A traits-based approach for prioritizing species for monitoring and surrogacy selection

Brenda M. Pracheil[1],*, **Ryan A. McManamay[1]**, **Mark S. Bevelhimer[1]**,
Chris R. DeRolph[1], **Glenn F. Čada[1],****

[1]**Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA**

ABSTRACT: The bar for justifying the use of vertebrate animals for study is being increasingly raised, thus requiring increased rigor for species selection and study design. Although we have power analyses to provide quantitative backing for the numbers of organisms used, quantitative backing for selection of study species is not frequently employed. This can be especially important when measuring the impacts of ecosystem alteration, when study species must be chosen that are both sensitive to the alteration and of sufficient abundance for study. Just as important is providing justification for designation of surrogate species for study, especially when the species of interest is rare or of conservation concern and selection of an appropriate surrogate can have legal implications. In this study, we use a combination of GIS, a fish traits database and multivariate statistical analyses to quantitatively prioritize species for study and to determine potential study surrogate species. We provide two case studies to illustrate our quantitative, traits-based approach for designating study species and surrogate species. In the first case study, we select broadly representative fish species to understand the effects of turbine passage on adult fishes based on traits that suggest sensitivity to turbine passage. In our second case study, we present a framework for selecting a surrogate species for an endangered species. We suggest that our traits-based framework can provide quantitative backing and added justification to selection of study species while expanding the inference space of study results.

KEY WORDS: Conservation · Life history traits · Monitoring · Threatened fish

INTRODUCTION

The standard for justifying the use of vertebrate animals for monitoring and study is being increasingly raised for researchers requiring increased rigor in designing studies. Because ecosystems may contain large numbers of species, and studying every species is not possible, the selection and prioritization of study species or target species for monitoring is critically important. Ideally, in selecting species for either of these purposes, the goal is to select the best umbrella species, that is, species that are representative of the largest number of taxa. The umbrella concept has been widely used in conservation planning

and monitoring, for instance, where study or monitoring of a small number of broadly representative species makes multi-species protection and assessment tractable (Fleishman et al. 2000, Hierl et al. 2008, Regan et al. 2008, Tucker et al. 2012). The idea here is that through using a combination of presumed sensitivity of species to disturbance and species occurrence and abundance records, conservation efforts applied to one or a small number of species will also provide the largest protective umbrella for conservation of non-target species (Roberge & Angelstam 2004). The umbrella species concept may also be a useful construct for understanding the species-level effects of ecosystem alteration; that is, the umbrella

*Corresponding author: pracheilbm@ornl.gov
**Retired

species concept can be extended to select one or a tractably small subset of species for study that allows for an understanding of ecological effects of perturbations, stressors or management actions. Although there is some debate as to whether the umbrella species construct is a useful or even a possible one (Andelman & Fagan 2000, Rubinoff 2001, Caro 2003), some studies have suggested that the umbrella concept can aid in conservation planning (Bell et al. 2015). A variety of species have been shown to serve as useful umbrella species, such as coho salmon *Oncorhyncus kisutch*, which serves as an umbrella for protection of aquatic invertebrate species due to similarities in habitat requirements (Branton & Richardson 2014), and sage grouse *Centrocerus urophasianus*, which serves as an umbrella for migratory mule deer *Odocoileus hemionus* due to a similar reliance on sagebrush *Artemesia* spp. (Copeland et al. 2014).

Determining the best ways to incorporate multiple data types related to species abundance, distribution, and physical and behavioral traits that can facilitate the selection of umbrella species or species that will serve as a surrogate for another species (hereafter, surrogate species) for detailed study is challenging. The dawn of the big data era has increased the public availability of species informatics databases such as the Global Invasive Species Database (GISD), Biodiversity Informatics Serving Our Nation Database (BISON), International Union for Conservation of Nature Red List Database (IUCN) and NatureServe, among many others. These databases often contain large amounts of information ranging from georeferenced species occurrence and abundance data to species trait data that can be used to inform selection of umbrella and surrogate species. Selection of appropriate species can be especially important when measuring the impacts of ecosystem alteration, when umbrella species must be chosen that are sensitive to the alteration, of sufficient abundance for study and representative of many other species in their responses (e.g. surrogate species; Sattler et al. 2014). Just as important is providing justification for designation of these species for study, especially when the species of interest is rare or of conservation concern. The designated umbrella or surrogate species in these cases will be used to guide regulation and conservation actions taken for the species of concern, making defensible justification for using a particular species of the utmost importance.

Although power analyses are available to provide quantitative justification for the numbers of organisms used in an experiment, such techniques are not as frequently employed for the selection of umbrella and surrogate species. There has been a variety of power analyses with different objectives reported in the literature (e.g. genetics: Cornuet & Luikart 1996; fisheries biology: Mathur et al. 1996; and social, behavioral and biomedical sciences: Faul et al. 2007), but few studies have focused on quantitative methods of species selection using species traits. Sattler et al. (2014) used a combination of species traits and biological diversity to quantitatively select umbrella species for ecosystem conservation and monitoring; however, their method requires species abundance data, which may not make it appropriate for all species-monitoring and surrogacy-selection study designs.

In this study, we use a combination of geographic information system (GIS)-based species distribution data, species traits and multivariate statistical analyses to step through 2 examples of a new, quantitative, traits-based approach for designating study species and surrogate species. As an example of the use of species informatics databases for community or ecosystem assessment and conservation, we offer a case study that focuses on incorporating fish species traits that indicate sensitivity to hydropower entrainment, that is, the likelihood of downstream movement of fish through hydroelectric turbines, and the susceptibility of these fish to turbine passage stressors. Artificial upstream and downstream fish passage measures are often employed to circumvent the dams and to ameliorate the impacts on fish populations of ecosystem fragmentation. Nonetheless, fish frequently pass through hydropower turbines, where they may encounter a variety of stressors such as extreme turbulence and shear forces, rapid pressure drops, and mechanical strike from turbine blades and other structures (Čada 2001). These stressors cause injuries and mortalities that can exceed 30% for some species (Pracheil et al. 2016). Incorporation of fish trait information into hydropower design and scheduling requires the understanding dose–response relationships between turbine stressors and fish injuries, which must be experimentally determined. Conducting experiments is labor- and resource-intensive, making it impossible to determine quantitative dose–response relationships for all species. As a result, our goal is to choose a few species that are broadly representative of many species – a selection of umbrella species for experimental studies – and whose traits are similar to many other species such that we can maximize our inferences about the effects of new turbine designs and operational strategies. Our case study stems from a study initiated by the US Department of Energy Hydropower Program, which seeks to incorporate biological information about fish into

the design of new hydropower turbines and the development of hydropower operating schemes with lower environmental impacts.

Moreover, there are situations where some species of interest are unavailable for study for logistical, permitting or other reasons (e.g. rare or protected species). In those cases, a quantitative approach to selecting a surrogate species for study is needed. In our second case study, we build on our first case study and present a framework for selecting a surrogate species for an endangered species. In particular, this case study uses a traits-based approach to identify a surrogate species for studying the effects of downstream passage in lieu of monitoring a protected species, the US federally endangered pallid sturgeon *Scaphirhynchus albus*. We selected pallid sturgeon for this analysis because there has been extensive documentation of surrogate study species for monitoring and experimentation of pallid sturgeon that allows us to evaluate how well our approach compares with existing recommendations of surrogate species (e.g. Quist et al. 2004) that are not based on explicitly quantitative, traits-based approaches. We suggest that our traits-based framework can provide quantitative backing and added justification to selection of study species while expanding the inference space of study results.

MATERIALS AND METHODS

Both our species prioritization and species surrogacy selection methods rely on similar approaches and incorporate the traits-based analysis framework presented by Laliberté & Legendre (2010) for a distance-based measure of trait functional diversity. Our specific case studies are based on the traits-based framework outlined by Čada & Schweizer (2012) for determining fish species' sensitivity to turbine passage. In brief, the methods we present here for monitoring species prioritization and surrogate species selection use adult fish life-history traits that indicate sensitivity or resilience to a particular perturbation, in this case to turbine passage. We use downstream hydropower passage stressors as inputs and then statistically group species based on these traits (Fig. 1).

