

# Lawrence Berkeley National Laboratory

## LBL Publications

### Title

Predicting responses to climate change using a joint species, spatially dependent physiologically guided abundance model

### Permalink

<https://escholarship.org/uc/item/11f4m7t8>

### Journal

Ecology, 105(8)

### ISSN

0012-9658

### Authors

Custer, Christopher A  
North, Joshua S  
Schliep, Erin M  
et al.

### Publication Date

2024-08-01

### DOI

10.1002/ecy.4362




### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

## STATISTICAL REPORT

# Predicting responses to climate change using a joint species, spatially dependent physiologically guided abundance model

Christopher A. Custer<sup>1</sup>  | Joshua S. North<sup>2</sup> | Erin M. Schliep<sup>3</sup> |  
 Michael R. Verhoeven<sup>4</sup> | Gretchen J. A. Hansen<sup>4</sup>  | Tyler Wagner<sup>5</sup> 

<sup>1</sup>Pennsylvania Cooperative Fish and Wildlife Research Unit, Department of Ecosystem Science and Management, The Pennsylvania State University, University Park, Pennsylvania, USA

<sup>2</sup>Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

<sup>3</sup>Department of Statistics, North Carolina State University, Raleigh, North Carolina, USA

<sup>4</sup>Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, St. Paul, Minnesota, USA

<sup>5</sup>U.S. Geological Survey, Pennsylvania Cooperative Fish and Wildlife Research Unit, The Pennsylvania State University, University Park, Pennsylvania, USA

### Correspondence

Christopher A. Custer  
 Email: [cac6877@psu.edu](mailto:cac6877@psu.edu)

### Funding information

National Science Foundation, Grant/Award Numbers: EF-163853, EF-163855, EF-1638554, EF-1638679; U.S. Geological Survey, Grant/Award Number: G20AC00096; Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy, Grant/Award Number: DE-AC02-05CH11231; Regional and Global Model Analysis Program area within the Earth and Environmental Systems Modeling Program

**Handling Editor:** Lorenzo Ciannelli

### Abstract

Predicting the effects of warming temperatures on the abundance and distribution of organisms under future climate scenarios often requires extrapolating species–environment correlations to climatic conditions not currently experienced by a species, which can result in unrealistic predictions. For poikilotherms, incorporating species’ thermal physiology to inform extrapolations under novel thermal conditions can result in more realistic predictions. Furthermore, models that incorporate species and spatial dependencies may improve predictions by capturing correlations present in ecological data that are not accounted for by predictor variables. Here, we present a joint species, spatially dependent physiologically guided abundance (jsPGA) model for predicting multispecies responses to climate warming. The jsPGA model uses a basis function approach to capture both species and spatial dependencies. We apply the jsPGA model to predict the response of eight fish species to projected climate warming in thousands of lakes in Minnesota, USA. By the end of the century, the cold-adapted species was predicted to have high probabilities of extirpation across its current range—with 10% of lakes currently inhabited by this species having an extirpation probability >0.90. The remaining species had varying levels of predicted changes in abundance, reflecting differences in their thermal physiology. Though the model did not identify many strong species dependencies, the variation in estimated spatial dependence across species suggested that accounting for both dependencies was important for predicting the abundance of these fishes. The jsPGA model provides a new tool for predicting changes in the abundance, distribution, and extirpation probability of poikilotherms under novel thermal conditions.

### KEYWORDS

joint species, poikilotherms, spatial dependence, thermal response

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Ecology* published by Wiley Periodicals LLC on behalf of The Ecological Society of America. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA.

## INTRODUCTION

Predicting changes in a species' abundance, distribution, and extinction risk is a critical component to understanding how organisms will respond to climate change (Ehrlén & Morris, 2015). A major challenge in these efforts is extrapolating relationships between species abundance or distributions and present-day or historical environmental conditions to future climate scenarios with environmental conditions that have not been experienced by past or contemporary species assemblages. These predictions under novel climatic conditions may be unreliable and unrealistic when derived from traditional modeling approaches, such as species distribution models (SDMs), that rely on the relationship between abundance and the current climate (Urban et al., 2012; Zurell et al., 2016). Incorporating some mechanisms into SDMs, including physiology, to create "hybrid correlative-mechanistic" models has been suggested as an important initial approach, particularly since detailed information is not available for most species to create mechanistic models that describe responses to climate change (Urban et al., 2016).

Poikilotherms, which rely on ambient temperature for thermoregulation and metabolic function, are among the most susceptible to climate change (Paaijmans et al., 2013). Because of the strong influence of environmental temperatures on the vital rates of poikilotherms (Deutsch et al., 2008), incorporating information on thermal physiology—such as thermal preference and tolerance—into hybrid correlative-mechanistic SDMs can lead to more realistic predictions under novel temperature conditions (Wagner et al., 2023). A recently developed statistical methodology, the physiologically guided abundance (PGA) model (Wagner et al., 2023), fuses a species' thermal response curve (and associated uncertainty) with a correlative niche model (CNM), such that abundance estimates and distributions are scaled based on species-specific thermal tolerances and preferences (see *The PGA model* for details). In other words, the thermal response curve incorporates a species' thermal physiology such that, with all other parameters held constant, abundance estimates are highest when ambient temperatures are at a species' thermal optimum ( $T_{opt}$ ) and decrease as temperatures approach the species' critical thermal maximum ( $CT_{max}$ ). Importantly, incorporating thermal response curves in this way enables predictions under novel thermal conditions that are informed by species physiology and allows for predictions of the probability of local extirpation events when thermal habitat exceeds the species'  $CT_{max}$  (under the assumption that the laboratory-derived  $CT_{max}$  represents lethal conditions in the wild). Although the

existing PGA model is a significant advancement in predicting the effects of climate change on poikilotherm abundance and distributions under novel thermal conditions, the model does not account for correlations among multiple species or spatial dependencies that are present in ecological data (Ovaskainen et al., 2016). A probabilistic model that leverages the dependencies across space, among species, and in relation to climate can lead to improved prediction and uncertainty estimation under future climates.

The importance of accounting for residual dependencies among species, especially when predicting species responses to environmental change, is reflected by the recent development and application of multispecies SDMs (e.g., Clark et al., 2014; Ovaskainen et al., 2010; Pichler & Hartig, 2021; Pollock et al., 2014; Thorson et al., 2016; Wilkinson et al., 2019). Such joint species distribution models (JSDMs) are a family of statistical methodologies that simultaneously model the distributions of multiple species by accounting for both abiotic environmental covariates and residual dependencies among species (Wilkinson et al., 2019). Species dependencies capture potential biotic interactions and missing environmental factors, thereby potentially improving predictive performance through the use of conditional predictions (Poggiato et al., 2021; Tikhonov et al., 2020). Accounting for species dependencies has been shown to improve predictive performance when modeling the distribution of fish communities (Clark et al., 2014; Warton et al., 2015). For freshwater fishes, Wagner et al. (2020) showed that conditional predictions (i.e., conditionally dependent on the occurrence of other species) of species occurrence significantly outperformed marginal predictions (i.e., independent of the occurrence of other species) for most species across stream and lake fish community data sets.

The prediction of species abundances and distributions across a landscape is inherently a spatial prediction problem (Latimer et al., 2006). Scientists often use SDMs to describe a species' occurrence or abundance with respect to some number of measured environmental constraints (Planque et al., 2011), essentially modeling a species' niche (Sillero et al., 2021). However, these environmental constraints do not always completely explain the distribution of a species across space. Ecological processes such as dispersal and reproduction often lead to spatial dependencies in species abundance, whereby sampled locations closer together tend to be more similar compared to sampled locations that are farther apart (Hefley et al., 2017; Latimer et al., 2006). Accounting for spatial dependence within a SDM can improve predictions of a species' distribution and abundance across space (Guélat & Kéry, 2018; Hefley et al., 2017; Record et al., 2013), as well

as affect the precision of the estimated environmental coefficients (Miller, 2012). Lake fish communities are particularly spatially autocorrelated when compared to other organisms and freshwater habitats (Shurin et al., 2009), suggesting that lakes that are in close proximity are more likely to have similar community compositions than lakes that are farther apart.

Multiple methodologies exist for incorporating spatial dependence into ecological models, and the optimal approach depends on the type of spatial data and the objective of the study. For example, conditional autoregressive models are a common approach for areal spatial data, whereas point spatial data are often modeled using geostatistical models and Gaussian processes (Paciorek, 2013). Much of the recent development of JSDMs has included efforts to incorporate spatial autocorrelation within their framework. This allows for the model to simultaneously account for both species and spatial dependencies alongside the included environmental covariates assumed to be driving species' distribution and abundance. Common approaches for modeling spatial dependencies within a JSDM framework have included latent variable models (LVMs) (e.g., Ovaskainen et al., 2016; Thorson et al., 2016) and basis function approaches (e.g., Hefley et al., 2017; Hui et al., 2023). While both approaches provide flexibility, the basis function approach has been shown to be more scalable and computationally efficient as it uses predefined spatial basis vectors within the model fitting process (Hui et al., 2023).

Here, we present a statistical method for incorporating physiological thermal tolerances, species dependencies, and spatial dependencies into a statistical model of the abundance of multiple species. This joint species, spatially dependent physiologically guided abundance (jsPGA) model provides a flexible approach to predicting the abundance, distribution, and extirpation probabilities of poikilotherm communities under future climate scenarios. The jsPGA model accounts for aspects of a species distribution not present within the modeled environmental factors (e.g., potential species interactions, random spatial processes, or important environmental constraints not included in the model). We begin with the construction of the jsPGA model, which we formulate under a Bayesian hierarchical framework. We then perform a simulation study for model validation and comparison across differing numbers of basis functions (DiRenzo et al., 2023). Finally, we apply the jsPGA to a case study of Minnesota lake fish communities to illustrate how our methodology can be used to predict future abundances, ranges, and probabilities of multiple species that vary in thermal physiology.

