# ECOGRAPHY

Diatoms (*Eunotia* sp.) growing on a red alga (*Batrachospermum* sp.). Algae form the base of the food chain in stream ecosystems and are instrumental in maintaining water quality, which necessitates a better understanding of the factors controlling their species and communities. Budnick et al. demonstrate that both algal species co-occurrence networks and community composition depend on climate and dispersal across nutrient supplies and ratios but to a different extent. Therefore, multi-level approaches are needed to predict how algae would respond to changes in their environment.

## Research

# The impacts of nutrient supply and imbalance on subcontinental co-occurrence networks and metacommunity composition of stream algae

William R. Budnick, Joseph L. Mruzek, Chad A. Larson and Sophia I. Passy

*W. R. Budnick (https://orcid.org/0000-0001-9288-6782), J. L. Mruzek (https://orcid.org/0000-0002-8067-3824) and S. I. Passy (https://orcid.org/0000-0002-8230-9380) ✉ (Sophia.passy@uta.edu), Dept of Biology, Univ. of Texas at Arlington, Arlington, TX, USA. – C. A. Larson (https://orcid.org/0000-0002-0329-8979), Washington State Dept of Ecology, Environmental Assessment Program, Lacey, WA, USA. WRB also at: Dept of Fisheries and Wildlife, Michigan State Univ., East Lansing, MI, USA.*

The amounts and ratios of nutrients (nitrogen and phosphorus) are important determinants of producer community biodiversity and composition and their responses to climate and dispersal. However, the nutrient effects on co-occurrence network topology, particularly in freshwaters, are understudied. Here, we investigate 1) whether nutrient supply and ratio constrain topological properties of algal co-occurrence networks in streams and 2) to what extent climate and space (a surrogate for dispersal) affect co-occurrence network topology versus metacommunity composition across nutrient supply and ratio contexts. We used a subcontinental dataset of benthic algae from 840 stream localities in the conterminous US. We constructed co-occurrence networks representing nutrient supply contexts (oligotrophic versus eutrophic) and nutrient ratio contexts (N-limited versus P-limited) and statistically assessed topological variability within each pair via randomization. We then used a null model framework and direct gradient analysis to ascertain the importance of climate and space in driving, respectively, network topology and metacommunity composition. Nutrient supply was only positively related to network size (species node counts), which was driven by motile species, while other topological differences were non-significant. Climatic and spatial variables had pronounced and for the most part comparable effects on network topology that further depended on nutrient context. A comparative assessment of topological versus compositional responses to climate and space across nutrient contexts identified both similarities and differences. While climate and space contributed to both network topology and metacommunity composition, space was a stronger predictor of compositional variability than climate, regardless of nutrient context. Our findings highlight the need for developing integrative multi-level approaches (from metacommunities to co-occurrence networks) to fully understand biological responses to complex and interactive abiotic forces.

Keywords: algae, climate, co-occurrence networks, dispersal, ecological guilds, metacommunity composition

## Introduction

The world's freshwater ecosystems are undergoing global changes in nutrient inputs, which are either increasing, causing eutrophication and deterioration of water quality (Dodds and Smith 2016, Stoddard et al. 2016) or decreasing, leading to oligotrophication and potentially restoration of the original state (Flaim et al. 2016, Verbeek et al. 2018). As anthropogenic factors continue to force the pendulum to shift between nutrient extremes, how species and communities respond is a topic of great research interest and environmental concern (Tilman and Isbell 2015, Wang et al. 2016).

Numerous investigations have shown that both availability and balance (i.e. relative proportions) of major nutrients, such as nitrogen and phosphorus, constrain biodiversity, composition and biomass production across ecosystems (Elser et al. 2007, Cardinale et al. 2009, Harpole et al. 2011, Lewandowska et al. 2016, Cook et al. 2018). However, we are just beginning to understand the impacts of nutrient supply and balance on the topology of species co-occurrence networks. Co-occurrence networks graphically represent the positive and negative pairwise species associations within a metacommunity, and the topological properties of these networks are sensitive to ecological gradients (Poisot et al. 2015). For example, eutrophication was shown to increase species connectance (inter-connectedness among species, or the mean number of neighbors) and decrease modularity (network subdivsion into modules; Cao et al. 2018) with potentially negative consequences for network resilience and susceptibility to disturbance. Nutrient balance effects (e.g. N-limitation versus P-limitation) on network parameters are much less studied, yet modularity and connectance in heterotrophic communities were demonstrated to respond to nutrient ratios (Larsen et al. 2019).

Notably, our knowledge of nutrient effects on the topology of co-occurrence networks comes primarily from terrestrial systems as research in aquatic environments (Hu et al. 2017, Qu et al. 2019, Chen et al. 2021) is just beginning. This is concerning as aquatic systems are disproportionately more susceptible to human influences because of their relative isolation within the surrounding landscape matrix (Woodward et al. 2010). Furthermore, it is not understood to what extent network topology in aquatic systems is driven by shared niches, dispersal or actual interspecific interactions, given that all of these factors could underlie species co-occurrence relationships (Morueta-Holme et al. 2016).

We provide the first study of how species co-occurrence networks respond to each of four nutrient contexts (oligotrophic versus eutrophic and N-limited versus P-limited) at a subcontinental scale, while controlling for climate and dispersal. We used as a model system benthic algae because their metacommunities are sensitive to nutrient, climatic and dispersal effects (Soininen 2007, Verleyen et al. 2009, Soininen et al. 2016, Leboucher et al. 2019) but it still remains unknown whether this sensitivity transcends to the level of to the level of co-occurrence networks. Given the backdrop of ongoing climate change, it is also poorly understood if climate effects on meta-community composition and network topology are enhanced or diminished by nutrient levels.

The goal of this investigation was therefore to determine how nutrient supply and imbalance affect properties of co-occurrence networks in stream benthic algae, growing under oligotrophic (low N and P concentrations), eutrophic (high N and P concentrations) and N- or P-limited conditions (intermediate N and P concentrations but Redfield ratio below or above 16:1, respectively). We focused on network size, i.e. number of nodes (species) and edges (species correlations), edge characteristics (mean proportion of positive edges), connectivity (mean shortest path length and connectance), clustering (local and global) and modularity. Edges represent spatial associations between species. Edges can be positive or negative in value and their relative proportions have implications for network stability (Mougi and Kondoh 2012, Suweis et al. 2014). Connectance among nodes reflects the fraction of possible edges that are realized in the network and is often examined as an indicator of the robustness of ecological associations (Dunne et al. 2002, Estrada 2007, Borrett et al. 2010). Clustering quantifies the tendency of nodes to form completely connected, distinct neighborhoods (local or global) (Shirley and Rushton 2005) and can influence network susceptibility to disturbance and propagation of effects (Watts and Strogatz 1998). Modularity reflects the overall subdivision of the network into discrete groups (modules), having the potential to represent functional groups or niche partitioning (Montoya et al. 2015). We thus pursued the following two objectives.

Our first objective was to quantify network topology under each nutrient context and test for differences. In addressing this objective, we test the hypothesis that nutrient contexts (supply and ratio) influence co-occurrence network topology. Greater algal biodiversity with nutrient addition (Hillebrand et al. 2007, Passy and Larson 2019) should translate into overall larger co-occurrence networks with more nodes and edges. Eutrophic conditions promote speciose overstory guilds (high profile and motile) that are sensitive to nutrient limitation, whereas understory low profile species dominate under nutrient limitation (Passy 2007, Marcel et al. 2017, Wu et al. 2017). Eutrophication is, therefore, likely to produce networks with greater local clustering, modularity and path lengths due to coexistence of different functional groups, while oligotrophic conditions may lead to greater network connectivity and global clustering among fewer and comparatively functionally uniform species. Algal functional responses to N:P ratios are varied and system-specific, making it difficult to predict how network topology may respond to N:P ratios, but some studies have reported a tendency of P-limitation to stimulate high profile diatoms, whereas N-limitation to favor cyanobacteria and/or motile diatoms (Smith 1983, Stelzer and Lamberti 2001, Qu et al. 2019).

