Community Ecology
https://doi.org/10.1007/s42974-024-00183-9

ORIGINAL ARTICLE

Quantifying the relative importance of biotic and abiotic factors in landscape-based models of stream fish distributions

Christopher A. Custer[1] · Douglas P. Fischer[2] · Geoffrey Smith[2] · Aaron Henning[3] · Megan Kepler Schall[4] · Matthew K. Shank[5] · Timothy A. Wertz[5] · Tyler Wagner[6]

Received: 17 July 2023 / Accepted: 9 February 2024

Abstract
Lotic fish species distributions are frequently predicted using remotely sensed habitat variables that characterize the adjacent landscape and serve as proxies for instream habitat. Recent advancements in statistical methodology, however, allow for leveraging fish assemblage data when predicting distributions. This is important because assemblage composition likely provides better information about instream habitat compared to landscape-derived metrics and therefore may improve predictions. To better understand the value of using multi-species fish data in species distribution modeling, we fit two conditional random fields (CRF) models to quantify the relative importance of fish assemblage co-occurrence, landscape-derived habitat variables, and interactions between these two predictor groups (i.e., effects of co-occurrence could be context-dependent) at over 1200 stream catchments in Pennsylvania, USA. We first compared predictive performance of CRF models against traditionally used single-species logistic regressions (generalized linear models; GLMs) and found that inclusion of fish assemblage data often improved predictive performance. The multi-species CRF models performed significantly better at predicting occurrence for 63% of species with an average percent increase in AUC of 25% compared to GLMs. Furthermore, the CRF identified species co-occurrences as more informative, and thus relatively more important, at predicting occurrence than the other effect types. The CRF also suggested that allowing these biotic effects to be context-dependent was important for predicting occurrence of many species. These findings illustrate the value of fish assemblage data for landscape-scale species distribution modeling and leveraging this information can improve predictions and inferences to help inform the management and conservation of freshwater fishes.

Keywords Lotic fish assemblages · Conditional random fields · Species distributions · Landscape-scale

Introduction

Predicting changes in species distributions can be a powerful tool for managers and conservationists as they attempt to understand how global change may impact freshwater ecosystems. Freshwater habitats are rich and diverse systems (Lundberg et al., 2000) and among the most threatened on the planet (Reid et al., 2019; Strayer & Dudgeon, 2010). Therefore, understanding the drivers and predicting the distributions of freshwater fishes is a global priority. Historically, species distribution models have often focused on the relationship between a single fish species and measures of abiotic habitat conditions (Olden et al., 2006). The focal species of such modeling efforts are most commonly managed—and therefore modeled—on a species-by-species basis, often focusing on important commercial and recreational species (e.g., Brandt et al., 2022; McKenna Jr &

✉ Christopher A. Custer
cac6877@psu.edu

[1] Intercollege Graduate Degree Program in Ecology, Pennsylvania Cooperative Fish and Wildlife Research Unit, Pennsylvania State University, University Park, PA, USA

[2] Pennsylvania Fish and Boat Commission, Bellefonte, PA, USA

[3] Susquehanna River Basin Commission, Harrisburg, PA, USA

[4] Biological Services, The Pennsylvania State University, Hazleton, PA, USA

[5] Pennsylvania Department of Environmental Protection, Harrisburg, PA, USA

[6] Pennsylvania Cooperative Fish and Wildlife Research Unit, U.S. Geological Survey, Pennsylvania State University, University Park, PA, USA

Published online: 30 April 2024

Johnson, 2011). Single-species models that use only abiotic landscape predictors implicitly make the assumption that landscape-derived predictors are reasonable proxies for instream habitat. For example, inclusion of the proportion of agricultural or urban land use in a stream catchment as a predictor variable in a species distribution model assumes that human-dominated landscapes result in degraded instream conditions and therefore are useful for prediction—which is often the case (Allan, 2004). Although such watershed-level metrics are useful in capturing broad-scale patterns in fish distributions (e.g., Kristensen et al., 2012), finer-scale heterogeneity in instream habitat that also affects species occurrences is likely not captured well. For example, DeWeber and Wagner (2015) developed a species distribution model for eastern brook trout *Salvelinus fontinalis* that focused on understanding abiotic drivers of brook trout occurrence across its native range in the eastern Unites States. They used the proportion of urban and agricultural land use as proxies for instream habitat and found that brook trout were less likely to occur in human altered landscapes—a finding that is consistent with expectations for altered instream habitat conditions and brook trout ecology. However, their approach—although providing valuable information about habitat factors influencing brook trout distributions—did not account for species co-occurrences that may also affect where brook trout occur. For example, the presence of non-native brown trout *Salmo trutta* can also affect the occurrence probability of native brook trout (Wagner et al., 2013). Therefore, the presence or absence of other species, such as brown trout in this example, may provide additional information about instream conditions that could improve predictions beyond those based solely on landscape-derived habitat metrics.

Although a single-species focus may be appropriate given specific research or management objectives, or necessary due to limited data, the paucity of multi-species studies that incorporate both community-wide co-occurrence information and abiotic drivers has been partly due to the lack of analytical frameworks that can accommodate such data. However, recent advancements in statistical methodologies allow for modeling multi-species data (e.g., fish assemblage occurrence) while accounting for species co-occurrences and abiotic environmental variables. For example, joint species distribution models (JSDMs; Clark et al., 2014; Ovaskainen et al., 2017; Wilkinson et al., 2019) that account for residual dependencies between species have seen rapid development in the past several years and are more commonly being applied to freshwater fisheries data (Inoue et al., 2017; Wagner et al., 2020). JSDMs capture species dependencies within a covariance matrix, after accounting for the effects of abiotic predictor variables included in the model, and leveraging this information can lead to improved predictive performance (e.g., Clark et al. 2014; Ovaskainen et al. 2017; Tikhonov et al., 2017; Vallé et al., 2023). Similarly to single-species models, using landscape-derived abiotic predictors within a multi-species framework assumes that they will act as reasonable proxies for instream habitat. However, in the case where these landscape-derived predictors are not sufficient proxies, JSDMs (and other multi-species approaches) may capture this fine-scale information through overlapping (or diverging) habitat requirements within the species dependencies. However, JSDMs allow for limited ability to make inferences regarding species associations as these pairwise species dependencies are not directly estimated, instead being inferred from the residual correlations (Poggiato et al., 2021). This also prevents their effect sizes from being directly compared to the abiotic covariates included in the model (Clark et al., 2018; Harris, 2016).

An alternative statistical approach called conditional random fields (CRF; a type of Markov network) simultaneously models both community-level species co-occurrence information and abiotic predictor effects on species distributions, while also allowing for the effects of species co-occurrences to vary along abiotic gradients (i.e., the effect of species co-occurrences in the community is allowed to be context-dependent). Furthermore, CRFs directly estimate the effect sizes of co-occurrences between species. Because these effect sizes are estimated jointly and at the same scale as those for the abiotic factors, they can then be used to calculate relative importance scores between all effect types. This allows for direct comparisons of effect sizes between abiotic factors, species co-occurrences, and their context dependency. These relative importance scores are calculated within each species of the community, allowing for inferences regarding how different effect types vary across the community. However, as with all modeling efforts, the choice of the modeling approach used is dependent on the research questions and available data. If a primary objective is to predict fish occurrence at unsampled locations, then CRF models would not be the preferred option because CRFs cannot predict to unsampled locations where community data are lacking. However, the application of CRF models can still be valuable tools for management and conservation efforts. For example, predictions may be leveraged to identify locations that are more or less susceptible to the spread of an invasive species, or those that are prime candidates for the reintroduction of a locally extirpated species—given both the abiotic habitat and occurrences of other species.

Here, we explore two objectives pertaining to the use of assemblage-level co-occurrence data and a relatively novel statistical methodology. First, we fit and compared the predictive performance of an assemblage-level (i.e., the assemblage represents fishes sampled and not an entire aquatic community) CRF model against a more traditional single-species generalized linear model (GLM). Previous research has suggested that including species co-occurrences can improve predictions of species distributions of freshwater fishes (e.g., through the use of JSDMs; Wagner et al., 2020). Thus, our hypothesis is that including species co-occurrence data within the CRF framework will capture additional information about instream habitat conditions compared with single-species GLMs that rely only on landscape-based predictors, which will lead to better predictions of stream fish occurrence across two regional watersheds. For example, species that share overlapping habitat requirements at a finer scale may appear as positive co-occurrences when such habitat information is not captured in coarser-scale abiotic variables (e.g., catchment land use). Second, within the CRF framework, we investigate the relative importance of the different effect types (i.e., species co-occurrence and environmental predictors) to understand their respective contributions to predicting species distributions. Previous research has suggested that the ecological processes structuring the assemblage of freshwater stream fishes is non-random and is largely governed by abiotic filters (Giam & Olden, 2016). Here, we test to see if this holds true when using data at a relatively large extent and coarse resolution through the use of relative importance measures.

Methods

Fish and landscape data

Fish assemblage data were obtained from three agencies that sample within Pennsylvania, USA: the Pennsylvania Department of Environmental Protection (PADEP), the Pennsylvania Fish and Boat Commission (PFBC), and the Susquehanna River Basin Commission (SRBC). Fish assemblage data were collected using standardized survey protocols over a 20-year period (2000–2020), with the majority of the sample sites (hereafter catchments) occurring within Pennsylvania state lines (Fig. 1). Only fishery surveys that sampled the entire fish assemblage were included in the analysis. A number of different gear types were used during the surveys and varied by agency. Gear type varied in order to effectively sample the fish assemblage given stream-specific characteristics (e.g., stream size), but the majority of samples were collected using electrofishing. Backpack and tow barge electrofishers were used for smaller, wadeable streams and boats for larger streams and rivers. Additional gear types included seine nets, trawling, and SCUBA/snorkeling visual identification. The fish data were summarized as presence–absence, where a species was considered present if it ever occurred at the sample site over the time period of record. We summarize site occurrence over time because the abiotic information used within our analysis was stationary (described below). In other words, we were not able to relate changes in environmental conditions to changes in occurrence over time. We also acknowledge that accounting for differences in sampling efficiency across gears and imperfect detection are important considerations when modeling fisheries catch data (e.g., Arreguín-Sánchez 1996, Ensign et al. 2002, Kennard et al. 2006, Peterson & Rabeni 1995); however, we assumed that summarizing data into presence–absence helped to reduce the influence of varying catchability on our analysis (i.e., the potential influence of varying catchability across gear types on inferences would

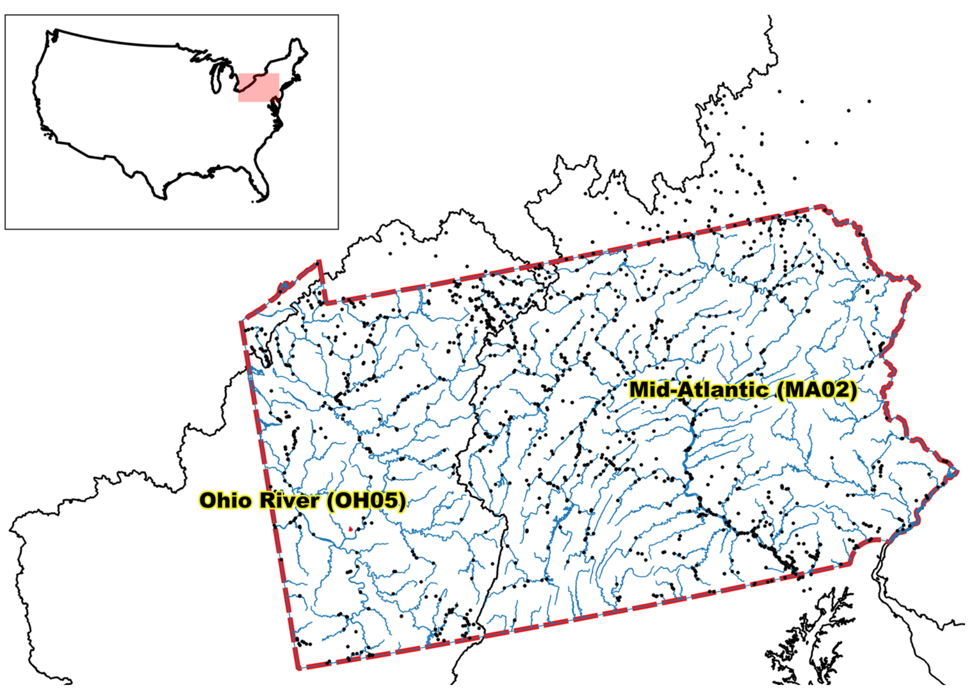

Fig. 1 Map of fish assemblage samples (dots) with HUC2 watersheds boundaries (solid lines with HUC2 code labels) and Pennsylvania, USA border (red dashed line). Blue lines represent major rivers of Pennsylvania. Inset shows location of Pennsylvania within contiguous United States

likely be greater when modeling relative abundance; King et al., 2023).

Catchments were grouped by location within level two hydrologic unit code (HUC; Seaber et al., 1987) regional watersheds to be modeled independently. The regional watersheds that fall within Pennsylvania and included within the study were the Mid-Atlantic (HUC2 = 02) and the Ohio River (HUC2 = 05). The full fish assemblage data were composed of 2222 samples from 1265 catchments, and included over 145 fish species. For each watershed, a species had to occur at 30 or more catchments to ensure a sufficient sample size of occurrences, reducing the full species list to 52 species for the Mid-Atlantic watershed model (hereafter the Mid-Atlantic model) and 46 species for the Ohio River watershed model (hereafter the Ohio River model). All fishes were identified to the species level when possible. However, if species-level identification was not possible then fish were identified at the genus level and grouped accordingly. For example, all species of the sculpin genus *Cottus* were grouped. Sculpins can be difficult to identify at the species level, but are quite common in Pennsylvania headwater streams (Stauffer et al., 2016). Both models included sculpins *Cottus spp.* Two additional genera were removed despite having high enough occurrence: unidentified *Lepomis spp.* in the Mid-Atlantic watershed, and unidentified lampreys from the Ohio River watershed. Unidentified *Lepomis spp.* were removed so that we could retain those identified at the species level rather than grouping them to genus. Unidentified lampreys were removed because much of their associated sample data did not include genus information and lampreys can differ significantly in their ecological niche (i.e., some are parasitic).

EcoSHEDS Northeast Catchment Delineation (NECD) hydrological catchments (Walker et al., 2015) were used as the base spatial unit for the analysis. The NECD catchments are similar to those of the National Hydrography Dataset (NHD) Plus product (Buto & Anderson, 2020), but have improved spatial resolution, with catchment delineation based on the high-resolution NHD flowlines. Sample locations were assigned to a catchment based on each sample site's spatial coordinates, resulting in 817 unique catchments in the Mid-Atlantic model and 412 in the Ohio River model. Most abiotic predictors were obtained from the NECD data product. This spatial dataset includes both local (information about the catchment itself) and upstream (the local and accumulated upstream catchments combined) summaries of each sample site's ecological context. For this analysis, all abiotic variables from this dataset were defined using the upstream summaries in an effort to capture information regarding the network influences on lotic systems. Importantly, as previously mentioned, the NECD data product is stationary and thus provides a single point estimate for each abiotic predictor within each hydrological catchment. In other words, these data do not capture changes in habitat characteristics over the sampled time period.

Abiotic factors included climate information (e.g., annual precipitation [mm]; PRISM 30-year normals; https://prism.oregonstate.edu/normals/), land use (National Land Cover Database; 2006 NLCD; Jin et al., 2013), nitrate deposition (National Atmospheric Deposition Program; https://nadp.slh.wisc.edu/), and land form (Tables 1 and 2). Stream temperature is known to be an important abiotic factor that affects the distribution of freshwater poikilotherms (Buisson et al., 2008; Wehrly et al., 2003), but is not measured across large spatial extents. Therefore, predicted stream water temperature (maximum summer 14-day mean) was sourced from DeWeber and Wagner (2014). All species presence and abiotic sample site data were joined to the NECD catchment using the `sf` package in R (Pebesma, 2018; R Core Team, 2022). Because the available abiotic data were landscape-derived variables (as opposed to instream habitat), we attempted to include as many of the NECD covariates as

Table 1 Summary of abiotic factors within the Mid-Atlantic watershed included in the model of stream fish assemblages. Stream temperature was sourced from DeWeber and Wagner (2014)

Variable	Min	Median (IQR)	Mean	Max
Agriculture (%)	0.00	21.61 (7.84–33.74)	23.57	88.24
Annual precipitation (mm)	883.21	1096.63 (1056.46–1141.43)	1098.40	1349.13
Developed (%)	0.00	5.77 (3.13–9.86)	10.17	99.90
Drainage area (km^2)	0.81	83.71 (16.9–388.95)	3921.82	70704.34
Forest (evergreen) (%)	0.00	3.22 (0.84–5.66)	3.98	36.54
Forest (mixed) (%)	0.00	8.69 (2.42–16.11)	10.79	63.41
Herbaceous (%)	0.00	2.16 (0.1–3.97)	2.73	22.35
Stream temperature (°C)	15.95	21.3 (19.95–22.65)	21.34	27.24
Nitrate deposition (kg/ha)	9.40	13.13 (12.54–13.77)	13.24	17.42
Sandy soil (%)	0.00	0 (0–2.36)	2.74	60.79
Slope (%)	1.37	13.58 (10.21–16.92)	13.72	40.30
Open water (%)	0.00	0.21 (0.02–0.93)	1.05	46.66
Wetland (%)	0.00	0.89 (0.05–2.36)	1.70	17.66

Table 2 Summary of abiotic factors within the Ohio River watershed included in the model of stream fish assemblages. Stream temperature was sourced from DeWeber and Wagner (2014)

Variable	Min	Median (IQR)	Mean	Max
Agriculture (%)	0.00	15.29 (5.67–29.68)	18.41	71.08
Annual precipitation (mm)	946.48	1145.76 (1082.44–1191.31)	1140.04	1410.10
Developed (%)	0.00	4.53 (1.88–7.27)	6.11	89.03
Drainage area (km^2)	0.90	47.32 (11.27–249.25)	1327.72	60871.06
Forest (evergreen) (%)	0.00	2.87 (0.9–4.77)	3.62	34.45
Forest (mixed) (%)	0.00	6.07 (0.98–11.61)	7.49	48.79
Herbaceous (%)	0.00	3.31 (0.98–6.25)	4.18	25.64
Stream temperature (°C)	15.88	20.17 (18.63–22.14)	20.33	26.26
Nitrate deposition (kg/ha)	9.31	13.08 (12.39–13.62)	12.98	16.00
Sandy soil (%)	0.00	1.75 (0.01–8.48)	6.36	56.47
Slope (%)	3.15	13.15 (7.81–16.67)	12.96	28.79
Open water (%)	0.00	0.07 (0–0.61)	0.55	10.15
Wetland (%)	0.00	0.67 (0.02–2.22)	1.89	27.10

possible. Variable selection was based on Pearson's pairwise correlations between variables such that no two variables had an estimated correlation greater than |0.6|. This variable selection process was performed separately within each watershed region, and the final set of covariates included in the model fitting process were those that met this criteria for both watersheds. All abiotic predictor variables shown in Tables 1 and 2 were standardized to have a mean zero and standard deviation of one. Predicted stream temperature was the only variable included in the model with missing values ($n = 60$ and $n = 7$ for the Mid-Atlantic and Ohio River models, respectively). Missing values for abiotic variables were set to the standardized mean.

Statistical analysis

We used conditional random fields (CRFs) to quantify the effect size and relative importance of species co-occurrences and abiotic factors structuring fish assemblages at the landscape scale and to determine if species co-occurrence patterns were context-dependent (i.e., vary across environmental gradients). CRFs are undirected graphical network models that allow for direct inferences of both species co-occurrence and abiotic factors structuring species distributions using presence–absence data (Harris, 2016). Briefly, the CRF model estimates the log-odds of observing species j given the occurrence of species k and covariate x by:

$$\log\left(\frac{P(y_j = 1|y_{\setminus j}, x)}{1 - P(y_j = 1|y_{\setminus j}, x)}\right) = \alpha_{j0} + \beta_j^T x + \sum_{k:k\neq j}(\alpha_{jk0} + \beta_{jk}^T x)y_k \quad (1)$$

where $\mathbf{y}_j$ is a vector of species observations (1 if present, 0 otherwise), $\mathbf{y}_k$ is a vector of observations for all other species not j, α_{j0} is the species-level intercept, β_j^T (superscript T means transposed) is the coefficient of covariate x on species j's occurrence probability, and α_{jk0} and β_{jk}^T represent the coefficients associated with species k's main effect and interaction effect with covariate x, respectively (see Clark et al. 2018, Harris 2016, for more details regarding Markov and conditional random fields). The full analysis was performed using the MRFcov R package (Clark et al., 2018; R Core Team, 2022). The model was run as a spatially-explicit model by providing coordinates (catchment centroid) of each sample site. This allowed the model to account for possible spatial autocorrelation via Gaussian Process spatial regression splines. Within each bootstrap iteration ($n = 400$), the model is fit across all species. Each model was fit separately for each watershed region. This resulted in models with 52 and 46 species (co-occurrence effects), 12 abiotic factors, and 624 and 552 interaction (species co-occurrence × abiotic) effects for the Mid-Atlantic and Ohio River models, respectively. To prevent overfitting with so many potential coefficients, the algorithm uses LASSO (least absolute shrinkage and selection operator) penalization to regularize the regressions. This regularization process is optimized automatically through the functions from MRFcov R package within each species' model fit and forces a number of coefficients to zero. The coefficient's estimates across all bootstrap iterations were then summarized into the mean coefficient value and 90% confidence intervals (i.e., 5% and 95% quantiles).

Predictive performance

We assessed the predictive performance of the CRF models against the single-species models using a 100-repetition 5-fold cross-validation. This was done by randomly partitioning each dataset into five different subsets. Four of these samples were used to fit the models and the fifth subset was used to test the predictive performance of that fitted model.

This was done five times so that each subset was used as testing data. This entire process was then repeated 100 times. For this portion of the analysis, each CRF model was fit using the `MRFcov_spatial` function from MRFcov R package. This function fits a single CRF model with spatial splines to account for spatial autocorrelation. Single-species models were fit using the same datasets and the model structure was developed to closely mirror that of the CRF. We fit generalized linear models (GLMs; logistic regression) with the `glm` function from the stats package in R using a logit link. Similar to the CRF models, the GLMs also included spatial splines to account for potential spatial autocorrelation which were calculated using the `smooth.construct2` function from the mgcv package in R. These approaches were chosen because they most closely mirror the algorithm used within the CRF model fitting process. A key difference here is that we did not use LASSO regularization for the single-species models. This was reasonable because we were only interested in using the GLM for prediction comparisons and not for making inferences about estimated coefficients.

To compare predictive performance between the multi-species CRF and the single-species GLM, the sensitivity (the ability of the model to correctly predict species presence), the estimated receiver operating characteristic (ROC) curve, and Youden index were calculated to determine the overall model's classification ability and the optimal threshold (the predicted probability cutoff level at which the model classifies a species as present) for classification (Khan & Brandenburger, 2020). The Youden index and ROC curve were then used to calculate the area under the curve (AUC) value. AUC values measure discriminatory capacity of the model where a higher AUC value represents better predictive performance (Jiménez-Valverde, 2012). These metrics were calculated for each species within each fold and repetition using the ROCit R package (Khan & Brandenburger, 2020), then summarized by their mean. The AUC was also summarized by the 5th and 95th percentiles (i.e., 90% bootstraped confidence intervals). The species-level CRF measures were then compared against their respective GLM measures to compare predictive performance of the two approaches and to determine if including assemblage co-occurrence data did in fact improve predictive performance. Models with AUC values > 0.9 are considered highly accurate and most useful for interpretation and prediction, and models with AUC values between 0.7 and 0.9 are considered moderately accurate (Manel et al., 2001). Note that the closer an AUC score is to 0.5, the closer the model is to determining presence/absence via "coin-flip."

Key coefficients

From the full multi-species CRF, estimated coefficients (hereafter referred to as key coefficients because they were the important estimated effects not forced to zero through regularization) for each species were used to make inferences about the relative importance of species co-occurrences and abiotic factors governing species distributions within each watershed. The full assemblage dataset was fit under a bootstrapped conditional random fields framework using the `bootstrapMRF` function within the MRFcov R package. This function also models spatial autocorrelation through splines and the bootstrapping allows for the estimation of uncertainty in the parameter estimates. Key coefficients were determined by a relative importance score (calculated $B^2/\Sigma B^2$ from all bootstrapped models across iterations, where the vector B is the regression coefficient for predictor variables) occurring above a default threshold (> 0.01). We focus our inferences on the use of these key coefficients rather than a significance level to focus on the relative strength of the effect size of the estimated associations. Key coefficients were also grouped according to effect type: co-occurrence main effect, abiotic main effect, and biotic-abiotic interaction effect (context-dependent species co-occurrences). Coefficients were ranked by their relative importance score within each species. The coefficient with the highest relative importance score (i.e., strongest absolute effect size) was ranked 1, and so on for each key coefficient. This was done to simplify comparisons of effect type's relative importance across all species. Network graphs were created using the igraph R package (Csardi & Nepusz, 2006; R Core Team, 2022).

Results

Across both watersheds, the 70 unique species of fish were from 14 families and included species such as brook trout—a coldwater stenotherm found in headwater streams (drainage area ranged from < 1km^2 to > 70,000 km^2)—to warmwater fishes such as bluegill *Lepomis macrochirus* and channel catfish *Ictalurus punctatus* found in larger, low gradient rivers (Appendix Tables 4 and 5). Sampled stream and river catchments varied in abiotic characteristics across watersheds and land form. For example, agriculture land use in the upstream catchment ranged from 0 to 71% in the Ohio River watershed (median = 22%) and from 0 to 88% (median = 15%) in the Mid-Atlantic watershed (Tables 1 and 2). However, percent slope ranged from 3 to 29% (median = 13%) in the Ohio River watershed and from 1 to 40% (median = 14%) in the Mid-Atlantic watershed.