## MODEL SPECIFICATION

### The PGA model

The PGA model developed by Wagner et al. (2023) employs a Poisson modeling framework to model the counts of a single species where the intensity is defined as the product of relative abundance constrained by the species' thermal physiology and an effort offset term. This effort offset term, or scaling, accounts for both the varying catchability and sampling effort. Modeling catchability and sampling effort allows the relative abundance to be estimated separately from the effect of sampling design to account for differences in sampling efficiencies across different gear types. Physiological information was incorporated into the relative abundance model through a thermal performance curve (see Appendix S1: Figure S1 for examples). Let  $C_{ijt}$  be the number of fish caught in lake  $i = 1, \dots, I$ , using sampling gear  $j = 1, \dots, J$  in year  $t = 1, \dots, T$ . The catch data are modeled as

$$C_{ijt} \sim \text{Pois}(\tilde{E}_{ijt}\lambda_{it}), \tag{1}$$

where  $\tilde{E}_{ijt}$  is the effort scaling associated with each sample, and  $\lambda_{it}$  captures the relative abundance in lake  $i$  for year  $t$ . The effort scaling is factored as  $\tilde{E}_{ijt} = E_{ijt}\theta_j$ , where  $E_{ijt}$  is the known effort and  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_J]'$ ,  $\theta_j > 0$  for all  $j$  is the catchability vector capturing differences in sampling efficiency across gear types. The catchability vector is constrained such that  $\sum_{j=1}^J \theta_j = J$ , whereby larger (or smaller) values signify a gear is more (or less) effective at catching fish. Note that  $\theta_j = 1$  for all  $j = 1, \dots, J$  implies a constant catchability across all gears and is a special case.

Relative abundance is modeled as

$$\lambda_{it} = P(T_{it}) \exp(\mathbf{X}'_{it}\boldsymbol{\beta}), \tag{2}$$

where  $\mathbf{X}_{it}$  is a vector of covariates hypothesized to explain the variation in species distribution and abundance for lake  $i$  and year  $t$ , and  $\boldsymbol{\beta}$  is the coefficient vector. The thermal performance scalar,  $P(T_{it})$ , is derived from the species-specific thermal performance function evaluated at temperature  $T_{it}$ . The values of the function  $P(T_{it})$  range from 0 (thermally unsuitable if temperatures exceed  $CT_{\max}$ ) to 1 (optimal performance at  $T_{\text{opt}}$ ).

The model assumed an asymmetric thermal performance curve that uses a Gaussian function to describe the ascending limb of the thermal performance curve up to  $T_{\text{opt}}$  and a quadratic decline to 0 at  $CT_{\max}$  for the descending limb (Gannon et al., 2014; Payne et al., 2018), although different thermal performance functions could



be implemented as needed (Padfield et al., 2021). The performance curve implemented was previously used in studies on the effects of changing temperatures on poikilotherms (Guo et al., 2020; Lear et al., 2019). The thermal performance curve was parameterized as

$$P(T) = \begin{cases} \exp\left(-\left(\frac{T-T_{\text{opt}}}{2\sigma}\right)^2\right) & T \leq T_{\text{opt}} \\ 1 - \left(\frac{T-T_{\text{opt}}}{T_{\text{opt}}-CT_{\text{max}}}\right)^2 & T_{\text{opt}} < T \leq CT_{\text{max}} \\ 0 & T > CT_{\text{max}}, \end{cases} \quad (3)$$

where  $\sigma$  is the scale parameter for the Gaussian portion of the curve and all other parameters are as described earlier. Although all parameters ( $T_{\text{opt}}$ ,  $CT_{\text{max}}$ , and  $\sigma$ ) can theoretically be estimated using abundance data, if sample locations do not span the entire temperature range of a species' current distribution, then estimated parameters may be biologically inaccurate (e.g., an underestimated  $CT_{\text{max}}$ ), highly uncertain, or both. This is likely the case for many poikilotherms, where abundance data are only available within a portion of a species' range. However, this issue is addressed by explicitly accounting for uncertainty in the thermal response curves (see *Uncertainty in thermal response curves*).

## The jsPGA model

We propose the jsPGA model as an extension to the PGA model. We reformulate the relative abundance model (Equation 2) to be applicable for multiple species by adding a random effect to account for species and spatial dependence. In addition, we model the catch data using a negative binomial (NB) with an overdispersion parameter in place of the Poisson model in Equation (1) since count data have the potential to be overdispersed (Stoklosa et al., 2022).

Let  $C_{ijk}$  denote the number of fish caught in lake  $i = 1, \dots, I$  using gear  $j = 1, \dots, J$  of species  $k = 1, \dots, K$  in year  $t = 1, \dots, T$ . We model the fish catch data as

$$C_{ijk} \sim \text{NB}(\tilde{E}_{ijk}\lambda_{itk}, \phi), \quad (4)$$

where  $\phi$  is the NB overdispersion parameter,  $\tilde{E}_{ijk} = E_{ijt}\theta_{jk}$ , and  $\theta_{jk}$  denotes the sampling efficiency of gear  $j$  in catching species  $k$ . Define  $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_K]$  as the  $J \times K$  catchability matrix for all gears and species where  $\boldsymbol{\theta}_k = [\theta_{1k}, \dots, \theta_{Jk}]'$ .

To account for multiple species, we extend the model for relative abundance,  $\lambda_{itk}$ , to

$$\lambda_{itk} = P_k(T_{it}) \exp(\mathbf{X}'_{it}\boldsymbol{\beta}_k + \omega_{ik}), \quad (5)$$

where now  $\boldsymbol{\beta}_k$  is the species-specific coefficient vector,  $\omega_{ik}$  is the random effect capturing the variation in abundance for lake  $i$  and species  $k$  not captured by the covariates, and  $P_k(T_{it})$  is the species-specific thermal performance scalar derived from Equation (3). Let  $\boldsymbol{\omega}$  denote the  $I \times K$  dimension matrix form of the random effects, where  $\omega_{ik}$  is the  $i, k$ th element. In modeling  $\boldsymbol{\omega}$ , we follow the basis function approach to capture spatial dependence (e.g., Hefley et al., 2017; Hui et al., 2023) and capture species dependence through the specification of the basis function coefficients. Thus, we define  $\boldsymbol{\omega}$  as

$$\boldsymbol{\omega} = \boldsymbol{\Psi}\mathbf{A}, \quad (6)$$

where  $\boldsymbol{\Psi}$  is an  $I \times M$  matrix of spatial basis vectors, and  $\mathbf{A}$  is an  $M \times K$  matrix of spatial basis coefficients such that  $M \in \{1, \dots, I\}$  (see *Spatial basis vectors* for more details). The basis coefficients  $\mathbf{A}$  are then defined such that

$$\mathbf{A} \sim \text{MN}(0, \mathbf{I}_M, \boldsymbol{\Pi}), \quad (7)$$

where MN is the matrix normal distribution and is parameterized such that  $\mathbf{I}_M$ , denoting an  $M \times M$  identity matrix, is the covariance among rows (i.e., spatial basis coefficients are independent *between* sites) and  $\boldsymbol{\Pi} = \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}$  is the covariance among columns (i.e., spatial basis coefficients are dependent *within* a site). We define  $\mathbf{T} = \text{diag}(\tau_1, \dots, \tau_K)$  as the species-specific scalings and  $\boldsymbol{\Sigma}$  as the  $K \times K$  species dependence (i.e., correlation) matrix. In other words, we capture spatial dependence through the spatial basis vectors in  $\boldsymbol{\Psi}$  and species dependence through the covariance structure  $\boldsymbol{\Pi}$  of the spatial basis coefficients in  $\mathbf{A}$ .

## Spatial basis vectors

Let  $N$  represent all (sampled  $I$  and unsampled  $N - I$ ) sites of interest. The spatial basis vectors are defined using the spectral decomposition of the Matérn kernel of the pairwise distances between all  $N$  sites, denoted by  $R_{\nu, \rho}(\mathbf{s}, \mathbf{s}')$ , where  $\nu$  and  $\rho$  are parameters that describe the differentiability and length scale of the implied stochastic process, respectively, and  $\mathbf{s}, \mathbf{s}'$  are two locations in the spatial domain. Let  $\mathbf{R}_{\nu, \rho}$  denote the  $N \times N$  Matérn covariance matrix calculated from the pairwise distances  $\mathbf{s}, \mathbf{s}'$  in the spatial domain. Then the spectral decomposition is such that  $\mathbf{R}_{\nu, \rho} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}$ , where  $\mathbf{Q}$  is an  $N \times N$  matrix where each column represents an eigenvector of  $\mathbf{R}_{\nu, \rho}$  and  $\boldsymbol{\Lambda}$  is a diagonal matrix where each element of the

diagonal represents the corresponding eigenvalue of  $\mathbf{R}_{\nu,\rho}$  (Hefley et al., 2017). We then define the full matrix of basis vectors as  $\Psi_{N \times N} = \mathbf{Q}\mathbf{\Lambda}^{1/2}$ . Finally, to reduce the dimension of the model, we retain only the first  $M$  columns of this matrix corresponding to the desired number of basis functions. Here,  $M \ll I$ , where  $I$  is the number of observed sites. Let  $\Psi$  denote the  $I \times M$  basis function matrix corresponding to the first  $I$  rows and  $M$  columns of  $\Psi_{N \times N}$ . Higher values of  $M$  allow for the basis vectors to potentially capture more spatial dependence and also increases the number of parameters to be estimated by the model. Recall that the basis coefficient matrix  $\mathbf{A}$  is of  $M \times K$  dimension, meaning every additional basis function requires  $K$  more basis coefficient parameters to be estimated. Within our analyses, we set  $\rho = 100$  km to be approximately one-sixth of the maximum distance between any two locations and  $\nu = 2.5$ , so the process is moderately smooth.

