



Cite this article: Knouft JH, Anthony MM.

2016 Climate and local abundance in freshwater fishes. *R. Soc. open sci.* **3**: 160093.
<http://dx.doi.org/10.1098/rsos.160093>

Received: 5 February 2016

Accepted: 24 May 2016

Subject Category:

Biology (whole organism)

Subject Areas:

ecology

Keywords:

abundant-centre, environmental niche,
geographical information systems,
population size, quantile regression

Author for correspondence:

Jason H. Knouft

e-mail: jknouft@slu.edu

Electronic supplementary material is available
at <http://dx.doi.org/10.1098/rsos.160093> or via
<http://rsos.royalsocietypublishing.org>.

Climate and local abundance in freshwater fishes

Jason H. Knouft¹ and Melissa M. Anthony²

¹Department of Biology, and ²Program in Integrated and Applied Sciences, Saint Louis University, 3507 Laclede Avenue, St Louis, Missouri 63103, USA

Identifying factors regulating variation in numbers of individuals among populations across a species' distribution is a fundamental goal in ecology. A common prediction, often referred to as the abundant-centre hypothesis, suggests that abundance is highest near the centre of a species' range. However, because of the primary focus on the geographical position of a population, this framework provides little insight into the environmental factors regulating local abundance. While range-wide variation in population abundance associated with environmental conditions has been investigated in terrestrial species, the relationship between climate and local abundance in freshwater taxa across species' distributions is not well understood. We used GIS-based temperature and precipitation data to determine the relationships between climatic conditions and range-wide variation in local abundance for 19 species of North American freshwater fishes. Climate predicted a portion of the variation in local abundance among populations for 18 species. In addition, the relationship between climatic conditions and local abundance varied among species, which is expected as lineages partition the environment across geographical space. The influence of local habitat quality on species persistence is well documented; however, our results also indicate the importance of climate in regulating population sizes across a species geographical range, even in aquatic taxa.

1. Introduction

Identification of the contemporary factors regulating local population abundance is a fundamental goal in ecology. Whittaker [1,2] provided formative evidence that variation in abundance within and among species is due to differential responses across habitat gradients, thus supporting the concept that species respond to environmental conditions based on tolerances to those conditions [3]. At larger spatial scales, studies of variation in population abundance have generally focused on documenting patterns across the geographical ranges of species (e.g. [4,5]). A common prediction, often referred to as the abundant-centre

hypothesis, suggests that population abundance is highest near the centre of a species' geographical range and decreases towards the edge of the range [5–7]. Although the abundant-centre prediction has been described as a general rule in biogeography [4] and driven by a common feature shared by all species [8], this framework provides limited opportunities for direct investigation of environmental factors regulating local abundance because of the primary focus on the geographical position of a population (e.g. geographical coordinates or linear distance from the centre of the range). Hawkins & Diniz-Filho [9] recognized a similar limitation in studies that relate species richness to latitude and noted that hypotheses based on geographical positioning provide minimal benefits for advancing our understanding of spatial variation in biodiversity because of the lack of quantification of environmental conditions (or other spatially varying factors) across the region of interest.

While multiple mechanisms have been proposed to explain the abundant-centre distribution [6,7], from a niche-based perspective, the abundant-centre prediction implies that the most suitable habitat for a species is found near the centre of the geographical distribution and habitat quality decreases towards the edge of the species' range until the absence of suitable habitat contributes to the inability of populations to persist [3,5,10]. However, the quantification of multivariate habitat characteristics across the species range has historically been a challenge, thus limiting the broad-scale assessment of the environmental conditions regulating local population abundance. Fortunately, the availability of global environmental geographical information systems (GISs) data has recently allowed for the quantification of climatic aspects of the niche at broad spatial scales (i.e. the environmental niche) [11].

