Journal of Biogeography

WILEY

Journal of Biogeography

RESEARCH ARTICLE OPEN ACCESS

The Influence of Spatial Extent and Distance on Macroinvertebrate-Environment Relationships in Conterminous USA Stream Sites

Bruno S. Godoy[1] | Tadeu Siqueira[2,3] | Robert M. Hughes[4,5]

[1]Instituto Amazônico de Agriculturas Familiares, Universidade Federal do Pará (UFPA), Belém, PA, Brazil | [2]Instituto de Biociências, Universidade Estadual Paulista (UNESP), Rio Claro, SP, Brazil | [3]School of Biological Sciences, University of Canterbury, Christchurch, New Zealand | [4]Amnis Opes Institute, Corvallis, Oregon, USA | [5]Department of Fisheries, Wildlife, & Conservation Sciences, Oregon State University, Corvallis, Oregon, USA

Correspondence: Bruno S. Godoy (bspacek@ufpa.br)

Received: 26 February 2025 | **Revised:** 28 May 2025 | **Accepted:** 3 June 2025

Funding: This work was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo (2019/04033-7) and Fulbright Brazil.

Keywords: aquatic insects | ecoregions | landscape ecology | lotic ecosystems | spatial extent | spatial resolution

ABSTRACT

Aim: The relationships between environmental variables and macroinvertebrate assemblages depend on spatial extent, but the relative importance of local and regional factors remains unclear. Understanding these scale-dependent relationships is crucial for improving biodiversity assessments and ecosystem management. We hypothesise that climate variables drive macroinvertebrate composition at broad spatial extents, whereas local water quality and habitat conditions are more influential at smaller spatial extents.

Location: A total of 5,788 streams across nine ecoregions in the conterminous United States, sampled by the United States Environmental Protection Agency (EPA).

Time Period: Samples collected in the years 2008, 2009, 2013, 2014, 2018, and 2019.

Major Taxa Studied: Freshwater macroinvertebrates, including 992 genera.

Methods: We analysed macroinvertebrate abundance data alongside 25 environmental variables categorised into four types: physical-habitat structure, water quality, watershed characteristics, and climate. We applied Multivariate Multiscale Codependency Analysis (MMCA) at both national and ecoregion-specific scales to determine which environmental variables were most strongly associated with assemblage composition at different spatial extents.

Results: Assemblage-environment relationships varied with spatial extents. Climate and watershed variables were the dominant drivers at broader spatial extents, whereas local water quality and physical habitat structure were more influential at smaller inter-site distances. Additionally, the environmental variables most strongly associated with macroinvertebrate composition varied across ecoregions, suggesting that regional factors such as land use and historical environmental conditions shape biotic responses.

Main Conclusions: Our findings underscore the importance of spatial scale in ecological studies and conservation planning. Recognising how key environmental predictors change with spatial extent enhances biodiversity assessments, informs targeted conservation strategies, and improves ecosystem management. The observed regional variation in assemblage-environment relationships highlights the need for regionally tailored management approaches to maintain and restore aquatic biodiversity.

1 | Introduction

Spatial extent (overall size of the study area) and distance (pairwise separation between sites) are key moderators of the relationships between aquatic communities and their environments (Burgazzi et al. 2020; Johnson et al. 2004; Lansac-Tôha et al. 2021). They represent the physical extent over which ecological processes are observed and can influence how patterns and processes are detected. For example, at small spatial extents, local factors such as substrate size, water quality, and hydrological conditions often determine assemblage composition at local sites (Doretto et al. 2021; Li et al. 2012). In addition, habitat structural heterogeneity, resource availability, and biotic interactions also influence assemblages at small extents (Boyero 2003; Burgazzi et al. 2020; Tabi et al. 2024). Largely natural local factors, such as benthic organic matter, flow velocity, and water depth, are critical in structuring macroinvertebrate assemblages, with seasonal variations influencing their relative importance (Burgazzi et al. 2020). For example, in perennial rivers, aquatic vegetation, nutrients, and conductivity may be primary influences, whereas in intermittent rivers, water temperature, dissolved oxygen, and canopy cover may be key factors (Herlihy et al. 2020; Karaouzas et al. 2019; Oliveira et al. 2024).

On the other hand, at larger spatial extents, broad-scale factors such as climate and catchment characteristics tend to become more dominant drivers (Tonn 1990; Zhou et al. 2020). For example, at larger spatial extents, variables such as latitude, forest cover, and temperature are important determinants of macroinvertebrate assemblage composition (Li et al. 2012; Schneck et al. 2022). These regional factors often interact with local factors to create complex distribution patterns (Linares et al. 2025; Roque et al. 2010). For example, Linares et al. (2025) found that relatively uncommon taxa tended to be more susceptible to local disturbances, whereas relatively common taxa tended to be more susceptible to catchment disturbances. Also, watershed and channel network conditions affect biota indirectly through their effects on local water quality and physical habitat conditions (Kaufmann et al. 2022a; Leitão et al. 2018). When attempting to identify key environmental drivers at multiple spatial extents and distances, it is essential to account for the spatial structure underlying the organisation of biological assemblages (Barreiros et al. 2023; Godoy et al. 2019). Spatial processes, such as dispersal between habitats, can overcome the influence of local environmental and climatic variables in structuring the beta diversity of macroinvertebrate assemblages (Siqueira et al. 2014; Wu et al. 2022). Yet, despite the recognised role of spatial processes, few studies have explicitly quantified the spatial extents at which local versus regional environmental drivers exert their maximal influence on macroinvertebrate assemblages, an analytical gap we address across multiple ecoregions.

To examine the differing importance of spatial extent and distance on environment-assemblage relationships, one must use multivariate and multiscale analyses. Such analyses can help us identify relationships between environmental variables and biological assemblages at different observation extents (Borcard and Legendre 2002; Guénard and Legendre 2018a). These tools are particularly useful for investigating how the relative importance of environmental variables can change with increasing spatial distance and extent of observation, providing critical information for the management and conservation of aquatic ecosystems (Godoy et al. 2017, 2022). In addition, multiscale analyses make it possible to identify emerging patterns that cannot be observed at smaller spatial extents, facilitating a more holistic understanding of ecological interactions. Integrating multiple sources of environmental and biological data into these analyses also broadens their applicability and contributes to the development of more robust predictive models. However, few studies have systematically tested how assemblage–environment relationships vary with inter-site distance across ecoregions in the contiguous USA.

Systematic observations covering different geographic extents are needed to define how different geographic extents affect environment-assemblage relationships. The United States Environmental Protection Agency (EPA) implemented standardised sampling protocols on thousands of river and stream sites (USEPA 2024) across the conterminous USA through the use of a probability survey design (Olsen and Peck 2008). That database can be used to study the effects of distance in the relationship between environmental predictors and macroinvertebrate responses. Previous evidence suggests that the strength and direction of assemblage-environment relationships vary across USA ecoregions (Herlihy et al. 2020; Patrick and Yuan 2019). For example, the strengths of assemblage-nutrient relationships have been reported to be weaker in ecoregions where all streams have high nutrient concentrations (Bini et al. 2014; Hughes et al. 2006). In addition, the strengths of the assemblage-environment relationships amongst USA ecoregions tended to show stronger relationships in the north (Thompson et al. 2020).

In our study, we used macroinvertebrate data from the EPA database to test the hypothesis that the environmental variables that affect macroinvertebrate assemblage composition vary with distance amongst sites. Therefore, we predicted that climate variables would drive variation in assemblage composition over large spatial extents (e.g., ecoregions), but that local (site-specific) variables would drive assemblage responses over small spatial extents. We expected that the set of environmental variables that correlated most strongly with macroinvertebrate assemblage composition would vary depending on the environmental conditions of the ecoregions. Therefore, we tested the effect of spatial extent for each USA ecoregion separately.

2 | Methods

2.1 | Database

We combined data from six survey years (2008, 2009, 2013, 2014, 2018, and 2019) of the EPA's National Rivers and Streams Assessment (NRSA), all collected under a fully standardised field and laboratory protocol (Paulsen, Hawkins, et al. 2008; Paulsen, Mayio, et al. 2008; USEPA 2024). Because (1) no sites were re-sampled across years and (2) identical sampling timing and methods were applied in each survey, we treated the pooled dataset as a spatial "snapshot." This maximised our geographic coverage without introducing detectable temporal bias. We acknowledge that temporal dynamics might influence assemblages at finer scales, but modelling interannual change is beyond the scope of the current spatial-extent focus. This