## Uncertainty in thermal response curves

Thermal response curves are themselves uncertain, as the true values of  $T_{\text{opt}(k)}$  and  $\text{CT}_{\text{max}(k)}$  are unknown, and there is considerable variability in their published values (Appendix S1: Table S1). The thermal response parameters were most commonly derived from growth rates (for  $T_{\text{opt}}$ ) and loss of equilibrium for  $\text{CT}_{\text{min}}$  and  $\text{CT}_{\text{max}}$ . Using the growth of individuals to quantify  $T_{\text{opt}}$  is useful for fishes because it is related to survival, reproductive potential, and other life history traits (Charnov, 2008; Charnov et al., 2013; Lester et al., 2004), while loss of equilibrium is commonly assumed to represent death under wild conditions (Turko et al., 2020). Furthermore, thermal response curve parameters are usually obtained through controlled laboratory experiments and likely do not accurately reflect the adaptive plasticity of organisms in the wild (Morgan et al., 2019). Since we do not have observed temperatures that span the full range of values for which the thermal response curves are estimated, we lack the information required to accurately estimate the thermal performance parameters ( $T_{\text{opt}}$  and  $\text{CT}_{\text{max}}$ ). Using the joint posterior of all parameters would be too heavily weighted by the likelihood given the data and overwhelm the information regarding these parameters derived from the literature. Therefore, we use numerical integration to incorporate uncertainty in these parameters. Specifically, we assign  $T_{\text{opt}(k)}$  and  $\text{CT}_{\text{max}(k)}$  truncated normal distributions where the mean and variance are derived from literature values. We truncate these distributions based on the maximum observed temperature of occurrence for each species to ensure that a random realization of  $\text{CT}_{\text{max}}$  is not lower than the observed maximum

temperature where the species is found in the wild (recall Equation 3 forces estimates of relative abundance to zero when ambient temperature exceeds  $\text{CT}_{\text{max}}$ ). Using the truncated distributions of thermal performance parameters, we randomly sampled 100 realizations of  $T_{\text{opt}(k)}$  and  $\text{CT}_{\text{max}(k)}$  (ensuring that  $T_{\text{opt}(k)} < \text{CT}_{\text{max}(k)}$ ) for each species and fit our jsPGA model using each of these realizations (see Appendix S1: Figure S1 for how the thermal performance curves vary across realizations). Therefore, we fit 100 separate models, each with a different pair of randomly sampled thermal response curve parameters—parameters that are treated as fixed values within each model fit. We then obtained the joint posterior distribution of all other model parameters from each of the 100 model fits before aggregating these into a single posterior distribution for each parameter (note we calculated a single value of  $\sigma_k$  and treated it as fixed across all model fits; see Wagner et al. [2023] for details).

## Bayesian model inference

The jsPGA model was fitted in a Bayesian framework using a Hamiltonian Monte Carlo sampling algorithm implemented in Stan (Carpenter et al., 2017). The hierarchical Bayesian framework allows us to directly incorporate species' physiological information (i.e., thermal performance parameters) and uncertainty into the model through the use of their literature-defined distributions as described in *Uncertainty in thermal response curves*. We then assigned prior distributions to all remaining model parameters. For each catchability parameter vector,  $\theta_k$ , we assigned a Dirichlet prior distribution and scaling such that  $\sum_{j=1}^J \theta_{jk} = J$  for each  $k = 1, \dots, K$ . We then assigned  $\beta_{lk} \sim N(0, 100)$  for  $l = 0, \dots, p$ , where  $p$  is the total number of environmental covariates. For the spatial basis coefficient covariance parameters, we assigned  $\tau_k \sim \text{Cauchy}(0, 2.5)$  and  $\Sigma \sim \text{LKJ}(2)$ , where LKJ is the Lewandowski–Kurowicka–Joe distribution (see provided Stan code for full details of optimization techniques used for estimation of  $\mathbf{A}$  and  $\Sigma$ ). All models were fitted using the `cmdstanr` package (Gabry & Češnovar, 2022) in R (R Core Team, 2022) on the Pennsylvania State University's high-performance research cloud (ROAR) using parallel computing. Nodes within the system have 256 GB of memory and 3.0 GHz base processing speed (turbo frequency 3.9 GHz).

## Simulation study

We performed a simulation study to assess the impact the choice of the number of basis functions  $M$  has on

parameter accuracy and precision. While reducing the number of basis functions can increase bias, it can also decrease variance and lead to a more parsimonious model. This simulation study helped to inform the choice of  $M$  for our case study.

The simulated data were informed by our case study of Minnesota lake fishes, though at a reduced size to decrease computation time (see *Fish and environmental data* for more detail). The simulated data set had  $K = 5$  species,  $I = 500$  lakes (we did not simulate unsampled lakes, so  $N = I$ ), and  $J = 2$  gears. The  $I = 500$  lakes were randomly sampled from the full Minnesota lake fish data set, and their corresponding information for effort, environmental factors, and spatial location were used to ensure a realistic gradient of variables across the entire sample region of the simulation. The full spatial basis vector matrix,  $\Psi_{N \times N}$ , was used in simulating the catch data. The models were fitted based on the truncated  $I \times M$   $\Psi$  basis function matrix for values of  $M = 16, 32, 64,$  and  $128$ . This allowed us to measure how reducing the rank (i.e., reducing the number of basis functions) of our spatial random effect influenced parameter estimation. These four models were fitted 100 times, once for each randomly sampled value of  $CT_{\max}$  and  $T_{\text{opt}}$ . That is, a model with  $M = 16, 32, 64,$  and  $128$  basis vectors was fitted for each realization of the pair  $CT_{\max}$  and  $T_{\text{opt}}$ . Using Stan, 2000 samples from the posterior were obtained for each pair  $CT_{\max}$  and  $T_{\text{opt}}$ . The first thousand iterations were discarded as burn-in, and every fifth sample was retained for posterior inference. Output from Stan gave no indication to suggest a lack of convergence, which was confirmed via graphical checks of model parameters.

We evaluated model performance based on the root mean square error (RMSE) and the continuous ranked probability score (CRPS; Gneiting & Raftery, 2007) for each model parameter. Each metric was then summarized across species, lake, and set of parameters for each of the  $M = 16, 32, 64,$  and  $128$  models. The evaluation metrics suggested that our choices of  $M$  did not significantly influence the parameter estimates (Table 1). That is to say, increasing  $M$  showed very minor differences within our parameter estimates as measured by RMSE and CRPS. The biggest improvement as measured by RMSE was seen in the estimates for  $\Pi$  going from  $M = 16$  to  $32$  with RMSE values of  $0.5237$  and  $0.4871$ , respectively. The biggest improvement as measured by CRPS was seen in the estimates for  $\omega$  going from  $M = 16$  to  $32$  with CRPS values of  $0.3059$  and  $0.2880$ , respectively. Given these minor differences, we opted for the most parsimonious model with  $M = 16$  for our case study analysis.

**TABLE 1** Summarized evaluation metrics for parameter estimation from a simulation study across increasing number ( $M$ ) of basis functions included within the joint species spatially dependent physiologically guided abundance model.

Parameter	$M = 16$	$M = 32$	$M = 64$	$M = 128$
<b><math>\beta</math></b>				
RMSE <sup>a</sup>	1.690	1.693	1.694	1.687
CRPS <sup>a</sup>	0.601	0.569	0.569	0.566
<b><math>\theta</math></b>				
RMSE <sup>b</sup>	7.694	7.649	7.655	7.655
CRPS <sup>b</sup>	3.474	3.452	3.436	3.425
<b><math>\omega</math></b>				
RMSE <sup>a</sup>	8.342	8.359	8.344	8.299
CRPS <sup>a</sup>	3.059	2.880	2.875	2.862
<b><math>\Pi</math></b>				
RMSE <sup>a</sup>	5.740	5.324	5.124	5.112
CRPS <sup>a</sup>	1.905	2.010	2.007	1.991

Note: Metrics were individually calculated for the full dimension of each parameter and then summarized into a single mean value for each model. A lower value represents a better estimate for both metrics.

Abbreviations: CRPS, continuous ranked probability score; RMSE, root mean square error.

<sup>a</sup> $\times 10^{-1}$ .

<sup>b</sup> $\times 10^{-3}$ .

## Case study: Minnesota lake fish

We illustrate the implementation of the jsPGA model using inland lake fish community data from Minnesota, USA. Lakes in this north-temperate region approximate the current northern and southern boundaries of many warm- and cold-water species, respectively, presenting an interesting case study for predicting shifts in the abundance and distribution of fish with differing thermal tolerances in response to a warming climate. For example, the case study of Minnesota lake fish using the single-species PGA model predicted that warming lake temperatures would lead to extirpation of cisco *Coregonus artedii*, a cold-water (cold-adapted) fish, in many Minnesota lakes, while bluegill *Lepomis macrochirus*, a warm-water (warm-adapted) fish, would see increases in abundance (Wagner et al., 2023). Here, we expand this analysis using the jsPGA model to jointly model eight species that represent a range of thermal tolerances (Appendix S1: Table S1), including black crappie *Pomoxis nigromaculatus*, bluegill, largemouth bass *Micropterus salmoides*, northern pike *Esox lucius*, smallmouth bass *Micropterus dolomieu*, walleye *Sander vitreus*, yellow perch *Perca flavescens*, and the cold-water species cisco.

