Skill testing oxygen data for distribution modeling of marine species

Julia Indivero[1*], Sean Anderson[2], Lewis Barnett[3], John Pohl[4], Sean Rohan[3,5], Sam Siedlecki[6], Eric Ward[7], and Tim Essington[1]

[1]School of Aquatic and Fisheries Sciences, University of Washington, Seattle, WA, USA

[2]Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, British Columbia, Canada

[3]Resource Assessment and Conservation Engineering Division, Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, Washington, USA

[4]Fishery Resource Analysis and Monitoring Division, Groundfish Ecology Program, Northwest Fishery Science Center, National Oceanic and Atmospheric Administration, Seattle, WA, USA

[5]Sustainable Fisheries Division, Alaska Region, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Juneau, AK, USA

[6]Department of Marine Sciences, University of Connecticut, Groton, CT, USA

[7]Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, Washington, USA

*Corresponding author: Julia.indivero@gmail.com

Data Availability Statement

The data and code are available at https://github.com/jindivero/o2-sdm-paper. They will also be archived at Zenodo upon publication.

Funding Statement

This publication was partially funded by Washington Sea Grant grant no. R/SFA-12, and from the Lowell E. Wakefield Professorship.

Conflict of Interest Disclosure

The authors have no conflicts of interest to declare.

Abstract

Spatial models that identify statistical relationships between environmental conditions and species distributional data are commonly used in fisheries research to evaluate habitat suitability and predict distributional shifts, such as those driven by changing ocean temperature and oxygen. However, a lack of environmental data—particularly dissolved oxygen—at the same temporal and spatial resolution as biological data can limit these analyses. We evaluate the ability to predict bottom dissolved oxygen via imputation and extrapolation and with biophysical oceanographic models in the northeastern Pacific Ocean (Aleutian Islands, Eastern Bering Sea, Gulf of Alaska, British Columbia, and California Current). Specifically, we measure predictive skill compared to *in situ* observations (measured concurrently with bottom trawl data) for 1) predictions from an empirical statistical model fit to integrated dissolved oxygen observations; and 2) a commonly used dynamical oceanographic model estimate of oxygen, the Global Oceanographic Biogeochemistry Hindcast (GOBH). Lastly, we evaluate how estimation and interpretation of a species distribution model is impacted by use of different oxygen data sources. Using leave-one-year-out cross-validation, we find that the empirical statistical model predicts bottom dissolved oxygen for fish catch sampling events with relatively high accuracy in only certain regions (California Current and British Columbia) (root mean squared error [RMSE] ~16-30 μmol kg^{-1}). Prediction skill was more than 2x lower in Alaska regions that did not have extensive data (~<0.075 observations km^{-2}), and this approach would likely not provide sufficiently accurate oxygen values for SDMs in these regions. An oceanographic model (the Copernicus Global Oceanographic Biogeochemistry Hindcast) had substantially lower prediction skill than the integrated statistical predictions (RMSE ~30-90 μmol kg^{-1}). When applied to species distribution models, the estimated dissolved oxygen thresholds differed by 20-50 μmol kg^{-1} when fit to different dissolved oxygen data sources. We focus on oxygen in the northeastern Pacific, yet our approach is generalizable to other variables and systems. We recommend increased attention to validating oceanographic models when operationalized to fisheries applications, and evaluating the robustness of conclusions to environmental covariate data sources.

Introduction

To evaluate environmental drivers of ecological systems, researchers often use retrospective spatial models that identify statistical relationships between historical environmental conditions and biological data. In marine systems, seabed features (such as depth and substrate) and ocean conditions (such as pH, temperature, and oxygen) are often included as covariates to assess their influence on habitat use and population dynamics, supporting sustainable fisheries management. For instance, spatial statistical models are used for designating essential habitat (Dambrine et al., 2021; Moore et al., 2016), identifying environmental impacts on individual traits (e.g., Lindmark et al., 2023; Oke et al., 2022), and for projecting population abundance and distribution under climate change (e.g. Grüss et al., 2021; Liu et al., 2023; Thompson, et al., 2023a). These models

rely on temporal and spatial variation in observations across time and space to provide informative contrasts and estimate the responses of species to environmental conditions. Including data with a greater range of environmental conditions—such as including longer time series and greater spatial coverage—improves estimation of environmental sensitivities and projections of the impacts of climate change by increasing contrasts in the data and reducing extrapolation beyond the range of observations (Brodie et al., 2022; Davies et al., 2023; Indivero et al., 2024).

There has been particular focus on how declining ocean oxygen may impact fish distribution (Deutsch et al., 2023; Pörtner & Knust, 2007; Pörtner, 2010; Pörtner et al., 2017; Rubalcaba et al., 2020; Verberk et al., 2016). Dissolved oxygen in the ocean has decreased globally by approximately 2% since the pre-industrial period due to increased temperatures and various associated biological and physical changes (Schmidtko et al., 2017), expanding oxygen minima and hypoxic dead zones (Diaz & Rosenberg, 2008; Stramma et al., 2010). Globally, dissolved oxygen is projected to continue to decline over the next century by around 1-7% (Keeling et al., 2010; Kwiatkowski et al., 2020; Matear & Hirst, 2003). Organisms' sensitivity to dissolved oxygen depends on temperature (Deutsch et al., 2015; Pörtner & Knust 2007; Vaquer-Sunyer & Duarte, 2011), and empirical biogeographic studies have demonstrated that these interacting constraints can explain marine species distributions (Deutsch et al., 2015, 2020; Morée et al., 2023). In addition to lethal effects of low dissolved oxygen, such as those associated with large-scale mortality events in the Gulf of Mexico (Altieri et al., 2017; Joyce, 2000), moderate declines in oxygen can also impact fish behavior and influence habitable area (Deutsch et al., 2015; Gray et al., 2002; Kim et al., 2023; Kramer, 1987; Vaquer-Sunyer & Duarte, 2008). Fish response to low oxygen is expected to be non-linear (Fry, 1971; Farrell and Richards 2009). Behavioral and physiological plasticity (Kramer, 1987) can allow an organism to tolerate low oxygen up to a certain threshold, below which there are sublethal impacts on growth and other core metabolic functions (Fry, 1971; Pörtner and Knust, 2007; Farrell and Richards, 2009). Eventually oxygen becomes too low to support routine metabolism and fish must shift to habitat with sufficient oxygen to survive (Kramer, 1987). Quantifying these thresholds is therefore necessary for predicting how declining ocean oxygen will impact habitat availability and distributional shifts.

Scientists have long conducted laboratory experiments to measure thresholds of oxygen tolerance (e.g. Ultsch et al., 1978, Beamish 1964; and see Rogers et al., 2016 for meta-analysis of 96 published studies). Yet there can be high variability in such thresholds between species and taxa, and data are only available for a handful of species (Vaquer-Sunyer & Duarte, 2008; Rogers et al., 2016). Additionally, thresholds of oxygen tolerance from laboratory-derived physiological measures have sometimes failed to show improvement in predicting fish distribution at a fine spatial scale (Bandara et al., 2023; Essington et al., 2022).

Due to these limitations of laboratory studies, there is a need for statistical spatiotemporal models that can estimate dissolved oxygen effects from field observations of species density (Bandara et al., 2023; Essington et al., 2022; Franco et al., 2022; Liu et al., 2023; Thompson et

al., 2023a; Thompson et al., 2023b). However, these analyses are often hampered by limited availability of dissolved oxygen data at the same temporal and spatial resolution as the standardized catch data that are used to monitor fish distribution and abundance.

Fishery-independent surveys sometimes concurrently collect *in situ* dissolved oxygen data, though this is often limited to only a subset of years or locations (Figure 1A), often due to logistical and financial constraints. For instance, while the U.S. West Coast Bottom Trawl Survey has consistently collected annual data since 2003, concurrent oxygen data are publicly available only for a subset of years (2009-2015, 2022-2023). And while stationary profilers and buoys are useful for monitoring environmental conditions, coverage on the continental slopes and shelf habitat is often insufficient to characterize oxygen at a scale that matches catch data (e.g. in the GOBAI model, Sharp et al., 2022; Breitburg et al., 2018). Similarly, regional and global oceanographic circulation models are generally only available in a coarser spatial and temporal resolution than biological data. For instance, bottom trawl survey observations are located at multiple specific points within the ¼° grid of the Global Ocean Biogeochemistry Hindcast (GOBH, Mercator-Ocean, 2023) (Figure 1B). Bottom trawl observations are also collected at specific points of time within a day, while oceanographic outputs like GOBH are typically only available at broader daily or 3-day averages. (Figure 1B). This mismatch in resolution means that oceanographic models may not capture the oxygen dynamics at the particular time and location of a biological observation. Oceanographic models also have faced challenges reproducing historical ocean oxygen conditions (e.g. Oschlies et al., 2018). The predicted decline in dissolved oxygen over the next century is not uniform between or within ocean basins, and is caused by different processes in different areas and depths: reduced oxygen solubility in warmer waters (Diaz & Breitburg, 2009; Keeling et al., 2010), increased stratification causing reduced subsurface ventilation (Keeling et al., 2010; Schmidtko et al., 2017), increased eutrophication and aerobic decomposition (Diaz & Breitburg, 2009), changes in wind and upwelling patterns (Sydeman et al., 2014), and long-term natural variability (Broecker et al., 1999). Including oxygen as a covariate in species distribution modeling is overall challenging due to the limited data availability and complexity of dissolved oxygen dynamics.

There is a trade-off between spatial and temporal coverage and spatiotemporal resolution in the different approaches taken to address this lack of dissolved oxygen data. Some modelers have opted to limit analyses to subsets of biological data with concurrent dissolved oxygen data available (e.g. Essington et al., 2022; Franco et al., 2022). This provides a close observation to the time and location of the actual bottom conditions at each sampling event, capturing fine-scale spatial and temporal variation. However, it removes a large quantity of biological data from analysis, and consequently possibly removes important inter- and intra-annual environmental contrasts, especially when entire years lack environmental data. For instance, 2021 was a particularly strong hypoxic event in the U.S. Pacific Northwest coast (Barth et al., 2024). If this year of sampling data is removed from model fitting because it lacks concurrent oxygen data, unique information on fish response in extreme low oxygen conditions may be lost. Other studies

instead use output from regional or global oceanographic models in place of any concurrent data, and extract the oceanographic output to sample locations via various methods, such as nearest neighbor matching or bilinear interpolation (e.g. Bandara et al., 2023; Liu et al., 2023). Such oceanographic models can cover a wider geographic area over a longer period of time, thereby providing a more complete time series to facilitate spatial coverage (Perruche et al., 2024). However, oceanographic models may not resolve the fine-scale spatial and temporal variation in fluctuations (e.g. Verezemskaya et al., 2021, Arevalo-Martínez et al., 2015; Breitburg et al., 2018; McCarthy et al., 2015) to which fish can respond to oxygen conditions (Kim et al., 2023; Kramer, 1987). Typically oceanographic model outputs have been validated only at coarser spatial and temporal scales than biological sampling (such as climatological means and seasonal patterns, e.g. Kristiansen et al., 2024; Perruche et al., 2024), yet there have been increasing efforts to validate for the fine-scale dynamics relevant to ecological modeling, such as for temperature (Kearney, 2021; Kearney & Porter, 2009) and zooplankton (Sullaway et al., 2024). Overall, neither using *in situ* observations or oceanographic models fully uses the information available to identify links between species response and environmental conditions.

Here, we aim to improve the use of environmental covariates in spatial modeling of marine species in three main ways. First, we test an empirical statistical model to expand the fine-resolution bottom oxygen data available for spatiotemporal modeling of biological data, building off Thompson et al. (2023a). This method combines multiple sources of *in situ* bottom dissolved oxygen observations from concurrent surveys and various independent CTD casts, fits spatiotemporal statistical models to the data, and predicts oxygen levels at biological sampling events. When there are gaps in oxygen observations of biological data, this can provide a way to estimate oxygen at a fine spatial and temporal scale by interpolating to unsampled dates and locations in a robust statistical framework. We conduct skill testing of the empirical model to predict the *in situ* data. Second, we compare a commonly used oceanographic model of oxygen to *in situ* data collected by bottom trawl surveys. Third, we evaluate how use of the different oxygen data sources (concurrent only, empirical statistical predictions, and dynamical oceanographic model output) in species distribution models impacts estimation of a threshold effect of oxygen with case studies of two species (sablefish *Anoplopoma fimbria* and Dover sole *Microstomus pacificus*) in two regions (California Current and British Columbia). Overall, we provide guidance to species distribution modelers on trade-offs among different sources of dissolved oxygen data by comparing these three approaches.

In the context of improving oxygen data availability and guidance for using in marine species spatial modeling, we evaluate:

1. To what extent can *in situ* bottom dissolved oxygen observations accurately predict values at unsampled biological sampling events (i.e. those without concurrent measurements)?
2. How does accounting for different spatial and temporal variation structures in fitting models to the *in situ* oxygen data impact predictive skill?

3. How does this approach compare to bottom dissolved oxygen values from a dynamical oceanographic model output?
4. How does the choice of oxygen data source in species distribution modeling impact estimation and interpretation of effects of oxygen on species population density?

Methods

We compared estimation of fish species distribution models when fit to three different sources of dissolved oxygen data: *in situ* concurrent observations, a dynamical oceanographic output, and predictions from an empirical statistical model. We used the northeastern Pacific Coast as a study system (from the Aleutian Islands and eastern Bering Sea through Gulf of Alaska, British Columbia, and California Current) (Figure 2). First, we validate the empirical statistical predictions compared to the concurrent *in situ* data to test an approach to expand the fine-scale bottom oxygen data available when fitting spatiotemporal models to fish catch data with environmental covariates. We then compare an oceanographic output to the *in situ* concurrent dissolved oxygen data. Lastly, we fit species distribution models from these three different oxygen data sources to case studies of two species in two regions.

Concurrent observations

Dissolved oxygen measurements collected at the time of bottom trawl sampling (i.e. concurrent) were obtained for the U.S. National Oceanic and Atmospheric Administration (NOAA) bottom trawl surveys in the Gulf of Alaska (Siple et al., 2023), eastern and northern Bering Sea continental shelf (Markowitz et al., 2024), eastern Bering Sea continental slope (Hoff, 2016), Aleutian Islands (Siple et al., 2025), and U.S. West Coast (Keller et al., 2017), and the Department of Fisheries and Oceans Canada (DFO) British Columbia Groundfish Concurrent Bottom Trawl Surveys in Queen Charlotte Sound, Hecate Strait, West Coast Vancouver Island, West Coast Haida Gwaii, and the Strait of Georgia (Anderson et al., 2019; DFO, 2024) (see Table 2). Measurements were similarly collected across all bottom trawl surveys, with dissolved oxygen sensors attached to the headline of trawls.

Predictions from empirical statistical model

We tested an empirical statistical approach to expand the fine-scale bottom oxygen data available when fitting spatiotemporal models to fish catch data with environmental covariates. This method is an extension of interpolation procedures to fill in missing data (see Li & Heap 2014, Nakagawa 2015, Little & Rubin 2020 for comprehensive reviews), expanding to integrate other data sources and rigorously skill test alternative model structures. In addition to the concurrent observations, we combined bottom dissolved oxygen observations from other scientific (i.e. fishery-independent) marine biological surveys (International Pacific Halibut Commission longline surveys, Joint U.S.-Canada Pacific Hake Acoustic Trawl Survey) and independent oxygen measurements from ocean monitoring programs (see Table 2 and Figure 1 for data included). Data were quality-controlled and standardized (see supplement and Fig S1 for details).