Both case study analyses were broadly interested in adult North American freshwater fish species and included all freshwater, anadromous and catadromous fish species from the NatureServe Explorer Freshwater Fish of North America database (NatureServe 2015). We then matched the fish distribution data with fish life history, morphological and trophic trait infor-
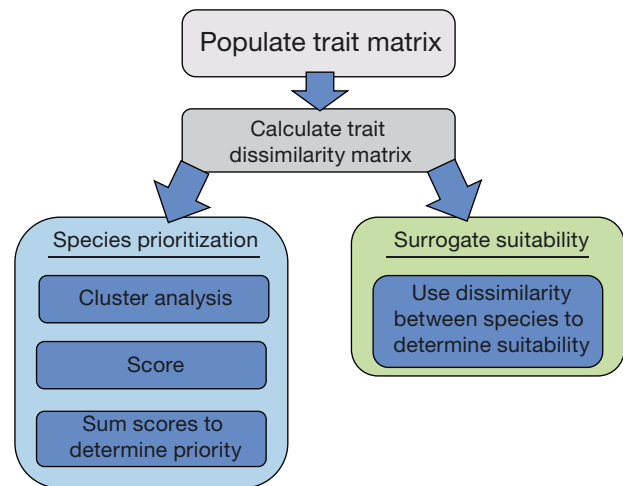

Fig. 1. Workflow for traits-based species prioritization and surrogate suitability analyses

mation for adult fishes from across the USA using 2 fish traits databases (Frimpong & Angermeier 2009, McManamay & Frimpong 2015) and identified traits indicating: (1) whether a species was vulnerable to entrainment (e.g. entrainment vulnerability; Table 1 and see Table S1 in Supplement 1 at www.int-res.com/articles/suppl/n031p243_supp1.pdf); (2) whether

Table 1. Hydropower turbine entrainment vulnerability from *k*-means clustering (*k* = 8). Entrainment risk ranking ranks groups from least vulnerable to turbine entrainment (1) to most vulnerable to turbine entrainment (8) based on expert opinion and published entrainment studies. Trait description provides the general description of the types of life-history traits included in groups identified from the cluster analysis. This dissimilarity matrix was constructed using fish traits and included water-column position (pelagic or not), habitat (river, lake, river and lake), mobility (migratory or not) and current velocity of habitat (slow, moderate or fast)

Entrainment risk ranking	Trait description
1	Benthic, river or lake, non-migratory, moderate-fast current
2	Benthic, river, non-migratory, slow-moderate current
3	Pelagic, river and lake, migratory, slow-moderate current
4	Big, benthic, river migrants
5	benthic, river, moderate-fast current
6	Benthic, river, fast current
7	Benthic, river or lake, non-migratory, slow current
8	Pelagic, river and lake habitat, migratory, moderate current

a species was vulnerable to injury during turbine passage (e.g. entrainment injury vulnerability; Table 2 and Table S2 in Supplement 1); and (3) the relative exposure of a species to turbines across its geographic range (e.g. turbine exposure). Relative exposure of a species to hydroelectric turbines across its range was determined by intersecting the National Hydropower Asset Assessment Program database — a spatial database containing location and plant attribute information such as turbine type and plant nameplate capacity for hydropower facilities in the USA (Hadjerioua et al. 2011) — and determining the number of hydropower turbines of all types across a species range. Our analyses included 450 extant North American freshwater fish species for which we could obtain both current distributional and traits information.

Our first case study, the species prioritization analysis, additionally incorporated species conservation status (e.g. secure, threatened, endangered; Table 3, see Table S3 in Supplement 1) using range reduction information from the Fish Traits Database and conservation status information from NatureServe. We did not include this information in the surrogacy analysis because the need for selection of surrogate species may be due to the fact a species of interest is too rare to study either in the field or in the laboratory. By incorporating population status into the analyses,

the method would be biased toward selecting species with similar population status. In other words, species of conservation concern may be shown as similar to other difficult-to-study species that may themselves be too rare to serve as a useful surrogate.

Case study 1: study species prioritization

We calculated functional trait diversity for each category of stressor using Gower dissimilarities (Gower 1971) for all species pairs based on the fish traits included in our analysis (see Supplement 1) using the R package Cluster (Maechler et al. 2014). Gower dissimilarities incorporate a mix of binary, continuous and discrete data in calculating the dissimilarity scores. We then used a Ward's agglomerative cluster-

Table 2. Hydropower turbine injury vulnerability from k-means clustering ($k = 8$). The injury risk ranking ranks taxa from least vulnerable to turbine injury (1) to most vulnerable to turbine injury (8) based on expert opinion and published turbine mortality rates. 'Taxa included' provides a general description of the types of fishes included in groups identified from the cluster analysis, but is not an all-encompassing description. This dissimilarity matrix was constructed using fish morphological traits and included maximum total length, scale type (scaleless, cycloid, ctenoid, ganoid, scutes) and swim bladder type (physoclistous or physostomous)

Injury risk ranking	Taxa included
1	*Scaphirhynchus* spp. and shortnose sturgeon
2	Small gar (e.g. shortnose, spotted)
3	Small catfishes (i.e. bullhead, madtom), lamprey
4	Cyprinids, catostomids, esocids, salmonids
5	All other sturgeons
6	Percids, centrarchids, mosquitofish
7	Giant gar, e.g. alligator and longnose
8	Big scaleless fish including paddlefish, blue catfish and sea lamprey

Table 3. Population vulnerability from k-means clustering ($k = 8$). The threat score is a ranking of species groups from least vulnerable to ecosystem alteration (1) to most vulnerable to ecosystem alteration (8). Trait description provides the general description of the types of species included in groups identified from the cluster analysis. This dissimilarity matrix was constructed using fish traits and species conservation statuses, and included habitat range reduction (minimal or significant), overexploitation (minimal or significant), anthropogenic threats (minimal or significant), range size (widespread or endemic), NatureServe conservation status, US Endangered Species Act status and International Union for the Conservation of Nature conservation status

Population risk	Description
1	Mix of game and non-game species; few species of concern. Opportunistic and periodic reproductive strategists
2	Mostly non-game fish, except catfish. Equilibrium reproductive strategists
3	Species of concern including alligator gar, blue sucker, paddlefish and green sturgeon. Periodic reproductive strategists
4	Desert sucker. Equilibrium reproductive strategist
5	Cyprinids, darters and shads, some of conservation concern. Opportunistic reproductive strategists.
6	Clupeids, cyprinids, lamprey of conservation concern, many non-game fish not of conservation concern. Equilibrium reproductive strategists.
7	Cyprinids and darters of conservation concern. Opportunistic reproductive strategists
8	Endangered sucker, sturgeon and chub. Periodic reproductive strategists

ing method (R base 3.0.2, 'Frisbee Sailing'; R Development Core Team 2014) to cluster species together based on Gower dissimilarities for each stressor category. Each cluster was defined as species with ≤20% dissimilarity to each other, which resulted in 8 clusters for entrainment vulnerability, entrainment injury and population vulnerability stressor categories. Each stressor category in our analysis was defined by 8 clusters (*k*-means clustering; *k* = 8), although this was a matter of chance based on dissimilarities among species rather than design.

After grouping species into clusters, ranks were assigned to each cluster by each of 4 of the authors (B.M.P., R.A.M., M.S.B., G.F.C.; hereafter labeled as A–D, in no particular order) based on relative vulnerability, where clusters were scored from 1 to 8 in order of increasing vulnerability. Ranks were determined based on turbine mortality rates presented in Pracheil et al. (2016) in addition to morphological, habitat and life history information associated with fish in each of the clusters. For example, the cluster containing small gar (i.e. shorter maximum total length [TL] as adults, specifically, longnose, shortnose and spotted gar) and sturgeons (i.e. *Scaphirhynchus* spp.) were ranked as resistant to downstream turbine passage because their heavy scales and integument presumably protect them from many injuries suffered during passage. In contrast, the cluster containing centrarchids (sunfishes and black bass) was ranked as vulnerable to injury because of high mortality rates in past entrainment studies (e.g. Pracheil et al. 2016). Ranks were then averaged across authors and each rank was assigned the nearest integer value, allowing for ties. We assigned all fish species within a cluster the rank of that cluster.