The jsPGA model provides a number of the aforementioned advantages over a traditional CNM and PGA model by accounting for a species thermal physiology (CNM) and species and spatial dependencies (CNM and PGA). The single-species PGA model showed that incorporating a species thermal physiology resulted in more realistic predictions of abundance for freshwater fishes (Wagner et al., 2023). For example, Wagner et al. (2023) revealed that using the observed relationship between environmental temperature and bluegill abundance led to unrealistic predictions under a warming climate. Specifically, making predictions using a model based solely on the observed positive abundance—temperature relationships led to unlimited bluegill population growth as lake temperatures warmed, even if future water temperatures exceeded the species  $CT_{max}$ . In contrast, the PGA model effectively scaled abundance predictions based on the bluegill's thermal physiology. The jsPGA model further improves predictions of abundance and distributions under a warming climate by accounting for species and spatial dependencies. In the case of bluegill, which are a common prey species for piscivorous fish, such as largemouth bass and walleye (Tomcko & Pierce, 2005), incorporating species dependencies into the modeling structure can potentially capture these predator–prey dynamics, an important ecological process regulating bluegill populations (Savino & Stein, 1982; Tomcko & Pierce, 2005). Species and spatial dependencies may also capture important habitat characteristics that were not available for inclusion as predictor variables in our model (see *Fish and environmental data* for details). For example, aquatic vegetation mediates the predator–prey dynamics between largemouth bass and bluegill (Savino & Stein, 1982; Trebitz et al., 1997). While there is a correlation between aquatic vegetation and water clarity (an environmental factor included in our analysis; Kosten et al., 2009), the observed relationship between fish abundance and measured water clarity may not fully capture the finer-scale effects of the vegetation biomass, distribution, and species composition of each lake. The relationship between predator, prey, and habitat may be more fully accounted for by the jsPGA model through the estimated parameters quantifying species dependencies. Furthermore, the distribution of vegetation has been shown to be spatially dependent, suggesting that incorporating spatial dependencies may further enhance our model's ability to compensate for missing abiotic factors (Miller et al., 2007). Vegetation and predator–prey dynamics are just two examples of important ecological factors for freshwater fishes that may be accounted for by using a model, such as the jsPGA, that incorporates both species and spatial dependencies.

## Fish and environmental data

Environmental covariates included water clarity (Secchi disk depth, in meters), lake area (in hectares), lake elevation (in meters), and proportion of agriculture and urban land use in the lake watershed. Water clarity was derived from annual median values of remotely sensed water clarity calibrated to Secchi disk depth (Max Gilnes, Rensselaer Polytechnic Institute, Troy, NY, USA, August 2023, written communication) and calculated as the 5-year rolling mean. Lake area, elevation, and proportion of land cover within a lake's watershed were obtained from LAGOS-US GEO version 1.0 (Smith et al., 2022). Proportions of land cover within the LAGOS-US data set are from the National Land Cover Database (NLCD; Homer et al., 2020) and include data from multiple years (2001, 2004, 2006, 2008, 2011, 2013, and 2016). Therefore, the nearest temporal land-use data were joined to each sampling record. Lake area was log-transformed and land use was logit-transformed prior to standardizing all covariates to a mean of zero and standard deviation (SD) of one. Lake surface water temperature data were obtained from process-based models of lake temperature profiles for contemporary and future climate conditions (Corson-Dosch et al., 2023). To fit the model, we used the daily lake surface temperature data generated using the General Lake Model (Hipsey et al., 2019) driven by contemporary climate drivers from the North American Land Data Assimilation System (NLDAS) and summarized into a 5-year rolling mean of average July surface temperatures with respect to the sampling year (Corson-Dosch et al., 2023).

Fish catch data were collected by the Minnesota Department of Natural Resources using standard sampling methodologies between 1998 and 2019 (Minnesota Department of Natural Resources (MNDNR), 2017). We restricted our analysis to those samples collected between 1 June and 30 September. All species were sampled using standard gill nets and trap nets, commonly used gear for assessing the relative abundance of fishes in littoral (nearshore) and pelagic (offshore) zones of inland lakes. One unit of sampling effort consisted of one net (gill net or trap net) deployed for a 24-h sampling period. A total of 1754 lakes were sampled. Unsampled lakes for predicting species responses to increased temperature included were lakes within Minnesota with a surface area greater than 0.40 ha that contained all environmental data described earlier, resulting in 3996 total lakes to be used for future predictions (2242 unsampled lakes and 1754 sampled for fish communities). Species thermal tolerance values,  $CT_{max}$ ,  $T_{opt}$ , and  $CT_{min}$  were obtained from the literature (Comte & Olden, 2017; Wismer & Christie, 1987, and references therein) and summarized



as means and SDs for each species (Appendix S1: Table S1). Values of  $CT_{\min}$  were used to estimate  $\sigma$  (Equation 3) following  $\sigma = T_{\text{opt}} - CT_{\min}/4$  because estimates for  $\sigma$  were not readily available in the literature (Deutsch et al., 2008). Estimates for  $CT_{\min}$  were also scarce in the literature for some species and supplemented with values of lower incipient lethal temperature (LILT).

## Future climate predictions

Future lake temperatures were simulated using the General Lake Model driven by downscaled climate drivers from six global climate models (GCMs) under a high-emissions scenario (Corson-Dosch et al., 2023). Projected lake temperature data were available for historical (1981–2000), mid-century (2040–2059), and late-century (2080–2099) surface temperature predictions (Corson-Dosch et al., 2023). To minimize bias introduced by the use of different source models for the predicted lake surface temperatures, we calculated lake-specific differences in predicted temperatures between historical and late-century eras from all six GCMs (Appendix S1: Figure S2) and added those differences to the corresponding 20-year average (1981–2000) NLDAS temperatures. This allowed us to maintain a consistent baseline source (NLDAS) for fitting the model and predicting in the future while taking advantage of the projected changes in temperatures provided by the GCMs. For this analysis, we focus on predictions under current conditions (using the most recent 20-year average, 2002–2021, of NLDAS temperatures to match the temporal resolution) and predictions under late-century conditions, as described earlier. Environmental conditions, other than water temperature, were held constant for climate predictions, so any predicted changes across time are due to predicted changes in surface water temperatures. In order to incorporate variability among GCM water temperature predictions, predictions for thermal performance scalars ( $P_k(T_{it})$  in Equation (5)) and relative abundance under future conditions were made using predicted temperatures from each of the six GCMs and then summarized into a single prediction by taking the mean of those predicted values. Posterior mean estimates of predicted relative abundance ( $\lambda_{itk}$  in Equation 5) across Minnesota lakes at different time periods were used to summarize the predicted response of each species to warming water temperatures. The probability of extirpation was estimated from the posterior distribution of the thermal performance scalars for each species at each lake. Predictions of local extirpations occurred when the

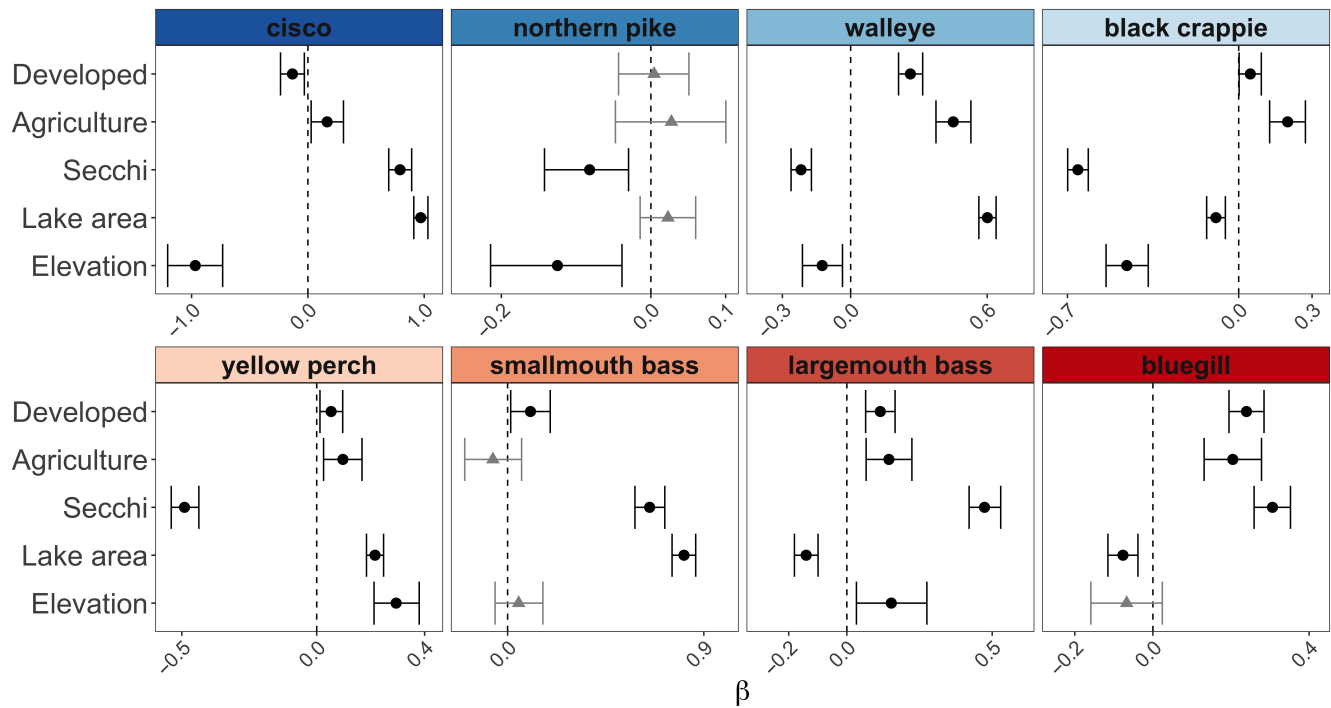
future lake surface water temperature exceeded a model's realization of  $CT_{\max}$  (which accounted for the uncertainty in  $CT_{\max}$ ), resulting in the thermal performance scalar ( $P_k(T_{it})$  of Equation (5)) to equal zero. Here, we focus on predictions at lakes where each species is currently found to emphasize the risk of extirpation where we know a species currently occurs.