In total, 26,617 bottom dissolved oxygen observations were included in model fitting. For each region, we fit generalized linear mixed effects models (GLMMs) to the combined oxygen dataset that considered spatial and temporal structure in different configurations using the package sdmTMB (Anderson et al., 2025). Fitted models were then used to predict bottom dissolved oxygen to sampling events from fish catch surveys. To evaluate how including different spatial and temporal structures in the model might improve prediction skill, we compared models with increasing levels of complexity (Table 2). We describe these models below.

The simplest model was a GLMM that estimated bottom dissolved oxygen O of observation i in region r following a Gaussian distribution based on depth D in meters and day of year C

$$O_{i,r} = s(\log(\mathrm{D}_i)) + s(\mathrm{C}_i), \qquad \text{Eq. 1}$$

where O_i is the oxygen measurement for sample i in μmol kg^{-1}, and $s(log(D))_i$ and $s(\mathrm{C})_i$ are smooth on log(D) and day-of-year (i.e. 1–365) (see supplement for additional model details). Specifically, we used penalized regression splines ("P-splines") for both smooths (Eilers & Marx, 2021). Depth and day-of-year were included to account for seasonal and depth-dependent variation in ocean dissolved oxygen. We then add a stationary spatial random effect ω_i to Eq. 1 that accounts for spatially structured Gaussian Markov random field (GMRF) latent variables that approximate a Gaussian random field with a Matérn correlation function (Eq. 4-6 in Table 1). We iteratively tested mesh size, and used a minimum allowed triangle edge length of 45 km to optimize number of knots in each region (~100-200) (see supplement for more details). The spatial field represents variation not explained by fixed effects, which includes both process variation (spatial patterning in oxygen levels shared across years) as well as sampling effects (slight differences in measurements from different surveys). We include temporal variability in two ways. One, we include a fixed effect of year t (as a factor) b_t to represent interannual variation in mean oxygen levels (Eq. 7-9). Second, we consider variation across space that evolves over time with a latent first-order autoregressive GMRF spatiotemporal processes ε_i (Eq. 10-12), which allows each year to have unique spatial patterning of oxygen levels.

For each of the four spatial and temporal structures described above (null, spatial only, annual+spatial, and spatial+spatiotemporal), we test three configurations of covariates—none, temperature only, and temperature and salinity—to determine the predictive ability of potential covariates that could be associated with oxygen concentration. Salinity (PSS-78, practical salinity units) was converted to potential density anomaly with a reference pressure of 0 dbar in kg m^{-3} (see supplement for more details). These covariates were considered because temperature and salinity are measured in trawl surveys more frequently than dissolved oxygen, and are likely to covary with dissolved oxygen due to shared physical drivers (e.g. upwelling, freshwater inputs). We also considered more complex model structures (e.g. spatially varying coefficients, interactions, bivariate splines, survey effects, and oxygen solubility) but found no improvement to predictive skill, so for brevity we do not include them here.

Table 1. Alternative models for fitting the empirical statistical model to combined dissolved oxygen data O_i, where i is used to index the observation. Four structures for accounting for latent spatial and temporal effects were considered (none, persistent spatial fields ω_i, annual effects in addition to persistent spatial fields $b_{t[i]}$, and persistent spatial fields in addition to latent spatiotemporal variation ε_i) for three covariate structures (no covariates, temperature only $s(T_i)$, and temperature and salinity $s(S_i)$ (potential density anomaly referenced to 0 dbar, kg m^{-3}). Components included in each model are marked with an X. Smoother functions were used on temperature and salinity. All models included smoothers on day-of-year $s(\mathrm{C}_i)$and depth $s(\log(D_i))$.

Equation Numbers	Spatial and Temporal Structure			Covariates		Equation
	Spatial	Annual	Spatiotemporal	Temp	Sal	
1						$O_i = s(\mathrm{C}_i) + s(log(\mathrm{D}_i))$
2				X		$+ s(T_i)$
3				X	X	$+s(T_i) + s(S_i)$
4	X					$+\omega_i$
5	X			X		$+\omega_i + s(T_i)$
6	X			X	X	$+\omega_i + s(T_i) + s(S_i)$
7	X	X				$+\omega_i + b_{year[i]}t_i$
8	X	X		X		$+\omega_i + b_{year[i]}t_i + s(T_i)$
9	X	X		X	X	$+\omega_i + b_{year[i]}t_i + s(T_i) + s(S_i)$
10	X		X			$+\omega_i + \varepsilon_i$
11	X		X	X		$+\omega_i + \varepsilon_i + s(T_i)$
12	X		X	X	X	$+\omega_i + \varepsilon_i + s(T_i) + s(S_i)$

To evaluate prediction skill of all models, we used a cross-validation procedure that iteratively removed one year of bottom trawl oxygen observations from the integrated oxygen dataset (i.e. "test year"), trained the model on the remaining years of data (which included data from non-concurrent sources in the test year, i.e. "training years"), used the fitted model to predict to concurrent sampling events from the test year, and compared the model predictions of dissolved oxygen to the observed concurrent oxygen in the test year. We are assessing whether we can accurately predict oxygen for an entire missing sample year, given concurrent data (i.e. the measurements collected with fish sampling) from the same survey available in other years and non-concurrent data from other sources in that missing year. Only years with more than 50 concurrent observations and independent data within that year were included as test years. For the California Current, this allowed nine test years (2009-2015, 2022-2023), six years for British Columbia (2017-2019, 2021-2023), but only two years for the Gulf of Alaska (2013, 2015), Eastern Bering Sea (2012, 2016), and Aleutian Islands (2014, 2016). The number of observations in each test year for each region ranged from 50—1314, and in training data from 680—10,592 (see Table S1df). We calculated root mean squared error (RMSE) of model predictions to observations in each excluded test year as a metric of predictive skill.

Table 2. Bottom dissolved oxygen data included in model fitting. "Concurrent" refers to oxygen observations measured *in situ* in the same survey protocol as bottom trawl sampling (either trawl-mounted or concurrent CTD cast). "Independent" refers to oxygen observations measured as part of surveys or research separate from bottom trawl surveys.

Name	Type	Region	Years	Number of observations	Source
West Coast Bottom Trawl	Concurrent	California Current	2009-2015, 2022-2023	5233	Keller et al., 2017
Alaska Bottom Trawl Survey	Concurrent	Gulf of Alaska	2013, 2015, 2023-2024	1316	Rohan et al., 2024; Siple et al., 2023
Alaska Bottom Trawl Survey	Concurrent	Aleutian Islands	2014, 2016, 2024	551	Rohan et al., 2024; Siple et al., 2025
Alaska Bottom Trawl Survey	Concurrent	Eastern Bering Sea	2012, 2016, 2023-2024	300	Rohan et al., 2024; Markowitz et al., 2023
DFO West Coast	Concurrent	British Columbia	2017-2023	1802	Anderson et al., 2019; DFO Canada, 2024
International Pacific Halibut Commission Long-Line Survey	Independent	California Current, British Columbia, Eastern Bering Sea (2 years), Gulf of Alaska, Aleutian Islands	2009-2022; Bering Sea 2006, 2015	7598	IPHC, 2022, 2024
Pacific Hake Acoustic Trawl Survey	Independent	California Current	2011-2013, 2015	368	NOAA, 2020
Newport Line	Independent	California Current	1998-2021	2425	Risien et al., 2022, 2023
CalCOFI	Independent	California Current	1949-2021	1176	CalCOFI, 2024
Line P	Independent	British Columbia, California Current	1990-2019	3	Franco et al., 2021
WCOA (NOAA West Coast Ocean Acidification) and Harmful Algal Bloom surveys	Independent	California Current	2007, 2011-2013, 2016-2017, 2021	964	NOAA, 2021b

CODAP	Independent	West Coast, Gulf of Alaska, Eastern Bering Sea	2007-2013, 2015-2017	1347	NOAA, 2021a
OCNMS (Olympic Coast National Marine Sanctuaries)	Independent	West Coast	2005-2023 2004-2015	2692	Risien et al., 2024b, 2024a

Oceanographic output

We also compared how a widely used dynamical oceanographic model product—the Copernicus Global Ocean Biogeochemistry Hindcast (GOBH)—compared to *in situ* bottom dissolved oxygen observations from concurrent fish surveys versus predictions from our approach. GOBH provides 3D dissolved oxygen values at a ¼ degree resolution and with 75 vertical levels (fixed levels for all points, except for the last level, where thickness is adapted to the bathymetry) from December 1992 through March 2024 at a daily resolution (Mercator-Ocean, 2024; Perruche et al., 2024). We selected GOBH because it provided comprehensive and consistent coverage of our entire study region and time period.

We extracted the bottom dissolved oxygen for each fish catch event from GOBH by using a statistically robust method to interpolate between the grid points and depths in GOBH. For each region in each year of fish catch data, we subset the GOBH model for the geographic domain of the region (including all depths) for the time period of the survey (May 1st-October 30th). We then fit the dissolved oxygen output for this entire GOBH subset to a generalized linear mixed effects model (GLMM) using sdmTMB, including smoother effects on depth and day-of-year, spatial random fields, and following a Gaussian distribution (Eq. 13). We then extracted dissolved oxygen as predictions to the specific day, depth, and location (coordinates) of each fish catch observation. We calculated RMSE of *in situ* oxygen observations as a metric of accuracy.

$$O_{i,year,region} = s(log(\text{depth}_i)) + s(\text{doy}_i) + \omega_i \quad \text{Eq. 13}$$

Species distribution modeling comparison

We fit case studies of fish distribution models using each of the three dissolved oxygen data sources and compared estimated effects of oxygen on fish density. We chose two benthic-associated species with commercial importance—one that has a deeper habitat preference, sablefish (*Anoplopoma fimbria*), and one shallower, Dover sole (*Microstomus pacificus*)—in the California Current and British Columbia. To estimate the response of fish density to dissolved oxygen, we implemented GLMMs in sdmTMB (Anderson et al, 2025) using total catch rate as a Tweedie distributed response (Eqs S1-S2). The model formula included fixed effects of quadratic logged depth, year, and dissolved oxygen, and a random spatial field (Eqs S1). Dissolved oxygen was modeled as a breakpoint function (Eq. S3). We focus on a breakpoint model of oxygen (see

supplement for more details) because fish response to dissolved oxygen is expected to be saturating (Kramer, 1987) and we are primarily interested in how different oxygen data sources could impact detection and estimation of an oxygen limitation on fish distribution. For the test of concurrent data only, we restricted model fitting to include data only where concurrent dissolved oxygen was available. For dissolved oxygen from empirical statistical predictions, we used the model that included spatiotemporal variation and temperature, which was the model with the highest predictive skill and would maximize the number of additional trawls (as salinity data was missing in many trawls).

Results

Empirical statistical predictions

The skill in predicting oxygen from the empirical statistical model varied between regions. Dissolved oxygen at bottom trawl sampling events in the California Current and British Columbia could be predicted with relatively high accuracy, while predictions were less skilled in all Alaska regions. For the model with the highest predictive skill in each region, the overall RMSE (across all test years) for the California Current was 16.2 μmol kg^{-1} and 25.1 μmol kg^{-1} for British Columbia (relative to observed range of 1-400 μmol kg^{-1} and 2-370 μmol $kg^{-1,}$ respectively), but was 36.7 μmol kg^{-1} for the eastern Bering Sea (relative to observations ranging from 16-450 μmol kg^{-1}), 55.1 μmol kg^{-1} for the Gulf of Alaska (relative observations ranging from 15-690 μmol kg^{-1}), and 55.2 μmol kg^{-1} for the Aleutian Islands (relative to observations ranging from 14-560 μmol kg^{-1}) (Table 3). At the resolution of sampling events, across all regions combined, 92% of test observations (n=13,581) were within 50 μmol kg^{-1} and 78% within 25 μmol kg^{-1} (Figure 3). Within each region, predictions from the model of bottom dissolved oxygen that included all covariates were more accurate and unbiased for the California Current, British Columbia, and Eastern Bering Sea, while predicted oxygen in the Gulf of Alaska and Aleutian Islands were consistently underestimated by the model (Figure 3).

In all regions, the model with the highest predictive skill in each region included both temperature and salinity, and a spatial and/or spatiotemporal random fields (Table 3). Including only temperature increased accuracy (i.e. reduced RMSE) over models with no covariates by ~2-8 μmol kg^{-1}, while including both temperature and salinity improved predictive skill by ~5-11 μmol kg^{-1} (Table 3). For models with both temperature and salinity, including latent spatial, temporal, or spatiotemporal structures usually improved predictive skill, but with smaller improvements than from adding temperature and salinity to the null model and by varying amounts across regions (Table 3). The strongest improvement was seen in the Eastern Bering Sea where RMSE decreased by 3.9 μmol kg^{-1} when spatial and spatiotemporal structures were added. However, in some cases, certain temporal and spatial structures increased RMSE compared to the temperature and salinity only models (e.g., Gulf of Alaska S+T vs. S+T+spatiotemporal in Table 3). For models with no observable covariates, including spatial, temporal, or

spatiotemporal structure only improved predictive skill over a null model by up to ~4 μmol kg^{-1}, except for in British Columbia where it improved accuracy by ~7 μmol kg^{-1}. The spatial or temporal structure that provided the most accurate predictions differed between regions. For the full covariate model (which was consistently more accurate than a model with no observable covariate regardless of spatial or temporal structure), the spatial and temporal structure with the highest predictive skill was spatial variation in the California Current, Gulf of Alaska, and Aleutian Islands, and both spatial and spatiotemporal variation in British Columbia and the Eastern Bering Sea. Overall, all the most accurate models included spatial and/or spatiotemporal structures, but these structures alone could not achieve the same predictive skill as including observable covariates.

Differences in accuracy between regions were likely driven by the amount and coverage of training data available. Accuracy was notably improved with increased density of observations (Figure S2). The densities of observations in the Gulf of Alaska and Aleutian Islands were only a quarter the density of observations in the California Current, and RMSE was three times higher. In the Aleutian Islands, for instance, there were fewer than 1,000 (ranging from 685—745 for each test year) bottom dissolved oxygen observations for fitting models in a survey area of over 70,000 km^2, while in the California Current there was an order of magnitude more—over 12,000 observations—for an area around 100,000 km^2 (Figure S2). British Columbia had ~5,000 observations and an area of 53,580 km^2, while the Gulf of Alaska had ~5,000 observations for an area of 320,202 km^2. Data coverage in some regions was patchier and more spatially unbalanced between years than other regions; for instance, the Eastern Bering Sea had highly varied spatial coverage between years (Figure S3C), while the California Current was more consistently covered (Figure S3A). The concurrent bottom trawl dissolved oxygen data in Gulf of Alaska and the Aleutian Islands also had overall higher oxygen values than the training datasets, but no bias was observed in any other region (Figure S4). Since only two years of concurrent bottom trawl data were available in each Alaska region for predictive skills testing, there were fewer years for comparison than the California Current and British Columbia. As dynamic oceanographic variables (such as temperature, oxygen, and salinity used here) and depth are highly correlated (Figure S6), distinct bathymetric and oceanographic features of each region likely contributed to regional differences in predictive skill and the best spatial and temporal structures. Yet overall there were not evident patterns in bias of predictions within each region across space (Figure S3), salinity, temperature, year, or day-of-year (Figure S5A). However, oxygen was generally over-predicted at shallower depths with higher predicted oxygen in the Alaska regions, while in British Columbia and the California Current there was underestimation of oxygen at shallow depths and low predicted oxygen (Figure 4).

