Global Ecology
and Biogeography

A Journal of
Macroecology

WILEY

Bioregions are predominantly climatic for fishes of northern lakes

Charlie J. G. Loewen[1] ID | Donald A. Jackson[1] ID | Cindy Chu[1,2] ID | Karen M. Alofs[3] ID |
Gretchen J. A. Hansen[4] ID | Andrew E. Honsey[4] ID | Charles K. Minns[1] ID | Kevin E. Wehrly[3,5]

[1]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada

[2]Ontario Ministry of Northern Development, Mines, Natural Resources and Forestry, Peterborough, Ontario, Canada

[3]School for Environment and Sustainability, University of Michigan, Ann Arbor, Michigan, USA

[4]Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, St. Paul, Minnesota, USA

[5]Institute for Fisheries Research, Michigan Department of Natural Resources, Ann Arbor, Michigan, USA

Correspondence
Charlie J. G. Loewen, Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, ON, Canada M5S 3B2.
Email: charlie.loewen@utoronto.ca

Present address
Andrew E. Honsey, Hammond Bay Biological Station, United States Geological Survey Great Lakes Science Center, Millersburg, Michigan, USA

Abstract

Aim: Recurrent species assemblages integrate important biotic interactions and joint responses to environmental and spatial filters that enable local coexistence. Here, we applied a bipartite (site–species) network approach to develop a natural typology of lakes sharing distinct fish faunas and provide a detailed, hierarchical view of their bioregions. We then compared the roles of key biogeographical factors to evaluate alternative hypotheses about how fish communities are assembled from the regional species pool.

Location: Ontario, Canada and the Upper Midwest, USA.

Time period: 1957–2017.

Major taxa studied: Freshwater fishes.

Methods: Bipartite modularity analysis was performed on 90 taxa from 10,016 inland lakes in the Southwestern Hudson Bay, Mississippi River and St. Lawrence River drainages, uncovering bioregionalization of North American fishes at a large, subcontinental scale. We then used a latent variable approach, pairing non-metric partial least-squares structural equation modelling with multiple logistic regression, to show differences in the biogeographical templates of each type of community. Indicators of contemporary and historical connectivity, climate and habitat constructs were estimated using a geographical information system.

Results: Fish assemblages reflected broad, overlapping patterns of postglacial colonization, climate and geological setting, but community differentiation was most linked to temperature, precipitation and, for certain groups, lake area and water quality. Bioregions were also marked by non-native species, showing broad-scale impacts of introductions to the Great Lakes and surrounding basins.

Main conclusions: The dominant effects of climate across broad spatial gradients indicate differing sensitivities of fish communities to rapidly accelerating climate change and opportunities for targeted conservation strategies. By assessing biological variation at the level of recurrent assemblages, we accounted for the non-stationarity of macroecological processes structuring different sets of species on the landscape and offer novel inference on the assembly of inland fish communities.

1 | INTRODUCTION

Species distributions and their assembly into communities reflect an array of contemporary and historical filters operating at multiple spatial scales (Dias et al., 2014; Ficetola et al., 2017; Jackson et al., 2001; Mantyka-Pringle et al., 2014). Studies seeking to understand sources of variation across ecological communities have often tested their compositional turnover or aggregate properties, such as species richness (e.g., Legendre & Legendre, 1984; Mandrak, 1995; Oikonomou et al., 2014). More recently, tools drawn from graph theory have been applied to reveal community typologies and natural boundaries of bioregions (i.e., geographical areas with similar biota) based on multispecies distribution patterns (e.g., Bloomfield et al., 2018; Carstensen & Olesen, 2009; McGarvey & Veech, 2018). Given that co-occurrence implies broadly shared niche characteristics and colonization history, repeated assemblages provide insight into formative ecological processes and a basis for setting habitat management units for conservation planning and targeted interventions (Brooks et al., 2006; Montalvo-Mancheno et al., 2020; Olden et al., 2010; Wehrly et al., 2012). Community differentiation is especially informative for freshwater organisms, such as fishes, which face strong dispersal constraints imposed by the hydrology of the watersheds they inhabit. Freshwater ecosystems exhibit high biodiversity and rates of endemism globally, in addition to generally greater threats and species imperilment than terrestrial or marine habitats (Reid et al., 2019). Nevertheless, in comparison to other realms, inland waters have received little attention paid to delineating recurrent assemblages and their underlying drivers at the broad spatial scales required to develop regional, cross-jurisdictional conservation and management strategies (Brooks et al., 2006; Heino et al., 2021; Oikonomou et al., 2014; Wehrly et al., 2012).

Our ability to predict changes in species diversity is challenged by the multitude of contributing spatial and environmental factors (e.g., Dormann et al., 2013; Loewen et al., 2020; Mantyka-Pringle et al., 2014). For instance, spatial patterns of northern fish communities may reflect historical imprints of glaciation (Bailey & Smith, 1981; Dias et al., 2014; Legendre & Legendre, 1984), in addition to present-day dispersal networks and regional habitat availability (e.g., Hitt & Angermeier, 2008). Given that most freshwater organisms are ectothermic, they are also intrinsically sensitive to temperature (Magnuson et al., 1979). Thus, although the dependence of fish on hydrological connections to reach new habitats limits their ability to move in response to interannual climatic variations, longer-term temperature conditions should constrain the distributional boundaries of species by restricting the fitness of edge populations (Alofs et al., 2014; Shuter & Post, 1990). Persistence of species arriving at a new location also depends on habitat quantity and quality. Larger

waterbodies may offer more space to more species (MacArthur & Wilson, 1963) and a greater diversity of habitats, including competitive or predatory refugia (Tonn & Magnuson, 1982). Aspects of lake basin morphometry, along with terrestrial (Heino et al., 2021) and geological setting (Conroy & Keller, 1976), also drive gradients in surface water quality and biological production (D'Arcy & Carignan, 1997). Thus, local habitat conditions combine with regional connectivity and climate filters to provide a multilayered template for fish species sorting along geographical gradients (Smith & Powell, 1971).

Our objectives were twofold. First, we leveraged a unique dataset of fish records across the Laurentian Shield and surrounding Lowlands of Ontario, Canada and the Upper Midwest, USA, to develop a site–species network graph. We then applied bipartite modularity analysis to detect community typology and provide a hierarchical view of freshwater fish bioregionalization. Second, we tested alternative hypotheses about the relative importance of biogeographical factors driving community differentiation, predicting varying roles of connectivity, climate and habitat filters across groups. For instance, we hypothesized a key role of climate for northern fishes, with contrasting responses to temperature and precipitation (owing to influence on lake thermal, chemical and hydrological regimes, including timing of ice-out and major runoff events) revealing sensitive regions and community types to target in future climate change adaptation strategies. We expected connectivity to constrain communities both within drainages and in relationship to deglaciation patterns, because glacial lakes provided an important means of dispersal, and southern sites might have had more time for postglacial colonization. Habitat filters were predicted to be of lesser importance at broad spatial scales, although we anticipated potential differentiation of communities linked to contrasting shield (metamorphic and igneous) and sedimentary lithologies as mediated by their effects on water chemistry (e.g., alkalinity and nutrient levels).