## DISCUSSION

### Parameter estimates

The jsPGA modeling framework enabled the prediction of changes in habitat suitability and relative abundance under future thermal conditions in north-temperate lakes by accounting for species physiology, while simultaneously accommodating both species and spatial dependencies. As expected, parameter estimates varied across species (Appendix S1: Table S2). Across all  $\beta$  coefficient parameters, (five environmental covariates across eight species), 34 (85%) had 95% credible intervals (CI) that did not contain zero, and each species had at least one environmental coefficient that had a 95% CI that did not contain zero, suggesting that the included environmental covariates were important determinants of the distributions and abundance of the eight fish species (Figure 1). Total agricultural and developed land cover were generally significantly positively related to the abundance of all species except cisco, suggesting that human land-use activities in a lake's watershed were associated with increased abundance of most species. While cisco were the exception and negatively associated with watershed development, they were more strongly negatively associated with elevation and positively associated with lake area and Secchi depth (i.e., the largest absolute magnitude of the coefficient estimates), suggesting that among the species included in our analysis, cisco had the most restrictive habitat requirements. Specifically, our analysis suggests that cisco prefer large, low-elevation lakes with high water clarity. Cisco are known to require cold, well-oxygenated waters and can be particularly susceptible to lake eutrophication (Lyons et al., 2018). High water clarity is associated with lower levels of productivity and, therefore, higher levels of dissolved oxygen. Furthermore, research has suggested that cisco undergo seasonal migrations between shallower nearshore habitat to deeper offshore habitat due to their oxythermal habitat requirements, which is likely reflected within our model through the strong, positive relationship between lake area and cisco abundance (Kao et al., 2020).

Estimates for  $\theta$  suggested that most of the species had higher catchability with gill nets compared with trap nets.



**FIGURE 1** Estimates for  $\beta$  for each environmental covariate for each Minnesota, USA, lake fish species. Points represent posterior medians, and bars represent 95% credible intervals. Black circles and bars represent estimates with credible intervals that do not overlap zero, and gray triangle and bar estimates represent intervals that do overlap zero. A horizontal dashed line is added at 0 to aid visualization. Species are arranged (left to right, top to bottom; cold  $\rightarrow$  warm colors) in order of increasing  $T_{opt}$ .

The exceptions were black crappie, which had nearly equal estimates of catchability across gear types, and bluegill, which had higher catchability using trap nets. This result highlights the importance of accounting for differences in sampling efficiency across gear types and species. Allowing for catchability to vary across gear and species prevents the model from over- or underestimating relative abundance due to difference in gear types. Here, with the exception of black crappie, we see that gill nets and trap nets were very different in their sampling efficiency across species, and therefore would likely have significant impacts on relative abundance estimates if not accounted for within our model.

The effect sizes associated with the spatial basis coefficients,  $\mathbf{A}$ , varied across species. For cisco, median estimates for the spatial basis coefficients were as large as 37.3 and  $-88.9$ , yet the largest values for northern pike were 1.1 and  $-1.6$ . The frequency at which 95% CIs overlapped zero also varied by species, with the lowest at 6.25% for cisco and highest at 68.8% for northern pike. Though the direct inferences of these effect sizes are of little interest, these results reflect the spatial variation in abundance across the region not accounted for by the covariates, emphasizing the importance of the random effect  $\omega$  (Equation 5) in our model. This variation across species was further reflected within the basis

coefficient's covariance  $\mathbf{\Pi}$  and its components,  $\mathbf{T}$  and  $\mathbf{\Sigma}$ . The species included in this analysis exhibited both negative and positive associations (Appendix S1: Figure S3). Bluegill and black crappie exhibited the strongest positive correlation (0.7; 95% CI 0.38, 0.87), while bluegill and smallmouth bass exhibited the largest estimated negative species correlation was  $-0.6$  ( $-0.81, -0.25$ ). Of the 28 pairwise species correlations, 7 (25%) had 95% CIs that did not overlap zero indicating significant residual dependence between these species pairs. Smallmouth bass accounted for nearly half of these significant species correlations and had negative associations with black crappie, bluegill, and cisco. These estimates may be capturing diverging habitat requirements not included within our environmental covariates or ecologically important biotic interactions. Smallmouth bass are known predators of cisco and are associated with expanding distributions within the north-temperate regions of North America in response to warming temperatures (Van Zuiden et al., 2016; Zanden et al., 2004) and expanding littoral zones (Robillard & Fox, 2006; Stasko et al., 2015). The estimated negative association between these two species may reflect the simultaneous influence of negative biotic interactions (predator-prey) and shifting habitat suitability across the region.

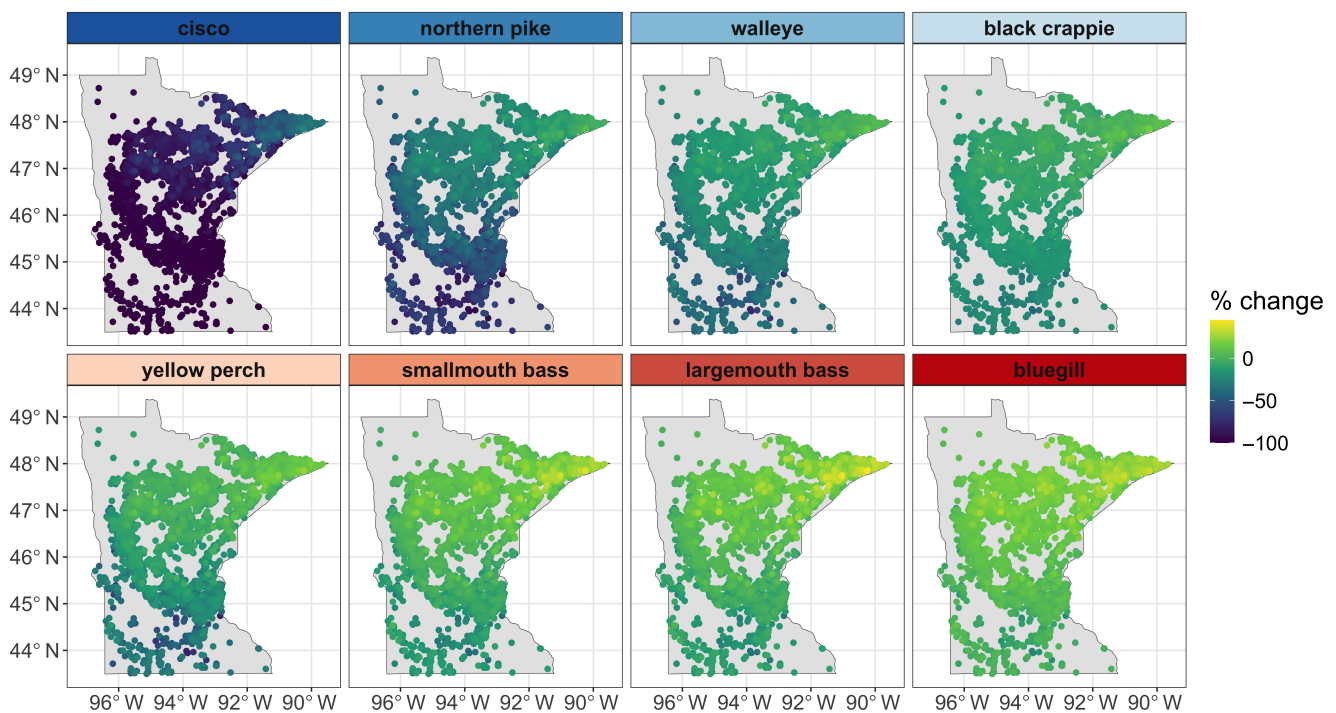


## Predictions of climate change responses

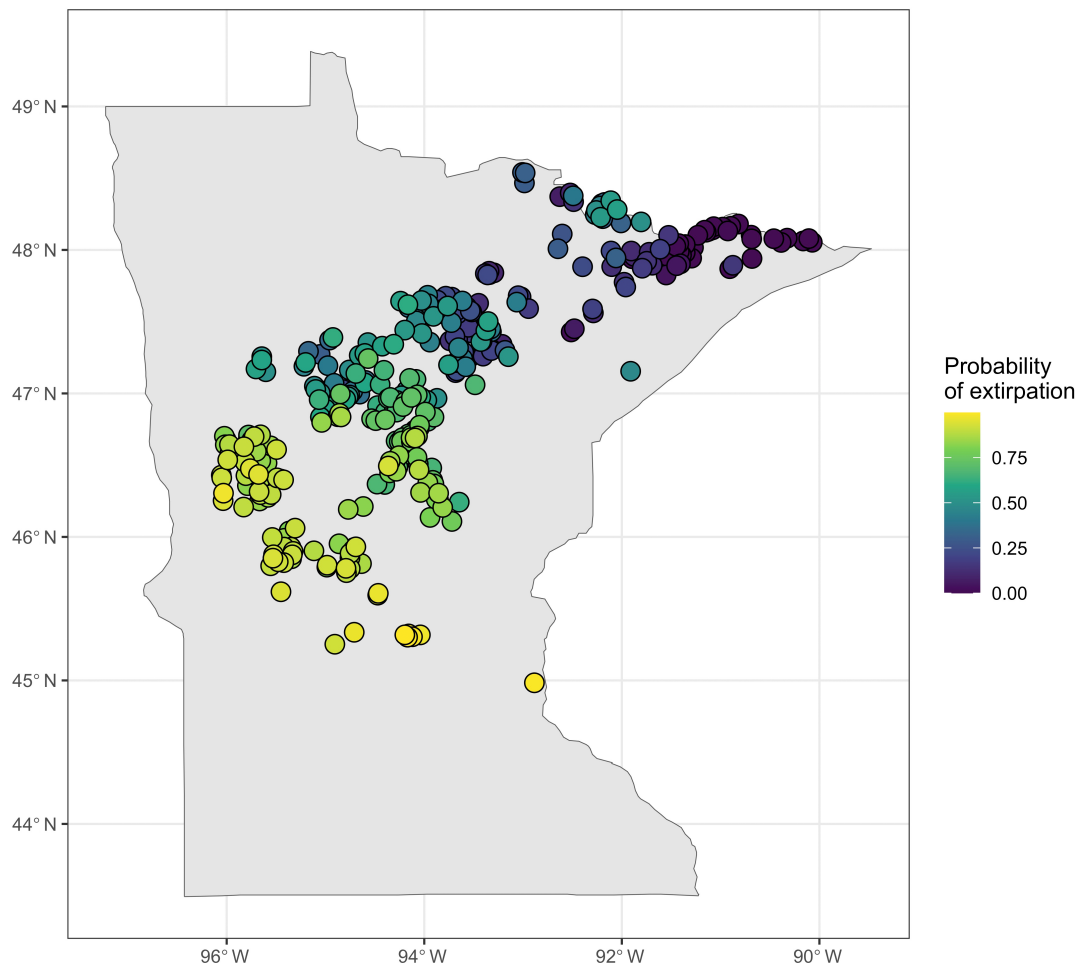
Yellow perch, black crappie, walleye, northern pike, and cisco were, on average, all predicted to decline in abundance by the late 21st century (Figure 2, Appendix S1: Table S3). Each of these species was also predicted to decrease in relative abundance across more than half of the lakes in our study. Conversely, bluegill, largemouth bass, and smallmouth bass were predicted to increase in abundance on average across all lakes. Climate change is likely to produce thermally unsuitable habitats in Minnesota lakes and therefore reduce abundances of fish species with lower thermal optima, whereas warming may result in increased thermal habitat and, thus, higher abundance of species with higher thermal optimums. Cisco and northern pike are likely to experience significant decreases in abundance across much of their current distribution in Minnesota, as the relative abundance of these species was predicted to decline in every lake within our study (sampled and unsampled), with average predicted percentage decreases of 87% and 38%, respectively. Black crappie and walleye were also predicted to decrease in abundance in nearly all lakes (95% and 98%, respectively), though their average percentage decreases (10.92% and 19.84%, respectively) were much smaller than those of cisco and northern pike. Generally, spatial