Our second objective was to determine the relative contributions of climate and space (surrogate of dispersal) to network topology versus meta-community composition and whether these factors differ between nutrient contexts. In

addressing this objective, we used a null model (Morueta-Holme et al. 2016) to test hypotheses that gradients in climate, dispersal or both structure metacommunity composition, which in turn constrains topologies of algal co-occurrence networks. In aquatic systems, climate drives local nutrient availability and ratio by affecting precipitation runoff and flows (Jeppesen et al. 2009, Özen et al. 2010), and metabolic requirements (Woodward et al. 2010). Climate–nutrient relationships in turn impact competitive and facilitative mechanisms in algae (Bestion et al. 2018, Marañón et al. 2018), which can have consequences for algal co-occurrence patterns, although this topic has not been studied with network methods. At the spatial scale of this study, we expected climatic factors to substantially constrain network topology across nutrient contexts due to their pronounced influence on algal species composition and distribution (Pajunen et al. 2016, Jyrkänkallio-Mikkola et al. 2017). However, stochastic mechanisms, which dominate in productive environments (Steiner and Leibold 2004, Chase 2010, Leboucher et al. 2019), could more strongly control eutrophic than oligotrophic network topology, but this question has not been explored so far. It is also unknown if or how networks differing in nutrient ratio would respond to climatic and spatial control. We therefore examined network topology as a function of climatic and spatial factors across nutrient contexts and tested whether these factors exercised a similar effect on metacommunity structure.

## Methods

### Datasets

We used data from 840 stream sites, where samples were taken from May to September between 1993 and 2015 by the National Water-Quality Assessment Program of the US Geological Survey. Site coordinates (in degrees latitudeand longitude) were reprojected into Cartesian coordinates with a Robinson projection (EPSG: 4326) for spatial analyses using R package 'rgdal' (Bivand et al. 2020). Algae were collected from a defined area in the richest-targeted habitats, which encompass hard substrates or macrophytes in faster currents. Generally, 25 cobbles, 5 woody snags or 5 macrophyte beds were sampled within a stream reach and the material was composited into a single sample. All taxa were identified mainly to species and their densities were measured as cells per $cm^2$. Species were categorized by ecological guild, following Passy (2007), Passy and Larson (2011) and Rimet and Bouchez (2012), including low profile (growing close to the substratum), high profile (extending into the biofilm matrix), motile (fast moving) and planktonic (sedimented from the water column).

Each sample was classified based on nutrient supply (eutrophic or oligotrophic) and ratio (N-limited or P-limited). Total nitrogen (TN) and total phosphorus (TP) concentrations were below 0.669 and 0.025 mg $l^{-1}$, respectively, in oligotrophic samples but above 1.499 and 0.075 mg $l^{-1}$, respectively,

in eutrophic samples (Dodds and Smith 2016). Nitrogen-limited samples were identified as those with Redfield N:P ratios less than 16:1, whereas P-limited samples had ratios above 16:1 (Redfield 1934). Nitrogen and phosphorus-limited sites had nutrient conditions that were either mesotrophic for both nutrients, or had one nutrient at oligotrophic or mesotrophic level and the other nutrient at eutrophic level. Only sites with pH $\geq$ 6.5 were considered to remove potentially confounding effects of high acidity. Using our classification and selection criteria, we identified 942 sites that could be used for analyses, however, they produced different sample sizes and different patterns of spatial clustering across nutrient categories. Thus, to equalize the number of samples and break up the geographic clustering, we randomly selected and removed sites from the most densely sampled areas. Our final datasets thus consisted of 193 oligotrophic, 193 eutrophic, 227 N-limited and 227 P-limited sites (Supporting information). All sites were unique to each nutrient context. Climatic predictors (bioclim variables 1–19, Fick and Hijmans 2017) were generated for each site.

### Co-occurrence networks

Prior to network construction (Fig. 1), we filtered the species by site metacommunity matrix for each nutrient context by removing rare species, which occurred in less than 10% of samples. We created weighted co-occurrence networks for each nutrient context by calculating a partial Spearman correlation matrix on relative species density, standardized with mean = 0 and standard deviation = 1. Partial Spearman correlations were used over raw correlations to account for the influence of indirect species associations and because the null model we employed requires their use (see Gradient effects
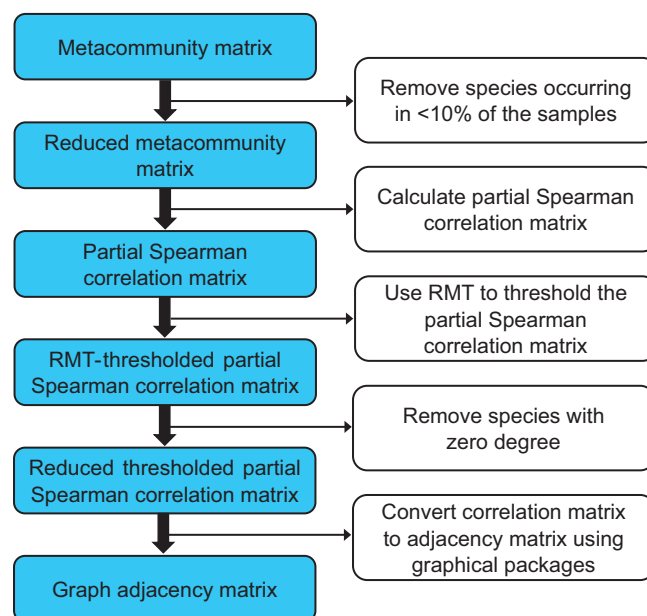


Figure 1. Flow chart showing the construction of nutrient supply and ratio networks.

section). We used a modified random matrix theory (RMT) method from R-package 'RMThreshold' (Menzel 2016) to objectively determine thresholds below which correlations were likely spurious and should be removed from further analyses (Supporting information).

After thresholding the partial correlation matrix with the RMT-selected correlation value, species with all correlations below the threshold were removed. The resulting correlation adjacency matrix was used to generate both a weighted and an unweighted network using R package 'igraph' (Csardi and Nepusz 2019). In these networks, nodes represented individual species, and the edges between the nodes represented species correlations. Edges between nodes within the unweighted network were binary, with 1 if the correlation was above the RMT-defined threshold and 0 (i.e. absent) if the correlation was below this threshold. In the weighted network, the edges represented the actual species correlations (again, only those above the RMT-defined threshold). It was created only for tracking edge sign in the calculation of proportion of positive edges, while the unweighted network was used for the remaining parameter calculations and the module clustering function. We used fast-greedy clustering function provided by this package with default settings (function: 'cluster_fast_greedy') to identify modules in the network. After using this function to assign nodes into their modules, we calculated network topology measures, including number of nodes, number of edges, mean node degree, mean shortest path length, clustering coefficients (a.k.a. transitivities), modularity and number of modules. Mean shortest path length, measuring on average how many nodes lie between any two nodes in the network, was calculated with a harmonic mean.

