

# ECOGRAPHY

## Research article

### Estimating a physiological threshold to oxygen and temperature from marine monitoring data reveals challenges and opportunities for forecasting distribution shifts

Julia Indivero<sup>1</sup>✉, Sean C. Anderson<sup>2</sup>, Lewis A. K. Barnett<sup>3</sup>, Timothy E. Essington<sup>1</sup> and Eric J. Ward<sup>4</sup>

<sup>1</sup>University of Washington, School of Aquatic and Fishery Sciences, Seattle, WA, USA

<sup>2</sup>Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC, Canada

<sup>3</sup>Resource Assessment and Conservation Engineering Division, Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA, USA

<sup>4</sup>Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA, USA

Correspondence: Julia Indivero ([jindiv@uw.edu](mailto:jindiv@uw.edu))

#### Ecography

2024: e07413

doi: [10.1111/ecog.07413](https://doi.org/10.1111/ecog.07413)

Subject Editor:

Michael Ray Kearney

Editor-in-Chief: Miguel Araújo

Accepted 31 October 2024



Species distribution modeling is increasingly used to describe and anticipate consequences of a warming ocean. These models often identify statistical associations between distribution and environmental conditions such as temperature and oxygen, but rarely consider the mechanisms by which these environmental variables affect metabolism. Oxygen and temperature jointly govern the balance of oxygen supply to oxygen demand, and theory predicts thresholds below which population densities are diminished. However, parameterizing models with this joint dependence is challenging because of the paucity of experimental work for most species, and the limited applicability of experimental findings in situ. Here we ask whether the temperature-sensitivity of oxygen can be reliably inferred from species distribution observations in the field, using the U.S. Pacific Coast as a model system. We developed a statistical model that adapted the metabolic index — a compound metric that incorporates these joint effects on the ratio of oxygen supply and oxygen demand by applying an Arrhenius equation — and used a non-linear threshold function to link the index to fish distribution. Through simulation testing, we found that our statistical model could not precisely estimate the parameters due to inherent features of the distribution data. However, the model reliably estimated an overall metabolic index threshold effect. When applied to case studies of real data for two groundfish species, this new model provided a better fit to spatial distribution of one species, sablefish *Anoplopoma fimbria*, than previously used models, but did not for the other, longspine thornyhead *Sebastolobus altivelis*. This physiological framework may improve predictions of species distribution, even in novel environmental conditions. Further efforts to combine insights from physiology and realized species distributions will improve forecasts of species' responses to future environmental changes.



[www.ecography.org](http://www.ecography.org)

© 2024 The Author(s). Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

## Introduction

Range shifts of aquatic and terrestrial species are widespread, although the velocity and magnitude vary with taxa and system (Lenoir et al. 2020). Anticipating how species distributions might shift under future climate conditions can enable more effective conservation and natural resource actions. Predicting these range changes can inform disease and parasite risk analysis (Polley and Thompson 2009, Rohr and Cohen 2020), invasive species control (Beaury et al. 2020, Smith et al. 2022), protected area design (Monzón et al. 2011, Hoffmann et al. 2019), and harvest management (Baudron et al. 2020, Bell et al. 2020).

However, there are limits to our ability to forecast species range shifts. For instance, correlative species distribution models (SDMs), a type of habitat suitability model (Hirzel and Le Lay 2008), have been a common approach to predict species responses to climate change (Franklin 2010, Warton and Shepherd, 2010, Swanson et al. 2013, Melo-Merino et al. 2020). These models identify historical associations between species densities and relevant environmental conditions to explain past shifts and forecast future shifts. Relying on correlations from the past may provide unskillful forecasts as climate change pushes environmental conditions far outside those seen in the past (Williams and Jackson 2007, Marshall et al. 2017, Rollinson et al. 2021, Smith et al. 2022). Ecological systems may have novel responses to these anomalous conditions (Williams and Jackson 2007, Hobbs et al. 2009) and often exhibit changing relationships over time (Myers 1998, Asch et al. 2022). For instance, the temperature-recruit relationship for Pacific sardine *Sardinops sagax* used in operational stock assessments was no longer significant when re-evaluated with more recent data (McClatchie et al. 2010). Predictions based on historical relationships may consequently break down when used in forecasting (Veloz et al. 2012, Muhling et al. 2020, Barnes et al. 2022), especially in novel environments (Waldock et al. 2022).

Incorporating mechanistic relationships into SDMs is one approach to include biological realism and improve inference and forecasting ability (Kearney and Porter 2009, Martínez et al. 2015, Urban et al. 2016, Urban 2019, Briscoe et al. 2023). For marine taxa, one possible mechanistic pathway for predicting responses to climate change is considering the joint effect of temperature and oxygen on metabolism (Methorst et al. 2017, Duncan et al. 2020). Due to climate change, ocean temperatures are projected to increase (Cheng et al. 2019), and extreme heatwave events are becoming more frequent (Frölicher et al. 2018, Laufkötter et al. 2020). There will also likely be widespread oxygen depletion (Diaz and Rosenberg 2008). Temperature governs metabolic rate and therefore oxygen demand (Gillooly et al. 2001, Brown et al. 2004). Temperature and oxygen may therefore jointly govern tolerances to environmental conditions through these metabolic requirements

(Pörtner and Knust 2007, Pörtner 2010, Verberk et al. 2016a, Pörtner et al. 2017, Leiva et al. 2019), though there is ongoing debate on the exact way and extent to which limitations may occur (Verberk et al. 2016b, 2021, Lefevre et al. 2018, Jutfelt et al. 2018, Seibel and Deutsch 2020, Scheuffele et al. 2021, Atkinson et al. 2022). Yet overall, parameterizing the interactive effect of temperature and oxygen in the metabolic index, rather than separately or linearly, can provide a framework rooted in physiological theory to explain historical effects of temperature and oxygen on distributions. This can help better identify habitats important for fish population processes (i.e. for designating essential fish habitat; Rosenberg et al. 2000, Moore et al. 2016) and predict future range shifts, contractions or expansions.

One way to capture how the interaction between oxygen and temperature shapes species' distribution is the 'metabolic index' ( $\Phi$ ) a compound metric that incorporates the joint effects of temperature and oxygen on the ratio of oxygen supply to oxygen demand (Deutsch et al. 2015). Theory predicts a ratio below which aquatic taxa are likely to avoid due to insufficient temperature–oxygen conditions for metabolic functioning (Deutsch et al. 2015). We recognize that there are other possible frameworks for combining the effects of oxygen and temperature on fish (e.g. the aerobic growth index in Clarke et al. 2021, 2022, Moreé et al. 2023, Ern 2019), yet comparison of different frameworks is beyond the scope of this study. The metabolic index was found to closely align with species distributions across biogeographic scales to define species' range limits (Deutsch et al. 2020, Sunday et al. 2022) and has been used to explain contemporary distributions (Franco et al. 2022, Penn and Deutsch 2024), past extinction events (Penn and Deutsch 2022), and predict future marine habitat shifts (Chen et al. 2024).

Using the metabolic index (or a similar ecophysiological metric) in statistical species distribution models may improve predictions of distribution shifts under future climate regimes. However, implementing this approach is often challenged by the paucity of information on how the ratio of oxygen demand and supply scale with temperature. Most data to estimate the metabolic index are derived from laboratory studies. Because of the time and logistical demands of such studies, such information is only available for a small number of marine taxa. A recent empirical meta-analysis used taxonomic imputation of data from 74 marine taxa to evaluate the parameters of the metabolic index across species (Essington et al. 2024). This approach (Penone et al. 2014, Debastiani et al. 2021, Thorson et al. 2021b) is a rigorous statistical model that accounts for phylogenetic relationships to predict traits for unmeasured species. The empirical meta-analysis found that the temperature-sensitivity of oxygen tolerance was highly variable across taxa and could not be precisely predicted for unmeasured species from existing data. Moreover, the applicability of experimental findings to explain distributions in situ is limited (Essington et al. 2022,

Bandara et al. 2023) owing to the wide range of physiological acclimations and potential for local adaptation.

Here, we evaluate the reliability of a new approach to estimate a physiologically-based (i.e. metabolic index) temperature-dependence of oxygen tolerance from species distribution data within a statistical model, using the U.S. Pacific coast as a model system. This approach aims to overcome limitations of laboratory studies by estimating this component of the metabolic index (the temperature-dependence of oxygen tolerance) from widely available data (e.g. species distribution data collected from the field). While our approach is not directly mechanistic, we aim to 1) improve existing methods by basing the response of fish to temperature and oxygen in a more physiologically and ecologically realistic relationship than considering temperature and oxygen independently or linearly, and 2) illustrate data limitations and gaps in estimating metabolic responses.

We first developed a statistical model that simultaneously estimates relevant metabolic index parameters and estimates the effect of temperature and oxygen (metabolic index) on local fish density (i.e. estimates parameters that describe this function). We then used simulation testing to evaluate whether our new model could accurately recover parameters, a standard practice when developing quantitative methods. For this, we simulated data of fish density and distribution from our model with specified parameters, fit our new model, and then compared estimated parameter estimates with the known values (i.e. those used in simulation). Finally, we fit the model to real-world data of two Pacific coast demersal fish species as case studies to evaluate model performance and behavior, and to demonstrate how the model can be used to evaluate the extent to which oxygen, temperature, or the metabolic index governs fish distributions.

## Methods

### Model structure

We modified a geostatistical generalized linear mixed effects model (GLMM) widely used for correlative SDMs that estimates local population densities as a function of observed and latent environmental variables. This approach is generic and could be extended to other models consistent with a GLMM framework. These modifications were twofold. First, we introduced an additional term that would allow for an asymptotic effect of oxygen on local abundance, which recognizes that oxygen is a limiting factor. Second, we accounted for the interactive effect of pO<sub>2</sub> and temperature as described by the metabolic index. We describe these steps in detail below and provide a graphical summary of the modeling framework in Fig. 1.

