

Joint species distribution models reveal taxon-specific sensitivities to potential anthropogenic alteration

Darin A. Kopp^{1,5}, John L. Stoddard^{2,6}, Ryan A. Hill^{2,7}, Jessie M. Doyle^{1,3,8}, Philip R. Kaufmann^{2,4,9}, Alan T. Herlihy^{4,10}, and Steven G. Paulsen^{2,11}

¹Oak Ridge Institute for Science and Education Participant c/o United States Environmental Protection Agency, Office of Research and Development, 200 Southwest 35th Street, Corvallis, Oregon 97333 USA

²United States Environmental Protection Agency, Office of Research and Development, 200 Southwest 35th Street, Corvallis, Oregon 97333 USA

³Oregon Department of Fish and Wildlife, 4034 Fairview Industrial Drive Southeast, Salem, Oregon 97302 USA

⁴Department of Fisheries, Wildlife, and Conservation Sciences, Oregon State University, Nash Hall, 2820 Southwest Campus Way, Corvallis, Oregon 97331 USA

Abstract: Taxon–environment relationships can elucidate a taxon’s tolerance or sensitivity to specific environmental conditions. We use a joint species distribution modeling framework to quantify relationships between ~1700 benthic macroinvertebrate assemblages in streams and rivers across the contiguous United States and several environmental gradients that are susceptible to human alteration (e.g., nutrients, salinity, physical habitat, and climate). We found that the predicted occurrence probability for sampling units where a taxon actually occurs was 0.15 to 0.24 greater than the predicted occurrence probability for sampling units where a taxon does not occur, and a relatively large percentage (32–58%) responded to gradients of substrate diameter, mean summer air temperature, or total P. At the assemblage level, genus richness could change along environmental gradients by as many as 5 to 17 taxa depending on the ecoregion. Often, the largest change in genus richness was associated with sediment diameter. We also investigated whether a suite of traits (i.e., clinger, scraper, pollution tolerance, and thermal optima) were related to a genus’ association with an environmental gradient and found that some traits are positively related to an organism’s occurrence along one environmental gradient but negatively related to its occurrence along another. For example, in several ecoregions, thermal preference was positively related to mean summer air temperature but negatively related to nutrient concentrations. Collectively, our results showcase a multivariate approach for modeling biotic assemblages that can integrate multiple sources of information (i.e., environmental factors, biological traits, phylogenetic relationships, and co-occurrences) that are routinely collected by biomonitoring programs.

Key words: National Rivers and Streams Assessment Program, benthic macroinvertebrates, joint species distribution modeling, hierarchical modeling of species communities, taxon–environment relationships, tolerance, sensitivity

Anthropogenic activities modify environmental gradients and create habitats with specific, and sometimes novel, conditions (Poff et al. 2007, Heino et al. 2013, Groffman et al. 2014). Identifying which organisms could disappear, persist, or appear under anthropogenic stress is a central objective of bioassessment (Clarke et al. 2003, Yuan 2006, Hawkins et al. 2010, 2015). Quantifying relationships between the occurrence of an organism and environmental gradients can reveal how the organism may respond if environmental conditions change (Elith and Leathwick 2009, Ovaskainen et al. 2017, Araújo et al. 2019). When anthropogenic activities

change abiotic conditions, taxon–environment relationships could guide predictions about which organisms could be sensitive or tolerant to those changing conditions (Yuan 2004, Domisch et al. 2011).

In rivers and streams, anthropogenic activities can alter chemical and physical gradients and drive changes in faunal occurrence (Bonada et al. 2006, Dolédec and Statzner 2010, Hawkins and Yuan 2016). Changes in benthic macroinvertebrate assemblage metrics are often used to quantify the magnitude of anthropogenic alterations (Karr 1981, Resh 2008, Buss et al. 2015), and taxon–environment relationships can

E-mail addresses: ⁵kopp.darin@epa.gov; ⁶stoddard.john@epa.gov; ⁷hill.ryan@epa.gov; ⁸jessie.m.doyle@odfw.oregon.gov; ⁹kaufmann.phil@epa.gov; ¹⁰alan.herlihy@oregonstate.edu; ¹¹Paulsen.Steve@epa.gov

Received 5 August 2022; Accepted 24 April 2023; Published online 19 July 2023. Associate Editor, Matthew Baker.

Freshwater Science, volume 42, number 3, September 2023. © 2023 The Society for Freshwater Science. All rights reserved. Published by The University of Chicago Press for the Society for Freshwater Science. <https://doi.org/10.1086/726283>

reveal which taxa could increase or decrease with anthropogenic disturbance (Yuan 2004, 2006). For example, if anthropogenic activities elevate nutrient concentrations, then taxa with a monotonically increasing relationship with nutrients would be expected to be tolerant to these elevated conditions, and taxa with a monotonically decreasing relationship would be expected to be sensitive. Taxa can also have unimodal taxon–environment relationships, but whether these taxa are sensitive, intermediately tolerant, or tolerant depends on the location of the optimum (i.e., maximum probability of occurrence along a gradient) relative to the value associated with anthropogenic alteration (Yuan 2004, 2006).

From a conservation standpoint, quantifying taxon–environment relationships is extremely useful (Strayer 2006, Collier et al. 2016), but the magnitude and direction of relationships can vary over geographic space (Osborne and Suárez-Seoane 2002, Osborne et al. 2007, Chang et al. 2014, Mao et al. 2019). Geographic variation can occur if a taxon's response to 1 environmental gradient depends on resource availability, behavior, or another environmental gradient that changes across the taxon's range (Bolnick et al. 2011, Heino et al. 2013, Zettler et al. 2013, Burner et al. 2021). For example, members of Coleoptera vary in their salinity tolerance because variation in water permanence causes species inhabiting temporary habitats to experience a wider range of salinities (Céspedes et al. 2013, Kefford et al. 2016). Alternatively, when regionalization is used to define discrete reporting units (Herlihy et al. 2008), taxon–environment relationships may vary among units because regional boundaries only contain a portion of the environmental conditions that limit an organism's distribution (Feld et al. 2016, Burner et al. 2021, Segurado et al. 2022). Because taxon–environment relationships can vary geographically, region-specific estimates may more accurately describe how organisms might respond to anthropogenic alteration and may facilitate the development of localized biological indices (Chang et al. 2014, Mao et al. 2019).

Joint species distribution models (JSDMs) are a multivariate extension of single species distribution models that evaluate taxon–environment relationships for all taxa simultaneously (Warton et al. 2015, Norberg et al. 2019, Ovaskainen and Abrego 2020). Typically, they are generalized linear mixed models that explicitly incorporate co-occurrence information (Warton et al. 2015, Ovaskainen and Abrego 2020). As such, these models evaluate taxon–environment relationships as fixed effects and use latent variables and co-occurrence data to model random effects (Warton et al. 2015, Ovaskainen et al. 2016a, Norberg et al. 2019, Ovaskainen and Abrego 2020). Including random effects accounts for nonindependence in residuals caused by missing covariates and can improve estimates for the fixed effects (Bolker et al. 2009, McElreath 2016). This attribute of JSDMs makes them uniquely suited for estimating region-specific taxon–environment relationships and revealing re-

gional variation because a consistent modeling structure can be used to estimate fixed effects in multiple regions.

In addition, JSDMs can integrate traits and phylogeny as hierarchical terms to evaluate their influence on taxon–environment relationships (Abrego et al. 2017). Traits can provide a mechanism for a taxon's persistence under certain environmental conditions, and trait-based approaches are increasingly used in ecological assessments (Statzner and Bêche 2010, Heino et al. 2013). However, for a trait to be a reliable indicator of an alteration, it should be consistently related to an organism's response to an environmental gradient (Statzner and Bêche 2010). In actuality, organisms are exposed to multiple environmental gradients, and a single trait that is associated with an organism's occurrence along one environmental gradient could be neutral to its occurrence along another (Finn and Poff 2005). Thus, variation in environmental conditions among biogeographic regions could cause organisms with the same trait to respond differently to their environment (Heino et al. 2013, Saito et al. 2020). Because JSDMs provide a comprehensive framework for quantifying taxon–environment relationships and assessing whether these relationships are influenced by the presence of certain traits, they could elucidate whether traits are reliable indicators of an organism's response to an environmental gradient.

Here, we used JSDMs to quantify taxon–environment relationships for benthic macroinvertebrate genera along environmental gradients commonly altered by anthropogenic activities and to determine whether these relationships are related to traits in 9 geographic regions of the contiguous United States (Fig. 1). Our primary objective was to quantify these relationships and elucidate the complexities of modeling entire taxonomic assemblages across multiple regions. We hypothesized that taxon–environment relationships vary regionally because regional boundaries do not contain the entire range of conditions that limit a taxon's distribution or because a taxon responds to other gradients that vary across large geographic extents (Herlihy et al. 2008, Yuan et al. 2008, Heino et al. 2013). If taxon–environment relationships vary regionally, tolerance and sensitivity values derived from one region may not be applicable in another (Chang et al. 2014). We also hypothesized that traits are related to an organism's association with an environmental gradient, and that these relationships are consistent among ecoregions because traits are mechanistically linked to an organism's behavior, tolerance, or preference for certain environmental conditions (Townsend and Hildrew 1994, Saito et al. 2020).

METHODS

Benthic macroinvertebrate assemblages

Since 2008, the United States Environmental Protection Agency's (USEPA) National Rivers and Streams Assessment

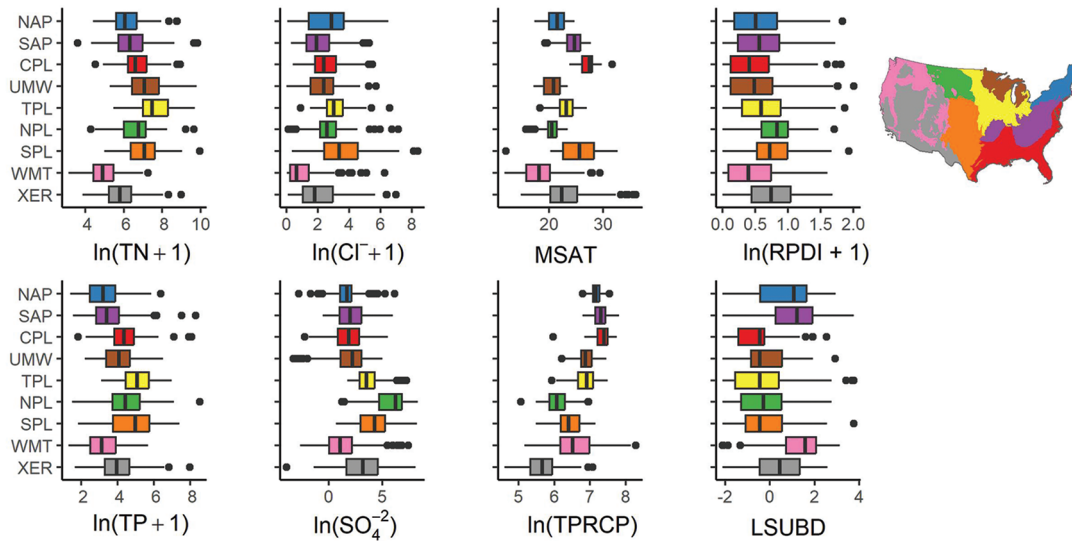


Figure 1. Boxplots of environmental gradients from the National Rivers and Streams Assessment (data from surveys done in 2018 and 2019) within 9 ecoregions of the contiguous United States. CPL = Coastal Plains, LSUBD = log mean substrate diameter, MSAT = mean summer air temperature, NAP = Northern Appalachians, NPL = Northern Plains, RPDI = riparian disturbance index, SAP = Southern Appalachians, SPL = Southern Plains, TN = total N, TP = total P, TPL = Temperate Plains, TPRCP = total annual precipitation, UMW = Upper Midwest, WMT = Western Mountains, XER = Xeric.