Oceanographic model

Accuracy of the oceanographic output, Copernicus Global Ocean Biogeochemistry Hindcast (GOBH), compared to the observed bottom dissolved oxygen at fish catch points also differed between regions. The overall error (i.e. RMSE) was 38.2 μmol kg^{-1} in the California Current and 43.7 μmol kg^{-1} in British Columbia, but almost doubled in Alaskan regions, with an RMSE of 66.5 μmol kg^{-1} in the Gulf of Alaska, 65.8 μmol kg^{-1} in the Eastern Bering Sea, and 89.4 μmol kg^{-1} in the Aleutian Islands. There was both a consistent bias in predictions underestimating observed oxygen and a high imprecision in estimates, as evident at the scale of individual fish catch sampling events (Figure 5). This inaccuracy in predicting oxygen at fish sampling events was similar when comparing GOBH to independent CTD cast observations (Figure S7). Overall the GOBH oxygen values were less accurate (i.e. a higher RMSE) than empirical statistical predictions in each region, with an RMSE ~15-30 μmol kg^{-1} higher than the integrated predictions in each region (Table 3).

Table 3. Overall root mean squared error (RMSE) of predictions empirical statistical predictions of bottom dissolved oxygen predictions for fish catch survey events withheld from training data for each region. Models with different spatial and temporal structures—none, persistent spatial fields ("Spatial") and plus annual fixed effects of year ("Annual") or plus spatiotemporal latent effects ("Spatiotemporal"— and including either no covariates ("None"), temperature ("T") or temperature and salinity ("T+S") (with salinity as sigma0) were fit to *in situ* dissolved bottom oxygen observations. Units are μmol kg^{-1}. RMSE is also shown for the Copernicus Global Ocean Biogeochemistry Hindcast (GOBH). an interpolation method that uses a spatial model to predict bottom dissolved oxygen from GOBH given depth, day-of-year, and latent spatial fields. Blue cells indicate models with the lowest out-of-sample RMSE.

		California Current			**British Columbia**			**Gulf of Alaska**			**Eastern Bering Sea**			**Aleutian Islands**		
		None	*T*	*S+T*	*None*	*T*	*S+T*	*None*	*T*	*S+T*	*None*	*T*	*S+T*	*None*	*T*	*S+T*
Empirical Statistical	*None*	23.4	20.7	17.4	36.9	33.3	25.5	64.4	65.0	57.9	47.2	40.1	40.6	68.6	63.6	56.6
	Spatial	21.1	19.1	16.2	28.3	27.7	25.6	56.9	59.5	55.1	47.8	40.3	40.6	69.1	57.9	55.2
	Annual	22.2	19.9	16.9	29.0	28.3	25.2	59.9	61.1	56.7	46.6	55.6	44.8	83.6	71.7	76.4
	Spatiotemporal	21.6	19.7	16.8	29.5	27.7	25.1	60.2	61.8	59.4	45.6	43.9	36.7	71.5	62.4	65.6
	GOBH		38.2			43.7			66.5			65.8			89.4	

Application to fish distribution model

In the case studies, the detection of a threshold—the biologically expected model—did not depend on the oxygen data used; however, if there was a threshold effect, the oxygen value at which fish densities were limited was sensitive to the type of oxygen data. For Dover sole in the California Current and sablefish in British Columbia, there was no threshold effect of oxygen detected (i.e. a flat line in the conditional effects plots) when fish density was fit to any of the oxygen data types (Figure 6). Similarly, for Dover sole in British Columbia and sablefish in the California Current, a threshold effect was detected regardless of the oxygen data type (Figure 6). However, the estimated threshold of this effect depended on data type. For instance, for sablefish in the California Current, when fit to only the concurrent *in situ* data, the estimated breakpoint (maximum likelihood estimate ± standard error) was 20.7 ± 60.8 μmol kg^{-1}, but 56.14 ± 69 μmol kg^{-1} when using empirical statistical predictions, and 73.6 ± 60.9 μmol kg^{-1} from the oceanographic model (Table S2). For Dover sole in British Columbia, the estimated threshold

was 22.7 ± 111.9 μmol kg^{-1} for the *in situ* data, 38.6 ± 114 μmol kg^{-1} for empirical statistical predictions, and 54.9 ± 130.8 μmol kg^{-1} for the oceanographic model. When these estimated dissolved oxygen thresholds are subsequently used to estimate habitat with dissolved oxygen concentrations higher than or less than the threshold, a much wider range (essentially all trawls across the California Current) would be identified as having insufficient oxygen (i.e. unsuitable) using the integrated predictions of oxygen (Figure 7). However, the model fit to concurrent data identifies only a narrow band at the (deeper) western edge of the range with unsuitable conditions (Figure 7).

The distribution of fish catch across depth and oxygen conditions in the different data types likely contributed to differences in estimated thresholds. For instance, in the California Current the concurrent data shows a much tighter range of oxygen conditions than predicted by the integrated model or in the oceanographic output at deeper depths (Figure 8). The difference in estimated effects are primarily due to differences in the oxygen predictions, rather than just the amount of data available. Here we showed the most realistic scenario, in which we used the integrated predictions or oceanographic model to predict at all trawl samples when concurrent *in situ* data was not available, which greatly expanded the number of years of fish data that were used in model fitting. When data used in model fitting is instead restricted to only the years where concurrent data is available, there are even greater differences in estimated effects. For British Columbia Dover sole (Figure S8), no threshold effect was detected when using the predictions from the integrated statistical model or the oceanographic output. In the California Current, no threshold is detected when using the oceanographic output (but is still detected from the integrated statistical predictions) (Figure S8). The differences in the oxygen values between data sources is therefore impacting threshold estimation.

Discussion

Here we demonstrate the necessity of carefully considering the source of covariate environmental data in species distribution modeling, focusing on dissolved oxygen in the northeastern Pacific. We found possible benefits and drawbacks to expanding the amount of data through empirical statistical predictions or with dynamical oceanographic model output. Using a relatively simple structure of the empirical statistical model to predict dissolved oxygen by integrating multiple existing datasets, we could reasonably predict oxygen conditions at the resolution of marine biological surveys in two of the five regions (only British Columbia and California Current) over the required duration. Yet achieving reasonably accurate predictions required an extensive amount of oxygen data. The Alaska regions, with limited data, produced large errors in predictions. In all regions, the dynamical oceanographic model output (GOBH) had even lower predictive skill at specific sampling events. The level of accuracy in both approaches still introduced sufficient variation in oxygen values to affect the outcome of estimated oxygen thresholds when used in species distributions models.

We found that empirical statistical predictions could be a viable option for expanding coverage of environmental data in only some cases. Leave-one-year-out cross-validation skill testing showed that the model could reliably predict dissolved oxygen at bottom trawl events in the California Current and British Columbia (with an RMSE ~16 μmol kg^{-1} and 22 μmol kg^{-1} respectively). And in all regions, this approach could fill in gaps in *in situ* concurrent observations with greater accuracy than using the oceanographic model output for the period of time it was trained to reproduce. In all regions here, the most accurate predictions could be generated from simple regression models of other covariates (temperature and salinity) that are driven by similar oceanographic processes as oxygen, plus spatial and sometimes spatiotemporal random fields. Since temperature has historically been more frequently available in trawl datasets than dissolved oxygen, this method may be readily applicable to other biological datasets. It can also be useful to fill in missing oxygen observations if measured simultaneously with temperature and salinity and sensors fail. If temperature and salinity are not available, accounting for latent spatial and temporal structure with random fields could still reasonably predict dissolved oxygen but with more error (an additional ~5-10 μmol kg^{-1}).

Despite the utility of empirical statistical approaches for resolving oxygen at appropriate scales, this method should only be used if data can provide sufficient spatial and temporal coverage. We find that regions with lower density of observations data available (<0.075 observations per km^2) had less accurate predictions. In Alaska regions, error was especially large (~35-50 μmol kg^{-1}), and ~2x greater than British Columbia and California Current. In the Gulf of Alaska, around four times more data than available (~20,000 observations; only ~5,000 available here) would be needed to achieve the same observational density as the California Current. Regions with fewer observations had both reduced years of coverage (e.g. less than 10 years of data with annual gaps in the Aleutian Islands vs. over 20 years of continuous annual data in the California Current) and were more spatially patchy data relative to the area of the survey region, and fewer independent sources of oxygen data (e.g. three sources of oxygen data in the Aleutian Islands versus eight in the California Current). The amount of data available relative to capturing the unique oceanographic characteristics within each region (such as particular bathymetry, river outflows, wind mixing, upwelling dynamics, and biogeochemical processes) may also impact predictive skill and needs to be considered specifically for each system.

Interpolating to fill in missing covariate data is commonly used in distribution modeling (e.g. Li & Heap, 2014; Thompson et al., 2023), yet we recommend considering empirical statistical predictions only after validating for the specific region of analysis with cross-validation techniques (Roberts et al., 2017) and using only if accuracy (e.g. RMSE) is within a tolerable range for the purpose of the analysis. For instance, in our case studies of groundfish distribution modeling, we would consider using this approach only for British Columbia and California Current, as the errors in the Alaska regions are too large relative to the low oxygen values that define hypoxic conditions.

There are additional limitations to consider when applying this empirical statistical method. We did not intend for this approach to mechanistically evaluate drivers of dissolved oxygen dynamics, but rather to maximize predictive skill for obtaining point oxygen value at biological sampling events to use as a covariate in species distribution models. In our case, simpler models with salinity and temperature produced more skillful predictions than models with only latent random fields. Combining both provided the highest accuracy. Testing alternative models of increasing levels of complexity—from simple models to highly parameterized models and with complex error structures—can help identify the most skillful model for the unique oceanographic conditions in other systems. In addition, we only evaluated skill for predicting entirely unsampled years. For using this method to fill in gaps in spatial coverage within a sampled year, one could conduct skill testing with spatial blocking (Roberts et al., 2017). Lastly, not unlike statistical predictions for ocean acidification (Carter et al. 2021; Alin et al. 2023) this statistical method does not include possible long-term change (i.e. deoxygenation signals) because of the short temporal duration (< 20 years) of observations in the training period. For ocean acidification reconstructions without time-varying anthropogenic effects, usage is often limited to 10 years around the training period (McGarry et al. 2021; Alin et al. 2020). Some regions may do well to predict contemporaneous oxygen, but it is likely that biases will emerge over the future years, particularly in regions with high uptake of anthropogenic CO_2 such as the North Atlantic (Gruber et al. 2019; Carter et al. 2021).

We validate that oceanographic model outputs for dissolved bottom oxygen can be a suitable alternative if sufficient *in situ* observations are not available, but a greater amount of error in dissolved oxygen should be expected. While the oceanographic model output used here (the Global Oceanographic Biogeochemical Hindcast) has been validated for capturing climatological and seasonal dynamics across space and depth (Perruche et al., 2024), it could not precisely describe fluctuations at the resolution of fish catch surveys. This inaccuracy was due in part to bias but also due to greater variation; adjusting with a simple constant (i.e. a delta-correction method) therefore would not sufficiently correct the discrepancy.

Overall, we recommend validating oceanographic model outputs for specific applications to species distribution modeling. Oceanographic model outputs are often validated at coarser spatial and temporal scales than biological sampling (such as climatological means and seasonal patterns, e.g. Kristiansen et al., 2024; Perruche et al., 2024; though see Kearney, 2021 and Kearney & Porter, 2009 for fine-scale temperature validation and Sullaway et al., 2024 for zooplankton). Our approach compares output to *in situ* observations for the specific biological survey being analyzed, and can therefore help quantify uncertainty in oceanographic models at the scale of biological observations. We did not intend to be a comprehensive survey of oceanographic data options. Rather, we aimed to validate the model most suitable for our purposes and provide a demonstration and framework that could be generalized for any oceanographic output when applying to species distribution models. Here, we used GOBH because it provided the most comprehensive spatial and temporal coverage for our study system.

However, because of the resolution it does not capture some local processes at smaller scales. In other studies, such as for a more constrained geographic area, oceanographic model outputs such as regionally downscaled models that have higher resolution of bathymetry and coastal dynamics (e.g. upwelling and eddies) may be available and could be validated using our approach. For instance, there are hindcasts from Regional Ocean Modeling System (ROMS) models (e.g. in the northeastern Pacific, Hauri et al., 2020; Pozo Buil et al., 2021), upcoming MOM6 models (Ross et al., 2023), GOBAI-O2 (Sharp et al., 2022), and ESPER-NN (Carter et al., 2021). In addition to validating oceanographic models, we also recommend to fit a spatial model to oceanographic output to interpolate values at specific location, depth, and day of each biological sampling event, rather than using the nearest oceanographic output grid location. Error of the oceanographic output was significantly greater when using nearest neighbor matching (~200% greater RMSE in the California Current and British Columbia, and 33% in the Alaskan regions). The interpolation method used here can greatly improve the accuracy of dissolved oxygen values by reducing the coarseness of the oceanographic model and better capturing the spatial resolution of biological data.

Our analysis spotlights that the level of error in both empirical statistical predictions and dynamical oceanographic output can be enough to impact the outcomes of species distributions models of demersal fishes. In our case studies, the detection of an effect of oxygen on fish distribution did not differ between oxygen data types. Yet if a threshold effect was detected (sablefish in the California Current and Dover sole in British Columbia), the estimated breakpoint—the oxygen value below which fish density is reduced—was sensitive to data type. Our results suggest that this is due to bias and variation in empirical statistical predictions and dynamical oceanographic values compared to *in situ* data. In estimating and projecting biological responses to the environment, there are multiple sources of uncertainty (e.g. Brodie et al., 2022; Davies et al., 2023) and decision points throughout the data and modeling process (e.g. Commander et al., 2023). Here we aim to highlight that the decisions of which environmental data source to use and how to account for uncertainty in this data are two other key decision points that needs to be carefully considered.

First, in deciding the input for an environmental covariate, modelers need to evaluate the tradeoff between coverage and resolution given the options available for the specific region and time period and the purpose of the model. There is no single “best” choice. Spatial models of biological data are used for various purposes, including projecting distributions of changing climate (e.g. Liu et al., 2023), designating essential fish habitat designation (e.g. Moore et al., 2016), and evaluating environmental drivers of characteristic such as growth (e.g. Lindmark et al., 2022a; Lindmark et al., 2022b). If the primary goal is historical ecological inference and precise empirical estimates, then prioritizing resolution over coverage and using *in situ* observations collected at the time of biological sampling would likely provide the most specific measure of conditions at the location and time of biological data. For instance in our case studies, we considered our baseline as expected oxygen based on the database of oxygen measurements.

