ORIGINAL RESEARCH

Ecology and Evolution
Open Access
WILEY

Using co-occurrence network topology in assessing ecological stress in benthic macroinvertebrate communities

Ariel Levi Simons[1] 🆔 | **Raphael Mazor[2]** 🆔 | **Susanna Theroux[2]** 🆔

[1]Dornsife College of Letters, Arts and Sciences, University of Southern California, Los Angeles, California

[2]Southern California Coastal Water Research Project, Costa Mesa, California

Correspondence
Ariel Levi Simons, Dornsife College of Letters, Arts and Sciences, University of Southern California, Los Angeles, CA 90089-2910.
Email: alsimons@usc.edu

Abstract

Ecological monitoring of streams has often focused on assessing the biotic integrity of individual benthic macroinvertebrate (BMI) communities through local measures of diversity, such as taxonomic or functional richness. However, as individual BMI communities are frequently linked by a variety of ecological processes at a regional scale, there is a need to assess biotic integrity of groups of communities at the scale of watersheds. Using 4,619 sampled communities of streambed BMIs, we investigate this question using co-occurrence networks generated from groups of communities selected within California watersheds under different levels of stress due to upstream land use. Building on a number of arguments in theoretical ecology and network theory, we propose a framework for the assessment of the biotic integrity of watershed-scale groupings of BMI communities using measures of their co-occurrence network topology. We found significant correlations between stress, as described by a mean measure of upstream land use within a watershed, and topological measures of co-occurrence networks such as network size ($r = -.81$, $p < 10^{-4}$), connectance ($r = .31$, $p < 10^{-4}$), mean co-occurrence strength ($r = .25$, $p < 10^{-4}$), degree heterogeneity ($r = -.10$, $p < 10^{-4}$), and modularity ($r = .11$, $p < 10^{-4}$). Using these five topological measures, we constructed a linear model of biotic integrity, here a composite of taxonomic and functional diversity known as the California Stream Condition Index, of groups of BMI communities within a watershed. This model can account for 66% of among-watershed variation in the mean biotic integrity of communities. These observations imply a role for co-occurrence networks in assessing the current status of biotic integrity for BMI communities, as well as their potential use in assessing other ecological communities.

KEYWORDS
co-occurrence network, ecological index, ecological stress, landscape ecology, stream ecosystems, topology

1 | INTRODUCTION

Humanity can be considered a global scale force for ecosystem engineering (Guerry et al., 2015; Laurance, Sayer, & Cassman, 2014; Vörösmarty et al., 2010). Subsequent to the rise of anthropogenic stressors on the environment, there has been the recognition of the need for ecological monitoring, which can match the scale of human activity (Bergseth, Russ, & Cinner, 2015; Corona, Chirici, McRoberts, Winter, & Barbati, 2011; Foley et al., 2013; Schmeller et al., 2015; Steenweg et al., 2017).

Among the ecosystems being monitored, streams have been of long-term and ongoing interest. Human activities are both dependent upon their ecological services (Anderson, Glibert, & Burkholder, 2002; Dudgeon et al., 2006) and dependent upon frequently a source of their environmental stress (Carpenter et al., 1998; Paerl et al., 2016). Human-dominated environments, such as farms, tend to cover large areas. For this reason, there have been efforts to monitor the state of streams across entire watersheds rather than individual streams (Grönroos et al., 2013; Socolar, Gilroy, Kunin, & Edwards, 2016), especially in light of the importance of regional versus local measures of habitat quality with stream communities (Stoll, Breyer, Tonkin, Früh, & Haase, 2016). These biomonitoring efforts have typically focused either on the presence of certain indicator taxa (Fausch, Lyons, Karr, & Angermeier, 1990; Vieira, Séneca, Sérgio, & Ferreira, 2012) or on comparing the composition of communities to an "undisturbed" reference (Kerans & Karr, 1994; Lakew & Moog, 2015; Masese, Raburu, & Muchiri, 2009; Mazor et al., 2015; Silva, Herlihy, Hughes, & Callisto, 2017; Vile & Henning, 2018).

Historically, bioassessments of stream have tended to be based on data sets composed on particular communities, such as BMIs (Cuffney, Brightbill, May, & Waite, 2010; Maxted et al., 2000), organized by morphological classifications. With the advent of high-throughput metagenomic sequencing, there now exists the potential for rapidly constructing a picture of community composition with greater breadth (Elbrecht, Vamos, Meissner, Aroviita, & Leese, 2017; Stein et al., 2014), taxonomic resolution, and reliability (Baird & Hajibabaei, 2012; Sweeney, Battle, Jackson, & Dapkey, 2011). There is an opportunity then to create a bioassessment framework for BMI communities in which one could readily incorporate community composition data constructed from metagenomic methods (Goodwin et al., 2017; Hering et al., 2018).

Here, we propose the use of co-occurrence networks to the task of ecological monitoring. These networks represent the likelihoods, represented by edges, of various unique categories of taxa, represented by nodes, co-occurring in a landscape defined by the spatial extent of the communities sampled and studied. These networks can be constructed from basic ecological data, such as the presence or absence of a set of taxonomic groups across a set of sites (Arita, 2016; Gotelli, 2000; Morueta-Holme et al., 2016).

Co-occurrence networks have been investigated as a means of inferring ecological patterns, particularly when direct measurement of ecological interactions proves infeasible. For example, the clustering of microbial species into distinct modules within co-occurrence networks has been used to infer physiochemical niches for various prokaryotic groups (Fuhrman & Steele, 2008; Larsen & Ormerod, 2014; Mandakovic et al., 2018; Ruan et al., 2006; Steele et al., 2011; Widder et al., 2014). In studies of larger organisms, topological measures of these networks have also been used to illustrate a loss in both diversity and the number of significant co-occurrences between reptiles in response to habitat degradation (Kay et al., 2018). The previous diversity of scenarios where co-occurrence network topology has been used in ecological analysis then implies that it could also be used to develop a framework for the assessment of the biotic integrity of streams across an entire catchment area (Ahn & Kim, 2017; Moyle & Randall, 1998; Smith & Lamp, 2008).

We hypothesized a number of relationships between ecological stress and five measures of co-occurrence network topology (Table 1). We chose these measures based on prior analyses of co-occurrence networks and their relationships with environmental stress, for example, spatial aggregation of species across an environmentally heterogeneous landscape due to variations in their ecological attributes (Bellisario, Cerfolli, & Nascetti, 2010; Borthagaray, Arim, & Marquet, 2014). These hypotheses were then tested using co-occurrence networks generated from the presence/absence data for benthic macroinvertebrates (BMIs) gathered in streams across the state of California. We used upstream land use as our measure of stress as it was consistently measured at every sample sited, and has been found to be a broad measure of anthropogenic stress in stream communities (Novotny, Bartošová, O'Reilly, & Ehlinger, 2005; Vander Laan, Hawkins, Olson, & Hill, 2013).

1.1 | Ecological stress and co-occurrence network topology

Prior observations of BMI communities under stress have shown two trends: first, a decline in taxonomic richness (Lenat & Crawford,

Topological measure	Ecological relevance	Hypothesized relationship with stress
Network size	The number of unique types of taxa across a set of communities.	(−)
Connectance	The fraction of significant co-occurrences realized compared to theoretical maximum for a network.	(+)
Mean co-occurrence strength	Correlation strength between unique types of taxa.	(+)
Modularity	How strongly patterns of co-occurrence are partitioned into subcommunities.	(+)
Degree heterogeneity	How skewed the distribution of the number of co-occurrences per unique type of taxa is in a community.	(−)

TABLE 1 Topological measures of co-occurrence networks, their ecological relevance, and predicted relationship with an increase in stress due to upstream land use: network size, connectance, mean co-occurrence strength, modularity, and degree heterogeneity

1994; Stepenuck, Crunkilton, & Wang, 2002; Voß & Schäfer, 2017); and second, a predominance of members of generalist groups with broad ecological niches (Büchi & Vuilleumier, 2014; Ducatez, Tingley, & Shine, 2014; Mykrä & Heino, 2017). Starting with these trends, we then hypothesized relationships between five topological measures and ecological stress.

1.1.1 | Network size

The sizes of our networks were determined by the number of unique BMI genera present within a given set of sampling sites. Ecological stress in BMI communities has been found to be associated with a decline in local taxonomic richness (Ourso & Frenzel, 2003; Paul & Meyer, 2001; Stepenuck et al., 2002; Wallace & Biastoch, 2016). This is especially the case where the stress is due to an increase in upstream land use (Allan, 2004; Sponseller, Benfield, & Valett, 2001; Stepenuck et al., 2002). Given then the correspondence between the taxonomic richness present in a group of sites, and the number of nodes in any resulting co-occurrence network, we expect network size to be negatively correlated with stress (H1).