## Gradient effects on co-occurrence networks

To explore whether species co-occurrence networks are structured by niche overlap along climatic gradients (climate), small scale environmental filtering and dispersal (space), or both climate and space, we used the null model outlined by Morueta-Holme et al. (2016). We implemented associated functions from the R-package 'netassoc' (Blonder and Morueta-Holme 2017), which we briefly describe (Supporting information). This procedure compares observed partial correlation values with a distribution of predicted partial correlation values and removes observed partial correlations that fall between the distribution extremes of the predicted. The predicted correlation distributions are produced by calculating partial correlations on a predicted metacommunity matrix that contains the same species as the observed species-by-site matrix, but with densities predicted using a species distribution model (SDM) framework. The SDM-predicted densities represent those expected, given climatic, spatial or both gradients. The null model procedure then samples individuals across sites using a weighted lottery model which constrains species' entries into sites proportional to their overall densities in the observed matrix. We then calculated a partial correlation matrix on the relative densities after the randomization. The observed partial correlation value is then statistically tested using the

generated null partial correlation values by counting the frequency of occurrences where the observed partial correlation was less than the predicted value, dividing the count by the number of randomizations (here 1000), and performing a Benjamani–Hochberg p-value adjustment. If significant, the observed value is retained in the matrix or otherwise deleted as it represents a shared response to a gradient.

Here, we modified the null model only to use partial Spearman correlation values rather than Pearson values to account for non-linear but monotonic correlations. We used the partial correlation matrix created by the RMT procedure as the observed correlation matrix. The null model procedure outlined by Morueta-Holme et al. (2016) can accommodate predicted abundances generated from a variety of SDM frameworks. For this study, densities for each species, predicted by climate or spatial variables, or both, were calculated separately using optimally-parameterized boosted regression trees (BRT) with a Bernoulli link-function and trained using 10-fold cross-validation (Elith et al. 2008). Spatial variables were generated from Robinson-projected site coordinates using Moran's eigenvector maps (MEM) with functions in R package 'adespatial' (Dray et al. 2016). MEMs are orthogonal variables whose values represent spatial autocorrelation patterns. They are generated as the eigenvectors of a spatial weighting matrix describing the strength of the spatial relationships between pairs of samples. Using function 'listw. candidates', we used all available combinations of spatial weighting matrix (except d-nearest neighbors) and spatial weighting definitions (concave up and concave down fitting functions set to default values of 0.5 and 5 respectively). We then used 'listw.select', which uses forward selection (global significance test set to 30 000 permutations) to find the optimal spatial weighting matrix, weight definitions and subset of MEM variables that best explained species presence/absence patterns (by adjusted $R^2$) (Bauman et al. 2018). Only MEM variables modelling positive autocorrelation patterns were retained.

BRTs for each species were optimized by iteratively finding the combination of tuning parameters (bag fractions ranging from 0.5 to 0.75, tree complexity parameters ranging from 1 to 10 and numbers of trees fit), which maximized area under the curve (AUC) values. Then, we refitted the models using the parameters corresponding to maximum AUC. For the BRT procedure with the combined climate and spatial predictors, we used all climatic and spatial predictors. The BRTs produced probability estimates owing to the specified Bernoulli link function, which we transformed to densities by multiplying each species' observed density by the species' probability for the sample. When a species' probabilities could not be predicted with any combination of tuning parameters, we used its observed abundances in the predicted matrix because the model fits were not better than a random expectation. We then combined the predicted abundances for each species into a single predicted metacommunity matrix and subsequently used it in the null model. We set the null model to run for 1000 iterations, transforming the randomized predicted densities to relative densities

for each iteration as described for network construction, and identified non-significant correlations (p > 0.05), which were then removed from the observed partial correlation matrix. We reconstructed the networks and calculated network parameters using these newly thresholded correlation matrices, which represented association networks free of co-occurrences due to climatic niche overlap and/or environmental filtering and dispersal. An important limitation of the lottery model is that the null model may identify and eliminate links between species for which we used their observed abundances. This could result in overestimating the number of climate- and/or space-influenced species associations. To eliminate this problem, we replaced the links between these species if they were removed during the null model analysis. After the null model analysis, we qualitatively assessed how climatic, spatial and climatic + spatial factors influenced overall network topologies when compared with the original networks by calculating the absolute proportional change in each measured network parameter.

## Statistical analyses

We tested our first hypothesis by using a randomization protocol to examine whether the ecological networks differed in topology between nutrient contexts. We note the oligotrophic and eutrophic data consisted of samples that were either N- or P-limited, therefore we could not make fair pairwise comparisons of these networks with the N- and P-limited networks. We thus separated our statistical comparisons and assessed differences between oligotrophic versus eutrophic networks and N-limited versus P-limited networks. For both pairs of stream networks, we pooled their sites together and then randomly reassigned them to either nutrient category. We then calculated all network parameters, recorded the difference for each network parameter value, and repeated the randomization 999 times. We then compared the observed difference in parameter values between the nutrient networks to their respective randomized mean difference and determined significant differences if the observed difference was more extreme than the 2.5 or the 97.5 percentile of the randomized difference values.

We used redundancy analysis to calculate the variance in Hellinger-transformed species relative abundances explained by climatic, spatial and climatic and spatial variables for each nutrient context. We then followed the analysis with variation partitioning. Climate variables were centered for these analyses, and to account for potential quadratic relationships, we incorporated the linear and quadratic forms of the centered climate variables. We then used the forward selection procedure provided in R package 'adespatial' ('forward.sel', Dray et al. 2016) to retain all significant climate predictors. Spatial factors were created using MEM, as described above for the BRT procedure, with significant MEMs retained using forward selection. The pure fractions of climate and space were tested using permutational analysis of variance (R-package 'vegan' function 'anova.rda', Oksanen et al. 2019). The variance in species composition explained by climatic, spatial and climatic and spatial variables was compared with the magnitude of the respective network controls to assess objective 2.

## Results

All four networks varied in size, ranging from 73 nodes in the oligotrophic network to 97 nodes in the eutrophic network, and from 86 nodes for the N-limited to 96 nodes in the P-limited network (Supporting information, Table 1). Examination of the edge counts and connectance revealed that all networks were sparse and poorly connected (Table 1). Positive network edges generally constituted the majority of edges in the networks, although their proportion was the highest in the P-limited network. All networks consisted primarily of diatoms (86–88% of nodes) with minor contribution by cyanobacteria and/or green algae depending on the nutrient context (4–8%). Altogether, the remaining algal groups (red algae, euglenoids and unclassified algal taxa) generally comprised < 4% of the nodes. With respect to guilds, motile taxa dominated all networks (43–47% of all classified nodes), except for the oligotrophic network, where high profile taxa prevailed instead (45% of all classified nodes). Planktonic taxa generally represented the lowest fraction of nodes (2–11%).