Using empirical statistical predictions or GOBH to fill in missing observations and expand the amount of data for distribution modeling did not provide enough benefit to outweigh information lost by introducing greater error in oxygen values. However for studies that need wider temporal or geographic coverage than that available from *in situ* data, empirical statistical predictions or a dynamical oceanographic model output would be more suitable. If using an oceanographic output, we recommend using the option available with the highest resolution possible that still covers the geographic extent of the study region, and following the validation and interpolation methods discussed above. Combining multiple downscaled oceanographic outputs to patch together across multiple regions (rather than using one more global model, as done with GOBH here) would require validating in each region and evaluating biases between regions to be able to compare biological models fit to different oceanographic outputs. For many species (such as Dover sole here), there are not laboratory estimates on oxygen sensitivity available. However considering the ecology and physiology of the species when available can provide context and biological realism when evaluating the sensitivity of the estimated biological response to possible error from the environmental data. For instance, the low end of hypoxia definitions (10 μmol kg^{-1}) and estimated loss-of-equilibrium for sablefish of 16 μmol kg^{-1} (at 12 C; $pO_{2,\,crit}$ of 44 μmol kg^{-1}) (Leeuwis et al., 2019) is within the RMSE of the empirical statistical and oceanographic models here even for regions with the highest predictive skill. It therefore may not be precise enough for studies evaluating precise empirical responses these low thresholds. Yet using these methods (for only the regions with higher predictive skill) may be more robust for species expected to have a higher threshold, such as Atlantic cod (52-58 μmol kg^{-1}·Claireaux et al., 2000) or Doryteuthis squid (51-68 μmol kg^{-1},Birk 2018).

There are other situations where an oceanographic model or empirical predictions may be more suitable. Oceanographic model outputs are a powerful tool for capturing biogeochemical and physical dynamics (Di Biagio et al., 2023; Li et al., 2024; Siedlecki et al. 2015; Pena et al., 2010) and are crucial for projecting impacts of future oxygen and temperature conditions (e.g. Liu et al., 2023; Thompson et al., 2023; Sunday et al. 2022). We did find that different data sources in our case studies were robust to overall categorical classification of whether a species did or did not have any response to oxygen conditions. Similarly, oceanographic model outputs would be useful for studies needing to summarize monthly, season, and annual indices and across broader geographic regions because they are often accurate at recreating trends and dynamics at these broader scales (e.g. Perruche et al., 2024). This includes annual recruitment models (e.g. Ward et al., 2024), characterizing marine heatwaves (e.g. Fredston et al., 2024), or evaluating spatially varying responses to global or regional conditions (e.g. Thorson 2019). Lastly, oceanographic model output, even if biased, is still effective for forecasting biological responses if they are used consistently. For instance, using the historical reconstruction and climate projections from the same oceanographic model will keep biases consistent between the data used to estimate the environmental effects and the projected environmental conditions.

A second decision point to consider, regardless of data source, is accounting for uncertainty in the environmental covariate data. Our framework—comparing species models fit to different environmental input data (e.g. oceanographic output versus *in situ* model) to quantify differences in estimated threshold due to data uncertainty—is one option. These range of thresholds could subsequently classify vulnerability of a species to low oxygen in a particular region by comparing the threshold range to predicted future changes in dissolved oxygen. Similar methods are already common for projecting ecological responses to climate forecasts from multiple different Earth System Models (e.g. Liu et al., 2023; Townhill et al., 2023). If only an oceanographic model hindcast and no *in situ* data are available in the region for comparison, using standard sensitivity analysis techniques (Cariboni et al., 2007; Steel et al., 2009), such as by simulating error in the oceanographic output, could quantify the robustness of estimated effects to error in the oceanographic output. Models could be combined via ensemble methods with equal or non-equal weights (e.g. Anderson et al., 2017; Harris et al., 2024). State-space models could explicitly build observation error in the environmental covariate into the model (Correa et al., 2023; Lindén & Knape, 2009; Thorson et al., 2021). Oceanographic output could be included in models as a separate covariate (Sullaway et al., 2024), or as joint models with observations as responses using shared latent variables (Thorson et al., 2021), to generate spatially resolved corrections (Hay et al., 2000). In some cases, it may be suitable to aggregate species density data into coarser scales, such as pooling data into the resolution of oceanographic output (Chen et al., 2023; Commander et al., 2022). Conversely, local biological response to regionally or globally scaled environmental variables (such as an oceanographic index) could be evaluated via spatially varying coefficients (e.g. Thorson 2019). Here, we focused on GLMMs for species distribution, but other approaches for estimating effects of oxygen on species populations, such as structural equation models (Thorson et al., 2024), simulating ecosystem dynamics in Ecopath and Ecosim (e.g. de Mutsert et al., 2016) or reconstructing environmental histories with otoliths (e.g. Valenza et al., 2023) may provide additional insights. Overall, considering uncertainty stemming from the environmental data will ensure more robust estimates of environmental drivers on ecological systems and projections of impacts under future climate change.

Our findings indicate that prioritizing measuring environmental data concurrently with biological sampling can reduce uncertainty in estimating environmental effects on marine species. We found that a global hindcast could produce greater uncertainty than concurrent data when using retrospective models to estimate environmental drivers of marine species distributions. Oceanographic models that have performed well at finer scales are not available at the large scales of time and space required for analyses like the one we performed here, due primarily to computational limitations. Concurrent observations are similarly prioritized in carbon cycling research (termed co-locating observations, e.g. Newman et al., 2019). Prioritizing environmental data can also improve oceanographic models. Dissolved oxygen is particularly challenging to capture at a fine scale in oceanographic models (Stramma et al., 2012). Yet dissolved oxygen

observations at bottom depths are less widely available (Wang et al., 2024); for instance, satellites do not detect bottom oxygen, and only a subset of Argo floats measure dissolved oxygen (Argo, 2024). Expanding advanced ocean observing instrumentation to regions with limited coverage, such as the Bering Sea and Aleutian Islands, and generating high-quality and long-term, historical reconstructions of subsurface environmental conditions will improve oceanographic modeling outputs (Breitburg et al., 2018). Collecting dissolved oxygen data concurrently requires time and financial resources that may be limited with possible declining budgets (NOAA, 2024b) and upcoming survey changes (NOAA, 2024a; NOAA Fisheries, 2024). Here, we demonstrate that these data are highly valuable for capturing fine-resolution environmental conditions at an ecologically relevant resolution, validating oceanographic outputs, and expanding coverage with predictions from empirical statistical approaches.

There is a growing need to bridge the gap between the methods and priorities of oceanographic and fisheries research communities to advance fisheries management (Berx et al., 2011, Drenkard et al., 2021). Here, we identify an opportunity for increasing alignment by validating oceanographic outputs for the resolution of biological sampling. We provide guidance on operationalizing *in situ* and oceanographic model outputs to be applicable to fisheries data and methods. In our case studies, relying on only *in situ* data or only an oceanographic output would have led to different interpretations of the vulnerability of species to decreased oxygen. The most suitable data source will depend on a study's purpose and the options available for the region and time period. We suggest greater attention to evaluating model sensitivity, quantifying uncertainty, and testing robustness of conclusions to environmental data inputs. We focused on bottom dissolved oxygen because it is less available in marine biological surveys and more difficult to capture in oceanographic models. Yet our approach and guidance can be similarly applied to other environmental characteristics often included in species modeling, such as temperature and phytoplankton. We also encourage the fisheries research community to develop standardized techniques for how environmental data (whether *in situ* or from oceanographic output) are processed in spatial models to facilitate comparison across systems.

Acknowledgements

We thank all survey participants of the NOAA West Coast, Gulf of Alaska, Bering Sea, and Aleutian Islands Bottom Trawl Surveys and the Department of Fisheries and Oceans Canada British Columbia Synoptic Bottom Trawl Surveys. We also thank Ole Shelton for helpful comments on an earlier review. This publication was partially funded by Washington Sea Grant grant no. R/SFA-12, and from the Lowell E. Wakefield Professorship.

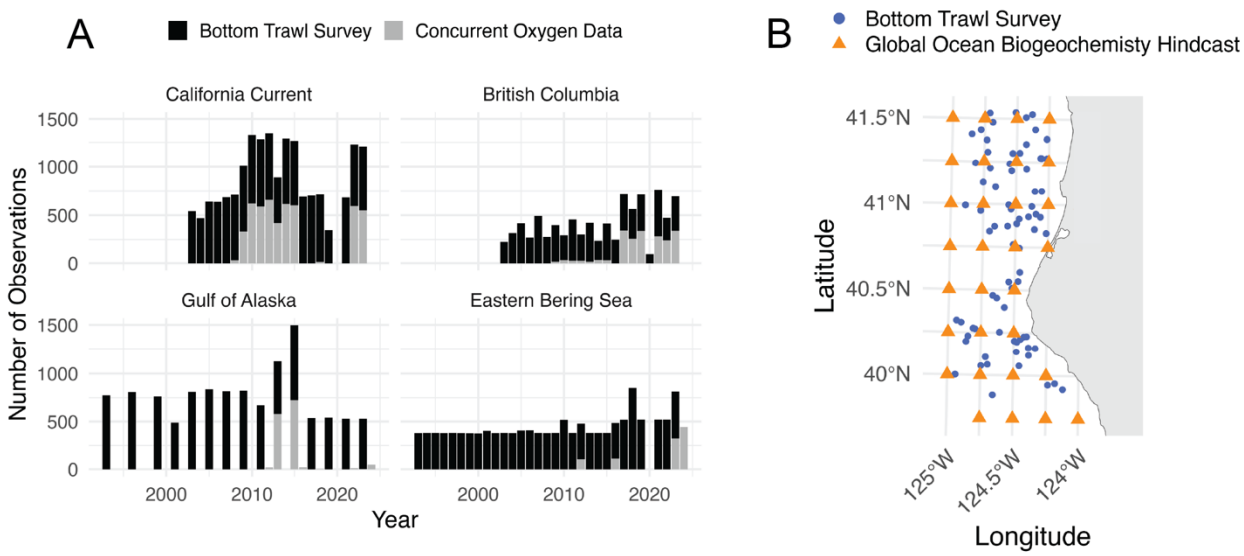


Figure 1. The trade-off between spatiotemporal coverage and resolution in oxygen data options for species distribution modeling. A) There are inter- and intra-annual gaps in oxygen data coverage of fish catch surveys. Number of observations from bottom trawl surveys (black) in the California Current (the NOAA Northwest Fisheries Science Center West Coast Bottom Trawl Survey), British Columbia (the Department of Fisheries and Oceans Canada West Coast Concurrent Trawl Surveys), the Gulf of Alaska and Eastern Bering Sea (NOAA Alaska Fisheries Science Center Gulf of Alaska and Eastern Bering Sea Bottom Trawl Surveys), and the number of trawls with *in situ* oxygen observations available (grey). B) The spatiotemporal resolution mismatch between oceanographic model output and fish catch surveys. Example of fish catch survey locations for a section of the West Coast Bottom Trawl Survey in the California Current in 2012) (blue points) and the Global Ocean Biogeochemistry Hindcast grid points (orange stars).

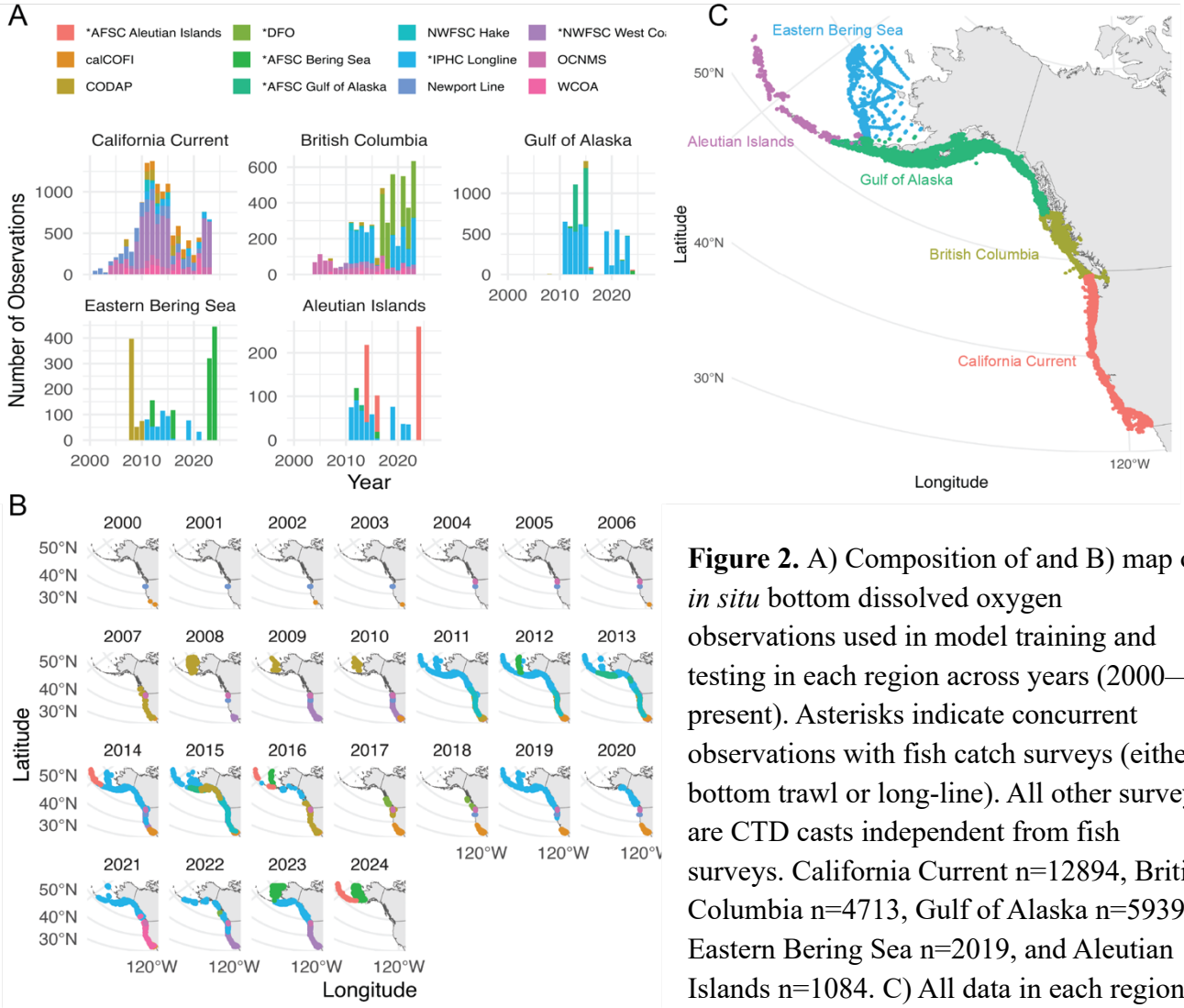
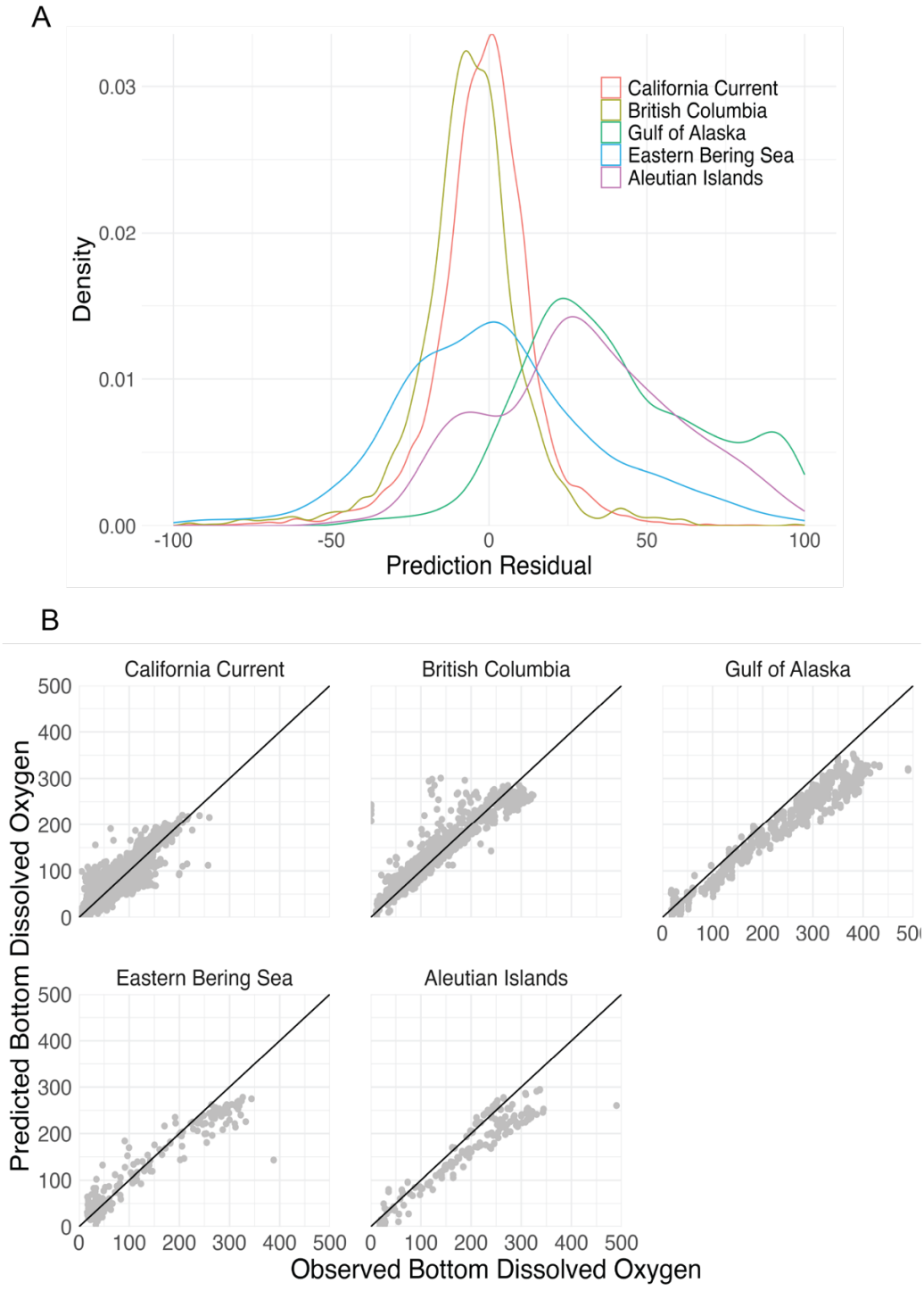


