

ECOGRAPHY

Research

Assessing the relative influences of abiotic and biotic factors on American eel *Anguilla rostrata* distribution using hydrologic, physical habitat, and functional trait data

Taylor Woods and Daniel J. McGarvey

T. Woods (<http://orcid.org/0000-0002-6277-1260>) (woodstaylorizbeth@gmail.com) and D. J. McGarvey (<http://orcid.org/0000-0002-1218-5896>), Center for Environmental Studies, Virginia Commonwealth Univ., Richmond, VA, USA.

Ecography

41: 2067–2079, 2018

doi: 10.1111/ecog.03782

Subject Editor: Carsten Dormann

Editor-in-Chief: Miguel Araújo

Accepted 20 March 2018

Species' distributions are influenced by abiotic and biotic factors but direct comparison of their relative importance is difficult, particularly when working with complex, multi-species datasets. Here, we compare the relative effects of hydrology, physical habitat, and co-occurring fish functional traits on the contemporary (1950–1990) distribution of the American eel *Anguilla rostrata* in six Mid-Atlantic (USA) rivers. To do so, we implement a null model approach that compares conditions at sites of known American eel presence to a random sample of sites throughout a broader landscape, allowing us to identify variables that may have the strongest influences on American eel distribution. Results suggest that, within this subset of the American eel's geographic range, the functional characteristics of locally co-occurring fishes and habitat fragmentation by dams may have the strongest influences on American eel distribution, compared to other predictor variables included in the analysis. Given the widespread distribution and complex biology of this species, we caution that our results may not apply to all American eel subpopulations or life stages. Nonetheless, the observed importance of co-occurring fish functional traits may inform American eel conservation and, more generally, provide a means to incorporate biotic influences in research on species' distributions.

Keywords: hydrology, physical habitat, biotic interactions

Introduction

Characterizing the effects of biotic and abiotic influences on species' distributions and community structure is a long-standing challenge in ecology (Hutchinson 1957, Chase and Leibold 2003). To make this endeavor more tractable, abiotic and biotic effects have often been conceptualized as distinct levels in a series of nested, hierarchal filters that sort regional species pools into local communities (Poff 1997, Jackson et al. 2001). Notably, abiotic effects tend to comprise the higher, large-scale levels of the hierarchal filter model, while biotic influences comprise the lower, local-scale levels.



In this way, species from a regional pool must navigate a series of relatively coarse-scale filters, such as physiological thermal constraints, before they can potentially take up residence at a particular locality. Then, to become a successful colonist, a species must navigate the final local-scale filter: coexistence within the previously established community of resident species.

Perhaps due to the inherent complexity of characterizing biotic interactions within natural communities, evidence for the role of abiotic filtering in nature has accumulated more quickly than for biotic filtering, particularly at landscape or regional scales (Lawton 1999, Morales-Castilla et al. 2015). But with global change and other large-scale perturbations rearranging species' distributions and creating novel communities, ecologists are now earnestly working to enhance understanding of biotic influences (Parmesan and Yohe 2003, Gallardo and Aldridge 2013). For instance, improved methods to incorporate biotic interactions in species distribution models are an active area of inquiry (Guisan and Thuiller 2005, Wisz et al. 2013). Yet despite this progress, two key developments remain that would greatly benefit basic and applied research on biotic filtering: 1) improved methods to efficiently and systematically characterize biotic influences within complex multi-species communities; and 2) analytical tools and/or databases that can be used to assess the relative importance of abiotic and biotic factors at common spatial scales.

Methods to characterize biotic interactions are often based upon records of pairwise species' co-occurrences (Connor and Simberloff 1979, Gotelli and McCabe 2002). For example, species distribution models seeking to incorporate biotic influences on the distribution of a focal species have made use of heterospecific occurrence records, abundances, and co-occurrence-based indices as biotic predictor variables (Leathwick and Austin 2001, Meier et al. 2011). Unfortunately, inferences based on co-occurrence data may be compromised by the fact that multiple processes can sometimes lead to the same co-occurrence patterns (Peres-Neto 2004, Cazelles et al. 2016). Direct observational evidence of pairwise interactions (positive, negative, or otherwise) can overcome these limitations, but such empirical data are difficult to obtain and only available for a relatively small number of species pairs (Bascompte and Jordano 2007, Connor et al. 2013). Furthermore, biotic influences may manifest as direct or indirect associations between many species that are not well-described by pairwise scenarios (Wootton 1994). Methods to efficiently characterize biotic interactions within complex, multi-species datasets (the 'biotic milieu' of McGill et al. 2006) are therefore needed (Tylianakis et al. 2008, Gallien et al. 2017).

The second development – direct comparisons of the relative importance of abiotic and biotic factors – is necessary because differing scales are often used to characterize abiotic and biotic influences (Wiens 2011, Staniczenko et al. 2017), which makes direct comparison difficult. Abiotic factors are commonly represented by coarse-scale environmental surveys or standardized data from a geographic information

system, while biotic interactions are evaluated at finer scales, often through experimental studies (Jackson et al. 2001, Ovaskainen et al. 2017). Also hindering direct comparisons are the assumptions and logistical hurdles that must be navigated when scaling up from fine-scale biotic data (e.g. point observations) to the larger scales at which abiotic data are most often available (Araújo and Luoto 2007, Funk et al. 2016). As noted above, a hierarchical filtering framework that a priori assumes abiotic habitat filtering is dominant at coarser scales while biotic influences are paramount at finer scales is often used to integrate abiotic and biotic data in a single, regional analysis (Pearson and Dawson 2003, Boulangeat et al. 2012). But large-scale species distribution and co-occurrence patterns may arise from habitat filtering, biotic interactions, or a combination of both processes (HilleRisLambers et al. 2012, Cadotte and Tucker 2017). Thus, direct comparisons of abiotic and biotic influences at common scales may help to resolve this issue.

In this study, we compare abiotic and biotic influences on the distribution of a focal freshwater fish species, using multiple data sources that have been carefully aligned to a common spatial scale. Specifically, we use a combination of hydrologic and physical habitat data to quantify abiotic conditions, then use fish functional trait data to characterize the effects of multi-species biotic influences. Functional trait analyses emphasize species' physiological, morphological, and behavioral characteristics, rather than their taxonomic identities (McGill et al. 2006, Frimpong and Angermeier 2010). They are ideal for assessing biotic influences in multi-species datasets because traits often regulate biotic interactions (e.g. resource competition among species with similar feeding behaviors; MacArthur and Levins 1967). Functional trait data can also be aggregated into assemblage- or community-level summaries of biotic influences (Carmona et al. 2016), then contrasted with abiotic variables. Finally, at very large regional to global scales where extensive turnover in species composition is likely, functional traits may facilitate direct comparison because similar or analogous trait profiles are often shared among taxonomically distinct biotas (Mims et al. 2010).

As an empirical context, we focus on the contemporary (records collected between 1950–1990) distribution of the American eel *Anguilla rostrata* in six Mid-Atlantic (USA) rivers. Freshwater eels (family: Anguillidae) occur in rivers around the world and have complex biological and life history characteristics that make them particularly suitable for a comparison of abiotic and biotic effects. Eels are catadromous (ocean spawning) fishes with tremendous migratory capabilities and are not believed to undergo a juvenile imprinting process (Gagnaire et al. 2012). This suggests that freshwater habitat selection by upstream migrating elvers (juvenile eels) is a 'real-time' function of an individual's perception of local habitat conditions and/or other locally occurring organisms, rather than a programmed or inherited response. Furthermore, while American eel spawning events have yet to be empirically observed, genetic evidence suggests this species is panmictic: adults spawn only once and do so within large

aggregations of conspecifics (Wirth and Bernatchez 2003). This ‘all eggs in one basket’ strategy minimizes the probability that selective forces will promote diversification in local habitat use among distinct populations or lineages (Oliveira 1999). Anguillids are also notable for their widespread distributions and generalist habitat tendencies. For instance, the North and South American distribution of the American eel ranges from Canada to Venezuela, with American eels known to inhabit most freshwater habitats within this range (Jenkins and Burkhead 1993).

