Barbero, L., Becker, S., Brown, P. J., Carter, B. R., da Cunha, L. C., Feely, R. A., Hoppema, M., Humphreys, M. P., Ishii, M., Jeansson, E., Jiang, L.-Q., Jones, S. D., Lo Monaco, C., Murata, A., Müller, J. D., Pérez, F. F., Pfeil, B., Schirnack, C., Steinfeldt, R., Suzuki, T., Tilbrook, B., Ulfso, A., Velo, A., Woosley, R. J., and Key, R. M.: GLODAPv2.2022: the latest version of the global interior ocean biogeochemical data product, *Earth Syst. Sci. Data*, 14, 5543–5572, <https://doi.org/10.5194/essd-14-5543-2022>, 2022a.
- Lauvset, S. K., Lange, N., Tanhua, T., Bittig, H. C., Olsen, A., Kozyr, A., Alin, S. R., Álvarez, M., Azetsu-Scott, K., Barbero, L., Becker, S., Brown, P. J., Carter, B. R., Cotrim da Cunha, L., Feely, R. A., Hoppema, M., Humphreys, M. P., Ishii, M., Jeansson, E., Jiang, L.-Q., Jones, S. D., Lo Monaco, C., Murata, A., Müller, J. D., Pérez, F. F., Pfeil, B., Schirnack, C., Steinfeldt, R., Suzuki, T., Tilbrook, B., Ulfso, A., Velo, A., Woosley, R. J., and Key, R. M.: Global Ocean Data Analysis Project version 2.2022 (GLODAPv2.2022) (NCEI Accession 0257247), NOAA National Centers for Environmental Information [data set], <https://doi.org/10.25921/1f4w-0t92>, 2022b.
- Limburg, K. E., Breitburg, D., Swaney, D. P., and Jacinto, G.: Ocean deoxygenation: A primer, *One Earth*, 2, 24–29, <https://doi.org/10.1016/j.oneear.2020.01.001>, 2020.
- Luz, B. and Barkan, E.: Assessment of oceanic productivity with the triple-isotope composition of dissolved oxygen, *Science*, 288, 2028–2031, 2000.
- Lyman, J. M. and Johnson, G. C.: Global High-Resolution Random Forest Regression Maps of Ocean Heat Content Anomalies Using in Situ and Satellite Data, *J. Atmos. Ocean. Tech.*, 40, 575–586, <https://doi.org/10.1175/JTECH-D-22-0058.1>, 2023.
- Maurer, T. L., Plant, J. N., and Johnson, K. S.: Delayed-Mode Quality Control of Oxygen, Nitrate, and pH Data on SOCCOM Biogeochemical Profiling Floats, *Front. Mar. Sci.*, 8, 683207, <https://doi.org/10.3389/fmars.2021.683207>, 2021.
- McDougall, T. J. and Barker, P. M.: Getting started with TEOS-10 and the Gibbs Seawater (GSW) oceanographic toolbox, *Scor/Iapso WG*, 127, 1–28, ISBN 978-0-646-55621-5, 2011.
- Nicholson, D. P. and Feen, M. L.: Air calibration of an oxygen optode on an underwater glider, *Limnol. Oceanogr.-Meth.*, 15, 495–502, <https://doi.org/10.1002/lom3.10177>, 2017.
- NOAA National Geophysical Data Center: 2 min Gridded Global Relief Data (ETOPO2) v2, NOAA National Centers for Environmental Information, <https://doi.org/10.7289/V5J1012Q>, 2006.
- Olsen, A., Key, R. M., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., Schirnack, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Pérez, F. F., and Suzuki, T.: The Global Ocean Data Analysis Project version 2 (GLODAPv2) – an internally consistent data product for the world ocean, *Earth Syst. Sci. Data*, 8, 297–323, <https://doi.org/10.5194/essd-8-297-2016>, 2016.
- Oschlies, A., Duteil, O., Getzlaff, J., Koeve, W., Landolfi, A., and Schmidtko, S.: Patterns of deoxygenation: sensitivity to natural and anthropogenic drivers, *Philos. T. Roy. Soc. A*, 375, 20160325, <https://doi.org/10.1098/rsta.2016.0325>, 2017.