We then calculated cumulative prioritization score for each species by summing the rank values from each stressor category with a standardized turbine passage exposure score T_E, which was calculated as:

$$T_E = T_R/T_{max} \times C \qquad (1)$$

where T_R is the number of turbines within a species' range, T_{max} is the maximum number of turbines reported within the range of any species in this study, and C is the standardizing coefficient (in this case, C = 8), because we wanted the weight of the turbine score to be equal to that of the other individual stressors, which were ranked from 1 to 8. Finally, we ranked species using these cumulative prioritization scores from highest to lowest, where species with the highest cumulative scores were determined to be the highest priority for study.

Although our method for using a traits-based approach for selecting study species provides quantitative backing for species selection, this prioritization approach has several components that rely on expert opinion, which can be somewhat subjective. Results of the prioritization exercise will, no doubt, be sensitive to these subjectivities. We provide some examples of how the subjective portions of our analyses can influence prioritization results on the upper 5% of species. The upper 5% of species were designated as those with the highest species scores, where the species scores were rounded down to the nearest whole integer. In cases where the 5% cut-off fell on species that had tied scores, all species with that score were included in the upper 5%.

We explicitly examined where and how variation in expert opinion influenced results in 4 ways. (1) We examined the sensitivity of species prioritization to expert opinion by comparing species lists obtained by the averaged expert opinion method we described above with species lists obtained using cluster rankings designated by each expert. We quantitatively compared differences in species lists using Jaccard's coefficient of community similarity — a metric that provides similarity in species composition for the upper 5% of species irrespective of ranking, where a coefficient of 1 indicates identical species in lists and a coefficient of 0 indicates no species in common between lists — for each pair of experts. (2) We examined sensitivity of the species lists to stressor categories chosen by conducting simple linear regressions between scores for each stressor category and species rank using the expert opinion averaged results. (3) We examined how sensitive overall species rank was to individual expert rankings and averaged expert rankings using a Pearson's product-moment correlation coefficient. These analyses correlated species scores from the entrainment vulnerability stressor category with averaged species rank calculated from all stressor categories. (4) We also used Pearson's product-moment correlation coefficients to test the sensitivity of the results to using raw species scores and ranks by correlating raw scores and ranks with averaged scores and ranks, respectively.

Case study 2: surrogacy analysis

For the species surrogacy analysis, we incorporated all traits from the entrainment vulnerability and entrainment injury vulnerability categories into a singular matrix and calculated Gower dissimilarities between pallid sturgeon and each of the 170 fish spe-

cies with ranges sympatric to that of the pallid sturgeon. Species with dissimilarity scores of 0 were perfectly not dissimilar (i.e. functionally identical) based on the traits matrix, indicating these species would be the best possible surrogates. Species with dissimilarity scores of 1 were perfectly dissimilar based on the traits matrix and would be the worst possible surrogates. We presented the lowest decile of species in addition to using multidimensional scaling (MDS) to display how species were related to each other based on the dissimilarity scores, where species pairs with lower dissimilarity scores were plotted closer together and species with higher dissimilarity scores were plotted farther apart (cmdscale; R base).

Spatial distribution of hydropower risk

To better understand where the greatest risk of hydropower turbines to fishes occur across the USA, according to our metrics, we calculated the turbine risk score for each watershed, R_w, as:

$$R_w = \sum T_E / S_w(N_T) \qquad (2)$$

where T_E is the standardized turbine passage exposure score calculated for a species (from Eq. 1) and S_w is the number of species in a watershed w multiplied by the number of turbines in a watershed, N_T. We then categorized watersheds into 1 of 7 equal percentiles such that each category equaled approximately 14.3% of the data, and mapped the watersheds using ArcGIS 10.3.1.

RESULTS

Because clusters for generating turbine entrainment injury vulnerability, vulnerability to entrainment and population vulnerability were informed using species traits; because traits such as habitat use, length and conservation status can vary within a fish taxon, clusters of fishes are not necessarily representative of phylogeny. In fact, only the hydropower turbine injury vulnerability traits formed clusters that resembled phylogeny, although only loosely so (Table 2). For instance, some of the primitive fishes, such as the *Scaphirynchus* spp. sturgeons and the smallest of the gar species, clustered together in our analyses based on similarities in characteristics of dermal armoring, scales and maximum total length. The other cluster analyses for hydropower entrainment vulnerability (Table 1) and population vulnerability (Table 3) were not generally structured in concordance with phylogeny.

Case study 1: species prioritization and selection

Supplement 2 provides expert rankings assigned to clusters and species scores (Tables S4–S7 at www.int-res.com/articles/suppl/n031p243_supp2.xlsx) used to inform prioritizations. The upper 5% of species identified in our traits-based species selection for selecting broadly representative species sensitive to hydropower system stressors had large species ranges (Fig. 2). Most of these fishes have current native ranges in the eastern portion of the USA, in particular, ranges that overlap with the high densities of hydropower facilities in the northeast and southeast (Fig. 2). Nearly all the species in the upper 5% list (Table 4) were not of any conservation concern, with the exception of the paddlefish *Polyodon spathula*, a species that was listed by the IUCN Red List as Vulnerable. None of the upper 5% species were listed as threatened or endangered by the US Endangered Species Act.

Analyses of sensitivity of our prioritization methods showed that, while we did find that some risk categories had higher correlations between average cluster score and species rank (e.g. some risk categories had r-values more than double those of other risk categories; Fig. 3A), when looking at an individual stressor category, correlations between individual expert assignments and final species prioritizations were more homogeneous; further, in cases where 4 experts were involved, very different rankings by 1 expert did not strongly influence the averaged opinion of the group (Fig. 3B). We found that relationships between individual species scores assigned by experts and composite species scores obtained using ranks from all experts were similar across experts, but once expert and composite scores were placed in ascending order and ranked, expert ranks and composite ranks were uncorrelated (Fig. 4). Concordance in top-20 species prioritizations varied between experts (Jaccard's coefficient of community similarity: 0.0566–0.2258; Table 5) and also between experts and the averaged species prioritizations (Jaccard's coefficient of community similarity: 0.0566–0.5200; Table 5).

Case study 2: designation of closely related surrogate species

Although the Gower dissimilarity matrix generated with our traits-based approach contains all pair-wise combinations of all species used in this analysis, we only present results for determining study surrogate
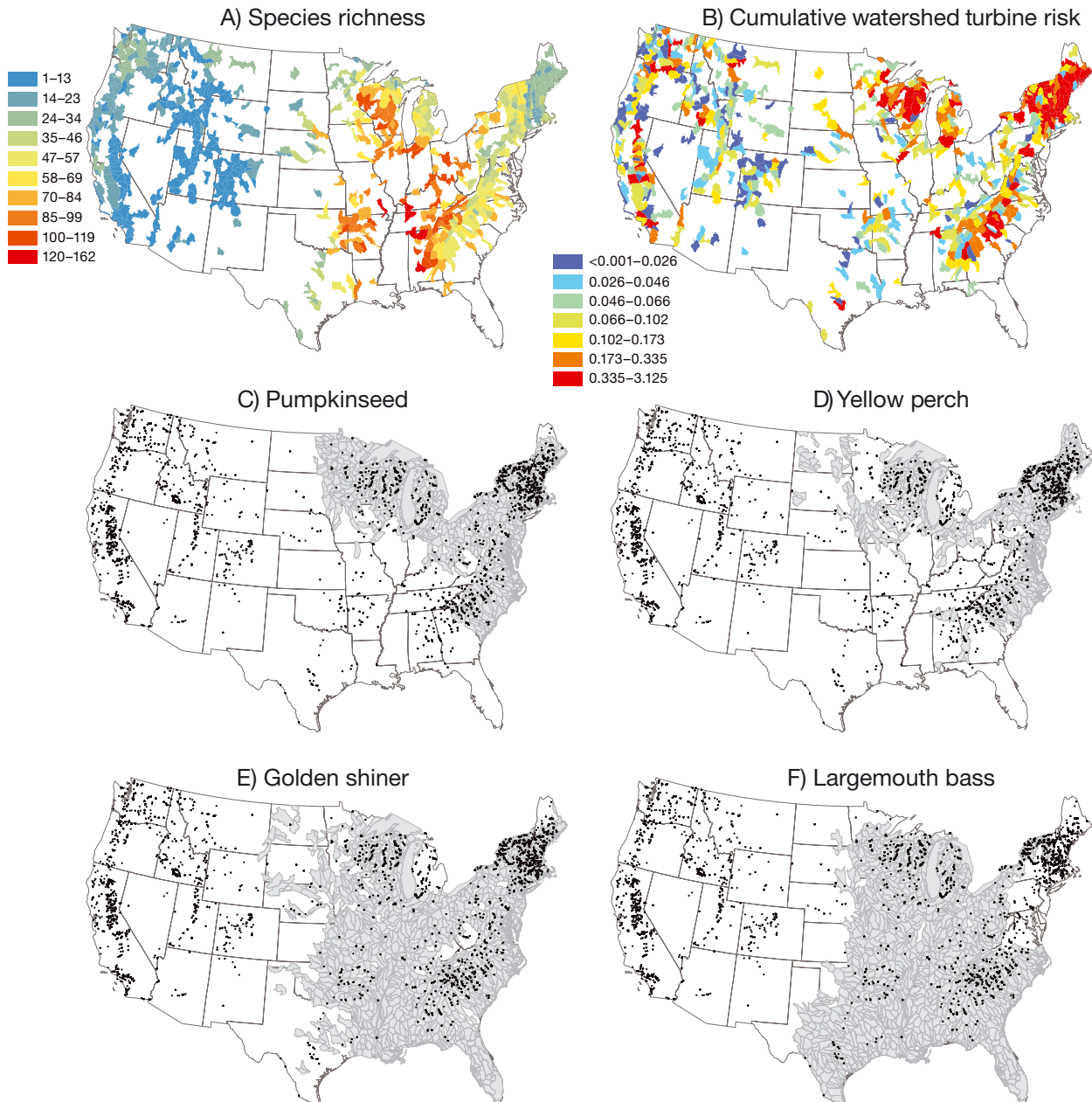
Fig. 2. Maps displaying fish (A) species richness (total number of fish species) in watersheds containing hydropower facilities (B) rarity-weighted turbine injury and mortality risk to fishes within a watershed containing hydropower turbines, and (C–F) distribution of selected fish species (shaded areas) from the top 5% priority species for assessing hydropower turbine-induced injury and mortality by watershed and occurrence of hydropower turbines (black dots)