Predictive performance

The CRF model was moderately accurate classifying occurrence for both watersheds. The mean AUC value (90% bootstrapped confidence interval) across all species, folds, and repetitions was 0.87 (0.65, 0.97) for the Mid-Atlantic

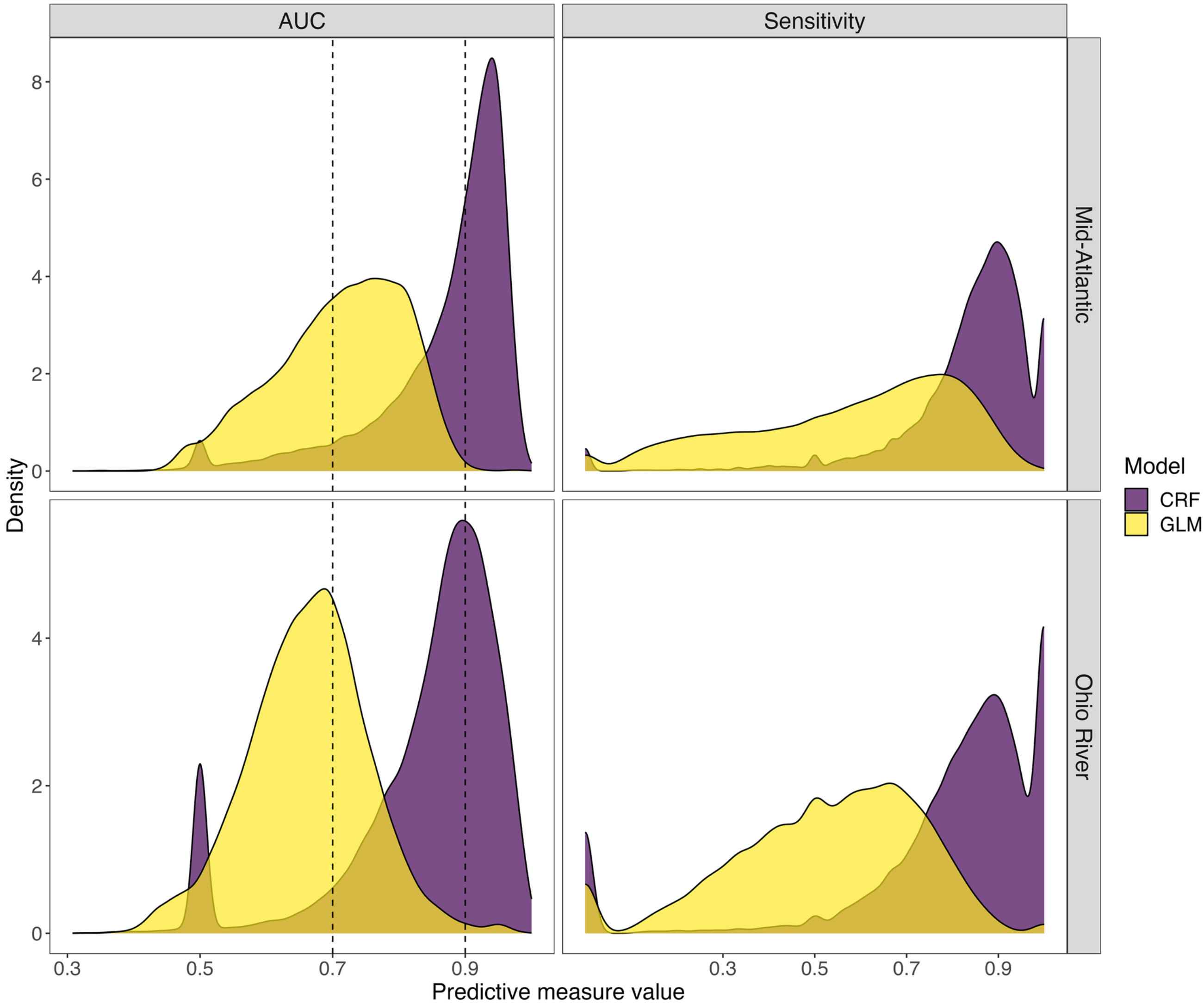

Fig. 2 Distribution of predictive performance metrics AUC and sensitivity across 70 freshwater fish species for a multi-species model that included both species co-occurrences and abiotic predictor variables (CRF) and single-species models that only included abiotic predictor variables (GLM) for two regional watersheds (Mid-Atlantic and Ohio River). Vertical dashed lines for AUC at 0.7 and 0.9 show cutoffs for low accuracy (AUC < 0.7), moderate accuracy (0.7 < AUC < 0.9), and high accuracy (AUC > 0.9) models

watershed and 0.84 (0.50, 0.97) for the Ohio River watershed. Furthermore, 53.6% of cross-validated CRF models had an AUC higher than 0.9 in the Mid-Atlantic watershed and 35.9% of cross-validated CRF models had an AUC higher than 0.9 in the Ohio River watershed. This accuracy is also reflected in the separate measures of predicting species presence (sensitivity = 0.84 and 0.79 for the Mid-Atlantic and Ohio River models, respectively).

Predictive performance of the CRF model varied among individual species across watersheds (Appendix Tables 4 and 5), but predicted occurrence with moderate to high accuracy for almost all species across both watersheds. The CRF model was highly accurate in predicting occurrence for more species in the Mid-Atlantic watershed than the Ohio River watershed. There were 27 (51.9%) species with an AUC value greater than 0.9 from the Mid-Atlantic CRF model, and 11 species (23.9%) from the Ohio River CRF model. There were only 4 (7.7%) species with an AUC value less than 0.7 from the Mid-Atlantic CRF model, and three species (6.5%) from the Ohio River CRF model (Fig. 2). In the Mid-Atlantic model, predictive performance for the CRF model was highest for *Sander vitreus* and *Clinostomus funduloides* (AUC = 0.96) and lowest for *Clinostomus elongatus* (AUC = 0.62). The full CRF model for the Ohio River model performed best for *Etheostoma zonale* (AUC = 0.96) and had the lowest AUC value for *Lota lota* (AUC = 0.50).

Many of the single-species GLMs performed moderately well as measured by AUC (Fig. 2; Appendix Table 4 and 5), though none of the cross-validated GLMs averaged higher than 0.9 for either watershed. Collectively, the mean AUC value (90% bootstrapped confidence interval) across all species, folds, and repetitions was 0.71 (0.54, 0.85) for the Mid-Atlantic watershed and 0.66 (0.51, 0.81) for the Ohio River watershed. The GLM models were moderately accurate, with an AUC value between 0.7 and 0.9, for 34 (65.4%) species from the Mid-Atlantic and 13 species (39.4%) from the Ohio River CRF model (Fig. 2). The remaining 18 (34.6%) and 33 (71.7%) species from the Mid-Atlantic and Ohio River watersheds, respectively, were predicted with low accuracy. Across the GLMs for the Mid-Atlantic species, the highest AUC value was 0.81 for *Micropterus dolomieu*, and the lowest was 0.55 for *Notemigonus crysoleucas*. Across the GLMs for the Ohio River watershed, the highest AUC was 0.78 for *Erimystax dissimilis*, and the lowest AUC was 0.52 for *Lepomis gibbosus*.

When comparing the predictive performance between models within the Mid-Atlantic, the multi-species CRF generally outperformed the single-species GLMs as measured by AUC and sensitivity where the GLM only did as well as the CRF for some species, but never better (Fig. 2). The CRF model had higher AUC estimates than the GLM for all species (Fig. 3), with 37 (71.2%) significantly different (i.e., non-overlapping) 90% confidence intervals (Table 4). Again, this is further reflected in the sensitivity measures, where the CRF also had higher values for predicting species presence for all species. The largest difference in estimated AUC values in favor of the CRF model in the Mid-Atlantic watershed was for *Carpiodes cyprinus* where the CRF model was highly accurate and the GLM predicted with low to moderate accuracy (CRF AUC = 0.93 [0.88, 0.96]; GLM AUC = 0.63, [0.54, 0.73]). In the Mid-Atlantic, the lowest difference in AUC point estimates in favor of the CRF was for *Clinostomus elongatus*, but both models were low accuracy with wide and overlapping 90% confidence intervals (CRF AUC = 0.62 [0.49, 0.86]; GLM AUC = 0.59 [0.47, 0.73]).

Within the Ohio River watershed, the CRF model again generally outperformed the single-species GLMs as measured by AUC and sensitivity, where the GLM only did as well as the CRF for some species, but never better. The CRF model had higher AUC estimates for all but two species and 25 (54.3%) had significantly different 90% confidence intervals. The CRF model had higher sensitivity scores than the GLMs for all but one species (Fig. 4). In the Ohio River watershed, the largest difference in estimated AUC values in favor of the CRF model was for *Notropis rubellus* (CRF AUC = 0.90 [0.84, 0.96]; GLM AUC = 0.61, [0.50, 0.72]), and the biggest difference in favor of the GLM was for *Lota lota* (GLM AUC = 0.69 [0.50, 0.87]; CRF AUC = 0.50, [0.50, 0.50]) suggesting the CRF model was no better than a coin flip. Despite having a higher point estimate for AUC, the wide confident intervals suggests that we cannot claim that the GLM did significantly better at predicting *Lota lota* occurrence.

CRF modeling: key coefficient summaries

Across all species, there were 712 key coefficients identified within the Mid-Atlantic model and 583 key coefficients identified within the Ohio River model (Appendix Tables 6 and 7 provide the full list of key coefficients for every species). Across both watersheds, the species co-occurrence main effects were the most frequently identified key coefficients (410 and 319 for the Mid-Atlantic and Ohio River, respectively; Figs. 5b and 6b), followed by context-dependent co-occurrence effects (297 and 261 for the Mid-Atlantic and Ohio River, respectively) and then abiotic main effects (5 and 3 for the Mid-Atlantic and Ohio River, respectively; Figs. 5a and 6a). When comparing these against their expected counts based on random chance (i.e., their respective percentages of total coefficients included in the model fitting multiplied by their respective total key coefficients), we saw that species co-occurrence effects (54 and 44 expected for the Mid-Atlantic and Ohio River models, respectively) over-performed (i.e., they were identified as key coefficients more often than expected), and abiotic main effects (12 and 11 expected for the Mid-Atlantic and Ohio River models, respectively) and interaction effects (646 and 528 expected for the Mid-Atlantic and Ohio River models, respectively) under-performed (i.e., identified as key coefficients less often than expected).

Both watershed models estimated a complex network of species associations. In the Mid-Atlantic model, all 52 species appeared as a key coefficient for predicting the occurrence of at least one other species, whereas 45 of the 46 species included in the Ohio River model appeared as a key coefficient for predicting the occurrence of at least one other species (Figs. 5b and 6b). The burbot *Lota lota* was the only species that did not appear as a key coefficient in the Ohio River model. The species most commonly identified as key coefficients were the creek chub *Semotilus atromaculatus* in the Mid-Atlantic and western blacknose dace *Rhinichthys obtusus* in the Ohio River watershed. Conversely, the rosyside dace *Clinostomus funduloides* in the Mid-Atlantic and brook trout *Salvelinus fontinalis* in the Ohio River watershed were species important for predicting the occurrence of relatively few other species (Fig. 6b). For species that occurred across both models, their effects as key coefficients varied. For example, the brook trout appeared as a key coefficient only once (as a negative main effect for *Pimephales notatus*) in the Ohio River model, but it was identified as a key coefficient 10 times (5 main effects and 5 interaction effect) in the Mid-Atlantic model. Species co-occurrences varied in their effect type across watersheds. In the Mid-Atlantic model, the fallfish *Semotilus corporalis* was the

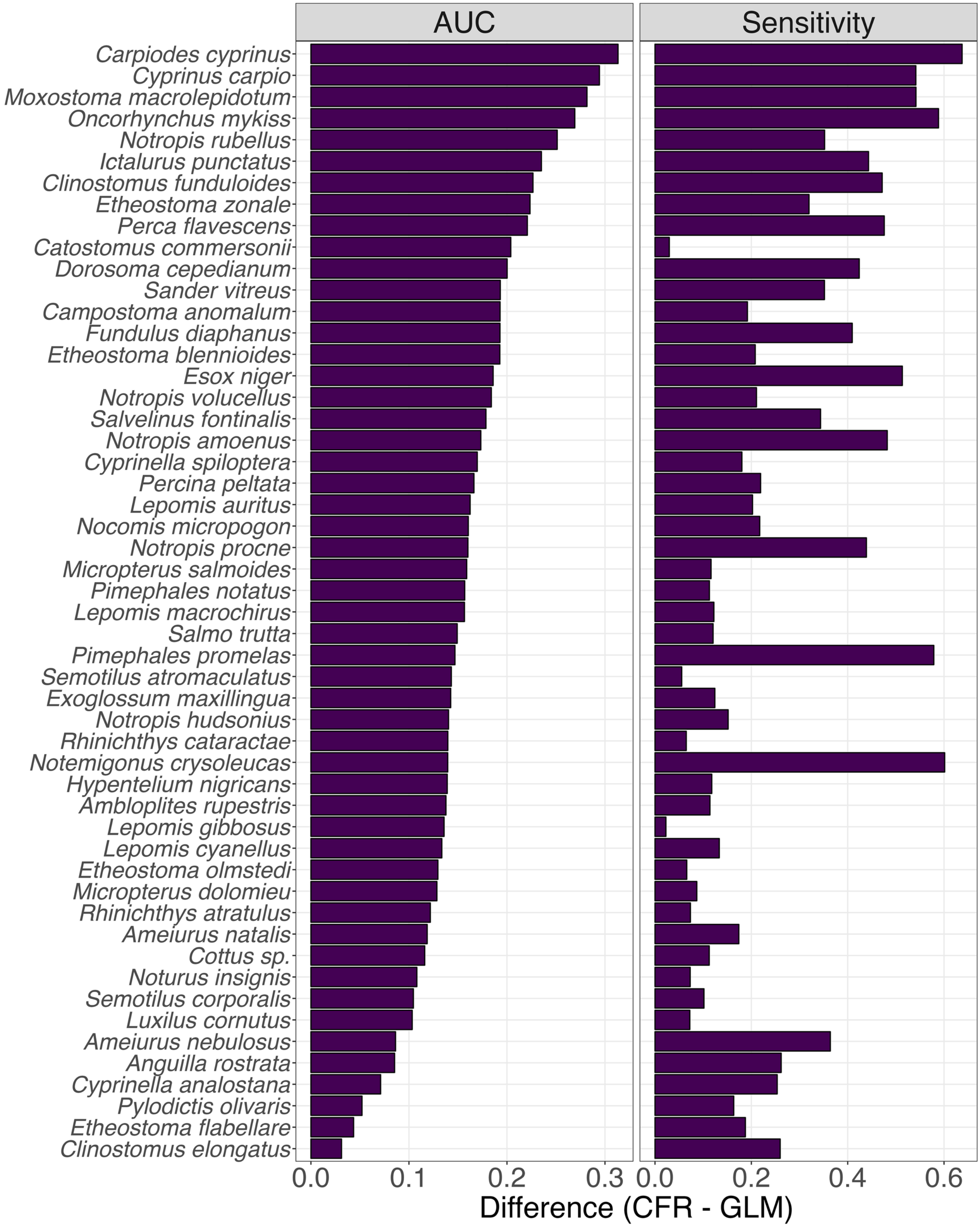

Fig. 3 Difference between the multi-species conditional random fields model (CRF) and the single-species generalized linear model (GLM) of the mean value of predictive performance metrics from a fivefold cross-validation repeated 100 times for the Mid-Atlantic watershed

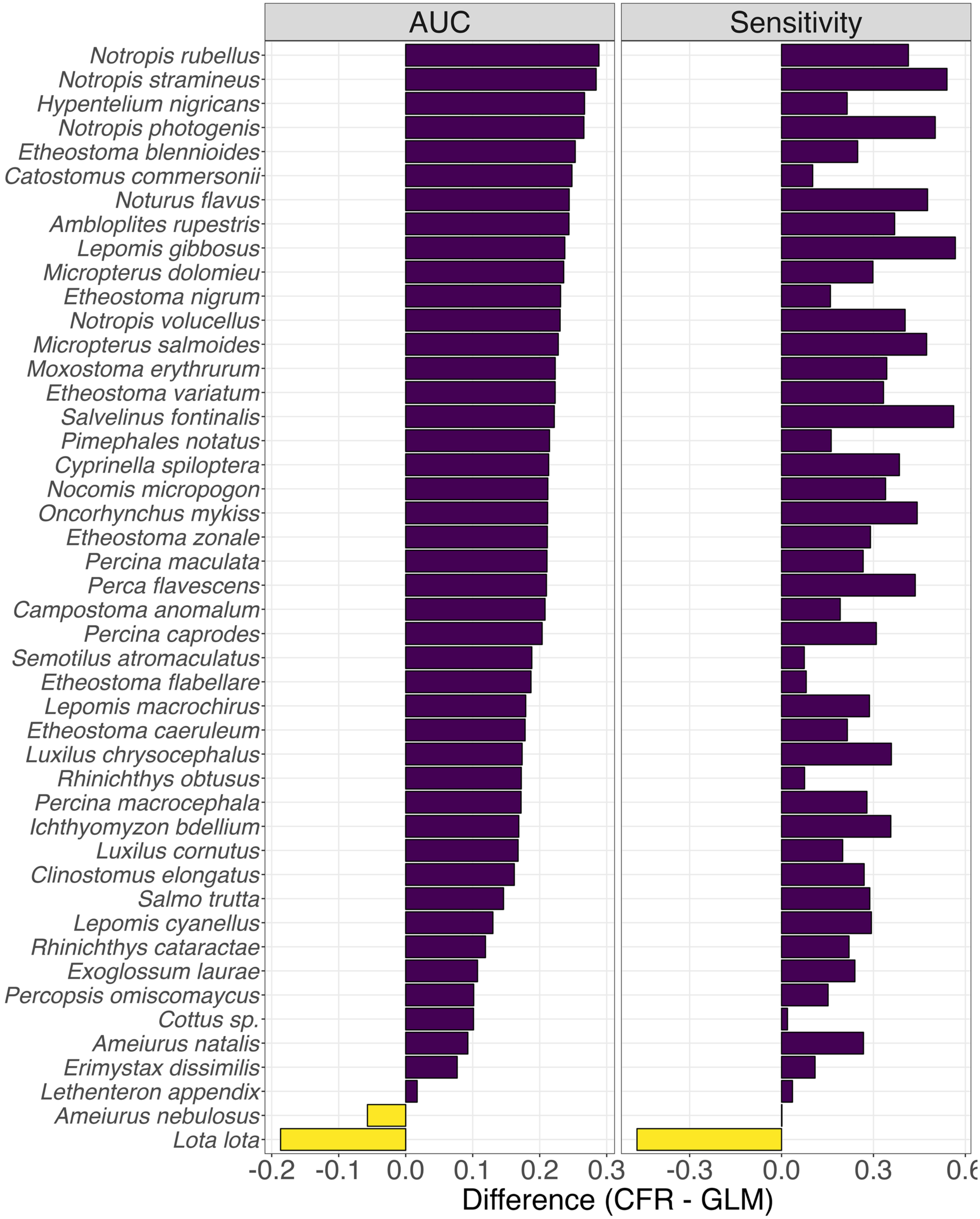

Fig. 4 Difference between the multi-species conditional random fields model (CRF) and the single-species generalized linear model (GLM) of the mean value of predictive performance metrics from a fivefold cross-validation repeated 100 times for the Ohio River watershed

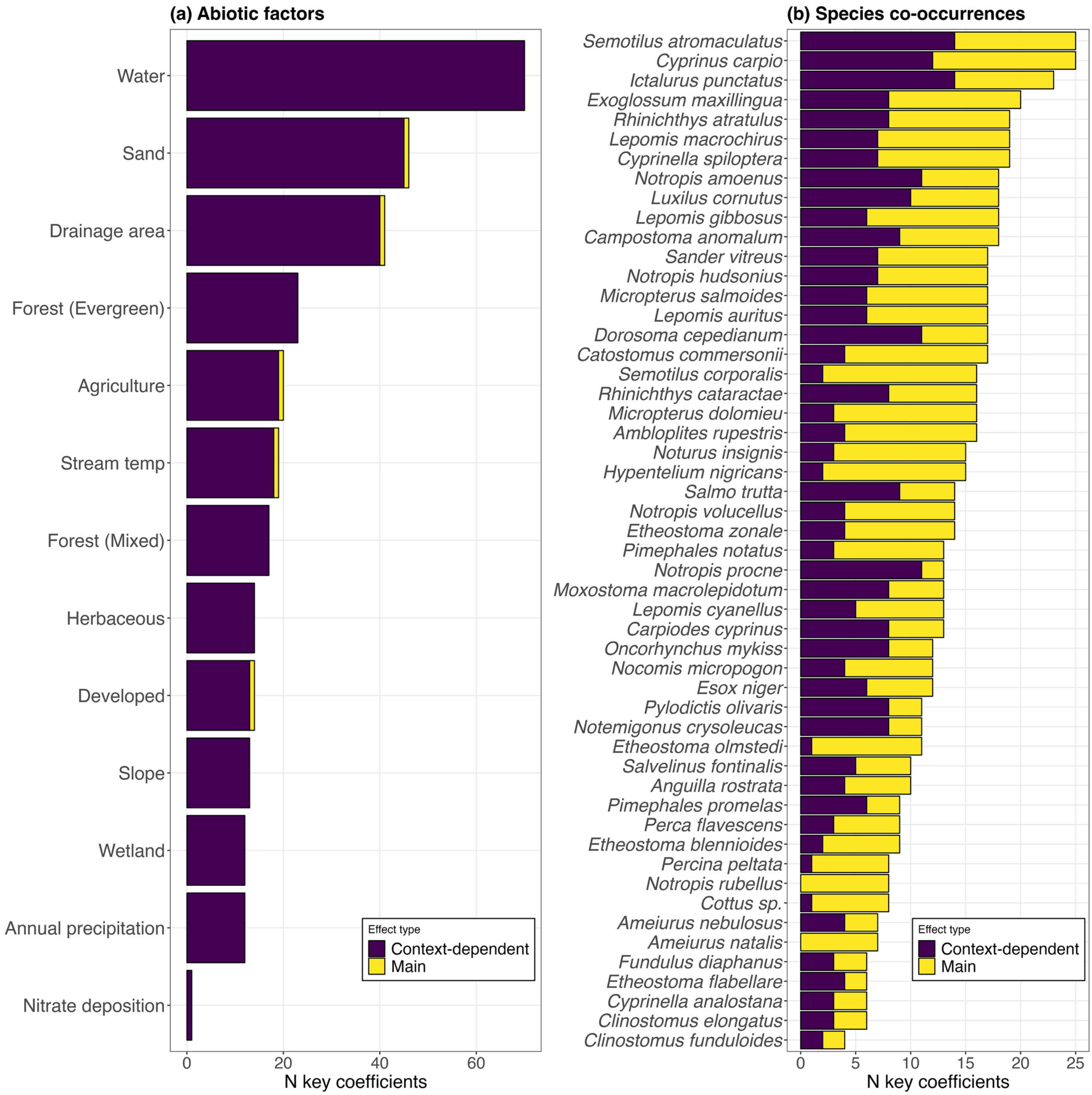

Fig. 5 Frequency of **a** abiotic factors and **b** species co-occurrences appearing as key coefficients from the conditional random fields model in the Mid-Atlantic. Fill color represents proportion of effect type for each factor

most frequently occurring species co-occurrence effect ($n =$ 14), whereas the greenside darter *Etheostoma blennioides* and the northern hogsucker *Hypentelium nigricans* tied for the most frequently occurring species co-occurrence effect ($n = 15$) in the Ohio River model. Additionally, the most frequently identified species co-occurrences among key context-dependent effects were the channel catfish *Ictalurus punctatus* and creek chub ($n = 14$) in the Mid-Atlantic model and the common shiner *Luxilus cornutus* ($n = 14$) in the Ohio River model.