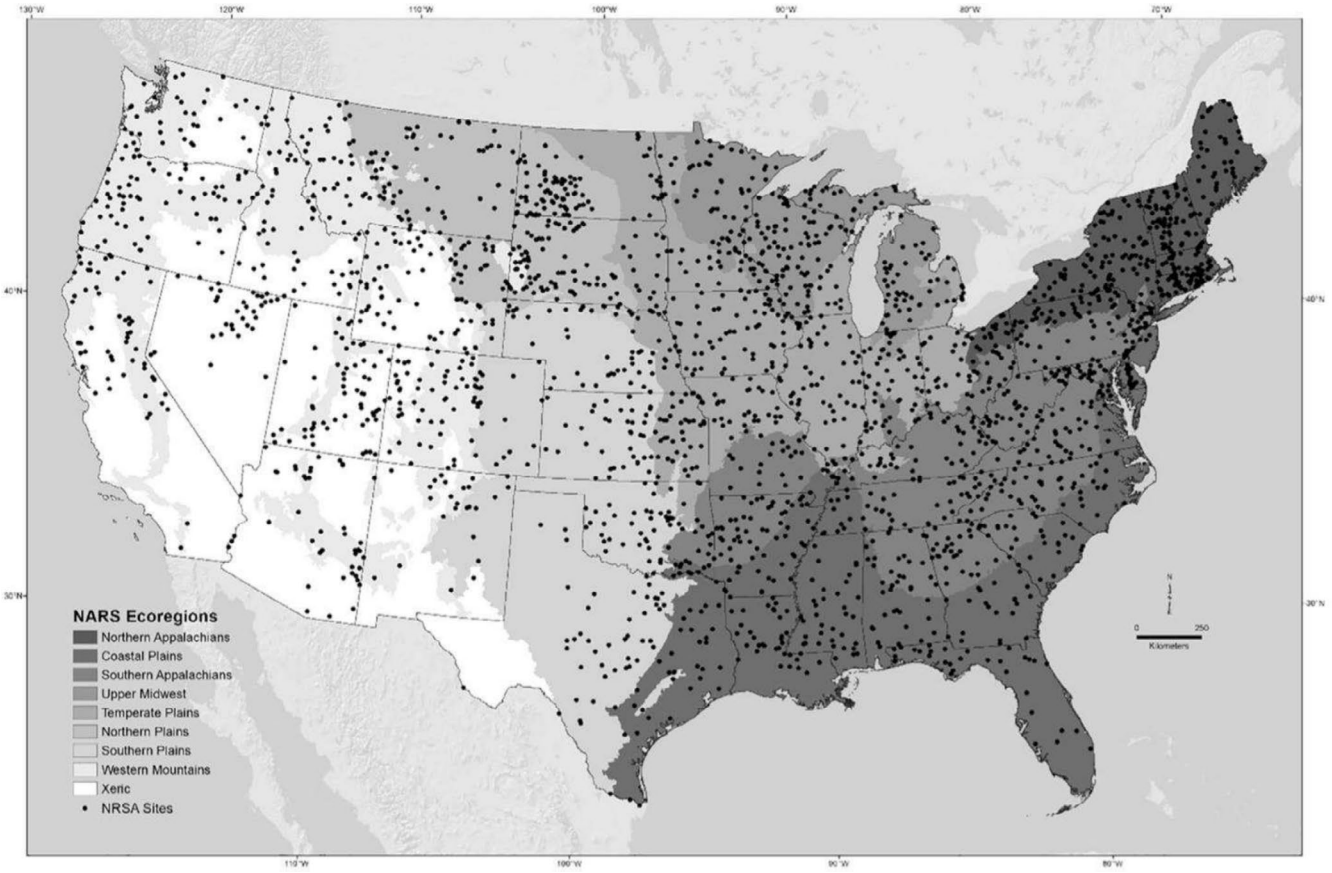

FIGURE 1 | Locations of the 5788 stream and river sites sampled between 2008 and 2019 in the nine study ecoregions (from Herlihy et al. 2020).

dataset allowed us to elaborate a hierarchy of spatial extents associated with varying spatial distances amongst sites because the sample covers different geographical units.

Before analysing the data, we selected compatible sampling units. First, we filtered taxonomic units for standardisation by removing taxa that were not identified to at least genus level, so that abundances would be expressed at the same taxonomic resolution. We excluded sites with fewer than 300 individuals because NRSA counts are based on non-uniform subsampling of sorted samples (Bini et al. 2014), and bioassessment guidelines recommend a minimum of ~300 organisms to ensure stable community metrics and minimise sampling variance (Barbour et al. 1999). This threshold helps guarantee that samples adequately capture the majority of taxa present at each site. We used a total of 992 macroinvertebrate genera, including Platyhelminthes, Mollusca, and Arthropoda. After those procedures, we retained 5788 sites distributed across nine ecoregions (Figure 1). The numbers of sites varied by ecoregion: Coastal plains (CPL) 686; Northern Appalachians (NAP) 700; Northern plains (NPL) 506; Southern Appalachians (SAP) 844; Southern plains (SPL) 494; Temperate plains (TPL) 680; Upper Midwest (UMW) 544; Western mountains (WMT) 748; and Xeric west (XER) 586.

We selected a focused suite of 25 habitat (e.g., substrate composition, riffle percentage, riparian cover), water-quality (nutrients, conductivity, DO, pH), and watershed (catchment area, forest cover, land-use composition) variables because these have been repeatedly identified as key drivers of stream macroinvertebrate communities (Fanelli et al. 2022; Herlihy et al. 2020). These variables, divided into four different categories, were selected for our analysis. The categories were: (1) variables describing site physical-habitat condition; (2) site water-quality variables; (3) watershed conditions; and (4) climate data (Table 1). The environmental variables resulting from EPA monitoring, the descriptions of how they were obtained, and standardisation details are freely available on the agency's website (see Supporting Information S1). Climate variables were 30-year normals (1991–2020) from PRISM, providing a stable baseline of average annual temperature and precipitation that aligns with, but is not tied to, our 2008–2019 sampling window (PRISM Climate Group 2023). We used those climate variables because previous studies indicated a relationship between macroinvertebrate assemblages and the seasonality of both temperature and rainfall (Heino 2002, 2009, 2011). The standardisation of which sites would be used, taxonomic unit, and environmental variables follows studies that showed relationships between macroinvertebrate assemblages (at this taxonomic resolution) and the selected environmental variables (Bini et al. 2014; Heino 2011).

2.2 | Data Analysis

Before the analyses, we standardised all the responses (site by genera abundance matrix) and the predictor (site by environmental variables matrix). For macroinvertebrate genera abundance data, we used the Hellinger transformation. Predictor variables

TABLE 1 | List of environmental variables and their categories and codes.

Category	Variable	Code	Unit
Physical habitat	Fine sediments	PCT_FN	%
	Bed stability	LRBS_USE	—
	Natural habitat	XFC_NAT	%
	Woody canopy	XCMGW	Proportion
	Human disturbance index	W1_HALL	—
	Human agricultural disturbance	W1_HAG	—
	Occurrence of pools	PCT_POOL	%
	Occurrence of fast waters	PCT_FAST	%
	Water surface slope	XSLOP	%
	Wetted width	XWIDTH	Meter
	Thalweg depth	XDEPTH	Meter
Water quality	Electrical conductivity	COND	μScm^{-1}
	Total phosphorus	PTL	μgL^{-1}
	Total nitrogen	NTL	μgL^{-1}
	Turbidity	TURB	NTU
	Chloride	CHLORIDE	μeqL^{-1}
	Dissolved organic carbon	DOC	μgL^{-1}
	Sulfate	SULFATE	μgL^{-1}
	pH	PH	—
Watershed conditions	Density of mines	MINEDENSWS	—
	Pollution discharges	NPDESDENSWS	—
	Cropland use	PCTCROP2019WS	%
	Herbaceous wetland land cover	PCTHBWET2013WS	%
	Imperviousness surface	PCTIMP2013WS	%
	Urban area	PCTURB2019WS	%
	Woody wetland land cover	PCTWDWED2013WS	%
	Population density	POPDEN2010WS	—
	Road density	RDDENSWS	Meters
	Catchment area	WSAREASQKM	km^2
Climate data	Annual precipitation	PRECIP8110WS	mm
	Annual maximum temperature	TMAX8110WS	°C
	Annual minimum temperature	TMIN8110WS	°C

were standardised using *z*-scores to achieve a mean of zero and a standard deviation of one. We used Multivariate Multiscale Codependency Analysis (MMCA; Guénard and Legendre 2018a) to identify the environmental variables most related to macroinvertebrate genera abundances at varying inter-site spatial distances. Instead of adding each single variable sequentially as in regression or random forest analyses, MMCA estimates the joint spatial dependence of an entire response matrix (i.e., a site by genera composition matrix) and an entire explanatory matrix (i.e., a site by environmental variables matrix) using a spatially explicit approach. Because predictors are never entered simultaneously into the same model, multicollinearity amongst variables cannot inflate or bias the scale-dependent relationships we investigate.

MMCA measures the strength of the relationships between the biotic composition matrix and the environmental variable matrix at various spatial distances for each site. The distances were represented by the axes of a Principal Coordinates

Analysis of a Neighbourhood Matrix (PCNM; Borcard and Legendre 2002) run on a matrix of geographical distances amongst sites. The number of eigenvectors calculated, each representing a spatial extent, is equivalent to the number of samples, with the first eigenvectors corresponding to larger distances and the last eigenvectors to smaller distances. These eigenvectors can be interpreted as spatial structures, as they create zones that have similar values separated by spatial distances.

Although straight-line (Euclidean) distance does not always mirror dendritic connectivity in stream networks, we used PCNM (a form of dbMEM) because its spatial eigenfunctions decompose variation into multiple scales and, when based on an appropriately constrained distance matrix, can approximate both connectivity and barriers (Borcard and Legendre 2002; Lacey et al. 2007). Empirical tests in stream macroinvertebrates show that Euclidean-based models perform comparably to, and sometimes better than, network-distance models, especially for taxa with aerial dispersal (Kärnä et al. 2015; Li, Chen, et al. 2021; Li, Heino, et al. 2021; Toskey et al. 2024). In a representative ecoregion (Southern plains), applying channel-network distances would have discarded ~20% of sites (> 500 m from mapped channels in the NHD) without altering the overall spatial-scaling patterns (Figure S1; Supporting Information S3). While acknowledging the ecological relevance of hydrologic-distance metrics, we opted for Euclidean-based dbMEM to maximise sample coverage and computational feasibility across 5700 sites, noting that our results are robust to either approach.