species when examining the impacts of downstream turbine passage for the federally endangered pallid sturgeon *S. albus*. Restricting this example to just one species allowed for ease of illustrating how to use our approach and interpret the results. The lower decile of dissimilarity scores (i.e. the species that are the most representative of pallid sturgeon) contained 17 species: 9 cyprinids, 4 catostomids, 2 sturgeon, 1 centrarchid and 1 percid (Table 6). The blue sucker

Cycleptus elongatus—an obligate, large river catostomid—had the lowest dissimilarity score with pallid sturgeon, closely followed by the shovelnose sturgeon *S. platorynchus*, a congener of the pallid sturgeon. Although 2 additional sturgeon species have species ranges sympatric with the pallid sturgeon—lake sturgeon *Acipenser fulvescens* and Atlantic sturgeon *A. oxyrinchus*—only the Atlantic sturgeon was listed in the lower decile of dissimilar-

Table 4. Upper 5% of umbrella species traits-based prioritization. These are the upper 5% of species that are most sensitive to turbine stressors while still being broadly representative of other fish species. Cumulative score shows the risk score calculated as the sum of entrainment vulnerability rank, entrainment injury rank and turbine exposure score for each species. No. turbines: number of turbines in each species' range

Rank	Common name	Scientific name	No. turbines	Cumulative score
1	Brook trout	*Salvelinus fontinalis*	2339	112
2	Yellow perch	*Perca flavescens*	2513	104
3	Pumpkinseed	*Lepomis gibbosus*	2857	84
4	White sucker	*Catostomus commersonii*	3100	72
5	Golden shiner	*Notemigonus crysoleucas*	3072	66
6	Largemouth bass	*Micropterus salmoides*	2071	65
7	Common shiner	*Luxilus cornutus*	2498	64
8	Walleye	*Sander vitreus*	1560	60
9	Redbreast sunfish	*Lepomis auritus*	1616	56
10	Spottail shiner	*Notropis hudsonius*	1755	56
11	Channel catfish	*Ictalurus punctatus*	1541	54
12	Banded killifish	*Fundulus diaphanus*	1803	52
13	Bluegill	*Lepomis macrochirus*	1802	52
14	Black crappie	*Pomoxis nigromaculatus*	1585	52
15	Longnose sucker	*Catostomus catostomus*	1572	52
16	Rockbass	*Ambloplites rupestris*	1608	48
17	Fathead minnow	*Pimephales promelas*	1703	48
18	Lake sturgeon	*Acipenser fulvescens*	1216	48
19	Gizzard shad	*Dorosoma cepedianum*	1810	44
20	Chain pickerel	*Esox niger*	1684	44
21	Redfin pickerel	*Esox americanus*	1568	44

ity scores. The lake sturgeon has a much higher dissimilarity score of 0.3812 (Fig. 5).

The MDS analysis using the Gower dissimilarity matrix depicts a graphical approximation of the dissimilarities of all fish species used in the analysis in relation to pallid sturgeon (Figs. 3 & 4). Interestingly, other primitive fishes sometimes used as surrogates for sturgeon species, such as the lake sturgeon and the closely related paddlefish *Polyodon spathula*, had relatively high dissimilarity scores with the pallid sturgeon.

Spatial distribution of hydropower risk

The Northeast, Great Lakes and Southeast USA comprised the areas of highest turbine-weighted mean watershed hydropower risk (Fig. 2B). There were also several watersheds
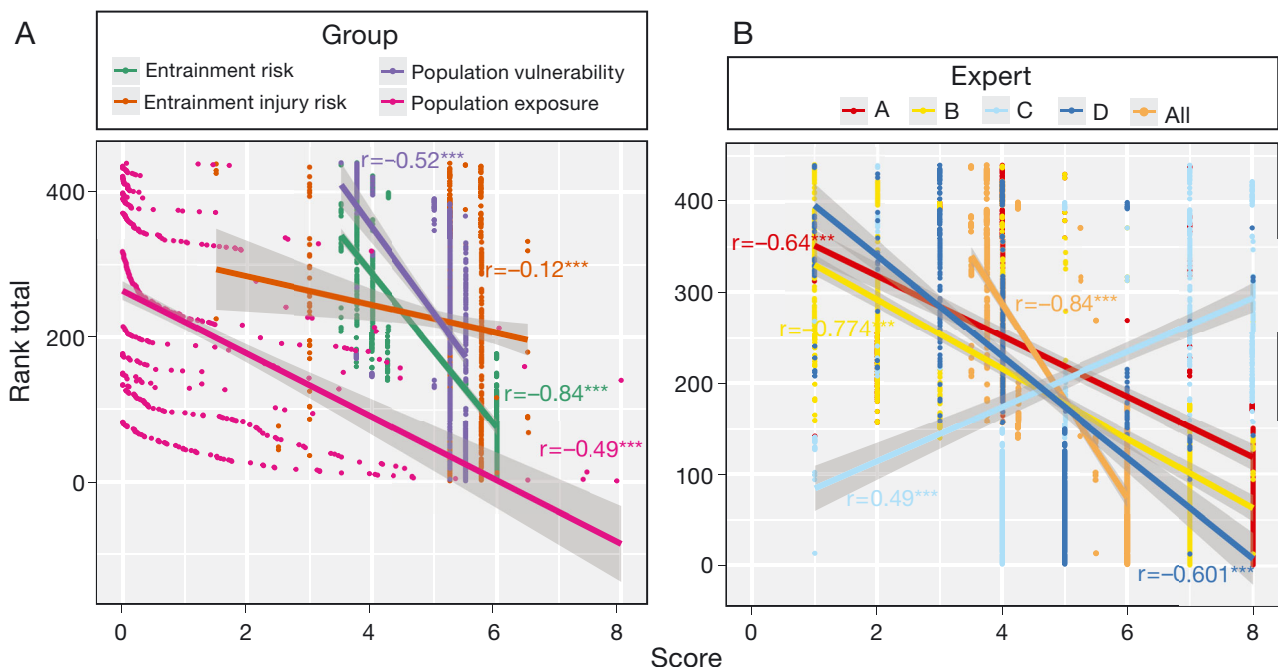