### Topological comparisons across contexts

The randomization tests indicated that overall network topologies generally did not differ significantly between nutrient

Table 1. Parameters of raw algal co-occurrence networks differing in nutrient availability (oligotrophic versus eutrophic) and nutrient ratio (N-limited versus P-limited).

| Parameter | Nutrient supply networks | | Nutrient ratio networks | |
| --- | --- | --- | --- | --- |
| | Oligotrophic | Eutrophic | N-limited | P-limited |
| Number of nodes | 73 | 97 | 86 | 96 |
| Number of edges | 108 | 159 | 108 | 119 |
| % positive edges | 0.65 | 0.61 | 0.64 | 0.66 |
| Local clustering | 0.10 | 0.12 | 0.07 | 0.08 |
| Global clustering | 0.11 | 0.10 | 0.11 | 0.09 |
| Modularity | 0.61 | 0.54 | 0.67 | 0.69 |
| Number of modules | 8 | 7 | 10 | 12 |
| Mean path length | 3.42 | 3.30 | 3.87 | 4.33 |
| Mean degree | 2.96 | 3.28 | 2.51 | 2.48 |
| Connectance | 0.04 | 0.03 | 0.03 | 0.03 |

contexts (Table 1). The only clear exception was network size (i.e. number of nodes), which was significantly greater (p < 0.001) in the eutrophic network compared to the oligotrophic network. Although edge counts were noticeably greater in the eutrophic versus the oligotrophic network, this difference was non-significant after randomization.

**Null model analysis and variance partitioning**

Qualitative examination of the networks produced by the null model procedure indicated that controlling for climatic, spatial and combined climatic and spatial effects changed overall network topology with mean absolute percent change in network parameters ranging between 46–67% on average (Fig. 2). Overall, the three controls had comparable effects on network topology, except for the P-limited networks, where climate and climate + space exerted the strongest effect. However, differences in overall topologies between controls within nutrient contexts were generally small.

General patterns were observed in some parameters after climate and spatial control, owing in part to increases in 0-degree nodes (i.e. nodes without any connections, Fig. 3, 4). Node counts, edge counts, connectance and mean node degree decreased in a correlated fashion across all controlled networks between 41% and 56%, whereas modularity increased by 21–35% and number of modules drastically increased between 225% and 330%. Notably, the effects of control depended on nutrient context, including number of modules, clustering and mean path lengths. Specifically, number of modules increased more strongly after control in the nutrient ratio networks compared with the supply networks, ranging as low as 266% in the controlled P-limited network to as high 330% within the N-limited controlled

networks. Local and global clustering decreased sharply in the eutrophic network, particularly when space was controlled for (individually or in conjunction with climate) but also showed highly variable patterns in response to controls in the other networks. Responses in mean path length to controlling variables were generally negative across networks, especially in the P-limited network, but became positive in the eutrophic network. Finally, changes in the fractions of positive edges showed no clear general trends with respect to any controlling variable.

Climatic, spatial and climate + spatial variables significantly explained community variation (Fig. 5a), with the majority of variance explained by spatial, followed by climatic + spatial variables. Variance partitioning analysis on the metacommunities revealed that climatic and spatial variables generally explained significant fractions of algal relative abundances, with spatial variables predominating regardless of nutrient context (Fig. 5b). When examining the metacommunity composed of species only found in the networks, we observed that climate + spatial variables explained slightly more variance than spatial variables alone across all network contexts (< 3%, results not shown), however the results for the variance partitioning did not change. The analysis further revealed that pure climatic effects generally explained between 2 and 5% of the total variability in composition in contrast to pure spatial factors, which explained usually between 9% and 19% of community variation. Covariance effects were also noticeably strong and explained between 8 and 15% of community variation, with the covariance fraction at times explaining more variation in meta-community composition than the spatial fraction (eutrophic and N-limited networks).

## Discussion

The dependence of algal composition and biomass production on nutrient supply and balance has been studied for decades and is generally well recognized (Hecky and Kilham 1988, Borchardt 1996, Bergström 2010, Hillebrand and Lehmpfuhl 2011, Hayes et al. 2015). Here we examined for the first time how the topology of algal co-occurrence networks is structured by nutrient context, consistent with our first objective. Networks differed significantly only in size, i.e. number of nodes, and only between nutrient supplies but not between nutrient ratios. Following objective 2, we found strong support for the prediction that climatic and spatial gradients constrain the topology of the networks and that the magnitude of this influence depends on nutrient context. We further elucidated the mechanisms that control metacommunity composition versus network topology.

Nutrient supply context influenced the size of co-occurrence networks but not their complexity and modular structure. Nutrient ratio, on the other hand, did not have any effect on network topology. The difference in size between the nutrient supply networks could be attributed to greater diversity of the most speciose guild, the motile guild, in the eutrophic network. The guild distribution was very similar
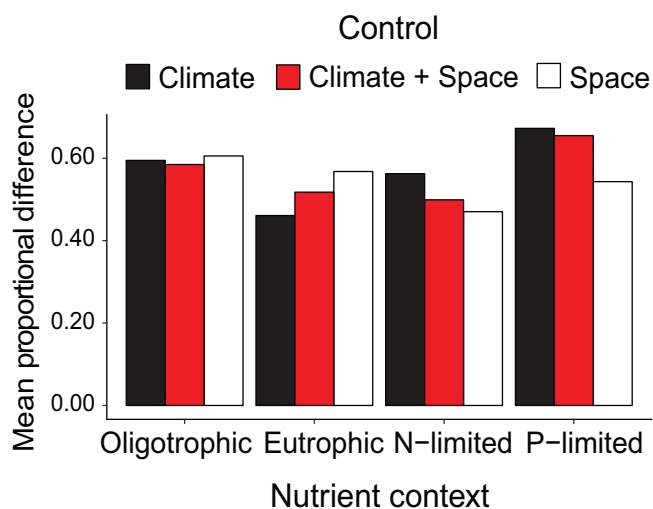


Figure 2. Mean proportional change in network topology from the original network across all network parameters after controlling for climatic, spatial and climatic + spatial variables. For each controlled network, the absolute value of the proportional change in each network parameter was calculated and then averaged across all values to obtain the mean proportional change.
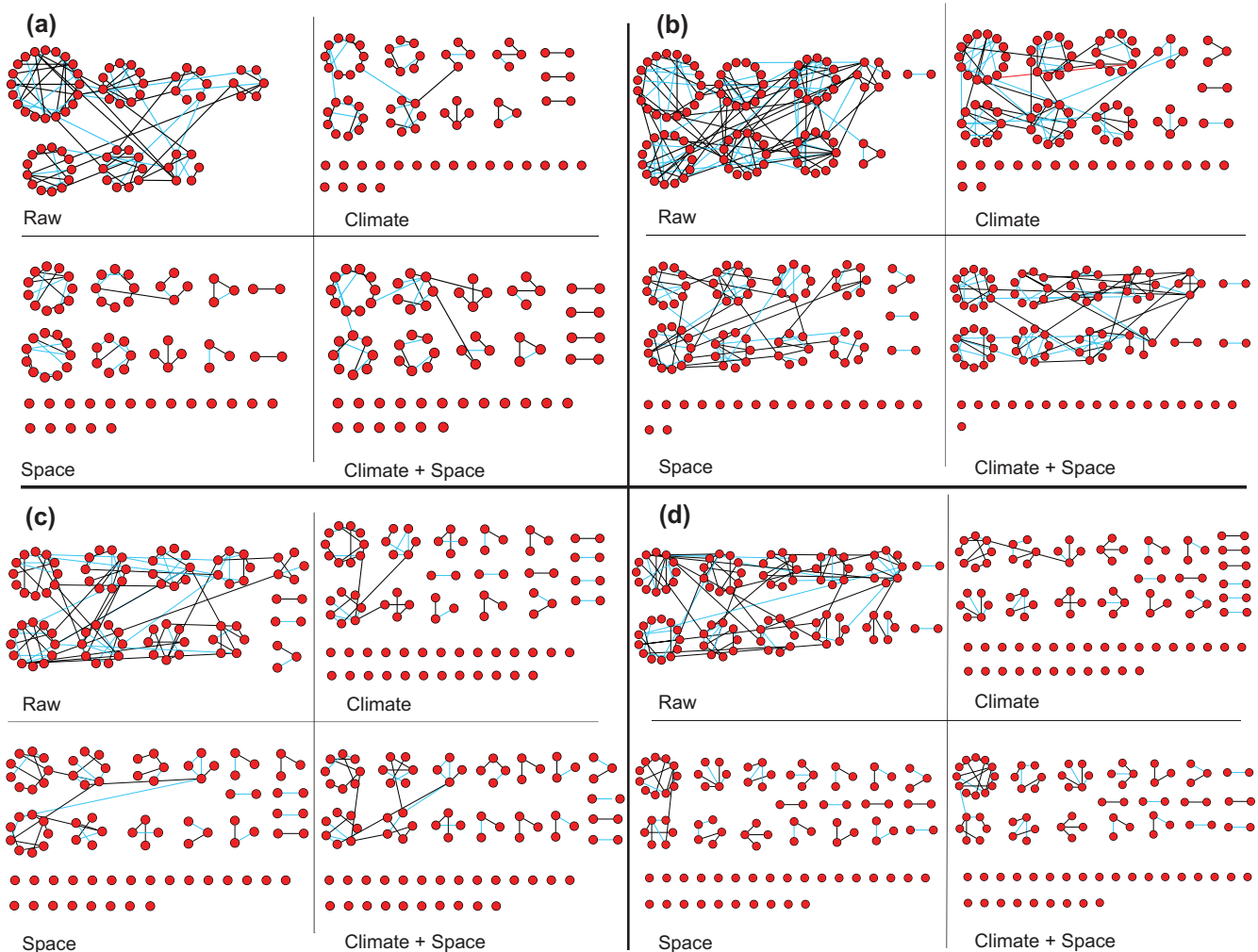
Figure 3. Raw networks and networks reconstructed after controlling for climate, space and Climate + Space in (a) oligotrophic, (b) eutrophic, (c) N-limited and (d) P-limited streams. Black lines = positive correlations, blue lines = negative correlations.