The SDM is a generalized linear mixed effects model (Eq. 1) that estimates the expected density  $\mu$  of observation  $i$  with a log link

$$\log(\mu_i) = b_{\text{year}[i]}t_i + b_2d_i + b_3d_i^2 + \omega_i \quad (1)$$

where  $b_{\text{year}}$  are independent year effects for each year  $t_i$ ,  $b_2$  and  $b_3$  are the estimated effects of scaled log(depth) ( $d_i$ ) and its square ( $d_i^2$ ), and  $\omega_i$  is a spatial random effect that accounts for spatially structured latent variables. Depth is included because fish have depth preferences separate from other environmental conditions (e.g. sablefish in Sogard and Berkely 2017). We model spatial effects using the stochastic partial differential equation (SPDE) approximation to Gaussian random fields via Gaussian Markov random fields (GMRFs) (Lindgren et al. 2011) as implemented in TMB (Kristensen et al. 2016). The random effects describing the spatial field are estimated at a set of vertices or knots and then projected to the locations of observed data with bilinear interpolation, assuming spatial covariance is modeled using a GMRF with a Matérn covariance function and anisotropy. The approach allows modelling a sparse precision matrix of a GMRF, which is considerably more efficient than modelling the covariance matrix of a Gaussian random field directly (Lindgren et al. 2011). Given that observations in catch-per-unit-effort (CPUE) data typically include positive continuous values and zeros, we assumed the observed process followed a Tweedie distribution (Shono 2008, Tweedie 1984) so that the  $i$ th observation is drawn from a Tweedie distribution with mean  $\mu_i$ , power parameter  $p$ , and scale parameter  $\sigma$ .

We made two adjustments to the linear structure of the above model (Eq. 1), depicted visually in Fig. 1. One, the model uses oxygen and temperature and estimates the parameter  $E_0$  to generate the portion of the metabolic index (the temperature-dependence of oxygen tolerance) that can be uniquely estimated in this statistical modelling framework ( $\Phi_{\text{eco},i}$ ) (Supporting information). The symbol  $\Phi_{\text{eco}}$  is defined as the terms of the metabolic index that depend on pO<sub>2</sub> and temperature (see the Supporting information for derivation):

$$\Phi_{\text{eco},i} = \text{pO}_{2,i} \exp \left[ \frac{E_0}{k_b} \left( \frac{1}{T_i} - \frac{1}{T_{\text{ref}}} \right) \right] \quad (2)$$

where  $E_0$  is an estimated parameter that describes the temperature dependence of the ratio of oxygen supply to demand,  $k_b$  is Boltzmann's constant, and  $T_{\text{ref}}$  is a reference temperature (here chosen as 12°C). Both  $T$  and  $T_{\text{ref}}$  are in Kelvin units in Eq. 2.

Second, we defined a function  $f(\Phi_{\text{eco},i})$  that takes the calculated  $\Phi_{\text{eco},i}$  and estimates parameters that governs the predicted mean density at a sample site, and is added to the suite of linear predictors in Eq. 1:

$$\log(\mu_i) = b_{\text{year}[i]}t_i + b_2d_i + b_3d_i^2 + \omega_i + f(\Phi_{\text{eco},i}) \quad (3)$$

Here,  $f(\Phi_{\text{eco},i})$ , when exponentiated, acts as a multiplier to diminish predicted mean density when oxygen and temperature conditions are estimated to be physiologically stressful. We expected this to be a monotonically increasing function that asymptotes at 1, i.e. there is no additional benefit of oxygen and temperature on density once oxygen and temperature exceed limiting values, following typical fish responses to

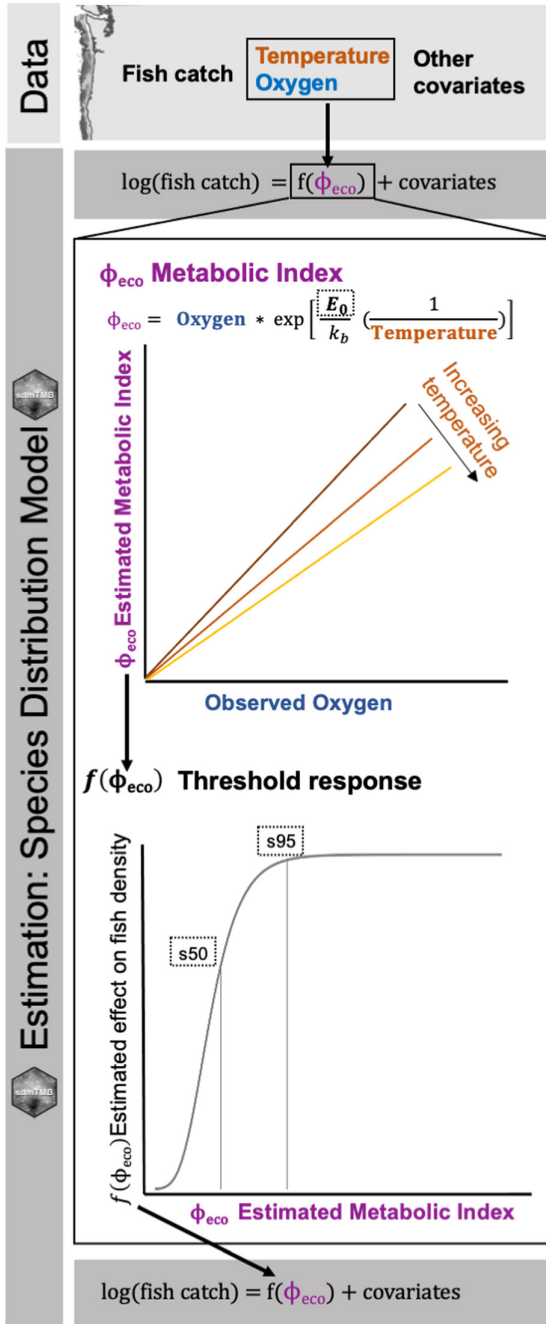


Figure 1. Conceptual diagram showing the data inputs and estimation model. The data required is spatially explicit temperature, oxygen, and fish catch, and any additional environmental covariates (e.g. depth). The model is a commonly used spatial generalized linear mixed model (GLMM), implemented using the R package ‘sdmTMB’ (Anderson et al. 2024). This model is flexible to allow optional temporal, spatial, and spatio-temporal random variation, as well as environmental covariates. We have added a feature within the standard GLMM estimation procedure that simultaneously estimates 1) a modified form of the metabolic index from temperature and oxygen data, and 2) a sigmoidal threshold function that estimates an asymptotic effect of oxygen on local abundance when oxygen becomes a limiting factor. Dashed line boxes indicate the estimated parameters.

environmental conditions (Farrell and Richards 2009). We therefore needed to choose a functional form for  $f(\Phi_{eco})$  such that  $e^{f(\Phi_{eco})}$  is asymptotic and bounded between 0 and 1. The following function meets these criteria:

$$f(\Phi_{eco}) = \psi \left( \frac{1}{1 + \exp(-\alpha_0 - \alpha_1 \Phi_{eco})} - 1 \right) \quad (4)$$

where  $\alpha_0$  and  $\alpha_1$  are parameters describing a sigmoidal curve (based on a modified logistic function) and  $\psi$  scales the function (example plot of this function shown in Fig. 1). We reparameterized Eq. 4 to both work in the log-linear model framework and to have more easily interpretable parameters (Supporting information): the value of  $\Phi_{eco}$  at which the exponentiated value of the function  $f(\Phi_{eco})$  equals 0.5,  $s50$  (i.e. the value of  $\Phi_{eco}$  that reduces fish density by 50%) and the value of  $\Phi_{eco}$  where the exponentiated value of the function equals 0.95,  $s95$  (i.e. the value of  $\Phi_{eco}$  that reduces fish density by 5%). All parameters of the model are estimated simultaneously using maximum marginal likelihood.

To implement this model, we have added a feature within the widely used ‘sdmTMB’ package (Anderson et al. 2024, [www.r-project.org](http://www.r-project.org)) that requires a simple addition to the model formula ([https://github.com/jindivero/estimating\\_mi\\_from\\_distribution2](https://github.com/jindivero/estimating_mi_from_distribution2)) for example vignette.

### Simulation testing

To evaluate whether our model could accurately estimate the metabolic-index parameter  $E_0$  and parameters of  $f(\Phi_{eco})$ , we used simulation testing, a standard practice for testing model performance. This involves specifying values of all model parameters in an operating model, simulating data as if the model were ‘true’, and passing simulated data to the estimation routine. In this way, bias and precision of parameter estimates can be evaluated because the ‘true’ parameter values are known. Our goals of this simulation were to ask: 1) how accurately could the model estimate key parameters relating local density to oxygen and temperature ( $E_0$  and  $s50$ )? 2) How accurately could the model estimate a response of local density to the combined temperature and oxygen function  $f(\Phi_{eco})$ ? 3) Does constraining model estimation with prior information on  $E_0$  improve accuracy? 4) How robust are parameter estimates to a hypothetical model mis-specification?

We simulated data for fish density and distribution, using the U.S. Pacific coast as a model system. Our objective was to replicate the distribution of a hypothetical fish species that inhabits a broad range of depths, specifically extending through its oxygen minimum zone. For instance, sablefish and longspine thornyhead can both be found in habitat up to 1250 m deep (Supporting information). We used year effects and a quadratic response to depth for sablefish based on the estimated parameters in Essington et al. (2022). We simulated datasets for two cases, reflecting contrasting effects of the temperature-dependence of oxygen tolerance. The first (i.e. ‘typical case’) represented a case where the simulated species’ temperature-dependence of oxygen tolerance (Eq. 2) was

close to the median value from the empirical meta-analysis for a generic teleost ( $E_0=0.3$ ) in [Essington et al. \(2024\)](#), as described above. The second (i.e. ‘unusual case’) represents a case in which the simulated species’ temperature-dependence of oxygen sensitivity deviates greatly from the expected value (i.e. the median of the meta-analysis) and is roughly equal to the upper 90% of the meta-analysis distribution ( $E_0=0.7$ ). A higher value of  $E_0$  indicates a steeper sensitivity of oxygen tolerance to changes in temperature.