Figure 2. A) Composition of and B) map of *in situ* bottom dissolved oxygen observations used in model training and testing in each region across years (2000—present). Asterisks indicate concurrent observations with fish catch surveys (either bottom trawl or long-line). All other surveys are CTD casts independent from fish surveys. California Current n=12894, British Columbia n=4713, Gulf of Alaska n=5939, Eastern Bering Sea n=2019, and Aleutian Islands n=1084. C) All data in each region.



679

680 **Figure 3.** A) Distribution of residuals (observations minus predictions) (μmol kg^{-1}) for
681 predictions of bottom dissolved oxygen for fish catch surveys (observations withheld from model
682 fitting) from the highest-skill model of *in situ* bottom dissolved oxygen observations for each
683 region. B) Predicted bottom dissolved oxygen and observed bottom dissolved oxygen in μmol
684 kg^{-1} (*in situ* observations withheld from training data) for each region. Black line indicates a 1:1
685 relationship (i.e. perfect accuracy).

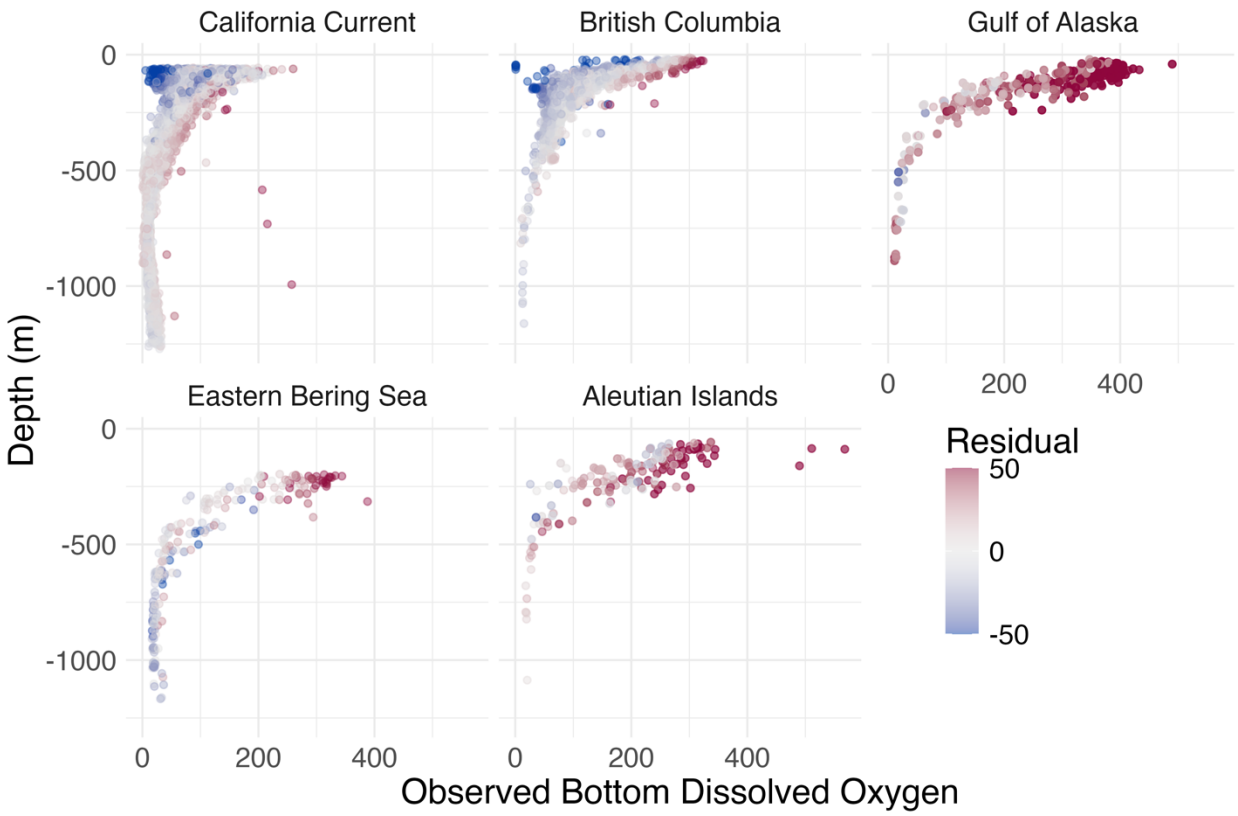
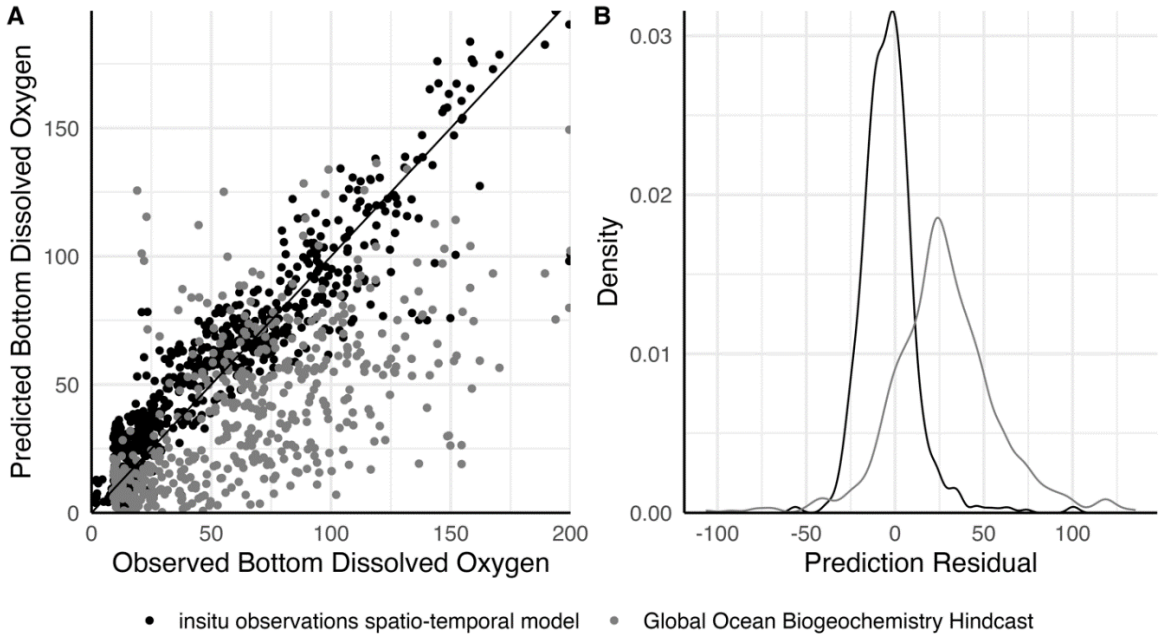


Figure 4. Residuals of predictions (observation—prediction) from all years of leave-one-year-out skills testing by observed bottom dissolved oxygen (μmol kg^{-1}) and depth of observation (m).



692

693 **Figure 5.** A) Predicted bottom dissolved oxygen and observed bottom dissolved oxygen in μmol
694 kg^{-1} (*in situ* observations withheld from training data) for an example year and region, the
695 California Current in 2012, for the spatiotemporal model with both temperature and salinity
696 covariates fit to in situ data (black) and the Global Ocean Biogeochemistry Hindcast model
697 interpolated to depth, day-of-year, and coordinates (grey). Black line indicates a 1:1 relationship
698 (i.e. perfect accuracy). B) Distribution of residuals (μmol kg^{-1}) for the predictions versus
699 observations from each.

700

701

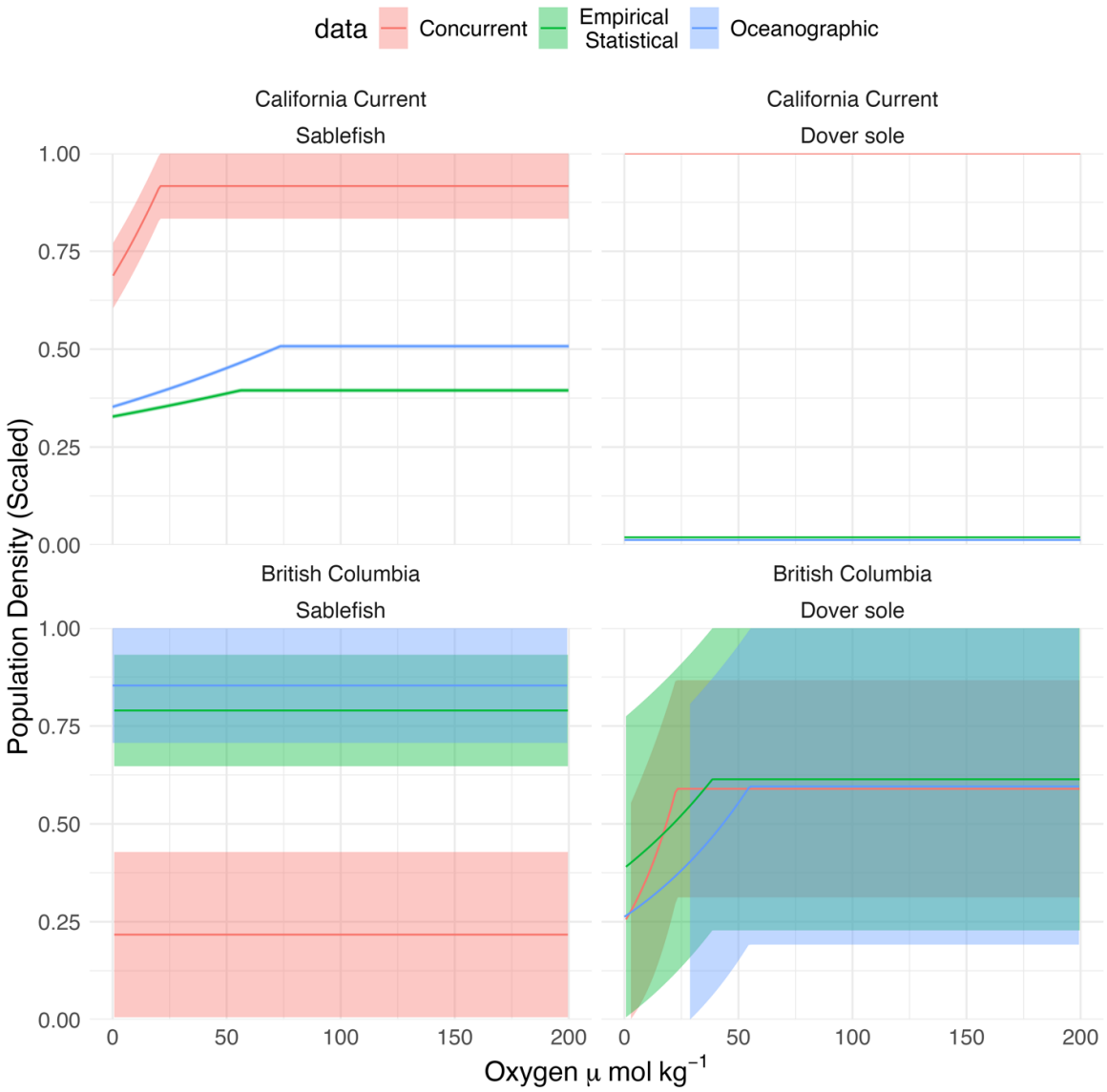


Figure 6. The response of fish density to dissolved oxygen for each species and region estimated from the three different sources of dissolved oxygen data (concurrent *in situ* measurements from the bottom trawl surveys, predictions from the empirical statistical model, and output from a global oceanographic model, the Global Oceanographic Biogeochemistry Hindcast). Response is shown as the conditional effect (maximum likelihood estimate ± standard error) of the dissolved oxygen breakpoint model. Conditional effects without standard errors are due to inherent computational challenges in maximum likelihood estimation when the second derivative of the log-likelihood is near zero or undefined, making the curvature-based (inverse Hessian) calculation of standard errors not reliably approximated. Dover sole in British Columbia and sablefish in the California Current show a threshold effect, but the specific breakpoint differs when estimated from different data sources. Dover sole in the California Current and sablefish in British Columbia show no effect of dissolved oxygen on fish density

regardless of data source. Conditional effects were calculated with depth effects set to 0 m, year at a reference year (2012 in California Current and 2019 in British Columbia), and no random effects. Population density (unscaled units of biomass kg km^{-2}) is shown scaled to the maximum value of the conditional effect on biomass density (i.e. 1) in each species and region.