The term 'environmental niche' is used to refer to the estimation of the n -dimensional hypervolume described by Hutchinson [12] using integrated habitat and species distribution data, particularly at landscape and larger spatial scales [13,14]. Characterization of the environmental niche frequently excludes information on biotic interactions; however, this approach is useful for relating available habitat and species distributions and diversity at broad spatial scales [11]. Using this GIS-based approach, habitat gradients can be quantified across large geographical areas to assess the relationship between regional climatic conditions and variation in local abundance. If a significant component of variation in species abundance is predicted by broad-scale environmental factors such as climate, from a hierarchical perspective, these broad-scale variables should have a primary effect on abundance, while local habitat quality and biotic interactions generate secondary effects that regulate abundance under the constraints of the regional environmental conditions. In this scenario, regional climate defines the arena in which local interactions take place and optimal local habitat will maximize abundance only to the degree allowable by climatic conditions.

While the relationship between the environmental niche and range-wide variation in local population abundance has been investigated in terrestrial species [15], a taxonomically broad assessment of the relationship between climate and range-wide variation in local abundance in freshwater taxa is apparently not available. Nevertheless, the use of climatic variables in correlative approaches to predict aquatic species distributions is common (e.g. [16–18]), and based on the untested assumption that climate is a consistent predictor of local species persistence in aquatic systems across broad geographical regions. The goal of this research is to determine if there is a relationship between climatic conditions and local abundance in freshwater fishes. We use data from a standardized sampling methodology across the ranges of 19 North American freshwater fish species and GIS-based climate data to test the hypothesis that a portion of the variation in local population abundance in freshwater taxa can be predicted with broad-scale environmental data.

2. Material and methods

2.1. Fish data

Fish abundance data were assembled from the United States Geological Survey (USGS) National Water-Quality Assessment (NAWQA) program (<http://water.usgs.gov/nawqa>) [19–21]. The NAWQA dataset contains data derived from over 15 000 fish, invertebrate and algae samples from 51 river basins across the United States and has been used to assess, for example, the relationship between land use and species assemblage structure [22], impacts of non-native species on native species diversity [23] and relationships between species distributions, body size and local abundance [24]. NAWQA sites were generally sampled during low-flow seasons based on standardized stream section length and geomorphological characteristics using electrofishing and beach seining methods to ensure a representative sample of the

Table 1. Results and variable coefficients from models predicting variation in local abundance for each species. '*N*' indicates the number of sites for each species. AICc is the Akaike's information criterion score corrected for small sample size. $R^2_{(\tau)}$ is the pseudo- R^2 statistic. s.e. is standard error. '—' indicates cases where climate does not explain a greater portion of variation in local abundance over the mean based on the Δ AICc > 2 criterion.