- 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.
- Piron, A., Thierry, V., Mercier, H., and Caniaux, G.: Argo float observations of basin-scale deep convection in the Irminger sea during winter 2011–2012, *Deep-Sea Res. Pt. I*, 109, 76–90, <https://doi.org/10.1016/j.dsr.2015.12.012>, 2016.
- Piron, A., Thierry, V., Mercier, H., and Caniaux, G.: Gyre-scale deep convection in the subpolar North Atlantic Ocean during winter 2014–2015, *Geophys. Res. Lett.*, 44, 1439–1447, <https://doi.org/10.1002/2016GL071895>, 2017.
- Pörtner, H. O. and Farrell, A. P.: Physiology and climate change, *Science*, 322, 690–692, <https://doi.org/10.1126/science.1163156>, 2008.
- Reuer, M. K., Barnett, B. A., Bender, M. L., Falkowski, P. G., and Hendricks, M. B.: New estimates of Southern Ocean biological production rates from O₂/Ar ratios and the triple isotope composition of O₂, *Deep-Sea Res. Pt. I*, 54, 951–974, <https://doi.org/10.1016/j.dsr.2007.02.007>, 2007.
- Ritter, R., Landschützer, P., Gruber, N., Fay, A. R., Iida, Y., Jones, S., and Zeng, J.: Observation-Based Trends of the Southern Ocean Carbon Sink, *Geophys. Res. Lett.*, 2, 339–348, <https://doi.org/10.1002/2017GL074837>, 2017.
- Roemmich, D. and Gilson, J.: The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program, *Prog. Oceanogr.*, 82, 81–100, <https://doi.org/10.1016/j.pocean.2009.03.004>, 2009.
- 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., Calil, P., Carse, F., Carval, T., Chai, F., Conchubhair, D. Ó., d’Ortenzio, F., Dall’Omo, G., Desbruyeres, D., Fennel, K., Fer, I., Ferrari, R., Forget, G., Freeland, H., Fujiki, T., Gehlen, M., Greenan, B., Hallberg, R., Hibiye, T., Hosoda, S., Jayne, S., Jochum, M., Johnson, G. C., Kang, K. R., Kolodziejczyk, N., Körzinger, A., Le Traon, P.-Y., Lenn, Y.-D., Maze, G., Mork, K. A., Morris, T., Nagai, T., Nash, J., Naveira Garbato, A., Olsen, A., Pattabhi, R. R., Prakash, S., Riser, S., Schmechtig, C., Schmid, C., Shroyer, E., Sterl, A., Sutton, P., Talley, L., Tanhua, T., Thierry, V., Thomalla, S., Toole, J., Troisi, A., Trull, T. W., Turton, J., Velez-Belchi, P. J., Walczowski, W., Wang, H., Wanninkhof, R., Waterhouse, A. F., Waterman, S., Watson, A., Wilson, C., Wong, A. P. S., Xu, J., and Yasuda, I.: On the future of Argo: A global, full-depth, multi-disciplinary array, *Frontiers in Marine Science*, 6, 1–28, <https://doi.org/10.3389/fmars.2019.00439>, 2019.
- Roemmich, D., Talley, L., Zilberman, N., Osborne, E., Johnson, K. S., Barbero, L., Bittig, H. C., Briggs, N., Fassbender, A. J., Johnson, G. C., King, B. A., McDonagh, E., Purkey, S., Riser, S., Suga, T., Takeshita, Y., Thierry, V., and Wijffels, S.: The technological, scientific, and sociological revolution of global subsurface ocean observing, *Oceanography*, 34, 2–8, <https://doi.org/10.5670/oceanog.2021.supplement.02-02>, 2021.
- Rydzewski, R. and Dunne, J.: A measured look at ocean chlorophyll trends, *Nature*, 472, E5–E6, <https://doi.org/10.1038/nature09952>, 2011.