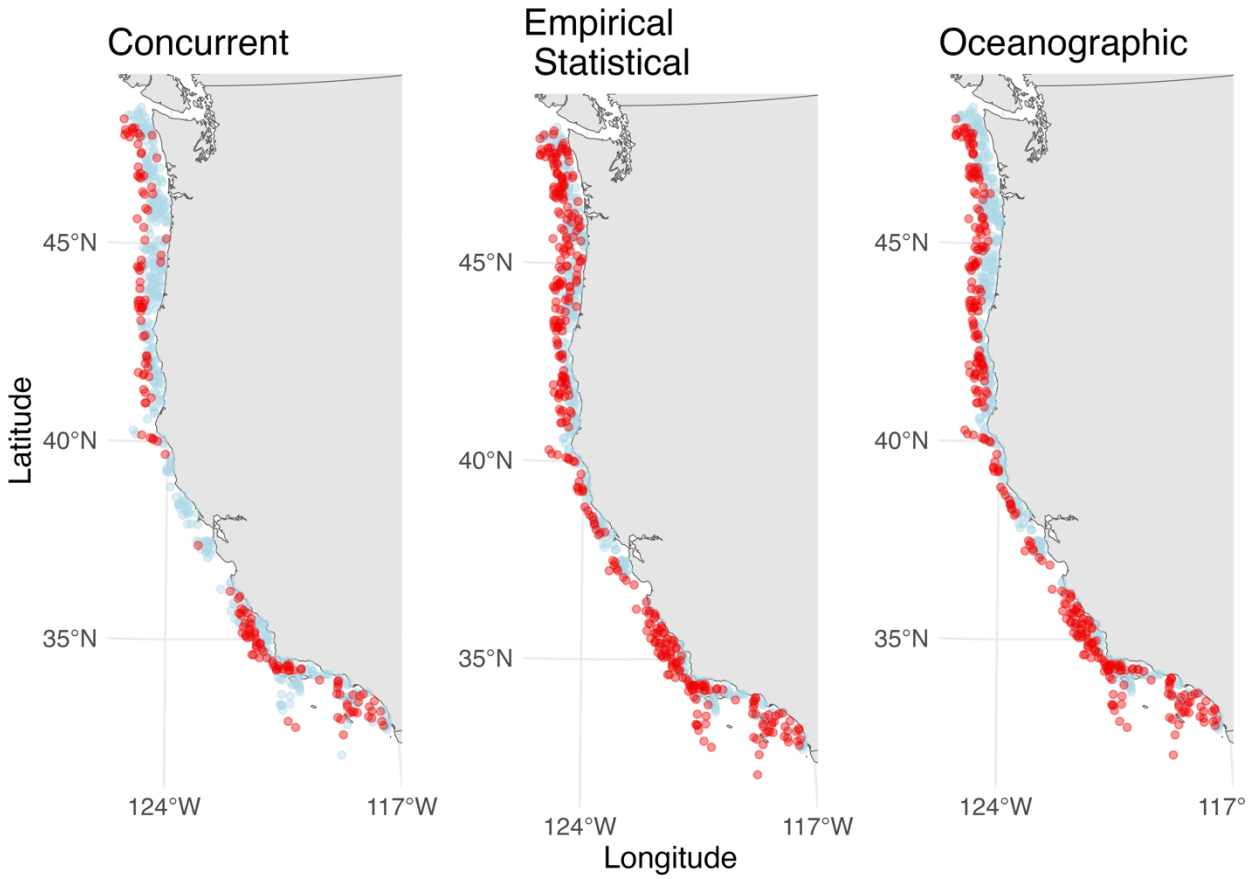


Figure 7. Trawl locations where dissolved oxygen conditions were estimated to be suitable (blue, no effect on fish density) and unsuitable (red, a reducing effect on fish density) for one example year (2023) for a case study species and region, sablefish in the California Current, using the estimated threshold from the dissolved oxygen breakpoint model fit to dissolved oxygen data from the different data sources: (left) concurrent *in situ* observations only, (center) predictions from the empirical statistical model, and (right) oceanographic model output, from the Global Oceanographic Biogeochemistry Hindcast.

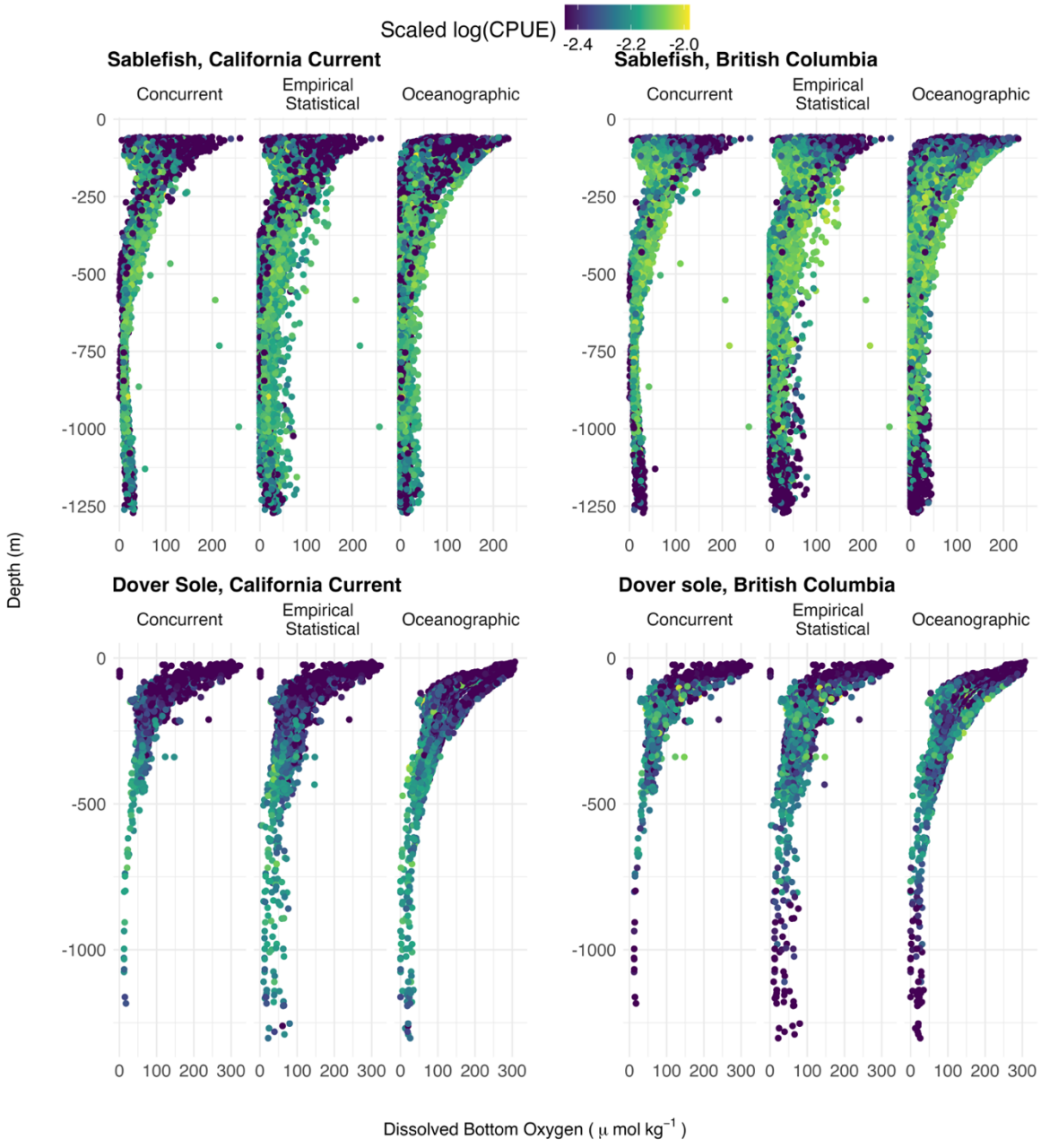


Figure 8. Dissolved oxygen (μmol kg^{-1}) and depth (m) for each of the three dissolved oxygen data sources (concurrent *in situ* observations measured at each bottom trawl survey, empirical statistical predictions, and the dynamical oceanographic output, Global Ocean Biogeochemistry Hindcast) at each haul for each case study species and region. Color indicates the fish catch per unit effort (CPUE) (kg km^{-2}) of each bottom trawl survey, scaled by the mean and standard deviation of log(CPUE) (i.e. the geometric mean and standard deviation). *In situ* observations have a narrower range of estimated dissolved oxygen concentrations across depth, especially at deeper locations, while there is much greater variation in dissolved oxygen in predictions from the empirical statistical model or oceanographic output.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Altieri, A. H., Harrison, S. B., Seemann, J., Collin, R., Diaz, R. J., & Knowlton, N. (2017). Tropical dead zones and mass mortalities on coral reefs. *Proceedings of the National Academy of Sciences*, *114*(14), 3660–3665. https://doi.org/10.1073/pnas.1621517114

Anderson, S. C., Cooper, A. B., Jensen, O. P., Minto, C., Thorson, J. T., Walsh, J. C., Afflerbach, J., Dickey-Collas, M., Kleisner, K. M., & Longo, C. (2017). Improving estimates of population status and trend with superensemble models. *Fish and Fisheries*, *18*(4), 732–741.

Anderson, S. C., Keppel, E. A., & Edwards, A. M. (2019). *A reproducible data synopsis for over 100 species of British Columbia groundfish*. https://www.dfo-mpo.gc.ca/csas-sccs/Publications/ResDocs-DocRech/2019/2019_041-eng.html

Anderson, S. C., Ward, E. J., English, P. A., Barnett, L. A. K., & Thorson, J.T. (2025). sdmTMB: an R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. In press at *Journal of Statistical Software*. *BioRxiv* preprint: 2022.03.24.485545. https://doi.org/10.1101/2022.03.24.485545

Anderson, Sean. A., Keppel, E. A., & Edwards, A. M. (2019). *A reproducible data synopsis for over 100 species of British Columbia groundfish*. publications.gc.ca/pub?id=9.884045&sl=0

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, *22*(1), 42–47.

Arevalo-Martínez, D. L., Kock, A., Löscher, C. R., Schmitz, R. A., & Bange, H. W. (2015). Massive nitrous oxide emissions from the tropical South Pacific Ocean. *Nature Geoscience*, *8*(7), 530–533.

Argo. (2024). *Argo float data and metadata from Global Data Assembly Centre (Argo GDAC)*. https://doi.org/10.17882/42182

Bandara, R. M. W. J., Curchitser, E., & Pinsky, M. L. (2023). The importance of oxygen for explaining rapid shifts in a marine fish. *Global Change Biology*, *n/a*(n/a), e17008. https://doi.org/https://doi.org/10.1111/gcb.17008

Barth, J. A., Pierce, S. D., Carter, B. R., Chan, F., Erofeev, A. Y., Fisher, J. L., Feely, R. A., Jacobson, K. C., Keller, A. A., & Morgan, C. A. (2024). Widespread and increasing near-bottom hypoxia in the coastal ocean off the United States Pacific Northwest. *Scientific Reports*, *14*(1), 3798.

774 Beamish, F.W.H. 1964. Respiration of fishes with special emphasis on standard oxygen
775 consumption: II. Influence of weight and temperature on respiration of several species.
776 *Canadian Journal of Zoology,* 42(2): 177-188. https://doi.org/10.1139/z64-016

777 Berx, B., Dickey-Collas, M., Skogen, M. D., De Roeck, Y.-H., Klein, H., Barciela, R., Forster, R.
778 M., Dombrowsky, E., Huret, M., & Payne, M. (2011). Does operational oceanography
779 address the needs of fisheries and applied environmental scientists? *Oceanography*, *24*(1),
780 166–171.

781 Birk, M.A. 2018. Ecophysiology of Oxygen Supply in Cephalopods. Dissertation. University of
782 South Florida. https://digitalcommons.usf.edu/etd/7265

783 Breitburg, D., Levin, L. A., Oschlies, A., Grégoire, M., Chavez, F. P., Conley, D. J., Garçon, V.,
784 Gilbert, D., Gutiérrez, D., Isensee, K., Jacinto, G. S., Limburg, K. E., Montes, I., Naqvi, S.
785 W. A., Pitcher, G. C., Rabalais, N. N., Roman, M. R., Rose, K. A., Seibel, B. A., & Zhang, J.
786 (2018). Declining oxygen in the global ocean and coastal waters. *Science*, *359*(6371),
787 eaam7240. https://doi.org/10.1126/science.aam7240

788 Brodie, S., Smith, J. A., Muhling, B. A., Barnett, L. A. K., Carroll, G., Fiedler, P., Bograd, S. J.,
789 Hazen, E. L., Jacox, M. G., Andrews, K. S., Barnes, C. L., Crozier, L. G., Fiechter, J.,
790 Fredston, A., Haltuch, M. A., Harvey, C. J., Holmes, E., Karp, M. A., Liu, O. R., … Kaplan,
791 I. C. (2022). Recommendations for quantifying and reducing uncertainty in climate
792 projections of species distributions. *Global Change Biology*, *28*(22), 6586–6601.
793 https://doi.org/10.1111/gcb.16371

794 Broecker, W. S., Sutherland, S., & Peng, T.-H. (1999). A possible 20th-century slowdown of
795 Southern Ocean deep water formation. *Science*, *286*(5442), 1132–1135.

796 Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a
797 metabolic theory of ecology. *Ecology*, *85*(7), 1771–1789. https://doi.org/10.1890/03-9000

798 CalCOFI. (2024). *CTD Cast Files, California Cooperative Ocean Fisheries Investigations2*.
799 https://calcofi.org/data/oceanographic-data/ctd-cast-files/

800 Cariboni, J., Gatelli, D., Liska, R., & Saltelli, A. (2007). The role of sensitivity analysis in
801 ecological modelling. *Ecological Modelling*, *203*(1), 167–182.
802 https://doi.org/https://doi.org/10.1016/j.ecolmodel.2005.10.045

803 Carter, B. R., Bittig, H. C., Fassbender, A. J., Sharp, J. D., Takeshita, Y., Xu, Y.-Y., Álvarez, M.,
804 Wanninkhof, R., Feely, R. A., & Barbero, L. (2021). New and updated global empirical
805 seawater property estimation routines. *Limnology and Oceanography: Methods*, *19*(12),
806 785–809. https://doi.org/https://doi.org/10.1002/lom3.10461Chen, Y., Shan, X., Gorfine, H.,
807 Dai, F., Wu, Q., Yang, T., Shi, Y., & Jin, X. (2023). Ensemble projections of fish distribution
808 in response to climate changes in the Yellow and Bohai Seas, China. *Ecological Indicators*,
809 *146*, 109759. https://doi.org/https://doi.org/10.1016/j.ecolind.2022.109759

Claireaux, G., Webber, D.M., Lagardère, J.P., and Kerr, S.R. 2000. Influence of water temperature and oxygenation on the aerobic metabolic scope of Atlantic cod (Gadus morhua). *Journal of Sea Research* 44(3-4): 257-265. https://doi.org/10.1016/S1385-1101(00)00053-8

Commander, C. J. C., Barnett, L. A. K., Ward, E. J., Anderson, S. C., & Essington, T. E. (2022). The shadow model: how and why small choices in spatially explicit species distribution models affect predictions. *PeerJ*, *10*, e12783. https://doi.org/10.7717/peerj.12783

Correa, G. M., Monnahan, C. C., Sullivan, J. Y., Thorson, J. T., & Punt, A. E. (2023). Modelling time-varying growth in state-space stock assessments. *ICES Journal of Marine Science*, *80*(7), 2036–2049. https://doi.org/10.1093/icesjms/fsad133

Dambrine, C., Woillez, M., Huret, M., & de Pontual, H. (2021). Characterising Essential Fish Habitat using spatio-temporal analysis of fishery data: A case study of the European seabass spawning areas. *Fisheries Oceanography*, *30*(4), 413–428.