species (<i>N</i>)							
family	AICc	$R^2_{(\tau)}$	variable	coefficient	s.e.	t_{stat}	<i>p</i> -value
<i>Campostoma anomalum</i> (151) Cyprinidae	265.42	0.146	constant	2.124	0.031	68.784	0.000
			PC2	−0.199	0.026	−7.574	0.000
			PC1 ²	−0.229	0.021	−10.749	0.000
			PC1 × PC2	−0.170	0.022	−7.568	0.000
<i>Campostoma oligolepis</i> (84) Cyprinidae	120.58	0.157	constant	1.991	0.038	52.582	0.000
			PC1	0.191	0.068	2.817	0.006
			PC2 ²	−0.051	0.005	−9.888	0.000
<i>Cottus caroliniae</i> (88) Cottidae	171.45	0.090	constant	1.777	0.083	21.420	0.000
			PC2 ²	−0.077	0.025	−3.111	0.003
			PC2 × PC3	0.134	0.044	3.058	0.003
<i>Cyprinella analostoma</i> (25) Cyprinidae	18.30	0.629	constant	1.577	0.060	26.165	0.000
			PC1	0.467	0.035	13.253	0.000
			PC2	0.228	0.035	6.462	0.000
			PC1 ²	−0.100	0.056	−1.792	0.088
<i>Cyprinella spiloptera</i> (117) Cyprinidae	218.13	0.153	constant	1.748	0.060	29.258	0.000
			PC1	−0.227	0.078	−2.905	0.004
			PC2	−0.110	0.094	−1.170	0.244
			PC1 ²	−0.128	0.027	−4.695	0.000
			PC1 × PC2	0.192	0.059	3.266	0.001
<i>Cyprinella venusta</i> (43) Cyprinidae	86.09	0.113	constant	1.940	0.067	28.786	0.000
			PC2	−0.161	0.068	−2.359	0.023
<i>Esox americanus</i> (44) Esocidae	100.48	0.150	constant	1.805	0.200	9.000	0.000
			PC1 ²	−0.454	0.116	−2.918	0.006
<i>Etheostoma blennioides</i> (102) Percidae	147.40	0.259	constant	1.195	0.079	15.145	0.000
			PC1	−0.324	0.074	−4.375	0.000
			PC1 × PC2	0.210	0.068	3.072	0.003
			PC2 ²	−0.097	0.051	−1.920	0.058
<i>Etheostoma caeruleum</i> (55) Percidae	103.46	0.074	constant	1.449	0.123	11.798	0.000
			PC1	0.114	0.124	0.920	0.362
<i>Etheostoma flabellare</i> (61) Percidae	99.87	0.186	constant	1.452	0.099	14.720	0.000
			PC1 ²	−0.144	0.046	−3.118	0.003
			PC1 × PC2	−0.085	0.088	−0.974	0.334
			PC2 ²	0.054	0.058	0.925	0.359
<i>Etheostoma olmstedi</i> (69) Percidae	116.86	0.163	constant	1.564	0.061	25.561	0.000
			PC1	−0.217	0.062	−3.524	0.001
			PC2	0.182	0.062	2.956	0.004

(Continued.)

Table 1. (Continued.)

species (<i>N</i>)							
family	AICc	$R^2_{(\tau)}$	variable	coefficient	s.e.	t_{stat}	<i>p</i> -value
<i>Fundulus olivaceus</i> (39)	22.05	0.509	constant	1.148	0.073	15.640	0.000
Fundulidae			PC2	−0.217	0.041	−5.287	0.000
			PC2 ²	−0.344	0.057	−5.991	0.000
			PC3	0.204	0.039	5.183	0.000
			PC3 ²	−0.052	0.019	−2.793	0.009
<i>Lepistosteus oculatus</i> (36)	27.66	0.225	constant	0.888	0.048	18.448	0.000
Lepistosteidae			PC1	0.181	0.048	3.736	0.001
			PC2 ²	−0.071	0.017	−4.297	0.000
<i>Lepistosteus osseus</i> (24)	18.27	0.555	constant	1.200	0.064	18.732	0.000
Lepistosteidae			PC1	0.414	0.045	9.271	0.000
			PC2	−0.101	0.040	−2.541	0.019
			PC2 ²	−0.520	0.053	−9.756	0.000
<i>Moxostoma duquesnei</i> (80)	116.45	0.071	constant	1.087	0.067	15.805	0.000
Catostomidae			PC2	−0.051	0.058	0.881	0.381
			PC2 ²	−0.085	0.044	−1.951	0.055
<i>Moxostoma erythrum</i> (100)	154.40	0.140	constant	1.042	0.078	13.392	0.000
Catostomidae			PC1	−0.101	0.078	−1.293	0.199
			PC2	−0.157	0.078	−2.013	0.047
<i>Noturus exilis</i> (20)	24.94	0.435	constant	1.271	0.078	16.192	0.000
Ictaluridae			PC1	0.481	0.104	4.629	0.000
			PC1 ²	−0.413	0.103	−4.015	0.001
			PC2 ²	0.057	0.065	0.868	0.398
<i>Percina nigrofasciata</i> (39)							
Percidae	—	—	—	—	—	—	—
<i>Percina sciera</i> (22)	0.08	0.738	constant	0.407	0.030	13.427	0.000
Percidae			PC1	0.480	0.023	20.982	0.000
			PC1 × PC2	0.085	0.023	3.725	0.002
			PC1 ²	0.111	0.024	4.648	0.000

fish assemblage [25,26]. Individuals were identified and counted at each locality to provide estimates of local abundance [21].