between the nutrient ratio networks, which could explain their network size similarity. Therefore, we found weak support for our first hypothesis, predicting that nutrient context would constrain network topology. A potential explanation is that topological responses to environmental factors may not manifest themselves at the sub-continental scale of this study due to low network connectance. Much work has shown that network topologies are strongly driven by the presence of highly connected nodes (Proulx et al. 2005, Poisot et al. 2015), which were generally absent in our networks. As a result, these networks were poorly connected with connectance values of 0.03–0.04, falling below the commonly reported values of 0.05–0.30 (Thompson et al. 2012). Consequently, differences between networks, originating from loss or gain of poorly connected nodes, had little consequence for network topology. An interesting avenue of future research will be to determine at what scale network topology of algae becomes sensitive to nutrient inputs.

Previous observations for mutualistic pollinator networks showed that greater diversity in networks may present with greater node clustering (Gómez et al. 2011). We found this not to be the case for the co-occurrence networks examined here as the differences in clustering between the larger eutrophic network and the smaller oligotrophic network were not significant. Ecologically, clustering in networks reflects the presence of redundant pathways (Karimi et al. 2017) and stronger organization of nodes into distinct subgroups (Girvan and Newman 2002). Much of our knowledge on the environmental dependence of clustering comes from research on soil microbiota, where clustering decreased with $CO_2$ concentrations (Zhou et al. 2011, Sauvadet et al. 2016) but showed no response to anthropogenic disturbance (Zappelini et al. 2015). We add to this knowledge by finding that the clustering patterns of stream co-occurrence networks were relatively robust to nutrient supply and ratio, at least at the scale of this study.

All of our networks were dominated by positive associations, as predicted for larger scales by multiple studies. Specific explanations include better detection of positive biotic interactions (Araújo and Rozenfeld 2014), increased
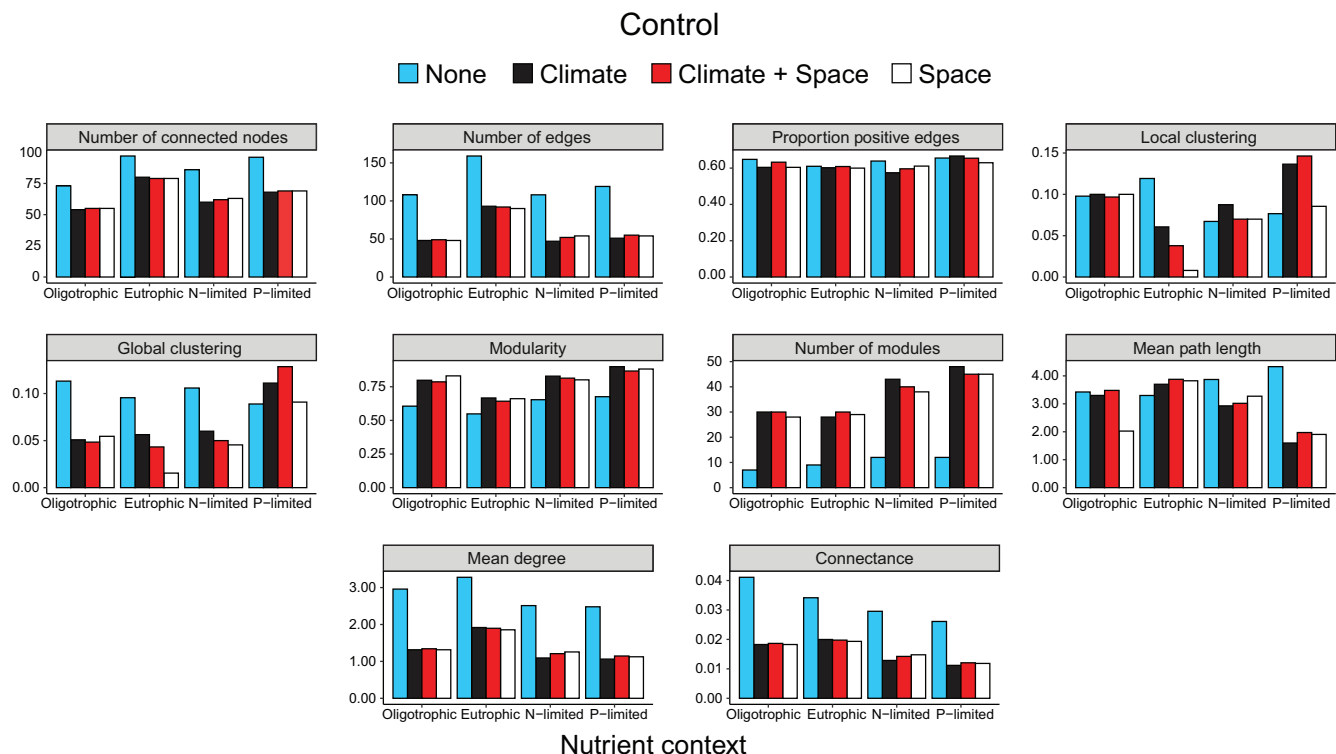
Figure 4. Parameter values across nutrient contexts in the raw networks versus networks controlled for climate, space and climate + space.

number of marginal environments that support aggregation of common and geographically broad species (Bar-Massada and Belmaker 2017), and stronger dispersal limitation, resulting in co-occurrences that reflect shared dispersal barriers (D'Amen et al. 2018). Proportions of positive connections in ecological interaction networks have been linked to stability (Mougi and Kondoh 2012, Lurgi et al. 2016, García-Callejas et al. 2018). However, we did not find evidence that positive co-occurrences may impart stability. Specifically, the proportion positive edges did not vary conspicuously across networks, but the network response to controls (described below) was widely variable. We thus conclude that positive co-occurrences do not necessarily relate to network topological stability, but further work in this area is necessary to determine the generality of this pattern and whether it is scale dependent.