For both cases, the  $f(\Phi_{eco})$  threshold function parameters were selected to generate a reasonably shaped sigmoidal curve, where ~ 30% of observations fell below the  $s50$  threshold (see the Supporting information for full values used in data generation and all equations and for simulated  $f(\Phi_{eco})$  response for typical and unusual cases). We used the actual latitude/longitude and temperature, oxygen, and depth of each haul from the West Coast Bottom Trawl Survey (WCBTS) data from 2010–2015, rather than simulating random locations, bathymetry and environmental conditions. We wanted to retain the covariances and confounding effects of environmental conditions and space from the real data, thus challenging the model’s ability to fit to data resembling real-world environmental and biological characteristics (see the Supporting information for covariation plots). The WCBTS is a standardized trawl survey of the U.S. portion of the California Current (spanning the Pacific coast waters of Washington, Oregon, and California). It uses a random stratified design and has had consistent sampling procedures since 2003 ([Keller et al. 2017](#)). Using this environmental data and the specified parameter values, fish densities were then generated following Eq. 2–4 using ‘sdmTMB’ ([Anderson et al. 2024](#)) for 250 unique datasets (Supporting information). Each iteration was a randomized realization (varying in the specific year effects, spatial variation, and observation error) of the same true structure. We evaluated the number of iterations to ensure that it was sufficient for stable results (Supporting information).

To evaluate how well parameters could be recovered, we then fit a perfectly specified model to each simulated dataset (i.e. the fitted model was identical to the data generating model, Eq. 3), using ‘sdmTMB’ ([Anderson et al. 2024](#)). To evaluate parameter estimation, we calculated overall accuracy (root mean squared error (RMSE) across all data iterations of each generating-fitting scenario); bias (difference between

the maximum likelihood estimate (MLE) and the true value), and precision (standard deviation of the MLEs across all data iterations). To evaluate the accuracy of the estimated temperature–oxygen effect compared to the true effect size used to simulate data, we calculated the conditional effect of  $f(\Phi_{eco})$  for each data iteration (Eq. 4) and the RMSE compared to the true value. We evaluated the support for the ‘true’ model in comparison to alternative models (see models in [Table 1](#)) by calculating and comparing the marginal Akaike information criterion (AIC) ([Akaike 1974](#)) for each iteration of generated data. This also provided an assessment of how well a standard model selection procedure could identify the correct model as the one most supported by the data.

To assess the impact of using empirical information about physiological parameters on model performance, we repeated the same model fitting procedure for each simulated dataset as described above (i.e. ‘unconstrained model’), but applied a prior probability via a likelihood penalty on estimated  $E_0$  (i.e. ‘constrained model’). The penalized likelihood was based on the distribution expected for teleosts from the empirical meta-analysis of laboratory data ([Essington et al. 2024](#)), which followed a normal distribution of mean=0.3422 and standard deviation=0.1455. This approach is one way to integrate information from multiple sources; if empirical data on a related species are available, that could provide an informative prior to constrain the values of  $E_0$  estimated from realized distribution data. By comparing model fit with and without a prior to both simulated species, we evaluated the ability of the data to estimate the  $f(\Phi_{eco})$  parameters versus the reliance on a prior and whether, and to what extent, a prior improves estimation.

We also evaluated the robustness of model performance against one hypothetical model mis-specification. In the first two model fittings above, the estimation model was perfectly matched to the covariate structure of the generating model. We challenged the model estimation by mis-specifying the covariate structure in Eq. 3 fit to the simulated datasets by using only a linear depth term rather than a quadratic (i.e. ‘mis-specified model’), using unconstrained  $E_0$  estimation (i.e. with no prior).

## Application to case studies

After the simulation testing allowed us to understand how the model performed with known parameters, we then

Table 1. Model selection for alternative models fit to sablefish and longspine thornyhead data.  $f(\Phi_{eco})$  refers to the new model described in Eq. 2–4, including degrees of freedom (df), log-likelihood, and  $\Delta$ AIC. Null refers to the base model only, which included independent means by year, a quadratic effect of depth, and a spatial random field.

Model	df	Sablefish		Longspine thornyhead	
		Log-likelihood	$\Delta$ AIC	Log-likelihood	$\Delta$ AIC
+ $f(\Phi_{eco})$	18	−14 477.61	0	−74441.266	0.8
+logistic( $pO_2$ )	17	−14 488.85	20.5	−7441.887	0
+breakpoint( $pO_2$ )	16	−14 495.00	32.7	−7469.342	52.9
+Temperature $\times$ $pO_2$	17	−14 500.74	50.9	−7460.328	46.1
+Temperature + $pO_2$	16	−14 505.06	51.9	−7466.307	44.3
+ temperature	15	−14 509.76	56.3	−7505.677	122.2
+ $pO_2$	15	−14 506.56	58.3	−7466.385	123.3
Null	14	−14 509.77	44.3	−7506.110	36.0

used empirical distribution data to assess how well this new model structure with the  $f(\Phi_{\text{eco}})$  term performed relative to other models that either did not consider oxygen or that considered oxygen and temperature as separate predictor variables. While this analysis does not allow us to assess precision or accuracy (because the true values of parameters are not known), it does illustrate the application of the method in a realistic context. For this assessment, we used species distribution data on two groundfish species – sablefish *Anoplopoma fimbria* and longspine thornyhead *Sebastolobus altivelis* – in the northeast Pacific coast that inhabit a wide range of habitats that include those with low oxygen over a range of temperatures and depths (Supporting information). There is not laboratory data sufficient to derive the metabolic index following Deutsch et al. (2015) for these or other demersal fish species important to the region. For each species, we modeled the observed CPUE ( $\text{kg km}^{-2}$ ) from the West Coast Bottom Trawl Survey data. We restricted analysis to the catch rate of intermediate-sized fish following the method in Essington et al. (2022) so that the data represented a narrow range of fish sizes. We did this because coast fish distributions are often size-structured, and physiological sensitivity to oxygen is also size-dependent (Rubalcaba et al. 2020).

We used multi-model inference to judge the weight of evidence for our new model (Eq. 3) compared to other parameterizations of oxygen or temperature, as well as not including them at all. Including a null model with no temperature and oxygen terms provided a comparison to evaluate if a species' distribution is not bound by these environmental conditions, rather than assuming they are a constraint. Models denoted with 'breakpoint' function use the following structure to relate a slope  $\beta$  on a predictor variable,  $x$ , up until the breakpoint of the predictor variable  $x_{\text{bp}}$ , on the expected value of the response variable:

$$g(x) = \begin{cases} x\beta & \text{if } x < x_{\text{bp}} \\ x_{\text{bp}}\beta & \text{otherwise} \end{cases} \quad (5)$$

The full suite of alternative models is provided in Table 1.

## Results

Our simulation testing revealed limited accuracy of estimated metabolic index parameters and thresholds when they were estimated from species distribution data. The parameter  $E_0$ , which defines the temperature-dependency of oxygen tolerance of the metabolic index (Eq. 2), was estimated with relatively low bias, but low precision (Fig. 2A). The average estimate of  $E_0$  was only 0.04 less (−13%) than the true value for the typical case ( $E_0 = 0.3$ ) and 0.03 higher (+4%) than that of the unusual case ( $E_0 = 0.7$ ). However, across the 250 iterations of generated data there was a wide range of estimates, for instance ranging from negative values to almost 1 for the typical case. A similar pattern was seen for the threshold parameter  $s50$  (Fig. 2B): the average

estimate was closely estimated to the true value, but there was wide variability in the estimate across data iterations.

A key reason for this low precision is the high covariance between these two parameters (Fig. 3A), caused in part by limited information in the data. Because the model can produce the same overall  $f(\Phi_{\text{eco}})$  shape by estimating multiple combinations of values  $E_0$  and  $s50$ , so that precision in any one parameter is low (Fig. 3B). Across all simulations, the estimated  $f(\Phi_{\text{eco}})$  were consequently close to the true function (Fig. 3B). There are two reasons underlying this covariance. One, the model internally estimates a covariate  $\Phi_{\text{eco}}$  (via  $E_0$ ) and the effect of that covariate  $f(\Phi_{\text{eco}})$  on the local density. Two, the environmental data do not have enough contrast in temperature and  $\text{pO}_2$  near the true threshold value (Fig. 3C). There was only a narrow band of temperatures (6.5–10°C) where observed  $\text{pO}_2$  values were above and below the threshold. However, as the overall function  $f(\Phi_{\text{eco}})$  was well estimated across a range of  $E_0$  estimates (Fig. 3B), predictions of fish response to novel climate conditions (e.g. a +1.5°C increase in temperature) were generally similar among simulations with different parameter estimates (Fig. 3D). The true model was also consistently identified as the most parsimonious model against the alternative models that were fit to the simulated data by a large margin ( $\Delta\text{AIC} = 20\text{--}300$ , Supporting information).

We then asked whether, and to what extent, constraining  $E_0$  based on empirical laboratory data on metabolic index traits would improve the accuracy of estimation. We specified the prior as the median value expected for a teleost from the empirical meta-analysis using taxonomic imputation (Essington et al. 2024), as this represents the best estimate for a generic fish with no laboratory data available. For a typical case, where the mean of this prior was close to the true value of the species, accuracy was improved (Fig. 4A). However, in an unusual case, where a fish had a temperature-dependence of oxygen tolerance far from that expected from empirical data, the overall accuracy was largely unchanged because precision improved while bias increased (Fig. 4A). For the unusual species, the median of the MLEs was 0.48, higher (+68%) than the prior median (0.331) and lower (−68%) than the true value (0.7) (Fig. 4A). In both cases, the distributions of maximum likelihood estimates were distinct from the prior (Fig. 4A).