The NAWQA dataset is limited to sites within the United States. To effectively examine range-wide variation in local abundance, only non-migratory species with distributions contained entirely within the United States were used for this study. To ensure robust statistical analysis, species were also excluded if they occurred at less than 20 localities in the NAWQA dataset and did not cover the full extent of the species' geographical distribution. The region of North America north of Mexico contains approximately 828 species of native freshwater fishes [27]. A total of 19 species, representing eight families, were available for analyses after the geographical and sample size constraints were imposed on the NAWQA dataset (table 1).

Although NAWQA samples were based on a standardized sampling methodology, peculiarities associated with site-specific sampling may influence abundance estimates. Rarefaction was used in this study to account for the influence of potential sampling bias on estimates of local population abundance [28]. Separate rarefaction estimates were generated for each of the 19 species. The sample size during rarefaction resampling was based on the total number of individuals (of all species) at the site where the

Table 2. Climate variables used in all analyses.

climate variable
mean annual temperature (K)
mean diurnal range (K) [<i>mean of monthly (maximum temperature – minimum temperature)</i>]
temperature seasonality (<i>standard deviation of monthly mean temperatures</i>)
maximum temperature of warmest month (K)
minimum temperature of coldest month (K)
temperature annual range (K)
annual precipitation (mm)
precipitation of wettest month (mm)
precipitation of driest month (mm)
precipitation seasonality (<i>coefficient of variation of monthly precipitation measures</i>)

least number of total individuals were collected [23]. For example, if a species occurred at 50 sites and the total number of individuals of all species at each site ranged from 100 to 1000, then 100 individuals were resampled from each of the 50 sites during rarefaction. Data from each site were resampled 1000 times without replacement to generate rarefied estimates of local abundance for each species occurring at the site (ECOSIM, v. 7.0, [29]) [23]. This process resulted in rarefied estimates of local population abundance for the focal species.

2.2. Climate variables

Data for 10 climate variables were assembled from the Worldclim dataset (www.worldclim.org) ([30]; table 2). Owing to the covariation of climate variables, climate data were subjected to a principal components analysis (PCA) for each species to produce a set of uncorrelated variables. All climate data, except temperature and precipitation seasonality, were \log_{10} -transformed prior to statistical modelling and separate analyses were carried out for each of the 19 fish species. Principal components with eigenvalues greater than 1.0 were retained for further analyses, resulting in two or three principal components with loadings that varied among species (electronic supplementary material, appendix SA).

2.3. Modelling individual species using linear quantile regression

The relationship between climate and $\log_{10}(\text{abundance})$ was examined using linear quantile regression. Quantile regression has been demonstrated to be preferable to ordinary least-squares (OLS) regression for analysing ecological data in cases where the limits of species abundance are a result of complex interactions between measured and unmeasured variables [31]. In this study, climate data are the measured variables and local habitat availability and biotic interactions are the unmeasured variables. In typical OLS regression, the effect of a set of predictors on the rate of change of the response is determined for the conditional mean of the response distribution. The conditional mean is the portion of the response distribution where the interactions between predictors of interest and other unmeasured predictors are most likely to occur. However, determining the rate of change of the conditional response near the upper boundary of its distribution has been demonstrated to provide a more accurate estimate of the limiting effect of measured predictors where they are most likely to be the primary factor limiting abundance [31]. A quantile regression approach can model any portion of the conditional response distribution, referred to as quantile (τ), which takes on values from 0 to 1. The model uses an optimization function that minimizes the sum of asymmetrically weighted absolute residuals based on the value of τ (for a more in-depth treatment of modelling using regression quantiles, see [31,32]). The limiting effect of climate for each fish species was modelled at the 95th percentile of the abundance, corresponding to $\tau = 0.95$ (similar to [33]).