With respect to objective 2, accounting for environmental (here climatic) and/or spatial effects strongly altered network topology, which supports results reported for North American tree communities (Morueta-Holme et al. 2016). In contrast to our findings for the raw networks, the controlled networks exhibited topological differences, which were qualitatively examined. It has long been argued that co-occurrence relationships contain much extraneous information (e.g. shared ecological responses and dispersal effects) that should be controlled for to identify relevant biological associations and/or interactions (Berlow 1999, Bairey et al. 2016, Blanchet et al. 2020). Our results supported this argument and revealed that after controlling for climate and/or space, the four networks

exhibited 46–67% change on average across all parameters. The eutrophic network retained 88–89% of their connected nodes and 56–60% of their edges, while the other three networks showed more drastic changes and retained 72–76% of their connected nodes and 40–50% of their edges. Plausibly, these remaining edge percentages could be attributed in part to biotic interactions, which may be most pronounced in the most diverse eutrophic network. The removal of large fractions of connected nodes and edges in our networks by the null model has important ramifications for co-occurrence network analyses because many network properties are highly sensitive to network size and sparseness of edges (Poisot and Gravel 2014). Our findings also highlight an opportunity worth exploring, as null model methods can reveal whether other factors, such as disturbance or herbivory, underlie algal co-occurrence patterns. There is a caveat with respect to the use of a single SDM framework because SDM predictions are sensitive to model choice (Araújo and Guisan 2006). Realistically, different machine learning methods could predict different densities, which in turn could produce different network topologies even with the same model predictors. We selected BRTs because of their good predictive performance and successful application in algal macroecology (Pajunen et al. 2016, Pound et al. 2021). Examining networks derived from different SDMs was beyond the scope of this paper but we recommend that future studies explore this topic.

On average, climate, space and climate + space imposed similar constraints on network topology. As predicted,
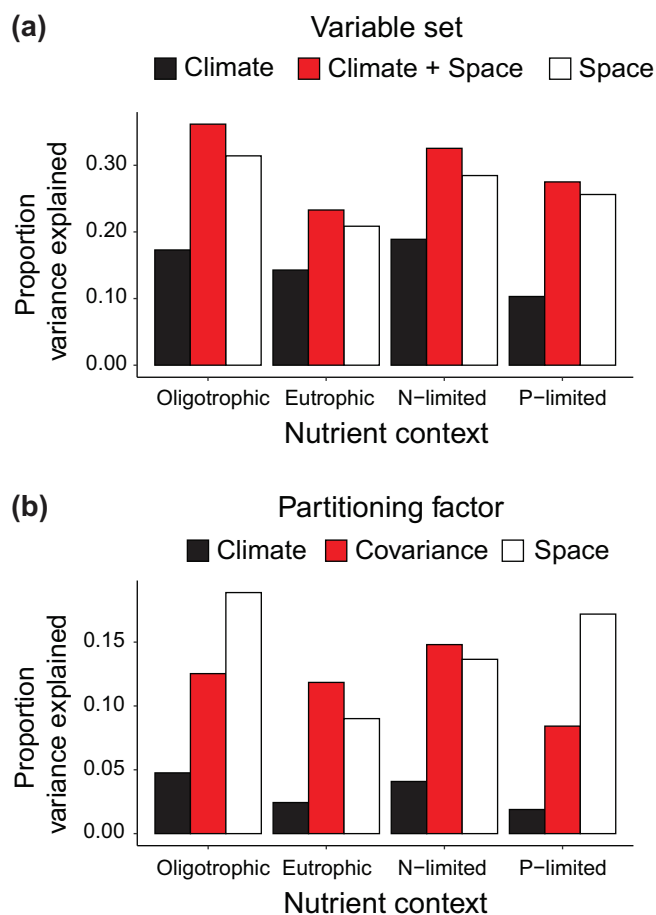
**(a)** Variable set

**(b)** Partitioning factor

Figure 5. (a) Adjusted $R^2$ obtained from redundancy analysis for fits of meta-community composition data by climatic, spatial and climatic + spatial variables. All model fits were significant as tested with permutational methods ($p < 0.01$). (b) Variance partitioning of algal metacommunity composition into fractions (as adjusted $R^2$) explained by pure climatic, pure spatial and covariance effects across nutrient contexts. The pure climatic and the pure spatial fraction within each nutrient context were significant sources of explained variance as confirmed by permutational methods.

climate strongly affected network topology across nutrient contexts. Space had a prominent role as well, but for some topological parameters, it had a stronger impact in the eutrophic than the oligotrophic context, as predicted, e.g. local and global clustering. However, for other parameters, space was a stronger driving force in the oligotrophic networks, e.g. mean path length. These important parameters describe at the network level whether a species influence is likely to be contained within groups (local clustering) or readily transmitted across groups (global clustering), and how quickly this may occur (mean path length). Therefore, our results indicate that dispersal may have pronounced but distinct effects on network parameters across nutrient contexts, which falls in line with conclusions made with similar metrics for biotic interaction networks under different dispersal contexts (Thompson and Gonzalez 2017). We thus recommend that

inferences about topological responses of co-occurrence networks to environmental contexts be made after appropriate controls are applied.

Across nutrient contexts, climatic and spatial factors had prominent effects on species composition and network properties. Prior reports as to whether ecological drivers of network topology and community composition coincide have been mixed, with some studies showing correspondence (Mokross et al. 2013), while others, divergence (Li et al. 2018). In this investigation, climatic and spatial variables had a substantial joint influence on metacommunity composition, consistent with other studies (Heino et al. 2012, Zorzal-Almeida et al. 2017). Additionally, for the most part, climate and space contributed to the network patterns together as much as they contributed alone. This indicated that spatially structured climate may be an important source of variation for both metacommunities and networks. However, there was also a notable difference – spatial factors outperformed climatic factors in explaining metacommunity composition, but generally exercised control over network topology similar to this of climate. This discrepancy could be explained with a scale mismatch in species versus metacommunity response. Individual species distributions, and by extension, pair-wise species co-occurrences in the networks, are more spatially restricted and thus may be strongly driven by climate, as shown for stream diatoms (Pajunen et al. 2016). Conversely, at the subcontinental scale of our metacommunities, geographic barriers among the discrete hydrologic systems that comprised their habitats, may have contributed to dispersal limitation subsuming some of the climatic effect. Although environmental factors predict better the variance in diatom metacommunities compared to space (Soininen and Teittinen 2019), these factors are generally associated with nutrient level (Soininen 2007), which in our case was accounted for by nutrient context. Thus, nutrient contexts apparently determined the contribution of dispersal processes to community variation or influenced spatially-structured climatic drivers. These results suggest that there are both similarities and differences in the drivers of network properties versus metacommunity composition, which to some extent depend on nutrient context.

Our study lays groundwork for connecting novel methodology with classic ecological research to generate new information explaining why species co-occur. At subcontinental scale, nutrient supply and ratio affected only weakly algal co-occurrence network properties but determined the magnitude of climatic and dispersal effects. The differential response of network topology and species composition to climate and dispersal calls for broader multi-level ecological approaches to better understand how the environment structures biological communities.