1.1.2 | Connectance

As the number of edges in a variety of ecological networks may be sensitive to the number of unique taxonomic groups (Dormann, Frund, Bluthgen, & Gruber, 2009; Goldwasser & Roughgarden, 1997; Nielsen & Bascompte, 2007), we then also calculated the connectance (Bell, King, Bohan, & Symondson, 2010). Generalists are expected to have a greater likelihood of co-occurring with a wider variety of organisms (Fridley, Vandermast, Kuppinger, Manthey, & Peet, 2007), and BMI communities in degraded environments tend to contain relatively more groups classified as generalist (von der Ohe & Goedkoop, 2013; Suga & Tanaka, 2013). We then expect stress to be positively correlated with the fraction of realized versus the potential number of edges (significant co-occurrences), that is, the connectance, of a resulting co-occurrence network (H2). With an expected increase in connectance associated with stress, as well as a decline in number of nodes (number of unique BMIs) (Blann, Anderson, Sands, & Vondracek, 2009; Shaver, Maxted, Curtis, & Carter, 1994), we also expect a decline in the number of edges.

1.1.3 | Mean co-occurrence strength

To make additional inferences on shifts in community co-occurrence patterns in relation to environmental stress, we then determined the mean strength of the co-occurrences found within each network (Araújo & Rozenfeld, 2014). For this value, we used the mean value of all of the significant correlations, as described by standardized effect-size scores (Morueta-Holme et al., 2016), within a network. For any two unique categories of organisms found in a group of communities, the standardized effect-size score represents the conditional probability, as compared to a null model, of observing one organism given the presence of the other. The mean strength of correlations

defining significant co-occurrences in a network has been observed to decline with the number of edges (Cazelles, Araújo, Mouquet, & Gravel, 2016). We then expect the number of edges in a co-occurrence network to decline with stress, as described by a mean measure of upstream land use within a watershed, along with a positive correlation between stress and the mean co-occurrence strength (H3).

1.1.4 | Modularity

Prior evidence suggests shifts in communities in response to environmental changes can be better illustrated not just from the number or strength of co-occurrences, but from their structural arrangement (Fortuna et al., 2010; Thébault & Fontaine, 2010; Tylianakis, Laliberté, Nielsen, & Bascompte, 2010). To measure these structural changes in our co-occurrence networks, we used the topological measures of modularity, defined here as the proportion of edges that occur within subnetworks less the expected proportion of such edges (Clauset, Newman, & Moore, 2004). With highly modular networks, this would be expected to lead to a co-occurrence network composed of sparsely interconnected subnetworks.

Prior observations of stressed watersheds have shown both a decline in local diversity and a rise in landscape diversity as a result of declining taxonomic similarities between individual stream communities (Hawkins, Mykrä, Oksanen, & Vander Laan, 2015; Simons, Mazor, Stein, & Nuzhdin, 2019). Given these prior observations, with regard to changes in patterns of diversity across watersheds in relation to stress, we expect the taxonomic "space" for co-occurrences to shrink with a rise in stress due to land use (Figure 1), and with it a trend toward the fracturing of assembled co-occurrence networks into weakly connected subnetworks. Similar relationships, between the modularity of co-occurrence networks and ecological stress, have also been observed in various ecological communities (Hu et al., 2017; Kay et al., 2018). To assess these trends, we use modularity, the degree to which networks are organized into clusters of weakly interconnected subnetworks (Barberán, Bates, Casamayor, & Fierer, 2012; Clauset et al., 2004). With stress expected to drive greater dissimilarity between communities, we then expect a positive relationship between the modularity of co-occurrence networks and the levels of stress experienced by the communities from which they are constructed (H4).

1.1.5 | Degree heterogeneity

To further investigate changes in the arrangement of co-occurrences, we determined the degree heterogeneity of each network, a measure of how skewed the distribution of edges per node in a network is toward the most connected nodes (Yan, Martinez, & Liu, 2017). The distribution of edges per node in ecological networks can be indicative of the structure of ecological communities, such as the likelihood of co-occurrence between generalist and specialist species (Dormann et al., 2009; Williams, 2011).
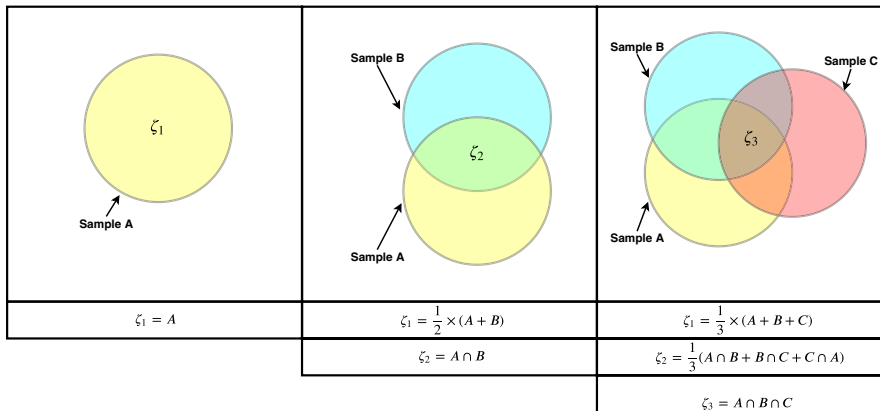
FIGURE 1 An example of a stressor reducing both the taxonomic richness of three communities, from an initial state (α_1, α_2, α_3) to a degraded state (α'_1, α'_2, α'_3), and the number of unique categories of taxa held in common between communities

Prior observations of co-occurrence networks assembled from communities at increasing levels of anthropogenic disturbance have shown a trend toward the preferential loss of taxa of low degree (Fournier, Mouly, & Gillet, 2016; Tulloch et al., 2016). Additional trends regarding ecological networks have also displayed trends toward the loss of highly keystone taxa due to environmental stresses (Araújo & Rozenfeld, 2014; Morriën et al., 2017). Given both of these trends, the loss of taxa of both high and low degrees, we then expect co-occurrence networks assembled by communities under stress have a narrow degree distribution and thus a low degree heterogeneity.

Using prior arguments regarding connectance (H2), we also make an additional argument regarding our expected trends in the degree with respect to land use. With ecological networks, connectance has been found to be negatively correlated with the skewness of their degree distributions (Poisot & Gravel, 2014). With degree heterogeneity being a measure of skew for the degree distribution of a network, this then implies that the stress experienced by the communities used to construct co-occurrence networks will be negatively correlated with their degree heterogeneity (H5).

2 | MATERIALS AND METHODS

2.1 | Sample scope

The initial scope of data covered in this analysis consists of 4,984 stream samples from 2,997 unique geographic locations across the state of California, constituting a 23-year period (1994–2016) (Mazor et al., 2015). Every sample contains the following data: BMIs enumerated and sorted to a standardized level (generally a genus-level identification except chironomids, which were identified to subfamily; Richards and Rogers, 2006), sample site altitude in meters, U.S. Geological Survey Hydrologic Unit Code 8 level watershed (Seaber, Kapinos, & Knapp, 1987), and the percent developed land use (agricultural, urban, and managed landscapes) within a 5 km clipped buffer of the watershed upstream of the sampling site, and a bioassessment index score (California Stream Condition Index [CSCI]) based on a composite of taxonomic and functional diversity within BMI assemblages (Mazor et al., 2015).

2.2 | Sample acquisitions and classifications

Approximately 55% of the BMI communities were sampled through a reach-wide protocol of Peck et al. (2006), with the remainder collected using a targeted riffle protocol, both of which produce comparable data (Gerth & Herlihy, 2006; Herbst & Silldorff, 2006; Rehn, Ode, & Hawkins, 2007). Taxa were identified to one of 334 genera, with each genus then assigned to one of eight functional feeding groups using CAMLnet (Ode, 2003). Of these eight functional feeding groups, we could conclusively assign five of them to either generalist or specialist categories (Barbour et al., 2006; De Castro et al., 2016; Feld & Hering, 2007; Mihuc, 1997; Rawer-Jost, Böhmer, Blank, & Rahmann, 2000). Using this information, we produced a measure of the number of generalist and specialist genera per sample site.

2.3 | Calculating the CSCI

Our measure of community biotic integrity at a given stream sample site was done using the CSCI. This index compares observed taxa and metrics to values expected under undisturbed reference conditions based on site-specific landscape-scale environmental variables, such as watershed area, geology, and climate (Mazor et al., 2015). This index comprises two sets of measurements using a standardized taxonomy for BMI communities (Richards and Rogers, 2006): the first being a ratio of observed-to-expected taxa (O/E), and the second a predictive multimetric index (pMMI) made of six metrics related to ecological structure and function of the BMI assemblage describing the composition of community within a site. The CSCI and its components were designed to have minimal influence from major natural gradients. This in turn has allowed for it to be used as a measure of biological conditions with a consistent meaning in different environmental settings (Reynoldson, Norris, Resh, Day, & Rosenberg, 2006; Hawkins, Olson, & Hill, 2010).