The strength of the relationship between an environmental variable and the genera composition data at a specific spatial distance was quantified using the codependence coefficient ($C_{u;Y,X}$), where u is the MEM vector, Y is the trace vector of the genera compositional matrix, and X is the environmental variable. This coefficient is calculated as the product of two elements: the square root of the R^2 of the regression between the MEM vector against the assemblage composition trace vector, and the Pearson correlation between the MEM vector and the environmental variable. There was then one coefficient for each environmental variable at each spatial distance. For simplicity, we only used the coefficient with the highest value (in this case, absolute, as it can be positive or negative) for each spatial scale. This decision enhances interpretability by highlighting the processes exerting the greatest influence at each spatial distance while filtering out weaker, potentially spurious associations. The proportion of variance in community structure explained by the selected predictors (R^2) is reported for each spatial eigenvector in Table S1. Values vary across eigenvectors (e.g., from ~5% up to ~30% for different scales), reflecting the amount of local-to-regional variation captured by the single strongest codependence at each level. Because each R^2 pertains to a distinct spatial distance, they are not additive and should not be summed across spatial distances.

Multivariate codependence coefficients were tested using the product of two Fisher–Snedecor F statistics (Guénard and Legendre 2018a). The individual predictor variables associated with the assemblages and the PCNM axes for which those associations were relevant are presented as results. The same variable can be related to assemblages more than once but be associated with different PCNM axes. This peculiarity of repetition is important because it facilitates the visualisation of relationships between the same facet of the environment and the assemblages at different distances.

We estimated the wavelength of each spatial PCNM predictor to determine its approximate spatial distance (Guénard et al. 2010). The wavelength for a PCNM variable with rank i was estimated using the formula: $2(L+\bar{s})(i+1)^{-1}$, where L is the spatial extent of the study area, and $\bar{s}$ represents the average of the site distances. We used the average distance between sites because the distances amongst sites in the NRSA survey differ. Using this equation, the calculated distances exhibit an inverse proportional decay relative to the rank of the PCNM axis. For example, the first PCNM axis had a wavelength of 8124.09 km ($L = 6322.56$ km and $\bar{s} = 1801.53$ km), whereas the 5788th axis had a wavelength of 2.81 km.

The analysis was conducted in two steps. First, we analysed all data across the conterminous USA by running the MMCA once for the entire site by the genera abundance matrix and the site by environment matrix. In a second step, we analysed the site by genera abundance matrix and the site by environment matrix for individual ecoregions separately (Herlihy et al. 2020). This analysis provided a more accurate picture of the ecoregions because smaller effect sizes become observable by excluding high variations when comparing across multiple ecoregions at the same time.

Given the high number of sites in our study (5788), we ended up with 654 codependence coefficients. We thus used a strategy to better visualise and interpret the results. We calculated the proportion of pre-defined environmental variable categories across the range of spatial scales represented by the MEM vectors. To statistically analyse the shift in variable category proportions with changing spatial scale, we used a multinomial regression. This model tested how the variations in the proportions of the number of explanatory variables classified in the four categories (physical habitat, water quality, watershed, and climate) related to the variation in the macroinvertebrate assemblage composition matrix. The explanatory variable of this multinomial regression was the spatial distance of the PCNM axis. We needed to use multinomial analyses because of the large number of axes selected in the MMCA, and we used the odds ratio test to determine the significance of the multinomial regression (for more details of the data analysis, see Supporting Information S2). We estimated the mean proportion of significant individual genera responses to the environmental variable matrices in each PCNM axis of the MMCA, grouped in the four categories of variables. These analyses revealed the effect size of the different environmental variable categories on the assemblages. We used the *vegan* and *codep* packages in R (Guénard and Legendre 2018b; Oksanen et al. 2020; R Core Team 2021).

3 | Results

3.1 | Watershed Area Distribution

The ecoregions had similar distributions of sampled watershed areas compared to those across the conterminous USA (Figure 2; Supporting Information S4). However, differences amongst

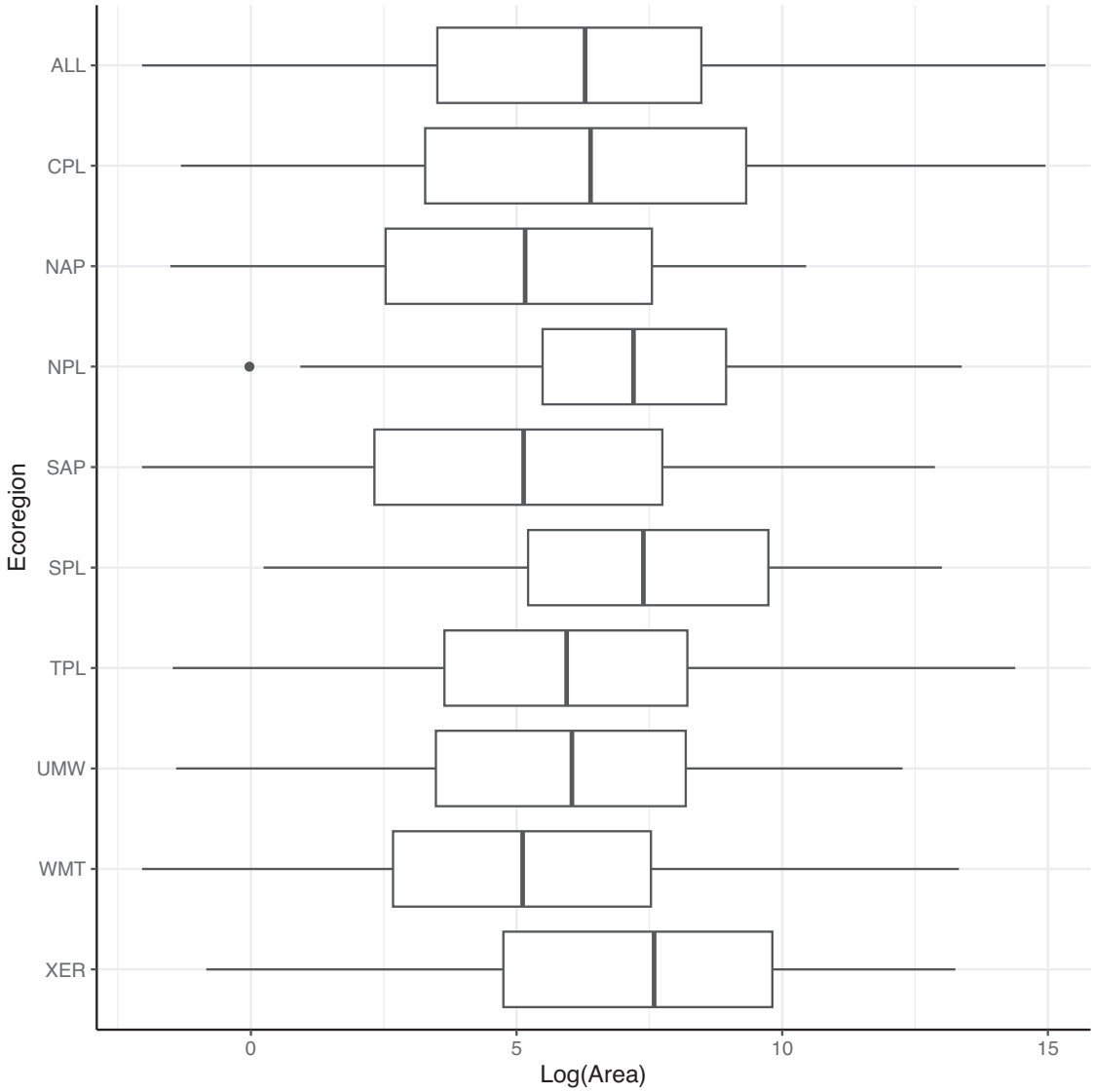

FIGURE 2 | Distributions of watershed areas for the conterminous USA rivers and streams sampled between 2008 and 2019, for all sites and in nine ecoregions.

ecoregions indicate a heterogeneous distribution of large and small water courses, with more sites with smaller watershed areas in the more mountainous ecoregions (NAP, SAP, WMT). On the other hand, the drier ecoregions (NPL, SPL, XER) tended to have sites with larger watershed areas.

3.2 | Spatial Autocorrelation of Environmental Categories

All the environmental variables used for analysis were correlated with the macroinvertebrate genera abundances to some degree, in at least one PCNM axis. The number of PCNMs selected for ecoregions differed because the number of PCNM axes is directly related to the number of observations. Despite this difference, the percentages of PCNM axes related to the environmental variables in the separate environmental categories were similar for all data and the nine ecoregions (Table 2, Supporting Information S4). The watershed variables had the largest number of correlations with macroinvertebrate genera abundances in all spatial extents, except in the NAP and NPL ecoregions.

The influence of habitat structure and water quality variables on macroinvertebrate genera composition generally decreased with increased distance, whereas the importance of climate variables increased. This pattern was consistent both across the conterminous USA (Figure 3) and within the nine ecoregions (Figure 4). Generally, we found the greatest proportions of the number of environmental variables correlated with the macroinvertebrate genera composition matrix at smaller distances, represented by habitat structure and water quality. This pattern was different only for water quality in the SAP and UMW and habitat structure in the CPL. The proportions of watershed variables were relevant at intermediate to large distances for all sites and for the nine ecoregions. The greatest proportions of climate variables correlated with the macroinvertebrate genera matrix were at the largest distances.