2.4. Model selection

Final models for each species were determined using stepwise variable selection. Along with the main effect of each principal component, quadratic terms and interaction terms were also considered for

inclusion in the final model to account for any curvilinear or synergistic effects. The retention of variables for the final model at each step of selection was determined using Akaike Information Criterion for small sample sizes (AICc). AICc is calculated using the negative log likelihood of the model, a penalty for including additional variables into the model to avoid over-fitting, and a bias correction term for small sample size, which results in an overall measure of lack of model fit to the data [34]. Stepwise selection began with a model containing only the intercept term. Single predictor variables were then added to the model and the variable generating the largest improvement based on the ΔAICc measure was retained. The cut-off for variable inclusion was an improvement in AICc of greater than two ($\Delta\text{AICc} > 2$) [34]. After the addition of each variable, variables already in the model were re-examined for inclusion because variables entered in a previous step may become superfluous at later steps due to their relationship with other variables [35]. This was done by deleting individual variables from the model and removing variables which no longer contributed to significant improvement of AICc (i.e. $\Delta\text{AICc} > 2$). This process was repeated until the addition or removal of variables no longer improved the model, resulting in the final model for each species. We chose a stepwise model selection approach as opposed to comparing the full model to candidate models based on *a priori* hypotheses for multiple reasons. First, the untested inclusion of interaction and polynomial terms in a full model results in an extremely complicated model with terms that are difficult to interpret. Moreover, the inclusion of interaction terms is dependent on the importance of the main effect of the interacting variables, thus the inclusion of interaction terms in *a priori* models can be overly speculative and challenging to conceptualize. Finally, the use of species-specific PCA produces univariate variables with varying characteristics for each taxon and makes the generation of consistent *a priori* hypotheses impossible.

Standard errors for parameters were calculated using the Markov chain marginal bootstrap (MCMB) developed by He & Hu [36]. For species with sample sizes less than 50, the sparsity method of confidence interval estimation was used because resampling methods become unstable at small sample sizes [37]. Goodness of fit of the final model for each species was determined using the $R^2_{(\tau)}$ statistic proposed in [38]. While traditional R^2 measures model performance on the conditional mean response, $R^2_{(\tau)}$ is a local measure of fit that characterizes model performance on the specified quantile of the response distribution [15]. $R^2_{(\tau)}$ is calculated for the specified quantile by first dividing the residuals of the full model by the residuals of the intercept-only (null) model. This quantity is then subtracted from 1.0, yielding a quantity interpretable as the per cent reduction in error obtained using the full model as opposed to the null model [38]. All analyses were carried out using the quantreg package [39] in R v. 3.1 software [40].

3. Results

Eighteen of 19 species models include variables that predict abundance data better than random (i.e. better than a model with only a constant = 0) using the ΔAICc (table 1). $R^2_{(\tau)}$ values among the final models ranged from 0.071 to 0.738 with an average of 0.240 among models (table 1). Species with smaller sample sizes tended to have larger $R^2_{(\tau)}$ values ($R^2 = 0.363$, $p = 0.001$), suggesting that the limited density of localities for certain species did not inhibit our ability to detect relationships among variables. In addition, the model selection approach and the limited number of predictor variables in each model suggest that model overfitting is probably not driving this relationship. A graphical example of the relationship between climate and local abundance for the longnose gar (*Lepisosteus osseus*) is presented in figure 1.

Based on principal component scores and loadings, temperature is positively correlated with local abundance in *Cyprinella analostoma*, *L. osseus*, *L. oculatus*, *Campostoma oligolepis*, *Cy. venusta* and *Etheostoma caeruleum*, and negatively correlated with local abundance in *E. blennioides*, *E. olmstedii* and *Moxostoma erythrurum*. Precipitation is positively correlated with local abundance in *Cy. analostoma* and *Ca. oligolepis* and negatively correlated with local abundance in *L. osseus*, *E. blennioides*, *E. olmstedii*, *Ca. anomalum*, and *M. erythrurum*. Fourteen of the 19 final models contained one or more squared terms, which indicates that local abundance is maximized at intermediate values of climate (table 1). Six of the 19 final models contained an interaction term, which indicates the strength of the association between particular climate variables and local abundance changes as a function of other climate variables (table 1).