(NRSA) surveys rivers and streams throughout the contiguous United States every 5 y to assess their ecological condition (USEPA 2016). For each NRSA survey, ~2000 locations are chosen as a spatially balanced probabilistic random sample throughout 9 ecoregions (Olsen and Peck 2008). Procedures for collecting macroinvertebrates are described in detail by Hughes and Peck (2008) and USEPA (2009, 2017). Briefly, each survey location was between 20 and 40 channel widths or a minimum of 150 m in length, and a D-frame kick net (500- μm mesh, 0.09- m^2 area) was used to collect macroinvertebrate samples along 11 transects equally spaced throughout the location. Wadeable sites were sampled in an alternating left, center, right order, and boatable sites were sampled by alternating between left and right bank locations within the wadeable margin. Samples from all 11 transects were combined into a single composite sample, preserved in ethanol, and then sent to a qualified taxonomist for identification (USEPA 2017).

In the laboratory, a fixed count of 500 ind. was identified to the lowest possible taxonomic level (USEPA 2017). Because it was not always possible to achieve 500 ind. from a site, all samples were rarified to 300 ind. for consistency. Our analysis focused on data collected during the 2018 and 2019 field campaign cycle and included all taxa that were identified to genus with a prevalence $\geq 10\%$ of sites within an ecoregion (Table 1, Fig. S1). Although a benefit of JSDBMs is their ability to model rare taxa, fitting nonlinear models with Bayesian inference can be challenging with large datasets, and we chose the prevalence threshold to ensure satisfactory convergence (Ovaskainen and Abrego 2020). We also ex-

cluded data collected during repeated visits. For each region we modeled 63 to 127 genera collected from 140 to 255 sites (Table 1).

Environmental gradients

NRSA also collects a suite of environmental variables at each location (USEPA 2017). From these data, we identified chemical, physical habitat, and climate variables that reflect environmental gradients commonly altered by human activities (Herlihy et al. 2008; Tables 2, S1). From the field data collected by NRSA, we selected total P (TP) and total N (TN) as indicators of nutrient gradients, Cl^- and SO_4^{-2} concentrations as indicators of salinity gradients, and log mean substrate diameter (LSUBD) and a riparian disturbance index (RPDI) as indicators of physical habitat gradients (USEPA 2020). TP and TN were measured by persulfate digestion and colorimetry, and Cl^- and SO_4^{-2} concentrations were measured in water samples by ion chromatography (USEPA 2008). LSUBD is calculated as the log of the frequency-weighted, geometric mean particle size, and RPDI is a composite index that reflects the intensity and proximity of anthropogenic activities (e.g., agriculture, roads, or buildings) in the riparian area adjacent to a stream (Kaufmann et al. 1999, 2009). For climate gradients, we extracted mean summer air temperature (MSAT; July and August) and total annual precipitation (TPRCP) for either 2018 or 2019 (depending on the year a site was surveyed) from the Parameter-elevation Regression on Independent Slopes Model Climate Group's Recent Years dataset (<https://prism.oregonstate.edu/>). Prior

Table 1. Number of sites and characteristics of genera included in the analysis. Sites were selected to have complete data for the environmental gradients and genera were selected to have a prevalence >0.1. CH = Chironomidae, CO = Coleoptera, CPL = Coastal Plains, D = Diptera, E = Ephemeroptera, M = Megaloptera, NAP = Northern Appalachians, NPL = Northern Plains, O = Odonata, P = Plecoptera, SAP = Southern Appalachians, SPL = Southern Plains, T = Trichoptera, TPL = Temperate Plains, UMW = Upper Midwest, WMT = Western Mountains, XER = Xeric.

Ecoregion	Sites (No.)	Genera (No.)	Not insects	Proportion of genera by taxonomic group								Mean genus prevalence (range)	Mean genus richness (range)
				Insects									
				CO	E	M	O	P	T	D			
						CH	Not CH						
NAP	227	127	0.14	0.08	0.13	0.02	0.02	0.06	0.17	0.34	0.06	0.25 (0.10–0.94)	31.69 (9–58)
SAP	255	114	0.19	0.10	0.18	0.03	0.01	0.03	0.11	0.31	0.06	0.26 (0.10–0.94)	29.95 (7–57)
CPL	187	80	0.19	0.07	0.10	0.01	0.04	0	0.07	0.48	0.04	0.25 (0.10–0.97)	19.96 (7–39)
UMW	199	107	0.18	0.04	0.16	0.01	0.02	0.02	0.12	0.39	0.07	0.27 (0.10–0.90)	29.23 (7–54)
TPL	202	78	0.21	0.06	0.17	0	0.01	0	0.08	0.45	0.03	0.28 (0.10–0.94)	21.52 (6–41)
NPL	140	79	0.15	0.08	0.19	0.01	0.01	0	0.09	0.43	0.04	0.26 (0.10–0.79)	20.66 (7–42)
SPL	141	63	0.21	0.08	0.13	0	0.02	0	0.10	0.44	0.03	0.29 (0.11–0.84)	18.01 (7–34)
WMT	221	94	0.13	0.05	0.15	0	0	0.09	0.15	0.38	0.05	0.28 (0.10–0.83)	26.59 (5–43)
XER	171	79	0.19	0.05	0.15	0	0.01	0.03	0.11	0.42	0.04	0.26 (0.11–0.70)	20.68 (8–40)

to modeling, we mean-centered and standardized the values for the environmental gradients within each region.

Benthic macroinvertebrate traits

Traits are measurable characteristics of an organism that should provide a mechanistic explanation for their relationship to an environmental gradient (Townsend and Hildrew 1994, Stutzner and Bêche 2010, Hamilton et al. 2020, Burner et al. 2021). We identified traits that we hypothesized to be related to a genus' occurrence with respect to ≥ 1 environmental gradient (scrapers, clingers, pollution tolerance value [PTV], and thermal optima [ThrmOptV]; Table 3). Scrapers, a functional feeding group, possess mouthparts to consume biofilms and macrophytes, and clingers have morphological features that allow them to attach to rocky substrates. Both trait states were treated as binary variables

(i.e., presence or absence of the trait). A taxon's PTV quantifies their ability to tolerate organic pollution, and their ThrmOptV quantifies their temperature preference. Both are continuous variables. By focusing on traits with simple a priori expectations to ≥ 1 niche component (Table 3), we evaluated consistency in these relationships among ecoregions and multiple environmental gradients. We obtained trait data from the autecological characteristics assembled by NRSA (USEPA 2020; <https://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>) and the Freshwater Biological Traits database (USEPA 2012).

JSDMs

We used the Hierarchical Modeling of Species Communities (HMSC) R package (*Hmsc*; version 3.0-13; Tikhonov et al. 2020) to fit separate JSDMs for each region. HMSC is a

Table 2. Environmental gradients and metrics used to quantify them. Where applicable we used data transformations to improve normality. Because the geometric mean of substrate diameter is already on the logarithmic scale, no transformation was needed.

Gradient	Metric	Units	Transformation
Nutrient	Total N	mg/L	$\ln(x + 1)$
	Total P	ug/L	$\ln(x + 1)$
Salinity	Cl ⁻	mg/L	$\ln(x + 1)$
	SO ₄ ⁻²	mg/L	$\ln(x)$
Physical	Log10 geometric mean substrate diameter	mm	–
	Riparian disturbance index	–	$\ln(x + 1)$
Climate	Mean summer air temperature	°C	–
	Total annual precipitation	mm	$\ln(x)$

Table 3. Traits used in analysis with a priori expectations for their relationships to environmental gradients (i.e., trait–niche relationships). Observed relationships are presented in Fig. 5. EPA = United States Environmental Protection Agency, NRSA = National Rivers and Streams Assessment.

Trait state	Description	Function	Hypothesis
Clinger	Morphological structures to attach to rocky surfaces (USEPA 2020)	Genera cling to hard substrates	Positively related to substrate diameter and total annual precipitation because this behavior may be beneficial in environments with larger substrates or prone to scouring events
Scraper	Mouthparts for consuming algae/biofilms (USEPA 2020)	Genera consume primary production	Positively related to nutrient gradients and riparian disturbance index because of increased primary production
Pollution tolerance value	A derived index related to organic and nutrient pollution tolerance (Source: NRSA autecology data)	Genera tolerate high-nutrient concentrations	Positively related to nutrient concentrations and riparian disturbance index because conditions are favorable for these taxa
Thermal optima	A weighted average indicating temperature preference (USEPA 2012)	Genera prefer an optimal temperature	Positively related to mean summer air temperatures because higher water temperatures favor higher thermal optima

hierarchical generalized linear mixed model that is described extensively by Ovaskainen and Abrego (2020). Mixed models consist of fixed and random effects. The fixed effects portion of the HMSC model estimates the relationship between a taxon's occurrence and a suite of environmental covariates as regression parameters (i.e., slopes and intercepts). The regression parameters quantify how the probability of occurrence for each taxon changes along the environmental gradients; therefore, the regression parameters are referred to as niche parameters (β) and interpreted as taxon–environment relationships (Ovaskainen and Abrego 2020). Random effects are used to account for nonindependence among residuals from the fixed effects portion of the model. In HMSC, site-level random effects are modeled from pairwise co-occurrences and latent variables because multiple taxa are collected from each site (i.e., residuals for taxa collected from the same site may not be independent; Ovaskainen et al. 2016b). The ability to use site-level random effects is a unique attribute of JSDMs that can advance modeling across large spatial extents because it improves model performance and accounts for the effects of unmeasured variables that contribute to a taxon's occurrence (Warton et al. 2015, Ovaskainen and Abrego 2020). By statistically accounting for unmeasured factors at each site, we were able to focus our analysis and interpretation on taxon–environment relationships for gradients commonly altered by anthropogenic activities.

An organism's response to environmental gradients could be influenced by the traits they express. HMSC uses a hierarchical (or multilevel) model structure to assess whether traits influence taxon–environment relationships. Conceptually, the model structure is analogous to using regression analysis with a taxon–environment relationship (i.e., estimated regression coefficients from the fixed effects portion of the HMSC model) as the dependent variable and a suite of traits as independent variables. The estimated coefficients

for each trait (γ) are then interpreted as the association between a trait and a taxon–environment relationship. We refer to these associations as trait–niche relationships. Because closely related taxa may have similar relationships to their environment, HMSC uses phylogenetic relatedness (i.e., pairwise distances) to model nonindependence among residuals from the trait–niche relationships as random effects (ρ). If the residuals from the trait–niche relationship have a phylogenetic structure, it suggests that closely related taxa respond similarly to their environment because of shared traits that were not included in the analysis (Ovaskainen and Abrego 2020). A directed acyclic graph of the HMSC model is provided as Fig. 4 in Ovaskainen et al. (2017).