2 | MATERIALS AND METHODS

2.1 | Study system

We conducted a large data synthesis to assess fish assemblages and their potential driving factors across a region of inland lakes in northeastern North America spanning > 13.4° of latitude (1,500 km), 21.9° of longitude (1,700 km) and 580 m in elevation (Figure 1a). Fish occurrence records and associated environmental measurements in portions of the Southwestern Hudson Bay (n = 3,637), St. Lawrence (n = 5,757) and Mississippi River (n = 622) drainages were compiled from existing
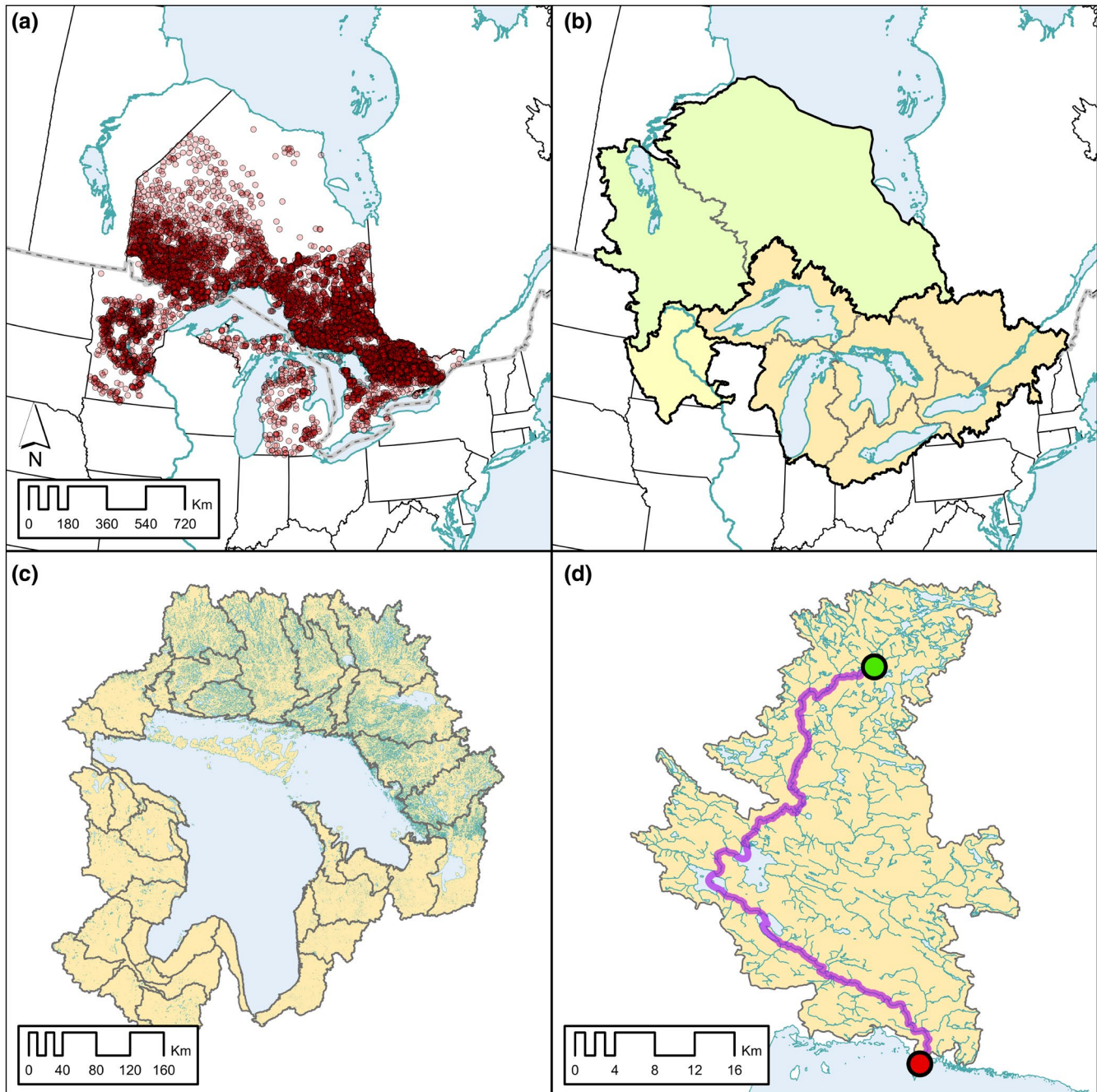
FIGURE 1 Maps presenting: (a) lake sampling locations across Ontario, Minnesota and Michigan; (b) portions of the major Southwestern Hudson Bay (green), Mississippi River (yellow) and St. Lawrence River (orange) drainage areas and their secondary drainages within our study region; (c) tertiary drainage areas (sub-sub-drainage or U.S. Geological Survey hydrological unit code 8) in the Lake Huron watershed showing regional habitat density; and (d) a single tertiary drainage area showing river network dispersal distance between a lake sampling location (green point) and its nearest secondary drainage feature (Lake Huron; red point)

datasets in Ontario (Aquatic Habitat Inventory; Dodge et al., 1987), Michigan (Status and Trends; Wehrly et al., 2021) and Minnesota (Index of Biotic Integrity; Drake & Pereira, 2002). To facilitate visualization of bioregions and certain other geospatial analyses, major drainage areas were divided further into secondary (Figure 1b) and smaller tertiary drainages (equivalent to sub-sub-drainages in Canada and hydrological unit code 8 drainages in the USA; Figure 1c).

Fishes were sampled from lakes, ponds, reservoirs and embayments (herein referred to as lakes or sites) during the open-water season over a 60-year period (1957–2017). Although we were unable to consider species turnover explicitly during this time (e.g., Cazelles et al., 2019; Keller et al., 2019; Lynch et al., 2016), we found that community differentiation was robust to minor changes in species occurrence (see results of our sensitivity analysis for non-native species in

section 3 below) and that temporal shifts in climate were generally small in comparison to spatial differences. For example, the mean annual air temperatures of sampled lakes (averaged across the study period) ranged from −4.1 to 9.7°C, whereas changes in mean temperatures between the 1950s and 2010s ranged from −.4 to 1.1°C across sites (Wang et al., 2016). Although potentially contributing to shifts in distribution or relative abundance for some populations (e.g., Finigan et al., 2018), warming is unlikely to have confounded multispecies associations and the delineation of bioregions at our scale of analysis.

Samples were collected using a combination of netting, trapping and electrofishing, and quantitative survey results were collapsed to species presence/absence to reduce biases associated with different methods (Jackson & Harvey, 1997). Records of 132 fish taxa were initially compiled across 11,112 lakes, but several species and lakes were excluded from analysis. Initially, we removed lakes lacking maximum depth or surface area measurements. We also omitted records for fishes with riverine life histories because they were likely to be transient when captured and therefore infrequently sampled in lake ecosystems. These undersampled species were identified based on expert opinion informed by comparisons to other sampling programmes in the region (e.g., Drake & Valley, 2005; Lester et al., 2003; Sandstrom et al., 2013) and aspects of organismal biology and encounter rates (relating to sampling methods/design; for a list of sampled fishes, see Supporting Information Table S1.1). Finally, because we were interested in patterns of interacting species, we removed sites with only a single species in addition to species with only a single occurrence. Our final dataset included 90 taxa across 10,016 lakes (mean species richness of eight).

2.2 | Bipartite network modularity analysis

We modelled an unweighted network graph composed of lakes and fishes as separate sets of nodes, with edges linking species to the sites at which they were observed. Although there are several approaches to grouping communities, most fail to exploit the bipartite nature of site–species networks fully to reveal both sites with similar species and species tending to form assemblages. Distance-based methods have conventionally been used to cluster sites based on their species composition and/or spatial configurations (e.g., Legendre & Legendre, 1984; Oikonomou et al., 2014); however, these procedures flatten bipartite networks, reducing species information to one of many site dissimilarity indices. Modularity analysis of site–species networks has proved to be an effective alternative to uncovering concise bioregions and their transitional zones (see participation coefficients; Guimerà & Amaral, 2005) across spatial scales (e.g., Bernardo-Madrid et al., 2019; Bloomfield et al., 2018; McGarvey & Veech, 2018).

We used the DIRTLPAb+ algorithm of Beckett (2016), which is a label propagation approach to identifying groups of sites and species that are linked together more densely than others by maximizing Barber's modularity (for further details, see Supporting Information

Appendix S1). Barber's index performs well for defining groups in bipartite networks (Thébault, 2013) and classifies both sets of nodes (sites and species) simultaneously. We applied the algorithm using the *computeModules* function in the "bipartite" package (Dormann et al., 2008), with default settings for random initializations (implemented with R v.3.6.1; R Core Team, 2019). This internally repeating procedure was itself replicated 10 times (with different seeds) to check consistency, and we recorded site and species assignments of the trial with the highest modularity (Supporting Information Table S1.2). The significance of resultant modularity was evaluated by one-sided randomization tests based on 100 null networks obtained from the efficient, sequentially swapping curveball algorithm (Strona et al., 2014). The curveball approach produces uniformly distributed null matrices, maintaining row and column sums, and was implemented with 1,000 burn-in and 1,000 thinning steps across 20 unique chains using the *nullmodel* and *oecosimu* functions in the "vegan" package (Oksanen et al., 2020). Simulations were checked for convergence before use. Participation coefficients measuring the among-module connectivity of each site were calculated using the *czvalues* function (Dormann et al., 2008). To reveal finer-scale structures, we invoked a second round of analysis restricted to each primary module. This multiscale bioregionalization was summarized spatially as the dominant primary (or sub) module across tertiary drainages (with ties adopting the most dominant type among neighbours).

2.3 | Environmental and spatial factors

We estimated environmental and spatial variables to capture key dimensions of biogeographical constructs (connectivity, climate and habitat; Supporting Information Table S1.3) hypothesized to determine natural site typology (see section 2.4 below for details of latent variable approach). A connectivity construct was formed by regional habitat density, river network dispersal distance, glacial lake/marine distance and time since glaciation. Variables were estimated from remote-sensing data using a geographical information system (ArcGIS Pro v.2.6.2; Esri, Redlands, CA, USA). Regional habitat density was calculated by summarizing all lake, pond and reservoir features within tertiary drainages and dividing the area of this combined habitat by the total area of the watershed (excluding the Great Lakes; 1:50,000 scale or better). River network dispersal distances estimated the lengths of potential contemporary dispersal routes to secondary drainage features in the Southwestern Hudson Bay drainage area (Lake Winnipeg or Hudson Bay), the St. Lawrence drainage area (Lake Superior, Lake Michigan, Lake Huron, Lake Erie, Lake Ontario or St. Lawrence River) or the Mississippi River (Figure 1d; for further details, see Supporting Information Appendix S1). Historical glacial lake/marine distance was calculated as the geodesic distance from each lake to the nearest glacial lake or marine extent polygon (Supporting Information Figure S1.1). The time since glaciation (i.e., age of the lake) was calculated by spatially joining sampling points near the centre of each lake with a time series of ice coverage

polygons and determining the most recent overlap. Estimated historical ice, lake and marine extents were obtained at *c.* 500-year resolution and 1:7,000,000 scale (Dyke, 2004).