Estimation accuracy was reduced by a hypothetical model mis-specification. When we omitted the quadratic effect of  $\log(\text{depth})$  from the estimation model, parameter estimates were highly biased in both the typical and unusual cases, leading to a 5–7 fold increase in RMSE (Supporting information). For instance, for the typical species,  $E_0$  was underestimated by 1.45 (−300%) (Supporting information), compared to 0.04 (+13%) in the correctly specified model (Fig. 2A). While the RMSE of the estimated function  $f(\Phi_{\text{eco}})$  also increased, the magnitude of change was relatively small with an RMSE of  $0.14 \pm 0.009$  for a typical case and  $0.26 \pm 0.14$  for an unusual case (Supporting information). Despite reduced accuracy, the model that included  $f(\Phi_{\text{eco}})$  was still most commonly identified as the

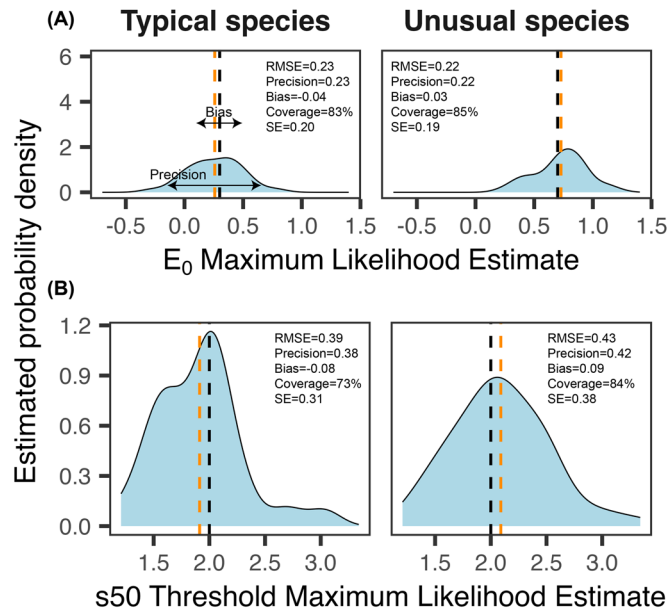


Figure 2. Comparison of maximum likelihood estimates for (A) the temperature-dependence of oxygen sensitivity parameter  $E_0$  from the derived metabolic index and (B)  $s_{50}$ , a parameter in the sigmoidal threshold function, for each of the two data generating scenarios (a typical case, left and an unusual case, right) for the unconstrained model (estimated with no prior on  $E_0$ ) across all 250 data simulation iterations. The typical case represents a simulated species' temperature-dependence of oxygen tolerance that is close to the median value from the empirical meta-analysis for a generic teleost ( $E_0=0.3$ ), and the unusual case where a simulated species' temperature-dependence of oxygen sensitivity deviates greatly from the expected value (i.e. the median of the meta-analysis) and is roughly equal to the upper 90% of the meta-analysis distribution ( $E_0=0.7$ ). The blue curves show a fitted kernel density smoother across the maximum likelihood estimates (MLEs) for all 250 simulated data iterations. The black dashed line shows the true parameter value specified to generate the simulated data, and the orange line shows the average maximum likelihood estimate across all iterations of simulated data. The upper left figure shows the interpretation of bias and precision. Bias is calculated as the difference between the average MLE and the true value, and precision is the standard deviation of MLEs across all iterations (proportional to the width of the blue kernel density). RMSE is the root mean squared error, which gives a single metric of accuracy.

most-supported model for the simulated data for the typical case and for the majority of simulated datasets for the unusual case.

### Estimation of temperature–oxygen synergistic effect in case studies

When applied to two groundfish species, we found mixed support for our new model compared to alternative models. For sablefish, our new model provided the most parsimonious fit by a large margin over a sigmoidal model of  $pO_2$  alone ( $\Delta AIC=20$ ) (Table 1). However, for longspine thornyhead the new model and the model with sigmoidal ( $pO_2$ ) as a predictor had roughly equivalent support ( $\Delta AIC$  of new model=0.77) (Table 1). For both species, the new models provided a substantially better fit than any other alternative model of temperature and/or oxygen and of the null model (i.e. oxygen linearly or as a breakpoint, temperature linearly, or both oxygen and temperature combined as separate linear terms or as an interactive term) (Table 1). Because the simulation testing showed that  $s_{50}$  and  $E_0$  could not be estimated precisely, we focus on the estimated  $f(\Phi_{eco})$  (Fig. 3B). (see the Supporting information for parameter estimates). Only a small percentage of longspine thornyhead observations were

below the  $f(\Phi_{eco})$   $s_{50}$  or  $s_{95}$  threshold and were therefore predicted to be restricted by low oxygen and high temperature (Fig. 5B). Yet for sablefish, 3% of hauls were below the  $s_{50}$  and 18% below  $s_{95}$  (Fig. 5A). A primary reason why the new model was strongly supported for sablefish, but not longspine thornyhead, is that there were more sablefish catches where density was constrained by the estimated  $\Phi_{eco}$ .

## Discussion

We developed a statistical model that estimates the temperature-dependence of oxygen tolerance (metabolic index) using the spatiotemporal distribution of a species, rather than laboratory data, by simultaneously estimating the value of the metabolic index and its effect on fish density. Using simulation testing, we evaluated the performance of the new model and then applied it to empirical data from two groundfish species. Our findings reveal both opportunities and challenges for this approach. Generally, historical distributions were estimated accurately, even when models were mis-specified. Additionally, these models were sufficient for forecasting responses to simulated changes in environmental conditions, even though the statistical model could not uniquely estimate

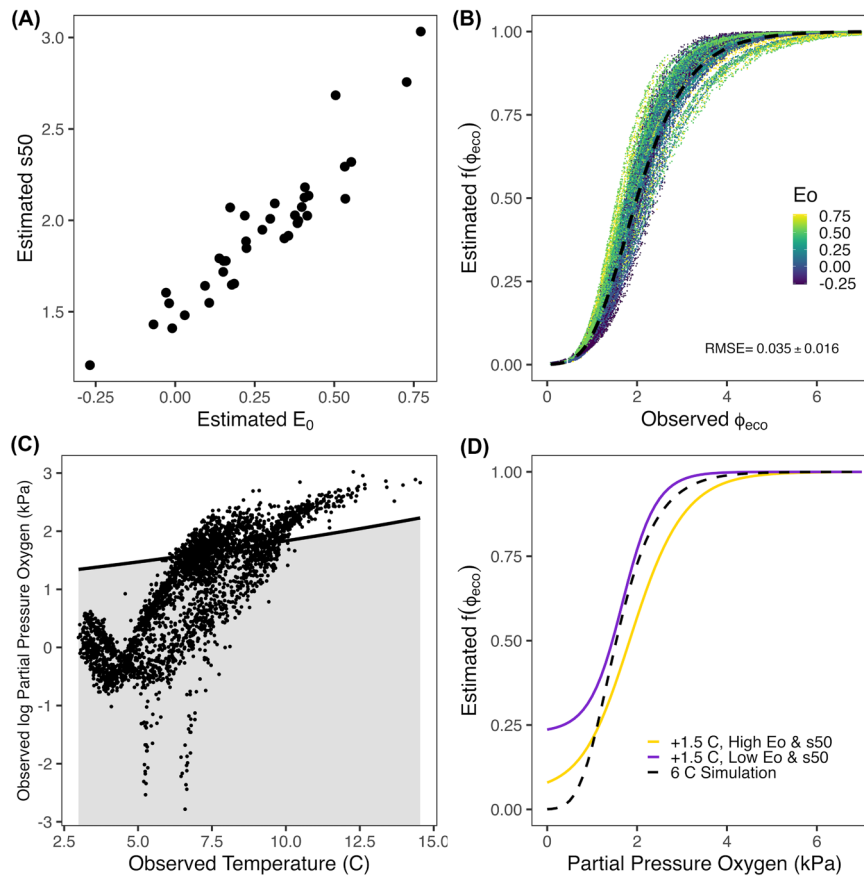


Figure 3. (A) Maximum likelihood estimates of derived metabolic index parameter  $E_0$  and sigmoidal threshold parameter  $s50$  for each simulated data set. (B) Estimated synergistic temperature–oxygen effect  $f(\Phi_{\text{ecc}})$  from each model fit to each of 250 simulated data sets, with color indicating the value of the maximum likelihood estimate of the derived metabolic index  $\Phi_{\text{ecc}}$  parameter  $E_0$ , and the dashed line the true effect specified in the simulated data. Root mean squared error, shown as the average  $\text{RMSE} \pm \text{SD}$  of iterations, shows the mean and spread of accuracy across iterations. RMSE for the unusual species was similar,  $0.032 \pm 0.016$ . (C) The temperature and  $\text{pO}_2$  values from the data (points), and the combinations of temperature and  $\text{pO}_2$  where  $\Phi_{\text{ecc}}$  equals the  $s50$  threshold value (black line). Grey shaded area depicts combinations below this threshold. (D) The estimated synergistic temperature–oxygen effect from two model fits, one high ( $E_0=0.77$ ,  $s50=3.03$ , yellow) and one low ( $E_0=0.03$ ,  $s50=1.48$ , blue) metabolic index and sigmoidal threshold parameters given a hypothetical  $1.5^\circ\text{C}$  increase in bottom temperature (compared to the true response at a base reference temperature of  $6^\circ\text{C}$ , dashed line) across a range of plausible observed oxygen values.

two covarying parameters: the temperature-dependence of oxygen sensitivity, and the threshold value of the index at which species densities decline. The addition of priors provided some benefits, but these came with the risk of increasing bias when taxa’s oxygen limits had unusual temperature dependence.