Abiotic main effects were infrequently identified across both models. In the Mid-Atlantic model, there were five abiotic main effects, one each for sandy soil cover, developed land cover, drainage area size, predicted stream temperature, and agricultural land cover. In the Ohio River model, only three abiotic main effects were identified, one each for

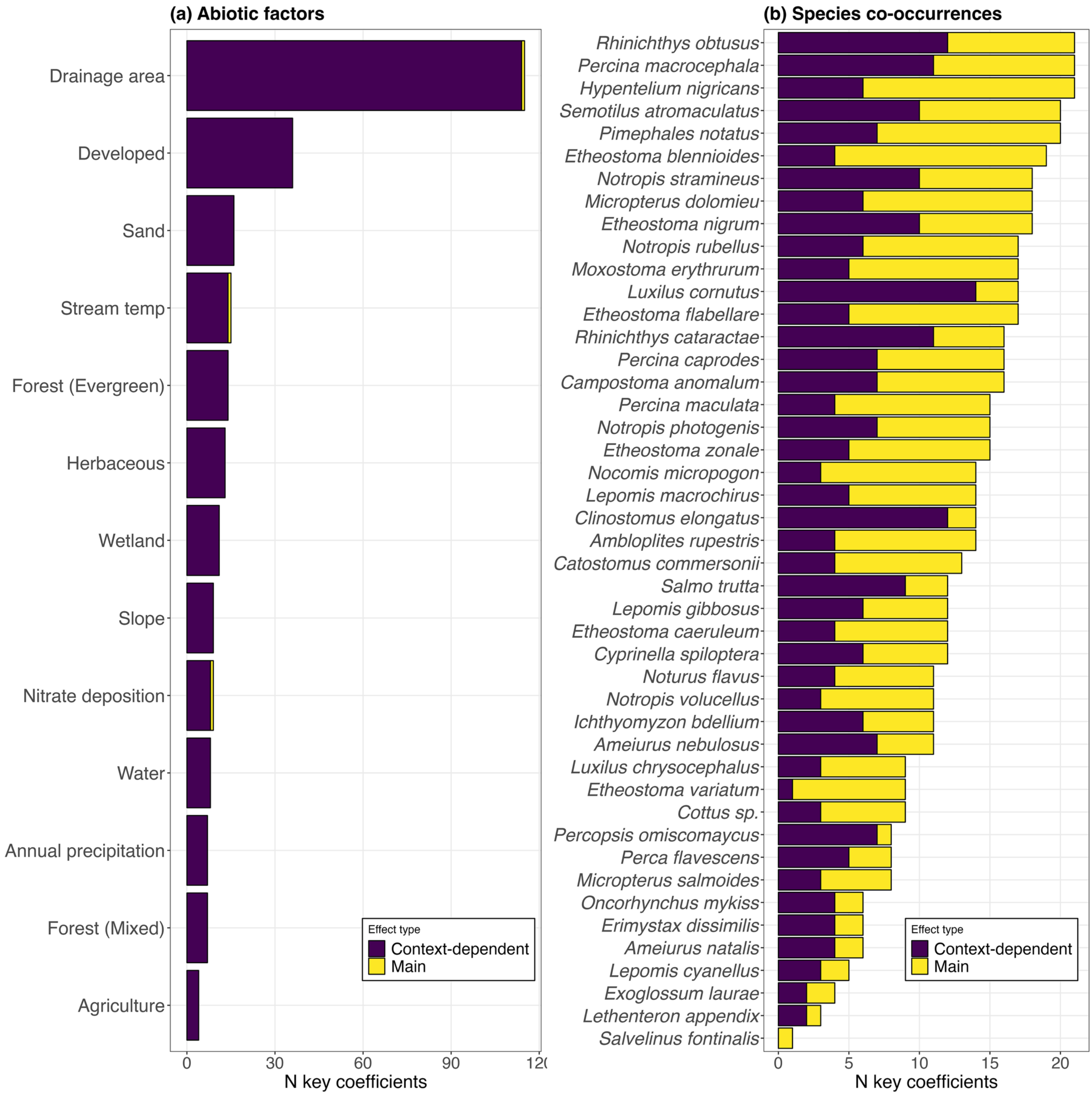

Fig. 6 Frequency of **a** abiotic factors and **b** species co-occurrences appearing as key coefficients from the conditional random fields model in the Ohio River watershed. Fill color represents proportion of effect type for each factor

predicted stream temperature, drainage area size, and nitrate deposition. The total frequency of abiotic factors appearing as key coefficients (main and interaction effects combined) varied by watershed. In the Mid-Atlantic model, the amount of open water was the most frequently identified abiotic factor, whereas drainage area size was the most frequently identified abiotic key coefficient in the Ohio River model. Across both models, drainage area size and sandy soil abiotic factors were in the top three most frequently identified abiotic key coefficients. Abiotic factors positively associated with species occurrence varied across watersheds (Appendix Tables 6 and 6). Across both watershed models, the proportion of all abiotic factors (main and interaction effects) with positive coefficients was approximately 25%. This suggests that species most often had negative associations (i.e., decreased probability of occurrence) with the included abiotic factors. In the Mid-Atlantic model, predicted stream temperature had the highest proportion (68.4%) of positive

Table 3 Summary of key coefficients for the fathead minnow within the Mid-Atlantic watershed

Variable	Mean coefficient value	Relative importance
Drainage area × *Semotilus atromaculatus*	− 1.18 (− 3.30, 0.00)	0.266
Salvelinus fontinalis	0.96 (0.00, 2.14)	0.175
Open water × *Ambloplites rupestris*	− 0.78 (− 3.75, 0.00)	0.117
Semotilus atromaculatus	0.64 (0.10, 1.42)	0.077
Sandy soil × *Notemigonus crysoleucas*	− 0.56 (− 2.43, 0.00)	0.060
Open water × *Salvelinus fontinalis*	0.39 (0.00, 2.03)	0.029
Wetland × *Oncorhynchus mykiss*	0.34 (0.00, 1.83)	0.022
Open water × *Luxilus cornutus*	− 0.32 (− 1.76, 0.00)	0.019
Open water × *Rhinichthys cataractae*	0.28 (0.00, 2.08)	0.015
Open water × *Notemigonus crysoleucas*	− 0.28 (− 1.92, 0.00)	0.015
Sandy soil × *Salvelinus fontinalis*	− 0.28 (− 1.16, 0.00)	0.015
Open water × *Semotilus atromaculatus*	− 0.26 (− 1.27, 0.00)	0.013
Agriculture × *Notemigonus crysoleucas*	0.26 (0.00, 1.14)	0.012
Notemigonus crysoleucas	0.24 (0.00, 0.92)	0.011

coefficient values, corresponding to a number of warmwater species within the region. Whereas, the Ohio River model slope had the highest proportion (100%) of positive values, suggesting a number of species used high gradient stream habitat.

The effect of co-occurrences on predicting species occurrence also varied along environmental gradients (i.e., context dependency) for 51 (98%) and 46 (100%) species in the Mid-Atlantic and Ohio River watersheds, respectively. This suggests that species co-occurrence effects varied along abiotic gradients and that this was important for almost every

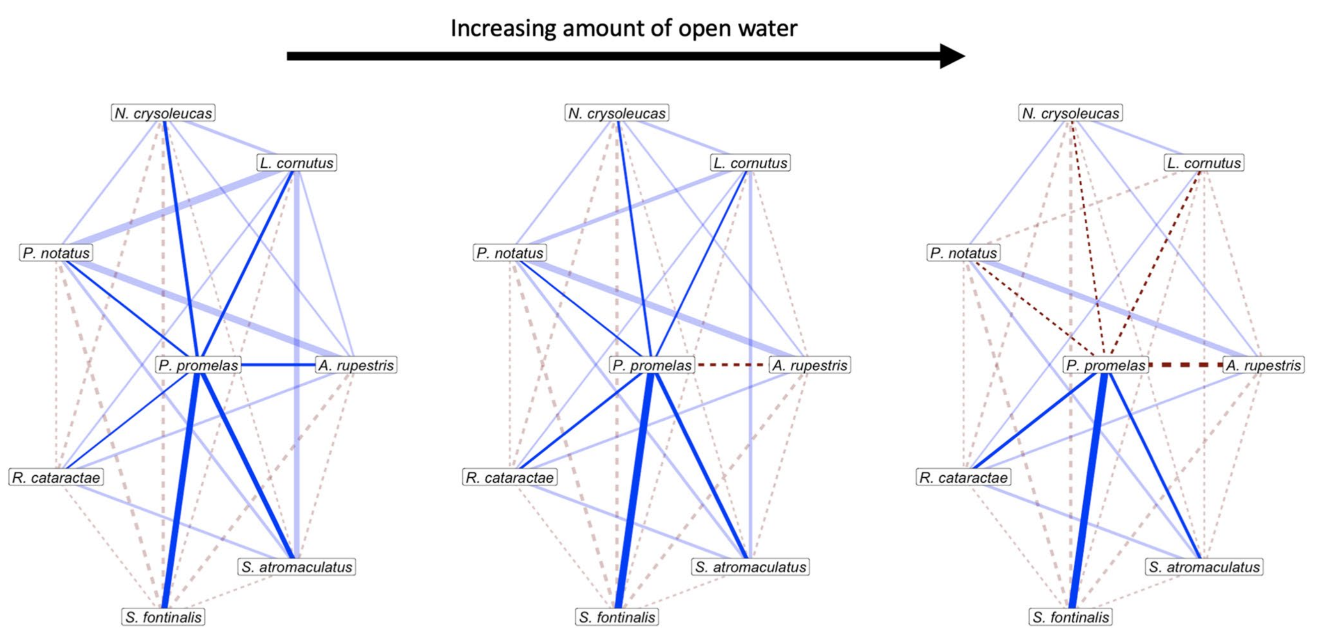

Fig. 7 An illustration of context dependency for the effects of species co-occurrences when predicting the occurrence of the fathead minnow *Pimephales promelas*. The co-occurrence network for fathead minnow (center node) is plotted along amount of open water gradient within the Mid-Atlantic watershed. The graphs to the left, middle, and right represent the estimated network structure at the minimum, midpoint, and maximum levels of open water where fathead minnows were sampled, respectively. Dashed-red and solid-blue lines represent negative and positive estimated co-occurrence patterns, respectively, with line weight representing relative strength of that association. Graphs are filtered to seven species with strongest overall estimated co-occurrence effects with the fathead minnow. Co-occurrence patterns between those seven species are transparent to emphasize co-occurrences that would be used to predict the occurrence of the fathead minnow

species. For example, predicting the occurrence of the fathead minnow *Pimephales promelas* in the Mid-Atlantic watershed estimated 11 context-dependent co-occurrences (Table 3). Figure 7 visualizes the context dependency of species co-occurrences as the amount of open water increases within the sampled catchment. Here, we see a number of species co-occurrences shift from a positive effect to a negative effect on predicting the occurrence of fathead minnow as the amount of open water increases. Of the seven species co-occurrences with the fathead minnow shown, four had positive associations at low levels of open water that became negative at high levels of open water. This suggests that in areas with less water, such as first order streams, it would be expected to find these species co-occurring and in areas with more water, such as higher order streams and large rivers, it would no longer be expected to find the fathead minnow co-occurring with those four species.

Discussion

We found that using fish assemblage co-occurrence data can improve predictions of stream fish distributions across regional watersheds compared to relying on remotely sensed landscape-derived environmental data alone. Here, we used a novel CRF modeling framework which offers a number of aforementioned advantages. Although there was variability in the magnitude of the differences among species and watersheds, the predictive performance (as measured by AUC) was consistently higher for the CRF models compared to landscape-based GLMs. Across both the Mid-Atlantic and Ohio River watersheds, the abiotic-only GLMs never outperformed the CRF models in predicting species occurrence. This aligns with previous efforts to model stream fish occurrence with assemblage data using JSDMs (Inoue et al., 2017; Rodríguez et al., 2021; Wagner et al., 2020). For example, Wagner et al. (2020) showed that including fish assemblage data through a JSDM framework improved conditional predictions of species occurrences for fishes in both stream and lake habitats. Their modeling efforts also showed that their abiotic variables performed poorly in predicting species occurrence. Similarly, Inoue et al. (2017) showed that when jointly modeling freshwater mussels and fishes, the residual correlations (i.e., species dependencies) were prevalent among fishes, whereas mussel occurrences were exclusively explained by abiotic factors. These studies, along with our analysis, suggest that modeling stream fish distributions with only remotely sensed abiotic factors, such as land cover, may lead to spurious inferences and incorporating the additional information provided by species co-occurrences can often improve predictions.

Scale dependency has been shown to play a significant role in species distribution models (Geheber & Geheber, 2016; König et al., 2021). Our study, like most regional modeling efforts, used landscape-scale abiotic factors (i.e., land use and cover) instead of habitat measurements at the local or micro-habitat scale. This resulted in our models relying heavily on species co-occurrences to accurately predict occurrence. The high frequency of important species co-occurrences is likely due to their ability to capture overlapping habitat requirements at a finer scale. For example, stream flow is well-known to be an important abiotic filter for stream fish (e.g., McManamay & Frimpong, 2015; Poff & Allan, 1995; Van Vliet et al. 2013). However, flow requires physical sampling or modeling and may not be readily available for landscape-scale modeling efforts. Thus, we relied on other factors to act as proxies for flow, such as drainage area size or the amount of open water. We did see both drainage area and the amount of open water appear among the most frequently estimated abiotic key coefficients, though rarely estimated as main effects. Instead, they were estimated as context-dependent co-occurrences, furthering the notion that the scale of our abiotic data was relatively coarse and not able to represent instream habitat as well as co-occurrence data. Our finding that species co-occurrences were relatively more important effect types within our models does not suggest that abiotic factors are unimportant filters for structuring stream fish assemblages. Instead, the assumption that our remotely sensed landscape, particularly land cover, variables would act as proxies for true instream habitat was shown to be insufficient for many species of stream fish. Context-dependent species co-occurrences were important effect types for improving the accuracy of species distribution predictions. Although they were estimated as key coefficients less frequently than expected across both models, they were still an important effect type for almost every species. This is important because even with the high relative importance of species co-occurrences, they were often dependent on the environmental conditions present in the stream's catchment. In other words, species co-occurrences were often capturing overlapping fine-scale habitat requirements between pairs of species, but the context dependency of these co-occurrences helped capture where these habitat needs diverged. However, these models are scale-dependent and inferences regarding the ecological processes structuring assemblages from these models should be done with caution.

Recall that a limitation of the CRF is that it cannot predict to unsampled locations. However, CRF predictions at sites with existing assemblage data are still useful for fisheries management as a means to identify high priority sites for potential invasions or re-introductions. A number of species within our study are invasive, such as flathead catfish *Pylodictis olivaris* and the banded darter *Etheostoma zonale* in the Mid-Atlantic watershed. The flathead catfish is of particular concern within the region due to their large size and piscivorous diet (Brown et al., 2005; Smith et al., 2021).

Fisheries managers could use the CRF model to identify areas with higher predicted probabilities of occurrence as habitats at risk for flathead catfish or banded darter range expansion. For flathead catfish, however, the analysis of predictive performance from the cross-validation suggested there was no significant difference between the two modeling approaches—suggesting that either the CRF or GLM could be used for this purpose. That said, the CRF model for flathead catfish does offer additional information that can help generate hypotheses about the potential effects of invasion in new locations. In the CRF analysis, we see a number of important species co-occurrences (both main and context-dependent effects) with comely shiner *Notropis amoenus* and channel catfish *Ictalurus punctatus* (Appendix Table 6). In fact, the frequency of key coefficients with the comely shiner and channel catfish may indicate shared, fine-scale habitat requirements not captured with our environmental data, or perhaps even important potential biotic interactions between these species. Importantly, the key coefficients associated with biotic factors may help identify which species face the biggest threat from this invasive species if they are consistently co-occurring. In contrast to the flathead catfish, we did see large improvements in predicting the occurrence of the banded darter using the CRF in the Mid-Atlantic watershed. Again, the occurrence of this species was not associated with a key abiotic main effect suggesting that the use of remotely sensed landscape data were poor predictors of their occurrence (Appendix Table 6). Such results may provide fisheries managers with an important foundation for developing management and research plans for invasive species. Although just one example, this contextualizes how fisheries managers may use CRF predictions and leverage available fish assemblage data to help inform management and conservation efforts. However, if predicting to unsampled locations is a main research objective, the aforementioned JSDMs may be a more suitable modeling framework. JSDMs still simultaneously model species dependencies with abiotic factors, but do not require a full assemblage sample thus allowing them to make unconditional predictions to unsampled locations.

Estimated coefficients from the CRF model may also be used to generate hypotheses about potential biotic interactions. It is important to first emphasize that species co-occurrences do not directly indicate true interactions and such inferences should be avoided (Poggiato et al., 2021). As previously noted, many, if not most, of these estimated co-occurrences are representative of missing abiotic information and capture overlapping habitat requirements (Zurell et al., 2018). However, they do represent patterns seen across assemblages and biotic interactions play an important role in structuring assemblages (Hutchinson, 1957; Ovaskainen et al., 2017). Therefore, key species co-occurrences can be used, with caution, as a baseline for developing hypotheses regarding potentially important biotic interactions. For example, previous studies have shown that brook trout and brown trout can have negative interactions (Hoxmeier & Dieterman, 2013; Wagner et al., 2013). Our Mid-Atlantic model estimated positive co-occurrences between the two species, an artifact of overlapping habitat requirements. This effect, though, was mediated by a negative, context-dependent co-occurrences with both drainage area and amount of open water. These could potentially represent either diverging habitat requirements or context-dependent biotic interactions.

Finally, we clarify key caveats of the assumptions of our analysis. Firstly, our abiotic predictor variables were limited to remotely sensed landscape data and predicted water temperature. It is possible that other landscape-scale variables, such as lithology, not included in our analysis act as better proxies for instream habitat factors structuring stream fish assemblages. Thus, our results should be interpreted within the context of the predictor variables we included in our models. Previous research has suggested that macroscale variables were sufficient proxies of instream characteristics for modeling stream fish distributions, albeit within a significantly different ecosystem (Brazilian Amazon basin; Frederico et al., 2014). Additionally, a study in the Piedmont Plateau region (which extends into a portion of southeastern Pennsylvania) of Georgia, USA found that reach-scale geomorphology factors were the best predictors for species composition (Walters et al., 2003). That same study, however, found that stream slope was a dominant factor whereas our analysis rarely identified it as a key coefficient across both watersheds. Furthermore, a study by Magalhaes et al. (2002) suggested that coarse-scale factors may better explain fish assemblage variation in Mediterranean streams than micro-habitat factors. Importantly, they hypothesized that these observed patterns could have been due to seasonal changes in water availability, such as summer droughts, an important factor in understanding fish distributions that was not captured within our analysis due to the stationarity of the data, which leads to another important caveat. Both biotic and abiotic data were summarized to single values across time due to the stationarity of the hydrological dataset (NECD). As previously mentioned, this prevented our analysis from relating any potential changes in habitat to potential changes in species occurrence. Thus, it is possible if a stream catchment underwent rapid changes in habitat, such as human development, the environmental predictors modeled may not accurately reflect the habitat conditions experienced by the fishes when sampled. We also made the assumption that summarizing the fish data into presence–absence reduced the effect of varying catchability and imperfect detection and thus did not directly incorporate this into our model. It is possible, however, that some species were still not accurately detected, which could introduce bias into our modeling efforts. That said, our assumption was reasonable for this analysis given the breadth of time and gear types used to detect species presence. Most instances of imperfect detection would likely occur for rarer species which were already not included in the model due to sample size restrictions. However, if accounting for imperfect detection is required for predicting species occurrence, there are multi-species models that incorporate imperfect detection within their framework (e.g., MacKenzie et al. 2004, Rota et al. 2016). Whereas our analysis suggested that using fish assemblage co-occurrence data can improve the predictive performance of stream fish species distribution modeling, there are numerous studies that show abiotic factors can adequately predict stream fish occurrence. Ultimately, the choice of species distribution model should be made with respect to available data and specific research objectives.

One of the most common approaches to modeling species distributions is to use landscape-based (often remotely sensed) habitat data as proxies for instream habitat conditions. We showed that predictions of fish assemblage distributions can be improved for many species by leveraging co-occurrence information. Furthermore, by taking advantage of the CRF methodology, we were able to directly compare the relative importance of species co-occurrences against abiotic factors, while allowing them to interact and be context-dependent. This information can help inform hypotheses about the effects of species range expansions on native fishes and the relative importance species co-occurrences and abiotic drivers of species distributions that can help motivate future research. Future research could also explore how modeling choices related to spatial scale and data resolution affect predictive performance for CRF models of stream fish assemblages. For example, comparing results from this analysis against efforts that include fine-scale, instream habitat data and/or relative abundance data would further improve our understanding of the ecological filters structuring stream fish assemblages. The predictive performance of the CRF, which maintains symmetric relationships between species, could also be compared against alternative multi-species modeling techniques that allow for asymmetric relationships between species.

Author statements

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Appendix

See Tables 4, 5, 6 and 7.

Table 4 List of all species in the Mid-Atlantic watershed with number of occurrences (*N*) and mean AUC values (90% bootstrapped confidence interval) from 100-repetition fivefold cross-validation

Species	*N*	CRF	GLM
*Ambloplites rupestris**	376	0.94 (0.91, 0.96)	0.80 (0.73, 0.86)
Ameiurus natalis	174	0.84 (0.78, 0.89)	0.72 (0.59, 0.82)
Ameiurus nebulosus	76	0.65 (0.50, 0.76)	0.56 (0.48, 0.65)
Anguilla rostrata	122	0.81 (0.73, 0.88)	0.72 (0.65, 0.81)
*Campostoma anomalum**	292	0.95 (0.92, 0.97)	0.75 (0.70, 0.81)
*Carpiodes cyprinus**	53	0.95 (0.88, 0.99)	0.64 (0.51, 0.78)
*Catostomus commersonii**	592	0.91 (0.87, 0.95)	0.71 (0.61, 0.83)
Clinostomus elongatus	37	0.62 (0.49, 0.86)	0.59 (0.47, 0.72)
*Clinostomus funduloides**	39	0.96 (0.91, 0.99)	0.73 (0.56, 0.88)
*Cottus spp.**	351	0.89 (0.84, 0.92)	0.77 (0.71, 0.83)
Cyprinella analostana	37	0.67 (0.47, 0.87)	0.60 (0.48, 0.74)
*Cyprinus carpio**	107	0.93 (0.88, 0.96)	0.63 (0.54, 0.73)
*Cyprinella spiloptera**	264	0.95 (0.92, 0.97)	0.78 (0.66, 0.88)
Dorosoma cepedianum	72	0.90 (0.81, 0.96)	0.70 (0.56, 0.83)
Esox niger	61	0.80 (0.69, 0.90)	0.61 (0.51, 0.72)
*Etheostoma blennioides**	295	0.95 (0.92, 0.97)	0.75 (0.70, 0.82)
Etheostoma flabellare	83	0.77 (0.61, 0.90)	0.73 (0.63, 0.83)
*Etheostoma olmstedi**	532	0.91 (0.88, 0.95)	0.78 (0.71, 0.84)
*Etheostoma zonale**	232	0.95 (0.92, 0.97)	0.72 (0.65, 0.80)
*Exoglossum maxillingua**	413	0.94 (0.91, 0.96)	0.79 (0.74, 0.84)
*Fundulus diaphanus**	122	0.90 (0.85, 0.94)	0.70 (0.61, 0.79)
*Hypentelium nigricans**	370	0.94 (0.92, 0.97)	0.81 (0.69, 0.88)
*Ictalurus punctatus**	97	0.94 (0.90, 0.98)	0.71 (0.61, 0.81)
*Lepomis auritus**	254	0.94 (0.91, 0.97)	0.78 (0.66, 0.87)
Lepomis cyanellus	272	0.85 (0.80, 0.90)	0.72 (0.60, 0.83)
*Lepomis gibbosus**	279	0.81 (0.75, 0.86)	0.67 (0.60, 0.74)
*Lepomis macrochirus**	333	0.87 (0.82, 0.91)	0.71 (0.65, 0.77)
*Luxilus cornutus**	312	0.90 (0.86, 0.93)	0.79 (0.73, 0.85)
*Micropterus dolomieu**	392	0.94 (0.91, 0.97)	0.81 (0.73, 0.87)
*Micropterus salmoides**	213	0.84 (0.79, 0.89)	0.68 (0.58, 0.76)
*Moxostoma macrolepidotum**	78	0.94 (0.89, 0.97)	0.66 (0.54, 0.78)
*Nocomis micropogon**	270	0.90 (0.86, 0.94)	0.74 (0.66, 0.85)
Notemigonus crysoleucas	64	0.69 (0.56, 0.80)	0.55 (0.47, 0.65)
Notropis amoenus	66	0.82 (0.72, 0.91)	0.64 (0.53, 0.76)

Table 4 (continued)

Species	*N*	CRF	GLM
*Notropis hudsonius**	232	0.92 (0.88, 0.95)	0.78 (0.66, 0.86)
Notropis procne	59	0.78 (0.65, 0.90)	0.62 (0.52, 0.73)
*Notropis rubellus**	244	0.93 (0.90, 0.95)	0.68 (0.61, 0.75)
*Notropis volucellus**	240	0.95 (0.92, 0.97)	0.76 (0.69, 0.83)
*Noturus insignis**	341	0.91 (0.87, 0.94)	0.80 (0.73, 0.85)
*Oncorhynchus mykiss**	84	0.85 (0.76, 0.92)	0.58 (0.49, 0.67)
*Perca flavescens**	99	0.83 (0.75, 0.91)	0.61 (0.52, 0.72)
*Percina peltata**	238	0.91 (0.88, 0.94)	0.75 (0.63, 0.87)
*Pimephales notatus**	380	0.93 (0.90, 0.96)	0.78 (0.66, 0.86)
Pimephales promelas	56	0.72 (0.58, 0.83)	0.57 (0.47, 0.70)
Pylodictis olivaris	33	0.72 (0.50, 0.98)	0.67 (0.49, 0.85)
*Rhinichthys atratulus**	503	0.93 (0.89, 0.95)	0.80 (0.74, 0.86)
*Rhinichthys cataractae**	471	0.95 (0.93, 0.97)	0.81 (0.69, 0.88)
Salmo trutta	202	0.88 (0.81, 0.93)	0.73 (0.59, 0.83)
*Salvelinus fontinalis**	149	0.90 (0.84, 0.94)	0.72 (0.64, 0.80)
*Sander vitreus**	102	0.96 (0.91, 0.98)	0.76 (0.67, 0.86)
*Semotilus atromaculatus**	424	0.89 (0.85, 0.92)	0.74 (0.68, 0.80)
*Semotilus corporalis**	309	0.91 (0.88, 0.95)	0.81 (0.73, 0.87)

CRF represents AUC estimates for predictive performance of the full multi-species conditional random fields model that incorporates species co-occurrences and allows for them to be context-dependent. GLM represents AUC estimates for single-species logistic regression models that only accounts for the occurrence of a species with respect to the landscape-derived abiotic predictors. Species marked with * represent those with non-overlapping confidence intervals

Table 5 List of all species in the Ohio River watershed with number of occurrences (*N*) and mean AUC values (90% bootstrapped confidence interval) from 100-repetition fivefold cross-validation