A climate construct was formed by mean warmest month temperature, mean coldest month temperature, May to September total precipitation, and total precipitation as snow. Climate variables were elevation-adjusted point estimates of average annual conditions at each lake over the entire study period. We extracted and downscaled estimates from moderate-resolution Climate Research Unit data (4 km gridded; Mitchell & Jones, 2005) using CLIMATENA v.6.40 (Wang et al., 2016; Supporting Information Figure S1.2).

Finally, a habitat construct was formed by surface area, maximum depth, Secchi depth and proportion shield lithology. Physical size measurements were retrieved from existing records. Secchi depth, which measures water clarity, provided a proxy for basin productivity given our lack of consistent water chemistry data. Secchi depths were measured at the majority (98%) of sites. For the 191 sites without direct measures, we estimated Secchi depths from their regional relationships to maximum depth. Here, separate linear regression models were constructed for each secondary drainage area stratified into depth categories (0–20, 20–40 and > 40 m) and used to predict missing values. We estimated proportion shield lithology to capture the dominant geological gradient in our study region (i.e., surface-exposed portions of igneous and metamorphic Laurentian Shield contrasting the sedimentary rocks of the Interior and Hudson Bay Lowlands). Surficial weathering of rocks provides a natural source of variation in water quality (e.g., ion chemistry) and biological production, especially between major lithological units (Conroy & Keller, 1976). Lake polygons were overlain with detailed lithological maps compiled by Harmann and Moosdoft (2012) at an average scale of 1:3,750,000 and used to calculate the proportions of each basin composed of either shield or sedimentary rock (Supporting Information Figure S1.3).
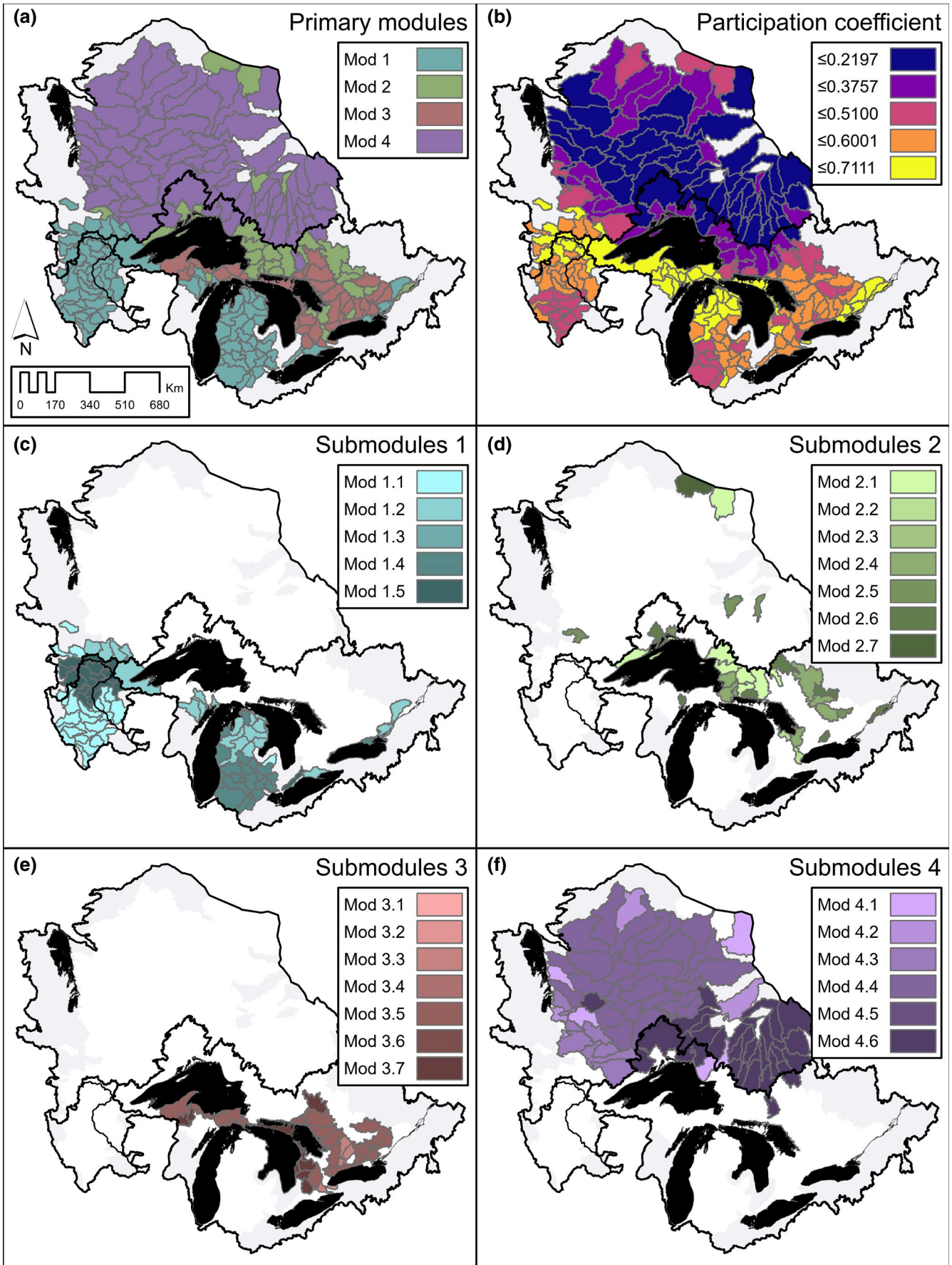
2.4 | Statistical analysis

To evaluate the roles of spatial and environmental constraints on fish communities, we applied non-metric partial least-squares structural equation modelling (NM-PLS-SEM, also called PLS path modelling) paired with multiple logistic regression (Petrarca et al., 2017). PLS-SEM is a variance-based approach to estimating unobserved latent variables and their relationships by a series of simple or multiple ordinary least-squares (OLS) regressions (Sanchez, 2013). A key application of PLS-SEM is for building latent variables as formative constructs using a composite measurement model. Here, latent variables (or constructs) are formed as linear combinations of their manifest variables (or observed indicators) by an iterative process of optimizing weights and paths in the outer measurement and inner structural equation models (Sanchez, 2013). Thus, the algorithm offers an intuitive inferential framework while also applying dimension reduction to develop proxies for multidimensional concepts, such as connectivity, climate and habitat (e.g., Loewen et al., 2020).

PLS methods can also be performed with (non-metric) nominal or ordinal indicators by introducing an optimal scaling step (Petrarca et al., 2017). Here, non-metric variables are assigned values on an interval scale by optimizing additional scaling parameters as part of the iterative PLS-SEM process.

We performed NM-PLS-SEM to assess the roles of biogeographical factors using the "path scheme" method and formative measurement (mode B) with the *plspm* function in the "plspm" package (Sanchez et al., 2017). Indicator variables were natural logaritmically or logit-transformed (for proportional data; adjusting by minimum observed values where necessary) and standardized (centred at zero and scaled to unit variance) before analysis. Connectivity, climate and habitat constructs were formed as linear combinations of their respective numerical indicators and used to predict individual modules. The approach was also applied to multilevel categorical outcomes of our modularity analysis. However, given that estimating a binomial response by OLS is less than ideal, even with the optimal scaling algorithm, we used a modified approach (Petrarca et al., 2017), initially obtaining latent variable scores from NM-PLS-SEM and then re-estimating inner model path coefficients by multiple logistic regression (generalized linear regression with binomial error distribution and a logit-link function; *glm* function in base R). The explanatory power of logistic models was measured by Tjur's R^2 using the *RsqGLM* function in the "modEvA" package (Barbosa et al., 2015). Tjur's coefficient of discrimination is an analogue to the coefficient of determination in linear regression but calculated as the difference in mean fitted values for successful and failed predictions (Tjur, 2009). Finally, because we anticipated potential collinearity among constructs, we used a variation-partitioning framework to reveal their unique, shared and total explanatory powers. Variation partitioning based on Tjur's R^2 from logistic regression was performed using the *varPart* function (Barbosa et al., 2015), and results were visualized by approximately area-proportional Euler diagrams using the *euler* function in the "eulerr" package (Larsson, 2020). To emphasize broad, biogeographical variation in fish assemblages, we limited our statistical analysis to primary modules. To test the influence of non-native species on bioregionalization patterns (Supporting Information Table S1.1), we conducted sensitivity analyses to determine changes to site and species assignments, in addition to interpretations of formative processes, either when non-native species were excluded from the species pool or when sites with non-natives were excluded from the site pool.