patterns of changes in predicted relative abundance followed a latitudinal gradient, as expected under climate change, where southern lakes experienced the highest probabilities of extirpation and the largest decreases in relative abundance compared to northern lakes, although considerable variability existed even within small spatial areas (Figure 2). This fine-scale heterogeneity in response to increasing temperatures illustrates how local lake and landscape characteristics may mediate fish responses to climate change. For example, the strong estimated relationships between cisco abundance and lake area, Secchi depth, and elevation suggested that large, low-elevation lakes with high water clarity may lessen the negative responses of cisco abundance to warming lake temperatures.

The estimated probability of extirpation was highest for the cold-adapted species, cisco. Across the 364 study lakes in Minnesota where cisco were observed, the average probability of extirpation at the end of the century was 0.53, and 35 (9.62%) lakes exceeded a probability of extirpation of 0.9 (Figure 3). No other species was predicted to have as high a probability of extirpation (i.e., >0.9) in any lake within its current distribution. The next highest predicted probability of extirpation for a single lake was 0.69 for the northern pike, the species with the second lowest thermal tolerance values in our



**FIGURE 2** Map of percentage change in late-century against current (2021) relative abundance predictions for eight species of lake fish across Minnesota, USA. Changes over time are the result of predicted changes in lake surface water temperature averaged across six global climate models and represent predicted changes in thermal suitability. Negative values indicate a decrease in thermal habitat suitability under future climate conditions. Species are arranged (left to right, top to bottom; cold → warm colors) in order of increasing  $T_{opt}$ .



**FIGURE 3** Map of predicted extirpation probabilities at end of century for cold-adapted fish species, cisco *Coregonus artedii*, across its current distribution in Minnesota, USA. Extirpations are predicted when the lake temperature exceeds the cisco's critical thermal maximum ( $CT_{max}$ ).

study. Conversely, bluegill, largemouth bass, smallmouth bass, and walleye all had zero probability of extirpation across their sampled distribution, and black crappie had near-zero probability, with all lakes having an extirpation probability  $<0.01$  (Appendix S1: Figure S4). In other words, none of the lakes where these species were observed were predicted to reach surface temperatures that exceeded their respective  $CT_{max}$  (accounting for the uncertainty in  $CT_{max}$ ) values by the end of the century.

The variability in species and lake-level predictions highlight how the jsPGA is able to predict changes in abundance and local extirpation events when future ambient temperatures become too warm based on the species' thermal physiology, particularly when these future thermal conditions exceed those currently observed. With predicted increases in surface temperatures as high as  $5.5^{\circ}\text{C}$  in some lakes by the end of the century, the jsPGA predicted negative impacts on abundance for most species, including black crappie, which had predicted decreases in abundance across 95% of the lakes in Minnesota despite

having the second highest mean value for  $CT_{max}$  (Appendix S1: Table S1). These predicted decreases were driven by their relatively low mean  $T_{opt}$  value. In fact, black crappie had the largest difference between these two thermal tolerance values, which resulted in a descending limb of their thermal performance curve that spanned a wide range of temperatures (Equation 3, Appendix S1: Figure S1). Thus, by incorporating the thermal physiology, the jsPGA was able to predict that lake temperatures would warm to suboptimal thermal habitat but not be warm enough to cause local extirpations for the observed black crappie populations in Minnesota.

### Caveats of model specification and model fitting

The goal of our case study was to illustrate the utility of the jsPGA model for making predictions about the effects of climate change on the distributions, abundance, and

extinction of poikilothermic animals and to highlight meaningful ecological inferences that could be made using this method. However, we note some caveats when interpreting the case study results.

First, we summarized the temperature experienced by fishes using surface water temperatures, which represents the warmest part of lakes and does not represent the total available thermal habitat—which will vary based on lake-specific characteristics (e.g., lake morphometry), along with species-specific life histories. This is important because an assumption of the jsPGA, as implemented in our case study, is that the modeled environmental temperature wholly represents the actual lake temperature experienced by organisms and does not account for thermal stratification common in north-temperate lakes. For example, cisco seek thermal refuge in deeper offshore waters to avoid the increased surface water temperatures during summer months (Kao et al., 2020), an important aspect of their life history not fully captured within the jsPGA model. To accommodate this aspect of cisco life history, new statistical methodology would be required to incorporate vertical lake water temperature profile data into the jsPGA modeling framework.

Second, our predictions incorporate predicted shifts in thermal habitat only and do not attempt to incorporate any changes in other environmental factors, such as land use or water clarity. Freshwater systems are under threat from a variety of global change issues, including land-use change, contaminants, and invasive species, in addition to a warming climate (Reid et al., 2019). Future modeling efforts could also aim to include potential changes in measured environmental covariates to develop a more robust understanding of the effects of global change on species abundance and persistence. However, given our results, it is likely that changes in thermal habitat conditions will play a dominant role in the future trajectories of Minnesota lake fishes.

Third, these predictions require the assumption that the estimated species–environmental relationships and species and spatial dependencies will remain constant over time (i.e., the assumption of temporal stationarity; Bueno de Mesquita et al., 2021). For example, we assume that the spatial basis coefficient estimates (which account for both species and spatial dependencies) are robust to changes in climate. This is an important assumption because these dependencies represent unknown mechanisms influencing the abundance and distribution of our focal species, such as biotic interactions or missing environmental covariates. That said, the assumption of stationarity is not unique to our modeling approach but rather represents a broader challenge for predicting the effects of climate change on biodiversity (Dormann et al., 2012).

Finally, it is important to account for the size of the data and available computing resources when fitting the jsPGA model. Numerical integration requires fitting a unique model realization for each randomly sampled pair of thermal response parameters. The average computing time for the models fitted in the simulation study was approximately 2 h—relatively quick when fitting these models in parallel. When applied to the Minnesota case study, computing time greatly increased due to the increase in the size ( $N$ ,  $I$ , and  $K$ ) of our data. Each model took an average of 23 h, making sequential model fitting unreasonable. Furthermore, though the PGA models are less data intensive than mechanistic models (i.e., mechanistic models often need detailed information about how a species' fitness relates to the environment, as well as estimations of phenotypic parameters across a diverse range of climatic conditions; Buckley et al., 2010; Wagner et al., 2023), they still require access to (typically) experimentally derived values of the thermal tolerances. The availability of this information will vary across taxa, and not all thermal tolerances are measured in the same way. If such data are not available, they could be informed based on closely related species or expert opinion.

## CONCLUSION

Climate change will have significant impacts on poikilotherms, and predicting temperature-driven changes in their distributions, abundance, and extinction probabilities is essential for informing climate adaptation policy and conservation and management efforts. Here, we developed a jsPGA model and showed its utility for predicting under future climate change scenarios. The jsPGA, and physiologically guided models in general, help to address issues associated with extrapolation under novel climate conditions and provide more realistic predictions of abundance compared with traditional physiologically naive SDM approaches that do not account for species-specific thermal performance and preferences (Wagner et al., 2023). Importantly, the jsPGA allows for the simultaneous estimation of potentially important species and spatial dependencies, which can improve the prediction performance of the model. Though our analysis focused on lake fishes, the jsPGA can be flexibly fit to model the abundance of poikilotherms across taxa and habitats. For example, efforts to understand the impacts of climate change on insect abundance, such as ecologically important pollinator species, may benefit from using the jsPGA model. Many of these species are already experiencing declines in populations globally (Johnson et al., 2023), and understanding how a warming climate may exacerbate these declines is likely of high importance. The jsPGA model is a new tool that can

be broadly applied to aquatic and terrestrial poikilotherms in an effort to understand and predict the effects of a warming climate on their abundance and distributions.

### AUTHOR CONTRIBUTIONS

Christopher A. Custer contributed to all aspects of the article. Joshua S. North contributed to conceptualization, methodology, software, formal analysis, writing—review and editing, and visualization. Erin M. Schliep contributed to conceptualization, methodology, formal analysis, and writing—review and editing. Michael R. Verhoeven contributed to data curation, software, and writing—review and editing. Gretchen J. A. Hansen contributed to conceptualization, methodology, supervision, project administration, writing—review and editing, and funding acquisition. Tyler Wagner contributed conceptualization, methodology, software, formal analysis, writing—review and editing, resources, supervision, project administration, and funding acquisition.