Species	*N*	CRF	GLM
*Ambloplites rupestris**	125	0.85 (0.79, 0.92)	0.61 (0.51, 0.71)
Ameiurus natalis	45	0.71 (0.50, 0.89)	0.62 (0.48, 0.76)
Ameiurus nebulosus	33	0.51 (0.37, 0.67)	0.57 (0.43, 0.75)
*Campostoma anomalum**	183	0.89 (0.83, 0.94)	0.68 (0.58, 0.77)
*Catostomus commersonii**	255	0.88 (0.83, 0.94)	0.64 (0.55, 0.72)
Clinostomus elongatus	98	0.83 (0.73, 0.92)	0.67 (0.57, 0.77)
Cottus spp.	267	0.84 (0.76, 0.91)	0.73 (0.64, 0.82)
*Cyprinella spiloptera**	67	0.91 (0.84, 0.97)	0.70 (0.59, 0.81)
Erimystax dissimilis	31	0.86 (0.50, 1.00)	0.78 (0.60, 0.96)
*Etheostoma blennioides**	190	0.95 (0.92, 0.98)	0.70 (0.62, 0.78)
*Etheostoma caeruleum**	166	0.89 (0.83, 0.93)	0.71 (0.62, 0.79)
*Etheostoma flabellare**	222	0.87 (0.82, 0.93)	0.69 (0.61, 0.77)
*Etheostoma nigrum**	205	0.89 (0.84, 0.94)	0.66 (0.58, 0.74)
*Etheostoma variatum**	79	0.95 (0.89, 0.99)	0.72 (0.62, 0.84)
*Etheostoma zonale**	115	0.96 (0.92, 0.98)	0.74 (0.64, 0.83)
Exoglossum laurae	34	0.82 (0.50, 0.96)	0.71 (0.53, 0.88)
*Hypentelium nigricans**	204	0.90 (0.85, 0.95)	0.63 (0.55, 0.72)
Ichthyomyzon bdellium	39	0.89 (0.75, 0.97)	0.72 (0.56, 0.87)
Lepomis cyanellus	62	0.81 (0.68, 0.91)	0.68 (0.56, 0.80)
*Lepomis gibbosus**	82	0.76 (0.65, 0.86)	0.52 (0.42, 0.63)
Lepomis macrochirus	117	0.77 (0.68, 0.85)	0.59 (0.49, 0.68)
Lethenteron appendix	39	0.63 (0.50, 0.87)	0.62 (0.46, 0.78)
Lota lota	32	0.50 (0.50, 0.50)	0.69 (0.50, 0.87)
Luxilus chrysocephalus	82	0.79 (0.69, 0.88)	0.62 (0.52, 0.72)
Luxilus cornutus	113	0.88 (0.80, 0.94)	0.71 (0.61, 0.80)
*Micropterus dolomieu**	125	0.93 (0.88, 0.98)	0.70 (0.60, 0.79)
*Micropterus salmoides**	65	0.81 (0.71, 0.92)	0.59 (0.48, 0.70)
*Moxostoma erythrurum**	97	0.90 (0.83, 0.95)	0.67 (0.57, 0.78)
*Nocomis micropogon**	118	0.88 (0.83, 0.93)	0.67 (0.57, 0.77)
*Notropis photogenis**	45	0.91 (0.82, 0.97)	0.64 (0.52, 0.79)
*Notropis rubellus**	103	0.90 (0.83, 0.96)	0.61 (0.50, 0.72)
Notropis stramineus	44	0.89 (0.77, 0.97)	0.61 (0.47, 0.77)
*Notropis volucellus**	79	0.86 (0.78, 0.92)	0.63 (0.53, 0.73)
*Noturus flavus**	56	0.89 (0.80, 0.95)	0.64 (0.53, 0.76)
Oncorhynchus mykiss	46	0.80 (0.67, 0.93)	0.59 (0.46, 0.73)
*Percina caprodes**	98	0.91 (0.86, 0.96)	0.71 (0.61, 0.80)
Perca flavescens	33	0.86 (0.68, 0.98)	0.65 (0.46, 0.81)
Percina macrocephala	49	0.94 (0.88, 0.99)	0.77 (0.64, 0.90)
*Percina maculata**	109	0.91 (0.86, 0.96)	0.70 (0.61, 0.79)
Percopsis omiscomaycus	31	0.72 (0.50, 0.98)	0.62 (0.45, 0.80)
*Pimephales notatus**	203	0.90 (0.85, 0.95)	0.68 (0.60, 0.76)
Rhinichthys cataractae	115	0.76 (0.68, 0.84)	0.65 (0.55, 0.74)
*Rhinichthys obtusus**	241	0.88 (0.81, 0.94)	0.70 (0.62, 0.79)
Salmo trutta	69	0.78 (0.62, 0.91)	0.63 (0.52, 0.75)
Salvelinus fontinalis	42	0.85 (0.75, 0.93)	0.63 (0.49, 0.80)
*Semotilus atromaculatus**	274	0.86 (0.80, 0.93)	0.68 (0.59, 0.76)

CRF represents AUC estimates for predictive performance of the full multi-species conditional random fields model that incorporates species co-occurrences and allows for them to be context-dependent. GLM represents AUC estimates for single-species logistic regression models that only accounts for the occurrence of a species with respect to the landscape-derived abiotic predictors. Species marked with * represent those with non-overlapping confidence intervals

Table 6 List of key coefficients for all species within the Mid-Atlantic watershed estimated from the bootstrapped conditional random fields model

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Ambloplites rupestris	*Lepomis auritus*	1.37 (0.63, 2.08)	0.20
	Micropterus dolomieu	1.18 (0.58, 1.79)	0.15
	Pimephales notatus	0.96 (0.26, 1.54)	0.10
	Open water × *Pimephales promelas*	− 0.78 (− 3.75, − 0.00)	0.07
	Lepomis cyanellus	0.75 (0.17, 1.39)	0.06
	Noturus insignis	0.64 (0.12, 1.20)	0.04
	Sander vitreus	0.59 (0.00, 1.39)	0.04
	Hypentelium nigricans	0.58 (0.02, 1.21)	0.04
	Stream temperature × *Exoglossum maxillingua*	0.50 (0.00, 1.18)	0.03
	Semotilus corporalis	0.43 (0.00, 1.02)	0.02
	Lepomis gibbosus	0.42 (0.00, 0.92)	0.02
	Rhinichthys atratulus	− 0.40 (− 0.80, − 0.04)	0.02
	Cyprinus carpio	0.39 (0.00, 1.11)	0.02
	Slope × *Lepomis cyanellus*	− 0.37 (− 0.97, 0.00)	0.01
	Nocomis micropogon	0.36 (0.00, 0.91)	0.01
	Exoglossum maxillingua	0.31 (0.00, 0.92)	0.01
Ameiurus natalis	*Lepomis auritus*	1.27 (0.65, 1.89)	0.35
	Micropterus salmoides	0.63 (0.08, 1.23)	0.09
	Micropterus dolomieu	0.55 (0.01, 1.19)	0.07
	Etheostoma olmstedi	0.51 (0.00, 1.22)	0.06
	Fundulus diaphanus	0.48 (0.00, 1.13)	0.05
	Semotilus atromaculatus	− 0.43 (− 1.07, 0.00)	0.04
	Catostomus commersonii	0.34 (0.00, 1.06)	0.03
	Etheostoma flabellare	0.30 (0.00, 1.03)	0.02
	Etheostoma zonale	0.29 (0.00, 0.82)	0.02
	Wetland × *Catostomus commersonii*	− 0.25 (− 0.84, 0.00)	0.01
	Wetland × *Notropis procne*	− 0.25 (− 1.09, 0.00)	0.01
	Cyprinella spiloptera	0.23 (0.00, 0.78)	0.01
	Cyprinus carpio	0.23 (0.00, 0.88)	0.01
	Forest (evergreen) × *Etheostoma flabellare*	0.22 (0.00, 0.99)	0.01
	Wetland × *Micropterus dolomieu*	− 0.22 (− 0.82, 0.00)	0.01
Ameiurus nebulosus	Open water × *Esox niger*	1.54 (0.00, 4.51)	0.52
	Open water × *Clinostomus elongatus*	− 0.49 (− 3.03, − 0.00)	0.05
	Esox niger	0.47 (0.00, 1.37)	0.05
	Lepomis gibbosus	0.41 (0.00, 1.06)	0.04
	Lepomis cyanellus	0.40 (0.00, 1.00)	0.04
	Sandy soil × *Lepomis cyanellus*	0.33 (0.00, 1.08)	0.02
	Developed × *Esox niger*	0.32 (0.00, 1.59)	0.02
	Lepomis macrochirus	0.25 (0.00, 0.89)	0.01
	Open water × *Lepomis gibbosus*	− 0.24 (− 1.21, 0.00)	0.01
	Salvelinus fontinalis	− 0.22 (− 0.98, 0.00)	0.01
	Sandy soil × *Anguilla rostrata*	0.22 (0.00, 1.37)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Anguilla rostrata	*Lepomis auritus*	1.57 (0.75, 2.36)	0.48
	Esox niger	0.67 (0.00, 1.58)	0.09
	Ictalurus punctatus	0.65 (0.00, 1.48)	0.08
	Agriculture × *Lepomis auritus*	− 0.37 (− 0.90, 0.00)	0.03
	Semotilus corporalis	0.34 (0.00, 0.97)	0.02
	Cottus spp.	− 0.31 (− 0.96, 0.00)	0.02
	Drainage area × *Esox niger*	− 0.31 (− 1.28, 0.00)	0.02
	Sandy soil × *Esox niger*	− 0.30 (− 1.28, 0.00)	0.02
	Herbaceous × *Lepomis auritus*	0.27 (0.00, 0.87)	0.01
	Forest (mixed) × *Lepomis auritus*	0.24 (0.00, 0.82)	0.01
	Wetland × *Notropis procne*	0.23 (0.00, 1.04)	0.01
	Nocomis micropogon	0.23 (0.00, 0.75)	0.01
Campostoma anomalum	*Rhinichthys cataractae*	2.19 (1.45, 2.94)	0.31
	Etheostoma blennioides	1.36 (0.73, 2.04)	0.12
	Open water × *Notropis volucellus*	− 1.20 (− 2.93, 0.00)	0.09
	Notropis volucellus	0.98 (0.34, 1.65)	0.06
	Open water × *Rhinichthys cataractae*	− 0.86 (− 2.54, 0.00)	0.05
	Luxilus cornutus	0.78 (0.25, 1.33)	0.04
	Hypentelium nigricans	0.75 (0.16, 1.41)	0.04
	Noturus insignis	0.74 (0.09, 1.31)	0.04
	Developed × *Luxilus cornutus*	− 0.74 (− 1.64, 0.00)	0.04
	Semotilus atromaculatus	0.73 (0.18, 1.32)	0.03
	Annual precipitation × *Semotilus atromaculatus*	− 0.66 (− 1.13, − 0.20)	0.03
	Developed × *Noturus insignis*	− 0.53 (− 1.69, 0.00)	0.02
	Open water × *Etheostoma flabellare*	− 0.49 (− 3.35, − 0.00)	0.02
	Open water × *Micropterus dolomieu*	− 0.41 (− 1.94, − 0.00)	0.01
Carpiodes cyprinus	*Sander vitreus*	1.43 (0.12, 2.85)	0.29
	Dorosoma cepedianum	1.00 (0.00, 2.28)	0.14
	Moxostoma macrolepidotum	0.90 (0.02, 2.03)	0.12
	Wetland × *Cyprinus carpio*	− 0.62 (− 1.86, 0.00)	0.05
	Sandy soil × *Dorosoma cepedianum*	− 0.61 (− 2.61, 0.00)	0.05
	Cyprinus carpio	0.54 (0.00, 1.45)	0.04
	Ictalurus punctatus	0.54 (0.00, 1.64)	0.04
	Agriculture × *Etheostoma zonale*	0.50 (0.00, 1.44)	0.04
	Annual precipitation × *Moxostoma macrolepidotum*	− 0.45 (− 1.72, 0.00)	0.03
	Annual precipitation × *Sander vitreus*	− 0.37 (− 1.17, 0.00)	0.02
	Developed × *Cyprinus carpio*	− 0.35 (− 1.94, 0.00)	0.02
	Stream temperature × *Moxostoma macrolepidotum*	0.34 (0.00, 1.34)	0.02
	Etheostoma zonale	0.27 (0.00, 0.81)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Catostomus commersonii	Drainage area × *Exoglossum maxillingua*	− 1.14 (− 2.84, 0.00)	0.13
	Etheostoma olmstedi	1.08 (0.48, 1.70)	0.11
	Rhinichthys cataractae	1.03 (0.32, 1.83)	0.10
	Rhinichthys atratulus	1.01 (0.37, 1.68)	0.10
	Micropterus salmoides	0.95 (0.23, 1.63)	0.09
	Exoglossum maxillingua	0.95 (0.29, 1.71)	0.09
	Salmo trutta	0.84 (0.14, 1.74)	0.07
	Hypentelium nigricans	0.79 (0.15, 1.50)	0.06
	Semotilus atromaculatus	0.64 (0.12, 1.21)	0.04
	Drainage area × *Salmo trutta*	− 0.51 (− 2.62, 0.00)	0.03
	Cyprinus carpio	0.38 (0.00, 1.03)	0.01
	Ameiurus natalis	0.34 (0.00, 1.06)	0.01
Clinostomus elongatus	Sandy soil × *Cottus spp.*	− 1.21 (− 2.69, − 0.00)	0.31
	Forest (mixed) × *Luxilus cornutus*	0.92 (0.00, 1.99)	0.18
	Open water × *Ameiurus nebulosus*	− 0.49 (− 3.03, − 0.00)	0.05
	Cottus spp.	0.44 (0.00, 1.38)	0.04
	Luxilus cornutus	0.44 (0.00, 1.27)	0.04
	Micropterus dolomieu	− 0.35 (− 1.28, 0.00)	0.03
	Sandy soil × *Luxilus cornutus*	− 0.32 (− 1.32, 0.00)	0.02
	Developed × *Semotilus atromaculatus*	− 0.31 (− 1.25, 0.00)	0.02
	Drainage area × *Campostoma anomalum*	− 0.30 (− 1.81, − 0.00)	0.02
	Rhinichthys atratulus	0.29 (0.00, 1.15)	0.02
	Stream temperature × *Semotilus atromaculatus*	− 0.29 (− 0.91, 0.00)	0.02
	Ambloplites rupestris	− 0.28 (− 0.98, 0.00)	0.02
	Exoglossum maxillingua	0.28 (0.00, 1.10)	0.02
	Semotilus atromaculatus	0.25 (0.00, 0.88)	0.01
	Forest (mixed) × *Ambloplites rupestris*	0.25 (− 0.00, 0.99)	0.01
	Open water × *Semotilus atromaculatus*	− 0.25 (− 1.82, 0.00)	0.01
Clinostomus funduloides	Drainage area × *Semotilus atromaculatus*	− 3.83 (− 6.17, − 1.14)	0.76
	Semotilus atromaculatus	1.36 (0.19, 2.47)	0.10
	Open water × *Semotilus atromaculatus*	− 1.33 (− 4.86, − 0.00)	0.09
	Drainage area × *Rhinichthys atratulus*	− 0.48 (− 2.20, 0.00)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Cottus spp.	Sandy soil × *Clinostomus elongatus*	− 1.21 (− 2.69, − 0.00)	0.24
	Lepomis macrochirus	− 0.82 (− 1.46, − 0.30)	0.11
	Lepomis auritus	− 0.77 (− 1.51, − 0.07)	0.10
	Rhinichthys atratulus	0.67 (0.06, 1.27)	0.07
	Sander vitreus	− 0.56 (− 1.42, 0.00)	0.05
	Drainage area × *Exoglossum maxillingua*	− 0.51 (− 1.90, − 0.00)	0.04
	Rhinichthys cataractae	0.48 (0.00, 1.11)	0.04
	Clinostomus elongatus	0.44 (0.00, 1.38)	0.03
	Dorosoma cepedianum	− 0.42 (− 1.06, 0.00)	0.03
	Sandy soil	0.38 (0.11, 0.65)	0.02
	Stream temperature × *Oncorhynchus mykiss*	0.38 (0.00, 1.50)	0.02
	Open water × *Oncorhynchus mykiss*	− 0.36 (− 2.39, -0.00)	0.02
	Open water × *Rhinichthys atratulus*	− 0.35 (− 1.41, 0.00)	0.02
	Slope × *Rhinichthys atratulus*	0.35 (0.00, 0.68)	0.02
	Anguilla rostrata	− 0.31 (− 0.96, 0.00)	0.02
	Drainage area × *Rhinichthys atratulus*	− 0.29 (− 1.87, 0.00)	0.01
	Drainage area × *Salmo trutta*	− 0.27 (− 1.46, 0.00)	0.01
	Nocomis micropogon	0.26 (0.00, 0.77)	0.01
	Open water × *Lepomis macrochirus*	− 0.25 (− 1.09, − 0.00)	0.01
Cyprinella analostana	*Notropis procne*	1.26 (0.00, 3.05)	0.40
	Open water × *Notropis procne*	− 0.82 (− 4.00, 0.00)	0.17
	Open water × *Dorosoma cepedianum*	0.54 (0.00, 4.50)	0.07
	Etheostoma olmstedi	0.46 (0.00, 1.21)	0.05
	Lepomis auritus	0.37 (0.00, 1.20)	0.04
	Sandy soil × *Fundulus diaphanus*	0.33 (0.00, 1.99)	0.03
	Sandy soil × *Ictalurus punctatus*	0.32 (0.00, 1.95)	0.03
	Forest (evergreen) × *Lepomis auritus*	− 0.30 (− 1.28, 0.00)	0.02
	Annual precipitation × *Notropis procne*	0.27 (0.00, 1.18)	0.02
	Ictalurus punctatus	0.27 (0.00, 1.10)	0.02
	Lepomis gibbosus	0.22 (0.00, 0.81)	0.01
	Sandy soil × *Notropis procne*	− 0.22 (− 1.66, 0.00)	0.01
	Agriculture × *Notropis procne*	− 0.20 (− 0.92, 0.00)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Cyprinus carpio	*Ictalurus punctatus*	1.35 (0.38, 2.36)	0.18
	Sandy soil × *Perca flavescens*	− 1.01 (− 3.32, − 0.00)	0.10
	Sander vitreus	0.91 (0.00, 1.83)	0.08
	Sandy soil × *Cyprinella spiloptera*	− 0.75 (− 2.19, 0.00)	0.06
	Lepomis gibbosus	0.74 (0.10, 1.56)	0.05
	Micropterus salmoides	0.69 (0.00, 1.53)	0.05
	Percina peltata	0.67 (0.00, 1.46)	0.04
	Wetland × *Carpiodes cyprinus*	− 0.62 (− 1.86, 0.00)	0.04
	Slope × *Ictalurus punctatus*	− 0.61 (− 1.75, 0.00)	0.04
	Carpiodes cyprinus	0.54 (0.00, 1.45)	0.03
	Cyprinella spiloptera	0.52 (0.00, 1.30)	0.03
	Lepomis cyanellus	0.48 (0.00, 1.19)	0.02
	Rhinichthys cataractae	− 0.48 (− 1.19, 0.00)	0.02
	Dorosoma cepedianum	0.42 (0.00, 1.25)	0.02
	Ambloplites rupestris	0.39 (0.00, 1.11)	0.02
	Catostomus commersonii	0.38 (0.00, 1.03)	0.01
	Perca flavescens	0.38 (0.00, 1.22)	0.01
	Sandy soil × *Sander vitreus*	− 0.35 (− 1.59, − 0.00)	0.01
	Developed × *Carpiodes cyprinus*	− 0.35 (− 1.94, 0.00)	0.01
	Agriculture × *Notropis hudsonius*	− 0.34 (− 1.27, − 0.00)	0.01
	Slope × *Dorosoma cepedianum*	− 0.33 (− 1.12, 0.00)	0.01
	Herbaceous × *Percina peltata*	− 0.33 (− 1.01, 0.00)	0.01
	Sandy soil × *Ictalurus punctatus*	− 0.32 (− 1.57, 0.00)	0.01
Cyprinella spiloptera	*Moxostoma macrolepidotum*	1.44 (0.50, 2.46)	0.18
	Pimephales notatus	1.40 (0.70, 2.13)	0.17
	Semotilus corporalis	1.29 (0.62, 2.02)	0.14
	Micropterus dolomieu	0.99 (0.29, 1.73)	0.09
	Lepomis auritus	0.90 (0.23, 1.55)	0.07
	Sandy soil × *Cyprinus carpio*	− 0.75 (− 2.19, 0.00)	0.05
	Notropis volucellus	0.73 (0.10, 1.40)	0.05
	Hypentelium nigricans	0.60 (0.00, 1.33)	0.03
	Percina peltata	0.55 (0.00, 1.23)	0.03
	Cyprinus carpio	0.52 (0.00, 1.30)	0.02
	Open water × *Notropis procne*	− 0.37 (− 2.84, 0.00)	0.01
	Open water × *Moxostoma macrolepidotum*	0.35 (0.00, 1.05)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Dorosoma cepedianum	*Carpiodes cyprinus*	1.00 (0.00, 2.28)	0.22
	Agriculture × *Perca flavescens*	0.63 (0.00, 1.61)	0.09
	Sandy soil × *Carpiodes cyprinus*	− 0.61 (− 2.61, 0.00)	0.08
	Open water × *Cyprinella analostana*	0.54 (0.00, 4.50)	0.07
	Developed × *Micropterus salmoides*	− 0.43 (− 1.65, − 0.00)	0.04
	Cyprinus carpio	0.42 (0.00, 1.25)	0.04
	Cottus spp.	− 0.42 (− 1.06, 0.00)	0.04
	Notropis hudsonius	0.40 (0.00, 1.35)	0.04
	Perca flavescens	0.37 (0.00, 1.27)	0.03
	Sandy soil × *Moxostoma macrolepidotum*	− 0.37 (− 3.24, − 0.00)	0.03
	Forest (mixed) × *Ictalurus punctatus*	− 0.34 (− 1.86, − 0.00)	0.03
	Forest (mixed) × *Sander vitreus*	− 0.34 (− 1.44, − 0.00)	0.03
	Slope × *Cyprinus carpio*	− 0.33 (− 1.12, 0.00)	0.02
	Drainage area × *Notropis amoenus*	0.30 (0.00, 1.04)	0.02
	Forest (evergreen) × *Notropis amoenus*	− 0.28 (− 1.50, 0.00)	0.02
	Notropis amoenus	0.26 (0.00, 1.03)	0.02
	Wetland × *Pylodictis olivaris*	− 0.23 (− 1.86, − 0.00)	0.01
	Forest (mixed) × *Cyprinus carpio*	− 0.23 (− 1.00, 0.00)	0.01
	Exoglossum maxillingua	− 0.22 (− 0.91, 0.00)	0.01
Esox niger	Open water × *Ameiurus nebulosus*	1.54 (0.00, 4.51)	0.39
	Lepomis gibbosus	0.81 (0.14, 1.57)	0.11
	Anguilla rostrata	0.67 (0.00, 1.58)	0.07
	Lepomis macrochirus	0.57 (0.00, 1.45)	0.05
	Hypentelium nigricans	0.49 (0.00, 1.22)	0.04
	Ameiurus nebulosus	0.47 (0.00, 1.37)	0.04
	Slope × *Perca flavescens*	0.36 (0.00, 1.38)	0.02
	Agriculture × *Lepomis macrochirus*	− 0.34 (− 1.22, 0.00)	0.02
	Wetland × *Lepomis macrochirus*	0.34 (0.00, 0.75)	0.02
	Perca flavescens	0.33 (0.00, 1.08)	0.02
	Developed × *Ameiurus nebulosus*	0.32 (0.00, 1.59)	0.02
	Drainage area × *Anguilla rostrata*	− 0.31 (− 1.28, 0.00)	0.02
	Sandy soil × *Anguilla rostrata*	− 0.30 (− 1.28, 0.00)	0.01
	Sandy soil × *Lepomis auritus*	0.30 (0.00, 1.30)	0.01
	Agriculture × *Lepomis gibbosus*	− 0.27 (− 0.94, 0.00)	0.01
	Developed × *Lepomis macrochirus*	− 0.26 (− 1.42, 0.00)	0.01
	Semotilus corporalis	0.26 (0.00, 0.95)	0.01
Etheostoma blennioides	*Etheostoma zonale*	3.18 (2.45, 3.92)	0.51
	Open water × *Etheostoma flabellare*	− 2.19 (− 4.90, − 0.00)	0.24
	Campostoma anomalum	1.36 (0.73, 2.04)	0.09
	Notropis rubellus	1.13 (0.44, 1.82)	0.06

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Etheostoma flabellare	Open water × *Etheostoma blennioides*	− 2.19 (− 4.90, − 0.00)	0.69
	Open water × *Semotilus atromaculatus*	− 0.61 (− 3.00, 0.00)	0.05
	Stream temperature × *Noturus insignis*	0.50 (0.00, 1.35)	0.04
	Open water × *Campostoma anomalum*	− 0.49 (− 3.35, − 0.00)	0.03
	Noturus insignis	0.44 (0.00, 1.02)	0.03
	Ameiurus natalis	0.30 (0.00, 1.03)	0.01
	Drainage area × *Semotilus atromaculatus*	− 0.30 (− 1.82, 0.00)	0.01
	Notropis rubellus	0.29 (0.00, 1.05)	0.01
Etheostoma olmstedi	*Exoglossum maxillingua*	1.54 (0.78, 2.24)	0.22
	Percina peltata	1.21 (0.50, 1.93)	0.13
	Notropis hudsonius	1.13 (0.35, 1.90)	0.12
	Pimephales notatus	1.12 (0.49, 1.72)	0.11
	Catostomus commersonii	1.08 (0.48, 1.70)	0.11
	Etheostoma zonale	0.87 (0.13, 1.74)	0.07
	Luxilus cornutus	0.77 (0.13, 1.40)	0.05
	Open water × *Exoglossum maxillingua*	− 0.70 (− 2.02, 0.00)	0.04
	Ameiurus natalis	0.51 (0.00, 1.22)	0.02
	Cyprinella analostana	0.46 (0.00, 1.21)	0.02
Etheostoma zonale	*Etheostoma blennioides*	3.18 (2.45, 3.92)	0.51
	Notropis volucellus	2.14 (1.47, 2.87)	0.23
	Etheostoma olmstedi	0.87 (0.13, 1.74)	0.04
	Percina peltata	0.84 (0.16, 1.61)	0.04
	Nocomis micropogon	0.82 (0.18, 1.49)	0.03
	Hypentelium nigricans	0.82 (0.20, 1.50)	0.03
	Forest (evergreen) × *Notropis volucellus*	− 0.74 (− 1.51, − 0.06)	0.03
	Agriculture × *Carpiodes cyprinus*	0.50 (0.00, 1.44)	0.01
	Micropterus salmoides	− 0.50 (− 1.08, − 0.02)	0.01
Exoglossum maxillingua	Drainage area × *Rhinichthys cataractae*	− 3.93 (− 5.50, − 2.49)	0.48
	Drainage area × *Luxilus cornutus*	− 1.73 (− 3.27, − 0.28)	0.09
	Etheostoma olmstedi	1.54 (0.78, 2.24)	0.07
	Noturus insignis	1.47 (0.81, 2.15)	0.07
	Luxilus cornutus	1.33 (0.66, 2.01)	0.05
	Drainage area × *Catostomus commersonii*	− 1.14 (− 2.84, 0.00)	0.04
	Rhinichthys cataractae	1.04 (0.56, 1.70)	0.03
	Catostomus commersonii	0.95 (0.29, 1.71)	0.03
	Hypentelium nigricans	0.91 (0.28, 1.59)	0.03
	Open water × *Etheostoma olmstedi*	− 0.70 (− 2.02, 0.00)	0.02