Model fitting and interpretation

For each region, we evaluated taxon–environment relationships for environmental gradients commonly altered by anthropogenic activities (i.e., TN, TP, Cl^- , SO_4^{2-} , MSAT, TPRCP, LSUBD, and RPD) as additive fixed effects. Each model included genus presence–absence as the response variable, environmental gradients as fixed effects, and sample-level random effects. We also included traits (i.e., scraper, clinger, PTV, and ThrmOptV) as additive fixed effects and taxonomic relatedness as random effects for the hierarchical structure of these models. We converted taxonomic levels to a phylogenetic tree with the *ape* package (version 5.6-2; Paradis and Schliep 2019) in R and assumed equal branch lengths between nodes as quantitative measures of phylogenetic distances separating all benthic macroinvertebrate genera.

We used Tjur R^2 to assess explanatory, predictive, and conditional predictive power of the models. Tjur R^2 is the mean predicted occurrence probability for sampling units where a taxon actually occurs minus the mean predicted occurrence probability for sampling units where a taxon does

not occur (Tjur 2009). Explanatory power assesses model predictions against the data used for fitting, predictive power assesses model predictions at new locations (i.e., using only fixed effects), and conditional predictive power assesses model predictions at new locations conditional on the presence of other taxa (Ovaskainen and Abrego 2020). To evaluate predictive power, we used 2-fold cross validation among sites, and to evaluate conditional predictive power, we used 2-fold cross validation among sites and 3-fold cross validation among genera.

We interpreted β as positive when 90% credible interval (CrI) > 0 and negative when 90% CrI < 0. For each region, we counted the number of positive and negative relationships identified among the entire regional taxa pool and several taxonomic groups (i.e., noninsects; Ephemeroptera, Plecoptera, Trichoptera [EPT]; and Diptera:Chironomidae) and report these counts as proportions. Proportional data facilitated comparisons across regions while accounting for differences in the number of taxa modeled. We interpreted γ as positive when 90% CrI > 0 and negative when 90% CrI < 0. For each ecoregion, we report the direction (either positive or negative) of the trait–niche relationships to visualize geographic differences. When ρ is positive (90% CrI > 0), it indicates that related taxa tend to have similar taxon–environment relationships because of shared, but unmeasured, traits (Ovaskainen and Abrego 2020).

We used predictions from the HMSC models to evaluate how genus, EPT, Chironomidae, and noninsect richness changed along the gradients. For each region, we iteratively selected a single gradient (i.e., focal gradient) and predicted the probability of occurrence for each taxon while holding the other gradients at a mean value, conditional on the focal gradient's value (Ovaskainen and Abrego 2020). We then visualized how the richness of each taxonomic group changed along the focal gradient by summing the predicted probabilities for each taxon at 25 equally spaced intervals (Moss et al. 1987). To quantitatively evaluate the trend in mean genus richness, we calculated the probability that the posterior predictive distribution of genus richness at the minimum value of a gradient was different from the posterior predictive distribution of genus richness at the maximum value. We determined that the 2 were sufficiently different when the probability was ≥ 0.95 .

Unimodal taxon–environment relationships are fundamental to niche theory (Yuan 2006) but are not accurately modeled using linear terms. Because including quadratic terms for each environmental gradient in a single HMSC model could potentially lead to overfitting, we also fitted a separate suite of models for each region to identify whether taxon–environment relationships were unimodal (Yuan 2004, Austin 2007). These unimodal models were intentionally simplistic and fitted using presence and absence data as the response variable and individual gradients with a quadratic term as the predictor variables. We also used sample-level random effects but excluded traits and phylogeny to avoid excessive computation time (Ovaskainen and Abrego 2020).

We identified unimodal relationships when the quadratic term for the environmental gradient was negative and the maximum of the parabola (P_{max}) was within the range of values used for fitting (Ovaskainen and Abrego 2020, Tikhonov et al. 2020). We calculated the maximum of the parabola as:

$$P_{max} = -(b/2a), \quad (\text{Eq. 1})$$

where b is the estimated linear coefficient, and a is the estimated quadratic coefficient. When a unimodal relationship was identified, we interpreted the peak as the taxon's optima for that gradient and report the mean and 90% CrI. We excluded taxa with unimodal relationships from our analysis of taxon–environment relationships estimated from additive fixed effects and richness because they were uncommon.

We fitted all models using a Dell Precision Tower 7910 (Dell Technologies, Round Rock, Texas) with an Intel® Xeon® processor (12 cores and 256 gigabytes of memory; Intel Corporation, Santa Clara, California), the default prior distributions, and 3 independent Markov chain Monte Carlo simulations (Ovaskainen and Abrego 2020). We sampled each chain 1000 \times and adjusted transient iterations and thinning to ensure convergence (Table S2). Transient iterations are discarded to reduce the influence of the starting values on Markov chain Monte Carlo simulations, and thinning reduces autocorrelation between sequential samples by specifying a number of iterations between each recorded sample (Gelman et al. 2013). Satisfactory convergence ensures a valid approximation of the posterior distribution and is evaluated by the potential scale reduction factor and effective sample size (Gelman et al. 2013, Tikhonov et al. 2020). We considered convergence satisfactory when the potential scale reduction factor was <1.1 and effective sample size was >1500 (Tikhonov et al. 2020; Table S2).

RESULTS

We modeled taxon–environment relationships for 1743 assemblages across 9 ecoregions. Most genera belonged to Chironomidae (Diptera, 31–48%), noninsects (13–21%), Ephemeroptera (10–19%), and Trichoptera (7–17%) (Table 1). The median explanatory Tjur R^2 across all taxa varied from 0.15 in the Coastal Plains (CPL) to 0.24 in the Western Mountains (WMT). The median conditional Tjur R^2 (range = 0.05–0.14) was higher than the median predictive Tjur R^2 (range = 0.03–0.08) in all regions (Fig. 2), but there was less difference in the Plains ecoregions (CPL, Upper Midwest [UMW], Northern Plains [NPL], Temperate Plains [TPL]) compared with others (Appendix S1).

As many as 32 to 58% of the genera in each region had a relationship with an environmental gradient (Fig. 3). The proportion of taxa that were associated with each

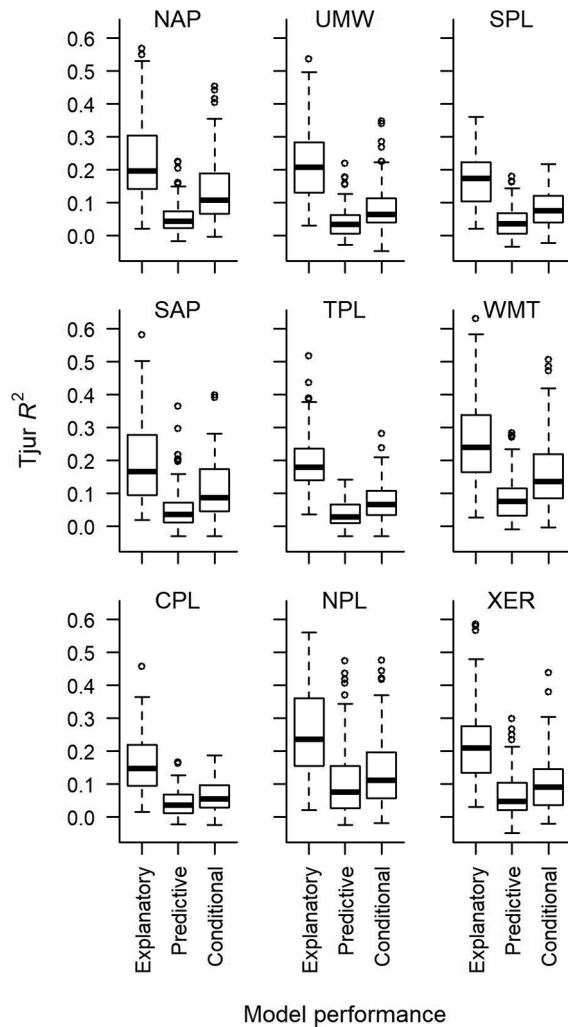


Figure 2. Boxplots showing performance metrics for Hierarchical Modeling of Species Communities models of aquatic macroinvertebrate assemblages in 9 ecoregions across the contiguous United States. Models included environmental gradients as additive fixed effects and sample level random effects. Explanatory power assesses model predictions against the data used for fitting, predictive power assesses model predictions at new locations, and conditional predictive power assesses model predictions at new locations conditional on the presence of other taxa. $T_{\text{jur}} R^2$ is the mean predicted occurrence probability where a genus occurs minus mean predicted occurrence probability where a genus does not occur. The box represents the 25th, 50th, and 75th percentiles of $T_{\text{jur}} R^2$ values for all genera in an ecoregion, and points are values that exceed $1.5 \times$ interquartile range. CPL = Coastal Plains, NAP = Northern Appalachians, NPL = Northern Plains, SAP = Southern Appalachians, SPL = Southern Plains, TPL = Temperate Plains, UMW = Upper Midwest, WMT = Western Mountains, XER = Xeric.

environmental gradient reveals the number of taxa that could potentially be affected if the gradient was altered within each ecoregion. Among the gradients, LSUBD had

a relationship with the highest percentage of genera in NPL (58%), WMT (51%), Northern Appalachians (NAP; 50%), CPL (36%), and UMW (29%), whereas MSAT had relationships with more genera in Southern Appalachians (SAP; 49%), Xeric (XER; 46%), and Southern Plains (SPL; 40%). In TPL, 35% of the genera were associated with TP. The direction of the taxon–environment relationships elucidate how taxa could change if a gradient is altered (Fig. 3). Negative taxon–environment relationships were most common with MSAT in XER (36%), SAP (35%), SPL (21%), CPL (20%), and UMW (18%) and with Cl^- in NPL (36%) and WMT (32%). In NAP, 23% of taxa were negatively associated with RPDI, and in TPL, 29% of taxa were negatively related to TP. Positive taxon–environment relationships were most common with LSUBD in NAP (45%), SAP (36%), WMT (35%), CPL (30%), XER (30%), and UMW (0.23%) and with SO_4^{-2} in NPL (35%) and SPL (22%). In TPL, 15% of the taxa had positive relationships with Cl^- .

Taxonomic groups varied in their association with the environmental gradients (Fig. 3). Most Chironomidae were associated with MSAT in SAP, UMW, WMT, and XER and with SO_4^{-2} in NPL and SPL. Alternatively, in CPL, NAP, and TPL, Chironomidae were typically associated with LSUBD, RPDI, and TP, respectively. Most EPT were associated with LSUBD in NAP, NPL, SAP, UMW, and XER and with MSAT in CPL, SPL, and WMT. In TPL, most EPT were associated with TN. Most noninsect taxa were associated with SO_4^{-2} in SAP, TPL, and UMW; with LSUBD in NPL and WMT; and with MSAT in SPL and XER. In CPL and NAP, most noninsect associations were with TP and Cl^- , respectively. Often, genera within a taxonomic group had different associations with the gradients (See Appendix S2 for mean, standard deviation, and 90% CrIs for the taxon–environment relationships).