Unfortunately, the American eel is, like most other freshwater eel species, in a state of decline (Dekker 2003). Key threats include commercial fishing, habitat fragmentation and turbine entrainment at dams, and invasive species (Chaput et al. 2014). Stocking and translocation of juvenile eels in upstream habitats has helped mitigate European eel declines (Moriarty and Dekker 1997). But these methods have less often been used with American eel populations; the first large-scale stocking program in North America was implemented in Canada in 2005 (Pratt and Threader 2011). Conservation of the American eel may ultimately require a better understanding of the factors that influence its distribution, particularly during the freshwater, inland stages of its life cycle (Smogor et al. 1995). We therefore believe that a large-scale comparison of abiotic and biotic factors that influence American eel occurrences within freshwater will benefit both basic ecology and applied conservation.

Using a combination of contemporary American eel occurrence records with data on local hydrology and physical habitat (abiotic influences), as well as functional trait data for local fish assemblages (biotic influences), our objectives are to: 1) use a combination of species’ occurrence and functional trait data to quantify potential biotic influences within stream networks; and 2) perform a direct comparison of the relative effects of abiotic and biotic factors on the regional distribution of the American eel.

Material and methods

Study area and spatial framework

Our study area includes six of the major Mid-Atlantic (USA) river basins within the native range of the American eel (Jenkins and Burkhead 1993): the Potomac, Rappahannock, York, James, Chowan, and Roanoke rivers (Supplementary material Appendix 1 Fig. A1). To facilitate direct comparisons among variables, all hydrologic, physical habitat, and fish data were aggregated within ‘subwatersheds’ or 12-digit hydrologic units from the USGS (2012) Watershed Boundary Dataset. A total of 1407 subwatersheds exist within the six major river basins included in our study, with a mean surface area of 89.3 km² (SD = 40.5 km²). They were chosen for this study because they maximized our ability to match American eel occurrence records with independent data on hydrology and physical habitat at local spatial scales.

Hydrologic data and subwatershed selection

Selection of subwatersheds to include in our analyses began with a survey of available hydrologic data, as these data were the most limiting of the three data classes (hydrologic, physical habitat, and biotic). Hydrologic metrics were calculated from daily discharge records, downloaded for all stream gauges located within the six study basins via the US Geological Survey’s Water Data for the Nation website (<<https://waterdata.usgs.gov/nwis>>). Within each subwatershed that contained at least one stream gauge, we used three criteria to determine which gauge data would be used to represent local hydrology (Fig. 1). First, we queried gauges with a minimum of 20 yr of nearly continuous (periodic instances of several days or weeks of missing data were acceptable) discharge records between the 1955–1985 water years (i.e. 1 October–30 September; $n=158$ gauges with sufficient records). This ensured that our hydrologic metrics would be broadly representative of contemporary flow conditions (Gan et al. 1991) and temporally consistent with the majority of the fish occurrence records (collected between 1950–1990 in our study basins; see ‘Fish and functional

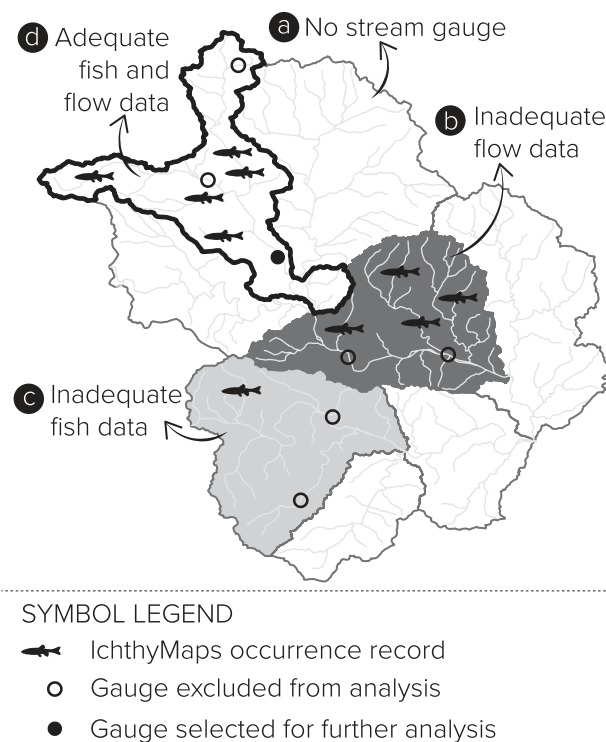


Figure 1. Map illustrating the process used to select subwatersheds used in this study. Four hypothetical scenarios are shown for subwatersheds that were: (a) not considered due to absence of an internal stream gauge (white, thin grey outline); (b) eliminated from further analysis due to inadequate flow data (e.g. internal gauge supplied flow data outside of water years 1960–1980; see text); (c) eliminated from further analysis due to inadequate fish sampling effort (≤ 10 th percentile of sample densities; see text); and (d) retained for further analysis, with adequate fish and flow data (white with heavy black outline).

trait data' section below). Second, we identified subwatersheds with more than one internal gauge ($n=5$), determined which gauge was closest to the downstream terminus or 'pour point' of each subwatershed, then removed gauges that were further upstream. This process narrowed the pool of suitable hydrologic units with representative flow data to 153 subwatersheds.

Hydrology in each of the remaining subwatersheds was then characterized with a subset of the indicators of hydrologic alteration (IHA) metrics (Richter et al. 1996). Sixteen IHA metrics were selected to represent the five primary flow regime components: magnitude, frequency, duration, timing, and rate of change (Olden and Poff 2003). For example, the median Julian date of annual minimum flow specifies the median day of the calendar year when streamflow is at an annual minimum, across all years of record. This index of the timing of extreme low flow events may be critical in regulating the occurrence of stream organisms that have low drought tolerances (Richter et al. 1996). IHA metrics were calculated with the 'IHA' package (Law 2013) in R (R Core Team), then appended to a subwatershed $\times$ IHA metrics matrix. From this matrix, we calculated Pearson's correlation coefficients and removed highly correlated metrics ($|r| \geq 0.7$), taking care to retain at least one variable from each flow regime component (Table 1).

Physical habitat data

Stream habitat within subwatersheds was represented by the 1:100 000 scale National Hydrography Dataset ver. 2 ('NHDv2'; McKay et al. 2012). All NHDv2 stream segments within each of the 153 selected subwatersheds (see above) were queried then appended with physical habitat characteristics from the original NHDv2 attribute tables and the StreamCat database (Hill et al. 2016). StreamCat variables represent local habitat conditions, including both natural and anthropogenic variables (e.g. land cover, urban land use, and geologic data), and are distributed at both catchment (i.e. the landscape that is immediately adjacent to a given stream segment, exclusive of landscapes further upstream) and watershed scales (i.e. the entire, cumulative landscape that is upstream of and contributing flow to a given segment; Hill et al. 2016). From StreamCat and NHDv2, we selected a suite of catchment-scale physical habitat metrics to represent instream fish habitat. For example, mean elevation indicates a subwatershed's longitudinal position within the stream continuum, which is strongly associated with channel morphology and substrate, as well as the ratio between autochthonous and allochthonous food resources (Vannote et al. 1980, Schlosser 1991). Catchment-scale metrics were selected from StreamCat, rather than watershed-scale, to better emphasize local conditions. For each of these physical habitat metrics, we calculated subwatershed means, medians, maximums, and coefficients of variation (CV), based upon the entire population of stream segments within a given subwatershed. Subwatershed summary statistics for each physical habitat metric were then appended

Table 1. Descriptions of the predictor variables used in the analysis following elimination of collinear variables. The statistic used to summarize each variable at the subwatershed scale (coefficient of variation [CV], mean [MN], median [MD], or proportion [%TRT]) is shown in parentheses, with units (if applicable). For hydrologic variables, the flow regime component represented by the predictor variable is shown in italics.