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Fundulus diaphanus	*Lepomis auritus*	1.11 (0.33, 1.84)	0.37
	Ameiurus natalis	0.48 (0.00, 1.13)	0.07
	Agriculture × *Pimephales notatus*	0.38 (0.00, 1.20)	0.04
	Open water × *Notropis procne*	− 0.38 (− 2.19, − 0.00)	0.04
	Developed	0.37 (0.00, 0.72)	0.04
	Notropis volucellus	0.37 (0.00, 0.99)	0.04
	Sandy soil × *Cyprinella analostana*	0.33 (0.00, 1.99)	0.03
	Cyprinella spiloptera	0.33 (0.00, 0.97)	0.03
	Slope × *Notropis hudsonius*	− 0.32 (− 1.12, − 0.00)	0.03
	Salmo trutta	− 0.26 (− 0.86, 0.00)	0.02
	Herbaceous × *Notropis amoenus*	0.24 (0.00, 1.40)	0.02
	Lepomis gibbosus	0.24 (0.00, 0.76)	0.02
	Forest (evergreen) × *Cyprinus carpio*	0.21 (0.00, 1.14)	0.01
	Etheostoma olmstedi	0.21 (0.00, 0.79)	0.01
	Herbaceous × *Notropis volucellus*	− 0.20 (− 0.72, 0.00)	0.01
	Rhinichthys atratulus	− 0.19 (− 0.59, 0.00)	0.01
	Micropterus salmoides	0.18 (0.00, 0.68)	0.01
Hypentelium nigricans	*Notropis rubellus*	1.50 (0.84, 2.19)	0.19
	Micropterus dolomieu	1.08 (0.41, 1.78)	0.10
	Notropis hudsonius	0.99 (0.30, 1.78)	0.08
	Sander vitreus	0.95 (0.14, 1.82)	0.08
	Exoglossum maxillingua	0.91 (0.28, 1.59)	0.07
	Etheostoma zonale	0.82 (0.20, 1.50)	0.06
	Catostomus commersonii	0.79 (0.15, 1.50)	0.05
	Campostoma anomalum	0.75 (0.16, 1.41)	0.05
	Forest (evergreen) × *Semotilus corporalis*	0.75 (0.00, 1.47)	0.05
	Semotilus corporalis	0.61 (0.02, 1.19)	0.03
	Cyprinella spiloptera	0.60 (0.00, 1.33)	0.03
	Noturus insignis	0.58 (0.00, 1.27)	0.03
	Ambloplites rupestris	0.58 (0.02, 1.21)	0.03
	Esox niger	0.49 (0.00, 1.22)	0.02
	Stream temperature × *Notropis hudsonius*	0.36 (0.00, 1.02)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Ictalurus punctatus	*Cyprinus carpio*	1.35 (0.38, 2.36)	0.18
	Forest (mixed) × *Pylodictis olivaris*	− 0.92 (− 2.48, 0.00)	0.08
	Sander vitreus	0.84 (0.00, 1.83)	0.07
	Pylodictis olivaris	0.80 (0.00, 2.08)	0.06
	Notropis volucellus	0.76 (0.00, 1.70)	0.06
	Agriculture × *Pylodictis olivaris*	0.68 (0.00, 2.28)	0.05
	Anguilla rostrata	0.65 (0.00, 1.48)	0.04
	Moxostoma macrolepidotum	0.65 (0.00, 1.76)	0.04
	Drainage area	0.64 (0.00, 1.35)	0.04
	Micropterus dolomieu	0.62 (0.00, 1.53)	0.04
	Slope × *Cyprinus carpio*	− 0.61 (− 1.75, 0.00)	0.04
	Carpiodes cyprinus	0.54 (0.00, 1.64)	0.03
	Sandy soil × *Sander vitreus*	− 0.53 (− 1.86, 0.00)	0.03
	Sandy soil × *Pylodictis olivaris*	− 0.48 (− 2.64, 0.00)	0.02
	Annual precipitation × *Notropis procne*	0.45 (0.00, 1.70)	0.02
	Sandy soil × *Moxostoma macrolepidotum*	− 0.43 (− 1.75, 0.00)	0.02
	Forest (mixed) × *Dorosoma cepedianum*	− 0.34 (− 1.86, − 0.00)	0.01
	Forest (evergreen) × *Notropis procne*	0.33 (0.00, 1.40)	0.01
	Rhinichthys atratulus	− 0.32 (− 0.82, 0.00)	0.01
	Forest (mixed) × *Notropis volucellus*	− 0.32 (− 1.25, 0.00)	0.01
Lepomis auritus	*Anguilla rostrata*	1.57 (0.75, 2.36)	0.20
	Ambloplites rupestris	1.37 (0.63, 2.08)	0.15
	Ameiurus natalis	1.27 (0.65, 1.89)	0.13
	Fundulus diaphanus	1.11 (0.33, 1.84)	0.10
	Sander vitreus	0.98 (0.17, 1.77)	0.08
	Cyprinella spiloptera	0.90 (0.23, 1.55)	0.07
	Cottus spp.	− 0.77 (− 1.51, − 0.07)	0.05
	Notropis hudsonius	0.63 (0.04, 1.35)	0.03
	Lepomis gibbosus	0.41 (0.00, 0.95)	0.01
	Cyprinella analostana	0.37 (0.00, 1.20)	0.01
	Agriculture × *Anguilla rostrata*	− 0.37 (− 0.90, 0.00)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Lepomis cyanellus	*Lepomis macrochirus*	0.78 (0.22, 1.30)	0.14
	Ambloplites rupestris	0.75 (0.17, 1.39)	0.13
	Cyprinus carpio	0.48 (0.00, 1.19)	0.06
	Lepomis gibbosus	0.42 (0.00, 0.91)	0.04
	Notropis volucellus	0.42 (0.00, 0.97)	0.04
	Ameiurus nebulosus	0.40 (0.00, 1.00)	0.04
	Slope × *Ambloplites rupestris*	− 0.37 (− 0.97, 0.00)	0.03
	Notemigonus crysoleucas	0.37 (0.00, 0.98)	0.03
	Pimephales notatus	0.34 (0.00, 0.87)	0.03
	Sandy soil × *Ameiurus nebulosus*	0.33 (0.00, 1.08)	0.03
	Clinostomus funduloides	− 0.31 (− 1.03, 0.00)	0.02
	Open water × *Notemigonus crysoleucas*	− 0.31 (− 1.98, 0.00)	0.02
	Forest (mixed) × *Etheostoma zonale*	0.29 (0.00, 0.95)	0.02
	Herbaceous × *Notemigonus crysoleucas*	0.28 (0.00, 0.88)	0.02
	Salmo trutta	− 0.27 (− 0.84, 0.00)	0.02
	Catostomus commersonii	0.26 (0.00, 0.82)	0.02
	Open water × *Ambloplites rupestris*	− 0.26 (− 0.78, 0.00)	0.02
	Etheostoma zonale	0.24 (0.00, 0.72)	0.01
	Forest (evergreen) × *Notemigonus crysoleucas*	− 0.24 (− 0.97, 0.00)	0.01
	Slope × *Carpiodes cyprinus*	0.23 (0.00, 1.17)	0.01
	Anguilla rostrata	0.21 (0.00, 0.66)	0.01
Lepomis gibbosus	*Lepomis macrochirus*	1.60 (1.14, 2.14)	0.36
	Esox niger	0.81 (0.14, 1.57)	0.09
	Cyprinus carpio	0.74 (0.10, 1.56)	0.08
	Micropterus salmoides	0.43 (0.00, 0.99)	0.03
	Lepomis cyanellus	0.42 (0.00, 0.91)	0.02
	Ambloplites rupestris	0.42 (0.00, 0.92)	0.02
	Slope × *Micropterus salmoides*	− 0.42 (− 0.94, − 0.00)	0.02
	Lepomis auritus	0.41 (0.00, 0.95)	0.02
	Notemigonus crysoleucas	0.41 (0.00, 1.10)	0.02
	Ameiurus nebulosus	0.41 (0.00, 1.06)	0.02
	Semotilus atromaculatus	0.40 (0.00, 0.87)	0.02
	Perca flavescens	0.35 (0.00, 0.95)	0.02
	Notropis hudsonius	0.29 (0.00, 0.79)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Lepomis macrochirus	*Micropterus salmoides*	1.66 (1.10, 2.23)	0.21
	Lepomis gibbosus	1.60 (1.14, 2.14)	0.20
	Drainage area × *Oncorhynchus mykiss*	− 1.48 (− 3.64, 0.00)	0.17
	Cottus spp.	− 0.82 (− 1.46, − 0.30)	0.05
	Lepomis cyanellus	0.78 (0.22, 1.30)	0.05
	Pimephales notatus	0.77 (0.17, 1.32)	0.05
	Esox niger	0.57 (0.00, 1.45)	0.03
	Oncorhynchus mykiss	0.56 (0.00, 1.49)	0.02
	Stream temperature × *Oncorhynchus mykiss*	− 0.52 (− 1.29, 0.00)	0.02
	Micropterus dolomieu	0.50 (0.01, 1.04)	0.02
	Perca flavescens	0.45 (0.00, 1.05)	0.02
	Rhinichthys atratulus	− 0.44 (− 0.96, 0.00)	0.02
	Notropis amoenus	0.44 (0.00, 1.19)	0.01
Luxilus cornutus	Drainage area × *Exoglossum maxillingua*	− 1.73 (− 3.27, − 0.28)	0.20
	Exoglossum maxillingua	1.33 (0.66, 2.01)	0.12
	Open water × *Pimephales notatus*	− 1.11 (− 2.38, 0.00)	0.08
	Sandy soil × *Notemigonus crysoleucas*	− 1.06 (− 2.51, − 0.00)	0.07
	Forest (mixed) × *Clinostomus elongatus*	0.92 (0.00, 1.99)	0.06
	Pimephales notatus	0.84 (0.27, 1.42)	0.05
	Campostoma anomalum	0.78 (0.25, 1.33)	0.04
	Etheostoma olmstedi	0.77 (0.13, 1.40)	0.04
	Developed × *Campostoma anomalum*	− 0.74 (− 1.64, 0.00)	0.04
	Noturus insignis	0.70 (0.16, 1.27)	0.03
	Open water × *Semotilus atromaculatus*	− 0.66 (− 2.02, 0.00)	0.03
	Notropis rubellus	0.59 (0.00, 1.33)	0.02
	Semotilus atromaculatus	0.59 (0.03, 1.21)	0.02
	Stream temperature × *Exoglossum maxillingua*	0.51 (0.08, 1.00)	0.02
	Clinostomus elongatus	0.44 (0.00, 1.27)	0.01
	Open water × *Notropis amoenus*	− 0.40 (− 1.76, 0.00)	0.01
Micropterus dolomieu	*Ambloplites rupestris*	1.18 (0.58, 1.79)	0.16
	Hypentelium nigricans	1.08 (0.41, 1.78)	0.14
	Cyprinella spiloptera	0.99 (0.29, 1.73)	0.12
	Semotilus corporalis	0.91 (0.30, 1.54)	0.10
	Percina peltata	0.80 (0.12, 1.49)	0.08
	Ictalurus punctatus	0.62 (0.00, 1.53)	0.05
	Ameiurus natalis	0.55 (0.01, 1.19)	0.04
	Lepomis macrochirus	0.50 (0.01, 1.04)	0.03
	Noturus insignis	0.45 (0.00, 1.08)	0.02
	Etheostoma blennioides	0.44 (0.00, 1.13)	0.02
	Open water × *Campostoma anomalum*	− 0.41 (− 1.94, − 0.00)	0.02
	Sander vitreus	0.40 (0.00, 1.18)	0.02
	Moxostoma macrolepidotum	0.35 (0.00, 1.09)	0.01
	Clinostomus elongatus	− 0.35 (− 1.28, 0.00)	0.01
	Notropis amoenus	0.35 (0.00, 1.21)	0.01
	Sandy soil × *Sander vitreus*	− 0.30 (− 1.61, 0.00)	0.01
	Agriculture × *Etheostoma blennioides*	0.29 (0.00, 0.77)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Micropterus salmoides	*Lepomis macrochirus*	1.66 (1.10, 2.23)	0.28
	Notropis amoenus	0.97 (0.16, 1.79)	0.10
	Catostomus commersonii	0.95 (0.23, 1.63)	0.09
	Perca flavescens	0.93 (0.17, 1.68)	0.09
	Cyprinus carpio	0.69 (0.00, 1.53)	0.05
	Ameiurus natalis	0.63 (0.08, 1.23)	0.04
	Etheostoma zonale	− 0.50 (− 1.08, − 0.02)	0.03
	Pimephales notatus	0.50 (0.00, 1.07)	0.03
	Developed × *Dorosoma cepedianum*	− 0.43 (− 1.65, − 0.00)	0.02
	Lepomis gibbosus	0.43 (0.00, 0.99)	0.02
	Slope × *Lepomis gibbosus*	− 0.42 (− 0.94, − 0.00)	0.02
	Herbaceous × *Notemigonus crysoleucas*	− 0.36 (− 1.63, − 0.00)	0.01
	Forest (evergreen) × *Pimephales notatus*	0.36 (0.00, 0.92)	0.01
	Stream temperature × *Notropis hudsonius*	0.32 (0.00, 0.92)	0.01
	Semotilus atromaculatus	0.32 (0.00, 0.80)	0.01
	Semotilus corporalis	0.31 (0.00, 0.87)	0.01
Moxostoma macrolepidotum	*Cyprinella spiloptera*	1.44 (0.50, 2.46)	0.37
	Carpiodes cyprinus	0.90 (0.02, 2.03)	0.14
	Ictalurus punctatus	0.65 (0.00, 1.76)	0.07
	Annual precipitation × *Carpiodes cyprinus*	− 0.45 (− 1.72, 0.00)	0.04
	Sandy soil × *Ictalurus punctatus*	− 0.43 (− 1.75, 0.00)	0.03
	Pylodictis olivaris	0.37 (0.00, 1.35)	0.02
	Sandy soil × *Dorosoma cepedianum*	− 0.37 (− 3.24, − 0.00)	0.02
	Micropterus dolomieu	0.35 (0.00, 1.09)	0.02
	Open water × *Cyprinella spiloptera*	0.35 (0.00, 1.05)	0.02
	Stream temperature × *Carpiodes cyprinus*	0.34 (0.00, 1.34)	0.02
	Forest (mixed) × *Notropis amoenus*	v0.33 (− 1.28, 0.00)	0.02
	Herbaceous × *Notropis amoenus*	− 0.33 (− 1.26, 0.00)	0.02
	Forest (evergreen) × *Cyprinella spiloptera*	− 0.32 (− 1.04, 0.00)	0.02
	Stream temperature × *Ictalurus punctatus*	0.31 (0.00, 1.04)	0.02
	Forest (mixed) × *Ictalurus punctatus*	− 0.27 (− 1.14, 0.00)	0.01
	Forest (mixed) × *Nocomis micropogon*	− 0.27 (− 1.15, 0.00)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Nocomis micropogon	*Notropis rubellus*	1.46 (0.81, 2.09)	0.24
	Noturus insignis	1.36 (0.71, 2.01)	0.21
	Open water × *Rhinichthys cataractae*	− 0.96 (− 3.60, 0.00)	0.10
	Etheostoma zonale	0.82 (0.18, 1.49)	0.08
	Open water × *Noturus insignis*	− 0.68 (− 2.08, 0.00)	0.05
	Drainage area × *Exoglossum maxillingua*	− 0.53 (− 1.70, − 0.00)	0.03
	Semotilus corporalis	0.41 (0.00, 0.98)	0.02
	Rhinichthys cataractae	0.36 (0.00, 1.02)	0.01
	Ambloplites rupestris	0.36 (0.00, 0.91)	0.01
	Percina peltata	0.34 (0.00, 0.89)	0.01
	Sandy soil × *Sander vitreus*	− 0.33 (− 1.85, 0.00)	0.01
	Campostoma anomalum	0.31 (0.00, 0.87)	0.01
	Drainage area × *Campostoma anomalum*	− 0.31 (− 0.93, − 0.00)	0.01
	Open water × *Luxilus cornutus*	− 0.30 (− 1.33, − 0.00)	0.01
Notemigonus crysoleucas	Sandy soil × *Luxilus cornutus*	− 1.06 (− 2.51, − 0.00)	0.28
	Sandy soil × *Pimephales promelas*	− 0.56 (− 2.43, 0.00)	0.08
	Lepomis gibbosus	0.41 (0.00, 1.10)	0.04
	Lepomis cyanellus	0.37 (0.00, 0.98)	0.03
	Herbaceous × *Micropterus salmoides*	− 0.36 (− 1.63, − 0.00)	0.03
	Sandy soil × *Lepomis macrochirus*	− 0.34 (− 1.27, − 0.00)	0.03
	Luxilus cornutus	0.32 (0.00, 1.04)	0.03
	Open water × *Lepomis cyanellus*	− 0.31 (− 1.98, 0.00)	0.02
	Noturus insignis	− 0.30 (− 0.99, 0.00)	0.02
	Herbaceous × *Lepomis cyanellus*	0.28 (0.00, 0.88)	0.02
	Open water × *Pimephales promelas*	− 0.28 (− 1.92, 0.00)	0.02
	Herbaceous × *Lepomis gibbosus*	0.26 (0.00, 0.95)	0.02
	Agriculture × *Pimephales promelas*	0.26 (0.00, 1.14)	0.02
	Open water × *Catostomus commersonii*	− 0.25 (− 1.68, − 0.00)	0.02
	Salvelinus fontinalis	− 0.25 (− 1.12, 0.00)	0.02
	Lepomis macrochirus	0.25 (0.00, 0.94)	0.02
	Catostomus commersonii	0.24 (0.00, 0.98)	0.01
	Pimephales promelas	0.24 (0.00, 0.92)	0.01
	Forest (evergreen) × *Lepomis cyanellus*	− 0.24 (− 0.97, 0.00)	0.01
	Sandy soil × *Lepomis gibbosus*	− 0.22 (− 0.91, 0.00)	0.01
	Oncorhynchus mykiss	0.22 (0.00, 0.91)	0.01
	Open water × *Lepomis gibbosus*	− 0.20 (− 1.48, 0.00)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Notropis amoenus	Sandy soil × *Pylodictis olivaris*	− 1.23 (− 4.62, 0.00)	0.24
	Micropterus salmoides	0.97 (0.16, 1.79)	0.15
	Sandy soil × *Notropis hudsonius*	− 0.78 (− 2.63, − 0.00)	0.10
	Notropis hudsonius	0.56 (0.00, 1.42)	0.05
	Pylodictis olivaris	0.48 (0.00, 1.37)	0.04
	Lepomis macrochirus	0.44 (0.00, 1.19)	0.03
	Stream temperature × *Pylodictis olivaris*	− 0.40 (− 1.35, − 0.00)	0.03
	Open water × *Luxilus cornutus*	− 0.40 (− 1.76, 0.00)	0.03
	Semotilus corporalis	0.40 (0.00, 1.05)	0.02
	Micropterus dolomieu	0.35 (0.00, 1.21)	0.02
	Forest (mixed) × *Moxostoma macrolepidotum*	− 0.33 (− 1.28, 0.00)	0.02
	Herbaceous × *Moxostoma macrolepidotum*	− 0.33 (− 1.26, 0.00)	0.02
	Forest (evergreen) × *Pylodictis olivaris*	− 0.31 (− 1.36, − 0.00)	0.02
	Forest (evergreen) × *Micropterus salmoides*	− 0.31 (− 1.19, 0.00)	0.02
	Forest (evergreen) × *Cyprinella spiloptera*	− 0.30 (− 1.31, 0.00)	0.01
	Drainage area × *Dorosoma cepedianum*	0.30 (0.00, 1.04)	0.01
	Forest (evergreen) × *Dorosoma cepedianum*	− 0.28 (− 1.50, 0.00)	0.01
	Wetland × *Pylodictis olivaris*	− 0.28 (− 1.74, − 0.00)	0.01
	Developed × *Cyprinella spiloptera*	0.26 (0.00, 0.88)	0.01
	Dorosoma cepedianum	0.26 (0.00, 1.03)	0.01
	Cyprinella spiloptera	0.26 (0.00, 0.93)	0.01
Notropis hudsonius	*Notropis procne*	1.37 (0.34, 2.40)	0.18
	Open water × *Notropis procne*	− 1.26 (− 4.76, 0.00)	0.15
	Etheostoma olmstedi	1.13 (0.35, 1.90)	0.12
	Hypentelium nigricans	0.99 (0.30, 1.78)	0.09
	Sandy soil × *Notropis amoenus*	− 0.78 (− 2.63, − 0.00)	0.06
	Lepomis auritus	0.63 (0.04, 1.35)	0.04
	Pimephales notatus	0.57 (0.00, 1.38)	0.03
	Notropis amoenus	0.56 (0.00, 1.42)	0.03
	Notropis rubellus	0.50 (0.00, 1.08)	0.02
	Dorosoma cepedianum	0.40 (0.00, 1.35)	0.02
	Semotilus corporalis	0.37 (0.00, 1.03)	0.01
	Annual precipitation × *Semotilus corporalis*	− 0.36 (− 0.92, − 0.00)	0.01
	Stream temperature × *Hypentelium nigricans*	0.36 (0.00, 1.02)	0.01
	Semotilus atromaculatus	0.34 (0.00, 0.92)	0.01
	Agriculture × *Cyprinus carpio*	− 0.34 (− 1.27, − 0.00)	0.01
	Stream temperature × *Micropterus salmoides*	0.32 (0.00, 0.92)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Notropis procne	*Notropis hudsonius*	1.37 (0.34, 2.40)	0.23
	Open water × *Notropis hudsonius*	− 1.26 (− 4.76, 0.00)	0.19
	Cyprinella analostana	1.26 (0.00, 3.05)	0.19
	Open water × *Cyprinella analostana*	− 0.82 (− 4.00, 0.00)	0.08
	Annual precipitation × *Ictalurus punctatus*	0.45 (0.00, 1.70)	0.03
	Open water × *Semotilus atromaculatus*	− 0.38 (− 2.03, 0.00)	0.02
	Open water × *Fundulus diaphanus*	− 0.38 (− 2.19, − 0.00)	0.02
	Open water × *Cyprinella spiloptera*	− 0.37 (− 2.84, 0.00)	0.02
	Forest (evergreen) × *Ictalurus punctatus*	0.33 (0.00, 1.40)	0.01
	Noturus insignis	0.29 (0.00, 0.91)	0.01
Notropis rubellus	*Hypentelium nigricans*	1.50 (0.84, 2.19)	0.23
	Nocomis micropogon	1.46 (0.81, 2.09)	0.22
	Etheostoma blennioides	1.13 (0.44, 1.82)	0.13
	Notropis volucellus	1.03 (0.38, 1.66)	0.11
	Semotilus corporalis	0.88 (0.22, 1.53)	0.08
	Luxilus cornutus	0.59 (0.00, 1.33)	0.04
	Notropis hudsonius	0.50 (0.00, 1.08)	0.03
Notropis volucellus	*Etheostoma zonale*	2.14 (1.47, 2.87)	0.36
	Open water × *Campostoma anomalum*	− 1.20 (− 2.93, 0.00)	0.11
	Notropis rubellus	1.03 (0.38, 1.66)	0.08
	Campostoma anomalum	0.98 (0.34, 1.65)	0.08
	Pimephales notatus	0.97 (0.25, 1.73)	0.07
	Ictalurus punctatus	0.76 (0.00, 1.70)	0.04
	Forest (evergreen) × *Etheostoma zonale*	− 0.74 (− 1.51, − 0.06)	0.04
	Cyprinella spiloptera	0.73 (0.10, 1.40)	0.04
	Semotilus corporalis	0.61 (0.01, 1.31)	0.03
	Etheostoma blennioides	0.45 (0.00, 1.13)	0.02
	Lepomis cyanellus	0.42 (0.00, 0.97)	0.01
	Fundulus diaphanus	0.37 (0.00, 0.99)	0.01
Noturus insignis	*Exoglossum maxillingua*	1.47 (0.81, 2.15)	0.23
	Nocomis micropogon	1.36 (0.71, 2.01)	0.20
	Rhinichthys cataractae	0.94 (0.20, 1.63)	0.09
	Campostoma anomalum	0.74 (0.09, 1.31)	0.06
	Luxilus cornutus	0.70 (0.16, 1.27)	0.05
	Open water × *Nocomis micropogon*	− 0.68 (− 2.08, 0.00)	0.05
	Ambloplites rupestris	0.64 (0.12, 1.20)	0.04
	Hypentelium nigricans	0.58 (0.00, 1.27)	0.04
	Developed × *Campostoma anomalum*	− 0.53 (− 1.69, 0.00)	0.03
	Stream temperature × *Etheostoma flabellare*	0.50 (0.00, 1.35)	0.03
	Micropterus dolomieu	0.45 (0.00, 1.08)	0.02
	Etheostoma flabellare	0.44 (0.00, 1.02)	0.02