TABLE 2 | Number of correlations between environmental variables and macroinvertebrate abundance mediated by PCNM, separated by variable category.

Subset	Habitat	Water quality	Watershed	Climate	Total
All	191 (29.20%)	192 (29.36%)	253 (38.69%)	18 (2.75%)	654
CPL	21 (25.93%)	24 (29.63%)	32 (39.51%)	4 (4.94%)	81
NAP	44 (34.65%)	45 (35.43%)	31 (24.41%)	7 (5.51%)	127
NPL	28 (32.56%)	23 (26.74%)	27 (31.40%)	8 (9.30%)	86
SAP	29 (19.46%)	49 (32.89%)	60 (40.27%)	11 (7.38%)	149
SPL	14 (21.54%)	24 (36.92%)	25 (38.46%)	2 (3.08%)	65
TPL	24 (27.91%)	20 (23.26%)	40 (46.51%)	2 (2.33%)	86
UMW	18 (27.27%)	17 (25.76%)	29 (43.94%)	2 (3.03%)	66
WMT	40 (27.97%)	27 (18.88%)	68 (47.55%)	8 (5.59%)	143
XER	40 (28.78%)	34 (24.46%)	59 (42.45%)	6 (4.32%)	139

Note: Values in parentheses are the percentages representing the total number of PCNM in each subset.

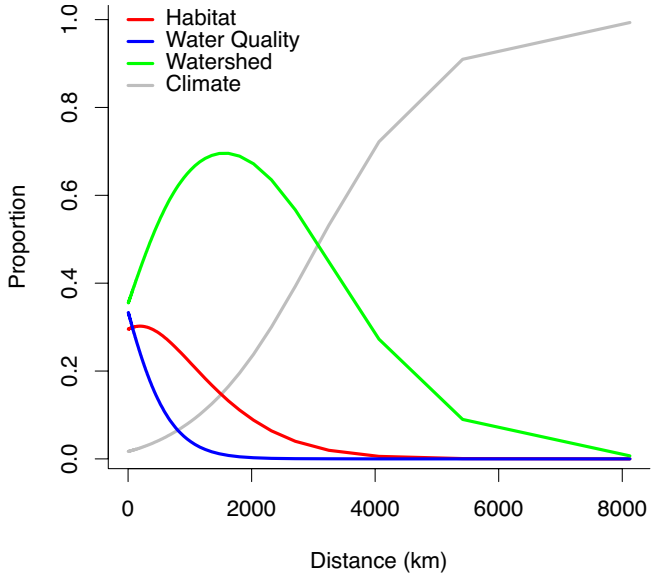

FIGURE 3 | Proportions of the number of environmental variable categories correlated with macroinvertebrate assemblage composition matrix related to distance for the 5788 sites sampled across the conterminous USA.

3.3 | Genus-Level Response by Category

For all sets of data (all sites and for the nine ecoregions), there was variation in the proportion of the number of variables related to spatial distance ($p < 0.05$ in the odds ratio test of the multinomial model). Additionally, at both the USA-wide and ecoregional scales, climate variables exhibited a greater number of genera correlated compared to the other environmental variables (Table 3).

4 | Discussion

The main novelty of our study lies in identifying three characteristic spatial extents at which different classes of environmental drivers peak in importance for explaining macroinvertebrate genus composition: (1) local extents (inter-site distance < 200 km), where water quality and physical habitat variables dominate; (2) intermediate extents (inter-site distance ~1000 km), driven primarily by watershed characteristics; and (3) broad extents (inter-site distance > 3500 km), where climate variables play the strongest role. This shifting pattern of assemblage–environment relationships with spatial distance aligns with classical distance-decay theory (Soininen et al. 2007) and underscores the scale-dependent processes structuring lotic ecosystems, demonstrating that environmental effects vary not only with the spatial extent of the study region (ecoregion vs. conterminous USA) but also with the type of variable measured.

Usually, studies relate assemblage responses with spatial extent as sets of environmental variables (e.g., local or regional) (Borcard and Legendre 2002; Dray et al. 2012; Leal et al. 2018; Macedo et al. 2014). However, the effect–distance relationship is important for understanding the drivers of assemblage dynamics (Moraga et al. 2019), determining optimal spatial extents for research (Patrick and Yuan 2019) for rational ecosystem management (Omernik and Griffith 2014; Wang et al. 2006). By separating the effects of water quality variables and physical habitat structure variables from those related to climate and watershed characteristics across varying spatial extents, we identified differing inter-site spatial distances at which different environmental predictors exert their influence.

The differing importance of subsets of environmental variables related to spatial distance may be explained by the intrinsic mechanisms associated with each variable. Water quality and physical habitat variables are locally important because they act as immediate environmental filters in streams and lakes, selecting individuals regardless of total assemblage density (Herlihy et al. 2020; Kaufmann et al. 2014; Tonn 1990). Only organisms that can tolerate those conditions can establish and maintain local populations. On the other hand, climate and watershed variables are related to assemblages over larger distances (Arita and Rodríguez 2004; Heino 2009). Climate variables are not susceptible to point changes in water bodies (e.g.,

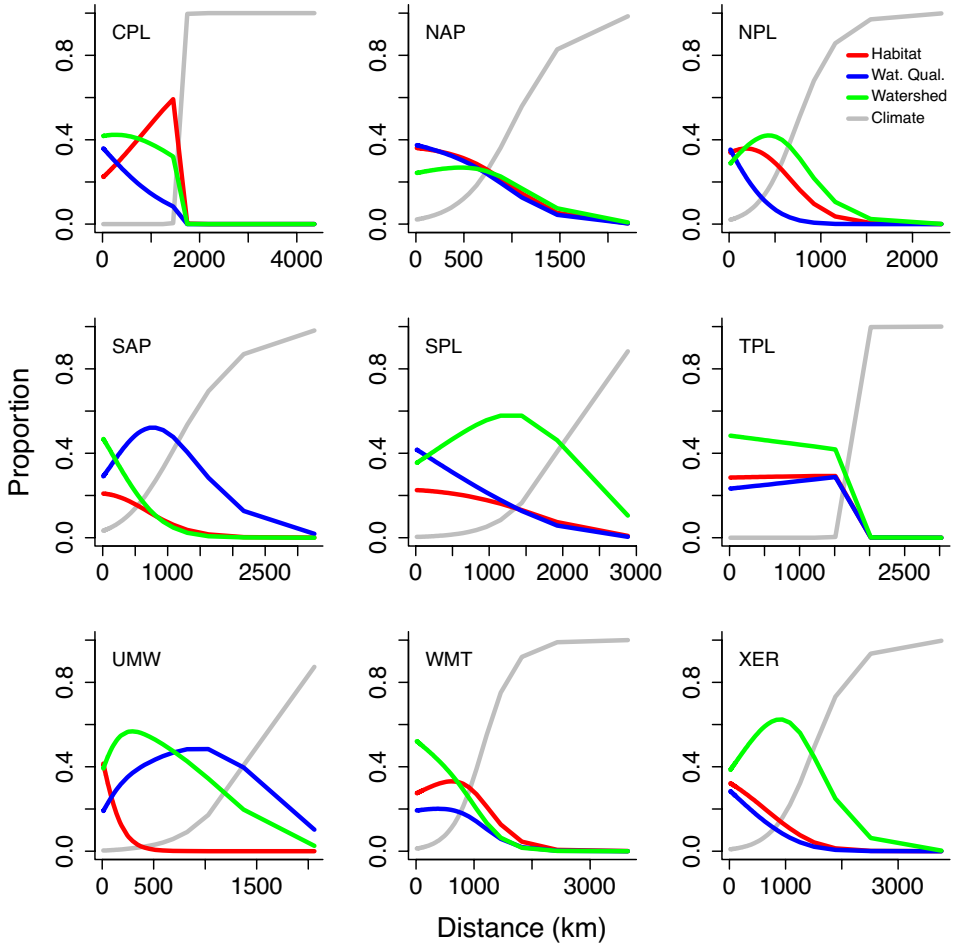

FIGURE 4 | Proportions of the number of environmental variable categories correlated with macroinvertebrate assemblage composition matrix related to distance for the nine ecoregions.

TABLE 3 | Average number of genera correlated with environmental variable categories for the USA and ecoregion spatial extents.

Subset	Habitat	Water quality	Watershed	Climate
All	52.12	71.97	118.76	521.94
CPL	31.33	55.29	47.94	372.25
NAP	28.84	47.60	75.68	197.57
NPL	30.32	30.96	43.56	209.50
SAP	30.00	50.80	42.08	175.00
SPL	34.50	33.25	56.38	305.50
TPL	24.54	44.40	51.25	420.50
UMW	29.17	59.18	50.86	335.00
WMT	32.33	98.04	60.54	269.88
XER	35.30	32.41	54.51	264.33

sewage discharges) and represent changes that distinguish ecoregions from others, even separating distinct species pools (Rohm et al. 1987; Whittier et al. 1988). Thus, climate may have both direct and indirect effects on macroinvertebrate assemblages, creating selective conditions or even affecting the productivity of environments (Heino 2011; Moi et al. 2023). However, groups that are found in nearby ecoregions with similar environments have a similar occupation history, both ecological and evolutionary, elevating the similarities in assemblages over large spatial distances and amongst ecoregions (McCormick et al. 2000; Van Sickle and Hughes 2000; Waite et al. 2000). Therefore, we suggest that assemblage variation over varying spatial distances can be linked to shifts in regional climatic conditions.