Fig. 3. Relationship between species rank (top panel; rank 1 = highest priority) and average risk group score obtained from all experts with Pearson's product moment correlation coefficients for the relationship between species rank and average risk group score. In panel (A), color of line, dots and r-value correspond to risk group as shown in key. In panel (B), color of line, dots and r-value correspond to risk score assigned to each species by each expert (denoted as A–D) and averaged from all experts (denoted as All) based on entrainment risk. In both panels, gray shaded regions surrounding linear best-fit lines show 95% confidence intervals. ***p < 0.001
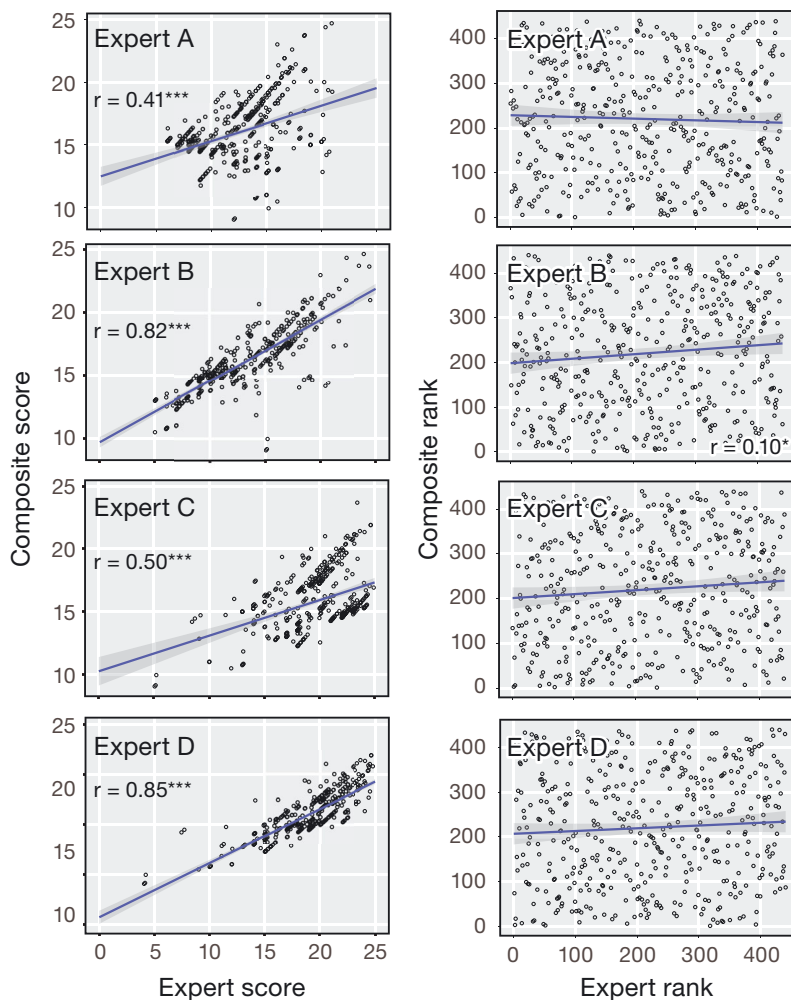
Fig. 4. Left panels: relationships between species score (highest score = highest priority) averaged from all experts and score based on each individual expert's input. Right panels: relationships between species rank (highest rank = highest priority) derived from score for all experts and rank derived from scores based on each individual expert's input. Expert A–D denotes which expert's input was used to derive the relationship. Pearson's product moment correlation coefficients are shown where correlations are significant at the $\alpha = 0.05$ level (***p < 0.001). In all panels, gray shaded regions surrounding the linear best-fit line show 95 % confidence intervals

Table 5. Jaccard's coefficient of community similarity for 20 fish species with the highest composite scores between each pair of experts. A coefficient of 1 indicates 2 experts had all of the same species in their lists of 20 species, and a coefficient of 0 indicates 2 experts had no species in common between their lists. Experts are designated A, B, C, D

	A	B	C	D	Average
A	1	0.1875	0.0556	0.0556	0.0556
B		1	0.1515	0.2258	0.5200
C			1	0.2258	0.1515
D				1	0.2667
Average					1

in the Pacific Northwest with high risk scores. Again, map legend categories were broken into 7 equal percentiles for ease of display, so although some categories have a large spread of values, they represent the range in a particular risk category.

DISCUSSION

Trait-based approaches have a long history rooted in community ecology and have been shown to be powerful for improving our understanding of biological communities as drivers of ecosystem function, process and turnover (McGill et al. 2006). Our approach for using a traits-based framework for defining broadly representative species for study and monitoring capitalizes on the power of quantitative traits-based approaches (Laliberté & Legendre 2010), further extending their utility to species conservation and management. Publically available traits databases such as those shown in Table 7 are becoming more common for many different types of species, making our approach also useful for prioritizing species for monitoring or designating appropriate surrogate species for nonfish taxa. Moreover, because our framework is flexible and non-specific to life stage, taxa or ecosystem, it can accommodate data from existing traits databases, new traits databases as they become available, or additional traits measured by researchers for any number of taxa, in any geographic area, and can even accommodate demographically specific traits. We did not weight any of the traits in our analyses and, conveniently, each measure of hydropower vulnerability separated into 8 clusters, making all traits within a hydropower vulnerability group (i.e. injury vulnerability, conservation status, etc.) equally weighted.

Although the framework we present here is quite flexible, it is important to note that our approach does have some limitations, and the specific case study results presented here are applicable only to the fish species and life stages we included in our analysis and only in rivers altered by hydropower. That is, we

Table 6. Lower 10% of traits-based surrogate species analysis for the endangered pallid sturgeon *Scaphirhynchus albus*. Rankings are based on Gower dissimilarities of fish species traits for species with ranges that overlap with that of pallid sturgeon. Lower dissimilarity scores mean that fish have traits that are most like those of the pallid sturgeon. The range overlap column shows the percentage of the pallid sturgeon range that is sympatric with each species, although it was not considered in our analysis

Rank	Common name	Scientific name	Dissimilarity	Range overlap (%)
1	Blue sucker	*Cycleptus elongatus*	0.1121	88
2	Shovelnose sturgeon	*Scaphirhynchus platorynchus*	0.1243	66
3	Blacktail redhorse	*Moxostoma poecilurum*	0.1879	8
4	Sicklefin chub	*Macrhybopsis meeki*	0.1986	60
5	Sturgeon chub	*Macrhybopsis gelida*	0.2009	71
6	Shoal chub	*Macrhybopsis hyostoma*	0.2013	32
7	Gravel chub	*Erimystax x-punctatus*	0.2017	6
8	River redhorse	*Moxostoma carinatum*	0.2083	6
9	Suckermouth minnow	*Phenacobius mirabilis*	0.2095	31
10	Silverjaw minnow	*Notropis buccatus*	0.2112	5
11	Steelcolor shiner	*Cyprinella whipplei*	0.2117	9
12	Bigeye chub	*Hybopsis amblops*	0.2156	3
13	Highfin carpsucker	*Carpiodes velifer*	0.2158	32
14	Atlantic sturgeon	*Acipenser oxyrinchus*	0.2497	4
15	Sand shiner	*Notropis stramineus*	0.2522	58
16	Shadow bass	*Ambloplites ariommus*	0.2651	3
17	Stargazing darter	*Percina uranidea*	0.2698	3