2.4 | Land use

The type and geographic extent of land use in the upstream vicinity of each sampling site data is derived from the National Land Cover Data set (NLCD) (Homer et al., 2007), with developed land

cover measured by the total percent of land cover in a designated area dedicated to agriculture, urbanization, or otherwise managed vegetative landscapes such as golf courses. The designated area for calculating percent developed land cover at each site is defined using a 5 km watershed-clipped buffer upstream of a stream sampling site using ArcGIS tools (version 10.3; Environmental Systems Research Institute) (Mazor et al., 2015). The values for land use were calculated from NLCD measurements acquired in the year 2000, though it should be noted that the sample sites in our study were located in areas where the percent developed land use was not significantly correlated with time over the duration of this study ($r = -.02$, $p = .27$).

2.5 | Sample group selection

We first filtered our initial data by selecting watersheds with 15 or more unique samples. This filtering reduced our overall data set from 4,984 to 4,619 unique samples in 2,694 unique geographic locations across 67 watersheds, while containing sample groups with sufficient data density for co-occurrence network construction. From these remaining samples, we then divided both upstream land use and sample site altitude into quintiles. Samples for network generation were then selected by randomly subsampling 10 samples within each watershed within quintiles of upstream land use and sample site altitude. For each group of 10 samples, we calculated the mean sample site altitude (altitude) and the mean percent developed upstream land use (land use), and for mean geographic separation distance in meters between samples (distance), we used the *distm* function within the R package *geosphere* (Hijmans, Williams, & Vennes, 2012). To obtain a measure of environmental heterogeneity within each sample group, we also calculated the standard deviations on altitude and land use.

2.6 | Network construction

Co-occurrence networks were then constructed using the R package *netassoc* (Morueta-Holme et al., 2016), with the presence/absence site by BMI genera as input. We chose to convert our abundance data to the presence/absence as the most conservative approach with representing our assembled database of BMI communities. Observed co-occurrences were compared against 100 randomized null communities with the same taxonomic richness as the observed community. The resulting edges were filtered so only correlations representing co-occurrences, as calculated by standardized effect-size scores, with a significance and false discovery rate less that 10^{-4} were kept. This process was repeated 100 times, with a set of 8,208 co-occurrence networks kept for analysis.

2.7 | Topological measures

Topological measures of our networks, such as size and connectance, were calculated using the packages *igraph* (v.1.2.2) (Csárdi & Nepusz, 2014) and *network* (v.1.13.0.1) (Butts, 2015) in R (v.3.5.1). The mean co-occurrence strength values were calculated, using the

R package *netassoc* (Morueta-Holme et al., 2016), as the network mean of the significant standardized effect-size scores. Modularity, defined as the proportion of edges that occur within subnetworks less the expected proportion of such edges, was calculated using the *modularity* function within the *igraph* package (Clauset et al., 2004). Degree heterogeneity was calculated as $\zeta = \frac{<k^2>}{<k>^2}$, where k represents the mean number of edges per node in a network (Yan et al., 2017).

2.8 | Modeled biotic integrity index

Using the *lm* function in the *stats* R package (v3.5.1, R Core Team, 2018), we constructed a best-fit linear model to predict the mean CSCI score of a set of samples, our measure of biotic integrity, given the topological measures of their co-occurrence networks. We then applied a backwards elimination method in order to select topological measures, which make a significant contribution to our model (Pearman, 1997; Snodgrass, 1997). In comparing the AIC scores of linear models after the removal of each topological measure, we found all five were significant. We calculated coefficients for our linear models using a 10-fold cross-validation, with 100 repeats, within the "train" function within the R package "caret" (Kuhn, 2008). To determine the relative importance of each topological measure in our linear models, and to adjust for any collinearity between measures as a result, the function *calc.relimp* was used within the *relaimpo* R package (Grömping, 2015). The relative importance of land use, altitude, and distance in describing variations in both the mean CSCI score and our modeled CSCI scores was also done using the *calc.relimp* function.

3 | RESULTS

In analyzing 8,208 co-occurrence networks, generated from communities collected from within-watershed groups with similar values for sample site land use and altitude, we found general support (Table 2) for our hypotheses (Table 1).

3.1 | Trends in co-occurrence network topology

The size of our co-occurrence networks declined significantly with a rise in land use ($r = -.81$, $p < 10^{-4}$). This reflects a general decline in both the number of genera found in an individual sampling site ($r = -.52$, $p < 10^{-4}$) and the number of functional feeding groups ($r = -.44$, $p < 10^{-4}$), in relation to land use.

While network size was found to have a strong negative correlation with land use, along with the mean number of edges per node ($r = -.56$, $p < 10^{-4}$), we still found that connectance tended to be larger in co-occurrence networks constructed from groups of stressed communities with a rise in land use ($r = .31$, $p < 10^{-4}$). This positive association between stress and connectance appears to reflect a greater relative decline in the number of nodes relative to land use ($r = -.81$, $p < 10^{-4}$) than with the number of edges ($r = -.70$, $p < 10^{-4}$).

TABLE 2 The relative importance of the topological measures used in our modeled stream health indices ($p < 10^{-4}$)

Topological measure	F (1, 8,208) (Model 1)	Relative importance (%) (Model 1)	F (1, 8,208) (Model 2)	Relative importance (%) (Model 2)
Network size	1.4×10^4	44.8	NA	NA
Connectance	704.3	10.0	2,086	19.3
Modularity	151.4	1.3	39.9	1.8
Mean co-occurrence strength	868.6	7.5	2,462	13.4
Degree heterogeneity	91.6	2.7	387.1	3.3

TABLE 3 Coefficients of sample site altitude and land use in linear models describing linear models of the percent of genera of BMIs per sample site per functional feeding group (All $p < 10^{-4}$ unless otherwise noted)

	Generalist functional feeding groups			Specialist functional feeding groups	
	Gatherers	Filterers	Omnivores	Scrapers	Shredders
Coefficient (land use)	1.6×10^{-3}	1.5×10^{-4}	3.5×10^{-4}	-9.0×10^{-5} ($p < 10^{-2}$)	-1.5×10^{-4}
Coefficient (altitude)	8.0×10^{-6} ($p < 10^{-2}$)	-1.3×10^{-5}	-5.8×10^{-6}	-9.0×10^{-5}	1.1×10^{-5}

These trends, a rise in connectance despite a decline in network size, may also reflect our observations regarding the relative abundance of unique genera classified by membership of a generalist or specialist functional feeding groups to land use. We found the proportion of genera from specialist functional feeding groups (e.g., shredders and scrapers) tended to decline with land use, while those of generalist functional feeding groups (e.g., gatherers, filterers, and omnivores) tended to increase with land use (Table 3).

In addition to a rise in connectance, networks assembled from communities with higher land use were on average found to have stronger co-occurrences ($r = .25, p < 10^{-4}$). We also found evidence of a negative relationship between both mean co-occurrence strength and the number of co-occurrences ($r = -.32, p < 10^{-4}$), and connectance ($r = -.24, p < 10^{-4}$). This potentially indicates a preferential loss of weak co-occurrences in networks assembled from communities under high levels of land use.

Weaker trends were observed with regard to variables, modularity and degree heterogeneity, which describe structural arrangements of co-occurrences. The mean modularity of our networks (0.35) was found to be both greater than the common modularity threshold of 0.3 (Newman & Girvan, 2004) and greater than that of our randomized null co-occurrence networks (0.22). Using a Wilcoxon signed-rank test, we found additional evidence for significant nonrandom structuring in our networks as their modularity values were significantly larger than their randomized null counterparts ($p < 10^{-4}$). However, despite evidence of significant network modularity there was only a relatively weak positive correlation with its value and land use ($r = .11, p < 10^{-4}$).

Across our watersheds, we observe a trend where land use is associated with a slight decline in degree heterogeneity ($r = -.10, p < 10^{-4}$). However, we did find support for our hypothesis that a decline in degree heterogeneity would be driven, at least in part,

by a rise in connectance ($r = -.60, p < 10^{-4}$). Similar to our results with modularity, we found evidence, using a Wilcoxon signed-rank test, for significantly higher values for degree heterogeneity in our co-occurrence networks than their randomized null counterparts ($p < 10^{-4}$). The higher mean degree heterogeneity of our co-occurrence networks (1.82), as compared to that of the null networks (1.14), indicates our networks are skewed more toward a relatively small number of highly connected nodes than what would be expected by chance.

3.2 | Linear models of watershed biotic integrity using co-occurrence network topology

Using five measures of co-occurrence network topology, network size (N), connectance (C), mean co-occurrence strength (S), modularity (M), and degree heterogeneity (ζ), a linear model was constructed to best predict the mean value of the CSCI score for a set of samples (Table 2). The relationship between these topological measures and our first modeled mean CSCI score per sample group is as follows:

$$\text{Mean CSCI} = 0.3 + 4.6 \times 10^{-3} \times N - 1.2 \times C - 1.8 \times 10^{-2} \times S + 0.3 \times M + 0.1 \times \zeta$$

This modeled index of watershed biotic integrity was found to be strongly correlated with the observed variation in the mean value of the CSCI score for a set of samples (Figure 2). After performing a 10-fold cross-validation, this model could still account for approximately 66% of the observed variation in the mean CSCI score. This modeled biotic integrity index was also found to vary in accordance with altitude, land use, and distance for a set of samples in a similar fashion as the mean CSCI score, although this first modeled index was less sensitive to altitude and the standard deviation on land use than the mean CSCI score (Table 4). We also observed that most of the variations observed in both the mean and our first modeled CSCI

scores were driven by land use and the standard deviation on land use (Table 4).