4. Discussion

Recent advances in the availability of GIS-based global environmental data have afforded the opportunity to examine variation in population sizes across a species' distribution, beyond simply

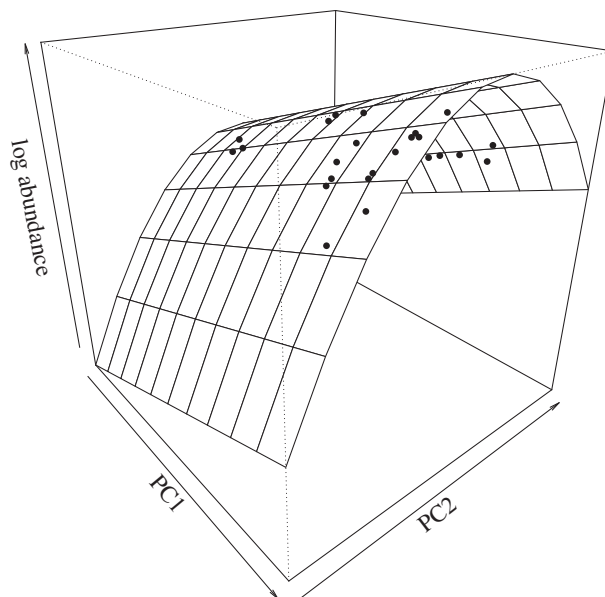


Figure 1. Relationship between principal component (PC) scores describing temperature and precipitation and $\log_{10}(\text{abundance})$ in the longnose gar (*Lepisosteus osseus*).

relating the size of a population to its position in the landscape. In particular, estimation of a climate-based environmental niche for a species provides the ability to examine variation in population size in the context of apparent species tolerances to environmental gradients. While relating population sizes to environmental gradients has been widely investigated in terrestrial taxa (e.g. [2]), this relationship has received less attention in freshwater species and generally focused on limited geographical areas [41–44], possibly due to limited amounts of standardized abundance data. Results from our niche-based approach using GIS environmental data support the hypothesis that climate is responsible for some aspect of local abundance in freshwater taxa, with climate variables explaining a portion of the variation in abundance in 18 of 19 species examined. While our quantile regression approach acknowledges the importance of local habitat and biotic interactions in the regulation of population size, results support the prediction that climate is important to population-level processes and not only at the edge of a species' distribution, but also throughout the entire range. Dispersal and stream network geometry are unlikely to influence variation in population sizes at our study scale due to the relatively limited intra-annual dispersal distances in most non-migratory North American fish species. However, these spatial influences are potentially important at smaller (e.g. watershed) scales.

The relationships between principal component (PC) climate variables and local abundance are different among species, which is expected as lineages partition the environment across geographical space [45]. Many species in this study exhibit linear relationships between climate and abundance. In particular, eight species exhibit either positive or negative correlations with temperature (based on principal component scores and loadings), suggesting that population abundance tends to peak near the warmest or coldest parts of the species' range. In these cases, physical features such as watershed boundaries may constrain species distributions as opposed to climatic conditions, particularly in warmer areas. For example, several species reach their highest abundances near the Gulf of Mexico (*Cy. venusta*, *L. oculatus*, *L. osseus*), which is an obvious barrier to expansion into warmer regions. Variation in abundance in 14 species is predicted by at least one squared variable, indicating a Gaussian-type relationship between local abundance and climate [46]. This suggests that particular aspects of climate may be more important in regulating these species' distributions than physical attributes of the landscape as species abundance peaks at an intermediate point along the species' perceived climatic range.