3 | RESULTS

We identified clear spatial patterns of fish community differentiation, which we summarized by tertiary watershed to identify tractable northern bioregions (Figure 2; Supporting Information Figures S1.4 and S1.5). Sites from module 1 (*n* = 1,258 lakes; referred to as the southern group) clustered in southern portions of the St. Lawrence and the Mississippi River drainages. Module 2 (*n* = 3,010; the trout group) and module 3 (*n* = 1,851; the central Great Lakes

(a) Primary modules

Mod 1
Mod 2
Mod 3
Mod 4

N

0 170 340 510 680 Km

(b) Participation coefficient

≤0.2197
≤0.3757
≤0.5100
≤0.6001
≤0.7111

(c) Submodules 1

Mod 1.1
Mod 1.2
Mod 1.3
Mod 1.4
Mod 1.5

(d) Submodules 2

Mod 2.1
Mod 2.2
Mod 2.3
Mod 2.4
Mod 2.5
Mod 2.6
Mod 2.7

(e) Submodules 3

Mod 3.1
Mod 3.2
Mod 3.3
Mod 3.4
Mod 3.5
Mod 3.6
Mod 3.7

(f) Submodules 4

Mod 4.1
Mod 4.2
Mod 4.3
Mod 4.4
Mod 4.5
Mod 4.6

FIGURE 2 Maps presenting network modularity results for site nodes summarized by tertiary drainages to identify bioregions. Bioregions were assigned to: (a) the dominant primary module; and (c–f) submodules across sites in each drainage. (b) Participation coefficients measure the diversity of connections between primary modules and were assigned to the mean across sites, with higher values indicating bioregional transition zones

FIGURE 3 Hierarchical edge bundling plot showing network modularity results for fish nodes to identify recurrent assemblages. Connections (edges) represent fish co-occurrences, with darker lines indicating more observations. Around the outside of the plot, node sizes indicate relative frequencies of fish occurrence, node colours indicate fish submodule assignment, and label colours indicate fish primary module assignment. Colours correspond to bioregions in Figure 2

group) were situated in more northern and east-central portions of the St. Lawrence drainage, respectively. Module 4 lakes (*n* = 3,897; the northern group) were dominant in northern areas draining into

the Southwestern Hudson Bay. However, none of the assemblages was restricted to any single drainage. Primary modules were internally consistent and had significant non-random structure, with

a modularity value (Q) of .332 (p < .01; Supporting Information Table S1.2). Participation coefficients showed that among-module connections were generally denser in the southern half of the study region, specifying a bioregional transition zone stretching from northern portions of the Mississippi River drainage to the centre of the Great Lakes Basin, in addition to adjacent to the St. Lawrence River (Figure 2b). Submodularity analysis, whereby procedures were repeated with restricted site and species pools, revealed further variation within primary modules (Figure 2c–f). Submodule assignments were weaker and more variable across trials but significantly non-random (p < .01; Supporting Information Table S1.2). Module 1 was shown to possess finer-scale north–south and east–west structure (Figure 2c), whereas module 4 revealed two central community types (in the north-west and south-east portions of the Southwestern Hudson Bay drainage), with greater heterogeneity to the south-west (Figure 2f).

In a similar manner, we found groups of associated fishes (Figure 3). Module 1 (n = 44; the southern group) contained the most species, including several warmwater fishes with limited occurrence near their northern range limits in our dataset (especially submodule 1.4). Module 2 (n = 15; the trout group) was characterized by several salmonids [i.e., lake trout (*Salvelinus namaycush*), brook trout (*Salvelinus fontinalis*), round whitefish (*Prosopium cylindraceum*) and rainbow trout (*Oncorhynchus mykiss*)] and a subset of leuciscid minnows. Module 3 (n = 8; the central Great Lakes group) contained species with known associations to the St. Lawrence River and Lake Ontario, such as American eel (*Anguilla rostrata*) and rainbow smelt (*Osmerus mordax*), along with certain warm- and coolwater game fishes, including muskellunge (*Esox masquinongy*), smallmouth bass (*Micropterus dolomieu*) and rock bass (*Ambloplites rupestris*). Module 4 (n = 23; the northern group) consisted of several species common to northern lakes, including yellow perch (*Perca flavescens*), white sucker (*Catostomus commersonii*) and the game fishes walleye (*Sander vitreus*), lake whitefish (*Coregonus clupeaformis*), cisco (*Coregonus artedi*) and northern pike (*Esox lucius*).

Climate was the dominant factor underlying community typology, while the importance of connectivity and habitat filters varied across groups. The influence of biogeographical constructs (modelled as latent variables; see meta-model in Figure 4) on the occurrence of each assemblage and the roles of individual environmental and spatial indicators can be interpreted from the model coefficients presented in Figure 5. Here, indicator weights and loadings describe the contributions and correlations of indicators with their respective latent variables in the outer measurement model (Figure 5a), while logistic path coefficients from the inner structural model show the magnitude and directionality (negative/positive) of relationships between latent variables and each of the primary modules (Figure 5b). For example, connectivity constructs were mostly driven by time since glaciation, which contributed positively to module 1 (the southern group), negatively to module 4 (the northern group) and negligibly to modules 2 and 3 (the trout and central Great Lakes groups). River network dispersal distance had a stronger negative association with the connectivity construct for module 3, but the small path coefficient for connectivity indicated that its effect was negligible. Likewise, error bars were inflated for indicators of connectivity contributing to module 2 because the construct had a minor effect, and coefficients flipped between positive and negative across bootstrapped samples. Climate indicators were generally correlated with one another (and time since glaciation; Supporting Information Table S1.4), and their individual contributions varied across modules. Precipitation as snow had a relatively strong negative correlation with module 1 and positive correlation with module 2, whereas modules 3 and 4 were more associated with higher and lower temperatures, respectively. Habitat indicators also had variable contributions and were most relevant to modules 2 and 4. Specifically, module 2 involved smaller lakes with greater Secchi depth, whereas module 4 was linked to greater surface area and proportion shield lithology. Categorical analysis indicated that climate was the principal driver of community differentiation overall (Supporting Information Tables S1.5 and S1.6).