Both network size and the CSCI, our measure of biotic integrity, represent measures based on the taxonomic diversity of sampled communities. To focus on the potential role of the characteristics and configuration of our co-occurrences, rather than measures of local diversity alone, we then generated a second model of the mean CSCI with network size removed from our list of topological measures (Tables 2 and 4). After performing a 10-fold cross-validation, we found this second linear model can account for 38% of the observed variation in the mean CSCI score per group of samples, and it exhibits a similar trend compared to the mean CSCI as with our first model (Figure 3).

4 | DISCUSSION

We found changes in patterns of co-occurrence between genera of BMIs can play a role in describing effects of land use on regional measures of biotic integrity. This is reflected in evidence supporting our hypotheses regarding the relationships between land use and the connectance of our co-occurrence networks (H2), and the mean strength of their co-occurrences (H3). Evidence supporting our hypothesis regarding a negative correlation between network size and land use (H1) reflects a well-established link between environmental stress and both the loss of biodiversity and measures of biotic integrity (Freeman & Schorr, 2004; Garie & McIntosh, 1986; Jun et al., 2016, 2011). We find the importance of network size, along with network connectance and co-occurrence strength, reinforces prior observations on the importance of both regional and local measures of environmental quality in BMI communities (Stoll et al., 2016).

Trends relating the arrangement of co-occurrences within of our networks, as described by our hypotheses regarding modularity (H4) and degree heterogeneity (H5), were less clear. This may reflect limitations in our use of co-occurrence rather than co-abundance networks. However, prior evidence from assessments of biotic integrity

FIGURE 2 A comparison of the first modeled CSCI and mean CSCI colored by land use ($r = .81$, $p < 10^{-4}$). CSCI, California Stream Condition Index

for stream communities of BMIs has shown a strong correlation between results generated using community data sets described through the abundance or presence/absence (Beentjes, Speksnijder, Schilthuizen, Schaub, & Hoorn, 2018). The more fundamental issue may stem from differences between networks assembled from co-occurrences rather than interactions.

Analyses of co-occurrence networks have been used to identify candidate keystone taxa (Berry & Widder, 2014), potential species interactions (Veech, 2013), and the simplification of communities under ecological stress (Araújo, Rozenfeld, Rahbek, & Marquet, 2011). Though inferring co-occurrences, rather than verifying interactions, is a far more tractable problem in complex ecological systems, we must acknowledge that co-occurrences do not necessarily imply interactions. An underlying caveat with analyses involving co-occurrences is that various types of ecological interactions, such as mutualism or similar environmental requirements, may produce similar patterns of co-occurrence (Ovaskainen, Hottola, & Siitonen, 2010). In the context of our study, we observed trends between land use and both network connectance and co-occurrence strength with our BMI communities, which may reflect changes in patterns of interaction between members of generalist genera. However, our co-occurrence networks may also be incorporating information beyond potential interactions between species, such as the tendency of organisms with similar ecological niches to form co-occurrences, or for dispersal limitation to tend to limit them (Morueta-Holme et al., 2016). While co-occurrence networks, such as the ones we have constructed, may only describe potential interactions, they can still provide useful indications of changes in ecological systems (Freilich, Wieters, Broitman, Marquet, & Navarrete, 2018).

Even with these limitations we found, a simple linear model composed of topological measures of co-occurrence networks could describe a significant portion of the observed variation in the biotic integrity of our BMI communities (Table 4). Analysis of these models also suggests the topology our networks reflect more than changes in local biodiversity. While network size contributes a sizeable portion of the observed variation in biotic integrity, its removal still leaves more than half of the remaining explanatory power of our linear model of the mean CSCI score (Table 4). This suggests that we are not simply observing a decline in local diversity in response to stress but a change in landscape diversity as well.

Variations in our models appeared to be driven more by both land use and the standard deviation on land use than either altitude or geographic separation distance (Table 4). This reflects our prior observations of this system (Simons et al., 2019), whereby both the mean and standard deviation of land use are strongly correlated with degree of taxonomic dissimilarity between communities. These results are also in agreement with studies of other BMI communities where measures of environmental heterogeneity, such as variations in upstream land use between sample sites (Astorga et al., 2012; Sponseller et al., 2001), appear to drive significant changes in patterns of co-occurrence (Heino, 2013; Larsen & Ormerod, 2014; Shostell & Williams, 2007; Zhang et al., 2014).
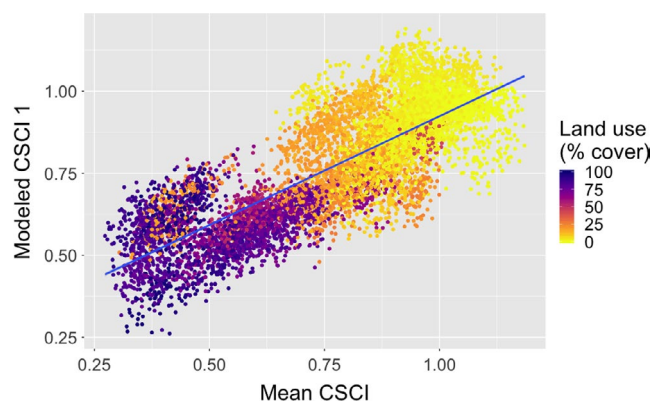
	CSCI	Modeled index 1	Modeled index 2
Proportion of variation due to altitude, F (1, 8,208)	690.5	9.0 ($p < 10^{-2}$)	189.2
Proportion of variation due to the standard deviation on altitude, F (1, 8,208)	26.6	51.8	302.9
Proportion of variation due to land use, F (1, 8,208)	2.2×10^4	1.6×10^4	3,082
Proportion of variation due to the standard deviation on land use, F (1, 8,208)	776.2	212.5	217.7
Proportion of variation due to distance, F (1, 8,208)	28.4	168.4	2.8
Relative importance of altitude (%)	11.6	5.2	7.3
Relative importance of the standard deviation on altitude (%)	7.4	4.1	2.7
Relative importance of land use (%)	33.6	35.2	11.9
Relative importance of the standard deviation on land use (%)	20.7	18.4	9.3
Relative importance of distance (%)	1.0	3.3	0.4
Proportion of variance explained by model (%)	74.2	66.3	31.6

TABLE 4 The relative importance of altitude, standard deviation on altitude, land use, standard deviation of land use, and distance in describing our linear models of the mean value of the CSCI and modeled index per HUC 8 watershed (All $p < 10^{-4}$ unless otherwise noted)

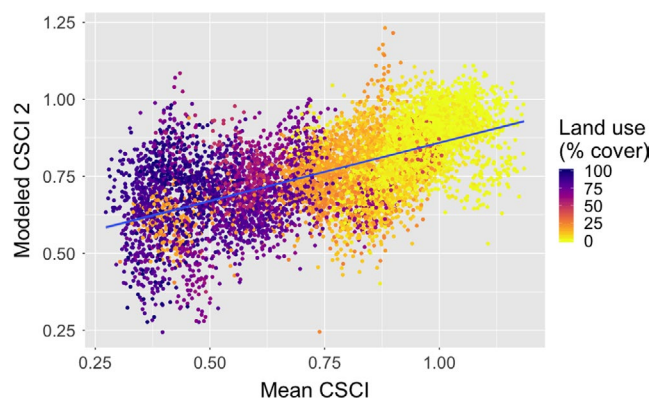Abbreviation: CSCI, California Stream Condition Index.

FIGURE 3 A comparison of the second modeled CSCI and mean CSCI colored by land use ($r = .61$, $p < 10^{-4}$). CSCI, California Stream Condition Index

These trends may reflect co-occurrence patterns unique to stream communities of BMIs. However, the framework we have used to test these hypotheses is not dependent upon the particular identities of taxa present in communities and may have the potential to be applied to other systems. Assessments of the biotic integrity of freshwater ecosystems have been carried out at a variety of spatial scales (Booth et al., 2004; King, Baker, Kazyak, & Weller, 2011; Pratt & Chang, 2012), regions (Jun, Won, Lee, Kong, & Hwang, 2012; Waite et al., 2010; Weigel & Dimick, 2011), and biological communities (Ferreira, Paiva, & Callisto, 2011; Fetscher et al., 2014; Zalack, Smucker, & Vis, 2010). With co-occurrence networks, we then assert the potential for the development of a more flexible framework for the monitoring of freshwater ecosystems, and find this direction warrants further research.