Variable code	Description
Hydrology	
AprCV	April stream flow (CV; <i>magnitude</i>)
AprMD	April stream flow (MD; $\text{ft}^3 \text{s}^{-1}$; <i>magnitude</i>)
BsflwCV	Baseflow index (CV; <i>magnitude</i>)
DtMnMD	Julian date of annual minimum flow (MD; <i>timing</i>)
HFlwCntCV	Number of flows above the 75th percentile (CV; <i>frequency</i>)
HFlwDurCV	Duration of flows periods above the 75th percentile (CV; <i>duration</i>)
RevCV	Number of transitions between rising and falling flow rates (CV; <i>rate of change</i>)
Physical habitat	
AreaMN	Catchment surface area (MN; km^2)
DmDnsMN	Density of dams per catchment area (MN; dams km^2)
DmStrMN	Volume of dams per catchment surface area (MN; $\text{m}^3 \text{km}^2$)
ElevCV	Catchment elevation (CV)
ElevMN	Catchment elevation (MN; m)
RnffCV	Catchment runoff (CV)
RnffMN	Catchment runoff (MN; mm)
Functional traits	
FecMN	Fecundity per female per spawning season (MN)
IncbTmMN	Incubation time from embryo fertilization to hatch (MN; hours)
SeasLenMN	Length of spawning season (MN; months)
Loc	Locomotion mode (metric indicating morphology and locomotion)
LocAccl	Accelerators: ambush predators with large caudal fins (%TRT)
LocAng	Anguilliform: swimmers with eel-like bodies (%TRT)
LocCrp	Creepers: bottom rovers with subterminal mouths (%TRT)
LocCrsr	Cruisers: active swimmers with streamlined body forms (%TRT)
LocHgr	Huggers: benthic fishes adapted to cling to substratum (%TRT)
LocMnvr	Maneuverers: laterally compressed body forms (%TRT)

to a subwatershed $\times$ habitat variable matrix and collinear variables ($|r| \geq 0.7$) were removed (Table 1).

Fish and functional trait data

Fish occurrence records were obtained from the IchthyMaps database (Frimpong et al. 2015) for each of the 153 subwatersheds. IchthyMaps is a compilation of contemporary fish records (collected between 1950–1990) that were assembled from point distribution maps in regional fish atlases and government agency surveys, then geo-referenced to their respective digital stream segments in the NHDv2. Within the study region, IchthyMaps provided a total of 32 463

fish occurrence records, distributed among 224 species. Of these, 559 were American eel records. Most American eel occurrence records were compiled from the text 'Freshwater fishes of Virginia' and were originally collected either with backpack electrofishing or seine netting surveys, between 1968–1986 (Jenkins and Burkhead 1993). Notably, these occurrence records should consist primarily of individuals at the yellow eel stage, rather than leptocephalus, glass eel, or silver eel stages; yellow eels are the most abundant life stage in freshwater, inland streams, where they become relatively sedentary and easy to capture with electrofishing equipment (Goodwin and Angermeier 2003). Therefore, we were reasonably certain that the American eel occurrences in our study represented permanent, resident habitat for American eels and were not unduly biased by inclusion of non-resident migratory life stages.

To account for potential sampling bias in the aggregated IchthyMaps data (i.e. under-sampled subwatersheds), we screened subwatersheds that contained low numbers of fish samples following McGarvey et al. (2017). Briefly, we calculated sample density as the total number of IchthyMaps observations (i.e. total number of occurrence records, regardless of species' identities) within a given subwatershed, divided by the total length of all stream segments within that subwatershed. This helped to detect subwatersheds that were vulnerable to under-sampling bias. Subwatersheds with fish sample densities below the 10th percentile (i.e. ≤ 3.0 occurrences per km of stream channel; median fish sample density = 11.3, CV = 1.3) were removed ($n = 16$) from the dataset.

Fish species' functional traits were then used to quantify biotic influences within each of the remaining subwatersheds ($n = 137$). We first compiled all IchthyMaps records within each subwatershed and converted the species list into a master species' presence matrix (subwatershed $\times$ species). Next, we collected functional trait data for each species ($n = 139$) through an extensive literature review, inclusive of the four data sources reported in Mims et al. (2010): 1) regional fish atlases (i.e. 'The fishes of...' texts); 2) primary and secondary literature publications; 3) the online FishBase (<www.fishbase.org>) and FishTraits (<www.fishtraits.info>) databases; and 4) keyword internet searches. For species' traits with multiple published values, we retained observations that were reported from localities that were closest to our study area. For instance, functional trait descriptions from the 'Freshwater fishes of Virginia' (Jenkins and Burkhead 1993) were prioritized over other sources for many of our trait values. Functional trait data were then compiled into a species $\times$ trait matrix and categorical variables were re-coded as binary dummy variables.

Finally, functional trait summaries were calculated for each subwatershed and used as proxies for species' potential interactions. For example, differences in fecundity or timing of spawning may facilitate coexistence among species with similar habitat requirements (Meador and Brown 2015). Mean, median, and maximum values were calculated for numeric traits while proportions were calculated for categorical traits.

Importantly, we omitted American eel traits from the subwatershed trait summaries because our objective was to detect interspecific trait influences on contemporary American eel occurrences; we sought only to determine which traits of co-occurring species may affect American eel habitat selection, not to define trait profiles for complete assemblages or to assess intraspecific effects. Trait summaries were compiled in a subwatershed $\times$ trait matrix and collinear traits ($|r| \geq 0.7$) were removed (Table 1).

Comparing sample and background distributions

Potential effects of the abiotic and biotic variables on American eel habitat selection were assessed with a null model approach. Specifically, we compared abiotic and biotic conditions at subwatersheds occupied by American eels with conditions observed in a random subset of subwatersheds, sampled from the complete population of subwatersheds throughout the landscape. This general approach to identifying variables with nonrandom effects on a species' distribution is useful when presence-only, rather than presence-absence, data are available (Hirzel et al. 2002). One particularly well-known application of this method is Maximum Entropy (MaxEnt; Phillips et al. 2006) species distribution modeling. MaxEnt compares a 'sample' distribution, representing conditions at sites of known occurrence, with a random 'background' sample that characterizes conditions throughout the potential range (Elith et al. 2011). A strong effect of a given variable on the focal species' distribution is indicated by sample and background distributions that exhibit modest or minimal overlap (i.e. conditions at known occurrence sites are non-random with respect to the entire landscape; see Fig. 1 in Merow et al. 2013).

We used the entire landscape of 137 subwatersheds as the background distribution for our analysis. We separated this background into a 'sample' population of subwatersheds with known American eel occurrences ($n = 24$) and an 'available' population that was within the geographic range of the American eel, but where its occurrence had not been documented and was therefore unknown ($n = 113$; Fig. 2, step 1).

Next, we implemented a random permutation algorithm to assess similarities between the sample and background distributions for each hydrologic, physical habitat, and functional trait variable. Permutations were used instead of direct, static comparisons of the complete sample and background distributions for each variable because we did not wish to overfit our results, leading to low transferability (Thomas and Bovee 1993, Chatfield 1995). In each of the 1000 permutations, we randomly selected (without replacement) 12 of the 24 sample units and 60 of the 137 background units (Fig. 2, step 2). These permuted sample sizes preserved the original, approximate ratio of sample-to-available units ($\sim 1:5$) and ensured that our results would not be biased by unbalanced representation of the sample vs. available units in any given permutation. We then compared the randomized sample and background distributions (in each permutation) for each of the predictor variables with a 2-step process (Fig. 2, step 3).