Multiple-Scale Codependency Analysis splits the partitioning of the total variation in the response matrix amongst different sets of predictor variables. As a result, we were unable to separately assess the influence of spatial distance from any single environmental variable on the assemblages studied. Although spatial distance was explicitly included in the model, we could not isolate its unique effect. However, previous studies (e.g., Bini et al. 2014) have shown a moderate degree of spatial autocorrelation (Moran's $I \approx 0.5$) in these assemblages at smaller spatial distances. This finding highlights the influence of riverine spatial structure on the distribution of assemblages. Therefore, we advise caution when interpreting our results, as some of the variation in assemblage composition may not have been fully accounted for in our analysis. Another point to caution is that we excluded sites with fewer than 300 individuals to ensure reliable

and representative community summaries (Barbour et al. 1999; Bini et al. 2014). However, we acknowledge that excluding low-abundance sites, which potentially reflect highly disturbed environments or unique local conditions, may lead to underrepresentation of extreme ecological scenarios.

The differences in the proportions of variable categories amongst ecoregions that were correlated with macroinvertebrate genera abundances related to spatial distance suggest that assemblage determinants likely interact with other processes, such as historical legacies and fundamental ecoregion differences. Different land use patterns, as well as the time over which this use has been occurring, may generate delayed assemblage responses to recent environmental changes (Fukami et al. 2016; Fukami and Nakajima 2011; Harding et al. 1998). In this scenario, assemblages would be subject to varied constraints that alter the history of extinction, selective pressures, and invasion by other species. Historical local contingencies (such as past colonisation sequences, disturbance legacies, or priority effects) can lead to alternative stable states in stream assemblages (Beisner et al. 2003; Santos et al. 2020; Schröder et al. 2005). These contingencies may produce patterns indistinguishable from neutral expectations (e.g., dispersal limitation, stochastic dominance) despite underlying niche-driven processes (Fukami 2015; Leibold and Chase 2018). In metacommunity theory, this interplay blurs the line between neutral and niche models. Limited dispersal or strong priority effects can mask environmental filtering, generating community similarity or turnover that appears randomly structured (Gravel et al. 2006; Vellend 2010). Conversely, niche differences may still underpin assembly even where neutral-like patterns emerge. Thus, our finding of scale-dependent driver shifts should be viewed in light of both deterministic filtering and potential neutral-like dynamics arising from historical legacies and dispersal constraints.

What is driving the ecoregional differences in environmental drivers of macroinvertebrate assemblage responses? We suggest that they originate from the fundamental differences in the ecoregions themselves. The original ecoregions were delineated by assessing differences in co-occurring patterns and apparent interactions of physiography, soils, potential natural vegetation, and predominant land uses, as well as occasional differences in crop types, hydrology, lithology, and climate (Omernik 1987; Omernik and Griffith 2014). However, as one ecoregion transitions into another, the boundary between the two may be relatively distinct or quite indistinct, and neighbouring ecoregions themselves vary in their distinctiveness and inherent variability. The relative ecoregion distinctiveness, variability, and boundary fuzziness affect the clarity of environment-assemblage responses for sites at or near those borders or in those ecoregions. The same is true for sites on rivers that drain two or more ecoregions because the water quality and substrate in the upstream ecoregion affect the biota of sites in the downstream ecoregion that receive that water and substrate (Hughes et al. 1994; Omernik et al. 2017; Omernik and Griffith 1991; Pinto et al. 2009). Our results indicate that these effects may occur at many kilometres' distance. Furthermore, the original ecoregions of the conterminous USA have been mapped at four different hierarchical levels (Omernik and Griffith 2014). Those used in our study are aggregates of Level-III ecoregions, further indicating that the environment-assemblage patterns that we observed can be aggregated or further subdivided at increasing or decreasing spatial extents, respectively.

In the USA, ecoregions have been employed for interpreting fish and macroinvertebrate assemblage patterns (Hughes et al. 1987; Larsen et al. 1986; Rohm et al. 1987; Whittier et al. 1988) as well as bird assemblage patterns (Edwards et al. 2016; Farrar 2004). USA environmental regulatory agencies have been using ecoregions for developing and implementing biological criteria for fish and macroinvertebrate assemblages (Davis et al. 1996; Stoddard 2004; Stoddard et al. 2008; Yoder and Rankin 1998). Those ecoregional assemblage patterns and aquatic biological criteria are driven by broad ecoregional patterns in water quality and physical habitat structure (Heiskary and Wilson 1989; Kaufmann et al. 2022a; Whittier et al. 1988). Although our study was based on data and ecoregions of the conterminous USA, differing environment-assemblage drivers and patterns have been documented elsewhere. Patterns differed in Brazil for neighbouring Amazonian and Atlantic Forest fish assemblages (Leitão et al. 2018; Pinto et al. 2009) and Cerrado macroinvertebrate assemblages (Agra et al. 2019; Martins et al. 2018; Saito et al. 2016), and nationwide for both fish and macroinvertebrates (Dala-Corte et al. 2020; Lansac-Tôha et al. 2021). Likewise, ecoregional typologies are important components in the European Union for developing and implementing biomonitoring indicators and programmes (Hering et al. 2010; Schmutz et al. 2007).

Although our results seem pertinent to stream ecology, monitoring, and management, they are only correlations, and we lacked data for all conceivable predictor variables. We did not suggest mechanisms for the predictor-response relationships that we reported. However, the water quality and habitat structure variables that we found to be important for predicting ecoregional and conterminous USA macroinvertebrate genus abundances were amongst the most important reported by the USEPA for explaining risks for poor macroinvertebrate index scores (Herlihy et al. 2020; Kaufmann et al. 2022a; Kaufmann et al. 2022b; Paulsen, Mayio, et al. 2008). They were also listed amongst the most important variables for estimating risk to macroinvertebrate assemblages (Jiménez-Valencia et al. 2014; Silva et al. 2018) and explaining macroinvertebrate alpha diversity in tropical and boreal streams (Heino et al. 2018; Macedo et al. 2014) and beta diversity in conterminous USA streams (Bini et al. 2014). Those relationships offer strong evidence for underlying mechanisms. We used a limited number of predictor environmental variables because those were the ones most consistently sampled. Nonetheless, they demonstrated the expected patterns: ecoregional differences, the importance of site variables, and effect-distance relationships. Although limited to four, the climate variables demonstrated the importance of ecoregional-extent predictors as well as the importance of considering effect-distance relationships. We suggest that additional regional-extent predictors would only add further evidence of those climate effect-distance relationships (Herlihy et al. 2020; Macedo et al. 2014; Silva et al. 2018).

Categorical differences in environmental drivers of macroinvertebrate composition amongst ecoregions are relevant not only

for ecological studies, but for the management and conservation of aquatic ecosystems as well. Identifying the correct research and monitoring spatial extents and the most important driving variables is essential for maintaining biological diversity and ecosystem services because they facilitate the development of increasingly better management strategies (Moraga et al. 2019; Patrick et al. 2021). Clearly, the strategies adopted must be consistent with the set of variables important for the region of interest. Because ecoregions present differences in these sets, monitoring and regulatory programs should be aware of their usefulness, limitations, and peculiarities, and calibrate biological expectations for natural environmental gradients within the ecoregions (Chen et al. 2019; Hawkins et al. 2000; Moya et al. 2011; Pont et al. 2006, 2009).

Author Contributions

Bruno S. Godoy: conceptualisation, data curation, formal analysis, methodology, software, supervision, visualisation, writing – original draft, writing – review and editing; Tadeu Siqueira: conceptualisation, validation, writing – review and editing; Robert M. Hughes: validation, writing – review and editing.

Acknowledgements

T.S. received grant #2019/04033-7 from the São Paulo Research Foundation (FAPESP). R.M.H. received a Fulbright Brazil distinguished scholar award. We also thank the two anonymous reviewers for their significant contributions, which helped improve the manuscript. The Article Processing Charge for the publication of this research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) (ROR identifier: 00x0ma614).

Ethics Statement

The study presented here did not require permits, as it was conducted using previously collected data from the US EPA.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data and R codes are available at Zenodo in: https://doi.org/10.5281/zenodo.14911674.

References

Agra, J. U. M., R. Ligeiro, D. R. Macedo, R. M. Hughes, and M. Callisto. 2019. "Ecoregions and Stream Types Help Us Understand Ecological Variability in Neotropical Reference Streams." *Marine and Freshwater Research* 70, no. 4: 594. https://doi.org/10.1071/MF18309.

Arita, H. T., and P. Rodríguez. 2004. "Local-Regional Relationships and the Geographical Distribution of Species." *Global Ecology and Biogeography* 13: 15–21.

Barbour, M. T., J. Gerritsen, B. Snyder, and J. Stribling. 1999. *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish*. Second ed. Project Officer. http://www.epa.gov/OWOW/monitoring/techmon.html.

Barreiros, N. M., T. Giarrizzo, and B. S. Godoy. 2023. "Beta Diversity of Ephemeroptera, Plecoptera and Trichoptera on Multiples Spatial Extents in Xingu River Rapids." *Acta Limnologica Brasiliensia* 35: e23. https://doi.org/10.1590/s2179-975x2923.