selected traits that would indicate susceptibility or resistance of a fish species to downstream passage through hydropower turbines, and applying the list of species and traits from our analyses to predict species sensitivity to a different set of environmental perturbations or in a different type of ecosystem would not provide meaningful results. For example, if researchers were looking to prioritize which fish species would be the best indicators of climate change impacts, they would not want to conduct this prioritization based on the suite of traits we use in our analyses. Rather, the researchers would want to select traits that may predict sensitivity to changing temperature and precipitation
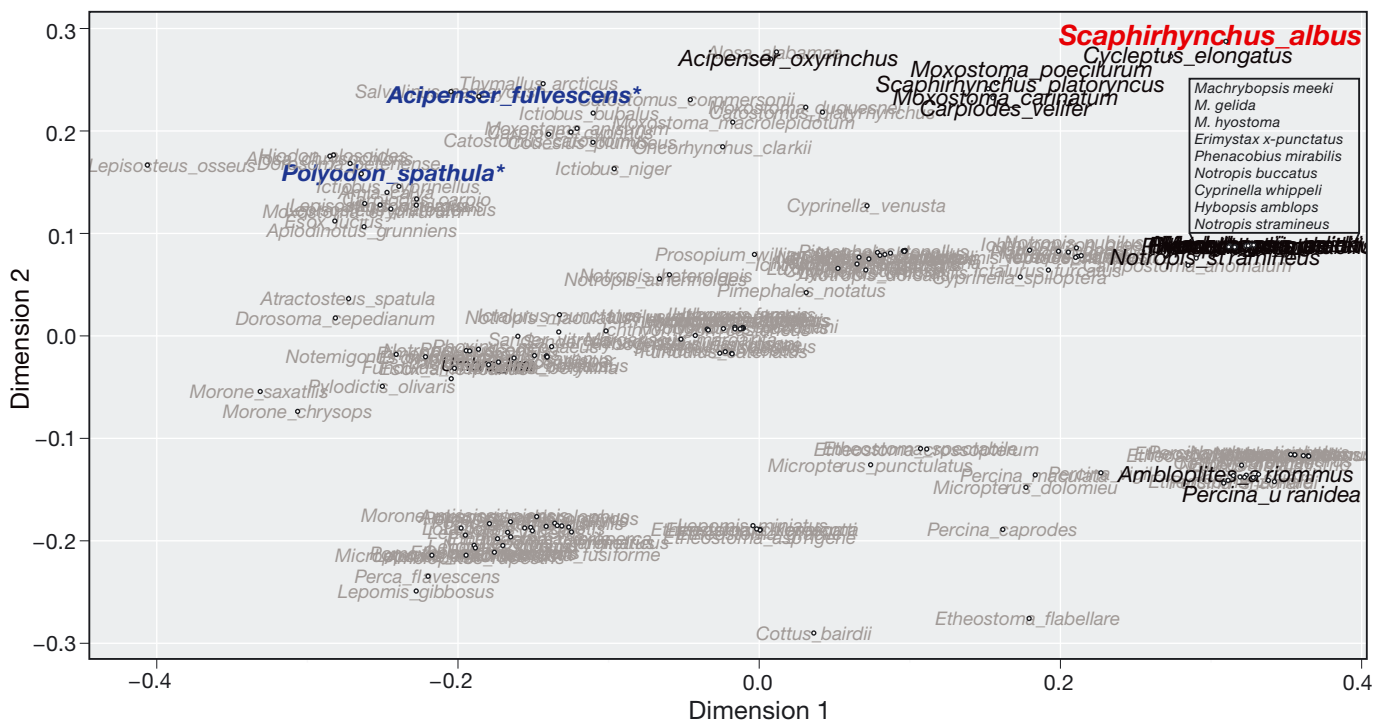

Fig. 5. Multidimensional scaling plot for species included in traits-based surrogacy analyses shown in reference to pallid sturgeon *Scaphirhynchus albus*. Fish species in black italic font are those in the lowest decile of similarity to pallid sturgeon (e.g. the most appropriate surrogates based on hydropower vulnerability traits). Fish species in the inset box are contained in the cluster of species just below the box. Fish species in blue followed by asterisks are closely related species that fall outside of the lowest decile of similar species

Table 7. Selected publically available traits databases for a variety of taxa and spatial extents including the database name, web address, spatial extent, description of traits included in the database, and the database citation

Database	Website	Spatial extent	Traits description	Citation
Plants				
BIOLFLOR	http://www2.ufz.de/biolflor/index.jsp	Germany	Life-history, dispersal and habitat use	Aguiar et al. (2013)
FLOWBASE	www.isa.ulisboa.pt/proj/flowbase/	Global	Morphology, phenology, reproduction and ecology	
TRY (Plant Trait Database)	https://www.try-db.org/TryWeb/Database.php	Global	Morphological, anatomical, biochemical, physiological and/or phenological traits	Varies based on the data used
Invertebrates				
Lotic Invertebrate Traits of North America	http://pubs.er.usgs.gov/publication/ds187	North America	Life-history, habitat use and reproductive strategy	Vieira et al. (2006)
Freshwater Biological Traits Database	www.epa.gov/ncea/global/traits/	North America	Habitat, life-history, mobility, morphology and ecological traits	US EPA (2012)
Fish				
Fish Traits	www.fishtraits.info/	North America	Tropic ecology, body size and reproductive ecology, habitat associations, and temperature and salinity tolerances	Frimpong & Angermeier (2009)
FishBase	www.fishbase.org	Global	Biology, ecology, taxonomy, life history, trophic features, population dynamics, and human uses	Froese & Pauly (2014)
Amphibians				
Life History Traits for European Amphibians	http://bdj.pensoft.net/articles.php?id=4123&display_type=list&element_type=5	Europe	Life history traits for 86 species of European frogs and salamanders	Trochet et al. (2014)
Reptiles				
Reptile Trait Database	http://scales.ckff.si/scaletool/index.php?menu=6&submenu=0	Europe	Behavior, bioenergetics, habitat use, morphology and phenology	Grimm et al. (2014)
Multiple taxa				
Freshwater Ecology Info	www.freshwaterecology.info/	Europe	Distributional, regional, habitat, life history and/or morphology for fish, macroinvertebrates, macrophytes, diatoms and phytoplankton	Schmidt-Kloiber & Hering (2015)
AnAge	http://genomics.senescence.info/species/	Global	Ageing and life history in animals with extensive longevity records	Depends on the data used, see: http://genomics.senescence.info/legal.html#citation
NatureServe Explorer	http://explorer.natureserve.org/	Canada & USA	Current and historic distribution, native and non-native distribution, and conservation status of plants, animals and ecosystems	NatureServe (2015)

patterns such as those associated with phenology, optimal growing conditions or use of habitat vulnerable to climate change impacts. Similarly, studies interested in prioritizing local- or regional-scale species should first restrict the species included in prioritization analyses to those present in the physical inference space for the study, as we did for the species surrogacy analysis. Also, while our approach is not specific to a particular life stage, the results of the specific analyses presented in this paper only apply to adults. The databases we used to populate our traits matrices (Frimpong & Angermeier 2009, McManamay & Frimpong 2015, NatureServe 2015) contained information on such life history traits as length and habitat use for adult fishes. It is therefore possible that using traits of very young or juvenile fishes or a particular age class may yield different results. Unfortunately, comparatively little is known about earlier life stages for most species of fish, and conducting prioritizations that incorporate them on a broad scale would be difficult.

Our species prioritization example deliberately did not prioritize threatened or endangered species. Alternatively, we selected traits for our analyses that would not only prioritize species at-risk to entrainment and injury but also those that are widely distributed so as to be able to draw inferences across a wide array of ecosystems. For example, our method did not contain any threatened or endangered salmonid (e.g. salmon and trout) or anguillid (e.g. eel) species that are often the source of hydropower mitigation requirements. This is due in no small part to our analysis 'rewarding' species that had large ranges and 'penalizing' species with more restricted ranges through incorporation of the number of turbines in a species' range. However, electing to study a species of concern over another species also generally requires little justification and the prioritization analysis we present would not be necessary.

We initially chose these case studies — prioritizing study species for understanding the effects of hydropower on fishes in the USA and choosing study surrogates for inferring effects of ecosystem alterations on a rare, endangered species — because of prior published efforts that used qualitative means to accomplish these objectives. In this way, we can compare previous qualitative assessments with our quantitative assessment. Our attempt to prioritize study species for monitoring, for example, focused on identifying fish species that are the most broadly representative of fish species impacted by hydropower in the USA, potentially also fish species that would be good targets for evaluating low-impact turbines to be installed across the USA. Peer-reviewed studies concur with the findings of our method, demonstrating that the species our method elevates as the highest priority for studying the effects of turbine injury and mortality (Table 4) are either very commonly studied in turbine entrainment studies (Pracheil et al. 2016) or very commonly entrained through hydropower turbines. For example, a systematic review of fish turbine entrainment by Pracheil et al. (2016) tallied multiple studies reporting entrainment of percids (mostly yellow perch *Perca flavescens*), paddlefish and acipenserids such as lake sturgeon *Acipenser fulvescens*. Similarly, Travnichek et al. (1993) and Michaud & Taft (2000) reported sunfishes (such as pumpkinseed *Lepomis gibbosus*), black bass (such as largemouth bass *Micropterus salmoides*) and logperch *Percina caproides* as being among the most common constituents of entrainment catches among hydropower facilities in the southern and midwestern USA. Gizzard shad ranks 19 in our prioritization results, is frequently entrained and is also among the least dissimilar species to American shad — a species that has been substantially impacted by hydropower (Haro et al. 1998, Castro-Santos & Haro 2015).