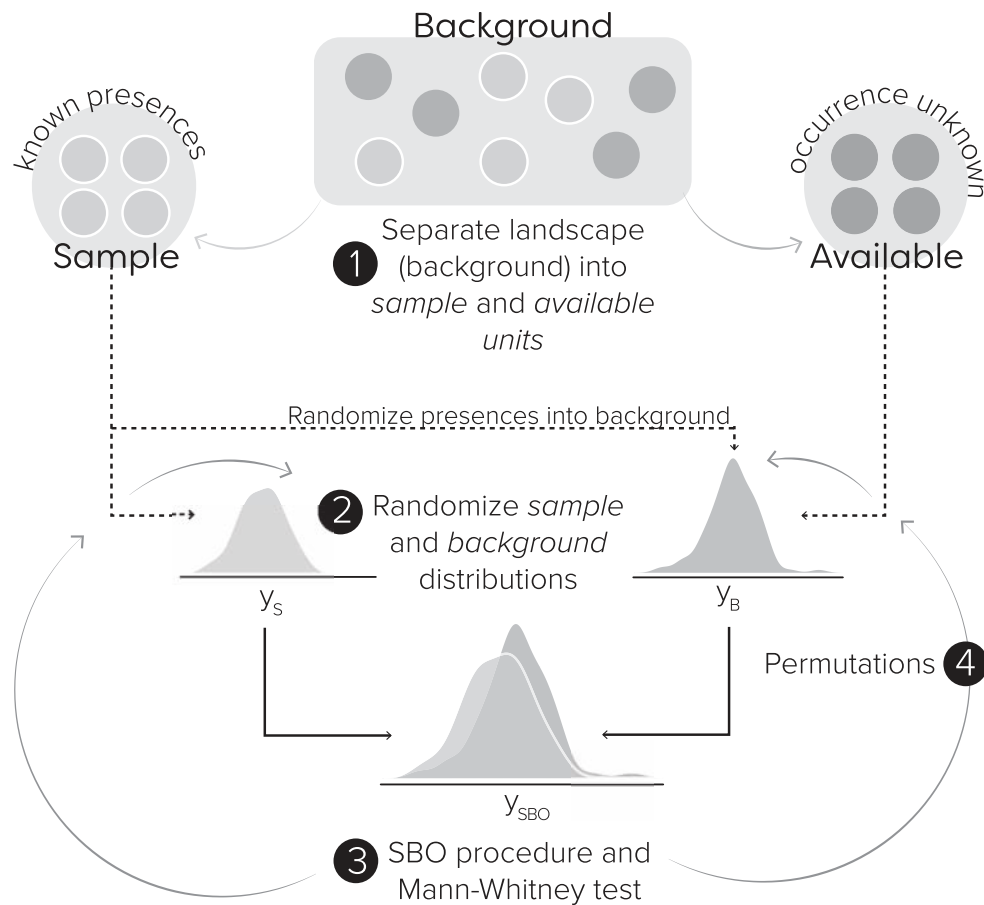


Figure 2. Workflow diagram of the procedure to compute permuted sample-background overlap (SBO) and Mann–Whitney U -statistics for each of the predictor variables included in this study. Major steps in the process are identified with numbers and explained in the Material and methods (main text).

First, we quantified sample-background overlap (SBO) as the literal area of overlap (i.e. superimposed probability distribution functions) between continuous sample and background distributions, with potential values ranging from 1 (perfect overlap) to 0 (no overlap). SBO values close to 1 suggest that sample and background units are both random samples from the complete landscape (i.e. sample values for the predictor variable do not differ with respect to the background values, H_0). Alternatively, SBO values much smaller than 1 imply that sample units are nonrandom with respect to the landscape (i.e. the sample values encompass a relatively small subset of the complete range of background values, indicative of a selective process, H_A ; Hirzel et al. 2002).

In each permutation, SBO was calculated for every predictor variable with Eq. 4 in Mouillot et al. (2005); their Eq. 4 was conceived as a metric of overlap in the trait densities of two co-occurring species, using kernel density functions (see also Mason et al. 2008, 2011). Our application of this metric was, however, different in one key regard; the original authors interpreted a high degree of overlap among kernel functions as evidence of potential competition, but we sought to identify variables with minimal overlap as evidence of a strong,

non-random influence on American eel distribution. Because many of the variables considered here were non-normal and/or contained zero values (e.g. dam density), we applied a $\ln(x+1)$ transformation to all continuous variables. All kernel density estimates were calculated with the ‘density’ function in the package ‘stats’ (Gaussian kernels and default bandwidth and n settings) in R.

In the second step, nonparametric Mann–Whitney U -statistics were used to determine whether permuted sample values were consistently higher or lower than permuted background values for each of the predictor variables. Two-sided statistics were used because we did not expect a priori that American eel sample values would be consistently higher or lower than background values for most of the predictor variables; we sought only to document whether the sample and background distributions were consistently different. In each permutation, we calculated and recorded the U -statistic from a 2-sided Mann–Whitney test for each variable, using the ‘wilcox.test’ function in R. U -statistic values were then interpreted relative to the magnitude of deviation from an equal ranking of sample and background values (i.e. $U=360$). U -statistics closer to zero indicated that sample distribution

values were consistently ranked higher or were larger than background values, whereas U -statistics approaching the maximum (i.e. $U=720$) indicated that sample values were consistently ranked lower or were smaller than background values for a given variable.

As a final step, we used a multivariate SBO approach to assess the collective effect of each of the three classes of variables on American eel distribution throughout the study rivers. For each class of predictor variable, nonmetric multidimensional scaling (nMDS) was used to build a 2-dimensional ordination of the sample and background data. Ninety-five percent confidence ellipses, or 'hypervolumes' (Blonder et al. 2014), were interpolated for the sample points and background points in each ordination plot. The regions defined by the background ellipses indicated the expected hypervolume for a random sample of the complete landscape. We then visually assessed the degree of overlap between sample and background hypervolumes to determine whether American eel occurrences were nonrandom with respect to the background hypervolume for each of the three predictor variable classes. We did not, however, use permutations to perform repeated comparisons of the sample and background hypervolumes. Instead, we used the complete sample and background data ($n=24$ and 137 , respectively) to perform a single multivariate SBO analysis for each of the three classes of predictor variables. For each predictor class, we calculated a Gower dissimilarity matrix (Gower 1971) that included all variables within the class (i.e. three independent subwatershed $\times$ subwatershed dissimilarity matrices), then used the dissimilarity matrices to perform nMDS. Gower dissimilarities were calculated with the 'FD' package (Laliberte et al. 2014), nMDS was performed with the 'vegan' package (Oksanen et al. 2017), and confidence ellipses were plotted with 'ggplot2' (Wickham 2009) in R.

Data deposition

The raw data and code for analyses are available from Figshare Digital Repository: <<https://doi.org/10.6084/m9.figshare.5481205.v4>> (Woods and McGarvey 2017).

Results

Permuted statistics for individual predictor variables

Of the three predictor variable classes, hydrologic variables generally exhibited the highest SBO values (median SBO=0.83, CV=0.06; Fig. 3a), suggesting that contemporary occurrences of the American eel are not strongly associated with specific hydrologic conditions. Only bsflwCV had low permuted SBO values (< 0.75 ; Fig. 3a) and U -statistics that strongly deviated from the line of equality ($U=360$), with sample values consistently larger than background values (Fig. 3b). In comparison, physical habitat variables appeared to have greater influence on the contemporary American eel distribution (median SBO=0.76, CV=0.16; Fig. 3a). In the physical habitat class, low permuted SBO values were

observed for dmDnsMN and elevMN (Fig. 3a). U -statistics showed that sample values were consistently larger than background values for dmDnsMN, but the opposite trend was observed for elevMN (Fig. 3b). Overall, the functional trait data class had the lowest SBO values (median SBO=0.69, CV=0.32), with particularly low permutation values for fecMN, incbTm, locCrsr, locMnvr, and seasLen (Fig. 3a). Functional trait U -statistics showed sample distribution values of fecMN, locMnvr, and seasLen were higher than background, whereas incbTm and locCrsr background values exceeded sample values (Fig. 3b).

Multivariate hypervolumes

Consistent with permutation results for individual variables, comparisons of the multivariate sample and background hypervolumes for the three data classes indicated that functional traits may have the overall strongest influence on American eel occurrence. Sample and background hypervolumes exhibited extensive overlap for hydrologic variables (Fig. 4a). Overlap between sample and background hypervolumes was intermediate for physical habitat variables (Fig. 4b). Hypervolume overlap was smallest for the functional trait data class, with approximately 50% overlap observed between the sample and background hypervolumes (Fig. 4c).

Discussion

Abiotic and biotic influence on American eel distribution

Within six Mid-Atlantic rivers, we compared abiotic and biotic conditions at confirmed American eel presence sites (the sample distribution) with a representative sample of conditions across the entire landscape (the background distribution). In general, results indicated that the sample distribution was more closely associated with dam density and the functional traits of co-occurring fishes than the other predictor variables included in our analysis. These results suggest that future research and conservation of inland American eel populations within the study region may benefit from a refined focus on the effects of habitat fragmentation and potential interactions with co-occurring fishes.