The predominance of climate in the bioregionalization of northern fishes was also supported by our partitioning of latent variable explanatory power (Figure 6). Here, we found that climate was the most predictive construct for each assemblage except module 2 (the trout
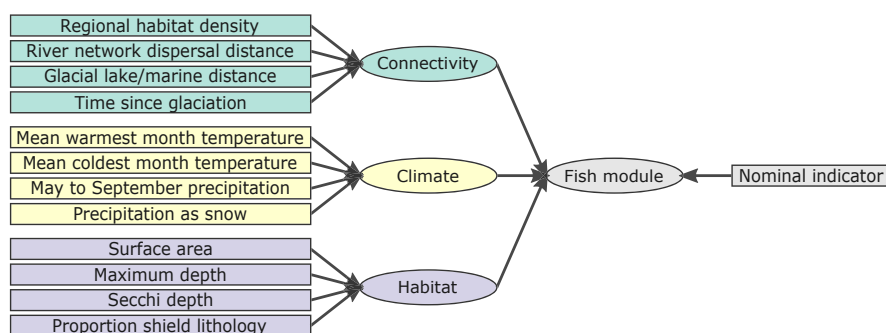
FIGURE 4 Non-metric partial least-squares structural equation (NM-PLS-SEM) meta-model. Major biogeographical constructs (connectivity, climate and habitat) were formed as linear combinations of observed spatial and environmental factors to explain variation in primary site module assignment. Biogeographical constructs and site modules were treated as composite and single-indicator latent variables (ellipses), respectively, for the inner structural model. Spatial and environmental factors (and nominal indicators of fish modules) were treated as formative indicators (rectangles) of their respective latent variables
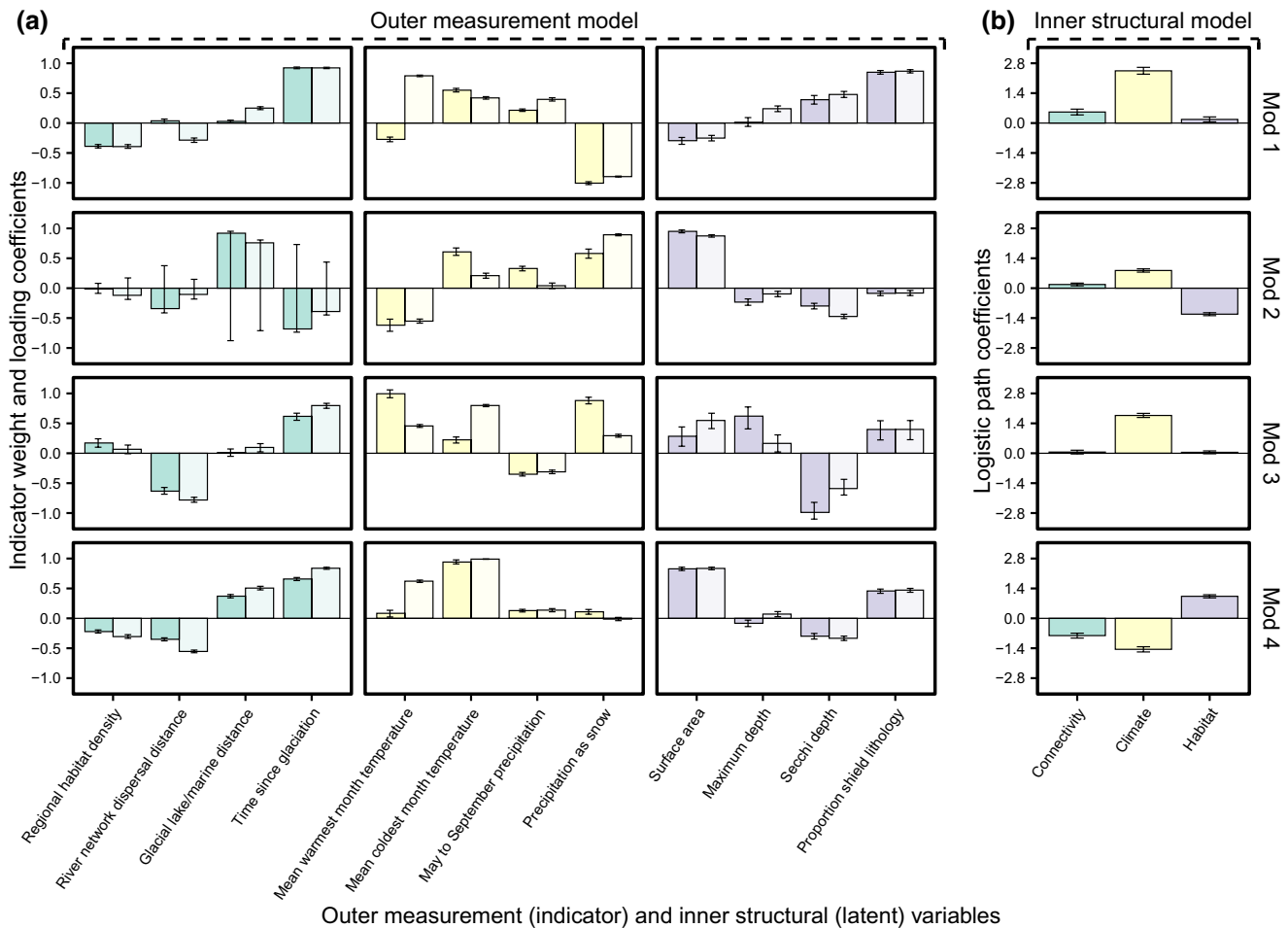
FIGURE 5 Results of non-metric partial least-squares structural equation (NM-PLS-SEM) and multiple logistic regression of primary site modules on biogeographical constructs. (a) Outer measurement model results are presented as weights (darker bars; representing multiple ordinary least-squares regression coefficients) and loadings (lighter bars; representing Pearson correlation coefficients) for standardized indicators. (b) Inner structural model results are presented as path coefficients from multiple logistic regression for latent variables. Error bars show the lower 2.5th and upper 97.5th percentiles of bootstrap and profile likelihood confidence intervals for outer and inner model coefficients, respectively. Detailed results are available in the Supporting Information (Tables S1.5–S1.8). Colours and model structure correspond to the meta-model presented in Figure 4

group), which was driven by lake area and water clarity. Modules 1 and 4 (the southern and northern groups) had the greatest classification success, with strong explanatory overlap between climate and connectivity (the latter associated with time since glaciation). Variation in connectivity and habitat were of lesser importance to module 3 (the Great Lakes group), contributing almost no additional prediction after accounting for temperature and precipitation effects (for detailed results, see Supporting Information Tables S1.7 and S1.8).

The most frequent non-natives were common carp (*Cyprinus carpio*; module 1), rainbow trout (module 2) and rainbow smelt (module 3; Supporting Information Table S1.1). We found that only two native species, spoonhead sculpin (*Cottus ricei*) and white sucker, changed modules when non-native species (*n* = 11) were excluded from the species pool (both from module 4 to 2). Assignments changed for 646 lakes (6.4%); the majority changing from module 4 to 2 (261) or 3 to 2 (134; Supporting Information Table S1.9). Changes to bioregions were few and mostly in transitional regions of southern

Ontario (Supporting Information Figure S1.6b). However, when sites with non-native species were excluded (*n* = 944; 474 from module 1, 243 from module 2, 197 from module 3, and 30 from module 4), assignments changed for nine native species (one from module 1 to 3, one from module 1 to 4, and seven from module 3 to 1) and 1,766 lakes (19.4%). Most site differences involved module 3 (94%), which largely merged with the southern group (Supporting Information Figure S1.6c), but interpretations of how communities were shaped by environmental and spatial processes were generally robust to the inclusion or exclusion of non-native species (Supporting Information Figures S1.7–S1.10).

4 | DISCUSSION

We leveraged exceptional sampling effort across our study region to present a detailed view of freshwater bioregions and their
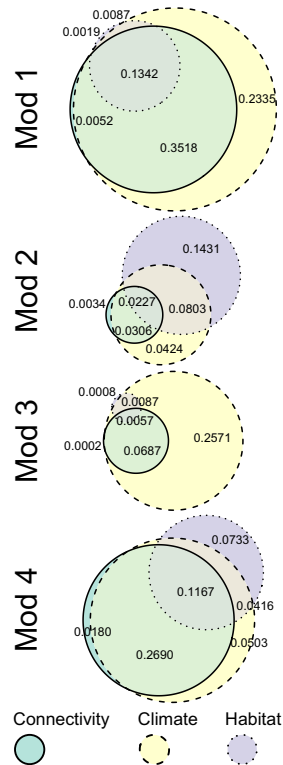
FIGURE 6 Approximately area-proportional Euler plots showing the total, shared and unique explanatory power of biogeographical constructs (connectivity, climate and habitat) in logistic regression. Fractions are expressed as Tjur's R^2 (coefficients of discrimination) from models of primary site modules (binary responses) predicted by latent variable scores (calculated by non-metric partial least-squares structural equation modelling). Colours correspond to the meta-model presented in Figure 4

underlying drivers at a broad, subcontinental scale. Given that bioregionalization is inherently taxa and scale dependent (McGarvey & Veech, 2018), we provide a novel classification of northern fish communities and advance understanding of their assembly from the regional species pool. Although global maps of fish biogeographical zones have been produced (Abell et al., 2008; Leroy et al., 2019), they have aggregated disparate surveys across larger watersheds and therefore have limited use for differentiating local communities or tractable management units. Alternative indices and module detection algorithms also exist (e.g., Guimerà & Amaral, 2005; Rosvall & Bergstrom, 2008), which may reveal different patterns (see Bloomfield et al., 2018; Thébault, 2013). Similar to the flow-based Infomap method (Edler et al., 2017), Beckett's (2016) label propagation identifies groups of highly associated sites and species simultaneously (without unipartite projection). Hence, our application of site–species network analysis builds on existing knowledge, revealing both lakes with similar fish assemblages and fishes tending to co-occur. By partitioning our network independently of spatial and environmental contexts, we were able to test the relative importance of key formative processes determining different sets of species and generate several hypotheses for future study.

Given that species sorting along geographical gradients involves multiple, often highly correlated factors, we evaluated drivers using a latent variable approach to capture broadly defined climate, connectivity and habitat filters (Dormann et al., 2013; Loewen et al., 2020). We discovered that climate had the greatest effect in differentiating fish communities, similar to recent reports for subcontinental patterns of terrestrial fauna (Ficetola et al., 2017). Thermal or hydrological regimes might also account for finer-scale differences, such as the submodular structure observed within the southern and northern groups (modules 1 and 4; Figure 2) or variation among lakes within tertiary catchments (e.g., Wehrly et al., 2012; Supporting Information Figures S1.4 and S1.5). We found that historical connectivity had a related, albeit lesser, influence on modules 1 and 4, pointing to differences in postglacial colonization opportunities between lakes in more southern and northern areas of our study region (Bailey & Smith, 1981; Dias et al., 2014; Legendre & Legendre, 1984). However, distances to glacial lakes do not necessarily capture dispersal routes, and given the degree of overlap between glacial recession and present temperature gradients (Supporting Information Figures S1.1 and S1.2), analysis of shifting recolonization pathways from glacial lakes and their outflows might offer clearer support (e.g., Mandrak & Crossman, 1992). Although habitat factors are often linked to species sorting at smaller scales (Jackson et al., 2001; Lansac-Tôha et al., 2021), we found evidence of broad-scale filtering by lake surface area, Secchi depth and shield lithology. For instance, the trout group (module 2) showed affinity to smaller lakes with greater Secchi depth, which serves as an integrative proxy for multiple terrestrial factors (e.g., sediment loading) and internal productivity (Tilzer, 1988). Smaller lakes with greater water clarity may therefore provide greater open-water foraging success for salmonids (Ortega et al., 2020). In contrast, the northern group (module 4) was linked to larger lakes with greater influence of shield lithology, in addition to those with reduced water clarity (as encountered in areas overlaying glacial lake beds). Here, species such as walleye might experience optimal foraging in habitats with lower light intensity (Lester et al., 2004). Bioregional transition zones also matched southern boundaries of the Laurentian Shield (Figure 2b; Supporting Information Figure S1.3), showing potential influence of geological setting via effects on water chemistry (Conroy & Keller, 1976; Neff & Jackson, 2012).