- Sampaio, E., Santos, C., Rosa, I. C., Ferreira, V., Pörtner, H., Duarte, C. M., Levin, L. A., and Rosa, R.: Impacts of hypoxic events surpass those of future ocean warming and acidification, *Nature Ecology and Evolution*, 5, 311–321, <https://doi.org/10.1038/s41559-020-01370-3>, 2021.

- Sarma, V. V. S. S. and Udaya Bhaskar, T. V. S.: Ventilation of oxygen to oxygen minimum zone due to anticyclonic eddies in the Bay of Bengal, *J. Geophys. Res.-Biogeosci.*, 123, 2145–2153, <https://doi.org/10.1029/2018JG004447>, 2018.
- Sarmiento, J. L. and Gruber, N.: *Ocean Biogeochemical Dynamics*, Princeton University Press, Princeton, NJ, ISBN 978-0691017075, 2006.
- Sauzède, R., Claustre, H., Jamet, C., Uitz, J., Ras, J., Mignot, A., and d’Ortenzio, F.: Retrieving the vertical distribution of chlorophyll a concentration and phytoplankton community composition from in situ fluorescence profiles: A method based on a neural network with potential for global-scale applications, *J. Geophys. Res.-Oceans*, 120, 451–470, <https://doi.org/10.1002/2014JC010355>, 2015.
- Sauzède, R., Claustre, H., Uitz, J., Jamet, C., Dall’Olmo, G., d’Ortenzio, F., Gentili, B., Poteau, A., and Schmechtig, C.: A neural network based method for merging ocean color and Argo data to extend surface bio optical properties to depth: Retrieval of the particulate backscattering coefficient, *J. Geophys. Res.-Oceans*, 121, 2552–2571, <https://doi.org/10.1002/2015JC011408>, 2016.
- Sauzède, R., Claustre, H., Pannimpullath, R., Uitz, J., and Guinehut, S.: New Global Vertical Distribution of Gridded Particulate Organic Carbon and Chlorophyll Concentration Using Machine Learning for CMEMs, in: 9th EuroGOOS International Conference, Brest, France 313–320, <https://hal.archives-ouvertes.fr/hal-03335370> (last access: 22 September 2023), 2021.
- 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.
- Schofield, O., Fassbender, A., Hood, M., Hill, K., and Johnson, K.: A global ocean biogeochemical observatory becomes a reality, *Eos*, 103, <https://doi.org/10.1029/2022EO220149>, 2022.
- Sharp, J. D.: GOBAI-O₂ Algorithms, Zenodo [data set], <https://doi.org/10.5281/zenodo.7747309>, 2023a.
- Sharp, J. D.: GOBAI-O₂ Training Data, Zenodo [data set], <https://doi.org/10.5281/zenodo.7747237>, 2023b.
- Sharp, J. D., Fassbender, A. J., Carter, B. R., Johnson, G. C., Schultz, C., and Dunne, J. P.: GOBAI-O₂: A Global Gridded Monthly Dataset of Ocean Interior Dissolved Oxygen Concentrations Based on Shipboard and Autonomous Observations (NCEI Accession 0259304). V2.0, NOAA National Centers for Environmental Information [data set], <https://doi.org/10.25921/z72m-yz67>, 2022.
- Stramma, L. and Schmidtko, S.: Global evidence of ocean deoxygenation, in: *Ocean Deoxygenation: Everyone’s Problem*, edited by: Laffoley, D. and Baxter, J. M., IUCN, Gland, Switzerland, 25–26, <https://doi.org/10.2305/IUCN.CH.2019.13.en>, 2019.
- Stramma, L. and Schmidtko, S.: Spatial and Temporal Variability of Oceanic Oxygen Changes and Underlying Trends, *Atmos. Ocean*, 59, 122–132, <https://doi.org/10.1080/07055900.2021.1905601>, 2021.
- Stukel, M. R. and Ducklow, H. W.: Stirring up the biological pump: vertical mixing and carbon export in the Southern Ocean, *Global Biogeochem. Cy.*, 31, 1420–1434, <https://doi.org/10.1002/2017GB005652>, 2017.