Dam density had one of the strongest influences on American eel distribution, suggesting that fishes in the study region have highly fragmented ranges. This result seemed intuitive because it is well-known that dams, which prevent migratory American eels from reaching upstream tributaries (Wiley et al. 2004, Machut et al. 2007), can have a stronger influence on American eel distribution than other abiotic factors (Hitt and Roberts 2012). However, the directional effect of dam density was surprising: rather than associating with low dam densities, the sample distribution suggested that American eels may be associated with relatively high dam

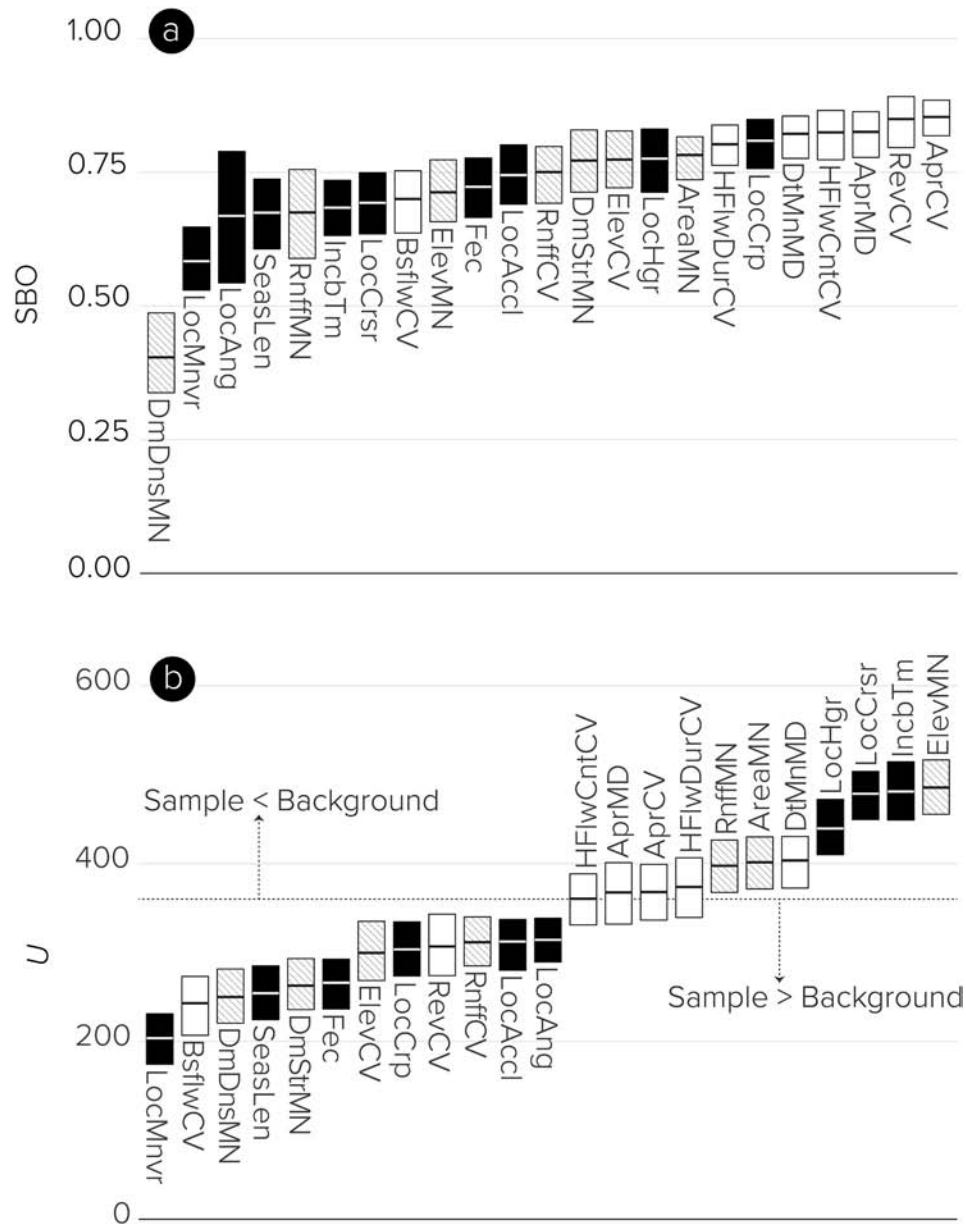


Figure 3. Boxplots (25th, 50th, and 75th quartiles) showing permutation results as sample-background overlap (SBO; panel a) and Mann-Whitney U -statistic (panel b) values. Boxplot colors denote the three classes of predictor variables: hydrologic (white, black outline), physical habitat (grey hatched, black outline), and functional traits (black, white outline). In panel (b), the horizontal line at $U=360$ represents rank equivalence between sample and background distribution values. Below the rank equivalence line, sample distribution values are, on average, larger than background values for the variable in question. Above the rank equivalence line, sample distribution values are smaller than background values for the variable in question.

densities. This counterintuitive result may be explained by a tendency for American eels to aggregate near dams. Prior studies have shown that American eel densities increase near dams and suggested that this may result from large congregations of American eels that gather downstream of impassable barriers (Machut et al. 2007). We mapped our American eel presence records against georeferenced dam locations and found that the sample distribution may have included multiple sites that were downstream of an impassable dam (these sites would have high dam density values) and therefore

represented the upstream limits of potential American eel migration (Supplementary material Appendix 1 Fig. A2). If so, the strong association between American eel presence and high dam density may be an artifact of higher probability of capture at these sites, rather than a preferential selection of subwatersheds with high dam densities per se. This observation is particularly concerning, given that heightened American eel densities below dams may amplify density-dependent processes and negatively impact these populations (Machut et al. 2007).

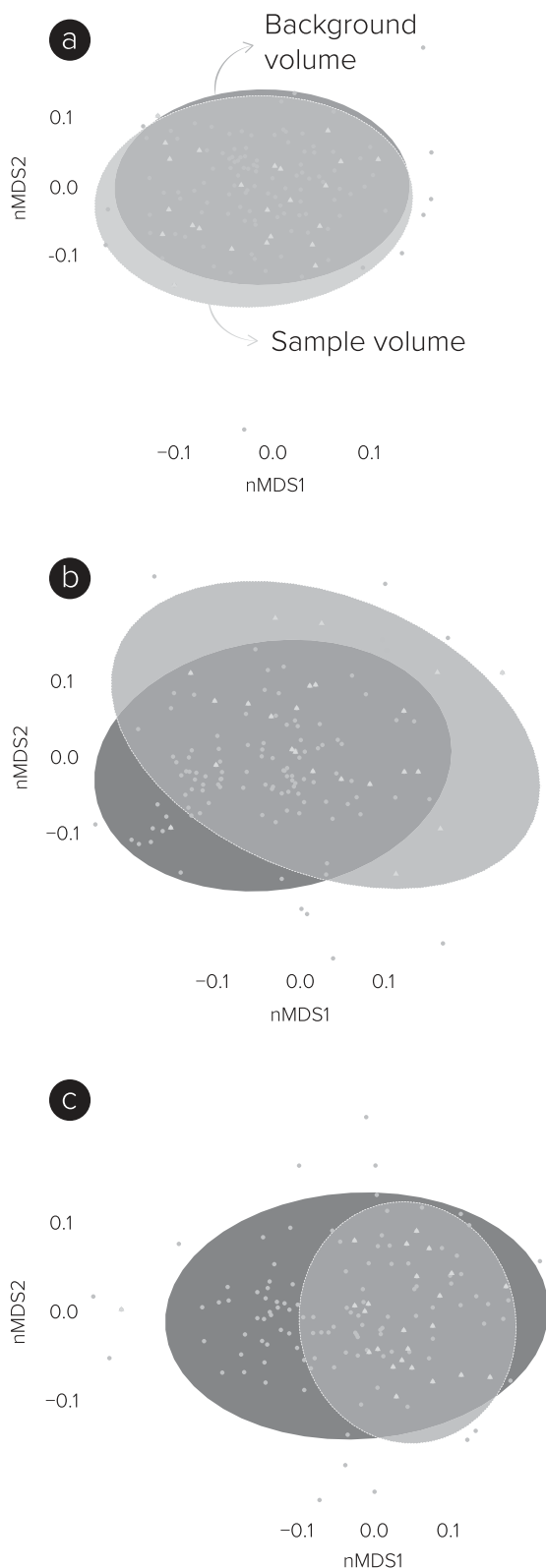


Figure 4. Nonmetric multidimensional scaling (nMDS) ordination plots for the hydrologic (a), physical habitat (b), and functional trait (c) data classes. Hypervolumes are shown as 95% confidence ellipses for sample (light grey) and background (dark grey) data. Points represent individual subwatersheds from the sample distribution (light grey triangles) and the background distribution (dark grey circles).