Percina nigrofasciata is the only species that does not exhibit a relationship between climate and local abundance. Other than Type II error, this lack of relationship may be a result of physical constraints on the species' distribution. The majority of the species range, which is relatively small compared with other species in this dataset, occurs in watersheds that drain directly to the Gulf of Mexico (e.g. the Mobile River basin) [27]. Thus, the environmental tolerances of the species may not be fully realized by

the species due to northern headwater and southern Gulf of Mexico distributional constraints. These constraints may limit detection of the response of abundance to the full climatic tolerances of the species.

The translation of climate to aquatic environmental conditions is a complex process and generally realized as hydrological regimes and water temperature. While flow, seasonal variability in flow and water temperature directly affect aquatic taxa and estimates of these variables would probably enhance our analyses, the generation of these types of data to address ecological patterns is constrained by available stream gauge monitoring (e.g. [47]) or requires a landscape-scale hydrological modelling approach [48,49]. However, prediction of local flow and water temperatures across a continental scale using this approach is computationally prohibitive; therefore, the application of non-aquatic data surrogates (such as climate) is necessary. Air temperature and water temperature are correlated at broad spatial scales [50,51], and precipitation and air temperature can serve as predictors of river flows due to their determinant roles in the hydrological cycle (e.g. rainfall contribution and evapotranspiration), at least at continental scales. However, the ability of climate variables to explain a portion of the variation in local abundance of aquatic taxa across the range of a species has not been addressed. Results from this study provide support that climate data can explain the distribution of freshwater taxa at broad spatial scales and improve the ability to predict local population sizes.

The abundant-centre hypothesis provides little more than a framework for speculation regarding factors that regulate population sizes over broad geographical areas and should be discarded in favour of more biologically meaningful hypotheses. Our results indicate that climate not only explains variation in abundance of aquatic species across their geographical range, but species vary in their response to gradients in climate in ways that are not consistent with predictions from the abundant-centre hypothesis. In particular, the abundant-centre hypothesis fails to consider physical aspects of the landscape in limiting the distribution of species, as well as the location of climatic optima, which can potentially occur in any portion of a species' geographical range. The differential response of species to temperature and precipitation along their geographical range indicates that climate is a meaningful predictor of local abundance of freshwater fishes. This, along with increasingly available GIS-based data resources on other biologically relevant environmental conditions will further improve the current understanding of factors regulating local abundance of freshwater species.

Data accessibility. Supporting data for this study can be accessed at <http://dx.doi.org/10.5061/dryad.1d1t5>.

Authors' contributions. J.H.K. was responsible for assembling species distribution and climate data. J.H.K. and M.M.A. were both responsible for study design, analyses, interpretation of results and writing of the manuscript.

Competing interests. We have no competing interests.

Funding. This work was supported by funding from the United States National Science Foundation to J.H.K. (DEB-0844644)

Acknowledgements. We thank Sophia Niu for helpful comments on a previous draft of this manuscript and Nicholas Caruso for assistance with data organization.