In addition to site typology, we revealed recurrent groups of fishes, implying broadly shared natural histories and niche space. We highlight several game fishes of interest to fisheries management but provide only a cursory examination of fish module composition. Future research should test for associations among species within and between groups (e.g., Cordero & Jackson, 2019), including their functional trait and phylogenetic similarities. For instance, the predominance of climate filters should sort species into thermal guilds (Magnuson et al., 1979), while connectivity constraints might produce assemblages related by their dispersal ability (i.e., swimming proficiency and tolerance to flow conditions) or common ancestry (Comte & Olden, 2018). Habitat filters should further aggregate species with similar life-history strategies, including

requirements for forage, reproduction and overwintering (e.g., Tonn & Magnuson, 1982). Explicit consideration of competition and predation among co-occurring species might yield additional mechanistic insights; however, as with habitat filters, we anticipate that species interactions will be most important at smaller scales of inquiry (Jackson et al., 2001).

Our findings have applications in conservation biogeography, revealing opportunities for targeted climate adaptation strategies. Setting conservation priorities for regional planning (e.g., protected areas or resource extraction) requires defining management units and understanding their formative processes (Brooks et al., 2006; Montalvo-Mancheno et al., 2020; Oikonomou et al., 2014). Applying space-for-time substitution, the central roles of temperature and precipitation provide evidence that fish communities are sensitive to climatic changes (Alofs et al., 2014; Lynch et al., 2016), but future impacts will depend on community type. We also found that connectivity does not appear to have been a major factor preventing lacustrine fishes from matching their temperature preferences within catchments. For instance, the southern group (module 1) was much more associated with warmer temperatures and less precipitation as snow than contemporary dispersal indicators, suggesting potential to expand northward as temperatures rise and connectivity permits. Indeed, there is evidence that certain warm- and coolwater game fishes are already expanding their northern ranges via multiple mechanisms (Alofs et al., 2014), including anthropogenic introductions. Coldwater fishes in the trout group (module 2) were linked to cooler summers and more snow, but habitat size and quality had a greater influence. Thus, this group might be relatively sensitive to hydrological changes and increasing water column turbidity or brownification. In contrast, the Great Lakes group (module 3) was clustered in warmer and snowier areas around the Great Lakes in south-central Ontario. These results suggest that Great Lakes effects create conditions for warm- and coolwater fishes that might offer partial buffering against climatic changes, although future climates are uncertain and will be regionally variable (Notaro et al., 2015). Finally, the northern group (module 4) was most sensitive to warming (especially in winter months) and potentially more dispersal limited (based on the greater importance of river network dispersal distance), indicating that these communities might require extra attention for climate change adaptation. However, our estimates of dispersal limitation reflect only the natural connectivity of sites within their respective drainages.

Establishment of introduced fishes outside of historical ranges indicates that natural colonization either lags changing conditions or is impeded by other environmental, connectivity or biotic factors (e.g., Allee effects). Fish movements are constrained by physical barriers (e.g., dams, culverts and waterfalls), and limited hydrological connections between catchments (especially along continental divides) may induce critical bottlenecks for migration (Jackson et al., 2001; Melles et al., 2015). Species also have thermal and depth requirements for dispersal, possibly limiting movement by coldwater species as climates warm. Although assemblages generally reflect shared colonization histories of dispersal along lake and river networks, effects of increasing fragmentation on future range shifts and persistence may vary (e.g., Herrera-R et al., 2020). Actions such as human-assisted migration or construction of fishways warrant consideration, weighing potential benefits with risks of unintended consequences from introducing harmful species (Olden et al., 2010), including pathogens.

Human introductions have already impacted the spatial structure of fish communities globally (Alofs et al., 2014; Olden et al., 2010). For instance, regional translocations contribute to biotic homogenization among lakes with higher angling pressure or nearer to urban centres (Olden et al., 2008). Stocking programmes and movement of bait fishes might also have contributed to a minor role of connectivity for the trout group (module 2), as introductions obscure how species are naturally filtered from, or serve as additions to, the regional pool (Alofs et al., 2014; Bernardo-Madrid et al., 2019; Cazelles et al., 2019). Although we were unable to account for regional translocations, the small effect of excluding records for non-native species (from outside the study region) on our understanding of community differentiation highlights the robustness of our results to potential sampling errors. Although biases associated with site selection (e.g., targeting lakes with fishes of greater management interest) or imperfect detection can confound species distribution patterns, errors were mitigated by our large sample size and the relative insensitivity of our analytical approach to missing data. However, removal of a species record does not remove its community impact, and our sensitivity analysis showed a major role of invaded communities in the differentiation of the southern and Great Lakes groups (modules 1 and 3). Whether driven by species interactions, loss of coverage or some other aspect of invaded sites, results excluding lakes with non-native species both revealed the imprint of introductions on fish communities in the region and upheld our interpretations of key formative processes and climatic sensitivities. Despite non-native species typically establishing by anthropogenic means (such as shipping, recreational stocking or escape from aquaculture/aquaria), they contribute to contemporary bioregions and are often management priorities as fisheries or biotic stressors on native communities (e.g., Loewen et al., 2020).

Finally, as we focus on natural drivers of community differentiation, important questions remain about how anthropogenic factors overlie observed patterns. For instance, regional land use can cause changes in downstream water quality and fish community structure (e.g., Mantyka-Pringle et al., 2014). In our study region, most urban and agricultural developments are concentrated at southern latitudes where fishes generally contend with heightened angling pressure (Lester et al., 2003). Many thousands of lakes in central Ontario have also experienced anthropogenic acidification, leaving lasting impacts on aquatic food webs (Keller et al., 2019) and contributing to a suite of waterscape stressors affecting freshwater ecosystems at multiple scales (Heino et al., 2021). Further integration of these and other human activities will be needed to reconcile natural patterns of freshwater biogeography with ongoing global changes and spread of invasive species.

AUTHOR CONTRIBUTIONS

All authors contributed to data compilation and manuscript revision. C.J.G.L. led study conception and design, conducted the analyses and composed the initial draft of the manuscript.

DATA AVAILABILITY STATEMENT

Processed data (including site and species modules and derived spatial and environmental indicator variables) and code supporting our results are archived in a dedicated GitHub repository (https://github.com/loewenecology/Fish-modularity-and-nmplssem). Raw fish data (with species at risk removed for their protection) are available from the Dryad Digital Repository (https://doi.org/10.5061/dryad.44j0z pcfc). Given that contractual requirements prevent us from releasing detailed sampling records for Ontario directly, these data are provided with approximate location coordinates. The full set of fish data (including species at risk) are readily available for non-commercial purposes from the Ministry of Northern Development, Mines, Natural Resources and Forestry in Ontario (https://geohub.lio.gov.on.ca/datasets/aquatic-resource-area-polygon-segment-/) and Departments of Natural Resources in Michigan (https://www.michigan.gov/dnr/0,4570,7-350-79137_79765_84005_84013---,00.html) and Minnesota (https://www.dnr.state.mn.us/waters/surfacewater_section/lake_ibi/index.html) through their standard data-sharing agreements.