Our findings demonstrate the possible benefit of this approach for predicting species’ responses to climate change. Our full model provided the most parsimonious fit to simulated data and for empirical Pacific coast sablefish distribution, and was a more parsimonious fit than a null model or models with temperature and oxygen alone for longspine thornyhead. The full model remained the most parsimonious fit in simulation testing even with a mis-specified covariate, though we did not test robustness against all possible mis-specifications. As our approach is a correlative model, it may still be subject to a common critique of correlative models,

namely that an improved fit to historical data does not necessarily translate to better prediction, especially when extrapolating to unobserved environmental conditions (Barnes et al. 2022, Brodie et al. 2022). However we found that our approach could reliably generate a unique effect of a change in oxygen for a given temperature change. A key benefit of our approach is that by rooting in physiology, it constrains extrapolation toward more biologically realistic responses by encompassing novel combinations of changes in two dimensions – oxygen and temperature – into a single value (the metabolic index) that can be interpolated along a known response curve. In our study region, bottom temperatures are projected to increase  $0.5\text{--}2.5^\circ\text{C}$  by 2100, and oxygen levels to drop by  $0\text{--}30 \text{ mmol m}^{-3}$  (Liu et al. 2023), producing a range of new conditions. These effects may be mapped to identify areas or habitats that are facing higher risks associated with climate change (Parouffe et al. 2023).



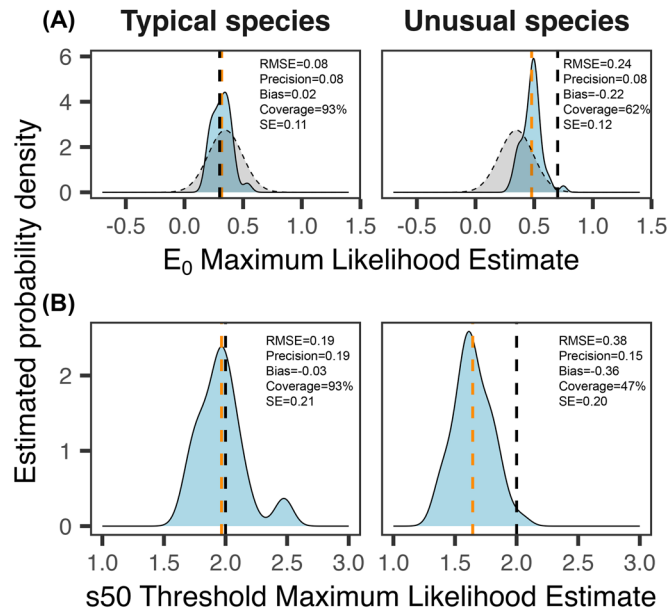


Figure 4. Comparison of maximum likelihood estimate for (A) the temperature-dependence of oxygen sensitivity parameter  $E_0$  from the derived metabolic index and (B)  $s_{50}$ , a parameter in the sigmoidal threshold function from each of the two data generating scenarios (a typical case, in the left column, and an unusual case, right). The typical case represents a simulated species' temperature-dependence of oxygen tolerance that is close to the median value from the empirical meta-analysis for a generic teleost ( $E_0=0.3$ ), and the unusual case where a simulated species' temperature-dependence of oxygen sensitivity deviates greatly from the expected value (i.e. the median of the meta-analysis) and is roughly equal to the upper 90% of the meta-analysis distribution ( $E_0=0.7$ ). Model estimation included a prior constraint on  $E_0$  via a penalized likelihood. The blue curves show the fitted kernel density smoother across the maximum likelihood estimates (MLEs) for all 250 simulated data iterations (narrower densities translates to more precision). Gray curves show the prior distribution used to constrain  $E_0$  based on the empirical model. The black dashed line shows the true parameter value specified to generate the simulated data, and the orange line is the average maximum likelihood estimate across all iterations of simulated data.

However, there are still limitations with the approach. We caution that each parameter ( $E_0$  and  $s_{50}$ ) should not be used in isolation for prediction outside of this model structure, since the covariance in estimation means that it did not reliably delineate between the temperature-sensitivity of oxygen (i.e.  $\Phi_{eco}$ ) versus the threshold of temperature-corrected oxygen (i.e.  $s_{50}$ ). Additionally, it is necessary to consider this

model in a suite of alternative configurations of temperature and oxygen, including a null model to ensure that a response to temperature and oxygen is best supported by the data, rather than assuming that they are constraining distribution. The estimated response also needs to be properly interpreted for the body size structure of each dataset and population, as metabolic rate scales with body size (Deutsch et al. 2015,

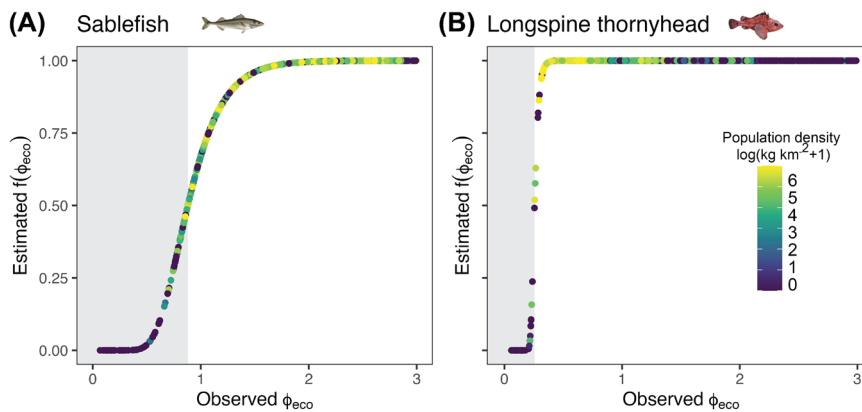


Figure 5. Estimated marginal effect of the synergistic temperature-oxygen term  $f(\Phi_{eco})$  on fish density for (A) sablefish *Anoplopoma fimbria* and (B) longspine thornyhead *Sebastolobus altivelis*, for each  $\Phi_{eco}$  calculated as the observed oxygen and temperature and estimated parameter  $E_0$ . Color indicates the observed catch per area sampled, and the grey shaded areas indicate the values  $\Phi_{eco}$  where fish would be restricted by low oxygen and high temperature.

Rubalcaba et al. 2020). For instance, biomass, rather than numerical abundance, data should be used, and the size structure of the fish catch needs to be considered. Here, we restricted fish densities to only biomass of intermediate-sized fish, and therefore interpret the metabolic index to be estimated and density predicted for intermediate-sized fishes. Overall properly quantifying and incorporating uncertainty in the metabolic index response (Brodie et al. 2022, Davies et al. 2023) will help reflect the parameter imprecision and range of possible fish distribution outcomes under future conditions.

Some of the estimation challenges we encountered were a result of inherent characteristics of the environmental data. For one, the range of environmental conditions present in the data was limited, providing insufficient contrast around the threshold. If the data had provided more combinations of temperature and  $pO_2$  that produced a  $\Phi_{eco}$  above and below the threshold (e.g. high temperature and low  $pO_2$ , or low temperature and high  $pO_2$ ), it is possible that the precision of the threshold estimation would have been improved. Truncated environmental data, with a restricted representation of possible conditions – particularly relative to the variation and extremes observed on shorter timescales – is an ongoing challenge in forecasting distributions (Thuiller et al. 2004, Hannemann et al. 2016, Charney et al. 2021). Seeking out datasets, or combining data across a broader geographic range, that can expand the range of training data to capture a wider range of conditions (Bandara et al. 2023) and reduce extrapolation (Brodie et al. 2022, Waldock et al. 2022, Davies et al. 2023) may allow better estimation of responses to temperature and oxygen. This would additionally help identify whether a species (such as longspine thornyhead here) with a poor fit of this model and with density estimated to be constrained only at a very low metabolic index value, is due to the dataset only including environmental conditions within a species' tolerance, versus a species whose distribution is truly not constrained by temperature and oxygen. This finding supports the need for increased monitoring to better characterize thresholds: advanced ocean observing instrumentation that can collect environmental data with greater spatial and temporal resolution and extent, and additional biological survey efforts that focus on sampling environmental extremes and complement existing stratified random sampling designs.

Second, there was spatial structure in – and covariation between – the environmental variables of interest (temperature and oxygen) and other environmental features, such as bathymetry. We recognize that the species distribution model (a generalized linear mixed model with latent spatial fields) and underlying data (point densities) we used here are only one of many possible frameworks for evaluating and predicting distributions in response to environmental change. Yet other model approaches and distribution data in other environmental systems likely face similar challenges, for instance as seen in: spatial confounding in Bayesian approaches (Mäkinen et al. 2022); spatial bias in presence-only data (Phillips et al. 2009, Baker et al. 2022); spatial dependence and bias in maximum entropy (Kramer-Schadt et al. 2013, Halvorsen et al. 2016); and the impact of spatial

autocorrelation in various methods including boosted regression trees, generalized additive models, mixed effects models, maximum entropy, and others (Santika and Hutchinson 2009, Crase et al. 2014, Brodie et al. 2020). Other efforts to apply the metabolic index by parameterizing directly from laboratory data have seen similar limitations when fitting to realized distributions, such as a restricted range of environmental conditions in the data (Bandara et al. 2023) and confounding spatial features (Essington et al. 2022). Estimating these temperature–oxygen relationships directly from the distribution data itself did not fully resolve these issues. Combining distribution data from a wider geographic area, such as across the entire northeastern Pacific range for sablefish, may provide enough contrast in the relationships between environmental variables to overcome challenges of spatial structure and covariance, since correlations between bathymetry, oxygen and temperature differ between regions.

Our findings highlight a well-recognized difficulty in identifying causal pathways from distribution data (Thorson et al. 2021a, Addicott et al. 2022). Here, we were able to flag uncertainty in the parameter estimates because our estimation model included latent spatial fields and other covarying environmental features (depth). If we had not included these effects, we may have wrongly identified a temperature–oxygen impact on distribution that was actually due to other factors. It is important to note that there are other possible impacts on density that were not included in the model (such as food availability or predator presence) that may not have been captured by the latent spatial field and may have impacted the estimated response to the metabolic index. Additionally, fish may show no threshold response to oxygen and temperature and not be bound by these environmental constraints. In some cases, there may be an advantage to not use spatial random fields, as they may absorb too much of the variation and dampen the effect of the environmental variable of interest (Hodges and Reich 2010). This reinforces the need to properly consider model structure to avoid spuriously attributing species distribution to environmental conditions (Bahn and McGill 2007, Chapman 2010, Brodie et al. 2022). By accounting for additional physical and biotic factors beyond temperature and oxygen that may exert more influence on population densities at fine spatiotemporal scales (Whittaker et al. 2001), our approach may improve the explanatory and predictive skill of the metabolic index, which has previously been challenging (Essington et al. 2022, Bandara et al. 2023). Analyses conducted at broader biogeographic scales to evaluate range limits have attributed much of the observed patterns to the metabolic index while ignoring additional drivers (Deutsch et al. 2015, Penn and Deutsch 2024). Considering multiple dimensions of control on species distribution, not solely the variables of interest, and comparing to more complex models to test robustness (Oster 2019) can ensure proper attribution to the temperature–oxygen response of interest.