Davies, S. C., Thompson, P. L., Gomez, C., Nephin, J., Knudby, A., Park, A. E., Friesen, S. K., Pollock, L. J., Rubidge, E. M., Anderson, S. C., Iacarella, J. C., Lyons, D. A., MacDonald, A., McMillan, A., Ward, E. J., Holdsworth, A. M., Swart, N., Price, J., & Hunter, K. L. (2023). Addressing uncertainty when projecting marine species' distributions under climate change. *Ecography*, *2023*(11), e06731. https://doi.org/https://doi.org/10.1111/ecog.06731

de Mutsert, K., Steenbeek, J., Lewis, K., Buszowski, J., Cowan Jr, J. H., & Christensen, V. (2016). Exploring effects of hypoxia on fish and fisheries in the northern Gulf of Mexico using a dynamic spatially explicit ecosystem model. *Ecological Modelling*, *331*, 142–150.

Deutsch, C., Ferrel, A., Seibel, B., Pörtner, H.-O., & Huey, R. B. (2015). Climate change tightens a metabolic constraint on marine habitats. *Science*, *348*(6239), 1132–1135.

Deutsch, C., Penn, J. L., & Lucey, N. (2023). Climate, Oxygen, and the Future of Marine Biodiversity. *Annual Review of Marine Science*. https://doi.org/10.1146/annurev-marine-040323-095231

Deutsch, C., Penn, J. L., & Seibel, B. (2020). Metabolic trait diversity shapes marine biogeography. *Nature*, *585*(7826), 557–562. https://doi.org/10.1038/s41586-020-2721-y

DFO Canada. (2024). *Groundfish Concurrent Bottom Trawl Surveys*. https://open.canada.ca/data/en/dataset/a278d1af-d567-4964-a109-ae1e84cbd24a

Di Biagio, V., Martellucci, R., Menna, M., Teruzzi, A., Amadio, C., Mauri, E., & Cossarini, G. (2023). Dissolved oxygen as an indicator of multiple drivers of the marine ecosystem: the southern Adriatic Sea case study. *State of the Planet*, *1*, 1–13.

Diaz, R. J., & Breitburg, D. L. (2009). Chapter 1 The Hypoxic Environment. In J. G. Richards, A. P. Farrell, & C. J. Brauner (Eds.), *Fish Physiology* (Vol. 27, pp. 1–23). Academic Press. https://doi.org/https://doi.org/10.1016/S1546-5098(08)00001-0

Diaz, R. J., & Rosenberg, R. (2008). Spreading Dead Zones and Consequences for Marine Ecosystems. *Science*, *321*(5891), 926–929. https://doi.org/10.1126/science.1156401

Drenkard, E.J., Stock, C., Ross, A.C., Dixon, K.W., Adcroft, A., Alexander, M., Balaji, V., Bograd, S.J., Butenschön, M., Cheng, W. & Curchitser, E. (2021). Next-generation regional ocean projections for living marine resource management in a changing climate. *ICES Journal of Marine Science,* 78(6), 1969-1987.

Eilers, P.H.C. & Marx, B.D. (2021) *Practical Smoothing: The Joys of P-splines.* Cambridge University Press.

Essington, T. E., Anderson, S. C., Barnett, L. A. K., Berger, H. M., Siedlecki, S. A., & Ward, E. J. (2022). Advancing statistical models to reveal the effect of dissolved oxygen on the spatial distribution of marine taxa using thresholds and a physiologically based index. *Ecography*, *2022*(8), e06249.

Farrell, A. P., & Richards, J. G. (2009). Defining hypoxia: an integrative synthesis of the responses of fish to hypoxia. In *Fish physiology* (Vol. 27, pp. 487-503). Academic Press.

Franco, A. C., Ianson, D., Ross, T., Hamme, R., Monahan, A., Christian, J., Davelaar, M., William, J., Miller, L., Robert, M., & Tortell, P. D. (2021). *A compilation of inorganic carbon system and other hydrographic and chemical discrete profile measurements obtained during the fifty five Line P cruises in the Northeast Pacific Ocean over the period from 1990 to 2019 (NCEI Accession 0234342)*. NOAA National Centers for Environmental Information. https://doi.org/10.25921/zrw8-kn24

Franco, A. C., Kim, H., Frenzel, H., Deutsch, C., Ianson, D., Sumaila, U. R., & Tortell, P. D. (2022). Impact of warming and deoxygenation on the habitat distribution of Pacific halibut in the Northeast Pacific. *Fisheries Oceanography*, *31*(6), 601–614. https://doi.org/10.1111/fog.12610

Fredston, A.L., Cheung, W.W., Frölicher, T.L., Kitchel, Z.J., Maureaud, A.A., Thorson, J.T., Auber, A., Mérigot, B., Palacios-Abrantes, J., Palomares, M.L.D. and Pecuchet, L., 2023. Marine heatwaves are not a dominant driver of change in demersal fishes. *Nature*, *621*(7978): 324-329. https://doi.org/10.1038/s41586-023-06449-y

Fry, F.E.J. 1971. The effect of environmental factors on the physiology of fish. *Fish Physiology* 6: 1-98.

Gillooly, J.F., Brown, J., West, G., Savage, V.M, & Charnov, E.L. (2001). Effects of size and temperature on metabolic rate. *Science,* 293(5538): 2248-2251. https:doi.org/10.1126/science.1061967

Gray, J. S., Wu, R. S., & Or, Y. Y. (2002). Effects of hypoxia and organic enrichment on the coastal marine environment. *Marine Ecology Progress Series*, *238*, 249–279.

Gruber, N., Clement, D., Carter, B., Feely, R., Van Heuven, S., Hoppema, M., Ishii, M., Key, R., Kozyr, A., Lauvset, S., Monaco, C., Mathis, J., Murata, A., Olsen, A., Perez, F., Sabine, C., and Wanninkhof, R. 2019. The oceanic sink for anthropogenic CO_2 from 1994 to 2007. Science, *363*(6432): 1193-1199. https://doi.org/ 10.1126/science.aau5153

Grüss, A., Thorson, J. T., Stawitz, C. C., Reum, J. C. P., Rohan, S. K., & Barnes, C. L. (2021). Synthesis of interannual variability in spatial demographic processes supports the strong influence of cold-pool extent on eastern Bering Sea walleye pollock (Gadus chalcogrammus). *Progress in Oceanography*, *194*. https://doi.org/10.1016/j.pocean.2021.102569

Hao, T., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2020). Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography*, *43*(4), 549–558.

Harris, J., Pirtle, J. L., Laman, E. A., Siple, M. C., & Thorson, J. T. (2024). An ensemble approach to species distribution modelling reconciles systematic differences in estimates of habitat utilization and range area. *Journal of Applied Ecology*, *61*(2), 351–364. https://doi.org/https://doi.org/10.1111/1365-2664.14559

Hauri, C., Schultz, C., Hedstrom, K., Danielson, S., Irving, B., Doney, S. C., Dussin, R., Curchitser, E. N., Hill, D. F., & Stock, C. A. (2020). A regional hindcast model simulating ecosystem dynamics, inorganic carbon chemistry, and ocean acidification in the Gulf of Alaska. *Biogeosciences*, *17*(14), 3837–3857.

Hay, L. E., Wilby, R. L., & Leavesley, G. H. (2000). A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States. *Journal of the American Water Resources Association*, 36(2), 387-397. https://doi.org/10.1111/j.1752-1688.2000.tb04276.x

Hoff, G. R. 2016. Results of the 2016 eastern Bering Sea upper continental slope survey of groundfish and invertebrate resources. U.S. Department of Commerce, NOAA Tech. Memo. NMFS-AFSC-339. http://doi.org/10.7289/V5/TM-AFSC-339

Indivero, J., Anderson, S., Barnett, L., Essington, T. E., & Ward, E. (2024). Estimating a physiological threshold to oxygen and temperature from marine monitoring data reveals challenges and opportunities for forecasting distribution shifts. Ecography, 2025(4): e07413https://doi.org/10.22541/au.172446072.20202938/v1

IPHC. (2022). *IPHC Water Column Profiler Deployment Procedures and Maintenance Guide*. https://www.iphc.int/uploads/pdf/manuals/2022/iphc-2022-wcpm-001.pdf

IPHC. (2024). *Water Column Profiler Data*. https://www.iphc.int/data/water-column-profiler-data/

Joyce, S. (2000). The dead zones: oxygen-starved coastal waters. *Environmental Health Perspectives*, *108*(3), A120–A125. https://doi.org/10.1289/ehp.108-a120

Kearney, K., Hermann, A., Cheng, W., Ortiz, I., & Aydin, K (2020). A coupled pelagic-benthic-sympagic biogeochemical model for the Bering Sea: documentation and validation of the BESTNPZ model (v2019.08.23) within a high-resolution regional ocean model. Geosci Model Dev 13:597–650. https://doi.org/10.5194/gmd-13-597-2020

Kearney, K. (2021). *Temperature Data from the Eastern Bering Sea Continental Shelf Bottom Trawl Survey as Used for Hydrodynamic Model Validation and Comparison*. https://doi.org/10.25923/e77k-gg40

Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, *12*(4), 334–350. https://doi.org/https://doi.org/10.1111/j.1461-0248.2008.01277.x

Keeling, R. F., Körtzinger, A., & Gruber, N. (2010). Ocean deoxygenation in a warming world. *Annual Review of Marine Science*, *2*(1), 199–229.

Keller, A. A., Wallace, J. R., & Methot, R. D. (2017). *The northwest fisheries science center's west coast groundfish bottom trawl survey: history, design, and description*.

Kim, H., Franco, A. C., & Sumaila, U. R. (2023). A selected review of impacts of ocean deoxygenation on fish and fisheries. *Fishes*, *8*(6), 316.

Kramer, D. L. (1987). Dissolved oxygen and fish behavior. *Environmental Biology of Fishes*, *18*(2), 81–92. https://doi.org/10.1007/BF00002597

Kristiansen, T., Butenschön, M., & Peck, M. A. (2024). Statistically downscaled CMIP6 ocean variables for European waters. *Scientific Reports*, *14*(1), 1209. https://doi.org/10.1038/s41598-024-51160-1

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., Dunne, J. P., Gehlen, M., Ilyina, T., & John, J. G. (2020). Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. *Biogeosciences*, *17*(13), 3439–3470.

Leeuwis, R.H., Nash, G.W., Sandrelli, R. M., Zanuzzo F.S., and Gamperl, A.K. (2019). The environmental tolerances and metabolic physiology of sablefish (*Anoplopoma fimbria*).

Comparative Biochemistry and Physiology, Part A, 231: 140-148. https://doi.org/10.1016/j.cbpa.2019.02.004.

Li, C., Huang, J., Liu, X., Ding, L., He, Y., & Xie, Y. (2024). The ocean losing its breath under the heatwaves. *Nature Communications*, *15*(1), 6840. https://doi.org/10.1038/s41467-024-51323-8

Li, J., & Heap, A. D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, *53*, 173–189. https://doi.org/10.1016/j.envsoft.2013.12.008

Lindén, A., & Knape, J. (2009). Estimating environmental effects on population dynamics: consequences of observation error. *Oikos*, *118*(5), 675–680. https://doi.org/https://doi.org/10.1111/j.1600-0706.2008.17250.x

Lindmark, M., Anderson, S. C., Gogina, M., & Casini, M. (2023). Evaluating drivers of spatiotemporal variability in individual condition of a bottom-associated marine fish, Atlantic cod (Gadus morhua). *ICES Journal of Marine Science*, *80*(5), 1539–1550. https://doi.org/10.1093/icesjms/fsad084

Lindmark, M., Audzijonyte, A., Blanchard, J. L., & Gårdmark, A. (2022a). Temperature impacts on fish physiology and resource abundance lead to faster growth but smaller fish sizes and yields under warming. *Global Change Biology*, *28*(21), 6239–6253. https://doi.org/10.1111/gcb.16341

Lindmark, M., Ohlberger, J., & Gårdmark, A. (2022b). Optimum growth temperature declines with body size within fish species. *Global Change Biology*, *28*(7), 2259–2271.Little, R. J. A., & Rubin, D. B. (2020). *Statistical analysis with missing data* (Third edition). Wiley.

Liu, O. R., Ward, E. J., Anderson, S. C., Andrews, K. S., Barnett, L. A. K., Brodie, S., Carroll, G., Fiechter, J., Haltuch, M. A., Harvey, C. J., Hazen, E. L., Hernvann, P. Y., Jacox, M., Kaplan, I. C., Matson, S., Norman, K., Buil, M. P., Selden, R. L., Shelton, A., & Samhouri, J. F. (2023). Species redistribution creates unequal outcomes for multispecies fisheries under projected climate change. *Science Advances*, *9*(33). https://doi.org/10.1126/sciadv.adg5468

Markowitz, E. H., Dawson, E. J., Wassermann, S. N., Anderson, C. B., Rohan, S. K., Charriere, N. E., & Stevenson, D. E. (2024). Results of the 2023 eastern and northern Bering Sea continental shelf bottom trawl survey of groundfish and invertebrate fauna. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-AFSC-487. https://doi.org/10.25923/2mry-yx09

Matear, R. J., & Hirst, A. C. (2003). Long-term changes in dissolved oxygen concentrations in the ocean caused by protracted global warming. *Global Biogeochemical Cycles*, *17*(4). https://doi.org/https://doi.org/10.1029/2002GB001997

McCarthy, M. J., Newell, S. E., Carini, S. A., & Gardner, W. S. (2015). Denitrification dominates sediment nitrogen removal and is enhanced by bottom-water hypoxia in the Northern Gulf of Mexico. *Estuaries and Coasts*, *38*, 2279–2294.

McGarry, K., Siedlecki, S., Salisbury, J., and Alin, S. 2020. Multiple linear regression models for reconstructing and exploring processes controlling the carbonate system of the Northeast UW from basic hydrographic data. *Journal of Geophysical Research: Oceans, 126*(2): e2020JC016480. https://doi.org/10.1029/2020JC016480

Mercator-Ocean. (2024). *Global Ocean Biogeochemistry Hindcast (v3.6_STABLE)*. Https://Doi.Org/10.48670/Moi-00019.