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Oncorhynchus mykiss	Drainage area × *Salmo trutta*	− 2.74 (− 6.03, 0.00)	0.41
	Salmo trutta	1.85 (1.00, 3.20)	0.19
	Drainage area × *Lepomis macrochirus*	− 1.48 (− 3.64, 0.00)	0.12
	Salvelinus fontinalis	1.15 (0.09, 2.16)	0.07
	Open water × *Salmo trutta*	− 0.61 (− 2.71, 0.00)	0.02
	Lepomis macrochirus	0.56 (0.00, 1.49)	0.02
	Open water × *Salvelinus fontinalis*	− 0.54 (− 2.78, 0.00)	0.02
	Stream temperature × *Lepomis macrochirus*	− 0.52 (− 1.29, 0.00)	0.01
	Herbaceous × *Salmo trutta*	− 0.49 (− 1.10, 0.00)	0.01
	Exoglossum maxillingua	0.48 (0.00, 1.08)	0.01
Perca flavescens	Sandy soil × *Cyprinus carpio*	− 1.01 (− 3.32, − 0.00)	0.24
	Micropterus salmoides	0.93 (0.17, 1.68)	0.20
	Agriculture × *Dorosoma cepedianum*	0.63 (0.00, 1.61)	0.09
	Lepomis macrochirus	0.45 (0.00, 1.05)	0.05
	Cyprinus carpio	0.38 (0.00, 1.22)	0.03
	Dorosoma cepedianum	0.37 (0.00, 1.27)	0.03
	Slope × *Esox niger*	0.36 (0.00, 1.38)	0.03
	Lepomis gibbosus	0.35 (0.00, 0.95)	0.03
	Exoglossum maxillingua	− 0.34 (− 0.99, 0.00)	0.03
	Esox niger	0.33 (0.00, 1.08)	0.03
	Sander vitreus	0.29 (0.00, 0.97)	0.02
	Stream temperature × *Sander vitreus*	0.25 (0.00, 0.93)	0.01
	Nitrate deposition × *Cyprinus carpio*	− 0.25 (− 0.95, 0.00)	0.01
	Annual precipitation × *Esox niger*	− 0.22 (− 0.82, 0.00)	0.01
	Catostomus commersonii	0.22 (0.00, 0.85)	0.01
Percina peltata	*Etheostoma olmstedi*	1.21 (0.50, 1.93)	0.24
	Etheostoma zonale	0.84 (0.16, 1.61)	0.11
	Semotilus corporalis	0.84 (0.19, 1.47)	0.11
	Micropterus dolomieu	0.80 (0.12, 1.49)	0.10
	Cyprinus carpio	0.67 (0.00, 1.46)	0.07
	Cyprinella spiloptera	0.55 (0.00, 1.23)	0.05
	Forest (evergreen) × *Etheostoma zonale*	− 0.35 (− 0.96, 0.00)	0.02
	Nocomis micropogon	0.34 (0.00, 0.89)	0.02
	Notropis volucellus	0.33 (0.00, 0.95)	0.02
	Herbaceous × *Cyprinus carpio*	− 0.33 (− 1.01, 0.00)	0.02
	Forest (evergreen) × *Cyprinella spiloptera*	0.33 (0.00, 0.89)	0.02
	Etheostoma blennioides	0.27 (0.00, 0.85)	0.01
	Drainage area × *Rhinichthys cataractae*	− 0.27 (− 1.47, 0.00)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Pimephales notatus	*Cyprinella spiloptera*	1.40 (0.70, 2.13)	0.18
	Etheostoma olmstedi	1.12 (0.49, 1.72)	0.12
	Open water × *Luxilus cornutus*	− 1.11 (− 2.38, 0.00)	0.12
	Notropis volucellus	0.97 (0.25, 1.73)	0.09
	Ambloplites rupestris	0.96 (0.26, 1.54)	0.09
	Luxilus cornutus	0.84 (0.27, 1.42)	0.07
	Lepomis macrochirus	0.77 (0.17, 1.32)	0.06
	Semotilus corporalis	0.63 (0.03, 1.25)	0.04
	Notropis hudsonius	0.57 (0.00, 1.38)	0.03
	Micropterus salmoides	0.50 (0.00, 1.07)	0.02
	Agriculture × *Fundulus diaphanus*	0.38 (0.00, 1.20)	0.01
	Forest (evergreen) × *Micropterus salmoides*	0.36 (0.00, 0.92)	0.01
	Lepomis auritus	0.35 (0.00, 1.01)	0.01
	Lepomis cyanellus	0.34 (0.00, 0.87)	0.01
	Campostoma anomalum	0.34 (0.00, 0.98)	0.01
Pimephales promelas	Drainage area × *Semotilus atromaculatus*	− 1.18 (− 3.30, 0.00)	0.27
	Salvelinus fontinalis	0.96 (0.00, 2.14)	0.17
	Open water × *Ambloplites rupestris*	− 0.78 (− 3.75, − 0.00)	0.12
	Semotilus atromaculatus	0.64 (0.10, 1.42)	0.08
	Sandy soil × *Notemigonus crysoleucas*	− 0.56 (− 2.43, 0.00)	0.06
	Open water × *Salvelinus fontinalis*	0.39 (0.00, 2.03)	0.03
	Wetland × *Oncorhynchus mykiss*	0.34 (0.00, 1.83)	0.02
	Open water × *Luxilus cornutus*	− 0.32 (− 1.76, 0.00)	0.02
	Open water × *Rhinichthys cataractae*	0.28 (0.00, 2.08)	0.02
	Open water × *Notemigonus crysoleucas*	− 0.28 (− 1.92, 0.00)	0.01
	Sandy soil × *Salvelinus fontinalis*	− 0.28 (− 1.16, 0.00)	0.01
	Open water × *Semotilus atromaculatus*	− 0.26 (− 1.27, 0.00)	0.01
	Agriculture × *Notemigonus crysoleucas*	0.26 (0.00, 1.14)	0.01
	Notemigonus crysoleucas	0.24 (0.00, 0.92)	0.01
Pylodictis olivaris	Sandy soil × *Notropis amoenus*	− 1.23 (− 4.62, 0.00)	0.29
	Forest (mixed) × *Ictalurus punctatus*	− 0.92 (− 2.48, 0.00)	0.16
	Ictalurus punctatus	0.80 (0.00, 2.08)	0.12
	Agriculture × *Ictalurus punctatus*	0.68 (0.00, 2.28)	0.09
	Sandy soil × *Ictalurus punctatus*	− 0.48 (− 2.64, 0.00)	0.04
	Notropis amoenus	0.48 (0.00, 1.37)	0.04
	Stream temperature × *Notropis amoenus*	− 0.40 (− 1.35, − 0.00)	0.03
	Moxostoma macrolepidotum	0.37 (0.00, 1.35)	0.03
	Forest (evergreen) × *Ictalurus punctatus*	− 0.32 (− 1.38, − 0.00)	0.02
	Forest (evergreen) × *Notropis amoenus*	− 0.31 (− 1.36, − 0.00)	0.02
	Etheostoma blennioides	0.28 (0.00, 0.98)	0.02
	Wetland × *Notropis amoenus*	− 0.28 (− 1.74, − 0.00)	0.01
	Noturus insignis	− 0.27 (− 0.86, 0.00)	0.01
	Agriculture × *Moxostoma macrolepidotum*	0.23 (0.00, 1.21)	0.01
	Wetland × *Dorosoma cepedianum*	− 0.23 (− 1.86, − 0.00)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Rhinichthys atratulus	Drainage area × *Semotilus atromaculatus*	– 3.50 (– 6.82, – 0.00)	0.37
	Drainage area × *Rhinichthys cataractae*	– 2.66 (– 11.61, 0.00)	0.21
	Rhinichthys cataractae	2.53 (0.00, 4.17)	0.19
	Drainage area × *Salmo trutta*	– 1.15 (– 3.39, 0.00)	0.04
	Catostomus commersonii	1.01 (0.37, 1.68)	0.03
	Semotilus atromaculatus	0.90 (0.00, 1.89)	0.02
	Cottus spp.	0.67 (0.06, 1.27)	0.01
	Sander vitreus	– 0.64 (– 1.45, 0.00)	0.01
	Open water × *Salmo trutta*	– 0.59 (– 2.38, 0.00)	0.01
Rhinichthys cataractae	Drainage area × *Exoglossum maxillingua*	– 3.93 (– 5.50, – 2.49)	0.38
	Drainage area × *Rhinichthys atratulus*	– 2.66 (– 11.61, 0.00)	0.17
	Rhinichthys atratulus	2.53 (0.00, 4.17)	0.16
	Campostoma anomalum	2.19 (1.45, 2.94)	0.12
	Exoglossum maxillingua	1.04 (0.56, 1.70)	0.03
	Catostomus commersonii	1.03 (0.32, 1.83)	0.03
	Open water × *Nocomis micropogon*	– 0.96 (– 3.60, 0.00)	0.02
	Noturus insignis	0.94 (0.20, 1.63)	0.02
	Open water × *Campostoma anomalum*	– 0.86 (– 2.54, 0.00)	0.02
	Drainage area × *Semotilus atromaculatus*	– 0.69 (– 2.28, 0.00)	0.01
Salmo trutta	Drainage area × *Oncorhynchus mykiss*	– 2.74 (– 6.03, 0.00)	0.30
	Drainage area × *Salvelinus fontinalis*	– 2.39 (– 5.88, – 0.00)	0.23
	Oncorhynchus mykiss	1.85 (1.00, 3.20)	0.14
	Salvelinus fontinalis	1.22 (0.00, 2.81)	0.06
	Drainage area × *Rhinichthys atratulus*	– 1.15 (– 3.39, 0.00)	0.05
	Open water × *Salvelinus fontinalis*	– 0.89 (– 3.55, 0.00)	0.03
	Catostomus commersonii	0.84 (0.14, 1.74)	0.03
	Drainage area × *Rhinichthys cataractae*	– 0.63 (– 1.88, 0.00)	0.02
	Open water × *Oncorhynchus mykiss*	– 0.61 (– 2.71, 0.00)	0.02
	Open water × *Rhinichthys atratulus*	– 0.59 (– 2.38, 0.00)	0.01
	Rhinichthys atratulus	0.57 (0.00, 1.41)	0.01
	Drainage area × *Catostomus commersonii*	– 0.51 (– 2.62, 0.00)	0.01
Salvelinus fontinalis	Drainage area × *Salmo trutta*	– 2.39 (– 5.88, – 0.00)	0.40
	Stream temperature	– 1.23 (– 1.84, – 0.71)	0.11
	Salmo trutta	1.22 (0.00, 2.81)	0.10
	Oncorhynchus mykiss	1.15 (0.09, 2.16)	0.09
	Pimephales promelas	0.96 (0.00, 2.14)	0.06
	Open water × *Salmo trutta*	– 0.89 (– 3.55, 0.00)	0.06
	Agriculture	– 0.72 (– 1.20, – 0.33)	0.04
	Open water × *Oncorhynchus mykiss*	– 0.54 (– 2.78, 0.00)	0.02
	Open water × *Pimephales promelas*	0.39 (0.00, 2.03)	0.01

Table 6 (continued)

Species	Key coefficient	Mean value (90% CI)	Rel. importance
Sander vitreus	*Carpiodes cyprinus*	1.43 (0.12, 2.85)	0.23
	Lepomis auritus	0.98 (0.17, 1.77)	0.11
	Hypentelium nigricans	0.95 (0.14, 1.82)	0.10
	Cyprinus carpio	0.91 (0.00, 1.83)	0.09
	Ictalurus punctatus	0.84 (0.00, 1.83)	0.08
	Rhinichthys atratulus	− 0.64 (− 1.45, 0.00)	0.05
	Ambloplites rupestris	0.59 (0.00, 1.39)	0.04
	Cottus spp.	− 0.56 (− 1.42, 0.00)	0.03
	Sandy soil × *Ictalurus punctatus*	− 0.53 (− 1.86, 0.00)	0.03
	Micropterus dolomieu	0.40 (0.00, 1.18)	0.02
	Annual precipitation × *Carpiodes cyprinus*	− 0.37 (− 1.17, 0.00)	0.02
	Sandy soil × *Cyprinus carpio*	− 0.35 (− 1.59, − 0.00)	0.01
	Herbaceous × *Lepomis auritus*	− 0.35 (− 1.00, 0.00)	0.01
	Forest (mixed) × *Dorosoma cepedianum*	− 0.34 (− 1.44, − 0.00)	0.01
	Sandy soil × *Nocomis micropogon*	− 0.33 (− 1.85, 0.00)	0.01
	Sandy soil × *Micropterus dolomieu*	− 0.30 (− 1.61, 0.00)	0.01
Semotilus atromaculatus	Drainage area × *Clinostomus funduloides*	− 3.83 (− 6.17, − 1.14)	0.37
	Drainage area × *Rhinichthys atratulus*	− 3.50 (− 6.82, − 0.00)	0.31
	Clinostomus funduloides	1.36 (0.19, 2.47)	0.05
	Open water × *Clinostomus funduloides*	− 1.33 (− 4.86, − 0.00)	0.04
	Drainage area × *Pimephales promelas*	− 1.18 (− 3.30, 0.00)	0.04
	Rhinichthys atratulus	0.90 (0.00, 1.89)	0.02
	Campostoma anomalum	0.73 (0.18, 1.32)	0.01
	Drainage area × *Rhinichthys cataractae*	− 0.69 (− 2.28, 0.00)	0.01
	Annual precipitation × *Campostoma anomalum*	− 0.66 (− 1.13, − 0.20)	0.01
	Open water × *Luxilus cornutus*	− 0.66 (− 2.02, 0.00)	0.01
	Catostomus commersonii	0.64 (0.12, 1.21)	0.01
	Pimephales promelas	0.64 (0.10, 1.42)	0.01
Semotilus corporalis	*Cyprinella spiloptera*	1.29 (0.62, 2.02)	0.21
	Micropterus dolomieu	0.91 (0.30, 1.54)	0.10
	Notropis rubellus	0.88 (0.22, 1.53)	0.10
	Percina peltata	0.84 (0.19, 1.47)	0.09
	Forest (evergreen) × *Hypentelium nigricans*	0.75 (0.00, 1.47)	0.07
	Pimephales notatus	0.63 (0.03, 1.25)	0.05
	Hypentelium nigricans	0.61 (0.02, 1.19)	0.05
	Notropis volucellus	0.61 (0.01, 1.31)	0.05
	Ambloplites rupestris	0.43 (0.00, 1.02)	0.02
	Nocomis micropogon	0.41 (0.00, 0.98)	0.02
	Notropis amoenus	0.40 (0.00, 1.05)	0.02
	Notropis hudsonius	0.37 (0.00, 1.03)	0.02
	Annual precipitation × *Notropis hudsonius*	− 0.36 (− 0.92, − 0.00)	0.02
	Anguilla rostrata	0.34 (0.00, 0.97)	0.01
	Micropterus salmoides	0.31 (0.00, 0.87)	0.01
	Exoglossum maxillingua	0.31 (0.00, 0.86)	0.01

Mean value represents mean value of coefficient estimates across all bootstraps. Ninety percent CI represents the 90% bootstrapped confidence interval. Relative importance was calculated based on absolute effect size against all nonzero coefficients within that species

Table 7 List of key coefficients for all species within the Ohio River watershed estimated from the bootstrapped conditional random fields model