4.1 | Synthesis and future directions

It is increasingly becoming feasible to characterize entire ecological communities, from prokaryotes through metazoa, through metagenomic approaches (Baird & Hajibabaei, 2012; Bohmann et al., 2014; Deiner, Fronhofer, Mächler, Walser, & Altermatt, 2016; Jackson et al., 2016). With the ability to generate such broad and deep pictures of multiple communities, there is a commensurate need to create a framework, which could evaluate in general patterns in ecological systems in order to evaluate the biotic integrity of ecosystems. Using these stream communities as an example, our study suggests significant relationships exist between ecological stress and the structure of co-occurrence networks. We found our co-occurrence networks reflected changes in the structure of within-watershed communities in accordance with both the mean and standard deviation on land use and altitude, potentially reflecting segregation in BMI communities with a rise in measures of environmental heterogeneity. Using only a small number of topological measures, we were also able to construct a simple linear model with a close correspondence to a well-accepted index of biotic integrity. These networks are based on patterns derived from the co-occurrences of unique organisms, rather than their identities. For such reasons, we believe co-occurrence networks may have the potential to describe the biotic integrity of a diverse array of ecological communities.

CONFLICT OF INTEREST

No conflict of interest exists in this manuscript.

AUTHORS CONTRIBUTIONS

A.L.S. both constructed the co-occurrence networks used in this study and performed the analyses involved with measures of their topology. R.M. and S.T. were involved in developing the California Stream Condition Index. A.L.S. wrote the first draft of the manuscript, with all authors contributing to subsequent revisions.

DATA AVAILABILITY STATEMENT

All original data used in this study can be found on the KNB Repository: https://doi.org/10.5063/F10K26VR.

ORCID

Ariel Levi Simons https://orcid.org/0000-0003-0492-5492

Raphael Mazor https://orcid.org/0000-0003-2439-3710

Susanna Theroux https://orcid.org/0000-0002-9812-7856

REFERENCES

Ahn, S. R., & Kim, S. J. (2017). Assessment of integrated watershed health based on the natural environment, hydrology, water quality, and aquatic ecology. *Hydrology and Earth System Sciences*, *21*(11), 5583. https://doi.org/10.5194/hess-21-5583-2017

Allan, J. D. (2004). Landscapes and riverscapes: The influence of land use on stream ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, *35*, 257–284. https://doi.org/10.1146/annurev.ecolsys.35.120202.110122

Anderson, D. M., Glibert, P. M., & Burkholder, J. M. (2002). Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences. *Estuaries*, *25*(4), 704–726. https://doi.org/10.1007/BF02804901

Araújo, M. B., & Rozenfeld, A. (2014). The geographic scaling of biotic interactions. *Ecography*, *37*(5), 406–415. https://doi.org/10.1111/j.1600-0587.2013.00643.x

Araújo, M. B., Rozenfeld, A., Rahbek, C., & Marquet, P. A. (2011). Using species co-occurrence networks to assess the impacts of climate change. *Ecography*, *34*(6), 897–908. https://doi.org/10.1111/j.1600-0587.2011.06919.x

Arita, H. T. (2016). Species co-occurrence analysis: Pairwise versus matrix-level approaches. *Global Ecology and Biogeography*, *25*(11), 1397–1400. https://doi.org/10.1111/geb.12418

Astorga, A., Oksanen, J., Luoto, M., Soininen, J., Virtanen, R., & Muotka, T. (2012). Distance decay of similarity in freshwater communities: Do macro- and microorganisms follow the same rules? *Global Ecology and Biogeography*, *21*(3), 365–375. https://doi.org/10.1111/j.1466-8238.2011.00681.x

Baird, D. J., & Hajibabaei, M. (2012). Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology*, *21*(8), 2039–2044. https://doi.org/10.1111/j.1365-294X.2012.05519.x

Barberán, A., Bates, S. T., Casamayor, E. O., & Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME Journal*, *6*(2), 343. https://doi.org/10.1038/ismej.2011.119

Barbour, M. T., Gerritsen, J., Griffith, G. E., Frydenborg, R., McCarron, E., White, J. S., & Bastian, M. L. (2006). A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society*, *15*(2), 185–211. https://doi.org/10.2307/1467948

Beentjes, K. K., Speksnijder, A. G. C. L., Schilthuizen, M., Schaub, B. E. M., & van der Hoorn, B. B. (2018). The influence of macroinvertebrate abundance on the assessment of freshwater quality in The Netherlands. *Metabarcoding and Metagenomics*, *1*, https://doi.org/10.3897/mbmg.2.26744

Bell, J. R., King, R. A., Bohan, D. A., & Symondson, W. O. C. (2010). Spatial co-occurrence networks predict the feeding histories of polyphagous arthropod predators at field scales. *Ecography*, *33*(1), 64–72. https://doi.org/10.1111/j.1600-0587.2009.06046.x

Bellisario, B., Cerfolli, F., & Nascetti, G. (2010). Spatial network structure and robustness of detritus-based communities in a patchy environment. *Ecological Research*, *25*(4), 813–821. https://doi.org/10.1007/s11284-010-0711-5

Bergseth, B. J., Russ, G. R., & Cinner, J. E. (2015). Measuring and monitoring compliance in no-take marine reserves. *Fish and Fisheries*, *16*(2), 240–258. https://doi.org/10.1111/faf.12051

Berry, D., & Widder, S. (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in Microbiology*, *5*, 219. https://doi.org/10.3389/fmicb.2014.00219

Blann, K. L., Anderson, J. L., Sands, G. R., & Vondracek, B. (2009). Effects of agricultural drainage on aquatic ecosystems: A review. *Critical Reviews in Environmental Science and Technology*, *39*(11), 909–1001. https://doi.org/10.1080/10643380801977966

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., ... de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution*, *29*(6), 358–367. https://doi.org/10.1016/j.tree.2014.04.003

Booth, D. B., Karr, J. R., Schauman, S., Konrad, C. P., Morley, S. A., Larson, M. G., & Burges, S. J. (2004). Reviving urban streams: Land use, hydrology, biology, and human behavior. *Journal of the American Water Resources Association*, *40*(5), 1351–1364. https://doi.org/10.1111/j.1752-1688.2004.tb01591.x

Borthagaray, A. I., Arim, M., & Marquet, P. A. (2014). Inferring species roles in metacommunity structure from species co-occurrence networks. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1792), 1425. https://doi.org/10.1098/rspb.2014.1425

Büchi, L., & Vuilleumier, S. (2014). Coexistence of specialist and generalist species is shaped by dispersal and environmental factors. *The American Naturalist*, *183*(5), 612–624. https://doi.org/10.1086/675756

Butts, C. T. (2015). network: A package for managing relational data in R. *Journal of Statistical Software*, *24*(2), 1–36. https://doi.org/10.18637/jss.v024.i02

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications*, *8*(3), 559–568. https://doi.org/10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2

Cazelles, K., Araújo, M., Mouquet, N., & Gravel, D. (2016). A theory for species co-occurrence in interaction networks. *Theoretical Ecology*, *9*(1), 39–48. https://doi.org/10.1007/s12080-015-0281-9

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *70*(6), 066111. https://doi.org/10.1103/PhysRevE.70.066111

Corona, P., Chirici, G., McRoberts, R. E., Winter, S., & Barbati, A. (2011). Contribution of large-scale forest inventories to biodiversity assessment and monitoring. *Forest Ecology and Management*, 262(11), 2061–2069. https://doi.org/10.1016/j.foreco.2011.08.044

Csárdi, G., & Nepusz, T. (2014). The igraph software package for complex network research. *Journal of Computer Applications*, 1695(5), 1–9. https://doi.org/10.3724/SP.J.1087.2009.02191

Cuffney, T. F., Brightbill, R. A., May, J. T., & Waite, I. R. (2010). Responses of benthic macroinvertebrates to environmental changes associated with urbanization in nine metropolitan areas. *Ecological Applications*, 20(5), 1384–1401. https://doi.org/10.1890/08-1311.1

De Castro, D. M. P., De Carvalho, D. R., Pompeu, P. D. S., Moreira, M. Z., Nardoto, G. B., & Callisto, M. (2016). Land use influences niche size and the assimilation of resources by benthic macroinvertebrates in tropical headwater streams. *PLoS ONE*, 11(3), e0150527. https://doi.org/10.1371/journal.pone.0150527

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J. C., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature Communications*, 7, 12544. https://doi.org/10.1038/ncomms12544

Dormann, C. F., Frund, J., Bluthgen, N., & Gruber, B. (2009). Indices, graphs and null models: Analyzing bipartite ecological networks. *The Open Ecology Journal*, 2(1), https://doi.org/10.2174/1874213000902010007

Ducatez, S., Tingley, R., & Shine, R. (2014). Using species co-occurrence patterns to quantify relative habitat breadth in terrestrial vertebrates. *Ecosphere*, 5(12), 1–12. https://doi.org/10.1890/ES14-00332.1

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., … Sullivan, C. A. (2006). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews of the Cambridge Philosophical Society*, 81(2), 163–182. https://doi.org/10.1017/S1464793105006950

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, 8(10), 1265–1275. https://doi.org/10.1111/2041-210X.12789

Fausch, K. D., Lyons, J., Karr, J. R., & Angermeier, P. L. (1990). Fish communities as indicators of environmental degradation. *American Fisheries Society Symposium*, 8(1), 123–144.