Beisner, B. E., D. T. Haydon, and K. Cuddington. 2003. "Alternative Stable States in Ecology." *Frontiers in Ecology and the Environment* 1: 376–382. https://doi.org/10.1890/1540-9295(2003)001[0376:ASSIE]2.0.CO;2.

Bini, L. M., V. L. Landeiro, A. A. Padial, T. Siqueira, and J. Heino. 2014. "Nutrient Enrichment Is Related to Two Facets of Beta Diversity for Stream Invertebrates Across the United States." *Ecology* 95, no. 6: 1569–1578. https://doi.org/10.1890/13-0656.1.

Borcard, D., and P. Legendre. 2002. "All-Scale Spatial Analysis of Ecological Data by Means of Principal Coordinates of Neighbour Matrices." *Ecological Modelling* 153, no. 1–2: 51–68. https://doi.org/10.1016/S0304-3800(01)00501-4.

Boyero, L. 2003. "Multiscale Patterns of Spatial Variation in Stream Macroinvertebrate Communities." *Ecological Research* 18: 365–379. https://doi.org/10.1046/j.1440-1703.2003.00562.x.

Burgazzi, G., A. Laini, O. Ovaskainen, M. Saccó, R. Stubbington, and P. Viaroli. 2020. "Communities in High Definition: Spatial and Environmental Factors Shape the Micro-Distribution of Aquatic Invertebrates." *Freshwater Biology* 65: 2053–2065. https://doi.org/10.1111/fwb.13599.

Chen, K., A. R. Rajper, R. M. Hughes, J. R. Olson, H. Wei, and B. Wang. 2019. "Incorporating Functional Traits to Enhance Multimetric Index Performance and Assess Land Use Gradients." *Science of the Total Environment* 691: 1005–1015. https://doi.org/10.1016/j.scitotenv.2019.07.047.

Dala-Corte, R. B., A. S. Melo, T. Siqueira, et al. 2020. "Thresholds of Freshwater Biodiversity in Response to Riparian Vegetation Loss in the Neotropical Region." *Journal of Applied Ecology* 57, no. 7: 1391–1402. https://doi.org/10.1111/1365-2664.13657.

Davis, W. S., B. D. Snyder, J. B. Stribling, and C. Stoughton. 1996. "Summary of State Biological Assessment Programs for Streams and Wadeable Rivers." EPA230-R-96-007.

Doretto, A., J. Receveur, E. Baker, M. Benbow, and K. Scribner. 2021. "Nested Analysis of Macroinvertebrate Diversity Along a River Continuum: Identifying Relevant Spatial Scales for Stream Communities." *River Research and Applications* 38: 334–344. https://doi.org/10.1002/rra.3889.

Dray, S., R. Pélissier, P. Couteron, et al. 2012. "Community Ecology in the Age of Multivariate Multiscale Spatial Analysis." *Ecological Monographs* 82, no. 3: 257–275. https://doi.org/10.1890/11-1183.1.

Edwards, L., J. Ambrose, and L. Kirkman. 2016. *The Natural Communities of Georgia*. University of Geoia Press.

Fanelli, R. M., M. J. Cashman, and A. J. Porter. 2022. "Identifying Key Stressors Driving Biological Impairment in Freshwater Streams in the Chesapeake Bay Watershed, USA." *Environmental Management* 70: 926–949. https://doi.org/10.1007/s00267-022-01723-7.

Farrar, J. 2004. "Birding Nebraska." *Nebraskaland* 82, no. 1: 178.

Fukami, T. 2015. "Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects." *Annual Review of Ecology, Evolution, and Systematics* 46: 1–23. https://doi.org/10.1146/annurev-ecolsys-110411-160340.

Fukami, T., E. A. Mordecai, and A. Ostling. 2016. "A Framework for Priority Effects." *Journal of Vegetation Science* 27, no. 4: 655–657. https://doi.org/10.1111/jvs.12434.

Fukami, T., and M. Nakajima. 2011. "Community Assembly: Alternative Stable States or Alternative Transient States?" *Ecology Letters* 14: 973–984. https://doi.org/10.1111/j.1461-0248.2011.01663.x.

Godoy, B. S., A. P. J. Faria, L. Juen, L. Sara, and L. G. Oliveira. 2019. "Taxonomic Sufficiency and Effects of Environmental and Spatial Drivers on Aquatic Insect Community." *Ecological Indicators* 107: 105624. https://doi.org/10.1016/j.ecolind.2019.105624.

Godoy, B. S., L. L. Queiroz, S. Lodi, and L. G. Oliveira. 2017. "Environment and Spatial Influences on Aquatic Insect Communities

in Cerrado Streams: The Relative Importance of Conductivity, Altitude, and Conservation Areas." *Neotropical Entomology* 46, no. 2: 151–158. https://doi.org/10.1007/s13744-016-0452-4.

Godoy, B. S., L. L. Queiroz, J. Simião-Ferreira, S. Lodi, L. M. Camargos, and L. G. Oliveira. 2022. "The Effect of Spatial Scale on the Detection of Environmental Drivers on Aquatic Insect Communities in Pristine and Altered Streams of the Brazilian Cerrado." *International Journal of Tropical Insect Science* 42, no. 3: 2173–2182. https://doi.org/10.1007/s42690-022-00738-1.

Gravel, D., C. D. Canham, M. Beaudet, and C. Messier. 2006. "Reconciling Niche and Neutrality: The Continuum Hypothesis." *Ecology Letters* 9, no. 4: 399–409.

Guénard, G., and P. Legendre. 2018a. "Bringing Multivariate Support to Multiscale Codependence Analysis: Assessing the Drivers of Community Structure Across Spatial Scales." *Methods in Ecology and Evolution* 9, no. 2: 292–304. https://doi.org/10.1111/2041-210X.12864.

Guénard, G., and P. Legendre. 2018b. "codep: Multiscale Codependence Analysis." R package version 0.9–1.

Guénard, G., P. Legendre, D. Boisclair, and M. Bllodeau. 2010. "Multiscale Codependence Analysis: An Integrated Approach to Analyze Relationships Across Scales." *Ecology* 91, no. 10: 2952–2964. https://doi.org/10.1890/09-0460.1.

Harding, J. S., E. F. Benfield, P. V. Bolstad, G. S. Helfman, and E. B. D. Jones. 1998. "Stream Biodiversity: The Ghost of Land Use Past." *Proceedings of the National Academy of Sciences of the United States of America* 95: 14843–14847. https://doi.org/10.1073/pnas.95.25.14843.

Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. W. Feminella. 2000. "Development and Evaluation of Predictive Models for Measuring the Biological Integrity of Streams." *Ecological Applications* 10: 1456–1477. https://doi.org/10.1890/1051-0761(2000)010[1456:DAEOPM]2.0.CO;2.

Heino, J. 2002. "Concordance of Species Richness Patterns Among Multiple Freshwater Taxa: A Regional Perspective." *Biodiversity and Conservation* 11, no. 1: 137–147. https://doi.org/10.1023/A:1014075901605.

Heino, J. 2009. "Biodiversity of Aquatic Insects: Spatial Gradients and Environmental Correlates of Assemblage-Level Measures at Large Scales." *Freshwater Reviews* 2, no. 1: 1–29. https://doi.org/10.1608/FRJ-2.1.1.

Heino, J. 2011. "A Macroecological Perspective of Diversity Patterns in the Freshwater Realm." *Freshwater Biology* 56, no. 9: 1703–1722. https://doi.org/10.1111/j.1365-2427.2011.02610.x.

Heino, J., A. S. Melo, J. Jyrkänkallio-Mikkola, et al. 2018. "Subtropical Streams Harbour Higher Genus Richness and Lower Abundance of Insects Compared to Boreal Streams, but Scale Matters." *Journal of Biogeography* 45, no. 9: 1983–1993. https://doi.org/10.1111/jbi.13400.

Heiskary, S. A., and C. B. Wilson. 1989. "The Regional Nature of Lake Water Quality Across Minnesota: An Analysis for Improving Resource Management." *Journal of the Minnesota Academy of Science* 55: 71–77.

Hering, D., A. Borja, J. Carstensen, et al. 2010. "The European Water Framework Directive at the Age of 10: A Critical Review of the Achievements With Recommendations for the Future." *Science of the Total Environment* 408, no. 19: 4007–4019. https://doi.org/10.1016/j.scitotenv.2010.05.031.

Herlihy, A. T., J. C. Sifneos, R. M. Hughes, D. V. Peck, and R. M. Mitchell. 2020. "The Relation of Lotic Fish and Benthic Macroinvertebrate Condition Indices to Environmental Factors Across the Conterminous USA." *Ecological Indicators* 112: 105958. https://doi.org/10.1016/j.ecolind.2019.105958.

Hughes, R. M., S. A. Heiskary, W. J. Matthews, and C. O. Yoder. 1994. "Use of Ecoregions in Biological Monitoring." In *Biological Monitoring of Freshwater Ecosystems*, edited by L. L. Stanford and A. Spacie, 1st ed., 125–151. Lewis Press.

Hughes, R. M., E. Rexstad, and C. E. Bond. 1987. "The Relationship of Aquatic Ecoregions, River Basins and Physiographic Provinces to the Ichthyogeographic Regions of Oregon." *Copeia* 1987, no. 2: 423. https://doi.org/10.2307/1445780.