Species	Key coefficient	Mean (90% CI)	RI
Ambloplites rupestris	*Micropterus dolomieu*	1.36 (0.40, 2.38)	0.42
	Drainage area × *Oncorhynchus mykiss*	− 0.72 (− 2.99, 0.00)	0.12
	Hypentelium nigricans	0.50 (0.00, 1.23)	0.06
	Notropis rubellus	0.47 (0.00, 1.25)	0.05
	Herbaceous × *Micropterus dolomieu*	− 0.46 (− 1.26, 0.00)	0.05
	Campostoma anomalum	0.43 (0.00, 1.20)	0.04
	Nocomis micropogon	0.42 (0.00, 1.19)	0.04
	Nitrate deposition × *Luxilus chrysocephalus*	0.40 (0.00, 1.40)	0.04
	Lepomis gibbosus	0.34 (0.00, 1.01)	0.03
	Oncorhynchus mykiss	0.29 (0.00, 0.97)	0.02
	Moxostoma erythrurum	0.29 (0.00, 0.98)	0.02
	Luxilus chrysocephalus	0.28 (0.00, 1.03)	0.02
	Drainage area × *Lepomis cyanellus*	− 0.26 (− 1.57, 0.00)	0.02
Ameiurus natalis	*Etheostoma nigrum*	0.38 (0.00, 1.18)	0.13
	Notropis stramineus	0.36 (0.00, 1.41)	0.12
	Drainage area × *Etheostoma nigrum*	− 0.34 (− 2.07, 0.00)	0.11
	Herbaceous × *Percina macrocephala*	0.32 (0.00, 1.55)	0.09
	Drainage area × *Notropis stramineus*	− 0.28 (− 2.00, − 0.00)	0.07
	Percina macrocephala	0.27 (0.00, 1.39)	0.07
	Drainage area × *Percina macrocephala*	− 0.25 (− 2.02, 0.00)	0.06
	Forest (mixed) × *Percina macrocephala*	− 0.23 (− 1.48, 0.00)	0.05
	Drainage area × *Percopsis omiscomaycus*	− 0.23 (− 2.01, 0.00)	0.05
	Herbaceous × *Notropis stramineus*	− 0.21 (− 1.16, 0.00)	0.04
	Herbaceous × *Micropterus salmoides*	− 0.17 (− 1.35, − 0.00)	0.03
	Forest (mixed) × *Micropterus salmoides*	− 0.15 (− 1.06, − 0.00)	0.02
	Drainage area × *Ameiurus nebulosus*	− 0.12 (− 0.59, 0.00)	0.01
	Drainage area × *Pimephales notatus*	− 0.12 (− 0.70, 0.00)	0.01
Ameiurus nebulosus	Drainage area × *Catostomus commersonii*	− 0.68 (− 3.90, 0.00)	0.25
	Drainage area × *Lepomis gibbosus*	− 0.51 (− 2.81, 0.00)	0.14
	Micropterus salmoides	0.40 (0.00, 1.44)	0.09
	Slope × *Lepomis gibbosus*	0.37 (0.00, 1.15)	0.07
	Drainage area × *Pimephales notatus*	− 0.35 (− 2.08, 0.00)	0.07
	Lepomis gibbosus	0.29 (0.00, 1.12)	0.05
	Catostomus commersonii	0.26 (0.00, 1.17)	0.04
	Drainage area × *Luxilus cornutus*	− 0.23 (− 1.83, − 0.00)	0.03
	Lepomis macrochirus	0.23 (0.00, 0.89)	0.03
	Drainage area × *Hypentelium nigricans*	− 0.21 (− 1.60, 0.00)	0.02
	Drainage area × *Campostoma anomalum*	− 0.16 (− 1.12, 0.00)	0.01
	Drainage area × *Lethenteron appendix*	− 0.16 (− 1.31, 0.00)	0.01
	Herbaceous × *Lepomis gibbosus*	0.15 (0.00, 0.76)	0.01
	Slope × *Hypentelium nigricans*	0.15 (0.00, 0.79)	0.01
	Agriculture × *Lepomis gibbosus*	− 0.15 (− 0.86, 0.00)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Campostoma anomalum	Drainage area × *Etheostoma blennioides*	− 1.06 (− 3.27, − 0.00)	0.16
	Etheostoma caeruleum	0.99 (0.25, 1.75)	0.14
	Pimephales notatus	0.95 (0.25, 1.65)	0.13
	Nocomis micropogon	0.83 (0.00, 1.64)	0.10
	Luxilus cornutus	0.83 (0.01, 1.76)	0.10
	Drainage area × *Semotilus atromaculatus*	− 0.61 (− 2.51, 0.00)	0.05
	Etheostoma flabellare	0.53 (0.00, 1.26)	0.04
	Etheostoma blennioides	0.48 (0.00, 1.21)	0.03
	Wetland × *Luxilus cornutus*	0.45 (0.00, 1.17)	0.03
	Luxilus chrysocephalus	0.45 (0.00, 1.40)	0.03
	Ambloplites rupestris	0.43 (0.00, 1.20)	0.03
	Hypentelium nigricans	0.39 (0.00, 1.07)	0.02
	Drainage area × *Luxilus cornutus*	− 0.38 (− 2.58, 0.00)	0.02
	Developed × *Luxilus cornutus*	− 0.37 (− 1.35, 0.00)	0.02
	Drainage area × *Rhinichthys cataractae*	− 0.33 (− 1.98, 0.00)	0.02
	Semotilus atromaculatus	0.28 (0.00, 0.87)	0.01
Catostomus commersonii	*Etheostoma nigrum*	1.48 (0.69, 2.24)	0.20
	Drainage area × *Semotilus atromaculatus*	− 1.40 (− 4.65, 0.00)	0.18
	Hypentelium nigricans	1.37 (0.58, 2.19)	0.17
	Semotilus atromaculatus	1.33 (0.36, 2.27)	0.16
	Drainage area × *Rhinichthys obtusus*	− 0.85 (− 3.16, 0.00)	0.07
	Drainage area × *Ameiurus nebulosus*	− 0.68 (− 3.90, 0.00)	0.04
	Rhinichthys obtusus	0.56 (0.00, 1.40)	0.03
	Cottus spp.	0.52 (0.00, 1.30)	0.03
	Drainage area × *Clinostomus elongatus*	− 0.39 (− 2.22, − 0.00)	0.01
	Lepomis gibbosus	0.38 (0.00, 1.09)	0.01
	Percina maculata	0.35 (0.00, 1.09)	0.01
	Pimephales notatus	0.35 (0.00, 0.92)	0.01
Clinostomus elongatus	Drainage area × *Semotilus atromaculatus*	− 1.60 (− 4.19, − 0.00)	0.46
	Drainage area × *Luxilus cornutus*	− 0.72 (− 4.39, − 0.00)	0.09
	Developed × *Luxilus cornutus*	− 0.61 (− 3.34, − 0.00)	0.07
	Drainage area × *Rhinichthys cataractae*	− 0.57 (− 2.46, 0.00)	0.06
	Drainage area × *Etheostoma nigrum*	− 0.47 (− 2.37, 0.00)	0.04
	Drainage area × *Catostomus commersonii*	− 0.39 (− 2.22, − 0.00)	0.03
	Etheostoma flabellare	0.38 (0.00, 1.06)	0.03
	Drainage area × *Rhinichthys obtusus*	− 0.36 (− 2.61, − 0.00)	0.02
	Stream temperature × *Etheostoma nigrum*	− 0.35 (− 1.09, − 0.00)	0.02
	Etheostoma zonale	− 0.35 (− 1.05, 0.00)	0.02
	Etheostoma caeruleum	0.33 (0.00, 1.20)	0.02
	Etheostoma nigrum	0.30 (0.00, 0.94)	0.02
	Micropterus dolomieu	− 0.28 (− 0.95, 0.00)	0.01
	Drainage area × *Etheostoma zonale*	− 0.26 (− 1.45, − 0.00)	0.01
	Semotilus atromaculatus	0.24 (0.00, 0.77)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Cottus spp.	Drainage area × *Rhinichthys obtusus*	− 1.84 (− 4.88, 0.00)	0.44
	Cyprinella spiloptera	− 1.22 (− 2.16, − 0.38)	0.19
	Rhinichthys obtusus	0.85 (0.23, 1.73)	0.09
	Stream temperature	− 0.59 (− 1.23, 0.00)	0.04
	Catostomus commersonii	0.52 (0.00, 1.30)	0.04
	Etheostoma flabellare	0.39 (0.00, 1.16)	0.02
	Drainage area × *Ichthyomyzon bdellium*	0.39 (0.00, 1.21)	0.02
	Drainage area × *Rhinichthys cataractae*	− 0.37 (− 1.81, 0.00)	0.02
	Ichthyomyzon bdellium	0.34 (0.00, 1.25)	0.01
	Etheostoma caeruleum	0.34 (0.00, 1.05)	0.01
Cyprinella spiloptera	*Notropis photogenis*	1.70 (0.30, 3.14)	0.37
	Cottus spp.	− 1.22 (− 2.16, − 0.38)	0.19
	Micropterus dolomieu	0.90 (0.00, 1.85)	0.10
	Drainage area × *Notropis stramineus*	− 0.68 (− 4.50, 0.00)	0.06
	Notropis stramineus	0.68 (0.00, 1.89)	0.06
	Percina caprodes	0.66 (0.00, 1.66)	0.05
	Forest (evergreen) × *Notropis stramineus*	− 0.61 (− 2.02, − 0.00)	0.05
	Notropis volucellus	0.38 (0.00, 1.12)	0.02
	Slope × *Notropis photogenis*	0.34 (0.00, 1.18)	0.01
	Sandy soil × *Notropis photogenis*	0.32 (0.00, 1.46)	0.01
	Agriculture × *Notropis photogenis*	− 0.29 (− 1.00, 0.00)	0.01
	Forest (evergreen) × *Micropterus dolomieu*	− 0.28 (− 1.64, 0.00)	0.01
Erimystax dissimilis	Developed × *Percina macrocephala*	− 1.63 (− 7.07, − 0.00)	0.58
	Notropis photogenis	0.69 (0.00, 2.08)	0.10
	Nitrate deposition × *Percina macrocephala*	0.49 (0.00, 2.12)	0.05
	Percina macrocephala	0.44 (0.00, 1.98)	0.04
	Sandy soil × *Notropis volucellus*	0.40 (0.00, 2.13)	0.04
	Sandy soil × *Notropis photogenis*	− 0.39 (− 2.29, 0.00)	0.03
	Etheostoma zonale	0.38 (0.00, 1.57)	0.03
	Wetland × *Percina macrocephala*	0.29 (0.00, 2.02)	0.02
	Drainage area × *Percina macrocephala*	0.26 (0.00, 1.22)	0.01
	Etheostoma variatum	0.25 (0.00, 1.13)	0.01
Etheostoma blennioides	*Etheostoma zonale*	3.35 (2.35, 4.33)	0.50
	Etheostoma caeruleum	1.69 (0.87, 2.56)	0.13
	Nocomis micropogon	1.51 (0.45, 2.59)	0.10
	Etheostoma flabellare	1.18 (0.39, 2.05)	0.06
	Drainage area × *Campostoma anomalum*	− 1.06 (− 3.27, − 0.00)	0.05
	Percina maculata	0.78 (0.00, 1.69)	0.03
	Notropis volucellus	0.69 (0.00, 1.66)	0.02
	Hypentelium nigricans	0.59 (0.00, 1.27)	0.02
	Notropis rubellus	0.54 (0.00, 1.50)	0.01
	Micropterus dolomieu	0.52 (0.00, 1.56)	0.01
	Campostoma anomalum	0.48 (0.00, 1.21)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Etheostoma caeruleum	*Etheostoma blennioides*	1.69 (0.87, 2.56)	0.43
	Campostoma anomalum	0.99 (0.25, 1.75)	0.15
	Notropis rubellus	0.80 (0.04, 1.65)	0.10
	Lepomis macrochirus	0.58 (0.00, 1.32)	0.05
	Etheostoma flabellare	0.58 (0.00, 1.30)	0.05
	Drainage area × *Rhinichthys cataractae*	− 0.44 (− 2.15, 0.00)	0.03
	Cottus spp.	0.34 (0.00, 1.05)	0.02
	Developed × *Percina maculata*	− 0.34 (− 1.43, 0.00)	0.02
	Developed × *Rhinichthys cataractae*	− 0.33 (− 1.64, 0.00)	0.02
	Clinostomus elongatus	0.33 (0.00, 1.20)	0.02
	Rhinichthys cataractae	0.29 (0.00, 0.86)	0.01
	Percina maculata	0.28 (0.00, 0.98)	0.01
	Drainage area × *Notropis stramineus*	− 0.27 (− 1.93, 0.00)	0.01
Etheostoma flabellare	Drainage area × *Semotilus atromaculatus*	− 3.50 (− 6.08, 0.00)	0.68
	Etheostoma blennioides	1.18 (0.39, 2.05)	0.08
	Hypentelium nigricans	0.71 (0.00, 1.43)	0.03
	Semotilus atromaculatus	0.60 (0.00, 1.51)	0.02
	Etheostoma caeruleum	0.58 (0.00, 1.30)	0.02
	Campostoma anomalum	0.53 (0.00, 1.26)	0.02
	Drainage area × *Luxilus cornutus*	− 0.51 (− 2.52, 0.00)	0.01
	Drainage area × *Salmo trutta*	− 0.49 (− 2.30, 0.00)	0.01
	Pimephales notatus	0.44 (0.00, 1.21)	0.01
Etheostoma nigrum	*Pimephales notatus*	1.90 (1.02, 2.80)	0.28
	Drainage area × *Semotilus atromaculatus*	− 1.79 (− 4.08, − 0.00)	0.25
	Catostomus commersonii	1.48 (0.69, 2.24)	0.17
	Percina maculata	1.26 (0.43, 2.14)	0.12
	Semotilus atromaculatus	0.50 (0.00, 1.33)	0.02
	Drainage area × *Clinostomus elongatus*	− 0.47 (− 2.37, 0.00)	0.02
	Drainage area × *Luxilus cornutus*	− 0.41 (− 2.12, 0.00)	0.01
	Ameiurus natalis	0.38 (0.00, 1.18)	0.01
	Etheostoma flabellare	0.37 (0.00, 1.08)	0.01
Etheostoma variatum	*Etheostoma zonale*	2.67 (1.64, 3.86)	0.43
	Percina macrocephala	1.98 (0.73, 3.15)	0.24
	Micropterus dolomieu	1.79 (0.72, 2.94)	0.20
	Notropis rubellus	0.87 (0.07, 1.84)	0.05
	Etheostoma blennioides	0.42 (0.00, 1.43)	0.01
	Developed × *Percina macrocephala*	− 0.42 (− 2.12, 0.00)	0.01
Etheostoma zonale	*Etheostoma blennioides*	3.35 (2.35, 4.33)	0.47
	Etheostoma variatum	2.67 (1.64, 3.86)	0.30
	Noturus flavus	1.22 (0.14, 2.24)	0.06
	Developed × *Noturus flavus*	− 0.99 (− 3.27, 0.00)	0.04
	Ichthyomyzon bdellium	0.90 (0.00, 2.17)	0.03
	Sandy soil × *Ichthyomyzon bdellium*	− 0.58 (− 1.64, − 0.00)	0.01
	Percina macrocephala	0.58 (0.00, 1.53)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Exoglossum laurae	Drainage area × *Luxilus cornutus*	− 1.49 (− 6.18, − 0.00)	0.60
	Nocomis micropogon	0.80 (0.00, 1.96)	0.17
	Wetland × *Nocomis micropogon*	0.50 (0.00, 1.62)	0.07
	Luxilus cornutus	0.40 (0.00, 1.42)	0.04
	Open water × *Nocomis micropogon*	− 0.24 (− 1.94, 0.00)	0.02
	Developed × *Luxilus cornutus*	− 0.22 (− 1.15, 0.00)	0.01
	Open water × *Luxilus cornutus*	− 0.21 (− 1.60, − 0.00)	0.01
Hypentelium nigricans	*Catostomus commersonii*	1.37 (0.58, 2.19)	0.22
	Nocomis micropogon	1.11 (0.19, 2.02)	0.14
	Pimephales notatus	0.80 (0.04, 1.58)	0.07
	Etheostoma flabellare	0.71 (0.00, 1.43)	0.06
	Rhinichthys cataractae	0.68 (0.02, 1.51)	0.05
	Rhinichthys obtusus	0.62 (0.00, 1.50)	0.04
	Notropis rubellus	0.61 (0.00, 1.45)	0.04
	Drainage area × *Rhinichthys cataractae*	− 0.61 (− 2.70, 0.00)	0.04
	Etheostoma blennioides	0.59 (0.00, 1.27)	0.04
	Micropterus dolomieu	0.55 (0.00, 1.39)	0.03
	Ambloplites rupestris	0.50 (0.00, 1.23)	0.03
	Percina caprodes	0.49 (0.00, 1.19)	0.03
	Moxostoma erythrurum	0.41 (0.00, 1.23)	0.02
	Developed × *Rhinichthys cataractae*	− 0.40 (− 1.88, 0.00)	0.02
	Campostoma anomalum	0.39 (0.00, 1.07)	0.02
	Stream temperature × *Rhinichthys obtusus*	0.38 (0.00, 0.76)	0.02
	Luxilus chrysocephalus	0.36 (0.00, 1.10)	0.02
	Herbaceous × *Micropterus dolomieu*	− 0.30 (− 0.90, 0.00)	0.01
Ichthyomyzon bdellium	*Etheostoma zonale*	0.90 (0.00, 2.17)	0.21
	Stream temperature × *Percina macrocephala*	0.59 (0.00, 1.98)	0.09
	Sandy soil × *Etheostoma zonale*	− 0.58 (− 1.64, − 0.00)	0.09
	Rhinichthys cataractae	0.51 (0.00, 1.55)	0.07
	Nocomis micropogon	0.49 (0.00, 1.57)	0.06
	Drainage area × *Semotilus atromaculatus*	− 0.47 (− 2.48, 0.00)	0.06
	Stream temperature × *Rhinichthys cataractae*	0.42 (0.00, 1.25)	0.05
	Sandy soil × *Percina macrocephala*	− 0.42 (− 1.80, − 0.00)	0.05
	Drainage area × *Cottus spp.*	0.39 (0.00, 1.21)	0.04
	Etheostoma variatum	0.35 (0.00, 1.31)	0.03
	Cottus spp.	0.34 (0.00, 1.25)	0.03
	Annual precipitation × *Percina macrocephala*	− 0.34 (− 1.59, − 0.00)	0.03
	Percina macrocephala	0.31 (0.00, 1.41)	0.02
	Semotilus atromaculatus	− 0.26 (− 1.09, 0.00)	0.02
	Lepomis macrochirus	− 0.24 (− 0.89, 0.00)	0.02
	Forest (mixed) × *Rhinichthys cataractae*	− 0.21 (− 1.00, 0.00)	0.01
	Etheostoma blennioides	0.20 (0.00, 0.91)	0.01
	Developed × *Moxostoma erythrurum*	− 0.20 (− 1.61, 0.00)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Lepomis cyanellus	*Lepomis macrochirus*	1.53 (0.47, 2.57)	0.85
	Drainage area × *Ambloplites rupestris*	− 0.26 (− 1.57, 0.00)	0.02
	Drainage area × *Lepomis gibbosus*	− 0.23 (− 1.29, 0.00)	0.02
	Lepomis gibbosus	0.22 (0.00, 0.81)	0.02
	Nitrate deposition × *Lepomis macrochirus*	0.19 (0.00, 0.64)	0.01
	Herbaceous × *Lepomis macrochirus*	− 0.19 (− 0.83, 0.00)	0.01
	Forest (evergreen) × *Lepomis gibbosus*	− 0.17 (− 0.93, 0.00)	0.01
Lepomis gibbosus	*Micropterus salmoides*	0.85 (0.03, 1.72)	0.24
	Lepomis macrochirus	0.83 (0.11, 1.58)	0.24
	Drainage area × *Ameiurus nebulosus*	− 0.51 (− 2.81, 0.00)	0.09
	Catostomus commersonii	0.38 (0.00, 1.09)	0.05
	Slope × *Ameiurus nebulosus*	0.37 (0.00, 1.15)	0.05
	Ambloplites rupestris	0.34 (0.00, 1.01)	0.04
	Ameiurus nebulosus	0.29 (0.00, 1.12)	0.03
	Moxostoma erythrurum	0.24 (0.00, 0.84)	0.02
	Annual precipitation × *Etheostoma nigrum*	− 0.24 (− 0.95, − 0.00)	0.02
	Drainage area × *Lepomis cyanellus*	− 0.23 (− 1.29, 0.00)	0.02
	Lepomis cyanellus	0.22 (0.00, 0.81)	0.02
	Etheostoma nigrum	0.21 (0.00, 0.87)	0.01
	Forest (evergreen) × *Etheostoma flabellare*	− 0.19 (− 1.05, − 0.00)	0.01
	Semotilus atromaculatus	0.19 (0.00, 0.82)	0.01
	Percina maculata	0.19 (0.00, 0.81)	0.01
	Developed × *Lepomis macrochirus*	0.18 (0.00, 0.66)	0.01
	Forest (evergreen) × *Lepomis cyanellus*	− 0.17 (− 0.93, 0.00)	0.01
Lepomis macrochirus	*Lepomis cyanellus*	1.53 (0.47, 2.57)	0.50
	Lepomis gibbosus	0.83 (0.11, 1.58)	0.15
	Etheostoma caeruleum	0.58 (0.00, 1.32)	0.07
	Micropterus salmoides	0.38 (0.00, 1.13)	0.03
	Noturus flavus	− 0.36 (− 1.11, 0.00)	0.03
	Moxostoma erythrurum	0.35 (0.00, 0.98)	0.03
	Forest (evergreen) × *Luxilus chrysocephalus*	0.34 (0.00, 1.10)	0.02
	Luxilus chrysocephalus	0.31 (0.00, 0.91)	0.02
	Ichthyomyzon bdellium	− 0.24 (− 0.89, 0.00)	0.01
	Ameiurus nebulosus	0.23 (0.00, 0.89)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Lethenteron appendix	Drainage area × *Percopsis omiscomaycus*	− 0.33 (− 2.70, − 0.00)	0.16
	Moxostoma erythrurum	0.32 (0.00, 1.19)	0.15
	Developed × *Notropis rubellus*	− 0.23 (− 1.80, − 0.00)	0.08
	Drainage area × *Clinostomus elongatus*	− 0.19 (− 1.48, − 0.00)	0.05
	Percina macrocephala	0.18 (0.00, 0.83)	0.05
	Open water × *Notropis stramineus*	− 0.17 (− 1.32, 0.00)	0.04
	Drainage area × *Etheostoma nigrum*	− 0.16 (− 1.35, 0.00)	0.04
	Notropis stramineus	0.16 (0.00, 0.83)	0.04
	Drainage area × *Ameiurus nebulosus*	− 0.16 (− 1.31, 0.00)	0.04
	Developed × *Notropis stramineus*	− 0.15 (− 1.14, 0.00)	0.03
	Forest (evergreen) × *Hypentelium nigricans*	0.14 (0.00, 1.05)	0.03
	Herbaceous × *Moxostoma erythrurum*	0.14 (0.00, 0.85)	0.03
	Open water × *Percopsis omiscomaycus*	− 0.14 (− 0.48, − 0.00)	0.03
	Hypentelium nigricans	0.14 (0.00, 0.74)	0.03
	Herbaceous × *Percina caprodes*	0.12 (0.00, 0.91)	0.02
	Annual precipitation × *Noturus flavus*	0.11 (0.00, 0.92)	0.02
	Developed × *Moxostoma erythrurum*	− 0.09 (− 0.60, 0.00)	0.01
	Forest (evergreen) × *Percopsis omiscomaycus*	0.08 (0.00, 0.62)	0.01
Lota lota	Nitrate deposition × *Clinostomus elongatus*	− 0.18 (− 1.66, − 0.00)	0.33
	Drainage area × *Etheostoma blennioides*	− 0.08 (− 0.75, 0.00)	0.07
	Forest (evergreen) × *Clinostomus elongatus*	− 0.07 (− 0.47, − 0.00)	0.05
	Rhinichthys obtusus	0.07 (0.00, 0.40)	0.04
	Developed × *Clinostomus elongatus*	− 0.06 (− 0.53, − 0.00)	0.04
	Wetland × *Clinostomus elongatus*	0.06 (0.00, 0.00)	0.04
	Wetland × *Semotilus atromaculatus*	− 0.06 (− 0.00, − 0.00)	0.03
	Drainage area × *Semotilus atromaculatus*	− 0.05 (0.00, 0.00)	0.03
	Drainage area × *Rhinichthys obtusus*	− 0.05 (− 0.08, 0.00)	0.03
	Drainage area × *Clinostomus elongatus*	− 0.05 (− 0.01, 0.00)	0.02
	Developed × *Rhinichthys obtusus*	− 0.05 (− 0.23, 0.00)	0.02
	Ameiurus nebulosus	− 0.05 (− 0.32, 0.00)	0.02
	Wetland × *Pimephales notatus*	0.04 (0.00, 0.00)	0.02
	Clinostomus elongatus	0.04 (0.00, 0.34)	0.02
	Nitrate deposition × *Pimephales notatus*	− 0.04 (− 0.00, 0.00)	0.02
	Developed × *Nocomis micropogon*	− 0.04 (0.00, 0.00)	0.02
	Open water × *Rhinichthys obtusus*	− 0.04 (− 0.21, 0.00)	0.02
	Nitrate deposition × *Lepomis macrochirus*	− 0.04 (− 0.00, 0.00)	0.01
	Ambloplites rupestris	− 0.04 (− 0.27, 0.00)	0.01
	Etheostoma blennioides	0.03 (0.00, 0.30)	0.01
	Drainage area × *Percopsis omiscomaycus*	− 0.03 (0.00, − 0.00)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Luxilus chrysocephalus	Sandy soil × *Noturus flavus*	− 0.70 (− 1.99, − 0.00)	0.19
	Notropis rubellus	0.65 (0.00, 1.50)	0.16
	Campostoma anomalum	0.45 (0.00, 1.40)	0.08
	Semotilus atromaculatus	0.43 (0.00, 1.28)	0.07
	Nitrate deposition × *Ambloplites rupestris*	0.40 (0.00, 1.40)	0.06
	Hypentelium nigricans	0.36 (0.00, 1.10)	0.05
	Forest (evergreen) × *Lepomis macrochirus*	0.34 (0.00, 1.10)	0.05
	Lepomis macrochirus	0.31 (0.00, 0.91)	0.04
	Noturus flavus	0.31 (0.00, 1.08)	0.04
	Ambloplites rupestris	0.28 (0.00, 1.03)	0.03
	Annual precipitation × *Etheostoma nigrum*	0.26 (0.00, 0.94)	0.03
	Stream temperature × *Semotilus atromaculatus*	0.21 (0.00, 0.87)	0.02
	Nocomis micropogon	0.20 (0.00, 0.84)	0.02
	Drainage area × *Salmo trutta*	− 0.18 (− 1.68, 0.00)	0.01
	Etheostoma nigrum	0.17 (0.00, 0.72)	0.01
	Pimephales notatus	0.17 (0.00, 0.74)	0.01
	Agriculture × *Campostoma anomalum*	0.16 (0.00, 0.74)	0.01
Luxilus cornutus	Drainage area × *Exoglossum laurae*	− 1.49 (− 6.18, − 0.00)	0.27
	Drainage area × *Rhinichthys obtusus*	− 0.87 (− 3.39, 0.00)	0.09
	Campostoma anomalum	0.83 (0.01, 1.76)	0.08
	Percina maculata	0.80 (0.04, 1.70)	0.08
	Drainage area × *Clinostomus elongatus*	− 0.72 (− 4.39, − 0.00)	0.06
	Drainage area × *Rhinichthys cataractae*	− 0.66 (− 3.28, − 0.00)	0.05
	Developed × *Clinostomus elongatus*	− 0.61 (− 3.34, − 0.00)	0.05
	Drainage area × *Percina maculata*	− 0.52 (− 2.74, 0.00)	0.03
	Drainage area × *Etheostoma flabellare*	− 0.51 (− 2.52, 0.00)	0.03
	Wetland × *Campostoma anomalum*	0.45 (0.00, 1.17)	0.03
	Etheostoma flabellare	0.42 (0.00, 1.10)	0.02
	Drainage area × *Etheostoma nigrum*	− 0.41 (− 2.12, 0.00)	0.02
	Exoglossum laurae	0.40 (0.00, 1.42)	0.02
	Drainage area × *Campostoma anomalum*	− 0.38 (− 2.58, 0.00)	0.02
	Rhinichthys obtusus	0.38 (0.00, 1.23)	0.02
	Developed × *Campostoma anomalum*	− 0.37 (− 1.35, 0.00)	0.02
	Etheostoma nigrum	0.36 (0.00, 1.02)	0.02
Micropterus dolomieu	*Etheostoma variatum*	1.79 (0.72, 2.94)	0.28
	Ambloplites rupestris	1.36 (0.40, 2.38)	0.16
	Percina caprodes	1.21 (0.12, 2.26)	0.13
	Moxostoma erythrurum	1.14 (0.21, 2.15)	0.11
	Cyprinella spiloptera	0.90 (0.00, 1.85)	0.07
	Noturus flavus	0.67 (0.00, 1.71)	0.04
	Hypentelium nigricans	0.55 (0.00, 1.39)	0.03
	Etheostoma blennioides	0.52 (0.00, 1.56)	0.02
	Herbaceous × *Ambloplites rupestris*	− 0.46 (− 1.26, 0.00)	0.02
	Notropis photogenis	0.44 (0.00, 1.39)	0.02
	Salmo trutta	− 0.38 (− 1.16, 0.00)	0.01
	Drainage area × *Salmo trutta*	− 0.35 (− 2.14, − 0.00)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Micropterus salmoides	*Lepomis gibbosus*	0.85 (0.03, 1.72)	0.38
	Pimephales notatus	0.56 (0.00, 1.36)	0.16
	Ameiurus nebulosus	0.40 (0.00, 1.44)	0.08
	Lepomis macrochirus	0.38 (0.00, 1.13)	0.08
	Sandy soil × *Perca flavescens*	− 0.30 (− 1.85, − 0.00)	0.05
	Forest (evergreen) × *Pimephales notatus*	− 0.25 (− 0.83, 0.00)	0.03
	Sandy soil × *Percina caprodes*	− 0.21 (− 1.37, − 0.00)	0.02
	Forest (mixed) × *Etheostoma nigrum*	− 0.21 (− 0.97, − 0.00)	0.02
	Perca flavescens	0.18 (0.00, 1.01)	0.02
	Herbaceous × *Ameiurus natalis*	− 0.17 (− 1.35, − 0.00)	0.02
	Forest (mixed) × *Ameiurus natalis*	− 0.15 (− 1.06, − 0.00)	0.01
	Notropis volucellus	0.14 (0.00, 0.63)	0.01
	Wetland × *Perca flavescens*	0.14 (0.00, 0.99)	0.01
Moxostoma erythrurum	*Notropis photogenis*	1.53 (0.41, 2.80)	0.34
	Micropterus dolomieu	1.14 (0.21, 2.15)	0.19
	Developed × *Perca flavescens*	0.94 (0.00, 2.37)	0.13
	Pimephales notatus	0.63 (0.00, 1.61)	0.06
	Perca flavescens	0.58 (0.00, 1.82)	0.05
	Drainage area × *Perca flavescens*	0.50 (0.00, 1.51)	0.04
	Hypentelium nigricans	0.41 (0.00, 1.23)	0.02
	Etheostoma zonale	0.40 (0.00, 1.16)	0.02
	Lepomis macrochirus	0.35 (0.00, 0.98)	0.02
	Notropis stramineus	0.33 (0.00, 1.18)	0.02
	Lethenteron appendix	0.32 (0.00, 1.19)	0.02
	Notropis volucellus	0.31 (0.00, 1.14)	0.01
	Percina caprodes	0.29 (0.00, 1.20)	0.01
	Ambloplites rupestris	0.29 (0.00, 0.98)	0.01
Nocomis micropogon	*Etheostoma blennioides*	1.51 (0.45, 2.59)	0.28
	Hypentelium nigricans	1.11 (0.19, 2.02)	0.15
	Campostoma anomalum	0.83 (0.00, 1.64)	0.09
	Notropis rubellus	0.81 (0.06, 1.62)	0.08
	Exoglossum laurae	0.80 (0.00, 1.96)	0.08
	Rhinichthys cataractae	0.57 (0.00, 1.40)	0.04
	Percina maculata	0.56 (0.00, 1.39)	0.04
	Wetland × *Exoglossum laurae*	0.50 (0.00, 1.62)	0.03
	Ichthyomyzon bdellium	0.49 (0.00, 1.57)	0.03
	Ambloplites rupestris	0.42 (0.00, 1.19)	0.02
	Percina macrocephala	0.36 (0.00, 1.09)	0.02
	Stream temperature × *Micropterus dolomieu*	0.30 (0.00, 0.95)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Notropis photogenis	*Cyprinella spiloptera*	1.70 (0.30, 3.14)	0.35
	Moxostoma erythrurum	1.53 (0.41, 2.80)	0.28
	Notropis volucellus	0.74 (0.00, 1.90)	0.07
	Erimystax dissimilis	0.69 (0.00, 2.08)	0.06
	Percina caprodes	0.58 (0.00, 1.66)	0.04
	Notropis rubellus	0.47 (0.00, 1.51)	0.03
	Micropterus dolomieu	0.44 (0.00, 1.39)	0.02
	Sandy soil × *Erimystax dissimilis*	− 0.39 (− 2.29, 0.00)	0.02
	Notropis stramineus	0.39 (0.00, 1.47)	0.02
	Slope × *Cyprinella spiloptera*	0.34 (0.00, 1.18)	0.01
	Sandy soil × *Cyprinella spiloptera*	0.32 (0.00, 1.46)	0.01
	Agriculture × *Cyprinella spiloptera*	− 0.29 (− 1.00, 0.00)	0.01
Notropis rubellus	*Percina caprodes*	1.07 (0.13, 1.95)	0.14
	Notropis volucellus	1.06 (0.24, 1.94)	0.14
	Etheostoma variatum	0.87 (0.07, 1.84)	0.09
	Nocomis micropogon	0.81 (0.06, 1.62)	0.08
	Etheostoma caeruleum	0.80 (0.04, 1.65)	0.08
	Stream temperature × *Noturus flavus*	0.73 (0.00, 1.71)	0.07
	Luxilus chrysocephalus	0.65 (0.00, 1.50)	0.05
	Hypentelium nigricans	0.61 (0.00, 1.45)	0.05
	Noturus flavus	0.61 (0.00, 1.52)	0.05
	Etheostoma blennioides	0.54 (0.00, 1.50)	0.04
	Ambloplites rupestris	0.47 (0.00, 1.25)	0.03
	Notropis photogenis	0.47 (0.00, 1.51)	0.03
	Drainage area × *Notropis stramineus*	− 0.47 (− 2.79, 0.00)	0.03
	Herbaceous × *Notropis volucellus*	0.45 (0.00, 1.05)	0.02
	Open water × *Percina caprodes*	− 0.36 (− 1.18, 0.00)	0.02
	Slope × *Percina caprodes*	0.33 (0.00, 1.01)	0.01
	Etheostoma zonale	0.30 (0.00, 1.07)	0.01
Notropis stramineus	*Notropis volucellus*	0.95 (0.00, 2.24)	0.19
	Drainage area × *Cyprinella spiloptera*	− 0.68 (− 4.50, 0.00)	0.10
	Cyprinella spiloptera	0.68 (0.00, 1.89)	0.10
	Pimephales notatus	0.64 (0.00, 1.57)	0.09
	Forest (evergreen) × *Cyprinella spiloptera*	− 0.61 (− 2.02, − 0.00)	0.08
	Drainage area × *Notropis rubellus*	− 0.47 (− 2.79, 0.00)	0.05
	Percina caprodes	0.44 (0.00, 1.35)	0.04
	Notropis photogenis	0.39 (0.00, 1.47)	0.03
	Ameiurus natalis	0.36 (0.00, 1.41)	0.03
	Moxostoma erythrurum	0.33 (0.00, 1.18)	0.02
	Drainage area × *Pimephales notatus*	− 0.33 (− 2.57, 0.00)	0.02
	Developed × *Notropis volucellus*	0.30 (0.00, 1.42)	0.02
	Drainage area × *Ameiurus natalis*	− 0.28 (− 2.00, − 0.00)	0.02
	Drainage area × *Etheostoma zonale*	− 0.28 (− 2.00, 0.00)	0.02
	Drainage area × *Etheostoma caeruleum*	− 0.27 (− 1.93, 0.00)	0.02
	Forest (evergreen) × *Etheostoma zonale*	− 0.26 (− 1.18, 0.00)	0.01
	Open water × *Notropis photogenis*	− 0.25 (− 1.61, 0.00)	0.01
	Salmo trutta	− 0.24 (− 0.85, 0.00)	0.01
	Forest (mixed) × *Notropis photogenis*	− 0.22 (− 1.07, − 0.00)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Notropis volucellus	*Notropis rubellus*	1.06 (0.24, 1.94)	0.24
	Notropis stramineus	0.95 (0.00, 2.24)	0.19
	Notropis photogenis	0.74 (0.00, 1.90)	0.12
	Etheostoma blennioides	0.69 (0.00, 1.66)	0.10
	Pimephales notatus	0.60 (0.00, 1.69)	0.07
	Herbaceous × *Notropis rubellus*	0.45 (0.00, 1.05)	0.04
	Sandy soil × *Erimystax dissimilis*	0.40 (0.00, 2.13)	0.03
	Cyprinella spiloptera	0.38 (0.00, 1.12)	0.03
	Moxostoma erythrurum	0.31 (0.00, 1.14)	0.02
	Developed × *Notropis stramineus*	0.30 (0.00, 1.42)	0.02
	Stream temperature × *Notropis photogenis*	0.27 (0.00, 1.01)	0.02
	Forest (evergreen) × *Etheostoma blennioides*	− 0.23 (− 1.08, 0.00)	0.01
Noturus flavus	*Etheostoma zonale*	1.22 (0.14, 2.24)	0.23
	Developed × *Etheostoma zonale*	− 0.99 (− 3.27, 0.00)	0.15
	Percina maculata	0.93 (0.08, 1.92)	0.13
	Stream temperature × *Notropis rubellus*	0.73 (0.00, 1.71)	0.08
	Sandy soil × *Luxilus chrysocephalus*	− 0.70 (− 1.99, − 0.00)	0.08
	Micropterus dolomieu	0.67 (0.00, 1.71)	0.07
	Notropis rubellus	0.61 (0.00, 1.52)	0.06
	Lepomis macrochirus	− 0.36 (− 1.11, 0.00)	0.02
	Etheostoma flabellare	0.35 (0.00, 1.20)	0.02
	Stream temperature × *Etheostoma flabellare*	0.33 (0.00, 1.14)	0.02
	Percina caprodes	0.33 (0.00, 1.16)	0.02
	Luxilus chrysocephalus	0.31 (0.00, 1.08)	0.02
Oncorhynchus mykiss	Drainage area × *Salmo trutta*	− 2.08 (− 7.48, 0.00)	0.43
	Salmo trutta	1.99 (0.76, 3.46)	0.39
	Drainage area × *Ambloplites rupestris*	− 0.72 (− 2.99, 0.00)	0.05
	Developed × *Salmo trutta*	− 0.68 (− 2.44, 0.00)	0.05
	Stream temperature × *Salmo trutta*	0.42 (0.00, 1.37)	0.02
Percina caprodes	*Micropterus dolomieu*	1.21 (0.12, 2.26)	0.26
	Notropis rubellus	1.07 (0.13, 1.95)	0.21
	Cyprinella spiloptera	0.66 (0.00, 1.66)	0.08
	Notropis photogenis	0.58 (0.00, 1.66)	0.06
	Hypentelium nigricans	0.49 (0.00, 1.19)	0.04
	Notropis stramineus	0.44 (0.00, 1.35)	0.04
	Perca flavescens	0.37 (0.00, 1.34)	0.03
	Open water × *Notropis rubellus*	− 0.36 (− 1.18, 0.00)	0.02
	Annual precipitation × *Perca flavescens*	0.34 (0.00, 1.14)	0.02
	Noturus flavus	0.33 (0.00, 1.16)	0.02
	Slope × *Notropis rubellus*	0.33 (0.00, 1.01)	0.02
	Percina macrocephala	0.32 (0.00, 1.15)	0.02
	Moxostoma erythrurum	0.29 (0.00, 1.20)	0.02
	Developed × *Micropterus dolomieu*	0.27 (0.00, 1.19)	0.01
	Etheostoma variatum	0.27 (0.00, 1.05)	0.01
	Developed × *Hypentelium nigricans*	0.26 (0.00, 0.69)	0.01
	Drainage area	0.25 (0.00, 0.64)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Perca flavescens	Developed × *Moxostoma erythrurum*	0.94 (0.00, 2.37)	0.37
	Moxostoma erythrurum	0.58 (0.00, 1.82)	0.14
	Drainage area × *Moxostoma erythrurum*	0.50 (0.00, 1.51)	0.10
	Percina caprodes	0.37 (0.00, 1.34)	0.06
	Etheostoma flabellare	− 0.36 (− 1.22, 0.00)	0.05
	Annual precipitation × *Percina caprodes*	0.34 (0.00, 1.14)	0.05
	Sandy soil × *Micropterus salmoides*	− 0.30 (− 1.85, − 0.00)	0.04
	Rhinichthys obtusus	− 0.26 (− 1.11, 0.00)	0.03
	Semotilus atromaculatus	− 0.24 (− 0.93, 0.00)	0.02
	Sandy soil × *Percina caprodes*	− 0.19 (− 1.04, 0.00)	0.02
	Forest (mixed) × *Cyprinella spiloptera*	− 0.19 (− 1.14, − 0.00)	0.02
	Micropterus salmoides	0.18 (0.00, 1.01)	0.01
	Pimephales notatus	0.16 (0.00, 0.85)	0.01
	Slope × *Percina caprodes*	0.16 (0.00, 0.88)	0.01
Percina macrocephala	*Etheostoma variatum*	1.98 (0.73, 3.15)	0.38
	Developed × *Erimystax dissimilis*	− 1.63 (− 7.07, − 0.00)	0.26
	Rhinichthys obtusus	− 0.68 (− 1.53, 0.00)	0.04
	Stream temperature × *Ichthyomyzon bdellium*	0.59 (0.00, 1.98)	0.03
	Etheostoma zonale	0.58 (0.00, 1.53)	0.03
	Nitrate deposition × *Erimystax dissimilis*	0.49 (0.00, 2.12)	0.02
	Drainage area × *Rhinichthys obtusus*	− 0.48 (− 2.30, − 0.00)	0.02
	Erimystax dissimilis	0.44 (0.00, 1.98)	0.02
	Sandy soil × *Ichthyomyzon bdellium*	− 0.42 (− 1.80, − 0.00)	0.02
	Developed × *Etheostoma variatum*	− 0.42 (− 2.12, 0.00)	0.02
	Nocomis micropogon	0.36 (0.00, 1.09)	0.01
	Percina maculata	0.36 (0.00, 1.14)	0.01
	Annual precipitation × *Ichthyomyzon bdellium*	− 0.34 (− 1.59, − 0.00)	0.01
Percina maculata	Drainage area × *Percopsis omiscomaycus*	− 1.35 (− 5.54, 0.00)	0.20
	Etheostoma nigrum	1.26 (0.43, 2.14)	0.18
	Noturus flavus	0.93 (0.08, 1.92)	0.10
	Developed × *Percopsis omiscomaycus*	− 0.88 (− 3.07, − 0.00)	0.09
	Luxilus cornutus	0.80 (0.04, 1.70)	0.07
	Etheostoma blennioides	0.78 (0.00, 1.69)	0.07
	Percopsis omiscomaycus	0.72 (0.00, 1.87)	0.06
	Nocomis micropogon	0.56 (0.00, 1.39)	0.03
	Drainage area × *Luxilus cornutus*	− 0.52 (− 2.74, 0.00)	0.03
	Etheostoma flabellare	0.42 (0.00, 1.24)	0.02
	Pimephales notatus	0.38 (0.00, 1.09)	0.02
	Etheostoma variatum	0.38 (0.00, 1.24)	0.02
	Percina macrocephala	0.36 (0.00, 1.14)	0.01
	Catostomus commersonii	0.35 (0.00, 1.09)	0.01
	Developed × *Etheostoma caeruleum*	− 0.34 (− 1.43, 0.00)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Percopsis omiscomaycus	Drainage area × *Percina maculata*	− 1.35 (− 5.54, 0.00)	0.51
	Developed × *Percina maculata*	− 0.88 (− 3.07, − 0.00)	0.22
	Percina maculata	0.72 (0.00, 1.87)	0.14
	Drainage area × *Lethenteron appendix*	− 0.33 (− 2.70, − 0.00)	0.03
	Drainage area × *Etheostoma nigrum*	− 0.33 (− 2.58, − 0.00)	0.03
	Drainage area × *Ameiurus natalis*	− 0.23 (− 2.01, 0.00)	0.01
Pimephales notatus	*Etheostoma nigrum*	1.90 (1.02, 2.80)	0.44
	Campostoma anomalum	0.95 (0.25, 1.65)	0.11
	Hypentelium nigricans	0.80 (0.04, 1.58)	0.08
	Notropis stramineus	0.64 (0.00, 1.57)	0.05
	Moxostoma erythrurum	0.63 (0.00, 1.61)	0.05
	Notropis volucellus	0.60 (0.00, 1.69)	0.04
	Micropterus salmoides	0.56 (0.00, 1.36)	0.04
	Etheostoma flabellare	0.44 (0.00, 1.21)	0.02
	Percina maculata	0.38 (0.00, 1.09)	0.02
	Drainage area × *Ameiurus nebulosus*	− 0.35 (− 2.08, 0.00)	0.02
	Catostomus commersonii	0.35 (0.00, 0.92)	0.01
	Etheostoma blennioides	0.34 (0.00, 1.03)	0.01
	Drainage area × *Notropis stramineus*	− 0.33 (− 2.57, 0.00)	0.01
	Salvelinus fontinalis	− 0.31 (− 1.08, 0.00)	0.01
Rhinichthys cataractae	Drainage area × *Rhinichthys obtusus*	− 1.19 (− 3.79, 0.00)	0.24
	Hypentelium nigricans	0.68 (0.02, 1.51)	0.08
	Drainage area × *Luxilus cornutus*	− 0.66 (− 3.28, − 0.00)	0.08
	Rhinichthys obtusus	0.66 (0.00, 1.50)	0.07
	Drainage area × *Hypentelium nigricans*	− 0.61 (− 2.70, 0.00)	0.06
	Nocomis micropogon	0.57 (0.00, 1.40)	0.06
	Drainage area × *Clinostomus elongatus*	− 0.57 (− 2.46, 0.00)	0.06
	Ichthyomyzon bdellium	0.51 (0.00, 1.55)	0.04
	Drainage area × *Etheostoma caeruleum*	− 0.44 (− 2.15, 0.00)	0.03
	Stream temperature × *Ichthyomyzon bdellium*	0.42 (0.00, 1.25)	0.03
	Developed × *Hypentelium nigricans*	− 0.40 (− 1.88, 0.00)	0.03
	Drainage area × *Cottus spp.*	− 0.37 (− 1.81, 0.00)	0.02
	Developed × *Etheostoma caeruleum*	− 0.33 (− 1.64, 0.00)	0.02
	Drainage area × *Campostoma anomalum*	− 0.33 (− 1.98, 0.00)	0.02
	Etheostoma caeruleum	0.29 (0.00, 0.86)	0.01
	Cottus spp.	0.26 (0.00, 0.89)	0.01
Rhinichthys obtusus	Drainage area × *Semotilus atromaculatus*	− 4.08 (− 9.38, − 0.00)	0.61
	Drainage area × *Cottus spp.*	− 1.84 (− 4.88, 0.00)	0.12
	Drainage area × *Rhinichthys cataractae*	− 1.19 (− 3.79, 0.00)	0.05
	Drainage area × *Luxilus cornutus*	− 0.87 (− 3.39, 0.00)	0.03
	Cottus spp.	0.85 (0.23, 1.73)	0.03
	Drainage area × *Catostomus commersonii*	− 0.85 (− 3.16, 0.00)	0.03
	Drainage area × *Salmo trutta*	− 0.69 (− 3.31, 0.00)	0.02
	Percina macrocephala	− 0.68 (− 1.53, 0.00)	0.02
	Rhinichthys cataractae	0.66 (0.00, 1.50)	0.02
	Hypentelium nigricans	0.62 (0.00, 1.50)	0.01
	Catostomus commersonii	0.56 (0.00, 1.40)	0.01