Feld, C. K., & Hering, D. (2007). Community structure or function: Effects of environmental stress on benthic macroinvertebrates at different spatial scales. *Freshwater Biology*, 52(7), 1380–1399. https://doi.org/10.1111/j.1365-2427.2007.01749.x

Ferreira, W., Paiva, L., & Callisto, M. (2011). Development of a benthic multimetric index for biomonitoring of a neotropical watershed. *Brazilian Journal of Biology*, 71(1), 15–25. https://doi.org/10.1590/s1519-69842011000100005

Fetscher, A. E., Stancheva, R., Kociolek, J. P., Sheath, R. G., Stein, E. D., Mazor, R. D., … Busse, L. B. (2014). Development and comparison of stream indices of biotic integrity using diatoms vs. non-diatom algae vs. a combination. *Journal of Applied Phycology*, 26(1), 433–450. https://doi.org/10.1007/s10811-013-0088-2

Foley, M. M., Armsby, M. H., Prahler, E. E., Caldwell, M. R., Erickson, A. L., Kittinger, J. N., … Levin, P. S. (2013). Improving ocean management through the use of ecological principles and integrated ecosystem assessments. *BioScience*, 63(8), 619–631. https://doi.org/10.1525/bio.2013.63.8.5

Fortuna, M. A., Stouffer, D. B., Olesen, J. M., Jordano, P., Mouillot, D., Krasnov, B. R., … Bascompte, J. (2010). Nestedness versus modularity in ecological networks: Two sides of the same coin? *Journal of Animal Ecology*, 79(4), 811–817. https://doi.org/10.1111/j.1365-2656.2010.01688.x

Fournier, B., Mouly, A., & Gillet, F. (2016). Multiple assembly rules drive the co-occurrence of orthopteran and plant species in grasslands: combining network, Functional and phylogenetic approaches. *Frontiers in Plant Science*, 7, 1224. https://doi.org/10.3389/fpls.2016.01224

Freeman, P. L., & Schorr, M. S. (2004). Influence of watershed urbanization on fine sediment and macroinvertebrate assemblage characteristics in tennessee ridge and valley streams. *Journal of Freshwater Ecology*, 19(3), 353–362. https://doi.org/10.1080/02705060.2004.9664908

Freilich, M. A., Wieters, E., Broitman, B. R., Marquet, P. A., & Navarrete, S. A. (2018). Species co-occurrence networks: Can they reveal trophic and non-trophic interactions in ecological communities? *Ecology*, 99(3), 690–699. https://doi.org/10.1002/ecy.2142

Fridley, J. D., Vandermast, D. B., Kuppinger, D. M., Manthey, M., & Peet, R. K. (2007). Co-occurrence based assessment of habitat generalists and specialists: A new approach for the measurement of niche width. *Journal of Ecology*, 95(4), 707–722. https://doi.org/10.1111/j.1365-2745.2007.01236.x

Fuhrman, J. A., & Steele, J. A. (2008). Community structure of marine bacterioplankton: Patterns, networks, and relationships to function. *Aquatic Microbial Ecology*, 53(1), 69–81. https://doi.org/10.3354/ame01222

Garie, H. L., & McIntosh, A. (1986). Distribution of benthic macroinvertebrates in a stream exposed to urban runoff. *Journal of the American Water Resources Association*, 22(3), 447–455. https://doi.org/10.1111/j.1752-1688.1986.tb01899.x

Gerth, W. J., & Herlihy, A. T. (2006). Effect of sampling different habitat types in regional macroinvertebrate bioassessment surveys. *Journal of the North American Benthological Society*, 25(2), 501–512. https://doi.org/10.1899/0887-3593(2006)25[501:eosdht]2.0.co;2

Goldwasser, L., & Roughgarden, J. (1997). Sampling effects and the estimation of food-web properties. *Ecology*, 78(1), 41–54. https://doi.org/10.1890/0012-9658(1997)078[0041:SEATEO]2.0.CO;2

Goodwin, K. D., Thompson, L. R., Duarte, B., Kahlke, T., Thompson, A. R., Marques, J. C., & Caçador, I. (2017). DNA sequencing as a tool to monitor marine ecological status. *Frontiers in Marine Science*, 4, 107. https://doi.org/10.3389/fmars.2017.00107

Gotelli, N. J. (2000). Null model analysis of species co-occurrence patterns. *Ecology*, 81(9), 2606–2621. https://doi.org/10.1890/0012-9658(2000)081[2606:NMAOSC]2.0.CO;2

Grömping, U. (2015). Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software*, 17(1), 1–27. https://doi.org/10.18637/jss.v017.i01

Grönroos, M., Heino, J., Siqueira, T., Landeiro, V. L., Kotanen, J., & Bini, L. M. (2013). Metacommunity structuring in stream networks: Roles of dispersal mode, distance type, and regional environmental context. *Ecology and Evolution*, 3(13), 4473–4487. https://doi.org/10.1002/ece3.834

Guerry, A. D., Polasky, S., Lubchenco, J., Chaplin-Kramer, R., Daily, G. C., Griffin, R., … Vira, B. (2015). Natural capital and ecosystem services informing decisions: From promise to practice. *Proceedings of the National Academy of Sciences*, 112(24), 7348–7355. https://doi.org/10.1073/pnas.1503751112

Hawkins, C. P., Mykrä, H., Oksanen, J., & Vander Laan, J. J. (2015). Environmental disturbance can increase beta diversity of stream macroinvertebrate assemblages. *Global Ecology and Biogeography*, 24(4), 483–494. https://doi.org/10.1111/geb.12254

Hawkins, C. P., Olson, J. R., & Hill, R. A. (2010). The reference condition: Predicting benchmarks for ecological and water-quality assessments. *Journal of the North American Benthological Society*, 29(1), 312–343. https://doi.org/10.1899/09-092.1

Heino, J. (2013). Environmental heterogeneity, dispersal mode, and co-occurrence in stream macroinvertebrates. *Ecology and Evolution*, 3(2), 344–355. https://doi.org/10.1002/ece3.470

Herbst, D. B., & Silldorff, E. L. (2006). Comparison of the performance of different bioassessment methods: Similar evaluations of biotic integrity from separate programs and procedures. *Journal of the*

North American Benthological Society, 25(2), 513–530. https://doi.org/10.1899/0887-3593(2006)25[513:cotpod]2.0.co;2

Hering, D., Borja, A., Jones, J. I., Pont, D., Boets, P., Bouchez, A., … Kelly, M. (2018). Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Research*, 138, 192–205. https://doi.org/10.1016/j.watres.2018.03.003

Hijmans, R. J., Williams, E., & Vennes, C. (2012). geosphere: Spherical Trigonometry. R package version 1.2–28. *Package Geosphere*.

Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., … Wickham, J. (2007). Completion of the 2001 National Land Cover Database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, 73(4), 337.

Hu, A., Ju, F., Hou, L., Li, J., Yang, X., Wang, H., … Yu, C. P. (2017). Strong impact of anthropogenic contamination on the co-occurrence patterns of a riverine microbial community. *Environmental Microbiology*, 19(12), 4993–5009. https://doi.org/10.1111/1462-2920.13942

Jackson, M. C., Weyl, O. L. F., Altermatt, F., Durance, I., Friberg, N., Dumbrell, A. J., … Woodward, G. (2016). Recommendations for the next generation of global freshwater biological monitoring tools. *Advances in Ecological Research*, 55, 615–636. https://doi.org/10.1016/bs.aecr.2016.08.008

Jun, Y. C., Kim, N. Y., Kim, S. H., Park, Y. S., Kong, D. S., & Hwang, S. J. (2016). Spatial distribution of benthic macroinvertebrate assemblages in relation to environmental variables in Korean nationwide streams. *Water (Switzerland)*, 8(1), 27. https://doi.org/10.3390/w8010027

Jun, Y.-C., Kim, N.-Y., Kwon, S.-J., Han, S.-C., Hwang, I.-C., Park, J.-H., … Hwang, S.-J. (2011). Effects of land use on benthic macroinvertebrate communities: Comparison of two mountain streams in Korea. *Annales De Limnologie - International Journal of Limnology*, 47(S1), S35–S49. https://doi.org/10.1051/limn/2011018

Jun, Y. C., Won, D. H., Lee, S. H., Kong, D. S., & Hwang, S. J. (2012). A multimetric benthic macroinvertebrate index for the assessment of stream biotic integrity in Korea. *International Journal of Environmental Research and Public Health*, 9(10), 3599–3628. https://doi.org/10.3390/ijerph9103599