The genus richness at the minimum and maximum value of a gradient was predicted to change by as many as 5 to 17 taxa (Fig. 4). LSUBD was associated with the greatest magnitude of change in genus richness in CPL (9 taxa), NAP (17 taxa), SAP (17 taxa), SPL (5 taxa), and UMW (9 taxa). In NPL, TPL, WMT, and XER, genus richness changed the most with RPDI (8 taxa), TP (7 taxa), SO_4^{-2} (15 taxa), and MSAT (10 taxa), respectively. Genus richness was not consistently related to the environmental gradients across all ecoregions (Fig. 4). Visually inspecting changes in each taxonomic subgroup revealed that EPT richness tended to change more steeply than Chironomidae and noninsect richness for most gradients.

In several ecoregions, trait–niche relationships were consistent with the hypotheses described in Table 3 (Fig. 5). PTV was positively related to β_{TN} in NAP, UMW, and NPL and to β_{RPDI} in UMW, SPL, and XER. Clingers were positively related to β_{LSUBD} in 5 ecoregions (NAP, UMW, N, WMT, and XER) and to β_{TPRCP} in 1 ecoregion (SAP). ThrmOptV was positively related to β_{MSAT} in all regions except UMW and NAP. We did not find a relationship between scrapers and

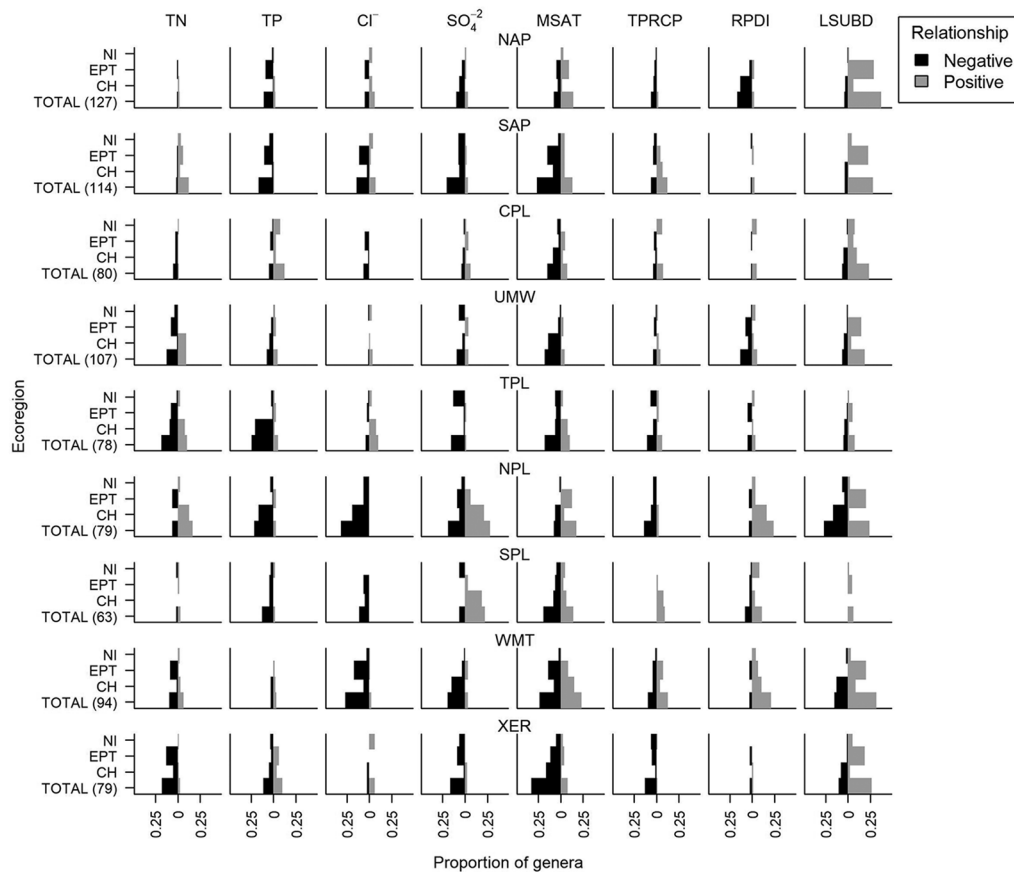


Figure 3. The proportion of aquatic macroinvertebrate genera where the 90% credible interval for their taxon–environment relationship does not overlap with 0. Black indicates that the relationship is negative. Gray indicates that the relationship is positive. Numbers in parenthesis are total number of genera included for each ecoregion. CH = Chironomidae, CPL = Coastal Plains, EPT = Ephemeroptera, Plecoptera, Trichoptera, LSUBD = log mean substrate diameter, MSAT = mean summer air temperature, NAP = Northern Appalachians, NI = noninsect, NPL = Northern Plains, RPDI = riparian disturbance index, SAP = Southern Appalachians, SPL = Southern Plains, TN = total N, TP = total P, TPL = Temperate Plains, UMW = Upper Midwest, WMT = Western Mountains, XER = Xeric.

β_{TN} or β_{TP} . In many instances, a single trait was related to multiple environmental gradients, and these relationships could be opposing. *ThrmOptV* was positively related to β_{MSAT} but negatively related to β_{TN} , β_{TP} , β_{SO_4} , β_{TPRCP} , and β_{LSUBD} . Similarly, clingers were positively related to β_{LSUBD} but negatively related to β_{MSAT} in WMT. We also found that residual variation in trait–niche relationships was phylogenetically structured (i.e., $\rho > 0$; Table S3), indicating that closely related taxa tend to have similar taxon–environment relationships.

We identified relatively few unimodal taxon–environment relationships (Table 4). When unimodal relationships were detected, they occurred along gradients of SO_4^{-2} concentration and LSUBD. Most unimodal relationships for SO_4^{-2} were detected in the UMW and XER ecoregions, whereas most unimodal relationships for LSUBD were detected in the TPL and CPL ecoregions (See Appendix S3 for mean and 90% CrIs of optima values). Importantly, unimodal relationships accounted for <26% (typically <15%) of the taxa included for each ecoregion (Table 4).

DISCUSSION

HMSC is a multivariate framework that can seamlessly integrate disparate information (i.e., environmental factors, biological traits, phylogenetic relationships, and co-occurrences) routinely collected by biomonitoring programs to quantify taxon–environment relationships. We used HMSC to quantify taxon–environment relationships for benthic macroinvertebrate genera in 9 ecoregions of the contiguous United States and investigated whether these relationships are related to a suite of traits. Foremost, our results are important from a conservation standpoint (Strayer 2006, Collier et al. 2016) because they elucidate how a taxon’s probability of occurrence could change if an environmental gradient is altered (Fig. S2). Generally, anthropogenic activities increase nutrient and salinity concentrations (Friberg et al. 2010, Kefford et al. 2016), MSAT (Vose et al. 2017), and RPDI (Kaufmann et al. 1999); decrease sediment diameter (Murphy 2020, Kaufmann et al. 2022); and have region-specific effects on TPRCP (Easterling et al. 2017). Given these expectations, we found that more taxon–environment relationships

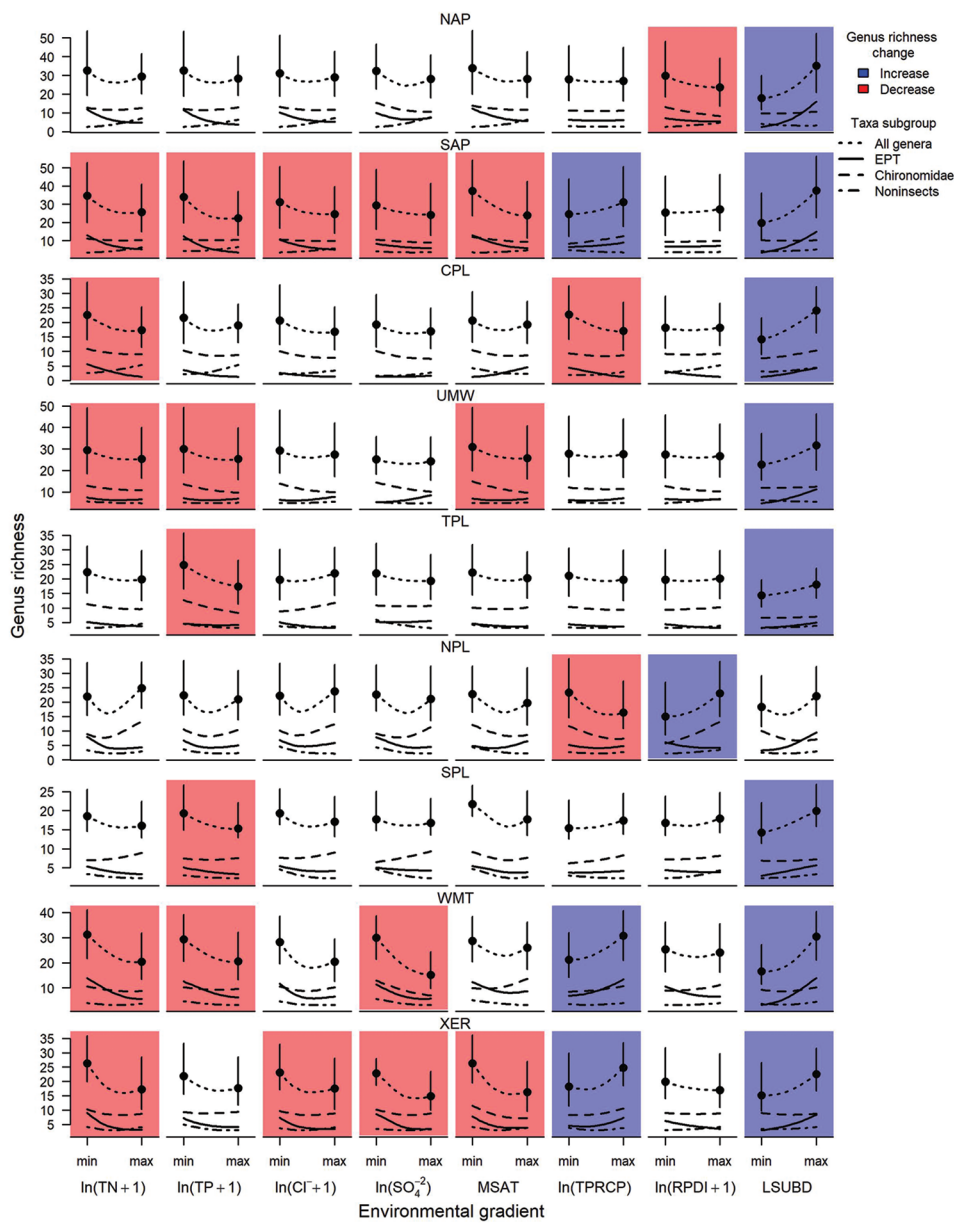


Figure 4. Predicted aquatic macroinvertebrate genus richness at the minimum and maximum value for environmental gradients in 9 ecoregions across the contiguous United States. Dots are mean genus richness estimates, and bars are 90% credible intervals. Red indicates that mean richness at the maximum gradient value was lower than mean richness at the minimum gradient value. Blue indicates that mean richness at the maximum gradient value was higher than mean richness at the minimum value. Lines indicate the trend in richness for all genera and taxonomic subgroups along gradients. All gradients were standardized to the same scale for visualization. CPL = Coastal Plains, EPT = Ephemeroptera, Plecoptera, Trichoptera, LSUBD = log mean substrate diameter, max = maximum, min = minimum, MSAT = mean summer air temperature, NAP = Northern Appalachians, NPL = Northern Plains, RPDI = riparian disturbance index, SAP = Southern Appalachians, SPL = Southern Plains, TN = total N, TP = total P, TPL = Temperate Plains, TPRCP = total annual precipitation, UMW = Upper Midwest, WMT = Western Mountains, XER = Xeric.

were negative in the direction of anthropogenic alteration, and genus richness often declined with increasing alteration. Because the number of taxon–environment relationships and the change in genus richness varied among gradients, our results suggest that changes to some gradients could have a

greater effect on macroinvertebrate assemblages than others. Further, we found trait–niche relationships were often inconsistent with a priori expectations, which suggests that traits may not consistently predict taxon–environment relationships in all ecoregions.