Improving methods for combining physiological responses and the information from realized distributions will support forecasting the effects of climate change on marine species

(Meineri et al. 2015, Peterson et al. 2015, Gamliel et al. 2020, Tourinho and Vale 2023). For the metabolic index, resolving the temperature-sensitivity of oxygen will be a particular challenge. Empirical laboratory data and spatial comparisons show high variation in  $E_0$  across species and regions (Essington et al. 2024, Penn and Deutsch 2024). Here, distribution data similarly did not provide a reliable estimate of  $E_0$  alone, but could estimate a combined threshold effect. Applying empirical data to inform a prior on the  $E_0$  parameter also did not resolve this challenge, as it can greatly bias estimates if species' realized behavior greatly differs from that theoretically expected. While we adopted the metabolic index for this study, other physiological frameworks of how fish respond to oxygen and temperature (Ern 2019, Clarke et al. 2021) and other mechanisms, could similarly be tested. Estimating sensitivities to oxygen and temperature from responses beyond distribution, such as growth and size (van Denderen et al. 2020, Dimarchopoulou and Tsikliras 2022), may also be informative. Limited availability or applicability of data for parameterization is a common problem for incorporating mechanistic responses into species distribution models (Gamliel et al. 2020). For identifying mechanistic parameters from distribution data in other cases, our study highlights the need for simulation testing (DiRenzo et al. 2023) to validate that the mechanistic parameters can be reliably estimated, given the unique spatial structure, range of environmental conditions, and confounding covariates specific to that dataset and model structure. Additionally, evaluating multiple distribution models specified with alternative parameter values spanning a plausible range informed by empirical data (e.g. similar to the 'ecotype' approach in Chen et al. 2024), rather than assuming a single parameterization, could help capture uncertainty in the mechanistic response (Brodie et al. 2022). Other modeling approaches – such as space- or time-varying coefficients (Barnett et al. 2021), a fully Bayesian method (Talluto et al. 2016, Gamliel et al. 2020, Morera-Pujol et al. 2023), empirical dynamic modeling (Wang et al. 2020) and structural equation modeling (Liu et al. 2005, Thorson et al. 2021a) – may also provide an improved framework to estimate mechanistic responses and combine insights from empirical laboratory work and realized species distributions.

*Acknowledgements* – We thank all West Coast Groundfish Bottom Trawl survey participants. We also thank Jenny Bigman for helpful comments.

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

## Author contributions

**Julia Indivero:** Conceptualization (equal); Formal analysis (lead); Methodology (equal); Writing - original draft (lead); Writing - review and editing (equal). **Sean C. Anderson:** Conceptualization (equal); Formal analysis (equal);

Methodology (equal); Software (equal); Writing - review and editing (equal). **Lewis A. K. Barnett:** Conceptualization (equal); Data curation (equal); Funding acquisition (equal); Methodology (equal); Supervision (equal); Writing - review and editing (equal). **Timothy E. Essington:** Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Methodology (equal); Project administration (equal); Supervision (lead); Writing - original draft (equal); Writing - review and editing (equal). **Eric J. Ward:** Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Software (equal); Writing - review and editing (equal).

## Transparent peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/ecog.07413>.

## Data availability statement

Data are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.47d7wm3pb> (Indivero et al. 2024). The data and code are also available at [https://github.com/jindivero/estimating\\_mi\\_from\\_distribution2](https://github.com/jindivero/estimating_mi_from_distribution2).

## Supporting information

The Supporting information associated with this article is available with the online version.

## References

- Addicott, E. T., Fenichel, E. P., Bradford, M. A., Pinsky, M. L. and Wood, S. A. 2022. Toward an improved understanding of causation in the ecological sciences. – *Front. Ecol. Environ.* 20: 474–480.
- Akaike, H. 1974. A new look at the statistical model identification. – *IEEE Trans. Autom. Control* 19: 716–723.
- Anderson, S. C., Ward, E. J., English, P. A., Barnett, L. A. K. and Thorson, J. T. 2024. sdmTMB: an R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. – *bioRxiv*, 2022.03.24.485545.
- Asch, R. G., Sobolewska, J. and Chan, K. 2022. Assessing the reliability of species distribution models in the face of climate and ecosystem regime shifts: small pelagic fishes in the California Current System. – *Front. Mar. Sci.* 9: 711522.
- Atkinson, D., Leighton, G. and Berenbrink, M. 2022. Controversial roles of oxygen in organismal responses to climate warming. – *Biol. Bull.* 243: 207–219.
- Bahn, V. and McGill, B. J. 2007. Can niche-based distribution models outperform spatial interpolation? – *Global Ecol. Biogeogr.* 16: 733–742.
- Baker, D. J., Maclean, I. M. D., Goodall, M. and Gaston, K. J. 2022. Correlations between spatial sampling biases and environmental niches affect species distribution models. – *Global Ecol. Biogeogr.* 31: 1038–1050.
- Bandara, R. M. W. J., Curchitser, E. and Pinsky, M. L. 2023. The importance of oxygen for explaining rapid shifts in a marine fish. – *Global Change Biol.* 30: e17008.

- Barnes, C. L., Essington, T. E., Pirtle, J. L., Rooper, C. N., Laman, E. A., Holsman, K. K., Aydin, K. Y. and Thorson, J. T. 2022. Climate-informed models benefit hindcasting but present challenges when forecasting species–habitat associations. – *Ecography* 2022: e06189.
- Barnett, L. A. K., Ward, E. J. and Anderson, S. C. 2021. Improving estimates of species distribution change by incorporating local trends. – *Ecography* 44: 427–439.
- Baudron, A. R., Brunel, T., Blanchet, M. A., Hidalgo, M., Chust, G., Brown, E. J., Kleisner, K. M., Millar, C., MacKenzie, B. R., Nikolioudakis, N., Fernandes, J. A. and Fernandes, P. G. 2020. Changing fish distributions challenge the effective management of European fisheries. – *Ecography* 43: 494–505.
- Beaury, E. M., Fusco, E. J., Jackson, M. R., Laginhas, B. B., Morelli, T. L., Allen, J. M., Pasquarella, V. J. and Bradley, B. A. 2020. Incorporating climate change into invasive species management: insights from managers. – *Biol. Invas.* 22: 233–252.
- Bell, R. J., Odell, J., Kirchner, G. and Lomonico, S. 2020. Actions to promote and achieve climate-ready fisheries: summary of current practice. – *Mar. Coast. Fish.* 12: 166–190.
- Briscoe, N. J., Morris, S. D., Mathewson, P. D., Buckley, L. B., Jusup, M., Levy, O., Maclean, I. M. D., Pincebourde, S., Riddell, E. A., Roberts, J. A., Schouten, R., Sears, M. W. and Kearney, M. R. 2023. Mechanistic forecasts of species responses to climate change: the promise of biophysical ecology. – *Global Change Biol.* 29: 1451–1470.
- Brodie, S. J., Thorson, J. T., Carroll, G., Hazen, E. L., Bograd, S., Haltuch, M. A., Holsman, K. K., Kotwicki, S., Samhouri, J. F., Willis-Norton, E. and Selden, R. L. 2020. Tradeoffs in covariate selection for species distribution models: a methodological comparison. – *Ecography* 43: 11–24.
- Brodie, S. et al. 2022. Recommendations for quantifying and reducing uncertainty in climate projections of species distributions. – *Global Change Biol.* 28: 6586–6601.
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. and West, G. B. 2004. Toward a metabolic theory of ecology. – *Ecology* 85: 1771–1789.
- Chapman, D. S. 2010. Weak climatic associations among British plant distributions. – *Global Ecol. Biogeogr.* 19: 831–841.
- Charney, N. D., Record, S., Gerstner, B. E., Merow, C., Zarnetske, P. L. and Enquist, B. J. 2021. A test of species distribution model transferability across environmental and geographic space for 108 western North American tree species. – *Front. Ecol. Evol.* 9: 689295.
- Chen, Z., Siedlecki, S., Long, M., Petrik, C. M., Stock, C. A. and Deutsch, C. A. 2024. Skillful multiyear prediction of marine habitat shifts jointly constrained by ocean temperature and dissolved oxygen. – *Nat. Commun.* 15: 900.
- Cheng, L., Abraham, J., Hausfather, Z. and Trenberth, K. E. 2019. How fast are the oceans warming? – *Science* 363: 128–129.
- Clarke, T. M., Wabnitz, C. C. C., Frölicher, T. L., Reygondeau, G., Pauly, D. and Cheung, W. W. L. 2022. Linking observed changes in pelagic catches to temperature and oxygen in the eastern Tropical Pacific. – *Fish Fish.* 23: 1371–1382.
- Clarke, T. M., Wabnitz, C. C. C., Striegel, S., Frölicher, T. L., Reygondeau, G. and Cheung, W. W. L. 2021. Aerobic growth index (AGI): an index to understand the impacts of ocean warming and deoxygenation on global marine fisheries resources. – *Prog. Oceanogr.* 195: 102588.
- Crase, B., Liedloff, A., Vesik, P. A., Fukuda, Y. and Wintle, B. A. 2014. Incorporating spatial autocorrelation into species distribution models alters forecasts of climate-mediated range shifts. – *Global Change Biol.* 20: 2566–2579.
- 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. and Hunter, K. L. 2023. Addressing uncertainty when projecting marine species' distributions under climate change. – *Ecography* 2023: e06731.
- Debastiani, V. J., Bastazini, V. A. G. and Pillar, V. D. 2021. Using phylogenetic information to impute missing functional trait values in ecological databases. – *Ecol. Inform.* 63: 101315.
- Deutsch, C., Ferrel, A., Seibel, B., Pörtner, H.-O. and Huey, R. B. 2015. Ecophysiology. Climate change tightens a metabolic constraint on marine habitats. – *Science* 348: 1132–1135.
- Deutsch, C., Penn, J. L. and Seibel, B. 2020. Metabolic trait diversity shapes marine biogeography. – *Nature* 585: 557–562.
- Diaz, R. J. and Rosenberg, R. 2008. Spreading dead zones and consequences for marine ecosystems. – *Science* 321: 926–929.
- Dimarchopoulou, D. and Tsikliras, A. C. 2022. Linking growth patterns to sea temperature and oxygen levels across European sardine (*Sardina pilchardus*) populations. – *Environ. Biol. Fishes* 105: 1335–1345.
- DiRenzo, G. V., Hanks, E. and Miller, D. A. W. 2023. A practical guide to understanding and validating complex models using data simulations. – *Methods Ecol. Evol.* 14: 203–217.
- Duncan, M. I., James, N. C., Potts, W. M. and Bates, A. E. 2020. Different drivers, common mechanism; the distribution of a reef fish is restricted by local-scale oxygen and temperature constraints on aerobic metabolism. – *Conserv. Physiol.* 8: coaa090.
- Ern, R. 2019. A mechanistic oxygen-and temperature-limited metabolic niche framework. – *Philos. Trans. R. Soc. B* 374: 20180540.
- Essington, T. E., Anderson, S. C., Barnett, L. A. K., Berger, H. M., Siedlecki, S. A. and 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: e06249.
- Essington, T. E., Thorson, J. T. and Deutsch, C. 2024. Remarkable similarity in oxygen tolerance among taxonomically diverse marine taxa revealed through hierarchical analysis. – *bioRxiv* 2024.08.23.606857.
- Farrell, A. P. and Richards, J. G. 2009. Chapter 11. Defining hypoxia: an integrative synthesis of the responses of fish to hypoxia. – In: Richards, J. G., Farrell, A. P. and Brauner, C. J. (eds), *Fish Physiology*. 27. Academic Press, pp. 487–503.
- Franco, A. C., Kim, H., Frenzel, H., Deutsch, C., Ianson, D., Sumaila, U. R. and Tortell, P. D. 2022. Impact of warming and deoxygenation on the habitat distribution of Pacific halibut in the northeast Pacific. – *Fish. Oceanogr.* 31: 601–614.
- Franklin, J. 2010. Statistical models – modern regression. In: *Mapping Species Distributions: Spatial Inference and Prediction*. Ecology, Biodiversity and Conservation. Cambridge Univ. Press.
- Frölicher, T. L., Fischer, E. M. and Gruber, N. 2018. Marine heatwaves under global warming. – *Nature* 560: 360–364.
- Gamliel, I., Buba, Y., Guy-Haim, T., Garval, T., Willette, D., Rilov, G. and Belmaker, J. 2020. Incorporating physiology into species distribution models moderates the projected impact of warming on selected Mediterranean marine species. – *Ecography* 43: 1090–1106.
- Gillooly, J. F., Brown, J. H., West, G. B., Savage, V. M. and Charnov, E. L. 2001. Effects of size and temperature on metabolic rate. – *Science* 293: 2248–2251.