Table 7 (continued)

Species	Key coefficient	Mean (90% CI)	RI
Salmo trutta	Drainage area × *Oncorhynchus mykiss*	− 2.08 (− 7.48, 0.00)	0.40
	Oncorhynchus mykiss	1.99 (0.76, 3.46)	0.36
	Drainage area × *Rhinichthys obtusus*	− 0.69 (− 3.31, 0.00)	0.04
	Developed × *Oncorhynchus mykiss*	− 0.68 (− 2.44, 0.00)	0.04
	Drainage area × *Etheostoma flabellare*	− 0.49 (− 2.30, 0.00)	0.02
	Rhinichthys obtusus	0.43 (0.00, 1.24)	0.02
	Stream temperature × *Oncorhynchus mykiss*	0.42 (0.00, 1.37)	0.02
	Micropterus dolomieu	− 0.38 (− 1.16, 0.00)	0.01
	Drainage area × *Micropterus dolomieu*	− 0.35 (− 2.14, − 0.00)	0.01
Salvelinus fontinalis	*Etheostoma blennioides*	− 0.44 (− 1.11, 0.00)	0.32
	Nitrate deposition	− 0.35 (− 1.17, 0.00)	0.20
	Pimephales notatus	− 0.31 (− 1.08, 0.00)	0.15
	Etheostoma zonale	− 0.18 (− 0.72, 0.00)	0.06
	Developed × *Salmo trutta*	− 0.18 (− 1.43, − 0.00)	0.05
	Semotilus atromaculatus	− 0.15 (− 0.76, 0.00)	0.04
	Wetland × *Pimephales notatus*	− 0.14 (− 0.80, − 0.00)	0.03
	Wetland × *Salmo trutta*	− 0.12 (− 1.04, − 0.00)	0.02
	Ambloplites rupestris	− 0.09 (− 0.51, 0.00)	0.01
	Slope × *Etheostoma blennioides*	0.08 (− 0.00, 0.39)	0.01
	Micropterus dolomieu	− 0.08 (− 0.45, 0.00)	0.01
	Drainage area × *Ameiurus nebulosus*	− 0.08 (− 0.23, 0.00)	0.01
Semotilus atromaculatus	Drainage area × *Rhinichthys obtusus*	− 4.08 (− 9.38, − 0.00)	0.41
	Drainage area × *Etheostoma flabellare*	− 3.50 (− 6.08, 0.00)	0.30
	Drainage area × *Etheostoma nigrum*	− 1.79 (− 4.08, − 0.00)	0.08
	Drainage area × *Clinostomus elongatus*	− 1.60 (− 4.19, − 0.00)	0.06
	Drainage area × *Catostomus commersonii*	− 1.40 (− 4.65, 0.00)	0.05
	Catostomus commersonii	1.33 (0.36, 2.27)	0.04

Mean value represents mean value of coefficient estimates across all bootstraps. Ninety percent CI represents the 90% bootstrapped confidence interval. Relative importance (RI) was calculated based on absolute effect size against all nonzero coefficients within that species

Acknowledgements The authors thank the staff of the Pennsylvania Department of Environmental Protection, Pennsylvania Fish and Boat Commission, and the Susquehanna River Basin Commission for their efforts in data collection required for this analysis. The authors thank Dr. Nicholas Clark for his assistance in using the CRF methodology within his MRFcov R package. The authors thank Dr. Jason Keagy for his review and feedback for this manuscript. The authors thank the anonymous reviewers for their insightful feedback and assistance in improving this manuscript.

Author contributions All authors contributed to conceptualization. CAC and TW contributed to all phases of the manuscript. DPF, AH, MK Shank, and TAW also contributed to data curation and writing—review and editing. MK Schall and GS also contributed to writing—review and editing.

Funding This research was funded by the Pennsylvania Sea Grant Program and the National Science Foundation (EF-1638679, EF-1638554, EF-1638539, and EF-1638550).

Data availability Raw data may be available by directly contacting the individual agencies responsible for data collection: Pennsylvania Department of Environmental Protection, Pennsylvania Fish and Boat Commission, and Susquehanna River Basin Commission.

Declarations

Conflict of interest The authors declare there are no Conflict of interest.

References

Allan, J. D. (2004). Landscapes and riverscapes: the influence of land use on stream ecosystems. *Annual Review of Ecology, Evolution, and Systematics, 35*, 257–284.

Arreguín-Sánchez, F. (1996). Catchability: A key parameter for fish stock assessment. *Reviews in fish biology and fisheries, 6*, 221–242.

Brandt, E. J., Feiner, Z. S., Latzka, A. W., & Isermann, D. A. (2022). Similar environmental conditions are associated with Walleye and Yellow Perch recruitment success in Wisconsin lakes. *North American Journal of Fisheries Management, 42*(3), 630–641.

Brown, J. J., Perillo, J., Kwak, T. J., & Horwitz, R. J. (2005). Implications of Pylodictis olivaris (flathead catfish) introduction into the Delaware and Susquehanna drainages. *Northeastern Naturalist, 12*, 473–484.

Buisson, L., Blanc, L., & Grenouillet, G. (2008). Modelling stream fish species distribution in a river network: The relative effects of temperature versus physical factors. *Ecology of Freshwater Fish, 17*, 244–257.

Buto, S. G., & Anderson, R. D. (2020). *NHDPlus high resolution (NHDPlus HR): A hydrography framework for the nation*. US Geological Survey: Technical report.

Clark, J. S., Gelfand, A. E., Woodall, C. W., & Zhu, K. (2014). More than the sum of the parts: Forest climate response from joint species distribution models. *Ecological Applications, 24*, 990–999.

Clark, N. J., Wells, K., & Lindberg, O. (2018). Unravelling changing interspecific interactions across environmental gradients using Markov random fields. *Ecology, 99*, 1277–1283.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems, 1695*, 1–9.

DeWeber, J. T., & Wagner, T. (2014). A regional neural network ensemble for predicting mean daily river water temperature. *Journal of Hydrology, 517*, 187–200.

DeWeber, J. T., & Wagner, T. (2015). Predicting brook trout occurrence in stream reaches throughout their native range in the eastern United States. *Transactions of the American Fisheries Society, 144*, 11–24.

Ensign, W. E., Temple, A. J., & Neves, R. J. (2002). Effects of fright bias on sampling efficiency of stream fish assemblages. *Journal of Freshwater Ecology, 17*, 127–139.

Frederico, R. G., De Marco Jr, P., & Zuanon, J. (2014). Evaluating the use of macroscale variables as proxies for local aquatic variables and to model stream fish distributions. *Freshwater Biology, 59*, 2303–2314.

Geheber, A. D., & Geheber, P. K. (2016). The effect of spatial scale on relative influences of assembly processes in temperate stream fish assemblages. *Ecology, 97*, 2691–2704.

Giam, X., & Olden, J. D. (2016). Environment and predation govern fish community assembly in temperate streams. *Global Ecology and Biogeography, 25*, 1194–1205.

Harris, D. J. (2016). Inferring species interactions from co-occurrence data with Markov networks. *Ecology, 97*, 3308–3314.

Hoxmeier, R. J. H., & Dieterman, D. J. (2013). Seasonal movement, growth and survival of brook trout in sympatry with brown trout in Midwestern US streams. *Ecology of Freshwater Fish, 22*, 530–542.

Hutchinson, G. E. (1957). Concluding remarks. In *Cold Spring Harbor Symposia on Quantitative Biology* (vol. 22, pp. 415–427), Cold Spring Harbor Laboratory Press.

Inoue, K., Stoeckl, K., & Geist, J. (2017). Joint species models reveal the effects of environment on community assemblage of freshwater mussels and fishes in European rivers. *Diversity and Distributions, 23*, 284–296.

Jiménez-Valverde, A. (2012). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography, 21*, 498–507.

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., & Xian, G. (2013). A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment, 132*, 159–175.

Kennard, M. J., Pusey, B. J., Harch, B. D., Dore, E., & Arthington, A. H. (2006). Estimating local stream fish assemblage attributes: Sampling effort and efficiency at two spatial scales. *Marine and Freshwater Research, 57*, 635–653.

Khan, M. R. A., & Brandenburger, T. (2020). ROCit: Performance assessment of binary classifier with visualization. https://CRAN.R-project.org/package=ROCit.

King, K. B., Giacomini, H. C., Wehrly, K., López-Fernández, H., Thomer, A. K., & Alofs, K. M. (2023). Using historical catch data to evaluate predicted changes in fish relative abundance in response to a warming climate. *Ecography, 2023*(8), e06798.

König, C., Wüest, R. O., Graham, C. H., Karger, D. N., Sattler, T., Zimmermann, N. E., & Zurell, D. (2021). Scale dependency of joint species distribution models challenges interpretation of biotic interactions. *Journal of Biogeography, 48*, 1541–1551.

Kristensen, E. A., Baattrup-Pedersen, A., & Andersen, H. E. (2012). Prediction of stream fish assemblages from land use characteristics: Implications for cost-effective design of monitoring programmes. *Environmental Monitoring and Assessment, 184*, 1435–1448.

Lundberg, J. G., Kottelat, M., Smith, G. R., Stiassny, M. L., & Gill, A. C. (2000). So many fishes, so little time: An overview of recent ichthyological discovery in continental waters. *Annals of the Missouri Botanical Garden, 87*, 26–62.

MacKenzie, D. I., Bailey, L. L., & Nichols, J. D. (2004). Investigating species co-occurrence patterns when species are detected imperfectly. *Journal of Animal Ecology, 73*, 546–555.

Magalhaes, M. F., Batalha, D. C., & Collares-Pereira, M. J. (2002). Gradients in stream fish assemblages across a Mediterranean landscape: Contributions of environmental factors and spatial structure. *Freshwater Biology, 47*, 1015–1031.

Manel, S., Williams, H. C., & Ormerod, S. J. (2001). Evaluating presence-absence models in ecology: The need to account for prevalence. *Journal of Applied Ecology, 38*, 921–931.

McKenna, J. E., Jr., & Johnson, J. H. (2011). Landscape models of brook trout abundance and distribution in lotic habitat with field validation. *North American Journal of Fisheries Management, 31*, 742–756.

McManamay, R. A., & Frimpong, E. A. (2015). Hydrologic filtering of fish life history strategies across the United States: Implications for stream flow alteration. *Ecological Applications, 25*, 243–263.

Olden, J. D., Joy, M. K., & Death, R. G. (2006). Rediscovering the species in community-wide predictive modeling. *Ecological Applications, 16*, 1449–1460.

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., & Abrego, N. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters, 20*, 561–576.

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal, 10*, 439–446.

Peterson, J. T., & Rabeni, C. F. (1995). Optimizing sampling effort for sampling warmwater stream fish communities. *North American Journal of Fisheries Management, 15*, 528–541.

Poff, N. L., & Allan, J. D. (1995). Functional organization of stream fish assemblages in relation to hydrological variability. *Ecology, 76*, 606–627.

Poggiato, G., Münkemüller, T., Bystrova, D., Arbel, J., Clark, J. S., & Thuiller, W. (2021). On the interpretations of joint modeling in community ecology. *Trends in Ecology & Evolution, 36*, 391–401.

R Core Team. (2022). R: A Language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., et al. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews, 94*, 849–873.

Rodríguez, M. A., Marselli, G., & Mandrak, N. E. (2021). Responses of vulnerable fishes to environmental stressors in the Canadian Great Lakes basin1. *Canadian Journal of Fisheries and Aquatic Sciences, 78*, 1278–1292.

Rota, C. T., Ferreira, M. A. R., Kays, R. W., Forrester, T. D., Kalies, E. L., McShea, W. J., Parsons, A. W., & Millspaugh, J. J. (2016). A multispecies occupancy model for two or more interacting species. *Methods in Ecology and Evolution, 7*, 1164–1173.

Seaber, P. R., Kapinos, F.P., & Knapp, G. L. (1987). Hydrologic unit maps.

Smith, G. D., Massie, D. L., Perillo, J., Wagner, T., & Pierce, D. (2021). Range expansion and factors affecting abundance of invasive flathead catfish in the Delaware and Susquehanna rivers, Pennsylvania, USA. *North American Journal of Fisheries Management, 41*, S205–S220.

Stauffer, J. R., Criswell, R. W., & Fischer, D. P. (2016). *The fishes of Pennsylvania*. Cichlid Press.

Strayer, D. L., & Dudgeon, D. (2010). Freshwater biodiversity conservation: Recent progress and future challenges. *Journal of the North American Benthological Society, 29*, 344–358.

Tikhonov, G., Abrego, N., Dunson, D., & Ovaskainen, O. (2017). Using joint species distribution models for evaluating how species-to-species associations depend on the environmental context. *Methods in Ecology and Evolution, 8*, 443–452.

Vallé, C., Poggiato, G., Thuiller, W., Jiguet, F., Princé, K., & Le Viol, I. (2023). Species associations in joint species distribution models: From missing variables to conditional predictions. *Journal of Biogeography, 51*(2), 311–324.

Van Vliet, M. T., Ludwig, F., & Kabat, P. (2013). Global streamflow and thermal habitats of freshwater fishes under climate change. *Climatic Change, 121*, 739–754.

Wagner, T., Deweber, J. T., Detar, J., & Sweka, J. A. (2013). Landscape-scale evaluation of asymmetric interactions between brown trout and brook trout using two-species occupancy models. *Transactions of the American Fisheries Society, 142*, 353–361.

Wagner, T., Hansen, G. J., Schliep, E. M., Bethke, B. J., Honsey, A. E., Jacobson, P. C., Kline, B. C., & White, S. L. (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*, 1540–1551.

Walker, J. D., O'Neil, K., & Letcher, B. (2015). EcoSHEDS Northeast Catchment Delineation. https://github.com/EcoSHEDS/necd.

Walters, D., Leigh, D., Freeman, M. C., Freeman, B., & Pringle, C. (2003). Geomorphology and fish assemblages in a Piedmont river basin, USA. *Freshwater Biology, 48*, 1950–1970.

Wehrly, K. E., Wiley, M. J., & Seelbach, P. W. (2003). Classifying regional variation in thermal regime based on stream fish community patterns. *Transactions of the American Fisheries Society, 132*, 18–38.

Wilkinson, D. P., Golding, N., Guillera-Arroita, G., Tingley, R., & McCarthy, M. A. (2019). A comparison of joint species distribution models for presence-absence data. *Methods in Ecology and Evolution, 10*, 198–211.

Zurell, D., Pollock, L. J., & Thuiller, W. (2018). Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments? *Ecography, 41*, 1812–1819.