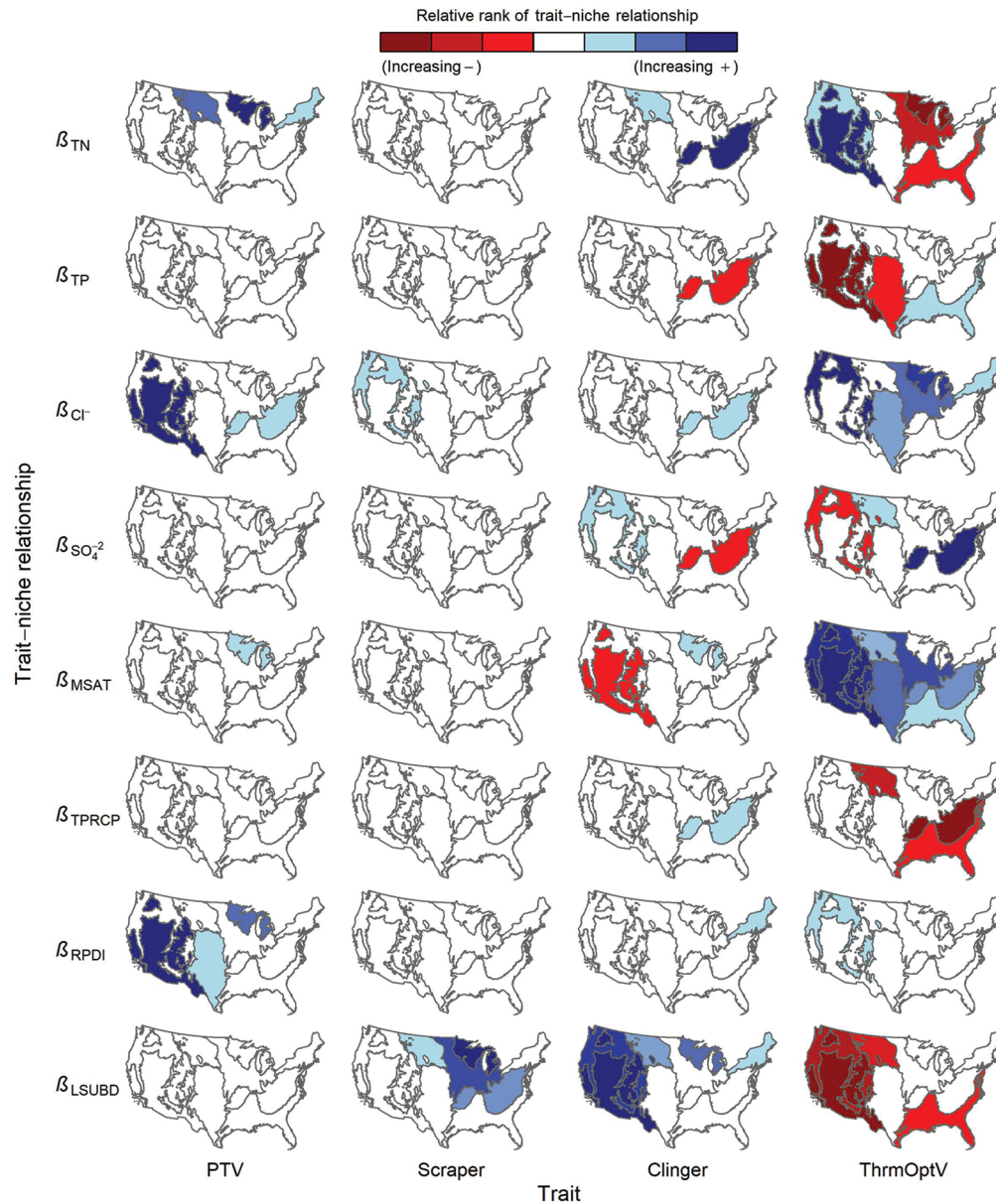


Figure 5. Trait–niche relationships for 9 ecoregions across the contiguous United States. Blue indicates that the 90% credible interval (CrI) for the relationship is >0 , and red indicates that the 90% CrI for relationship is <0 . White indicates that the 90% CrI for the relationship contains 0. For each map, color shading differentiates ecoregions when multiple trait–niche relationships were detected. Darker colors indicate stronger relationships, relative to other ecoregions. CPL = Coastal Plains, LSUBD = log mean substrate diameter, MSAT = mean summer air temperature, NAP = Northern Appalachians, NPL = Northern Plains, PTV = pollution tolerance value, RPDI = riparian disturbance index, SAP = Southern Appalachians, SPL = Southern Plains, ThrmOptV = thermal optima value, TN = total N, TP = total P, TPL = Temperate Plains, UMW = Upper Midwest, WMT = Western Mountains, XER = Xeric.

Classifying taxon-specific sensitivity and tolerance

Taxon-specific sensitivity and tolerance values can be assigned using expert opinion (Hawkes 1997), cumulative frequency distributions (Ashton et al. 2014), weighted averaging (Whittier and Van Sickle 2010), or univariate-regression approaches (Yuan 2004, Segurado et al. 2011). Expert opinion can be subjective and not supported by empirical data, and cumulative frequency distributions and

weighted averaging do not account for natural variability (Segurado et al. 2011). Univariate regression-based approaches can account for natural factors and have been used to categorize taxa as sensitive or tolerant from the shape of their response along a stressor gradient (Yuan 2004) or the difference between a taxon's occurrence and model predictions (Segurado et al. 2011). HMSC could advance regression-based approaches because co-occurrences are used to

Table 4. Proportion of taxon–environment relationships identified as unimodal. Unimodal relationships were not detected for the other gradients we assessed. CPL = Coastal Plains, LSUBD = log geometric mean substrate diameter, NAP = Northern Appalachians, NPL = Northern Plains, SAP = Southern Appalachians, SPL = Southern Plains, TPL = Temperate Plains, UMW = Upper Midwest, WMT = Western Mountains, XER = Xeric.

Region	SO ₄ ⁻²	LSUBD
NAP	0.15	0.12
SAP	0.00	0.02
CPL	0.06	0.16
UMW	0.23	0.07
TPL	0.00	0.26
NPL	0.00	0.01
SPL	0.00	0.05
WMT	0.10	0.06
XER	0.22	0.05

estimate sample-level random effects (Warton et al. 2015, Norberg et al. 2019, Ovaskainen and Abrego 2020), which account for unmeasured variables and improve estimates for the fixed effects (Bolker et al. 2009, McElreath 2016). Sample-level random effects cannot be estimated for univariate models because they are completely confounded with residual variation (Ovaskainen and Abrego 2020). Further, HMSC is fitted using Bayesian inference such that measures of uncertainty (i.e., 90% CrI) are generated for each taxon–environment relationship.

Because taxon–environment relationships can assess how taxa could change if the environment changes (Domisch et al. 2011, Funk et al. 2013, Bennetsen et al. 2016), the direction of the relationship can indicate whether they are sensitive or tolerant (Yuan 2004). Indeed, Yuan (2004) identified taxa with a monotonically decreasing relationship along a stressor gradient as sensitive and demonstrated that sensitive taxa richness decreased with increasing anthropogenic disturbance. Our results could be used in a similar fashion to quantify the number of sensitive taxa that occur at a site. In addition, we provide the magnitude of each taxon–environment relationship (i.e., estimated regression coefficient). These values could be used to assess the relative importance of each environmental gradient for a taxon’s occurrence (Van Sickle et al. 2006) or to rank taxa within an ecoregion according to their sensitivity or tolerance to a specific gradient.

We define taxon–environment relationships across the entire environmental gradient observed within an ecoregion such that sensitivity and tolerance can be adapted to situations where alteration is site specific or in a different direction than general trends (Fig. S2). For example, some prairie streams are historically turbid (Matthews 1988), and a reduction in fine sediments could be considered a stressor

for taxa that would occur there naturally. Similarly, waterbodies that have naturally high nutrient (Carpenter 1981) or salinity (Kefford et al. 2016) levels may have taxa that could decline if nutrient or salinity concentrations decreased. Indeed, macroinvertebrate assemblages are often used to infer difficult-to-measure abiotic conditions (Stevenson et al. 2008, Hubler et al. 2016), and acquiring an expectation of how anthropogenic activities change environmental conditions could be challenging. However, many studies have generated site-specific estimates for least-disturbed abiotic conditions (sensu Stoddard et al. 2006) from hindcast modeling (Dodds and Oakes 2004, Hawkins et al. 2010, Olson and Hawkins 2012, Hill and Hawkins 2014, Kaufmann et al. 2022) or regionally representative reference sites (Reynoldson et al. 1997, Stoddard et al. 2006, 2008, Herlihy et al. 2008), and historical climate data are available (<https://prism.oregonstate.edu/historical/>). Thus, the taxon–environment relationships we report could be used in tandem with efforts to model expected, desired, or historical abiotic conditions, but it is important to be cognizant of the range of environmental values used to fit JSDMs because the taxon–environment relationship could not be valid if extrapolated.

Taxon–environment relationships and changes in genus richness

We found that the proportion of taxa associated with each gradient and the change in genus richness along the gradients varied among ecoregions. One potential explanation for these differences is that the ecoregion boundaries enclose different portions of the environmental gradients (Fig. 1). If an organism’s range extends beyond these boundaries, it is conceivable that 1 ecoregion contains values at the upper limit of a taxon’s distribution, and another contains values at the lower limit. Accordingly, regions that contain extreme ends of a gradient could have a greater proportion of the taxa with positive or negative associations than regions that only contain intermediate values. Others have proposed that inconsistencies in gradient length among regions could contribute to context dependency in an organism’s response to a stressor (Feld et al. 2016, Segurado et al. 2022), and future efforts could evaluate potential effects of gradient lengths by comparing the length of a gradient across a taxon’s range to the length of the gradient within an ecoregion.