## Author contributions

**William Budnick**: Data curation (lead); Formal analysis (lead); Methodology (lead); Software (lead); Writing – original draft (lead); Writing – review and editing (lead). **Joseph Mruzek**: Data curation (supporting); Formal analysis (supporting); Methodology (supporting); Software (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Chad Larson**: Data curation (supporting); Formal analysis (supporting); Methodology (supporting); Software (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Sophia I. Passy**: Conceptualization (lead); Funding acquisition (lead); Investigation (lead); Project administration (lead); Resources (lead); Supervision (lead); Validation (lead); Writing – original draft (equal); Writing – review and editing (equal).

## Transparent Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1111/ecog.05307>.

## Data availability statement

Data available from Dryad Digital Repository: <https://dx.doi.org/10.5061/dryad.dv41ns1xw> (Budnick et al. 2021).

## References

Araújo, M. B. and Guisan, A. 2006. Five (or so) challenges for species distribution modelling. – J. Biogeogr. 33: 1677–1688.

Araújo, M. B. and Rozenfeld, A. 2014. The geographic scaling of biotic interactions. – Ecography 37: 406–415.

Bairey, E. et al. 2016. High-order species interactions shape ecosystem diversity. – Nat. Commun. 7: 12285.

Bar-Massada, A. and Belmaker, J. 2017. Non-stationarity in the co-occurrence patterns of species across environmental gradients. – J. Ecol. 105: 391–399.

Bauman, D. et al. 2018. Disentangling good from bad practices in the selection of spatial or phylogenetic eigenvectors. – Ecography 41: 1638–1649.

Bergström, A. K. 2010. The use of TN:TP and DIN:TP ratios as indicators for phytoplankton nutrient limitation in oligotrophic lakes affected by N deposition. – Aquat. Sci. 72: 277–281.

Berlow, E. L. 1999. Strong effects of weak interactions in ecological communities. – Nature 398: 330–334.

Bestion, E. et al. 2018. Nutrient limitation constrains thermal tolerance in freshwater phytoplankton. – Limnol. Oceanogr. Lett. 3: 436–443.

Bivand, R. et al. 2020. rgdal: bindings for the 'geospatial' data abstraction library. R package Ver. 1.5–16. – <https://CRAN.R-project.org/package=rgdal>

Blanchet, F. G. et al. 2020. Co-occurrence is not evidence of ecological interactions. – Ecol. Lett. 23: 1050–1063.

Blonder, B. and Morueta-Holme, N. 2017. netassoc: inference of species associations from co-occurrence data. R package Ver. 0.6.3. – <https://CRAN.R-project.org/package=netassoc>.

Borchardt, M. A. 1996. Nutrients. – In: Stevenson, R. J. et al. (eds), Algal ecology: freshwater benthic ecosystems. Academic Press, pp. 183–227.

Borrett, S. R. et al. 2010. Rapid development of indirect effects in ecological networks. – Oikos 119: 1136–1148.

Budnick, W. R. et al. 2021. Data from: The impacts of nutrient supply and imbalance on subcontinental co-occurrence networks and metacommunity composition of stream algae. – Dryad Digital Repository, <https://dx.doi.org/10.5061/dryad.dv41ns1xw>.

Cao, X. et al. 2018. Heterogeneity of interactions of microbial communities in regions of Taihu Lake with different nutrient loadings: a network analysis. – Sci. Rep. 8: 8890.

Cardinale, B. J. et al. 2009. Separating the influence of resource 'availability' from resource 'imbalance' on productivity–diversity relationships. – Ecol. Lett. 12: 475–487.

Chase, J. M. 2010. Stochastic community assembly causes higher biodiversity in more productive environments. – Science 328: 1388–1391.

Chen, J. et al. 2021. Sedimentary microeukaryotes reveal more dispersal limitation and form networks with less connectivity than planktonic microeukaryotes in a highly regulated river. – Freshwater Biol. 66: 826–841.

Cook, S. C. et al. 2018. Freshwater eutrophication drives sharp reductions in temporal beta diversity. – Ecology 99: 47–56.

Csardi, G. and Nepusz, T. 2019. igraph: network analysis and visualization. R package Ver. 1.2.6. – <https://CRAN.R-project.org/package=igraph>.

D'Amen, M. et al. 2018. Disentangling biotic interactions, environmental filters and dispersal limitation as drivers of species co-occurrence. – Ecography 41: 1233–1244.

Dodds, W. K. and Smith, V. H. 2016. Nitrogen, phosphorus and eutrophication in streams. – Inl. Waters 6: 155–164.

Dray, A. S. et al. 2016. adespatial: multivariate multiscale spatial analysis. R package Ver. 0.0-7. – <https://CRAN.R-project.org/package=adespatial>.

Dunne, J. A. et al. 2002. Network structure and biodiversity loss in food webs: robustness increases with connectance. – Ecol. Lett. 5: 558–567.

Elith, J. et al. 2008. A working guide to boosted regression trees. – J. Anim. Ecol. 77: 802–813.

Elser, J. J. et al. 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. – Ecol. Lett. 10: 1135–1142.

Estrada, E. 2007. Food webs robustness to biodiversity loss: the roles of connectance, expansibility and degree distribution. – J. Theor. Biol. 244: 296–307.

Fick, S. E. and Hijmans, R. J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. – Int. J. Climatol. 37: 4302–4315.

Flaim, G. et al. 2016. Effects of re-oligotrophication and climate change on lake thermal structure. – Freshwater Biol. 61: 1802–1814.

García-Callejas, D. et al. 2018. The effect of multiple biotic interaction types on species persistence. – Ecology 99: 2327–2337.

Girvan, M. and Newman, M. E. J. 2002. Community structure in social and biological networks. – Proc. Natl Acad. Sci. USA 99: 7821–7826.

Gómez, J. M. et al. 2011. The functional consequences of mutualistic network architecture. – PLoS One 6: e16143.

Harpole, W. S. et al. 2011. Nutrient co-limitation of primary producer communities. – Ecol. Lett. 14: 852–862.

Hayes, N. M. et al. 2015. Climate and land use interactively affect lake phytoplankton nutrient limitation status. – Ecology 96: 392–402.

Hecky, R. E. and Kilham, P. 1988. Nutrient limitation of phytoplankton in freshwater and marine environments: a review of recent evidence on the effects of enrichment. – Limnol. Oceanogr. 33: 796–822.

Heino, J. et al. 2012. Context dependency and metacommunity structuring in boreal headwater streams. – Oikos 121: 537–544.

Hillebrand, H. and Lehmpfuhl, V. 2011. Resource stoichiometry and consumers control the biodiversity–productivity relationship in pelagic metacommunities. – Am. Nat. 178: 171–181.

Hillebrand, H. et al. 2007. Consumer versus resource control of producer diversity depends on ecosystem type and producer community structure. – Proc. Natl Acad. Sci. USA 104: 10904–10909.

Hu, A. et al. 2017. Strong impact of anthropogenic contamination on the co-occurrence patterns of a riverine microbial community. – Environ. Microbiol. 19: 4993–5009.

Jeppesen, E. et al. 2009. Climate change effects on runoff, catchment phosphorus loading and lake ecological state, and potential adaptations. – J. Environ. Qual. 38: 1930–1941.

Jyrkänkallio-Mikkola, J. et al. 2017. Disentangling multi-scale environmental effects on stream microbial communities. – J. Biogeogr. 44: 1512–1523.

Karimi, B. et al. 2017. Microbial diversity and ecological networks as indicators of environmental quality. – Environ. Chem. Lett. 15: 265–281.