- Takeshita, Y., Martz, T. R., Johnson, K. S., Plant, J. N., Gilbert, D., Riser, S. C., Neill, C., and Tilbrook, B.: A climatology-based quality control procedure for profiling float oxygen data, *J. Geophys. Res.-Oceans*, 118, 5640–5650, <https://doi.org/10.1002/jgrc.20399>, 2013.
- Talley, L. D., Pickard, G. L., Emery, W. J., and Swift, J. H.: *Descriptive physical oceanography: an introduction*, Academic Press, ISBN 978-0-7506-4552-2, 2011.
- Testor, P., de Young, B., Rudnick, D. L., Glenn, S., Hayes, D., Lee, C. M., Pattiaratchi, C., Hill, K., Heslop, E., Turpin, V., Alenius, P., Barrera, C., Barth, J. A., Beaird, N., Bécu, G., Bosse, A., Bourrin, F., Brearley, J. A., Chao, Y., Chen, S., Chigiato, J., Coppola, L., Crout, R., Cummings, J., Curry, B., Curry, R., Davis, R., Desai, K., DiMarco, S., Edwards, C., Fielding, S., Fer, I., Frajka-Williams, E., Gildor, H., Goni, G., Gutierrez, D., Haugan, P., Hebert, D., Heiderich, J., Henson, S., Heywood, K., Hogan, P., Houpert, L., Huh, S., E. Inall, M., Ishii, M., Ito, S., Itoh, S., Jan, S., Kaiser, J., Karstensen, J., Kirkpatrick, B., Klymak, J., Kohut, J., Krahnemann, G., Krug, M., McClatchie, S., Marin, F., Mauri, E., Mehra, A., P. Meredith, M., Meunier, T., Miles, T., Morell, J. M., Mortier, L., Nicholson, S., O’Callaghan, J., O’Conchubhair, D., Oke, P., Pallàs-Sanz, E., Palmer, M., Park, J., Perivoliotis, L., Poulain, P.-M., Perry, R., Queste, B., Rainville, L., Rehm, E., Roughan, M., Rome, N., Ross, T., Ruiz, S., Saba, G., Schaeffer, A., Schönau, M., Schroeder, K., Shimizu, Y., Sloyan, B. M., Smeed, D., Snowden, D., Song, Y., Swart, S., Tenreiro, M., Thompson, A., Tintore, J., Todd, R. E., Toro, C., Venables, H., Wagawa, T., Waterman, S., Watlington, R. A., and Wilson, D.: OceanGliders: A component of the integrated GOOS, *Front. Mar. Sci.*, 6, 422, <https://doi.org/10.3389/fmars.2019.00422>, 2019.
- Udaya Bhaskar, T. V. S., Sarma, V. V. S. S., and Pavan Kumar, J.: Potential mechanisms responsible for spatial variability in intensity and thickness of oxygen minimum zone in the Bay of Bengal, *J. Geophys. Res.-Biogeo.*, 126, e2021JG006341, <https://doi.org/10.1029/2021JG006341>, 2021.
- Wang, Z., Garcia, H. E., Boyer, T. P., Reagan, J., and Cebrian, J.: Controlling factors of the climatological annual cycle of the surface mixed layer oxygen content: A global view, *Front. Mar. Sci.*, 9, 1001095, <https://doi.org/10.3389/fmars.2022.1001095>, 2022.
- Winkler, L. W.: Die Bestimmung des im Wasser gelösten Sauerstoffes, *Chem. Ber.*, 21, 2843–2855, <https://doi.org/10.1002/cber.188802102122>, 1888.
- Wolf, M. K., Hamme, R. C., Gilbert, D., Yashayaev, I., and Thierry, V.: Oxygen saturation surrounding deep water formation events in the Labrador Sea from Argo-O₂ data, *Global Biogeochem. Cy.*, 32, 635–653, <https://doi.org/10.1002/2017GB005829>, 2018.
- Yang, B., Emerson, S. R., and Bushinsky, S. M.: Annual net community production in the subtropical Pacific Ocean from in situ oxygen measurements on profiling floats, *Global Biogeochem. Cy.*, 31, 728–744, <https://doi.org/10.1002/2016GB005545>, 2017.