Mignot, A., D'Ortenzio, F., Taillandier, V., Cossarini, G., & Salon, S. (2019). Quantifying Observational Errors in Biogeochemical-Argo Oxygen, Nitrate, and Chlorophyll a Concentrations. *Geophysical Research Letters*, *46*(8), 4330–4337. https://doi.org/https://doi.org/10.1029/2018GL080541

Moore, C., Drazen, J. C., Radford, B. T., Kelley, C., & Newman, S. J. (2016). Improving essential fish habitat designation to support sustainable ecosystem-based fisheries management. *Marine Policy*, *69*, 32–41.

Morée, A. L., Clarke, T. M., Cheung, W. W. L., & Frölicher, T. L. (2023). Impact of deoxygenation and warming on global marine species in the 21st century. *Biogeosciences*, *20*(12), 2425–2454.

Nakagawa, S. 2015. Chapter 4 Missing data: mechanisms, methods, and messages, in Gordon A. Fox, Simoneta Negrete-Yankelevich, and Vinicio J. Sosa (eds), *Ecological Statistics: Contemporary theory and application*: 81-105. https://doi-org.offcampus.lib.washington.edu/10.1093/acprof:oso/9780199672547.003.0005, accessed 6 June 2025.

Newman, L., Heil, P., Trebilco, R., Katsumata, K., Constable, A., Van Wijk, E., Assmann, K., Beja, J., Bricher, P., Coleman, R. and Costa, D. (2019). Delivering sustained, coordinated, and integrated observations of the Southern Ocean for global impact. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00433

NOAA. (2020). *Oceanographic and Ecosystem Sampling During the Pacific Hake Survey*. https://www.fisheries.noaa.gov/west-coast/science-data/oceanographic-and-ecosystem-sampling-during-pacific-hake-survey#measuring-conductivity-temperature-and-depth

NOAA. (2021a). *Coastal Ocean Data Analysis Product in North America (CODAP-NA, Version 2021) (NCEI Accession 0219960)*. Coastal Ocean Data Analysis Product in North America (CODAP-NA, Version 2021) (NCEI Accession 0219960)

NOAA. (2021b). *NOAA West Coast Ocean Acidification Cruises (WCOA) and Harmful Algal Blooms cruise (NOAA HABs)*. https://www.ncei.noaa.gov/access/ocean-carbon-acidification-data-system/oceans/Coastal/WCOA.html

NOAA. (2024a). *Integrated West Coast Pelagics Survey*. https://www.fisheries.noaa.gov/west-coast/science-data/integrated-west-coast-pelagics-survey

NOAA. (2024b). *National Oceanic and Atmospheric Administration Budget Estimates Fiscal Year 2025*. https://www.noaa.gov/sites/default/files/2024-03/NOAA_FY25_Congressional_Justification.pdf

NOAA Fisheries. (2024). *NOAA Fisheries Announces Changes in its Alaska Survey Portfolio*. https://www.fisheries.noaa.gov/feature-story/noaa-fisheries-announces-changes-its-alaska-survey-portfolio

Oke, K. B., Mueter, F., & Litzow, M. A. (2022). Warming leads to opposite patterns in weight-at-age for young versus old age classes of Bering Sea walleye pollock. *Canadian Journal of Fisheries and Aquatic Sciences*, *79*(10), 1655–1666. https://doi.org/10.1139/cjfas-2021-0315

Oschlies, A., Brandt, P., Stramma, L., & Schmidtko, S. (2018). Drivers and mechanisms of ocean deoxygenation. *Nature Geoscience*, *11*(7), 467–473. https://doi.org/10.1038/s41561-018-0152-2

Pena, M. A., Katsev, S., Oguz, T., & Gilbert, D. (2010). Modeling dissolved oxygen dynamics and hypoxia. *Biogeosciences*, *7*(3), 933–957.

Perruche, C., Szczypta, C., Paul, J., & Drévillon, M. (2024). *Quality Information Document Global Production Centre GLOBAL_MULTIYEAR_BGC_001_029 Issue 1.2*. https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-GLO-QUID-001-029.pdf

Pörtner, H. O., & Knust, R. (2007). Climate Change Affects Marine Fishes Through the Oxygen Limitation of Thermal Tolerance. *Science*, *315*(5808), 95–97. https://doi.org/10.1126/science.1135471

Pörtner, H.-O. (2010). Oxygen-and capacity-limitation of thermal tolerance: a matrix for integrating climate-related stressor effects in marine ecosystems. *Journal of Experimental Biology*, *213*(6), 881–893.

Pörtner, H.-O., Bock, C., & Mark, F. C. (2017). Oxygen- and capacity-limited thermal tolerance: bridging ecology and physiology. *Journal of Experimental Biology*, *220*(15), 2685–2696. https://doi.org/10.1242/jeb.134585

Pozo Buil, M., Jacox, M. G., Fiechter, J., Alexander, M. A., Bograd, S. J., Curchitser, E. N., Edwards, C. A., Rykaczewski, R. R., & Stock, C. A. (2021). A dynamically downscaled

ensemble of future projections for the California current system. *Frontiers in Marine Science*, *8*, 612874. https://doi.org/10.3389/fmars.2021.612874

Risien, C., Hough, K., Waddell, J., Fewings, M., & Cervantes, B. (2024a). *Conductivity–Temperature–Depth (CTD) and dissolved oxygen profile data from shipboard surveys collected within Olympic Coast National Marine Sanctuary, 2005-2023 (v1.0)*. [Data Set]. Zenodo. https://doi.org/10.5281/zenodo.10466124

Risien, C., Hough, K., Waddell, J., Fewings, M., & Cervantes, B. (2024b). *Shipboard Conductivity–Temperature–Depth (CTD) and dissolved oxygen profile data collected during hypoxia surveys along six hydrographic sampling lines within Olympic Coast National Marine Sanctuary, 2004–2015 (v1.0)*. [Data Set]. Zenodo. https://doi.org/10.5281/zenodo.11167853

Risien, C. M., Cervantes, B. T., Fewings, M. R., Barth, J. A., & Kosro, P. M. (2023). A stitch in time: Combining more than two decades of mooring data from the central Oregon shelf. *Data in Brief*, *48*, 109041. https://doi.org/https://doi.org/10.1016/j.dib.2023.109041

Risien, C. M., Fewings, M. R., Fisher, J. L., Peterson, J. O., & Morgan, C. A. (2022). Spatially gridded cross-shelf hydrographic sections and monthly climatologies from shipboard survey data collected along the Newport Hydrographic Line, 1997–2021. *Data in Brief*, *41*, 107922. https://doi.org/https://doi.org/10.1016/j.dib.2022.107922

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929. https://doi.org/https://doi.org/10.1111/ecog.02881

Rohan, S. K., Charriere, N. E., Riggle, B., O'Leary, C. A., & Raring, N. W. (2024). *A Flexible approach for processing data collected using trawl-mounted CTDs during Alaska bottom-trawl surveys*. https://repository.library.noaa.gov/view/noaa/55614

Ross, A. C., Stock, C. A., Adcroft, A., Curchitser, E., Hallberg, R., Harrison, M. J., Hedstrom, K., Zadeh, N., Alexander, M., & Chen, W. (2023). A high-resolution physical-biogeochemical model for marine resource applications in the Northwest Atlantic (MOM6-COBALT-NWA12 v1. 0). *Geoscientific Model Development Discussions*, *2023*, 1–65. https://doi.org/10.5194/gmd-16-6943-2023

Rubalcaba, J. G., Verberk, W. C. E. P., Hendriks, A. J., Saris, B., & Woods, H. A. (2020). Oxygen limitation may affect the temperature and size dependence of metabolism in aquatic ectotherms. *Proceedings of the National Academy of Sciences*, *117*(50), 31963–31968. https://doi.org/10.1073/pnas.2003292117

Schmidtko, S., Stramma, L., & Visbeck, M. (2017). Decline in global oceanic oxygen content during the past five decades. *Nature*, *542*(7641), 335–339. https://doi.org/10.1038/nature21399

Sharp, J. D., Fassbender, A. J., Carter, B. R., Johnson, G. C., Schultz, C., & Dunne, J. P. (2022). GOBAI-O 2: temporally and spatially resolved fields of ocean interior dissolved oxygen over nearly two decades. *Earth System Science Data Discussions*, *2022*, 1–46. https://doi.org/10.5194/essd-15-4481-2023

Siple, M. C., von Szalay, P. G., Raring, N. W., Dowlin, A. N., and Riggle, B. C. (2024). Data Report: 2023 Gulf of Alaska bottom trawl survey. AFSC Processed Rep. 2024-09. Alaska Fisheries Science Center, NOAA, National Marine Fisheries Service. https://doi.org/10.25923/gbb1-x748

Steel, E. A., McElhany, P., Yoder, N. J., Purser, M. D., Malone, K., Thompson, B. E., Avery, K. A., Jensen, D., Blair, G., Busack, C., Bowen, M. D., Hubble, J., & Kantz, T. (2009). Making the Best Use of Modeled Data: Multiple Approaches to Sensitivity Analysis of a Fish-Habitat Model. *Fisheries*, *34*(7), 330–339. https://doi.org/https://doi.org/10.1577/1548-8446-34.7.330

Stramma, L., Oschlies, A., & Schmidtko, S. (2012). Mismatch between observed and modeled trends in dissolved upper-ocean oxygen over the last 50 yr. *Biogeosciences*, *9*(10), 4045–4057. https://doi.org/10.5194/bg-9-4045-2012

Stramma, L., Schmidtko, S., Levin, L. A., & Johnson, G. C. (2010). Ocean oxygen minima expansions and their biological impacts. *Deep Sea Research Part I: Oceanographic Research Papers*, *57*(4), 587–595. https://doi.org/https://doi.org/10.1016/j.dsr.2010.01.005

Sullaway, G., Cunningham, C. J., Kimmel, D., Pilcher, D. J., & Thorson, J. T. (2025). Evaluating the performance of a system model in predicting zooplankton dynamics: Insights from the Bering Sea ecosystem. *Fisheries Oceanography*, 34(1), e12691. https://doi.org/10.1111/fog.12691

Sydeman, W. J., García-Reyes, M., Schoeman, D. S., Rykaczewski, R. R., Thompson, S. A., Black, B. A., & Bograd, S. J. (2014). Climate change and wind intensification in coastal upwelling ecosystems. *Science*, *345*(6192), 77–80. http://www.jstor.org.offcampus.lib.washington.edu/stable/24744817

Thompson, P. L., Nephin, J., Davies, S. C., Park, A. E., Lyons, D. A., Rooper, C. N., Angelica Peña, M., Christian, J. R., Hunter, K. L., & Rubidge, E. (2023a). Groundfish biodiversity change in northeastern Pacific waters under projected warming and deoxygenation. *Philosophical Transactions of the Royal Society B*, *378*(1881), 20220191. https://doi.org/10.1098/rstb.2022.0191

Thompson, P. L., Rooper, C. N., Nephin, J., Park, A. E., Christian, J. R., Davies, S. C., Hunter, K., Lyons, D. A., Peña, M. A., Proudfoot, B., Rubidge, E. M., & Holdsworth, A. M. (2023b). Response of Pacific halibut (Hippoglossus stenolepis) to future climate scenarios in the Northeast Pacific Ocean. *Fisheries Research*, *258*, 106540. https://doi.org/https://doi.org/10.1016/j.fishres.2022.106540

Thorson, J.T. (2019). Measuring the impact of oceanographic indices on species distribution shifts: The spatially varying effect of cold-pool extent in the eastern Bering Sea. *Limnology and Oceanography, 64*(6): 2632-2645. https://doi.org/10.1002/lno.11238

Thorson, J.T., Arimitsu, M.L., Barnett, L.A.K., Cheng, W., Eisner, L.B., Haynie, A. C., Hermann, A.J., Holsman, K., Kimmel, D. G., Lomas, M. W., Richar, J., & Siddon, E. (2021). Forecasting community reassemblage using climate-linked spatio-temporal ecosystem models. *Ecography, 44*(4): 612-625. https://doi.org/10.1111/ecog.05471

Thorson, J. T., Andrews, A. G. III, Essington, T. E., & Large, S. I. (2024). Dynamic structural equation models synthesize ecosystem dynamics constrained by ecological mechanisms. *Methods in Ecology and Evolution* 15: 744–755. https://doi.org/10.1111/2041-210X.14289

Townhill, B. L., Couce, E., Tinker, J., Kay, S., & Pinnegar, J. K. (2023). Climate change projections of commercial fish distribution and suitable habitat around north western Europe. *Fish and Fisheries*, *24*(5), 848–862. https://doi.org/https://doi.org/10.1111/faf.12773

Ultsch, G.R., Boschung, H., and Ross, M. 1978. Metabolism, critical oxygen tension, and habitat selection in Darters (Etheostoma). *Ecology*, 59(1):99-107. https://doi.org/10.2307/1936635

Valenza, A. N., Altenritter, M. E., & Walther, B. D. (2023). Reconstructing consequences of lifetime hypoxia exposure on metabolism of demersal fish in the northern Gulf of Mexico using otolith chemistry. *Environmental Biology of Fishes*, *106*(11), 2045–2057. https://doi.org/10.1007/s10641-023-01483-1

Vaquer-Sunyer, R., & Duarte, C. M. (2008). Thresholds of hypoxia for marine biodiversity. *Proceedings of the National Academy of Sciences*, *105*(40), 15452–15457. https://doi.org/10.1073/pnas.0803833105

Vaquer-Sunyer, R., & Duarte, C. M. (2011). Temperature effects on oxygen thresholds for hypoxia in marine benthic organisms. *Global Change Biology*, *17*(5), 1788–1797. https://doi.org/https://doi.org/10.1111/j.1365-2486.2010.02343.x

Verberk, W. C. E. P., Overgaard, J., Ern, R., Bayley, M., Wang, T., Boardman, L., & Terblanche, J. S. (2016). Does oxygen limit thermal tolerance in arthropods? A critical review of current evidence. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *192*, 64–78.

Verezemskaya, P., Barnier, B., Gulev, S. K., Gladyshev, S., Molines, J., Gladyshev, V., Lellouche, J., & Gavrikov, A. (2021). Assessing Eddying (1/12) ocean reanalysis GLORYS12 using the 14-yr instrumental record from 59.5 N section in the Atlantic. *Journal of Geophysical Research: Oceans*, *126*(6), e2020JC016317.

Wang, Z., Xue, C., & Ping, B. (2024). A Reconstructing Model Based on Time–Space–Depth Partitioning for Global Ocean Dissolved Oxygen Concentration. *Remote Sensing*, *16*(2), 228. https://doi.org/10.3390/rs16020228

Ward, E. J., Hunsicker, M. E., Marshall, K. N., Oken, K. L., Semmens, B. X., Field, J. C., Haltuch, M. A., Johnson, K. F., Taylor, I. G., Thompson, A. R., & Tolimieri, N. (2024). Leveraging ecological indicators to improve short term forecasts of fish recruitment. *Fish and Fisheries*, 25, 895–909. https://doi.org/10.1111/faf.12850