ORCID

Charlie J. G. Loewen (iD) https://orcid.org/0000-0002-4389-4134
Donald A. Jackson (iD) https://orcid.org/0000-0002-6107-0753
Cindy Chu (iD) https://orcid.org/0000-0002-1914-3218
Karen M. Alofs (iD) https://orcid.org/0000-0002-4161-3554
Gretchen J. A. Hansen (iD) https://orcid.org/0000-0003-0241-7048
Andrew E. Honsey (iD) https://orcid.org/0000-0001-7535-1321
Charles K. Minns (iD) https://orcid.org/0000-0003-2249-1624

REFERENCES

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., Coad, B., Mandrak, N., Balderas, S. C., Bussing, W., Stiassny, M. L. J., Skelton, P., Allen, G. R., Unmack, P., Naseka, A., Ng, R., Sindorf, N., Robertson, J., Armijo, E., … Petry, P. (2008). Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *BioScience*, *58*(5), 403–414. https://doi.org/10.1641/B580507

Alofs, K. M., Jackson, D. A., & Lester, N. P. (2014). Ontario freshwater fishes demonstrate differing range-boundary shifts in a warming climate. *Diversity and Distributions*, *20*(2), 123–136. https://doi.org/10.1111/ddi.12130

Bailey, R. M., & Smith, G. R. (1981). Origin and geography of the fish fauna of the Laurentian Great Lakes Basin. *Canadian Journal of Fisheries and Aquatic Sciences*, *38*(12), 1539–1561. https://doi.org/10.1139/f81-206

Barbosa, A. M., Real, R., Muñoz, A.-R., & Brown, J. A. (2015). New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Diversity and Distributions*, *19*(10), 1333–1338. https://doi.org/10.1111/ddi.12100

Beckett, S. J. (2016). Improved community detection in weighted bipartite networks. *Royal Society Open Science*, *3*(1), 140536. https://doi.org/10.1098/rsos.140536

Bernardo-Madrid, R., Calatayud, J., González-Suárez, M., Rosvall, M., Lucas, P. M., Rueda, M., Antonelli, A., & Revilla, E. (2019). Human activity is altering the world's zoogeographical regions. *Ecology Letters*(8), *22*, 1297–1305. https://doi.org/10.1111/ele.13321

Bloomfield, N. J., Knerr, N., & Encinas-Viso, F. (2018). A comparison of network and clustering methods to detect biogeographical regions. *Ecography*, *41*(1), 1–10. https://doi.org/10.1111/ecog.02596

Brooks, T. M., Mittermeier, R. A., da Fonseca, G. A. B., Gerlach, J., Hoffmann, M., Lamoreux, J. F., Mittermeier, C. G., Pilgrim, J. D., & Rodrigues, A. S. L. (2006). Global biodiversity conservation priorities. *Science*, *313*(5783), 58–61. https://doi.org/10.1126/science.1127609

Carstensen, D. W., & Olesen, J. M. (2009). Wallacea and its nectarivorous birds: Nestedness and modules. *Journal of Biogeography*, *36*(8), 1540–1550. https://doi.org/10.1111/j.1365-2699.2009.02098.x

Cazelles, K., Bartley, T., Guzzo, M. M., Brice, M.-H., MacDougall, A. S., Bennett, J. R., Esch, E. H., Kadoya, T., Kelly, J., Matsuzaki, S.-I., Nilsson, K. A., & McCann, K. S. (2019). Homogenization of freshwater lakes: Recent compositional shifts in fish communities are explained by gamefish movement and not climate change. *Global Change Biology*, *25*(12), 4222–4233. https://doi.org/10.1111/gcb.14829

Comte, L., & Olden, J. D. (2018). Evidence for dispersal syndromes in freshwater fishes. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1871), 20172214. https://doi.org/10.1098/rspb.2017.2214

Conroy, N., & Keller, W. (1976). Geological factors affecting biological activity in Precambrian shield lakes. *The Canadian Mineralogist*, *14*, 62–72.

Cordero, R. D., & Jackson, D. A. (2019). Species-pair associations, null models, and test of mechanisms structuring ecological communities. *Ecosphere*, *10*(7), e02797. https://doi.org/10.1002/ecs2.2797

D'Arcy, P., & Carignan, R. (1997). Influence of catchment topography on water chemistry in southeastern Québec Shield lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, *54*(10), 2215–2227. https://doi.org/10.1139/f97-129

Dias, M. S., Oberdorff, T., Hugueny, B., Leprieur, F., Jézéquel, C., Cornu, J.-F., Brosse, S., Grenouillet, G., & Tedesco, P. A. (2014). Global imprint of historical connectivity on freshwater fish biodiversity. *Ecology Letters*, *17*(9), 1130–1140. https://doi.org/10.1111/ele.12319

Dodge, D. P., Goodchild, G. A., Tilt, J. C., Waldriff, D. G., & MacRitchie, I. (1987). *Manual of instructions: Aquatic habitat inventory surveys*. Ontario Ministry of Natural Resources.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013).

Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Dormann, C. F., Gruber, B., & Fründ, J. (2008). Introducing the bipartite package: Analysing ecological networks. *R News*, *8*, 8–11.

Drake, M. T., & Pereira, D. L. (2002). Development of a fish-based index of biotic integrity for small inland lakes in Central Minnesota. *North American Journal of Fisheries Management*, *22*(4), 1105–1123.

Drake, M. T., & Valley, R. D. (2005). Validation and application of a fish-based index of biotic integrity for small central Minnesota lakes. *North American Journal of Fisheries Management*, *25*(3), 1095–1111. https://doi.org/10.1577/M04-128.1

Dyke, A. S. (2004). *An outline of North American deglaciation with emphasis on central and northern Canada*. Geological Survey of Canada.

Edler, D., Guedes, T., Zizka, A., Rosvall, M., & Antonelli, A. (2017). Infomap bioregions: Interactive mapping of biogeographical regions from species distributions. *Systematic Biology*, *66*(2), 197–204.

Ficetola, G. F., Mazel, F., & Thuiller, W. (2017). Global determinants of zoogeographical boundaries. *Nature Ecology and Evolution*, *1*(4), 0089. https://doi.org/10.1038/s41559-017-0089

Finigan, P. A., Mandrak, N. E., & Tufts, B. L. (2018). Large-scale changes in the littoral fish communities of lakes in southeastern Ontario, Canada. *Canadian Journal of Zoology*, *96*(7), 753–759. https://doi.org/10.1139/cjz-2017-0080

Guimerà, R., & Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature*, *433*(7028), 895–900. https://doi.org/10.1038/nature03288

Harmann, J., & Moosdoft, N. (2012). The new global lithological map database GLiM: A representation of rock properties at the Earth surface. *Geochemistry, Geophysics, Geosystems*, *13*, Q12004.

Heino, J., Alahuhta, J., Bini, L. M., Cai, Y., Heiskanen, A.-S., Hellsten, S., Kortelainen, P., Kotamäki, N., Tolonen, K. T., Vihervaara, P., Vilmi, A., & Angeler, D. G. (2021). Lakes in the era of global change: Moving beyond single-lake thinking in maintaining biodiversity and ecosystem services. *Biological Reviews*, *96*(1), 89–106. https://doi.org/10.1111/brv.12647

Herrera-R, G. A., Oberdorff, T., Anderson, E. P., Brosse, S., Carvajal-Vallejos, F. M., Frederico, R. G., Hidalgo, M., Jézéquel, C., Maldonado, M., Maldonado-Ocampo, J. A., Ortega, H., Radinger, J., Torrente-Vilara, G., Zuanon, J., & Tedesco, P. A. (2020). The combined effects of climate change and river fragmentation on the distribution of Andean Amazon fishes. *Global Change Biology*, *26*(10), 5509–5523. https://doi.org/10.1111/gcb.15285

Hitt, N. P., & Angermeier, P. L. (2008). Evidence for fish dispersal from spatial analysis of stream network topology. *Journal of the North American Benthological Society*, *27*(2), 304–320. https://doi.org/10.1899/07-096.1

Jackson, D. A., & Harvey, H. H. (1997). Qualitative and quantitative sampling of lake fish communities. *Canadian Journal of Fisheries and Aquatic Sciences*, *54*(12), 2807–2813. https://doi.org/10.1139/f97-182

Jackson, D. A., Peres-Neto, P. R., & Olden, J. D. (2001). What controls who is where in freshwater fish communities – the roles of biotic, abiotic, and spatial factors. *Canadian Journal of Fisheries and Aquatic Sciences*, *58*, 157–170. https://doi.org/10.1139/f00-239

Keller, W. B., Heneberry, J., & Edwards, B. A. (2019). Recovery of acidified Sudbury, Ontario, Canada, lakes: A multi-decade synthesis and update. *Environmental Reviews*, *27*(1), 1–16. https://doi.org/10.1139/er-2018-0018

Lansac-Tôha, F. M., Bini, L. M., Heino, J., Meira, B. R., Segovia, B. T., Pavanelli, C. S., Bonecker, C. C., Deus, C. P., Benedito, E., Alves, G. M., Manetta, G. I., Dias, J. D., Vieira, L. C. G., Rodrigues, L. C., Carmo Roberto, M., Brugler, M. R., Lemke, M. J., Tessler, M., DeSalle, R., … Velho, L. F. M. (2021). Scale-dependent patterns of metacommunity structuring in aquatic organisms across floodplain systems. *Journal of Biogeography*(4), *48*, 872–885. https://doi.org/10.1111/jbi.14044

Larsson, J. (2020). *eulerr: Area-proportional Euler and Venn diagrams with ellipses*. R package version 6.1.0. https://CRAN.R-project.org/package=eulerr

Legendre, P., & Legendre, V. (1984). Postglacial dispersal of freshwater fishes in the Québec Peninsula. *Canadian Journal of Fisheries and Aquatic Sciences*, *41*(12), 1781–1802. https://doi.org/10.1139/f84-220

Leroy, B., Dias, M. S., Giraud, E., Hugueny, B., Jézéquel, C., Leprieur, F., Oberdorff, T., & Tedesco, P. A. (2019). Global biogeographical regions of freshwater fish species. *Journal of Biogeography*, *46*(11), 2407–2419. https://doi.org/10.1111/jbi.13674