Larsen, M. L. et al. 2019. Nutrient stoichiometry shapes microbial coevolution. – Ecol. Lett. 22: 1009–1018.

Leboucher, T. et al. 2019. Diatom β-diversity in streams increases with spatial scale and decreases with nutrient enrichment across regional to sub-continental scales. – J. Biogeogr. 46: 734–744.

Lewandowska, A. M. et al. 2016. The influence of balanced and imbalanced resource supply on biodiversity–functioning relationship across ecosystems. – Phil. Trans. R. Soc. B 371: 20150283.

Li, D. et al. 2018. Homogenization of species composition and species association networks are decoupled. – Global Ecol. Biogeogr. 27: 1481–1491.

Lurgi, M. et al. 2016. The effects of space and diversity of interaction types on the stability of complex ecological networks. – Theor. Ecol. 9: 3–13.

Marañón, E. et al. 2018. Nutrient limitation suppresses the temperature dependence of phytoplankton metabolic rates. – ISME J. 12: 1836–1845.

Marcel, R. et al. 2017. Modelling diatom life forms and ecological guilds for river biomonitoring. – Knowl. Manage. Aquat. Ecosyst. 418: 1.

Menzel, U. 2016. 'RMThreshold'. R package Ver. 1.1. – <https://CRAN.R-project.org/package=RMThreshold>.

Mokross, K. et al. 2013. Decay of interspecific avian flock networks along a disturbance gradient in Amazonia. – Proc. R. Soc. B 281: 20132599.

Montoya, D. et al. 2015. Functional group diversity increases with modularity in complex food webs. – Nat. Commun. 6: 1–9.

Morueta-Holme, N. et al. 2016. A network approach for inferring species associations from co-occurrence data. – Ecography 39: 1139–1150.

Mougi, A. and Kondoh, M. 2012. Diversity of interaction types and ecological community stability. – Science 337: 349–351.

Oksanen, J. et al. 2019. vegan: community ecology package. R package Ver. 2.5-4. – <https://CRAN.R-project.org/package=vegan>.

Özen, A. et al. 2010. Drought-induced changes in nutrient concentrations and retention in two shallow Mediterranean lakes subjected to different degrees of management. – Hydrobiologia 646: 61–72.

Pajunen, V. et al. 2016. Climate is an important driver for stream diatom distributions. – Global Ecol. Biogeogr. 25: 198–206.

Passy, S. I. 2007. Diatom ecological guilds display distinct and predictable behavior along nutrient and disturbance gradients in running waters. – Aquat. Bot. 86: 171–178.

Passy, S. I. and Larson, C. A. 2011. Succession in stream biofilms is an environmentally driven gradient of stress tolerance. – Microb. Ecol. 62: 414–424.

Passy, S. I. and Larson, C. A. 2019. Niche dimensionality and herbivory control stream algal biomass via shifts in guild composition, richness and evenness. – Ecology 100: e02831.

Poisot, T. and Gravel, D. 2014. When is an ecological network complex? Connectance drives degree distribution and emerging network properties. – PeerJ 2: e251.

Poisot, T. et al. 2015. Beyond species: why ecological interaction networks vary through space and time. – Oikos 124: 243–251.

Pound, K. L. et al. 2021. Current distributions and future climate-driven changes in diatoms, insects and fish in US streams. – Global Ecol. Biogeogr. 30: 63–78.

Proulx, S. R. et al. 2005. Network thinking in ecology and evolution. – Trends Ecol. Evol. 20: 345–353.

Qu, X. et al. 2019. Networks and ordination analyses reveal the stream community structures of fish, macroinvertebrate and benthic algae, and their responses to nutrient enrichment. – Ecol. Indic. 101: 501–511.

Redfield, A. C. 1934. On the proportions of organic derivatives in sea water and their relation to the composition of plankton. – In: James Johnnstone Memorial Volume. Univ. Press of Liverpool, pp. 176–192.

Rimet, F. and Bouchez, A. 2012. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. – Knowl. Manage. Aquat. Ecosyst. 406: 1.

Sauvadet, M. et al. 2016. The dynamics of soil micro-food web structure and functions vary according to litter quality. – Soil Biol. Biochem. 95: 262–274.

Shirley, M. D. F. and Rushton, S. P. 2005. The impacts of network topology on disease spread. – Ecol. Complex. 2: 287–299.

Smith, V. H. 1983. Low nitrogen to phosphorus ratios favor dominance by blue-green algae in lake phytoplankton. – Science 221: 669–671.

Soininen, J. 2007. Environmental and spatial control of freshwater diatoms – a review. – Diatom Res. 22: 473–490.

Soininen, J. and Teittinen, A. 2019. Fifteen important questions in the spatial ecology of diatoms. – Freshwater Biol. 64: 2071–2083.

Soininen, J. et al. 2016. Global patterns and drivers of species and trait composition in diatoms. – Global Ecol. Biogeogr. 25: 940–950.

Steiner, C. F. and Leibold, M. A. 2004. Cyclic assembly trajectories and scale-dependent productivity–diversity relationships. – Ecology 85: 107–113.

Stelzer, R. S. and Lamberti, G. A. 2001. Effects of N:P ratio and total nutrient concentration on stream periphyton community structure, biomass and elemental composition. – Limnol. Oceanogr. 46: 356–367.

Stoddard, J. L. et al. 2016. Continental-scale increase in lake and stream phosphorus: are oligotrophic systems disappearing in the United States? – Environ. Sci. Technol. 50: 3409–3415.

Suweis, S. et al. 2014. Disentangling the effect of hybrid interactions and of the constant effort hypothesis on ecological community stability. – Oikos 123: 525–532.

Thompson, P. L. and Gonzalez, A. 2017. Dispersal governs the reorganization of ecological networks under environmental change. – Nat. Ecol. Evol. 1: 0162.

Thompson, R. M. et al. 2012. Freshwater food webs: towards a more fundamental understanding of biodiversity and community dynamics. – Freshwater Biol. 57: 1329–1341.

Tilman, D. and Isbell, F. 2015. Biodiversity: recovery as nitrogen declines. – Nature 528: 336.

Verbeek, L. et al. 2018. Warming and oligotrophication cause shifts in freshwater phytoplankton communities. – Global Change Biol. 24: 4532–4543.

Verleyen, E. et al. 2009. The importance of dispersal related and local factors in shaping the taxonomic structure of diatom metacommunities. – Oikos 118: 1239–1249.

Wang, J. et al. 2016. Nutrient enrichment modifies temperature–biodiversity relationships in large-scale field experiments. – Nat. Commun. 7: 13960.

Watts, D. J. and Strogatz, S. H. 1998. Collective dynamics of 'small-world' networks. – Nature 393: 440–442.

Woodward, G. et al. 2010. Ecological networks in a changing climate. – In: Woodward, G. (ed.), Advances in ecological research. Elsevier, pp. 71–138.

Wu, N. et al. 2017. Using river microalgae as indicators for freshwater biomonitoring: review of published research and future directions. – Ecol. Indic. 81: 124–131.

Zappelini, C. et al. 2015. Diversity and complexity of microbial communities from a chlor-alkali tailings dump. – Soil Biol. Biochem. 90: 101–110.

Zhou, J. et al. 2011. Phylogenetic molecular ecological network of soil microbial communities in response to elevated $CO_2$. – MBio 2: e00122-11.

Zorzal-Almeida, S. et al. 2017. Local environment and connectivity are the main drivers of diatom species composition and trait variation in a set of tropical reservoirs. – Freshwater Biol. 62: 1551–1563.