### ACKNOWLEDGMENTS

The authors thank Holly K. Masui for her assistance with data preparation. This research was supported by the U.S. Geological Survey Midwest Climate Adaptation Science Center Grant No. G20AC00096 and the National Science Foundation (EF-1638679, EF-1638554, EF-1638539, and EF-1638550). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. government. This research was supported by the director, Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 and by the Regional and Global Model Analysis Program area within the Earth and Environmental Systems Modeling Program. This document was prepared as an account of work sponsored by the U.S. government. While this document is believed to contain correct information, neither the U.S. government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. government or any agency thereof or the Regents of the University of California.

### CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

### DATA AVAILABILITY STATEMENT

Data (Kundel et al., 2023) are available from the U.S. Geological Survey ScienceBase data release at <https://doi.org/10.21429/307n-mf74>. Code (Custer et al., 2023) is available from the U.S. Geological Survey software release at <https://doi.org/10.5066/P959EMT5>.

### ORCID

Christopher A. Custer  <https://orcid.org/0000-0003-2072-1414>

Gretchen J. A. Hansen  <https://orcid.org/0000-0003-0241-7048>

Tyler Wagner  <https://orcid.org/0000-0003-1726-016X>

### REFERENCES

- Buckley, L. B., M. C. Urban, M. J. Angilletta, L. G. Crozier, L. J. Rissler, and M. W. Sears. 2010. “Can Mechanism Inform Species’ Distribution Models?” *Ecology Letters* 13: 1041–54.
- Bueno de Mesquita, C. P., C. T. White, E. C. Farrer, L. M. Hallett, and K. N. Suding. 2021. “Taking Climate Change into Account: Non-Stationarity in Climate Drivers of Ecological Response.” *Journal of Ecology* 109: 1491–1500.
- Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and A. Riddell. 2017. “Stan: A Probabilistic Programming Language.” *Journal of Statistical Software* 76: 1–32.
- Charnov, E. L. 2008. “Fish Growth: Bertalanffy  $k$  Is Proportional to Reproductive Effort.” *Environmental Biology of Fishes* 83: 185–87.
- Charnov, E. L., H. Gislason, and J. G. Pope. 2013. “Evolutionary Assembly Rules for Fish Life Histories.” *Fish and Fisheries* 14: 213–224.
- Clark, J. S., A. E. Gelfand, C. W. Woodall, and K. Zhu. 2014. “More than the Sum of the Parts: Forest Climate Response from Joint Species Distribution Models.” *Ecological Applications* 24: 990–99.
- Comte, L., and J. D. Olden. 2017. “Climatic Vulnerability of the World’s Freshwater and Marine Fishes.” *Nature Climate Change* 7: 718–722.
- Corson-Dosch, H. R., W. A. Mcaliley, L. R. Platt, J. A. Padilla, and J. S. Read. 2023. “Daily Water Column Temperature Predictions for Thousands of Midwest U.S. Lakes between 1979–2022 and under Future Climate Scenarios.” U.S. Geological Survey Data Release. <https://doi.org/10.5066/P9EQQER7>.
- Custer, C., J. S. North, E. M. Schliep, and T. Wagner. 2023. “Predicting Fish Responses to Climate Change using a Joint Species, Spatially Dependent Physiologically Guided Abundance Model.” U.S. Geological Survey Software Release. <https://doi.org/10.5066/P959EMT5>.
- Deutsch, C. A., J. J. Tewksbury, R. B. Huey, K. S. Sheldon, C. K. Ghalambor, D. C. Haak, and P. R. Martin. 2008. “Impacts of Climate Warming on Terrestrial Ectotherms across Latitude.” *Proceedings of the National Academy of Sciences of the United States of America* 105: 6668–72.



- DiRenzo, G. V., E. Hanks, and D. A. Miller. 2023. "A Practical Guide to Understanding and Validating Complex Models Using Data Simulations." *Methods in Ecology and Evolution* 14: 203–217.
- Dormann, C. F., S. J. Schymanski, J. Cabral, I. Chuine, C. Graham, F. Hartig, M. Kearney, et al. 2012. "Correlation and Process in Species Distribution Models: Bridging a Dichotomy." *Journal of Biogeography* 39: 2119–31.
- Ehrlén, J., and W. F. Morris. 2015. "Predicting Changes in the Distribution and Abundance of Species under Environmental Change." *Ecology Letters* 18: 303–314.
- Gabry, J., and R. Češnovar. 2022. "cmdstanr: R Interface to 'CmdStan'." <https://mc-stan.org/cmdstanr/>, <https://discourse.mc-stan.org>.
- Gannon, R., M. D. Taylor, I. M. Suthers, C. A. Gray, D. E. van der Meulen, J. A. Smith, and N. L. Payne. 2014. "Thermal Limitation of Performance and Biogeography in a Free-Ranging Ectotherm: Insights from Accelerometry." *Journal of Experimental Biology* 217: 3033–37.
- Gneiting, T., and A. E. Raftery. 2007. "Strictly Proper Scoring Rules, Prediction, and Estimation." *Journal of the American Statistical Association* 102: 359–378.
- Guélat, J., and M. Kéry. 2018. "Effects of Spatial Autocorrelation and Imperfect Detection on Species Distribution Models." *Methods in Ecology and Evolution* 9: 1614–25.
- Guo, F., B. Guénard, E. P. Economo, C. A. Deutsch, and T. C. Bonebrake. 2020. "Activity Niches Outperform Thermal Physiological Limits in Predicting Global Ant Distributions." *Journal of Biogeography* 47: 829–842.
- Hefley, T. J., K. M. Brooms, B. M. Brost, F. E. Buderman, S. L. Kay, H. R. Scharf, J. R. Tipton, P. J. Williams, and M. B. Hooten. 2017. "The Basis Function Approach for Modeling Autocorrelation in Ecological Data." *Ecology* 98: 632–646.
- Hipsey, M. R., L. C. Bruce, C. Boon, B. Busch, C. C. Carey, D. P. Hamilton, P. C. Hanson, et al. 2019. "A General Lake Model (GLM 3.0) for Linking with High-Frequency Sensor Data from the Global Lake Ecological Observatory Network (GLEON)." *Geoscientific Model Development* 12: 473–523. <https://gmd.copernicus.org/articles/12/473/2019/>.
- Homer, C., J. Dewitz, S. Jin, G. Xian, C. Costello, P. Danielson, L. Gass, et al. 2020. "Conterminous United States Land Cover Change Patterns 2001–2016 from the 2016 National Land Cover Database." *ISPRS Journal of Photogrammetry and Remote Sensing* 162: 184–199. <https://doi.org/10.1016/j.isprsjprs.2020.02.019>.
- Hui, F. K., D. I. Warton, S. D. Foster, and C. R. Haak. 2023. "Spatiotemporal Joint Species Distribution Modelling: A Basis Function Approach." *Methods in Ecology and Evolution* 14: 2150–64.
- Johnson, M. G., J. R. Glass, M. E. Dillon, and J. F. Harrison. 2023. "Chapter One – How Will Climatic Warming Affect Insect Pollinators?" In *Environmental Threats to Pollinator Health and Fitness*. Advances in Insect Physiology, Vol. 64, edited by J. F. Harrison, 1–115. Cambridge, MA: Academic Press. <https://doi.org/10.1016/bs.aiip.2023.01.001>.
- Kao, Y.-C., D. B. Bunnell, R. L. Eshenroder, and D. N. Murray. 2020. "Describing Historical Habitat Use of a Native Fish—Cisco (*Coregonus artedii*)—In Lake Michigan between 1930 and 1932." *PLoS One* 15: e0231420.
- Kosten, S., G. Lacerot, E. Jeppesen, D. da Motta Marques, E. H. van Nes, N. Mazzeo, and M. Scheffer. 2009. "Effects of Submerged Vegetation on Water Clarity across Climates." *Ecosystems* 12: 1117–29.
- Kundel, H., M. R. Verhoeven, and G. J. A. Hansen. 2023. "Observations of 12 Fish Species from Minnesota Department of Natural Resources Fish Surveys of Inland Lakes between 1968–2022." U.S. Geological Survey Data Release. <https://doi.org/10.21429/307n-mf74>.
- Latimer, A. M., S. Wu, A. E. Gelfand, and J. A. Silander, Jr. 2006. "Building Statistical Models to Analyze Species Distributions." *Ecological Applications* 16: 33–50.
- Lear, K. O., N. M. Whitney, D. L. Morgan, L. R. Brewster, J. M. Whitty, G. R. Poulakis, R. M. Scharer, T. L. Guttridge, and A. C. Gleiss. 2019. "Thermal Performance Responses in Free-Ranging Elasmobranchs Depend on Habitat Use and Body Size." *Oecologia* 191: 829–842.
- Lester, N., B. Shuter, and P. Abrams. 2004. "Interpreting the von Bertalanffy Model of Somatic Growth in Fishes: The Cost of Reproduction." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271: 1625–31.
- Lyons, J., T. P. Parks, K. L. Minahan, and A. S. Ruesch. 2018. "Evaluation of Oxythermal Metrics and Benchmarks for the Protection of Cisco (*Coregonus artedii*) Habitat Quality and Quantity in Wisconsin Lakes." *Canadian Journal of Fisheries and Aquatic Sciences* 75: 600–608.
- Miller, J., J. Franklin, and R. Aspinall. 2007. "Incorporating Spatial Dependence in Predictive Vegetation Models." *Ecological Modelling* 202: 225–242.
- Miller, J. A. 2012. "Species Distribution Models: Spatial Autocorrelation and Non-Stationarity." *Progress in Physical Geography* 36: 681–692.
- Minnesota Department of Natural Resources (MNDNR). 2017. "Manual of Instructions for Lake Survey." Minnesota Department of Natural Resources, Special Publication No. 180, St. Paul, Minnesota (version 1.04, released January 2019).
- Morgan, R., J. Sundin, M. H. Finnøen, G. Dresler, M. M. Vendrell, A. Dey, K. Sarkar, and F. Jutfelt. 2019. "Are Model Organisms Representative for Climate Change Research? Testing Thermal Tolerance in Wild and Laboratory Zebrafish Populations." *Conservation Physiology* 7: coz036.
- Ovaskainen, O., J. Hottola, and J. Siitonen. 2010. "Modeling Species Co-Occurrence by Multivariate Logistic Regression Generates New Hypotheses on Fungal Interactions." *Ecology* 91: 2514–21.
- Ovaskainen, O., D. B. Roy, R. Fox, and B. J. Anderson. 2016. "Uncovering Hidden Spatial Structure in Species Communities with Spatially Explicit Joint Species Distribution Models." *Methods in Ecology and Evolution* 7: 428–436.
- Paaijmans, K. P., R. L. Heinig, R. A. Seliga, J. I. Blanford, S. Blanford, C. C. Murdock, and M. B. Thomas. 2013. "Temperature Variation Makes Ectotherms More Sensitive to Climate Change." *Global Change Biology* 19: 2373–80.
- Paciorek, C. J. 2013. "Spatial Models for Point and Areal Data Using Markov Random Fields on a Fine Grid." *Electronic Journal of Statistics* 7: 946–972. <https://doi.org/10.1214/13-EJS791>.
- Padfield, D., H. O'Sullivan, and S. Pawar. 2021. "rTPC and nls.multstart: A New Pipeline to Fit Thermal Performance Curves in R." *Methods in Ecology and Evolution* 12: 1138–43.