Interestingly, our surrogacy analysis suggests that phylogeny is not necessarily a good indicator of what species would best serve as surrogates, and reinforces the power of a traits-based approach. This was interesting, although unsurprising because hydropower entrainment vulnerability would not have been a selecting factor during the evolution of fishes, and more directly because clustering was based on traits such as habitat use that do not necessarily map onto phylogeny. In our example, many of the species most closely related to pallid sturgeon, such as sturgeon, paddlefish and gar, are deemed less suitable surrogates than blue sucker and some chubs. This is important because we may intuitively bias our selection of surrogate species towards phylogenetically related species even though there may be more appropriate species based on traits that are sensitive to the perturbation of interest. Paddlefish, for instance, have been named as an indicator of pallid sturgeon spawning success and habitat use (Dryer & Sandvol 1993). Our analysis suggests that they would not be equally sensitive to hydropower system stressors. In fact, neither paddlefish nor lake sturgeon would be good surrogates in this case, likely because of their different habitat uses. For example, pallid sturgeon are found only in rivers while paddlefish and lake sturgeon are both found in rivers and lakes (Fig. 6B). Also, both paddlefish and lake sturgeon were associated with slow current velocities while the pallid
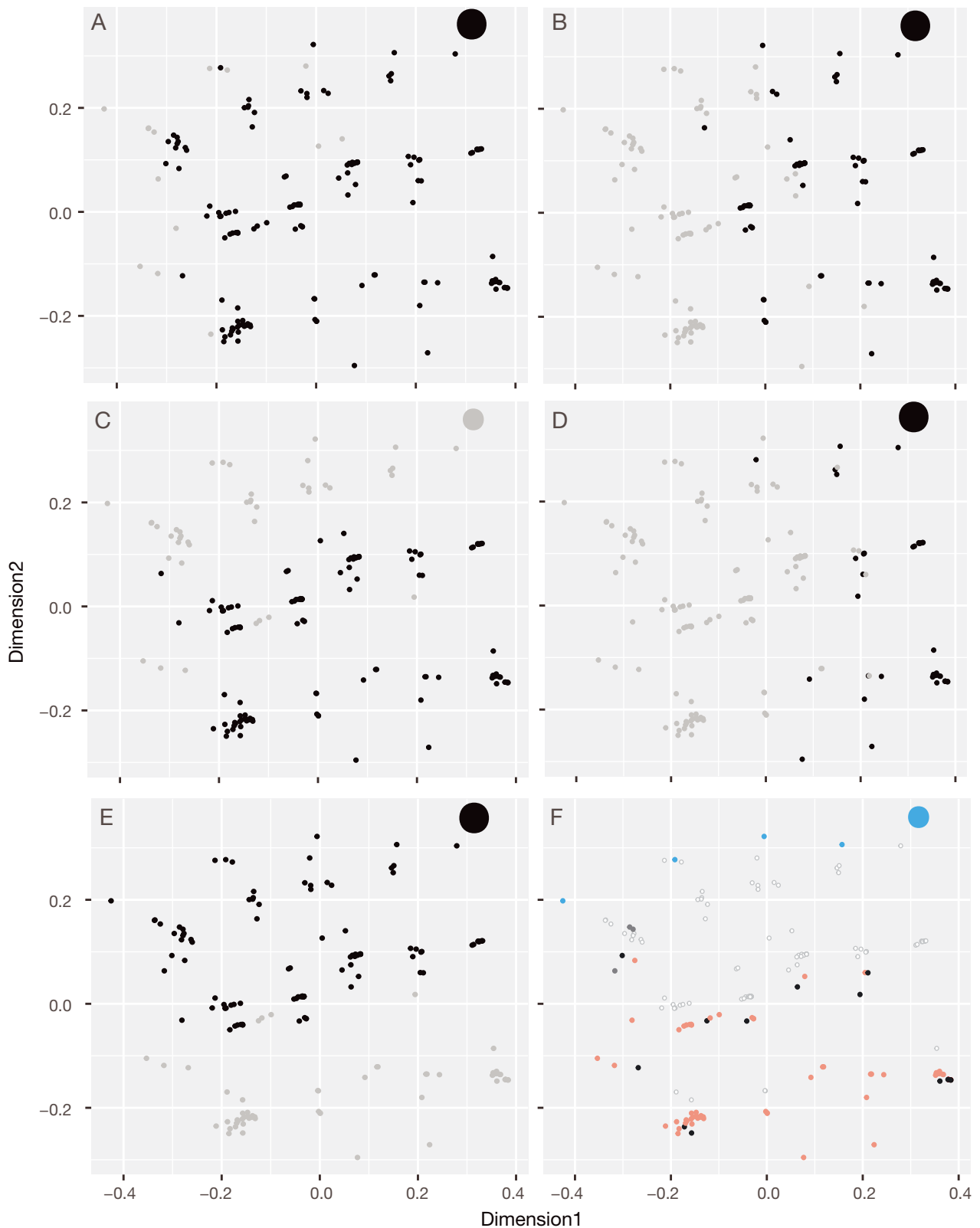
Fig. 6. Multidimensional scaling plots (see also Fig. 5) showing patterns of clustering for species characteristics used to populate the trait matrix. Characteristics shown are as follows: (A) pelagic (gray: yes, black: no), (B) river or lake habitats (gray: both; black: river), (C) migratory (gray: yes; black: no), (D) slow current (gray: yes; black: no), (E) swim bladder type (gray: physoclistous; black: physostomous), and (F) scale type (blue: scutes; gray: ganoid; white: cycloid; pink: ctenoid; black: none). The large dot in the upper right of each panel indicates the position of pallid sturgeon

sturgeon was associated with fast current velocities (Fig. 6D). Interestingly, our surrogacy analysis finds blue sucker as the species with the lowest dissimilarity values to pallid sturgeon, although shovelnose sturgeon — a congener of the pallid sturgeon — has been generally thought of as the best and most appropriate study surrogate (Dryer & Sandvol 1993). However, blue sucker has also been named as a study surrogate species for pallid sturgeon because of their similarities in benthic habitat use, spawning cues and migratory behavior (Quist et al. 2004). It must be reiterated, however, that our methods of designating surrogate species differ from those of Dryer & Sandvol (1993) and Quist et al. (2004). First, we used a quantitative approach for designating surrogates, while the previous assessments used qualitative approaches, although they were keen to consider traits in their surrogate designations. Second, while we were using basic life history and morphological traits in our analysis, we honed in on traits that would be indicative of susceptibility or resistance to entrainment, and the addition or subtraction of certain traits may enable a different solution. Third, it is possible that we did not arrive at shovelnose sturgeon as the best surrogate for pallid sturgeon because our prioritization did not include traits information about hydrodynamics of body shape and aspects of life history such as spawning periodicity that may make shovelnose sturgeon a superior surrogate. As a result, we presented the entire lower decile of dissimilarity scores. Fourth, we did not apply weights to any of the traits. It may be that some traits are more important than others for determining whether organisms are similar. For instance, if we had weighted some scale types over others because scutes — the sharp, bony plates on the backs of sturgeons — make sturgeons much more resistant to scale loss and other injury than fish with other scale types (Amaral et al. 2015), or weighted scale type as being more important to entrainment injury prediction than other traits such as habitat, our surrogacy analysis would have provided a different prioritization. The differences in our findings and the findings of Dryer & Sandvol (1993) and Quist et al. (2004) highlight the importance of carefully choosing the traits that are incorporated into the analysis because inclusion or omission of traits, or over-/underweighting of any trait, can inadvertently provide differing results and contribute to misleading interpretations and poor decisions.