We found that hydrology may play a minimal role in regulating inland American eel distribution in the study region, but we caution that our results may only be interpreted with respect to the particular American eel life stages that were represented in the IchthyMaps database. For instance, if the occurrence records included silver eel observations, we would expect a strong hydrologic signal because seaward migration of silver eels is strongly influenced by river discharge (Jansen et al. 2007). We did find that the CV of baseflow index showed an effect on American eel distribution and this variable seemed to suggest that the sample distribution may be associated with habitat characterized by variable or unpredictable flows. These results are surprising because migratory fishes might be expected to associate with specific, predictable flow conditions (Bunn and Arthington 2002). One possible reason American eels may select streams with variable flow is that this species' ability to sustain prolonged exposure to air may give American eels a competitive advantage in these habitats.

In general, we acknowledge that our abiotic results may not be characteristic of American eel habitat selection throughout its inland range. Habitat-mediated, regional variation in life history characteristics has been documented for the American eel (Oliveira 1999) and it is likely that other regional subpopulations will respond differently to hydrology and physical habitat. Nevertheless, our results are consistent with previous studies that failed to identify strong associations between American eel populations and abiotic variables (Smogor et al. 1995, Wiley et al. 2004).

Of the three predictor variable classes, the functional traits of co-occurring fishes were the most effective in differentiating sample and background distributions, suggesting that inland distribution of the American eel in our study region may be a function of biotic influences. To better understand the effects that individual traits may have on American eel occurrence, we group the trait variables into two categories: reproductive behavior and locomotion. Reproductive trait results indicate that the sample distribution contains fishes that have relatively short incubation times, high fecundities, and/or long spawning seasons. Thus, the American eel seems to associate with heterospecifics that have high reproductive capacities. This could be advantageous to American eels because fishes with high reproductive capacities produce an abundance of ichthyoplankton, which in turn provides a rich food resource to predatory American eels (Helfman and Winkelman 1991).

Locomotion traits suggest that American eels associate with species that exhibit 'maneuvering' locomotion (sluggish swimmers with laterally compressed body forms typified by sunfishes), but not with 'cruising' (active swimmers with streamlined bodies typified by salmonids) species. This may indicate a preference for local assemblages of heterospecifics in which interspecific competition is minimized and prey availability is maximized. American eels may avoid streamlined, cruising species that are more effective predators than slower anguilliform eels (Sinha and Jones 1967). Alternatively, American eels may favor coexistence with maneuvering species because these fishes are less adept

competitors and easier prey. Coexistence with water column-dwelling, maneuvering species may also reduce interspecific competition through vertical habitat partitioning, as eels are primarily benthic during their freshwater residency (Ross 1986).

Functional traits as proxies for biotic interactions

Pairwise species approaches have traditionally been used to study biotic interactions, with methods ranging from basic null models to more recent developments in network analysis. However, these approaches can be computationally impractical because the number of species pairs that may potentially interact grows quickly as the number of locally co-occurring species increases (Morales-Castilla et al. 2015). In the present context, information on 138 pairwise associations would be needed to fully account for all direct biotic interactions between the American eel and co-occurring heterospecifics.

As an alternative, we used a general, null model approach to infer the effects of biotic interactions from community functional trait data. This is just one example of a broader effort to use functional traits to generalize processes that are logistically difficult to quantify on a species $\times$ species basis (McGill et al. 2006). For instance, body size may be used to estimate dispersal ability in metacommunity frameworks (De Bie et al. 2012), multidimensional trait classifications can be used to link species to their ecosystem functions (Winemiller et al. 2015), and knowledge of trait performance in relation to the environment may be used to predict community responses to abiotic filtering (Webb et al. 2010).

In the present example, two specific benefits of a generalized, traits-based approach to studying biotic interactions are evident. First, as noted above, the community trait analyses were more computationally feasible and less data-limited than a biotic filtering analysis that relied upon large numbers of pairwise species contrasts would have been. Without using the trait data, it is not clear that we could have compared abiotic and biotic factors at a common spatial scale. Second, because we did not focus explicitly on taxonomic identity, our results may be compared to other freshwater eels, such as the European eel *A. anguilla* or New Zealand longfin eel *A. dieffenbachia*. These fishes exhibit similar life histories to the American eel (Haro et al. 2000) but must coexist with very different fish assemblages throughout their freshwater life stages. Thus, by focusing on the functional relationships that eels have with conspecifics, rather than the taxonomic identities of co-occurring fishes, it may be possible to distinguish biotic influences that are germane to all freshwater eels from influences that are unique to a given species or population.

Caveats and opportunities

This study suggests that biotic predictor variables may be key determinants of American eel occurrence within Mid-Atlantic

ivers. But for several reasons, we urge caution when interpreting the results. First, although the final set of predictor variables included in our analysis was selected to represent a broad range of abiotic and biotic influences, it is unlikely that we incorporated all relevant factors. For example, water pollution is known to have direct lethal effects, as well as indirect behavioral effects, on the American eel (Haro et al. 2000). However, due to data limitations, we were unable to incorporate robust indicators of pollution in our analysis. Similarly, we were unable to incorporate recent or legacy effects of fishing harvest in our study. This is important because the American eel has historically been exposed to variable yet continuous fishing pressure throughout its Atlantic Coast range (Chaput et al. 2014, ASMFC 2017). We therefore suggest that new information on anthropogenic stressors, as well as other natural influences, might enhance future research on the American eel and potentially alter our conclusions.

Second, our analyses are subject to a constraint that necessarily arises from the use of combined survey data across large spatial extents. In large-scale studies where the use of aggregate occurrence data from multiple sources (such as IchthyMaps) is necessary, it may be difficult or impossible to retroactively distinguish among species' life stages, or to determine resident versus migrant status. Strong inference (sensu Platt 1964) regarding biotic interactions is therefore beyond the scope of these data (Elton 1946). This point is certainly relevant for the American eel, with a life history that includes five distinct developmental stages, each of which utilizes different combinations of habitat and trophic resources. For example, if the IchthyMaps data were inclusive of large numbers of glass eel observations, we might expect that local environmental conditions (e.g. water temperature and turbidity; Harrison et al. 2014), rather than biotic interactions, would drive selection of estuarine habitat for this non-feeding, migratory life stage. Previously, we explained that yellow eels should comprise the majority of the IchthyMaps records for the American eel and why uncertainty regarding life stage should not have a strong influence on our results (see 'Fish and functional trait data' above). We also note that many freshwater fishes experience ontogenetic shifts in feeding behavior or habitat use (Mittelbach and Persson 1998) that are not well-represented by coarse-scale occurrence data. Therefore, the data limitations discussed here are germane to a large number of species.

Third, we recognize that the inferred biotic effects (i.e. functional traits with low permuted SBO values and high magnitude *U*-statistics) could be artifacts of abiotic filtering processes that regulate species' presences and, by extension, perceived functional trait patterns. For instance, our observation that the American eel is highly associated with maneuvering (locomotion mode) heterospecifics could be an artifact or secondary effect of dams having a strong influence on the presence of maneuvering fishes. In our particular case, we do not believe that the biotic filtering results are spurious. Using linear regression, we found no evidence that the proportion of maneuvering fishes was a function of dam density

($F_{1,111} = 1.19$, $p = 0.278$). Nevertheless, we caution that when using functional traits to infer biotic interactions, post-hoc tests to confirm that a given functional trait result is not an artifact of covariance with another abiotic variable will often be necessary, either for discrete pairs of variables (e.g. linear regression) or entire networks of variables (e.g. structural equation modeling).