- Halvorsen, R., Mazzoni, S., Dirksen, J. W., Næsset, E., Gobakken, T. and Ohlson, M. 2016. How important are choice of model selection method and spatial autocorrelation of presence data for distribution modelling by MaxEnt? – *Ecol. Modell.* 328: 108–118.
- Hannemann, H., Willis, K. J. and Macias-Fauria, M. 2016. The devil is in the detail: unstable response functions in species distribution models challenge bulk ensemble modelling. – *Global Ecol. Biogeogr.* 25: 26–35.
- Hirzel, A. H. and Le Lay, G. 2008. Habitat suitability modelling and niche theory. – *J. Appl. Ecol.* 45: 1372–1381.
- Hobbs, R. J., Higgs, E. and Harris, J. A. 2009. Novel ecosystems: implications for conservation and restoration. – *Trends Ecol. Evol.* 24: 599–605.
- Hodges, J. S. and Reich, B. J. 2010. Adding spatially-correlated errors can mess up the fixed effect you love. – *Am. Stat.* 64: 325–334.
- Hoffmann, S., Irl, S. D. H. and Beierkuhnlein, C. 2019. Predicted climate shifts within terrestrial protected areas worldwide. – *Nat. Commun.* 10: 4787.
- Indivero, J., Anderson, S. C., Barnett, L. A. K., Essington, T. E. and Ward, E. J. 2024. Data from: Estimating a physiological threshold to oxygen and temperature from marine monitoring data reveals challenges and opportunities for forecasting distribution shifts. – Dryad Digital Repository, <https://doi.org/10.5061/dryad.47d7wm3pb>.
- Jutfelt, F., Norin, T., Ern, R., Overgaard, J., Wang, T., McKenzie, D. J., Lefevre, S., Nilsson, G. E., Metcalfe, N. B. and Hickey, A. J. R. 2018. Oxygen-and capacity-limited thermal tolerance: blurring ecology and physiology. – *J. Exp. Biol.* 221: jeb169615.
- Kearney, M. and Porter, W. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. – *Ecol. Lett.* 12: 334–350.
- Keller, A. A., Wallace, J. R. and Methot, R. D. 2017. The northwest fisheries science center's west coast groundfish bottom trawl survey: history, design, and description. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-NWFSC-136.
- Kramer-Schadt, S. et al. 2013. The importance of correcting for sampling bias in MaxEnt species distribution models. – *Divers. Distrib.* 19: 1366–1379.
- Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H. and Bell, B. M. 2016. TMB: automatic differentiation and Laplace approximation. – *J. Stat. Softw.* 70: 1–21.
- Laufkötter, C., Zscheischler, J. and Frölicher, T. L. 2020. High-impact marine heatwaves attributable to human-induced global warming. – *Science* 369: 1621–1625.
- Lefevre, S., McKenzie, D. J. and Nilsson, G. E. 2018. In modelling effects of global warming, invalid assumptions lead to unrealistic projections. – *Global Change Biol.* 24: 553–556.
- Leiva, F. P., Calosi, P. and Verberk, W. C. E. P. 2019. Scaling of thermal tolerance with body mass and genome size in ectotherms: a comparison between water-and air-breathers. – *Philos. Trans. R. Soc. B* 374: 20190035.
- Lenoir, J., Bertrand, R., Comte, L., Bourgeaud, L., Hattab, T., Murienne, J. and Grenouillet, G. 2020. Species better track climate warming in the oceans than on land. – *Nat. Ecol. Evol.* 4: 1044–1059.
- Lindgren, F., Rue, H. and Lindström, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. – *J. R. Stat. Soc. B* 73: 423–498.
- Liu, X., Wall, M. M. and Hodges, J. S. 2005. Generalized spatial structural equation models. – *Biostatistics* 6: 539–557.
- Liu, O. R. et al. 2023. Species redistribution creates unequal outcomes for multispecies fisheries under projected climate change. – *Sci. Adv.* 9: eadg5468.
- Mäkinen, J., Numminen, E., Niittynen, P., Luoto, M. and Vanhatalo, J. 2022. Spatial confounding in Bayesian species distribution modeling. – *Ecography* 2022: e06183.
- Marshall, K. N., Kaplan, I. C., Hodgson, E. E., Hermann, A., Busch, D. S., McElhany, P., Essington, T. E., Harvey, C. J. and Fulton, E. A. 2017. Risks of ocean acidification in the California Current food web and fisheries: ecosystem model projections. – *Global Change Biol.* 23: 1525–1539.
- Martínez, B., Arenas, F., Trilla, A., Viejo, R. M. and Carreño, F. 2015. Combining physiological threshold knowledge to species distribution models is key to improving forecasts of the future niche for macroalgae. – *Global Change Biol.* 21: 1422–1433.
- McClatchie, S., Goericke, R., Auad, G. and Hill, K. 2010. Re-assessment of the stock–recruit and temperature–recruit relationships for Pacific sardine (*Sardinops sagax*). – *Can. J. Fish. Aquat. Sci.* 67: 1782–1790.
- Meineri, E., Deville, A.-S., Grémillet, D., Gauthier-Clerc, M. and Béchet, A. 2015. Combining correlative and mechanistic habitat suitability models to improve ecological compensation. – *Biol. Rev.* 90: 314–329.
- Melo-Merino, S. M., Reyes-Bonilla, H. and Lira-Noriega, A. 2020. Ecological niche models and species distribution models in marine environments: a literature review and spatial analysis of evidence. – *Ecol. Modell.* 415: 108837.
- Methorst, J., Böhning-Gaese, K., Khaliq, I. and Hof, C. 2017. A framework integrating physiology, dispersal and land-use to project species ranges under climate change. – *J. Avian Biol.* 48: 1532–1548.
- Monzón, J., Moyer-Horner, L. and Palamar, M. B. 2011. Climate change and species range dynamics in protected areas. – *BioScience* 61: 752–761.
- Moore, C., Drazen, J. C., Radford, B. T., Kelley, C. and Newman, S. J. 2016. Improving essential fish habitat designation to support sustainable ecosystem-based fisheries management. – *Mar. Policy* 69: 32–41.
- Morée, A. L., Clarke, T. M., Cheung, W. W. L. and Frölicher, T. L. 2023. Impact of deoxygenation and warming on global marine species in the 21st century. – *Biogeosciences* 20: 2425–2454.
- Morera-Pujol, V., Mostert, P. S., Murphy, K. J., Burkitt, T., Coad, B., McMahon, B. J., Nieuwenhuis, M., Morelle, K., Ward, A. I. and Ciuti, S. 2023. Bayesian species distribution models integrate presence-only and presence–absence data to predict deer distribution and relative abundance. – *Ecography* 2023: e06451.
- Muhling, B. A., Brodie, S., Smith, J. A., Tommasi, D., Gaitan, C. F., Hazen, E. L., Jacox, M. G., Auth, T. D. and Brodeur, R. D. 2020. Predictability of species distributions deteriorates under novel environmental conditions in the California Current system. – *Front. Mar. Sci.* 7: 589.
- Myers, R. A. 1998. When do environment–recruitment correlations work? – *Rev. Fish Biol. Fish.* 8: 285–305.
- Oster, E. 2019. Unobservable selection and coefficient stability: theory and evidence. – *J. Bus. Econ. Stat.* 37: 187–204.
- Parouffe, A., Garçon, V., Dewitte, B., Paulmier, A., Montes, I., Parada, C., Mecho, A. and Veliz, D. 2023. Evaluating future climate change exposure of marine habitat in the South East Pacific based on metabolic constraints. – *Front. Mar. Sci.* 9: 1055875.