Regardless of the underlying reason, regional variation provides insights into the proportion of a regional assemblage that could respond if the gradient changes and into the magnitude of change in genus richness between the gradient’s extreme values. We found that substrate diameter and MSAT were often associated with the greatest proportion of taxa. Indeed, substrate size may affect benthic macroinvertebrates by providing refugia, trapping detritus, increasing surface area (Flecker and Allan 1984), or enhancing structural heterogeneity (but see Palmer et al.

2010). On the other hand, MSAT is a major determinant of streamwater temperature, which, in turn, could influence the distribution of stream biota (Hill and Hawkins 2014, Hill et al. 2014). Our finding that relatively more taxa were associated with these gradients could suggest that changes to them could have a relatively large effect on benthic macroinvertebrates.

We also found that genus richness could change by as many as 5 to 17 genera and was typically lower in the expected direction of anthropogenic disturbance. For example, we predicted genus richness to be higher when mean substrate diameter was larger, which is consistent with our finding of proportionally more taxa being associated with this gradient. Interestingly, genus richness was often lower in habitats with higher nutrient concentrations, even though relatively few taxa were associated with nutrient gradients. The decrease in genus richness along nutrient gradients suggests that a few taxa with strong negative associations can change assemblage-level metrics. Collectively, these changes in genus richness could support the notion that increases in fine sediment inputs (i.e., decreasing mean substrate diameter) and nutrient inputs could decrease macroinvertebrate genus richness (Wood and Armitage 1997, Jones et al. 2012, Stoddard et al. 2016, Lin et al. 2021).

In addition, we report the taxon–environment relationships and changes in genus richness for taxonomic subgroups because Chironomidae, EPT, and noninsects are widely used indices in biological assessments (USEPA 2020). By providing information about individual taxonomic groups, it could elucidate the importance of taxon-specific responses in the behavior of aggregated metrics (King and Baker 2010). Most genera we modeled belonged to Chironomidae and, therefore, accounted for most of the taxon–environment relationships. This family is often considered an indicator of poor water quality (Serra et al. 2017), but we found that the direction of taxon–environment relationships varied among genera for many gradients. Similarly, we found variation among taxon–environment relationships for genera belonging to EPT, even though this group is often used to indicate high-quality streams (King and Baker 2010). Regardless of the mixed associations at the genus level, we found that EPT genera richness typically changed more sharply along the gradients compared with the other taxonomic groups and seemingly contributed more to the overall pattern in genus richness. Indeed, this result is perhaps unsurprising given many successful applications of EPT metrics in biological assessment.

Trait–niche relationships

An organism's traits can provide a mechanistic link to their occurrence along an environmental gradient (Statzner and Bêche 2010, Heino et al. 2013), and we found that simple a priori expectations were supported for some trait–niche relationships. In ecoregions where expectations were cor-

roborated, our results support the notion that the expression of a trait is related to an organism's association with an environmental gradient (Townsend et al. 1997, Ovaskainen and Abrego 2020). For example, organisms with clinging habits typically have a positive response to substrate size and are, therefore, potentially less likely to tolerate sedimentation (i.e., trait is presumed to be beneficial in habitats with coarser substrates). Similarly, organisms with high ThrmOptV or PTV typically respond positively to MSAT and nutrient concentrations. As a consequence, organisms possessing these traits could become more prevalent under warming temperatures or increasing levels of organic pollution. In ecoregions where we confirmed that traits are related to taxon–environment relationships, knowledge of an organism's traits may help provide a mechanistic explanation for their response if the environmental gradient was altered.

However, not all a priori trait–niche relationships were supported. For example, a genus' response to nutrient concentrations was typically not related to whether they were a scraper or to their pollution tolerance level. Scrapers could have positive responses to nutrient concentrations via indirect effects on primary production (Feminella and Hawkins 1995, Stelzer and Lamberti 2001), and PTV could have a positive association when TN is associated with organic pollution. The lack of a strong relationship for both scrapers and pollution tolerance may indicate that another gradient could determine whether these traits are advantageous. For example, scrapers tend to consume biomass that attaches to coarse sediments (Yao et al. 2017), and a relatively open canopy is needed for nutrients to enhance primary production. If high nutrient concentrations do not coincide with appropriate substrates or an open canopy, then the trait may be decoupled from an organism's response to nutrient gradients. In addition, the relatively coarse assignment of genera to trait values could also affect our results (Yuan 2010), and a fuzzy-coding approach to assignments may better capture plasticity among individuals and differences in life cycles (Silva et al. 2017). Understanding the conditions and contingencies in trait relationships requires further investigation.

Unimodal relationships

Unimodal taxon–environment relationships are fundamental to niche theory (Yuan 2006) but are not accurately modeled by linear relationships. Generally, unimodal relationships are expected when abiotic stress limits the distribution of taxa at both ends of an environmental gradient (Normand et al. 2009). These unimodal relationships were uncommon, which could indicate that the range of values for environmental gradients within an ecoregion often did not include the full range of conditions that limit each taxon's distribution. For example, we did not detect unimodal relationships with MSAT, and this gradient was relatively narrow within each ecoregion whereas the few unimodal relationships we detected were along relatively long gradients of LSUBD and

SO_4^{-2} (Fig. 1). Still, for some relatively long gradient lengths (e.g. RPDI), we found a low number of unimodal relationships. This result could suggest that some organisms occur only in habitats without anthropogenic activity or that are degraded and do not have a minimum or maximum requirement for an environmental gradient (Yuan 2004, 2006).

Future application

HMSC uses a single model to provide a comprehensive framework to quantify taxon–environment relationships, investigate the association between traits and these relationships, and predict how assemblages could change along environmental gradients. This framework is rich with opportunities for future analysis (Warton et al. 2015, Ovaskainen et al. 2017, Norberg et al. 2019, Ovaskainen and Abrego 2020).

Foremost, conditional predictions are possible because HMSC models all taxa simultaneously and allows the model to predict the occurrence of a taxon based on the known occurrence of other taxa. We found that conditional prediction improved model performance for all ecoregions, which suggests that future efforts could consider fewer covariates (fixed effects) without sacrificing predictive performance if some taxa are considered known (Fig. 2). Conditional prediction could also be useful for incorporating biotic interactions into predictions (White et al. 2021). For example, in a restoration context, it may be possible to assess whether a taxon's probability of occurrence at a site increases or decreases based on the presence of another taxon. In the present study, we found that conditional predictions varied among ecoregions, which could suggest that pairwise associations between taxa are weaker in some regions (Ovaskainen and Abrego 2020). Interestingly, anthropogenic disturbance can weaken co-occurrences among macroinvertebrates (Larsen and Ormerod 2014), and we found that the ecoregions suspected to have the most anthropogenic disturbance (i.e., plains) had the least improvement in conditional performance, but this pattern requires further investigation.

HMSC can also partition variance in occurrences among groups of environmental variables and random processes (Ovaskainen and Abrego 2020). In a supplementary analysis, we confirmed that the fixed effects explained more variation in occurrences than the random effects for all regions except UMW, and the proportion of variance explained by a group of variables differed among regions (Fig. S3). However, although variance partitioning can be applied to any model, we refrain from interpreting these results extensively because probit models use a nonlinear link function to convert the linear predictor into the scale of the data (Ovaskainen and Abrego 2020). Thus, although this feature could potentially offer further insight into factors underlying regional variation, further analysis is needed.

Finally, HMSC can be complex and computationally intensive to fit. This process required us to use different model

variations to identify unimodal relationships and quantify the effects of multiple gradients as additive fixed effects. However, the effects of multiple stressors are not always additive, and understanding the interactions among stressors is important (Townsend et al. 2008, Clements et al. 2012, Feld et al. 2016, Segurado et al. 2022). We recommend future research incorporate interactions by focusing on fewer environmental gradients and smaller spatial extents to avoid overfitting and computational limitations. Nonetheless, this framework permitted us to evaluate relationships between benthic macroinvertebrate genera and a suite of environmental gradients commonly altered by human activities for multiple ecoregions. Importantly, our results indicate that these relationships can be geographically dependent and suggest that a regionalized perspective is necessary to develop more localized estimates of taxon-specific sensitivities and tolerances.

ACKNOWLEDGEMENTS

Author contributions: DAK, JLS, RAH, and SGP conceived and designed the study. DAK analyzed the data. DAK, JLS, PRK, RAH, ATH, and JMD interpreted the data. DAK wrote the 1st draft of the manuscript. DAK, JLS, RAH, PRK, ATH, and JMD contributed to subsequent revisions and editing.

We would like to thank L. Riato, E. Fergus, K. Blocksom, and D. Peck for valuable discussions throughout the development of this project. We would also like to thank L. Yuan, N. Schumaker, P. Meeks, and 2 anonymous reviewers for providing insightful comments on previous versions of this manuscript. The information in this document has been funded entirely by the United States Environmental Protection Agency (USEPA), in part through appointments to the USEPA's Internship/Research Participation Program at the Office of Research and Development administered by the Oak Ridge Institute for Science and Education through an inter-agency agreement. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the USEPA. Any mention of trade names, products, or services does not imply an endorsement by the United States Government or the USEPA. The USEPA does not endorse any commercial products, services, or enterprises.

Open research statement: Data and code are available at United States Environmental Protection Agency ScienceHub (<https://doi.org/10.23719/1527632>).

LITERATURE CITED

- Abrego, N., A. Norberg, and O. Ovaskainen. 2017. Measuring and predicting the influence of traits on the assembly processes of wood-inhabiting fungi. *Journal of Ecology* 105:1070–1081.
- Araújo, M. B., R. P. Anderson, A. M. Barbosa, C. M. Beale, C. F. Dormann, R. Early, R. A. Garcia, A. Guisan, L. Maiorano, B. Naimi, R. B. O'Hara, N. E. Zimmermann, and C. Rahbek. 2019. Standards for distribution models in biodiversity assessments. *Science Advances* 5:4858–4874.
- Ashton, M. J., R. P. Morgan, and S. Stranko. 2014. Relations between macroinvertebrates, nutrients, and water quality criteria in wadeable streams of Maryland, USA. *Environmental Monitoring and Assessment* 186:1167–1182.