These caveats are all noteworthy and underscore the fact that our results do not ‘prove’ that American eel occurrence is regulated by biotic interactions with conspecifics, rather than abiotic factors. Field trials and direct observational information would be necessary to show this conclusively. But we contend that the inherent limitations of our environmental and species’ occurrence data are not a fundamental problem. Large-scale studies of species’ distributions often make use of data from disparate sources, without precise information on species’ life stages or the degree of spatial or temporal matching between species occurrence and environmental data. Instances where fine-scale species’ occurrence and environmental data were synchronously collected, over large spatial extents, are rare (but see Al-Chokhachy et al. 2013 for a notable example). Yet much has been learned about the factors that govern species’ distributions, using relatively coarse occurrence and environmental data (Elith and Leathwick 2009). We cannot say with certainty that the IchthyMaps American eel samples are representative of individuals at the sedentary yellow eel stage (or another life stage), but this would be a general concern with any species that is highly vagile through part of its life history. We do know that, unlike most primary freshwater fishes that may hatch and take up residence in the same habitats without first navigating a suite of diverse abiotic conditions and local species assemblages, catadromous eels must invest considerable energetic resources to reach their resident freshwater habitats. Thus, it is logical to assume that the American eel is a good candidate for tests of abiotic and biotic influences on freshwater species distributions.

Finally, we emphasize that our analyses were conducted entirely with freely available data. We downloaded all hydrologic, physical habitat, and fish occurrence data from public archives. Only the functional trait data required manual assembly and proofing, and this information is now archived with the code needed to reproduce our analyses (see ‘Data deposition’ above). By using these existing resources or appending them with new variables (hundreds of additional attributes can be freely accessed from StreamCat or other sources, then cross-referenced to our study sites using the NHDv2 ‘COMID’ field), our study could be expanded to assess the effects of other environmental factors on American eel occurrence. For instance, historical air temperature and precipitation records from WorldClim (<www.worldclim.org>) could be appended to our data and used to better understand how the American eel may respond to future climate change in freshwater ecosystems. Alternatively, similar analyses could search for abiotic and biotic effects on other freshwater fishes within the Mid-Atlantic region.

Distinguishing abiotic from biotic influences remains a central challenge in ecology, but by leveraging these tremendous data resources in novel ways, significant and rapid progress is possible.

Acknowledgements – We are grateful to Carsten Dormann, Seth Wenger, and two anonymous reviewers for helpful comments on an early draft of the manuscript.

Funding – TW was supported by a Rice Rivers Center Graduate Research Award from Virginia Commonwealth Univ. DJM was supported by a grant from the U.S. National Science Foundation (DEB-1553111).

Author contributions – DJM and TW designed the study. TW prepared the data and performed the analyses. TW and DJM wrote the manuscript.

References

- Al-Chokhachy, R. et al. 2013. Characterizing the thermal suitability of instream habitat for salmonids: a cautionary example from the Rocky Mountains. – *Trans. Am. Fish. Soc.* 142: 793–801.
- Araújo, M. B. and Luoto, M. 2007. The importance of biotic interactions for modelling species distributions under climate change. – *Global Ecol. Biogeogr.* 16: 743–753.
- ASMFC 2017. 2017 American eel stock assessment update. – Atlantic States Marine Fisheries Commission, Arlington, VA, USA, <www.asmfc.org/species/american-eel>.
- Bascompte, J. and Jordano, P. 2007. Plant–animal mutualistic networks: the architecture of biodiversity. – *Annu. Rev. Ecol. Syst.* 38: 567–593.
- Blonder, B. et al. 2014. The n-dimensional hypervolume. – *Global Ecol. Biogeogr.* 23: 595–609.
- Boulangeat, I. et al. 2012. Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. – *Ecol. Lett.* 15: 584–593.
- Bunn, S. E. and Arthington, A. H. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. – *Environ. Manage.* 30: 492–507.
- Cadotte, M. W. and Tucker, C. M. 2017. Should environmental filtering be abandoned? – *Trends Ecol. Evol.* 32: 1–9.
- Carmona, C. P. et al. 2016. Traits without borders: integrating functional diversity across scales. – *Trends Ecol. Evol.* 31: 382–394.
- Cazelles, K. et al. 2016. A theory for species co-occurrence in interaction networks. – *Theor. Ecol.* 9: 39–48.
- Chaput, G. et al. 2014. Recovery potential assessment for the American eel (*Anguilla rostrata*) for eastern Canada: description and quantification of threats. – DFO Can. Sci. Advis. Sec. Res. Doc. 2013/135, vi + 90 p.
- Chase, J. M. and Leibold, M. A. 2003. Ecological niches: linking classical and contemporary approaches. – Univ. of Chicago Press.
- Chatfield, C. 1995. Model uncertainty, data mining and statistical inference. – *J. R. Stat. Soc. Ser. A* 158: 419–466.
- Connor, E. F. and Simberloff, D. 1979. The assembly of species communities: chance or competition? – *Ecology* 60: 1132–1140.
- Connor, E. F. et al. 2013. The checkered history of checkerboard distributions. – *Ecology* 94: 2403–2414.

- De Bie, T. et al. 2012. Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. – *Ecol. Lett.* 15: 740–747.
- Dekker, W. 2003. Worldwide decline of eel resources necessitates immediate action. Québec Declaration of Concern. – *Fisheries* 28: 28–30.
- Elith, J. and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. – *Annu. Rev. Ecol. Evol. Syst.* 40: 677–697.
- Elith, J. et al. 2011. A statistical explanation of MaxEnt for ecologists. – *Divers. Distrib.* 17: 43–57.
- Elton, C. 1946. Competition and the structure of ecological communities. – *J. Anim. Ecol.* 15: 54–68.
- Frimpong, E. and Angermeier, P. 2010. Trait-based approaches in the analysis of stream fish communities. – *Am. Fish. Soc. Symp.* 73: 109–136.
- Frimpong, E. et al. 2015. Historical stream fish distribution database for the conterminous United States (1950–1990): IchthyMaps. – U.S. Geol. Surv.
- Funk, J. L. et al. 2016. Revisiting the Holy Grail: using plant functional traits to understand ecological processes. – *Biol. Rev.* 92: 1156–1173.
- Gagnaire, P. A. et al. 2012. The genetic consequences of spatially varying selection in the panmictic American eel (*Anguilla rostrata*). – *Genetics* 190: 725–736.
- Gallardo, B. and Aldridge, D. C. 2013. Evaluating the combined threat of climate change and biological invasions on endangered species. – *Biol. Conserv.* 160: 225–233.
- Gallien, L. et al. 2017. The effects of intransitive competition on coexistence. – *Ecol. Lett.* 20: 791–800.
- Gan, K. C. et al. 1991. Analysis of periodicity in streamflow and rainfall data by Colwell's indices. – *J. Hydrol.* 123: 105–118.
- Goodwin, K. and Angermeier, P. 2003. Demographic characteristics of American eel in the Potomac River drainage, Virginia. – *Trans. Am. Fish. Soc.* 132: 524–535.
- Gotelli, N. J. and McCabe, D. J. 2002. Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. – *Ecology* 83: 2091–2096.
- Gower, J. 1971. A general coefficient of similarity and some of its properties. – *Biometrics* 27: 857–871.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – *Ecol. Lett.* 8: 993–1009.
- Haro, A. et al. 2000. Population decline of the American eel: implications for research and management. – *Fisheries* 25: 7–16.
- Harrison, A. J. et al. 2014. A review of glass eel migratory behaviour, sampling techniques and abundance estimates in estuaries: implications for assessing recruitment, local production and exploitation. – *Rev. Fish Biol. Fish.* 24: 967–983.
- Helfman, G. and Winkelman, D. 1991. Energy trade-offs and foraging mode choice in American eels. – *Ecology* 72: 310–318.
- Hill, R. A. et al. 2016. The stream-catchment (StreamCat) dataset: a database of watershed metrics for the conterminous United States. – *J. Am. Water Resour. Assoc.* 52: 120–128.
- HilleRisLambers, J. et al. 2012. Rethinking community assembly through the lens of coexistence theory. – *Annu. Rev. Ecol. Evol. Syst.* 43: 227–248.
- Hirzel, A. H. et al. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? – *Ecology* 83: 2027–2036.
- Hitt, N. P. and Roberts, J. H. 2012. Hierarchical spatial structure of stream fish colonization and extinction. – *Oikos* 121: 127–137.
- Hutchinson, G. E. 1957. Concluding remarks. – *Cold Spring Harb. Symp. Quant. Biol.* 22: 415–427.
- Jackson, D. A. et al. 2001. What controls who is where in freshwater fish communities – the roles of biotic, abiotic, and spatial factors. – *Can. J. Fish. Aquat. Sci.* 58: 157–170.
- Jansen, H. M. et al. 2007. Just go with the flow? Route selection and mortality during downstream migration of silver eels in relation to river discharge. – *ICES J. Mar. Sci.* 64: 1437–1443.
- Jenkins, R. and Burkhead, N. 1993. Freshwater fishes of Virginia. – American Fisheries Society.
- Laliberte, E. et al. 2014. FD: measuring functional diversity (FD) from multiple traits, and other tools for functional ecology. – R package ver. 1.0-12.
- Law, J. 2013. IHA: this package implements the nature conservancy's indicators of hydrologic alteration software in R. – <<https://rdr.io/rforge/IHA/>>.
- Lawton, J. H. 1999. Are there general laws in ecology? – *Oikos* 84: 177–192.
- Leathwick, J. R. and Austin, M. P. 2001. Competitive interactions between tree species in New Zealand old-growth indigenous forests. – *Ecology* 82: 2560–2573.
- MacArthur, R. and Levins, R. 1967. The limiting similarity, convergence, and divergence of coexisting species. – *Am. Nat.* 101: 377–385.
- Machut, L. S. et al. 2007. Anthropogenic impacts on American eel demographics in Hudson River tributaries, New York. – *Trans. Am. Fish. Soc.* 136: 1699–1713.
- Mason, N. W. H. et al. 2008. Does niche overlap control relative abundance in French lacustrine fish communities? A new method incorporating functional traits. – *J. Anim. Ecol.* 77: 661–669.
- Mason, N. W. H. et al. 2011. Niche overlap reveals the effects of competition, disturbance and contrasting assembly processes in experimental grassland communities. – *J. Ecol.* 99: 788–796.
- McGarvey, D. J. et al. 2017. On the use of climate covariates in aquatic species distribution models: are we at risk of throwing out the baby with the bath water? – *Ecography* doi:10.1111/ecog.03134
- McGill, B. J. et al. 2006. Rebuilding community ecology from functional traits. – *Trends Ecol. Evol.* 21: 178–185.
- McKay, L. et al. 2012. NHDPlus version 2: user guide (data model version 2.1). – <ftp://ftp.horizon-systems.com/NHDPlus/NHDPlusV21/Documentation/NHDPlusV2_User_Guide.pdf>.
- Meador, M. R. and Brown, L. M. 2015. Life history strategies of fish species and biodiversity in eastern USA streams. – *Environ. Biol. Fishes* 98: 663–677.
- Meier, E. S. et al. 2011. Co-occurrence patterns of trees along macro-climatic gradients and their potential influence on the present and future distribution of *Fagus sylvatica* L. – *J. Biogeogr.* 38: 371–382.
- Merow, C. et al. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. – *Ecography* 36: 1058–1069.
- Mims, M. C. et al. 2010. Life history trait diversity of native freshwater fishes in North America. – *Ecol. Freshwater Fish* 19: 390–400.