- Penn, J. L. and Deutsch, C. 2022. Avoiding ocean mass extinction from climate warming. – *Science* 376: 524–526.
- Penn, J. L. and Deutsch, C. 2024. Geographical and taxonomic patterns in aerobic traits of marine ectotherms. – *Philos. Trans. R. Soc. B* 379: 20220487.
- Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E., Graham, C. H. and Costa, G. C. 2014. Imputation of missing data in life-history trait datasets: which approach performs the best? – *Methods Ecol. Evol.* 5: 961–970.
- Peterson, A. T., Papeş, M. and Soberón, J. 2015. Mechanistic and correlative models of ecological niches. – *Eur. J. Ecol.* 1: 28–38.
- Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J. and Ferrier, S. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. – *Ecol. Appl.* 19: 181–197.
- Polley, L. and Thompson, R. C. A. 2009. Parasite zoonoses and climate change: molecular tools for tracking shifting boundaries. – *Trends Parasitol.* 25: 285–291.
- Pörtner, H. O. 2010. Oxygen- and capacity-limitation of thermal tolerance: a matrix for integrating climate-related stressor effects in marine ecosystems. – *J. Exp. Biol.* 213: 881–893.
- Pörtner, H. O. and Knust, R. 2007. Climate change affects marine fishes through the oxygen limitation of thermal tolerance. – *Science* 315: 95–97.
- Pörtner, H.-O., Bock, C. and Mark, F. C. 2017. Oxygen- and capacity-limited thermal tolerance: bridging ecology and physiology. – *J. Exp. Biol.* 220: 2685–2696.
- Rohr, J. R. and Cohen, J. M. 2020. Understanding how temperature shifts could impact infectious disease. – *PLoS Biol.* 18: e3000938.
- Rollinson, C. R., Finley, A. O., Alexander, M. R., Banerjee, S., Dixon Hamil, K.-A., Koenig, L. E., Locke, D. H., DeMarche, M. L., Tingley, M. W., Wheeler, K., Youngflesh, C. and Zipkin, E. F. 2021. Working across space and time: nonstationarity in ecological research and application. – *Front. Ecol. Environ.* 19: 66–72.
- Rosenberg, A., Bigford, T., Leathery, S., Hill, R., and Bickers, K. 2000. Ecosystem approaches to fishery management through essential fish habitat. – *Bull. Mar. Sci.* 66: 535–542.
- Rubalcaba, J. G., Verberk, W. C. E. P., Hendriks, A. J., Saris, B. and Woods, H. A. 2020. Oxygen limitation may affect the temperature and size dependence of metabolism in aquatic ectotherms. – *Proc. Natl Acad. Sci. USA* 117: 31963–31968.
- Santika, T. and Hutchinson, M. F. 2009. The effect of species response form on species distribution model prediction and inference. – *Ecol. Modell.* 220: 2365–2379.
- Scheuffele, H., Jutfelt, F. and Clark, T. D. 2021. Investigating the gill-oxygen limitation hypothesis in fishes: intraspecific scaling relationships of metabolic rate and gill surface area. – *Conserv. Physiol.* 9: coab040.
- Seibel, B. A. and Deutsch, C. 2020. Oxygen supply capacity in animals evolves to meet maximum demand at the current oxygen partial pressure regardless of size or temperature. – *J. Exp. Biol.* 223: jeb210492.
- Shono, H. 2008. Application of the Tweedie distribution to zero-catch data in CPUE analysis. – *Fish. Res.* 93: 154–162.
- Smith, J. A., Pozo Buil, M., Fiechter, J., Tommasi, D. and Jacox, M. G. 2022. Projected novelty in the climate envelope of the California Current at multiple spatial-temporal scales. – *PLoS Clim.* 1: e0000022.
- Sogard, S. M. and Berkeley, S. A. 2017. Patterns of movement, growth, and survival of adult sablefish (*Anoplopoma fimbria*) at contrasting depths in slope waters off Oregon. – *Fish. Bull.* 115: 233–251.
- Sunday, J. M., Howard, E., Siedlecki, S., Pilcher, D. J., Deutsch, C., MacCready, P., Newton, J. and Klinger, T. 2022. Biological sensitivities to high-resolution climate change projections in the California Current marine ecosystem. – *Global Change Biol.* 28: 5726–5740.
- Swanson, A. K., Dobrowski, S. Z., Finley, A. O., Thorne, J. H. and Schwartz, M. K. 2013. Spatial regression methods capture prediction uncertainty in species distribution model projections through time. – *Global Ecol. Biogeogr.* 22: 242–251.
- Talluto, M. V., Boulangeat, I., Ameztegui, A., Aubin, I., Berteaux, D., Butler, A., Doyon, F., Drever, C. R., Fortin, M.-J., Franceschini, T., Liénard, J., McKenney, D., Solarik, K. A., Strigul, N., Thuiller, W. and Gravel, D. 2016. Cross-scale integration of knowledge for predicting species ranges: a metamodelling framework. – *Global Ecol. Biogeogr.* 25: 238–249.
- Thorson, J. T., Hermann, A. J., Siwicke, K., Zimmermann, M. and Grand, M. 2021a. Grand challenge for habitat science: stage-structured responses, nonlocal drivers, and mechanistic associations among habitat variables affecting fishery productivity. – *ICES J. Mar. Sci.* 78: 1956–1968.
- Thorson, J., Salguero-Gómez, R., Jones, O. R., Childs, D. Z. and Beckerman, A. P. 2021b. Bridging gaps in demographic analysis with phylogenetic imputation. – *Conserv. Biol.* 35: 1210–1221.
- Thuiller, W., Brotons, L., Araújo, M. B. and Lavorel, S. 2004. Effects of restricting environmental range of data to project current and future species distributions. – *Ecography* 27: 165–172.
- Tourinho, L. and Vale, M. M. 2023. Choosing among correlative, mechanistic and hybrid models of species' niche and distribution. – *Integr. Zool.* 18: 93–109.
- Tweedie, M. C. K. 1984. An index which distinguishes between some important exponential families. In statistics: applications and new directions. – In: Gosh, J. K. and Roy, J. (eds), *Proc. Indian statistical institute golden jubilee international conference*. Indian Statistical Institute, Calcutta, pp. 579–604.
- Urban, M. C. 2019. Projecting biological impacts from climate change like a climate scientist. – *WIREs Clim. Change* 10: e585.
- Urban, M. C. et al. 2016. Improving the forecast for biodiversity under climate change. – *Science* 353: aad8466.
- van Denderen, D., Gislason, H., van den Heuvel, J. and Andersen, K. H. 2020. Global analysis of fish growth rates shows weaker responses to temperature than metabolic predictions. – *Global Ecol. Biogeogr.* 29: 2203–2213. <https://doi.org/10.1111/geb.13189>
- Veloz, S. D., Williams, J. W., Blois, J. L., He, F., Otto-Bliesner, B. and Liu, Z. 2012. No-analog climates and shifting realized niches during the late quaternary: implications for 21st-century predictions by species distribution models. – *Global Change Biol.* 18: 1698–1713.
- Verberk, W. C. E. P., Durance, I., Vaughan, I. P. and Ormerod, S. J. 2016a. Field and laboratory studies reveal interacting effects of stream oxygenation and warming on aquatic ectotherms. – *Global Change Biol.* 22: 1769–1778.
- Verberk, W. C., Overgaard, J., Ern, R., Bayley, M., Wang, T., Boardman, L. and Terblanche, J. 2016b. Does oxygen limit thermal tolerance in arthropods? A critical review of evidence. – *Comp. Biochem. Physiol.* 192: 64–78.
- Verberk, W. C. E. P., Atkinson, D., Hoefnagel, K. N., Hirst, A. G., Horne, C. R. and Siepel, H. 2021. Shrinking body sizes in response to warming: explanations for the temperature–size rule

- with special emphasis on the role of oxygen. – *Biol. Rev.* 96: 247–268.
- Waldock, C., Stuart-Smith, R. D., Albouy, C., Cheung, W. W. L., Edgar, G. J., Mouillot, D., Tjiputra, J. and Pellissier, L. 2022. A quantitative review of abundance-based species distribution models. – *Ecography* 2022: e05694.
- Wang, J.-Y., Kuo, T.-C. and Hsieh, C.-H. 2020. Causal effects of population dynamics and environmental changes on spatial variability of marine fishes. – *Nat. Commun.* 11: 2635.
- Warton, D. I. and Shepherd, L. C. 2010. Poisson point process models solve the pseudo-absence problem for presence-only data in ecology. – *Ann. Appl. Stat.* 4: 1383–1402.
- Whittaker, R. J., Willis, K. J. and Field, R. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. – *J. Biogeogr.* 28: 453–470.
- Williams, J. W. and Jackson, S. T. 2007. Novel climates, no-analog communities and ecological surprises. – *Front. Ecol. Environ.* 5: 475–482.