Hughes, R. M., L. Wang, and P. W. Sellbach. 2006. *Landscape Influences on Stream Habitats and Biological Assemblages* (Hughes, R. M., Wang, L., & Seelbach, P. W., Eds.; 48th ed.). American Fisheries Society. https://doi.org/10.47886/9781888569766.

Jiménez-Valencia, J., P. R. Kaufmann, A. Sattamini, R. Mugnai, and D. F. Baptista. 2014. "Assessing the Ecological Condition of Streams in a Southeastern Brazilian Basin Using a Probabilistic Monitoring Design." *Environmental Monitoring and Assessment* 186, no. 8: 4685–4695. https://doi.org/10.1007/s10661-014-3730-9.

Johnson, R., W. Goedkoop, and L. Sandin. 2004. "Spatial Scale and Ecological Relationships Between the Macroinvertebrate Communities of Stony Habitats of Streams and Lakes." *Freshwater Biology* 49: 1179–1194. https://doi.org/10.1111/J.1365-2427.2004.01262.X.

Karaouzas, I., C. Theodoropoulos, A. Vourka, K. Gritzalis, and N. Skoulikidis. 2019. "Stream Invertebrate Communities Are Primarily Shaped by Hydrological Factors and Ultimately Fine-Tuned by Local Habitat Conditions." *Science of the Total Environment* 665: 290–299. https://doi.org/10.1016/j.scitotenv.2019.02.134.

Kärnä, O. M., M. Grönroos, H. Antikainen, et al. 2015. "Inferring the Effects of Potential Dispersal Routes on the Metacommunity Structure of Stream Insects: As the Crow Flies, as the Fish Swims or as the Fox Runs?" *Journal of Animal Ecology* 84: 1342–1353. https://doi.org/10.1111/1365-2656.12397.

Kaufmann, P. R., R. M. Hughes, S. G. Paulsen, et al. 2022a. "Physical Habitat in Conterminous US Streams and Rivers, Part 2: A Quantitative Assessment of Habitat Condition." *Ecological Indicators* 141: 109047. https://doi.org/10.1016/j.ecolind.2022.109047.

Kaufmann, P. R., R. M. Hughes, S. G. Paulsen, et al. 2022b. "Physical Habitat in Conterminous US Streams and Rivers, Part 1: Geoclimatic Controls and Anthropogenic Alteration." *Ecological Indicators* 141: 109046. https://doi.org/10.1016/j.ecolind.2022.109046.

Kaufmann, P. R., R. M. Hughes, T. R. Whittier, S. A. Bryce, and S. G. Paulsen. 2014. "Relevance of Lake Physical Habitat Indices to Fish and Riparian Birds." *Lake and Reservoir Management* 30, no. 2: 177–191. https://doi.org/10.1080/10402381.2013.877544.

Lacey, R. W. J., P. Legendre, and A. G. Roy. 2007. "Spatial-Scale Partitioning of In Situ Turbulent Flow Data Over a Pebble Cluster in a Gravel-Bed River." *Water Resources Research* 43, no. 3: W03416. https://doi.org/10.1029/2006WR005044.

Lansac-Tôha, F. M., L. M. Bini, J. Heino, et al. 2021. "Scale-Dependent Patterns of Metacommunity Structuring in Aquatic Organisms Across Floodplain Systems." *Journal of Biogeography* 48, no. 4: 872–885. https://doi.org/10.1111/JBI.14044.

Larsen, D. P., J. M. Omernik, R. M. Hughes, et al. 1986. "Correspondence Between Spatial Patterns in Fish Assemblages in Ohio Streams and Aquatic Ecoregions." *Environmental Management* 10: 815–828.

Leal, C. G., J. Barlow, T. A. Gardner, et al. 2018. "Is Environmental Legislation Conserving Tropical Stream Faunas? A Large-Scale Assessment of Local, Riparian and Catchment-Scale Influences on Amazonian Fish." *Journal of Applied Ecology* 55, no. 3: 1312–1326. https://doi.org/10.1111/1365-2664.13028.

Leibold, M. A., and J. M. Chase. 2018. *Metacommunity Ecology*. 1st ed. Princeton University Press.

Leitão, R. P., J. Zuanon, D. Mouillot, et al. 2018. "Disentangling the Pathways of Land Use Impacts on the Functional Structure of Fish Assemblages in Amazon Streams." *Ecography* 41, no. 1: 219–232. https://doi.org/10.1111/ecog.02845.

Li, F., N. Chung, M. Bae, Y. Kwon, and Y. Park. 2012. "Relationships Between Stream Macroinvertebrates and Environmental Variables at Multiple Spatial Scales." *Freshwater Biology* 57: 2107–2124. https://doi.org/10.1111/J.1365-2427.2012.02854.X.

Li, Z., X. Chen, X. Jiang, J. D. Tonkin, Z. Xie, and J. Heino. 2021. "Distance Decay of Benthic Macroinvertebrate Communities in a Mountain River Network: Do Dispersal Routes and Dispersal Ability Matter?" *Science of the Total Environment* 758: 143630. https://doi.org/10.1016/j.scitotenv.2020.143630.

Li, Z., J. Heino, X. Chen, et al. 2021. "Understanding Macroinvertebrate Metacommunity Organization Using a Nested Study Design Across a Mountainous River Network." *Ecological Indicators* 121: 107188. https://doi.org/10.1016/j.ecolind.2020.107188.

Linares, M. S., D. R. Macedo, J. C. Marques, R. M. Hughes, and M. Callisto. 2025. "Biodiversity Spatial Distribution of Benthic Macroinvertebrate Assemblages Is Influenced by Anthropogenic Disturbances at Multiple Spatial Extents." *Science of the Total Environment* 960: 178365. https://doi.org/10.1016/J.SCITOTENV.2024.178365.

Macedo, D. R., R. M. Hughes, R. Ligeiro, et al. 2014. "The Relative Influence of Catchment and Site Variables on Fish and Macroinvertebrate Richness in Cerrado Biome Streams." *Landscape Ecology* 29, no. 6: 1001–1016. https://doi.org/10.1007/s10980-014-0036-9.

Martins, I., R. Ligeiro, R. M. Hughes, D. R. Macedo, and M. Callisto. 2018. "Regionalisation Is Key to Establishing Reference Conditions for Neotropical Savanna Streams." *Marine and Freshwater Research* 69, no. 1: 82. https://doi.org/10.1071/MF16381.

McCormick, F. H., D. V. Peck, and D. P. Larsen. 2000. "Comparison of Geographic Classification Schemes for Mid-Atlantic Stream Fish Assemblages." *Journal of the North American Benthological Society* 19, no. 3: 385–404. https://doi.org/10.2307/1468102.

Moi, D. A., M. B. Gómez, M. Burwood, et al. 2023. "Human Land-Uses Homogenize Stream Assemblages and Reduce Animal Biomass Production." *Journal of Animal Ecology* 92, no. 6: 1176–1189. https://doi.org/10.1111/1365-2656.13924.

Moraga, A. D., A. E. Martin, and L. Fahrig. 2019. "The Scale of Effect of Landscape Context Varies With the Species' Response Variable Measured." *Landscape Ecology* 34, no. 4: 703–715. https://doi.org/10.1007/s10980-019-00808-9.

Moya, N., R. M. Hughes, E. Domínguez, F.-M. Gibon, E. Goitia, and T. Oberdorff. 2011. "Macroinvertebrate-Based Multimetric Predictive Models for Evaluating the Human Impact on Biotic Condition of Bolivian Streams." *Ecological Indicators* 11, no. 3: 840–847. https://doi.org/10.1016/j.ecolind.2010.10.012.

Oksanen, J., F. G. Blanchet, M. Friendly, et al. 2020. "Community Ecology Package." R package version 2.5–6.

Oliveira, J. S., R. M. Hughes, and B. d. F. Terra. 2024. "Fish and Macroinvertebrates Respond Differently to Seasonal Drying in Tropical Non-Perennial Streams." *Austral Ecology* 49, no. 7: e13558. https://doi.org/10.1111/AEC.13558.

Olsen, A. R., and D. V. Peck. 2008. "Survey Design and Extent Estimates for the Wadeable Streams Assessment." *Journal of the North American Benthological Society* 27, no. 4: 822–836. https://doi.org/10.1899/08-050.1.

Omernik, J. M. 1987. "Ecoregions of the Conterminous United States." *Annals of the Association of American Geographers* 77, no. 1: 118–125. https://doi.org/10.1111/j.1467-8306.1987.tb00149.x.

Omernik, J. M., and G. E. Griffith. 1991. "Ecological Regions Versus Hydrologic Units: Frameworks for Managing Water Quality." *Journal of Soil and Water Conservation* 48: 334–340.

Omernik, J. M., and G. E. Griffith. 2014. "Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework." *Environmental Management* 54, no. 6: 1249–1266. https://doi.org/10.1007/s00267-014-0364-1.

Omernik, J. M., G. E. Griffith, R. M. Hughes, J. B. Glover, and M. H. Weber. 2017. "How Misapplication of the Hydrologic Unit Framework Diminishes the Meaning of Watersheds." *Environmental Management* 60, no. 1: 1–11. https://doi.org/10.1007/s00267-017-0854-z.

Patrick, C. J., K. E. Anderson, B. L. Brown, et al. 2021. "The Application of Metacommunity Theory to the Management of Riverine Ecosystems." *Wiley Interdisciplinary Reviews: Water* 8, no. 6: e1557. https://doi.org/10.1002/WAT2.1557.