- Austin, M. 2007. Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling* 200:1–19.
- Bennetsen, E., S. Gobeyn, and P. L. M. Goethals. 2016. Species distribution models grounded in ecological theory for decision support in river management. *Ecological Modelling* 325:1–12.
- Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. 2009. Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution* 24:127–135.
- Bolnick, D. I., P. Amarasekare, M. S. Araújo, R. Bürger, J. M. Levine, M. Novak, V. H. W. Rudolf, S. J. Schreiber, M. C. Urban, and D. A. Vasseur. 2011. Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution* 26:183–192.
- Bonada, N., N. Prat, V. H. Resh, and B. Statzner. 2006. Developments in aquatic insect biomonitoring: A comparative analysis of recent approaches. *Annual Review of Entomology* 51:495–523.
- Burner, R. C., J. G. Stephan, L. Drag, T. Birkemoe, J. Muller, T. Snäll, O. Ovaskainen, M. Potterf, J. Siitonen, O. Skarpaas, I. Doerfler, M. M. Gossner, P. Schall, W. W. Weisser, and A. Sverdrup-Thygeson. 2021. Traits mediate niches and co-occurrences of forest beetles in ways that differ among bioclimatic regions. *Journal of Biogeography* 48:3145–3157.
- Buss, D. F., D. M. Carlisle, T.-S. Chon, J. Culp, J. S. Harding, H. E. Keizer-Vlek, W. A. Robinson, S. Strachan, C. Thirion, and R. M. Hughes. 2015. Stream biomonitoring using macroinvertebrates around the globe: A comparison of large-scale programs. *Environmental Monitoring and Assessment* 187:4132.
- Carpenter, S. R. 1981. Submersed vegetation: An internal factor in lake ecosystem succession. *The American Naturalist* 118:372–383.
- Céspedes, V., S. Pallarés, P. Arribas, A. Millán, and J. Velasco. 2013. Water beetle tolerance to salinity and anionic composition and its relationship to habitat occupancy. *Journal of Insect Physiology* 59:1076–1084.
- Chang, F. H., J. E. Lawrence, B. Rios-Touma, and V. H. Resh. 2014. Tolerance values of benthic macroinvertebrates for stream biomonitoring: Assessment of assumptions underlying scoring systems worldwide. *Environmental Monitoring and Assessment* 186:2135–2149.
- Clarke, R. T., J. F. Wright, and M. T. Furse. 2003. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. *Ecological Modelling* 160:219–233.
- Clements, W. H., C. W. Hickey, and K. A. Kidd. 2012. How do aquatic communities respond to contaminants? It depends on the ecological context. *Environmental Toxicology and Chemistry* 31:1932–1940.
- Collier, K. J., P. K. Probert, and M. Jeffries. 2016. Conservation of aquatic invertebrates: Concerns, challenges and conundrums. *Aquatic Conservation: Marine and Freshwater Ecosystems* 26:817–837.
- Dodds, W., and R. Oakes. 2004. A technique for establishing reference nutrient concentrations across watersheds affected by humans. *Limnology and Oceanography: Methods* 2:333–341.
- Dolédéc, S., and B. Statzner. 2010. Responses of freshwater biota to human disturbances: Contribution of J-NABS to developments in ecological integrity assessments. *Journal of the North American Benthological Society* 29:286–311.
- Domisch, S., S. C. Jähnig, and P. Haase. 2011. Climate-change winners and losers: Stream macroinvertebrates of a submontane region in Central Europe. *Freshwater Biology* 56:2009–2020.
- Easterling, D. R., J. R. Arnold, T. Knutson, K. E. Kunkel, A. N. LeGrande, L. R. Leung, R. S. Vose, D. E. Waliser, and M. F. Wehner. 2017. Precipitation change in the United States. Pages 207–230 in D. J. Wuebbles, D. W. Fahey, K. A. Hibbard, D. J. Dokken, B. C. Stewart, and T. K. Maycock (editors). *Climate science special report: Fourth National climate assessment Volume I. US Global Change Research Program, Washington, DC.*
- Elith, J., and J. R. Leathwick. 2009. Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40:677–697.
- Feld, C. K., P. Segurado, and C. Gutiérrez-Cánovas. 2016. Analysing the impact of multiple stressors in aquatic biomonitoring data: A ‘cookbook’ with applications in R. *Science of the Total Environment* 573:1320–1339.
- Feminella, J. W., and C. P. Hawkins. 1995. Interactions between stream herbivores and periphyton: A quantitative analysis of past experiments. *Journal of the North American Benthological Society* 14:465–509.
- Finn, D. S., and N. L. Poff. 2005. Variability and convergence in benthic communities along the longitudinal gradients of four physically similar Rocky Mountain streams. *Freshwater Biology* 50:243–261.
- Flecker, A. S., and D. J. Allan. 1984. The importance of predation, substrate and spatial refugia in determining lotic insect distributions. *Oecologia* 64:306–313.
- Friberg, N., J. Skriver, S. E. Larsen, M. L. Pedersen, and A. Buffagni. 2010. Stream macroinvertebrate occurrence along gradients in organic pollution and eutrophication. *Freshwater Biology* 55:1405–1419.
- Funk, A., C. Gschöpf, A. P. Blaschke, G. Weigelhofer, and W. Reckendorfer. 2013. Ecological niche models for the evaluation of management options in an urban floodplain-conservation vs. restoration purposes. *Environmental Science and Policy* 34:79–91.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. *Bayesian data analysis*. 3rd edition. CRC Press, Boca Raton, Florida.
- Groffman, P. M., J. Cavender-Bares, N. D. Bettez, J. M. Grove, S. J. Hall, J. B. Heffernan, S. E. Hobbie, K. L. Larson, J. L. Morse, C. Neill, K. Nelson, J. O’Neil-Dunne, L. Ogden, D. E. Pataki, C. Polsky, R. R. Chowdhury, and M. K. Steele. 2014. Ecological homogenization of urban USA. *Frontiers in Ecology and the Environment* 12:74–81.
- Hamilton, A. T., R. B. Schäfer, M. I. Pyne, B. Chessman, K. Kakouei, K. S. Boersma, P. F. M. Verdonschot, R. C. M. Verdonschot, M. Mims, K. Khamis, B. Bierwagen, and J. Stamp. 2020. Limitations of trait-based approaches for stressor assessment: The case of freshwater invertebrates and climate drivers. *Global Change Biology* 26:364–379.
- Hawkes, H. A. 1997. Origin and development of the biological monitoring working party score system. *Water Research* 32:964–968.
- Hawkins, C. P., H. Mykrä, J. Oksanen, and J. J. vander Laan. 2015. Environmental disturbance can increase beta diversity of

- stream macroinvertebrate assemblages. *Global Ecology and Biogeography* 24:483–494.
- Hawkins, C. P., J. R. Olson, and R. A. Hill. 2010. The reference condition: Predicting benchmarks for ecological and water-quality assessments. *Journal of the North American Benthological Society* 29:312–343.
- Hawkins, C. P., and L. L. Yuan. 2016. Multitaxon distribution models reveal severe alteration in the regional biodiversity of freshwater invertebrates. *Freshwater Science* 35:1365–1376.
- Heino, J., D. Schmera, and T. Eros. 2013. A macroecological perspective of trait patterns in stream communities. *Freshwater Biology* 58:1539–1555.
- Herlihy, A. T., S. G. Paulsen, J. Van Sickle, J. L. Stoddard, C. P. Hawkins, and L. L. Yuan. 2008. Striving for consistency in a national assessment: The challenges of applying a reference-condition approach at a continental scale. *Journal of the North American Benthological Society* 27:860–877.
- Hill, R. A., and C. P. Hawkins. 2014. Using modelled stream temperatures to predict macro-spatial patterns of stream invertebrate biodiversity. *Freshwater Biology* 59:2632–2644.
- Hill, R. A., C. P. Hawkins, and J. Jin. 2014. Predicting thermal vulnerability of stream and river ecosystems to climate change. *Climatic Change* 125:399–412.
- Hubler, S., D. D. Huff, P. Edwards, and Y. Pan. 2016. The Biological Sediment Tolerance Index: Assessing fine sediments conditions in Oregon streams using macroinvertebrates. *Ecological Indicators* 67:132–145.
- Hughes, R. M., and D. V. Peck. 2008. Acquiring data for large aquatic resource surveys: The art of compromise among science, logistics, and reality. *Journal of the North American Benthological Society* 27:837–859.
- Jones, J. I., J. F. Murphy, A. L. Collins, D. A. Sear, P. S. Naden, and P. D. Armitage. 2012. The impact of fine sediment on macroinvertebrates. *River Research and Applications* 28:1055–1071.
- Karr, J. R. 1981. Assessment of biotic integrity using fish communities. *Fisheries* 6:21–27.
- Kaufmann, P. R., R. M. Hughes, S. G. Paulsen, D. V. Peck, C. W. Seeliger, M. H. Weber, and R. M. Mitchell. 2022. Physical habitat in conterminous US streams and rivers, Part 1: Geoclimatic controls and anthropogenic alteration. *Ecological Indicators* 141:109046.
- Kaufmann, P. R., D. P. Larsen, and J. M. Faustini. 2009. Bed stability and sedimentation associated with human disturbances in Pacific Northwest streams. *Journal of the American Water Resources Association* 45:434–459.
- Kaufmann, P. R., P. Levine, E. G. Robison, C. Seeliger, and D. V. Peck. 1999. Quantifying physical habitat in wadeable streams. Environmental Monitoring and Assessment Program, National Health and Environmental Effects Research Laboratory, Office of Research and Development, United States Environmental Protection Agency, Research Triangle Park, North Carolina.
- Kefford, B. J., D. Buchwalter, M. Cañedo-Argüelles, J. Davis, R. P. Duncan, A. Hoffmann, and R. Thompson. 2016. Salinized rivers: Degraded systems or new habitats for salt-tolerant faunas? *Biology Letters* 12:20151072.
- King, R. S., and M. E. Baker. 2010. Considerations for analyzing ecological community thresholds in response to anthropogenic environmental gradients. *Journal of the North American Benthological Society* 29:998–1008.
- Larsen, S., and S. J. Ormerod. 2014. Anthropogenic modification disrupts species co-occurrence in stream invertebrates. *Global Change Biology* 20:51–60.
- Lin, J., J. E. Compton, R. A. Hill, A. T. Herlihy, R. D. Sabo, J. R. Brooks, M. Weber, B. Pickard, S. G. Paulsen, and J. L. Stoddard. 2021. Context is everything: Interacting inputs and landscape characteristics control stream nitrogen. *Environmental Science & Technology* 55:7890–7899.
- Mao, F., X. Zhao, P. Ma, S. Chi, K. Richards, D. M. Hannah, and S. Krause. 2019. Revision of biological indices for aquatic systems: A ridge-regression solution. *Ecological Indicators* 106.
- Matthews, W. J. 1988. North American Prairie streams as systems for ecological study. *Journal of the North American Benthological Society* 7:387–409.
- McElreath, R. 2016. *Statistical rethinking: A Bayesian course with examples in R and Stan*. 1st edition. CRC Press, Boca Raton, Florida.
- Moss, D., M. T. Furse, J. F. Wright, and P. D. Armitage. 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. *Freshwater Biology* 17:41–52.
- Murphy, J. C. 2020. Changing suspended sediment in United States rivers and streams: Linking sediment trends to changes in land use/cover, hydrology and climate. *Hydrology and Earth System Sciences* 24:991–1010.
- Norberg, A., N. Abrego, F. G. Blanchet, F. R. Adler, B. J. Anderson, J. Anttila, M. B. Araújo, T. Dallas, D. Dunson, J. Elith, S. D. Foster, R. Fox, J. Franklin, W. Godsoe, A. Guisan, B. O'Hara, N. A. Hill, R. D. Holt, F. K. C. Hui, M. Husby, J. A. Kålås, A. Lehtikoinen, M. Luoto, H. K. Mod, G. Newell, I. Renner, T. Roslin, J. Soininen, W. Thuiller, J. Vanhatalo, D. Warton, M. White, N. E. Zimmermann, D. Gravel, and O. Ovaskainen. 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs* 89:e01370.
- Normand, S., U. A. Treier, C. Randin, P. Vittoz, A. Guisan, and J. C. Svenning. 2009. Importance of abiotic stress as a range-limit determinant for European plants: Insights from species responses to climatic gradients. *Global Ecology and Biogeography* 18:437–449.
- Olsen, A. R., and D. V. Peck. 2008. Survey design and extent estimates for the Wadeable Streams Assessment. *Journal of the North American Benthological Society* 27:822–836.
- Olson, J. R., and C. P. Hawkins. 2012. Predicting natural base-flow stream water chemistry in the western United States. *Water Resources Research* 48:W02504.
- Osborne, P. E., G. M. Foody, and S. Suárez-Seoane. 2007. Non-stationarity and local approaches to modelling the distributions of wildlife. *Diversity and Distributions* 13:313–323.
- Osborne, P. E., and S. Suárez-Seoane. 2002. Should data be partitioned spatially before building large-scale distribution models? *Ecological Modelling* 157:249–259.
- Ovaskainen, O., and N. Abrego. 2020. *Joint species distribution modelling with applications in R*. Cambridge University Press, Cambridge, United Kingdom.
- Ovaskainen, O., N. Abrego, P. Halme, and D. Dunson. 2016a. Using latent variable models to identify large networks of