- Payne, N. L., C. G. Meyer, J. A. Smith, J. D. Houghton, A. Barnett, B. J. Holmes, I. Nakamura, et al. 2018. "Combining Abundance and Performance Data Reveals how Temperature Regulates Coastal Occurrences and Activity of a Roaming Apex Predator." *Global Change Biology* 24: 1884–93.
- Pichler, M., and F. Hartig. 2021. "A New Joint Species Distribution Model for Faster and More Accurate Inference of Species Associations from Big Community Data." *Methods in Ecology and Evolution* 12: 2159–73.
- Planque, B., C. Loots, P. Petitgas, U. Lindstrøm, and S. Vaz. 2011. "Understanding What Controls the Spatial Distribution of Fish Populations Using a Multi-Model Approach." *Fisheries Oceanography* 20: 1–17.
- Poggiato, G., T. Münkemüller, D. Bystrova, J. Arbel, J. S. Clark, and W. Thuiller. 2021. "On the Interpretations of Joint Modeling in Community Ecology." *Trends in Ecology & Evolution* 36: 391–401.
- Pollock, L. J., R. Tingley, W. K. Morris, N. Golding, R. B. O'Hara, K. M. Parris, P. A. Vesik, and M. A. McCarthy. 2014. "Understanding Co-occurrence by Modelling Species Simultaneously with a Joint Species Distribution Model (JSDM)." *Methods in Ecology and Evolution* 5: 397–406.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Record, S., M. C. Fitzpatrick, A. O. Finley, S. Veloz, and A. M. Ellison. 2013. "Should Species Distribution Models Account for Spatial Autocorrelation? A Test of Model Projections across Eight Millennia of Climate Change." *Global Ecology and Biogeography* 22: 760–771.
- Reid, A. J., A. K. Carlson, I. F. Creed, E. J. Eliason, P. A. Gell, P. T. Johnson, K. A. Kidd, et al. 2019. "Emerging Threats and Persistent Conservation Challenges for Freshwater Biodiversity." *Biological Reviews* 94: 849–873.
- Robillard, M., and M. Fox. 2006. "Historical Changes in Abundance and Community Structure of Warmwater Piscivore Communities Associated with Changes in Water Clarity, Nutrients, and Temperature." *Canadian Journal of Fisheries and Aquatic Sciences* 63: 798–809.
- Savino, J. F., and R. A. Stein. 1982. "Predator-Prey Interaction between Largemouth Bass and Bluegills as Influenced by Simulated, Submersed Vegetation." *Transactions of the American Fisheries Society* 111: 255–266.
- Shurin, J. B., K. Cottenie, and H. Hillebrand. 2009. "Spatial Autocorrelation and Dispersal Limitation in Freshwater Organisms." *Oecologia* 159: 151–59.
- Sillero, N., S. Arenas-Castro, U. Enriquez-Urzelai, C. G. Vale, D. Sousa-Guedes, F. Martinez-Freira, R. Real, and A. M. Barbosa. 2021. "Want to Model a Species Niche? A Step-by-Step Guideline on Correlative Ecological Niche Modelling." *Ecological Modelling* 456: 109671.
- Smith, N. J., K. E. Webster, L. K. Rodriguez, K. S. Cheruvilil, and P. A. Soranno. 2022. "LAGOS-US GEO v1.0: Data Module of Lake Geospatial Ecological Context at Multiple Spatial and Temporal Scales in the Conterminous U.S." <https://doi.org/10.6073/pasta/e5c2fb8d77467d3f03de4667ac2173ca>
- Stasko, A. D., T. A. Johnston, and J. M. Gunn. 2015. "Effects of Water Clarity and Other Environmental Factors on Trophic Niches of Two Sympatric Piscivores." *Freshwater Biology* 60: 1459–74.
- Stoklosa, J., R. V. Blakey, and F. K. Hui. 2022. "An Overview of Modern Applications of Negative Binomial Modelling in Ecology and Biodiversity." *Diversity* 14: 320.
- Thorson, J. T., J. N. Ianello, E. A. Larsen, L. Ries, M. D. Scheuerell, C. Szuwalski, and E. F. Zipkin. 2016. "Joint Dynamic Species Distribution Models: A Tool for Community Ordination and Spatio-Temporal Monitoring." *Global Ecology and Biogeography* 25: 1144–58.
- Tikhonov, G., Ø. H. Opedal, N. Abrego, A. Lehtikainen, M. M. de Jonge, J. Oksanen, and O. Ovaskainen. 2020. "Joint Species Distribution Modelling with the R-package Hmsc." *Methods in Ecology and Evolution* 11: 442–47.
- Tomcko, C. M., and R. B. Pierce. 2005. "Bluegill Recruitment, Growth, Population Size Structure, and Associated Factors in Minnesota Lakes." *North American Journal of Fisheries Management* 25: 171–79.
- Trebitz, A., S. Carpenter, P. Cunningham, B. Johnson, R. Lillie, D. Marshall, T. Martin, et al. 1997. "A Model of Bluegill-Largemouth Bass Interactions in Relation to Aquatic Vegetation and Its Management." *Ecological Modelling* 94: 139–156.
- Turko, A. J., C. B. Nolan, S. Balshine, G. R. Scott, and T. E. Pitcher. 2020. "Thermal Tolerance Depends on Season, Age and Body Condition in Imperilled Redside Dace *Clinostomus elongatus*." *Conservation Physiology* 8: coaa062.
- Urban, M. C., G. Bocedi, A. P. Hendry, J.-B. Mihoub, G. Pe'er, A. Singer, J. Bridle, et al. 2016. "Improving the Forecast for Biodiversity under Climate Change." *Science* 353: aad8466.
- Urban, M. C., J. J. Tewksbury, and K. S. Sheldon. 2012. "On a Collision Course: Competition and Dispersal Differences Create no-Analogue Communities and Cause Extinctions during Climate Change." *Proceedings of the Royal Society B: Biological Sciences* 279: 2072–80.
- Van Zuiden, T. M., M. M. Chen, S. Stefanoff, L. Lopez, and S. Sharma. 2016. "Projected Impacts of Climate Change on Three Freshwater Fishes and Potential Novel Competitive Interactions." *Diversity and Distributions* 22: 603–614.
- Wagner, T., G. J. Hansen, E. M. Schliep, B. J. Bethke, A. E. Honsey, P. C. Jacobson, B. C. Kline, and S. L. White. 2020. "Improved Understanding and Prediction of Freshwater Fish Communities through the Use of Joint Species Distribution Models." *Canadian Journal of Fisheries and Aquatic Sciences* 77: 1540–51.
- Wagner, T., E. M. Schliep, J. S. North, H. Kundel, C. A. Custer, J. K. Ruzich, and G. J. Hansen. 2023. "Predicting Climate Change Impacts on Poikilotherms Using Physiologically Guided Species Abundance Models." *Proceedings of the National Academy of Sciences of the United States of America* 120: e2214199120.
- Warton, D. I., F. G. Blanchet, R. B. O'Hara, O. Ovaskainen, S. Taskinen, S. C. Walker, and F. K. Hui. 2015. "So Many Variables: Joint Modeling in Community Ecology." *Trends in Ecology & Evolution* 30: 766–779.
- Wilkinson, D. P., N. Golding, G. Guillera-Aroita, R. Tingley, and M. A. McCarthy. 2019. "A Comparison of Joint Species Distribution Models for Presence–Absence Data." *Methods in Ecology and Evolution* 10: 198–211.
- Wismer, D. A., and A. E. Christie. 1987. *Temperature Relationships of Great Lakes Fishes: A Data Compilation*. Number 87. Ann Arbor, MI: Great Lakes Fishery Commission.



- Zanden, M. J. V., J. D. Olden, J. H. Thorne, and N. E. Mandrak. 2004. "Predicting Occurrences and Impacts of Smallmouth Bass Introductions in North Temperate Lakes." *Ecological Applications* 14: 132–148.
- Zurell, D., W. Thuiller, J. Pagel, J. S. Cabral, T. Münkemüller, D. Gravel, S. Dullinger, et al. 2016. "Benchmarking Novel Approaches for Modelling Species Range Dynamics." *Global Change Biology* 22: 2651–64.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Custer, Christopher A., Joshua S. North, Erin M. Schliep, Michael R. Verhoeven, Gretchen J. A. Hansen, and Tyler Wagner. 2024. "Predicting Responses to Climate Change Using a Joint Species, Spatially Dependent Physiologically Guided Abundance Model." *Ecology* e4362. <https://doi.org/10.1002/ecy.4362>