Lester, N. P., Dextrase, A. J., Kushneriuk, R. S., Rawson, M. R., & Ryan, P. A. (2004). Light and temperature: Key factors affecting walleye abundance and production. *Transactions of the American Fisheries Society*, *133*(3), 588–605. https://doi.org/10.1577/T02-111.1

Lester, N. P., Marshall, T. R., Armstrong, K., Dunlop, W. I., & Ritchie, B. (2003). A broad-scale approach to management of Ontario's recreational fisheries. *North American Journal of Fisheries Management*, *23*(4), 1312–1328. https://doi.org/10.1577/M01-230AM

Loewen, C. J. G., Strecker, A. L., Gilbert, B., & Jackson, D. A. (2020). Climate warming moderates the impacts of introduced sportfish on multiple dimensions of prey biodiversity. *Global Change Biology*, *26*(9), 4937–4951. https://doi.org/10.1111/gcb.15225

Lynch, A. J., Myers, B. J. E., Chu, C., Eby, L. A., Falke, J. A., Kovach, R. P., Krabbenhoft, T. J., Kwak, T. J., Lyons, J., Paukert, C. P., & Whitney, J. E. (2016). Climate change effects on North American inland fish populations and assemblages. *Fisheries*, *41*(7), 346–361. https://doi.org/10.1080/03632415.2016.1186016

MacArthur, R. H., & Wilson, E. O. (1963). An equilibrium theory of insular zoogeography. *Evolution*, *17*(4), 373–387. https://doi.org/10.1111/j.1558-5646.1963.tb03295.x

Magnuson, J. J., Crowder, L. B., & Medvick, P. A. (1979). Temperature as an ecological resource. *American Zoology*, *19*(1), 331–343. https://doi.org/10.1093/icb/19.1.331

Mandrak, N. E. (1995). Biogeographic patterns of fish species richness in Ontario lakes in relation to historical and environmental factors. *Canadian Journal of Fisheries and Aquatic Sciences*, *52*(7), 1462–1474. https://doi.org/10.1139/f95-141

Mandrak, N. E., & Crossman, E. J. (1992). Postglacial dispersal of freshwater fishes into Ontario. *Canadian Journal of Zoology*, *70*(11), 2247–2259. https://doi.org/10.1139/z92-302

Mantyka-Pringle, C. S., Martin, T. G., Moffatt, D. B., Linke, S., & Rhodes, J. R. (2014). Understanding and predicting the combined effects of climate change and land-use change on freshwater macroinvertebrates and fish. *Journal of Applied Ecology*(3), *51*, 572–581. https://doi.org/10.1111/1365-2664.12236

McGarvey, D. J., & Veech, J. A. (2018). Modular structure in fish co-occurrence networks: A comparison across spatial scales and grouping methodologies. *PLoS One*, *13*(12), e0208720. https://doi.org/10.1371/journal.pone.0208720

Melles, S. J., Chu, C., Alofs, K. M., & Jackson, D. A. (2015). Potential spread of Great Lakes fishes given climate change and proposed dams: An approach using circuit theory to evaluate invasion risk. *Landscape Ecology*, *30*(5), 919–935. https://doi.org/10.1007/s10980-014-0114-z

Mitchell, T. D., & Jones, P. D. (2005). An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, *25*(6), 693–712. https://doi.org/10.1002/joc.1181

Montalvo-Mancheno, C. S., Ondei, S., Brook, B. W., & Buettel, J. C. (2020). Bioregionalization approaches for conservation: Methods, biases, and their implications for Australian biodiversity. *Biodiversity*

and Conservation, 29(1), 1–17. https://doi.org/10.1007/s10531-019-01913-6

Neff, M. R., & Jackson, D. A. (2012). Geology as a structuring mechanism of stream fish communities. *Transactions of the American Fisheries Society*, 141(4), 962–974. https://doi.org/10.1080/00028487.2012.676591

Notaro, M., Bennington, V., & Vavrus, S. (2015). Dynamically downscaled projections of lake-effect snow in the Great Lakes Basin. *Journal of Climate*, 28(4), 1661–1684. https://doi.org/10.1175/JCLI-D-14-00467.1

Oikonomou, A., Leprieur, F., & Leonardos, I. D. (2014). Biogeography of freshwater fishes of the Balkan Peninsula. *Hydrobiologia*, 738(1), 205–220. https://doi.org/10.1007/s10750-014-1930-5

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2020). *vegan: Community ecology package*. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Olden, J. D., Kennard, M. J., Leprieur, F., Tedesco, P. A., Winemiller, K. O., & García-Berthou, E. (2010). Conservation biogeography of freshwater fishes: Recent progress and future challenges. *Diversity and Distributions*, 16(3), 496–513. https://doi.org/10.1111/j.1472-4642.2010.00655.x

Olden, J. D., Kennard, M. J., & Pusey, B. J. (2008). Species invasions and the changing biogeography of Australian freshwater fishes. *Global Ecology and Biogeography*, 17(1), 25–37. https://doi.org/10.1111/j.1466-8238.2007.00340.x

Ortega, J. C. G., Figueiredo, B. R. S., da Graça, W. J., Agostinho, A. A., & Bini, L. M. (2020). Negative effect of turbidity on prey capture for both visual and non-visual aquatic predators. *Journal of Animal Ecology*, 89(11), 2427–2439. https://doi.org/10.1111/1365-2656.13329

Petrarca, F., Russolillo, G., & Trinchera, L. (2017). Integrating non-metric data in partial least squares path models: Methods and application. In H. Latan, & R. Noonan (Eds.), *Partial least squares path modeling* (pp. 259–279). Springer International Publishing.

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., Smol, J. P., Taylor, W. W., Tockner, K., Vermaire, J. C., Dudgeon, D., & Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*(3), 94, 849–873. https://doi.org/10.1111/brv.12480

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, 105(4), 1118–1123. https://doi.org/10.1073/pnas.0706851105

Sanchez, G. (2013). *PLS path modeling with R*. Trowchez Editions.

Sanchez, G., Trinchera, L., & Russolillo, G. (2017). *plspm: Tools for partial least squares path modeling (PLS-PM)*. R package version 0.4.9. https://CRAN.R-project.org/package=plspm

Sandstrom, S., Rawson, M., & Lester, N. P. (2013). *Manual of instructions for broad-scale fish community monitoring using North American (NA1) and Ontario small mesh (ON2) gillnets*. Ontario Ministry of Natural Resources.

Shuter, B. J., & Post, J. R. (1990). Climate, population viability, and the zoogeography of temperate fishes. *Transactions of the American Fisheries Society*, 119(2), 314–336.

Smith, C. L., & Powell, C. R. (1971). *The summer fish communities of Brier Creek, Marshall County, Oklahoma*. No. 2458. American Museum of Natural History. http://hdl.handle.net/2246/2666

Strona, G., Nappo, D., Boccacci, F., Fattorini, S., & San-Miguel-Ayanz, J. (2014). A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications*, 5(1), 4114. https://doi.org/10.1038/ncomms5114

Thébault, E. (2013). Identifying compartments in presence–absence matrices and bipartite networks: Insights into modularity measures. *Journal of Biogeography*, 40(4), 759–768. https://doi.org/10.1111/jbi.12015

Tilzer, M. M. (1988). Secchi disk — chlorophyll relationships in a lake with highly variable phytoplankton biomass. *Hydrobiologia*, 162(2), 163–171. https://doi.org/10.1007/BF00014539

Tjur, T. (2009). Coefficients of determination in logistic regression models—a new proposal: The coefficient of discrimination. *The American Statistician*, 63(4), 366–372. https://doi.org/10.1198/tast.2009.08210

Tonn, W. M., & Magnuson, J. J. (1982). Patterns in the species composition and richness of fish assemblages in northern Wisconsin lakes. *Ecology*, 63(4), 1149–1166. https://doi.org/10.2307/1937251

Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS One*, 11(6), e0156720. https://doi.org/10.1371/journal.pone.0156720

Wehrly, K. E., Breck, J. E., Wang, L., & Szabo-Kraft, L. (2012). A landscape-based classification of fish assemblages in sampled and unsampled lakes. *Transactions of the American Fisheries Society*, 141(2), 414–425. https://doi.org/10.1080/00028487.2012.667046

Wehrly, K. E., Carter, G. S., & Breck, J. E. (2021). *Standardized sampling methods for the inland lakes status and trends program*. Fisheries Special Report. Michigan Department of Natural Resources.

BIOSKETCH

Charlie J. G. Loewen is a community and macroecologist. His research takes a quantitative approach to understanding key processes and dynamics of freshwater ecosystems, with a focus on their response to compounding environmental changes.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Loewen, C. J. G., Jackson D. A., Chu C., Alofs K. M., Hansen G. J. A., Honsey A. E., Minns C. K., & Wehrly K. E. (2022). Bioregions are predominantly climatic for fishes of northern lakes. *Global Ecology and Biogeography*, 31, 233–246. https://doi.org/10.1111/geb.13424