- species-to-species associations at different spatial scales. *Methods in Ecology and Evolution* 7:549–555.
- Ovaskainen, O., D. B. Roy, R. Fox, and B. J. Anderson. 2016b. Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models. *Methods in Ecology and Evolution* 7:428–436.
- Ovaskainen, O., G. Tikhonov, A. Norberg, F. G. Blanchet, L. Duan, D. Dunson, T. Roslin, and N. Abrego. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters* 20:561–576.
- Palmer, M. A., H. L. Menninger, and E. Bernhardt. 2010. River restoration, habitat heterogeneity and biodiversity: A failure of theory or practice? *Freshwater Biology* 55:205–222.
- Paradis, E., and K. Schliep. 2019. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35:526–528.
- Poff, N. L., J. D. Olden, D. M. Merritt, and D. M. Pepin. 2007. Homogenization of regional river dynamics by dams and global biodiversity implications. *Proceedings of the National Academy of Sciences* 104:5732–5737.
- Resh, V. H. 2008. Which group is best? Attributes of different biological assemblages used in freshwater biomonitoring programs. *Environmental Monitoring and Assessment* 138:131–138.
- Reynoldson, T. B., R. H. Norris, V. H. Resh, K. E. Day, and D. M. Rosenberg. 1997. The reference condition: A comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society* 16:833–852.
- Saito, V. S., T. Siqueira, L. M. Bini, R. Costa-Pereira, E. P. Santos, and S. Pavoiné. 2020. Comparing taxon- and trait-environment relationships in stream communities. *Ecological Indicators* 117: 106625.
- Segurado, P., C. Gutiérrez-Cánovas, T. Ferreira, and P. Branco. 2022. Stressor gradient coverage affects interaction identification. *Ecological Modelling* 472:110089.
- Segurado, P., J. M. Santos, D. Pont, A. H. Melcher, D. G. Jalon, R. M. Hughes, and M. T. Ferreira. 2011. Estimating species tolerance to human perturbation: Expert judgment versus empirical approaches. *Ecological Indicators* 11:1623–1635.
- Serra, S. R. Q., M. A. S. Graça, S. Dolédec, and M. J. Feio. 2017. Chironomidae traits and life history strategies as indicators of anthropogenic disturbance. *Environmental Monitoring and Assessment* 189:326.
- Silva, D. R. O., A. T. Herlihy, R. M. Hughes, and M. Callisto. 2017. An improved macroinvertebrate multimetric index for the assessment of wadeable streams in the neotropical savanna. *Ecological Indicators* 81:514–525.
- Statzner, B., and L. A. Bêche. 2010. Can biological invertebrate traits resolve effects of multiple stressors on running water ecosystems? *Freshwater Biology* 55:80–119.
- Stelzer, R. S., and G. A. Lamberti. 2001. Effects of N:P ratio and total nutrient concentration on stream periphyton community structure, biomass, and elemental composition. *Limnology and Oceanography* 46:356–367.
- Stevenson, R. J., Y. Pan, K. M. Manoylov, C. A. Parker, D. P. Larsen, and A. T. Herlihy. 2008. Development of diatom indicators of ecological conditions for streams of the western US. *Journal of the North American Benthological Society* 27:1000–1016.
- Stoddard, J. L., A. T. Herlihy, D. V. Peck, R. M. Hughes, T. R. Whittier, and E. Tarquinio. 2008. A process for creating multimetric indices for large-scale aquatic surveys. *Journal of the North American Benthological Society* 27:878–891.
- Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: The concept of reference condition. *Ecological Applications* 16:1267–1276.
- Stoddard, J. L., J. Van Sickle, A. T. Herlihy, J. Brahney, S. Paulsen, D. V. Peck, R. Mitchell, and A. I. Pollard. 2016. Continental-scale increase in lake and stream phosphorus: Are oligotrophic systems disappearing in the United States? *Environmental Science & Technology* 50:3409–3415.
- Strayer, D. L. 2006. Challenges for freshwater invertebrate conservation. *Journal of the North American Benthological Society* 25:271–287.
- Tikhonov, G., Ø. H. Opedal, N. Abrego, A. Lehtikoinen, M. M. J. de Jonge, J. Oksanen, and O. Ovaskainen. 2020. Joint species distribution modelling with the R-package *Hmsc*. *Methods in Ecology and Evolution* 11:442–447.
- Tjur, T. 2009. Coefficients of determination in logistic regression models—A new proposal: The coefficient of discrimination. *American Statistician* 63:366–372.
- Townsend, C. R., and A. G. Hildrew. 1994. Species traits in relation to a habitat templet for river systems. *Freshwater Biology* 31:265–275.
- Townsend, C. R., M. R. Scarsbrook, and S. Dolédec. 1997. Species traits and species richness in relation to temporal and spatial heterogeneity in streams: A test of habitat templet theory. *Freshwater Biology* 37:367–387.
- Townsend, C. R., S. S. Uhlmann, and C. D. Matthaei. 2008. Individual and combined responses of stream ecosystems to multiple stressors. *Journal of Applied Ecology* 45:1810–1819.
- USEPA (United States Environmental Protection Agency). 2008. National Rivers and Streams Assessment: Laboratory methods manual. EPA-841-B-07-010. Office of Water, Office of Research and Development, United States Environmental Protection Agency, Washington, DC. (Available from: https://www.epa.gov/sites/default/files/2013-11/documents/nrsa200809_lab_methodsmanual_20091229_2_web.pdf)
- USEPA (United States Environmental Protection Agency). 2009. National Rivers and Streams Assessment: Field operations manual. EPA-841-B-07-009. Office of Water, Office of Research and Development, United States Environmental Protection Agency, Washington, DC. (Available from: <https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P100AVOF.TXT>)
- USEPA (United States Environmental Protection Agency). 2012. Freshwater biological traits database (final report). EPA-600-R-11-038F. United States Environmental Protection Agency, Washington, DC. (Available from: https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NCEA&dirEntryId=241813)
- USEPA (United States Environmental Protection Agency). 2016. National Rivers and Streams Assessment 2008-2009: Technical report. EPA-841-R-16-008. Office of Wetlands, Oceans and Watersheds, Office of Research and Development, United States Environmental Protection Agency, Washington, DC. (Available from: <https://www.epa.gov/national-aquatic-re-source-surveys/national-rivers-and-streams-assessment-2008-2009-technical-report>)

- USEPA (United States Environmental Protection Agency). 2017. National Rivers and Streams Assessment 2018/19: Field operations manual wadeable. EPA-841-B-17-003a. Office of Water, United States Environmental Protection Agency, Washington, DC. (Available from: https://www.epa.gov/sites/default/files/2019-05/documents/nrsa_1819_fom_wadeable_version_1.2_0.pdf)
- USEPA (United States Environmental Protection Agency). 2020. National Rivers and Streams Assessment 2013-2014 technical support document. EPA 843-R-19-001. Office of Water, Office of Research and Development. Washington, DC. (Available from: https://www.epa.gov/sites/default/files/2020-12/documents/nrsa_2013-14_final_tsd_12-15-2020.pdf)
- Van Sickle, J., J. L. Stoddard, S. G. Paulsen, and A. R. Olsen. 2006. Using relative risk to compare the effects of aquatic stressors at a regional scale. *Environmental Management* 38:1020–1030.
- Vose, R. S., D. R. Easterling, K. E. Kunkel, A. N. LeGrande, and M. F. Wehner. 2017. Temperature changes in the United States. Pages 185–206 *in* D. J. Wuebbles, D. W. Fahey, K. A. Hibbard, D. J. Dokken, B. C. Stewart, and T. K. Maycock (editors). *Climate science special report: Fourth national climate assessment. Volume I. United States Global Change Research Program*, Washington, DC. (Available from: https://science2017.globalchange.gov/downloads/CSSR_Ch6_Temperature.pdf)
- Warton, D. I., F. G. Blanchet, R. B. O'Hara, O. Ovaskainen, S. Taskinen, S. C. Walker, and F. K. C. Hui. 2015. So many variables: Joint modeling in community ecology. *Trends in Ecology & Evolution* 30:766–779.
- White, B. E., A. R. McIntosh, C. M. Febria, and H. J. Warburton. 2021. The potential role of biotic interactions in stream restoration. *Restoration Ecology* 29:e13396.
- Whittier, T. R., and J. Van Sickle. 2010. Macroinvertebrate tolerance values and an assemblage tolerance index (ATI) for western USA streams and rivers. *Journal of the North American Benthological Society* 29:852–866.
- Wood, P. J., and P. D. Armitage. 1997. Biological effects of fine sediment in the lotic environment. *Environmental Management* 21:203–217.
- Yao, J., F. Colas, A. G. Solimini, T. J. Battin, S. Gafny, M. Morais, M. Puig, E. Martí, M. T. Pusch, C. Voreadou, F. Sabater, F. Julien, J. M. Sánchez-Pérez, S. Sauvage, P. Vervier, and M. Gerino. 2017. Macroinvertebrate community traits and nitrate removal in stream sediments. *Freshwater Biology* 62:929–944.
- Yuan, L. L. 2004. Assigning macroinvertebrate tolerance classifications using generalised additive models. *Freshwater Biology* 49:662–677.
- Yuan, L. L. 2006. Estimation and application of macroinvertebrate tolerance values. EPA-600-P-04-116F. Office of Research and Development, United States Environmental Protection Agency, Washington, DC. (Available from: <https://cfpub.epa.gov/ncea/risk/hhra/recordisplay.cfm?deid=154869>)
- Yuan, L. L. 2010. Estimating the effects of excess nutrients on stream invertebrates from observational data. *Ecological Applications* 20:110–125.
- Yuan, L. L., C. P. Hawkins, and J. Van Sickle. 2008. Effects of regionalization decisions on an O/E index for the US national assessment. *Journal of the North American Benthological Society* 27: 892–905.
- Zettler, M. L., C. E. Proffitt, A. Darr, S. Degraer, L. Devriese, C. Greathead, J. Kotta, P. Magni, G. Martin, H. Reiss, J. Speybroeck, D. Tagliapietra, G. van Hoey, and T. Ysebaert. 2013. On the myths of indicator species: Issues and further consideration in the use of static concepts for ecological applications. *PLoS ONE* 8:e78219.