- Mittelbach, G. G. and Persson, L. 1998. The ontogeny of piscivory and its ecological consequences. – *Can. J. Fish. Aquat.* 55: 1454–1465.
- Morales-Castilla, I. et al. 2015. Inferring biotic interactions from proxies. – *Trends Ecol. Evol.* 30: 347–356.
- Moriarty, C. and Dekker, W. 1997. Management of the European eel. – *Fish. Bull.* 15: 1–110.
- Mouillot, D. et al. 2005. Niche overlap estimates based on quantitative functional traits: a new family of non-parametric indices. – *Oecologia* 145: 345–353.
- Oksanen, J. et al. 2017. Vegan: community ecology package. – <<http://cran.r-project.org/web/packages/vegan/index.html>>.
- Olden, J. D. and Poff, N. L. 2003. Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. – *River Res. Appl.* 19: 101–121.
- Oliveira, K. 1999. Life history characteristics and strategies of the American eel, *Anguilla rostrata*. – *Can. J. Fish. Aquat.* 56: 795–802.
- Ovaskainen, O. et al. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. – *Ecol. Lett.* 20: 561–576.
- Parmesan, C. and Yohe, G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. – *Nature* 421: 37–42.
- Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – *Global Ecol. Biogeogr.* 12: 361–371.
- Peres-Neto, P. R. 2004. Patterns in the co-occurrence of fish species in streams: the role of site suitability, morphology and phylogeny versus species interactions. – *Oecologia* 140: 352–360.
- Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – *Ecol. Model.* 190: 231–259.
- Platt, J. R. 1964. Strong inference. – *Science* 146: 347–353.
- Poff, N. L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. – *J. North Am. Benthol. Soc.* 16: 391–409.
- Pratt, T. C. and Threader, R. W. 2011. Preliminary evaluation of a large-scale American eel conservation stocking experiment. – *North Am. J. Fish. Manage.* 31: 619–628.
- Richter, B. D. et al. 1996. A method for assessing hydrologic alteration within ecosystems. – *Conserv. Biol.* 10: 1163–1174.
- Ross, S. T. 1986. Resource partitioning in fish assemblages: a review of field studies. – *Copeia* 1986: 352–388.
- Schlosser, I. J. 1991. Stream fish ecology: a landscape perspective. – *Bioscience* 41: 704–712.
- Sinha, V. R. P. and Jones, J. W. 1967. On the food of the freshwater eels and their feeding relationship with the salmonids. – *J. Zool.* 153: 119–137.
- Smogor, R. A. et al. 1995. Distribution and abundance of American eels in Virginia streams: tests of null models across spatial scales. – *Trans. Am. Fish. Soc. Am. Fish. Soc.* 124: 789–803.
- Staniczenko, P. P. A. et al. 2017. Linking macroecology and community ecology: refining predictions of species distributions using biotic interaction networks. – *Ecol. Lett.* 20: 693–707.
- Thomas, J. A. and Bovee, K. D. 1993. Application and testing of a procedure to evaluate transferability of habitat suitability criteria. – *Regul. River.* 8: 285–294.
- Tylianakis, J. M. et al. 2008. Global change and species interactions in terrestrial ecosystems. – *Ecol. Lett.* 11: 1351–1363.
- USGS 2012. Federal standards and procedures for the National Watershed Boundary Dataset (WBD): techniques and methods 11-A3. – U.S. Geological Survey and U.S. Dept of Agriculture, Natural Resources Conservation Service, Reston, VA.
- Vannote, R. L. et al. 1980. The river continuum concept. – *Can. J. Fish. Aquat. Sci.* 37: 130–137.
- Webb, C. T. et al. 2010. A structured and dynamic framework to advance traits-based theory and prediction in ecology. – *Ecol. Lett.* 13: 267–283.
- Wickham, H. 2009. *ggplot2: elegant graphics for data analysis*. – Springer.
- Wiens, J. J. 2011. The niche, biogeography and species interactions. – *Phil. Trans. R. Soc. B* 366: 2336–2350.
- Wiley, D. J. et al. 2004. Relations between physical habitat and American eel abundance in five river basins in Maryland. – *Trans. Am. Fish. Soc.* 133: 515–526.
- Winemiller, K. O. et al. 2015. Functional traits, convergent evolution, and periodic tables of niches. – *Ecol. Lett.* 18: 737–751.
- Wirth, T. and Bernatchez, L. 2003. Decline of North Atlantic eels: a fatal synergy? – *Proc. R. Soc. B* 270: 681–688.
- Wisz, M. S. et al. 2013. The role of biotic interactions in shaping distributions and realized assemblages of species: implications for species distribution modelling. – *Biol. Rev.* 88: 15–30.
- Woods, T. and McGarvey, D. J. 2017. Data and code from: Assessing the relative influences of abiotic and biotic factors on American eel *Anguilla rostrata* distribution using hydrologic, physical habitat, and functional trait data. – Figshare Digital Repository, <<https://doi.org/10.6084/m9.figshare.5481205.v4>>.
- Wootton, T. 1994. Putting the pieces together: testing the independence of interactions among organisms. – *Ecology* 75: 1544–1551.

Supplementary material (Appendix ECOG-03782 at <www.ecography.org/appendix/ecog-03782>). Appendix 1.