Selecting species for study, monitoring or management activities can be somewhat subjective in the absence of a rigorous quantitative framework, some-

times leaving these choices as the focus of scrutiny during litigation. Some portions of our method rely on expert opinion — deciding which traits to include, assigning ranking scores to clusters of species — which can bear heavily on the results and introduce measures of subjectivity into the quantitative analysis. First of all, it is important to remember that species selection decisions are typically made solely on what are the subjective portions of our analyses without incorporating an unbiased, quantitative method for comparing species and tracking species choice in the decision-making process. In contrast, it is also important to remember that our process does not replace expert knowledge. Our framework is able to provide this quantitative backing to the choice of species and to serve as a proxy for biological impact to an ecosystem (in the case of selecting a species for ecosystem monitoring) or to the relationship between a surrogate species and a threatened or endangered one. Our methodology capitalizes on advances in data availability, coverage and computing power to provide quantitative justification for a traditionally qualitative process that can result in more easily defensible monitoring and conservation plans for impacted ecosystems. Our method did not deliberately weight any risk categories or expert opinions against one another; however, risk categories with higher correlations between average cluster score and overall species rank had higher levels of consensus among experts, whereas risk categories with higher disagreement had lower correlations between average cluster score and overall species rank. We do urge caution, however, in strict interpretations of ranks issued from composite species risk scores. Our sensitivity analyses do suggest that while composite species scores are fairly good representations of values assigned by individual experts, translating these scores into ranks can distort relationships between original expert opinions and the composite scores derived from combining all opinions (Fig. 4). Nevertheless, other portions of our sensitivity analyses show that there can be at least some overlap in the top species designated by all experts (Table 6). Prioritizing these species in common in the upper percentiles for all or several experts may be one way to help conserve expert opinion in final species selection.

LITERATURE CITED

Aguiar FC, Fabião AM, Bejarano MD, Merritt D, Nilsson C, Martins MJ (2013) FLOWBASE – a riparian plant trait-base. Centro de Estudos Florestais, Instituto Superior de Agronomia, Universidade de Lisboa, Lisbon

Amaral SV, Bevelhimer MS, Čada GF, Giza DJ, Jacobson PT, McMahon BJ, Pracheil BM (2015) Evaluation of behavior and survival of fish exposed to an axial-flow hydrokinetic turbine. N Am J Fish Manage 35:97–113

Andelman SJ, Fagan WF (2000) Umbrellas and flagships: efficient conservation surrogates, or expensive mistakes? Proc Natl Acad Sci USA 97:5954–5959

Bell D, Hjältén J, Nilsson C, Jørgensen D, Johansson T (2015) Forest restoration to attract a putative umbrella species, the white-backed woodpecker, benefited saproxylic beetles. Ecosphere 6:art278

Branton MA, Richardson JS (2014) A test of the umbrella species approach in restored floodplain ponds. J Appl Ecol 51:776–785

Čada GF (2001) The development of advanced hydroelectric turbines to improve fish passage survival. Fisheries 26: 14–23

Čada GF, Schweizer PE (2012) The application of traits-based assessment approaches to estimate the effects of hydroelectric turbine passage on fish populations ORNL/TM-2012/110. Oak Ridge National Laboratory, Oak Ridge, TN

Caro TM (2003) Umbrella species: critique and lessons from East Africa. Anim Conserv 6:171–181

Castro-Santos T, Haro A (2015) Survival and behavioral effects of exposure to a hydrokinetic turbine on juvenile Atlantic salmon and adult American shad. Estuaries Coasts 38:203–214

Copeland HE, Sawyer H, Monteith KL, Naugle DE, Pocewicz A, Graf N, Kauffman MJ (2014) Conserving migratory mule deer through the umbrella of sage-grouse. Ecosphere 5:1–16

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001–2014

Dryer MP, Sandvol AJ (1993) Recovery plan for the pallid sturgeon (*Scaphirhynchus albus*). US Fish and Wildlife Service, Bismark, ND

Faul F, Erdfelder E, Lang AG, Buchner A (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 39:175–191

Fleishman E, Murphy DD, Brussard PF (2000) A new method for selection of umbrella species for conservation planning. Ecol Appl 10:569–579

Frimpong EA, Angermeier PL (2009) Fish traits: a database of ecological and life-history traits of freshwater fishes of the United States. Fisheries 34:487–495

Froese R, Pauly D (eds) (2016) FishBase. www.fishbase.org

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27:857–871

Grimm A, Prieto Ramírez AM, Moulherat S, Reynaud J, Henle K (2014) Life-history trait database of European reptile species. J Nat Conserv 9:45–67

Hadjerioua B, Kao S, Sale M, Wei Y and others (2011) National Hydropower Asset Assessment Project, Oak Ridge National Laboratory. US Department of Energy, Oak Ridge National Laboratory, Oak Ridge, TN

Haro A, Odeh M, Noreika J, Castro-Santos T (1998) Effect of water acceleration on downstream migratory behavior and passage of Atlantic salmon smolts and juvenile American shad at surface bypasses. Trans Am Fish Soc 127:118–127

Hierl LA, Franklin J, Deutschman DH, Regan HM, Johnson BS (2008) Assessing and prioritizing ecological communities for monitoring in a regional habitat conservation plan. Environ Manage 42:165–179

Laliberté E, Legendre P (2010) A distance-based framework for measuring functional diversity from multiple traits. Ecology 91:299–305

Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K, Studer M, Roudier P (2014) Cluster: cluster analysis basics and extensions. R package version 1.15.3

Mathur D, Heisey PG, Euston ET, Skalski JR, Hays S (1996) Turbine passage survival estimation for chinook salmon smolts (*Oncorhynchus tshawytscha*) at a large dam on the Columbia River. Can J Fish Aquat Sci 53:542–549

McGill BJ, Enquist BJ, Weiher E, Westoby M (2006) Rebuilding community ecology from functional traits. Trends Ecol Evol 21:178–185

McManamay RA, Frimpong EA (2015) Hydrologic filtering of fish life history strategies across the United States: implications for stream flow alteration. Ecol Appl 25:243–263

Michaud DT, Taft EP (2000) Recent evaluations of physical and behavioral barriers for reducing fish entrainment at hydroelectric plants in the upper Midwest. Environ Sci Policy 3:499–512

NatureServe (2015) NatureServe Explorer: an online encyclopedia of life.Version 70. NatureServe, Arlington, VA, USA. http://explorer.natureserve.org (accessed 15 November 2014)

Pracheil BM, DeRolph CR, Schramm MP, Bevelhimer MS (2016) A fish-eye view of riverine hydropower systems: the current understanding of the biological response to turbine passage. Rev Fish Biol Fish 2016:1–15

Quist MC, Boelter AM, Lovato JM, Korfanta NM and others (2004) Research and assessment needs for pallid sturgeon recovery in the Missouri River. Final Report to the US Geological Survey, US Army Corps of Engineers, US Fish and Wildlife Service and US Environmental Protection Agency. William D. Ruckelshaus Institute of Environment and Natural Resources, University of Wyoming, Laramie, WY

R Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. www.R-project.org/

Regan HM, Hierl LA, Franklin J, Deutschman DH, Schmalbach HL, Winchell CS, Johnson BS (2008) Species prioritization for monitoring and management in regional multiple species conservation plans. Divers Distrib 14:462–471

Roberge JM, Angelstam PER (2004) Usefulness of the umbrella species concept as a conservation tool. Conserv Biol 18:76–85

Rubinoff D (2001) Evaluating the California gnatcatcher as an umbrella species for conservation of southern California coastal sage scrub. Conserv Biol 15:1374–1383

Sattler T, Pezzatti GB, Nobis MP, Obrist MK, Roth T, Moretti M (2014) Selection of multiple umbrella species for functional and taxonomic diversity to represent urban biodiversity. Conserv Biol 28:414–426

Schmidt-Kloiber A, Hering D (2015) www.freshwaterecology.info — an online tool that unifies, standardises and codifies more than 20,000 European freshwater organisms and their ecological preferences. Ecol Indic 53: 271–282

Travnichek VH, Zale AV, Fisher WL (1993) Entrainment of ichthyoplankton by a warmwater hydroelectric facility. Trans Am Fish Soc 122:709–716

Trochet A, Moulherat S, Calvez O, Stevens V, Clobert J, Schmeller D (2014) A database of life-history traits of European amphibians. Biodivers Data J 2:e4123

Tucker CM, Cadotte MW, Davies TJ, Rebelo TG (2012) Incorporating geographical and evolutionary rarity into conservation prioritization. Conserv Biol 26:593–601

US EPA (United States Environmental Protection Agency) (2012) Freshwater biological traits database (Final Report). EPA/600/R-11/038F, US Environmental Protection Agency, Washington, DC

Viera NKM, Poff NL, Carlisle NM, Moulton SR, Koski ML, Kondratieff BC (2006) A database of iotic invertebrate traits for North America. US Geological Survey Data Series 187. US Geological Survey, US Department of the Interior, Reston, VA