Kay, G. M., Tulloch, A., Barton, P. S., Cunningham, S. A., Driscoll, D. A., & Lindenmayer, D. B. (2018). Species co-occurrence networks show reptile community reorganization under agricultural transformation. *Ecography*, 41(1), 113–125. https://doi.org/10.1111/ecog.03079

Kerans, B. L., & Karr, J. R. (1994). A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee valley. *Ecological Applications*, 4(4), 768–785. https://doi.org/10.2307/1942007

King, R. S., Baker, M. E., Kazyak, P. F., & Weller, D. E. (2011). How novel is too novel? Stream community thresholds at exceptionally low levels of catchment urbanization. *Ecological Applications*, 21(5), 1659–1678. https://doi.org/10.1890/10-1357.1

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5), 1–26. https://doi.org/10.18637/jss.v028.i05

Lakew, A., & Moog, O. (2015). A multimetric index based on benthic macroinvertebrates for assessing the ecological status of streams and rivers in central and southeast highlands of Ethiopia. *Hydrobiologia*, 751(1), 229–242. https://doi.org/10.1007/s10750-015-2189-1

Larsen, S., & Ormerod, S. J. (2014). Anthropogenic modification disrupts species co-occurrence in stream invertebrates. *Global Change Biology*, 20(1), 51–60. https://doi.org/10.1111/gcb.12355

Laurance, W. F., Sayer, J., & Cassman, K. G. (2014). Agricultural expansion and its impacts on tropical nature. *Trends in Ecology and Evolution*, 29(2), 107–116. https://doi.org/10.1016/j.tree.2013.12.001

Lenat, D. R., & Crawford, J. K. (1994). Effects of land use on water quality and aquatic biota of three North Carolina Piedmont streams. *Hydrobiologia*, 294(3), 185–199. https://doi.org/10.1007/BF00021291

Mandakovic, D., Rojas, C., Maldonado, J., Latorre, M., Travisany, D., Delage, E., … González, M. (2018). Structure and co-occurrence patterns in microbial communities under acute environmental stress reveal ecological factors fostering resilience. *Scientific Reports*, 8(1), 5875. https://doi.org/10.1038/s41598-018-23931-0

Masese, F. O., Raburu, P. O., & Muchiri, M. (2009). A preliminary benthic macroinvertebrate index of biotic integrity (B-IBI) for monitoring the Moiben River, Lake Victoria Basin, Kenya. *African Journal of Aquatic Science*, 34(1), 1–4. https://doi.org/10.2989/AJAS.2009.34.1.1.726

Maxted, J. R., Barbour, M. T., Gerritsen, J., Poretti, V., Primrose, N., Silvia, A., … Renfrow, R. (2000). Assessment framework for mid-Atlantic coastal plain streams using benthic macroinvertebrates. *Journal of the North American Benthological Society*, 19(1), 128–144. https://doi.org/10.2307/1468286

Mazor, R. D., Rehn, A. C., Ode, P. R., Engeln, M., Schiff, K. C., Stein, E. D., … Hawkins, C. P. (2015). Bioassessment in complex environments: Designing an index for consistent meaning in different settings. *Freshwater Science*, 35(1), 249–271. https://doi.org/10.1086/684130

Mihuc, T. B. (1997). The functional trophic role of lotic primary consumers: Generalist versus specialist strategies. *Freshwater Biology*, 37(2), 455–462. https://doi.org/10.1046/j.1365-2427.1997.00175.x

Morriën, E., Hannula, S. E., Snoek, L. B., Helmsing, N. R., Zweers, H., de Hollander, M., … van der Putten, W. H. (2017). Soil networks become more connected and take up more carbon as nature restoration progresses. *Nature Communications*, 8, 14349. https://doi.org/10.1038/ncomms14349

Morueta-Holme, N., Blonder, B., Sandel, B., McGill, B. J., Peet, R. K., Ott, J. E., … Svenning, J.-C. (2016). A network approach for inferring species associations from co-occurrence data. *Ecography*, 39(12), 1139–1150. https://doi.org/10.1111/ecog.01892

Moyle, P. B., & Randall, P. J. (1998). Evaluating the biotic integrity of watersheds in the Sierra Nevada, California. *Conservation Biology*, 12(6), 1318–1326. https://doi.org/10.1111/j.1523-1739.1998.97161.x

Mykrä, H., & Heino, J. (2017). Decreased habitat specialization in macroinvertebrate assemblages in anthropogenically disturbed streams. *Ecological Complexity*, 31, 181–188. https://doi.org/10.1016/j.ecocom.2017.07.002

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 69(2), 026113. https://doi.org/10.1103/PhysRevE.69.026113

Nielsen, A., & Bascompte, J. (2007). Ecological networks, nestedness and sampling effort. *Journal of Ecology*, 95(5), 1134–1141. https://doi.org/10.1111/j.1365-2745.2007.01271.x

Novotny, V., Bartošová, A., O'Reilly, N., & Ehlinger, T. (2005). Unlocking the relationship of biotic integrity of impaired waters to anthropogenic stresses. *Water Research*, 39(1), 184–198. https://doi.org/10.1016/j.watres.2004.09.002

Ode, P. R. (2003). CAMLnet: List of California macroinvertebrate taxa and standard taxonomic effort. Aquatic Bioassessment Laboratory, Rancho Cordova. Retrieved September 10, 2014.

Ourso, R. T., & Frenzel, S. A. (2003). Identification of linear and threshold responses in streams along a gradient of urbanization in Anchorage, Alaska. *Hydrobiologia*, 501(1–3), 117–131. https://doi.org/10.1023/A:1026211808745

Ovaskainen, O., Hottola, J., & Siitonen, J. (2010). Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. *Ecology*, 91(9), 2514–2521. https://doi.org/10.1890/10-0173.1

Paerl, H. W., Gardner, W. S., Havens, K. E., Joyner, A. R., McCarthy, M. J., Newell, S. E., … Scott, J. T. (2016). Mitigating cyanobacterial harmful algal blooms in aquatic ecosystems impacted by climate change and anthropogenic nutrients. *Harmful Algae*, 54, 213–222. https://doi.org/10.1016/j.hal.2015.09.009

Paul, M. J., & Meyer, J. L. (2001). Streams in the urban landscape. *Annual Review of Ecology and Systematics*, 32(1), 333–365. https://doi.org/10.1146/annurev.ecolsys.32.081501.114040

Pearman, P. B. (1997). Correlates of amphibian diversity in an altered landscape of Amazonian Ecuador. *Conservation Biology*, 11(5), 1211–1225. https://doi.org/10.1046/j.1523-1739.1997.96202.x

Peck, D. V., Herlihy, A. T., Hill, B. H., Hughes, R. M., Kaufmann, P. R., Klemm, D. J., …Cappaert, M. (2006). *Environmental Monitoring and Assessment Program—Surface Waters Western Pilot study: Field operations manual for wadeable streams*. EPA/620/R-06/003. Corvalis, OR. Washington, D.C: U.S. Environmental Protection Agency, Office of Research and Development.

Poisot, T., & Gravel, D. (2014). When is an ecological network complex? Connectance drives degree distribution and emerging network properties. *PeerJ*, 2, e251. https://doi.org/10.7717/peerj.251

Pratt, B., & Chang, H. (2012). Effects of land cover, topography, and built structure on seasonal water quality at multiple spatial scales. *Journal of Hazardous Materials*, 209, 48–58. https://doi.org/10.1016/j.jhazmat.2011.12.068

Rawer-Jost, C., Böhmer, J., Blank, J., & Rahmann, H. (2000). Macroinvertebrate functional feeding group methods in ecological assessment. *Assessing the Ecological Integrity of Running Waters*, 422(423), 225–232. https://doi.org/10.1007/978-94-011-4164-2_18

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rehn, A. C., Ode, P. R., & Hawkins, C. P. (2007). Comparisons of targeted-riffle and reach-wide benthic macroinvertebrate samples: Implications for data sharing in stream-condition assessments. *Journal of the North American Benthological Society*, 26(2), 332–348. https://doi.org/10.1899/0887-3593(2007)26[332:cotarb]2.0.co;2

Reynoldson, T. B., Norris, R. H., Resh, V. H., Day, K. E., & Rosenberg, D. M. (2006). The reference condition: A comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society*, 16(4), 833–852. https://doi.org/10.2307/1468175

Richards, A. B., & Rogers, D. C. (2006). *List of freshwater macroinvertebrate taxa from California and adjacent states including standard taxonomic effort levels*. Retrieved from http://www.safit.org/ste.html

Ruan, Q., Dutta, D., Schwalbach, M. S., Steele, J. A., Fuhrman, J. A., & Sun, F. (2006). Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics*, 22(20), 2532–2538. https://doi.org/10.1093/bioinformatics/btl417

Schmeller, D. S., Julliard, R., Bellingham, P. J., Böhm, M., Brummitt, N., Chiarucci, A., … Belnap, J. (2015). Towards a global terrestrial species monitoring program. *Journal for Nature Conservation*, 25, 51–57. https://doi.org/10.1016/j.jnc.2015.03.003

Seaber, P. R., Kapinos, F. P., & Knapp, G. L. (1987). *Hydrologic units maps*. U. S. Geological Survey Water-Supply Paper 2294, 63 p.