Patrick, C. J., and L. L. Yuan. 2019. "The Challenges That Spatial Context Present for Synthesizing Community Ecology Across Scales." *Oikos* 128, no. 3: 297–308. https://doi.org/10.1111/oik.05802.

Paulsen, S. G., C. P. Hawkins, J. Van Sickle, L. L. Yuan, and S. M. Holdsworth. 2008. "An Invitation to Apply National Survey Data to Ecological Research." *Journal of the North American Benthological Society* 27: 1017–1018. https://doi.org/10.1899/08-123.1.

Paulsen, S. G., A. Mayio, D. V. Peck, et al. 2008. "Condition of Stream Ecosystems in the US: An Overview of the First National Assessment." *Journal of the North American Benthological Society* 27: 812–821. https://doi.org/10.1899/08-098.1.

Pinto, B. C. T., F. G. Araujo, V. D. Rodrigues, and R. M. Hughes. 2009. "Local and Ecoregion Effects on Fish Assemblage Structure in Tributaries of the Rio Paraíba Do Sul, Brazil." *Freshwater Biology* 54, no. 12: 2600–2615. https://doi.org/10.1111/j.1365-2427.2009.02269.x.

Pont, D., R. M. Hughes, T. R. Whittier, and S. Schmutz. 2009. "A Predictive Index of Biotic Integrity Model for Aquatic-Vertebrate Assemblages of Western U.S. Streams." *Transactions of the American Fisheries Society* 138: 292–305.

Pont, D., B. Hugueny, U. Beier, et al. 2006. "Assessing River Biotic Condition at a Continental Scale: A European Approach Using Functional Metrics and Fish Assemblages." *Journal of Applied Ecology* 43, no. 1: 70–80. https://doi.org/10.1111/j.1365-2664.2005.01126.x.

PRISM Climate Group. 2023. "PRISM Gridded Climate Data." Https://Prism.Oregonstate.Edu.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Rohm, C. M., J. W. Giese, and C. C. Bennett. 1987. "Evaluation of an Aquatic Ecoregion Classification of Streams in Arkansas." *Journal of Freshwater Ecology* 4, no. 1: 127–140. https://doi.org/10.1080/02705060.1987.9665169.

Roque, F. O., T. Siqueira, L. M. Bini, et al. 2010. "Untangling Associations Between Chironomid Taxa in Neotropical Streams Using Local and Landscape Filters." *Freshwater Biology* 55, no. 4: 847–865. https://doi.org/10.1111/J.1365-2427.2009.02314.X.

Saito, V. S., M. V. Cianciaruso, T. Siqueira, A. A. Fonseca-Gessner, and S. Pavoine. 2016. "Phylogenies and Traits Provide Distinct Insights About the Historical and Contemporary Assembly of Aquatic Insect Communities." *Ecology and Evolution* 6, no. 9: 2925–2937. https://doi.org/10.1002/ece3.2081.

Santos, E. P., H. H. Wagner, S. F. B. Ferraz, and T. Siqueira. 2020. "Interactive Persistent Effects of Past Land-Cover and Its Trajectory on Tropical Freshwater Biodiversity." *Journal of Applied Ecology* 57, no. 11: 2149–2158. https://doi.org/10.1111/1365-2664.13717.

Schmutz, S., I. G. Cowx, G. Haidvogl, and D. Pont. 2007. "Fish-Based Methods for Assessing European Running Waters: A Synthesis." *Fisheries Management and Ecology* 14, no. 6: 369–380. https://doi.org/10.1111/j.1365-2400.2007.00585.x.

Schneck, F., L. M. Bini, A. S. Melo, et al. 2022. "Catchment Scale Deforestation Increases the Uniqueness of Subtropical Stream Communities." *Oecologia* 199, no. 3: 671–683. https://doi.org/10.1007/S00442-022-05215-7.

Schröder, A., L. Persson, and A. M. De Roos. 2005. "Direct Experimental Evidence for Alternative Stable States: A Review." *Oikos* 110, no. 1: 3–19. https://doi.org/10.1111/j.0030-1299.2005.13962.x.

Silva, D. R. O., A. T. Herlihy, R. M. Hughes, D. R. Macedo, and M. Callisto. 2018. "Assessing the Extent and Relative Risk of Aquatic Stressors on Stream Macroinvertebrate Assemblages in the Neotropical Savanna." *Science of the Total Environment* 633: 179–188. https://doi.org/10.1016/j.scitotenv.2018.03.127.

Siqueira, T., L. D. Durães, and F. de Oliveira Roque. 2014. "Predictive Modelling of Insect Metacommunities in Biomonitoring of Aquatic Networks." In *Ecological Modelling Applied to Entomology*, edited by C. Ferreira and W. Godoy, vol. 109–126. Springer International Publishing. https://doi.org/10.1007/978-3-319-06877-0_5.

Soininen, J., R. McDonald, and H. Hillebrand. 2007. "The Distance Decay of Similarity in Ecological Communities." *Ecography* 30: 3–12. https://doi.org/10.1111/j.0906-7590.2007.04817.x.

Stoddard, J. L. 2004. "Use of Ecological Regions in Aquatic Assessments of Ecological Condition." *Environmental Management* 34, no. S1: S61–S70. https://doi.org/10.1007/s00267-003-0193-0.

Stoddard, J. L., A. T. Herlihy, D. V. Peck, R. M. Hughes, T. R. Whittier, and E. Tarquinio. 2008. "A Process for Creating Multimetric Indices for Large-Scale Aquatic Surveys." *Journal of the North American Benthological Society* 27, no. 4: 878–891. https://doi.org/10.1899/08-053.1.

Tabi, A., T. Siqueira, and J. D. Tonkin. 2024. "Species Interactions Drive Continuous Assembly of Freshwater Communities in Stochastic Environments." *Scientific Reports* 14, no. 1: 1–8. https://doi.org/10.1038/s41598-024-72405-z.

Thompson, P. L., L. M. Guzman, L. De Meester, et al. 2020. "A Process-Based Metacommunity Framework Linking Local and Regional Scale Community Ecology." *Ecology Letters* 23: 1314–1329. https://doi.org/10.1111/ele.13568.

Tonn, W. M. 1990. "Climate Change and Fish Communities: A Conceptual Framework." *Transactions of the American Fisheries Society* 119, no. 2: 337–352. https://doi.org/10.1577/1548-8659(1990)119<0337:CCAFCA>2.3.CO;2.

Toskey, E. K., S. M. Bollens, P. M. Kiffney, K. D. Martens, and G. Rollwagen-Bollens. 2024. "The Relative Importance of Abiotic, Biotic, and Spatial Factors in Structuring the Stream Macroinvertebrate Metacommunity in a Temperate Rainforest." *Aquatic Sciences* 86: 110. https://doi.org/10.1007/s00027-024-01122-6.

USEPA. 2024. *National Rivers and Streams Assessment 2008–2019: A Collaborative Survey*. USEPA Office of Water.

Van Sickle, J., and R. M. Hughes. 2000. "Classification Strengths of Ecoregions, Catchments, and Geographic Clusters for Aquatic Vertebrates in Oregon." *Journal of the North American Benthological Society* 19, no. 3: 370–384. https://doi.org/10.2307/1468101.

Vellend, M. 2010. "Conceptual Synthesis in Community Ecology." *Quarterly Review of Biology* 85: 183–206. https://doi.org/10.1086/652373.

Waite, I. R., A. T. Herlihy, D. P. Larsen, and D. J. Klemm. 2000. "Comparing Strengths of Geographic and Nongeographic Classifications of Stream Benthic Macroinvertebrates in the Mid-Atlantic Highlands, USA." *Journal of the North American Benthological Society* 19, no. 3: 429–441. https://doi.org/10.2307/1468105.

Wang, L., P. W. Sellbach, and R. M. Hughes. 2006. "Introduction to Influences of Landscape on Stream Habitat and Biological Assemblages." In *Landscape Influences on Stream Habitat and Biological Assemblages*, edited by R. M. Hughes, L. Wang, and P. W. Sellbach. American Fisheries Society.

Whittier, T. R., R. M. Hughes, and D. P. Larsen. 1988. "Correspondence Between Ecoregions and Spatial Patterns in Stream Ecosystems in Oregon." *Canadian Journal of Fisheries and Aquatic Sciences* 45, no. 7: 1264–1278. https://doi.org/10.1139/f88-149.

Wu, N., G. Liu, M. Zhang, Y. Wang, W. Peng, and X. Qu. 2022. "Spatial Factors Outperform Local Environmental and Geo-Climatic Variables in Structuring Multiple Facets of Stream Macroinvertebrates' β-Diversity." *Animals: An Open Access Journal From MDPI* 12: 2648. https://doi.org/10.3390/ani12192648.

Yoder, C. O., and E. T. Rankin. 1998. "The Role of Biological Indicators in a State Water Quality Management Process." *Environmental Monitoring and Assessment* 51: 61–88.

Zhou, S.-C., N. Wu, M. Zhang, et al. 2020. "Local Environmental, Geo-Climatic and Spatial Factors Interact to Drive Community Distributions and Diversity Patterns of Stream Benthic Algae, Macroinvertebrates and Fishes in a Large Basin, Northeast China." *Ecological Indicators* 117: 106673. https://doi.org/10.1016/j.ecolind.2020.106673.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.