Shaver, E., Maxted, J., Curtis, G., Carter, D., (1994). Watershed protection using an integrated approach. In *Stormwater NPDES related monitoring needs*, (pp. 435–459). ASCE.

Shostell, J. M., Williams, B. S. (2007). Habitat complexity as a determinate of benthic macroinvertebrate community structure in cypress tree reservoirs. *Hydrobiologia*, 575(1), 389–399. https://doi.org/10.1007/s10750-006-0385-8

Silva, D. R. O., Herlihy, A. T., Hughes, R. M., & Callisto, M. (2017). An improved macroinvertebrate multimetric index for the assessment of wadeable streams in the neotropical savanna. *Ecological Indicators*, 81, 514–525. https://doi.org/10.1016/j.ecolind.2017.06.017

Simons, A. L., Mazor, R., Stein, E. D., & Nuzhdin, S. (2019). Using alpha, beta, and zeta diversity in describing the health of stream-based benthic macroinvertebrate communities. *Ecological Applications*, 29(4), e01896. https://doi.org/10.1002/eap.1896

Smith, R. F., & Lamp, W. O. (2008). Comparison of insect communities between adjacent headwater and main-stem streams in urban and rural watersheds. *Journal of the North American Benthological Society*, 27(1), 161–175. https://doi.org/10.1899/07-071.1

Snodgrass, J. W. (1997). Temporal and spatial dynamics of beaver-created patches as influenced by management practices in a South-Eastern North American Landscape. *Journal of Applied Ecology*, 34(4), 1043–1056. https://doi.org/10.2307/2405293

Socolar, J. B., Gilroy, J. J., Kunin, W. E., & Edwards, D. P. (2016). How should beta-diversity inform biodiversity conservation? *Trends in Ecology and Evolution*, 31(1), 67–80. https://doi.org/10.1016/j.tree.2015.11.005

Sponseller, R. A., Benfield, E. F., & Valett, H. M. (2001). Relationships between land use, spatial scale and stream macroinvertebrate communities. *Freshwater Biology*, 46(10), 1409–1424. https://doi.org/10.1046/j.1365-2427.2001.00758.x

Steele, J. A., Countway, P. D., Xia, L. I., Vigil, P. D., Beman, J. M., Kim, D. Y., … Fuhrman, J. A. (2011). Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISME Journal*, 5, 1414–1425. https://doi.org/10.1038/ismej.2011.24

Steenweg, R., Hebblewhite, M., Kays, R., Ahumada, J., Fisher, J. T., Burton, C., … Rich, L. N. (2017). Scaling-up camera traps: Monitoring the planet's biodiversity with networks of remote sensors. *Frontiers in Ecology and the Environment*, 15(1), 26–34. https://doi.org/10.1002/fee.1448

Stein, E. D., White, B. P., Mazor, R. D., Jackson, J. K., Battle, J. M., Miller, P. E., … Sweeney, B. W. (2014). Does DNA barcoding improve performance of traditional stream bioassessment metrics? *Freshwater Science*, 33(1), 302–311. https://doi.org/10.1086/674782

Stepenuck, K. F., Crunkilton, R. L., & Wang, L. (2002). Impacts of urban landuse on macroinvertebrate communities in southeastern Wisconsin streams. *Journal of the American Water Resources Association*, 38(4), 1041–1051. https://doi.org/10.1111/j.1752-1688.2002.tb05544.x

Stoll, S., Breyer, P., Tonkin, J. D., Früh, D., & Haase, P. (2016). Scale-dependent effects of river habitat quality on benthic invertebrate communities - Implications for stream restoration practice. *Science of the Total Environment*, 553, 495–503. https://doi.org/10.1016/j.scitotenv.2016.02.126

Suga, C. M., & Tanaka, M. O. (2013). Influence of a forest remnant on macroinvertebrate communities in a degraded tropical stream. *Hydrobiologia*, 703(1), 203–213. https://doi.org/10.1007/s10750-012-1360-1

Sweeney, B. W., Battle, J. M., Jackson, J. K., & Dapkey, T. (2011). Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *Journal of the North American Benthological Society*, 30(1), 195–216. https://doi.org/10.1899/10-016.1

Thébault, E., & Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, 329(5993), 853–856. https://doi.org/10.1126/science.1188321

Tulloch, A. I. T., Chadès, I., Dujardin, Y., Westgate, M. J., Lane, P. W., & Lindenmayer, D. (2016). Dynamic species co-occurrence networks require dynamic biodiversity surrogates. *Ecography*, 39(12), 1185–1196. https://doi.org/10.1111/ecog.02143

Tylianakis, J. M., Laliberté, E., Nielsen, A., & Bascompte, J. (2010). Conservation of species interaction networks. *Biological Conservation*, 143(10), 2270–2279. https://doi.org/10.1016/j.biocon.2009.12.004

Vander Laan, J. J., Hawkins, C. P., Olson, J. R., & Hill, R. A. (2013). Linking land use, in-stream stressors, and biological condition to infer causes of regional ecological impairment in streams. *Freshwater Science*, 32(3), 801–820. https://doi.org/10.1899/12-186.1

Veech, J. A. (2013). A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography*, 22(2), 252–260. https://doi.org/10.1111/j.1466-8238.2012.00789.x

Vieira, C., Séneca, A., Sérgio, C., & Ferreira, M. T. (2012). Bryophyte taxonomic and functional groups as indicators of fine scale ecological gradients in mountain streams. *Ecological Indicators*, *18*, 98–107. https://doi.org/10.1016/j.ecolind.2011.10.012

Vile, J. S., & Henning, B. F. (2018). Development of indices of biotic integrity for high-gradient wadeable rivers and headwater streams in New Jersey. *Ecological Indicators*, *90*, 469–484. https://doi.org/10.1016/j.ecolind.2018.03.027

von der Ohe, P. C., & Goedkoop, W. (2013). Distinguishing the effects of habitat degradation and pesticide stress on benthic invertebrates using stressor-specific metrics. *Science of the Total Environment*, *444*, 480–490. https://doi.org/10.1016/J.SCITOTENV.2012.12.001

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., … Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature*, *467*, 555–561. https://doi.org/10.1038/nature09440

Voß, K., & Schäfer, R. B. (2017). Taxonomic and functional diversity of stream invertebrates along an environmental stress gradient. *Ecological Indicators*, *81*, 235–242. https://doi.org/10.1016/j.ecolind.2017.05.072

Waite, I. R., Brown, L. R., Kennen, J. G., May, J. T., Cuffney, T. F., Orlando, J. L., & Jones, K. A. (2010). Comparison of watershed disturbance predictive models for stream benthic macroinvertebrates for three distinct ecoregions in western US. *Ecological Indicators*, *10*(6), 1125–1136. https://doi.org/10.1016/j.ecolind.2010.03.011

Wallace, A. M., & Biastoch, R. G. (2016). Detecting changes in the benthic invertebrate community in response to increasing chloride in streams in Toronto, Canada. *Freshwater Science*, *35*(1), 353–363. https://doi.org/10.1086/685297

Weigel, B. M., & Dimick, J. J. (2011). Development, validation, and application of a macroinvertebrate-based Index of Biotic Integrity for nonwadeable rivers of Wisconsin. *Journal of the North American Benthological Society*, *30*(3), 665–679. https://doi.org/10.1899/10-161.1

Widder, S., Besemer, K., Singer, G. A., Ceola, S., Bertuzzo, E., Quince, C., … Battin, T. J. (2014). Fluvial network organization imprints on microbial co-occurrence networks. *Proceedings of the National Academy of Sciences*, *111*(35), 12799–12804. https://doi.org/10.1073/pnas.1411723111

Williams, R. J. (2011). Biology, methodology or chance? The degree distributions of bipartite ecological networks. *PLoS ONE*, *6*(3), e17645. https://doi.org/10.1371/journal.pone.0017645

Yan, G., Martinez, N. D., & Liu, Y. Y. (2017). Degree heterogeneity and stability of ecological networks. *Journal of the Royal Society Interface*, *131*(14), https://doi.org/10.1098/rsif.2017.0189

Zalack, J. T., Smucker, N. J., & Vis, M. L. (2010). Development of a diatom index of biotic integrity for acid mine drainage impacted streams. *Ecological Indicators*, *10*(2), 287–295. https://doi.org/10.1016/j.ecolind.2009.06.003

Zhang, Y., Liu, L., Cheng, L., Cai, Y., Yin, H., Gao, J., & Gao, Y. (2014). Macroinvertebrate assemblages in streams and rivers of a highly developed region (Lake Taihu Basin, China). *Aquatic Biology*, *23*(1), 15–28. https://doi.org/10.3354/ab00600