References

- Whittaker RH. 1952 A study of summer foliage insect communities in the Great Smoky Mountains. *Ecol. Monogr.* **22**, 1–44. (doi:10.2307/1948527)
- Whittaker RH. 1956 Vegetation of the Great Smoky Mountains. *Ecol. Monogr.* **26**, 1–80. (doi:10.2307/1943577)
- Shelford VE. 1931 Some concepts of biogeology. *Ecology* **12**, 455–467. (doi:10.2307/1928991)
- Hengeveld R, Haeck J. 1982 The distribution of abundance. I. Measurements. *J. Biogeogr.* **9**, 303–316. (doi:10.2307/2844717)
- Brown JH. 1984 On the relationship between abundance and distribution of species. *Am. Nat.* **124**, 255–279. (doi:10.1086/284267)
- Sagarin RD, Gaines SD. 2002 The 'abundant centre' distribution: to what extent is it a biogeographical rule? *Ecol. Lett.* **5**, 137–147. (doi:10.1046/j.1461-0248.2002.00297.x)
- Sagarin RD, Gaines SD, Gaylord B. 2006 Moving beyond assumptions to understand distributions across the ranges of species. *Trends Ecol. Evol.* **21**, 524–530. (doi:10.1016/j.tree.2006.06.008)
- Guo Q, Taper M, Schoenberger M, Brandl J. 2005 Spatial-temporal population dynamics across species range: from centre to margin. *Oikos* **108**, 47–57. (doi:10.1111/j.0030-1299.2005.13149.x)
- Hawkins BA, Diniz-Filho JAF. 2004 'Latitude' and geographic patterns in species richness. *Ecography* **27**, 268–272. (doi:10.1111/j.0906-7590.2004.03883.x)
- MacArthur RH. 1972 *Geographical ecology: patterns in the distribution of species*. Princeton, NJ: Princeton University Press.
- Kozak KH, Graham CH, Wiens JJ. 2008 Integrating GIS-based environmental data into evolutionary biology. *Trends Ecol. Evol.* **23**, 141–148. (doi:10.1016/j.tree.2008.02.001)
- Hutchinson GE. 1957 Concluding remarks. *Cold Spring Harb. Symp. Quant. Biol.* **22**, 415–427. (doi:10.1101/SQB.1957.022.01.039)
- Austin M. 2007 Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecol. Model.* **200**, 1–19. (doi:10.1016/j.ecolmodel.2006.07.005)
- Austin MP, Meyers JA. 1996 Current approaches to modeling the environmental niche of eucalypts: implications for management of forest biodiversity. *Forest Ecol. Manage.* **85**, 95–106. (doi:10.1016/S0378-1127(96)03753-X)
- Vanderwal J, Shoo LP, Johnson CN, Williams SE. 2009 Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *Am. Nat.* **174**, 282–291. (doi:10.1086/600087)
- McNyset KM. 2009 Ecological niche conservatism in North American freshwater fishes. *Biol. J. Linn. Soc.* **96**, 282–295. (doi:10.1111/j.1095-8312.2008.01121.x)
- Reshetnikov AN, Ficetola GF. 2011 Potential range of the invasive fish rotan (*Percottus glenii*) in the Holarctic. *Biol. Invasions* **13**, 2967–2980. (doi:10.1007/s10530-011-9982-1)
- Logez M, Bady P, Pont D. 2012 Modelling the habitat requirement of riverine fish species at the European scale: sensitivity to temperature and precipitation and associated uncertainty. *Ecol. Freshw. Fish* **21**, 266–282. (doi:10.1111/j.1600-0633.2011.00545.x)

19. Hirsch RM, Alley WM, Wilber WG. 1988 Concepts for a national water-quality assessment program. *U.S. Geological Survey Circular* **1021**, 1–42.
20. Leahy PP, Rosenshein JS, Knopman DS. 1990 Implementation plan for the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report: 90–174, 10 p.
21. Walsh SJ, Meador MR. 1998 Guidelines for quality assurance and quality control of fish taxonomic data collected as part of the National Water-Quality Assessment Program: U.S. Geological Survey Water-Resources Investigations Report 98–4239, 33 p.
22. Carlisle DM, Hawkins CP. 2008 Land use and the structure of western U.S. stream invertebrate assemblages: predictive models and ecological traits. *J. N. Am. Benthol. Soc.* **27**, 986–999. (doi:10.1899/07-176.1)
23. Mitchell AL, Knouft JH. 2009 Non-native fishes and native species diversity in freshwater fish assemblages across the United States. *Biol. Invasions* **11**, 1441–1450. (doi:10.1007/s10530-008-9352-9)
24. Passy SI. 2008 Species size and distribution jointly and differentially determine diatom densities in U.S. streams. *Ecology* **89**, 475–484. (doi:10.1890/07-0405.1)
25. Fitzpatrick FW, D'Arconte PJ, Meador MR, Maupin MA, Gurts ME. 1998 Revised methods for characterizing stream habitat in the National Water-Quality Assessment Program. U.S. Geological Survey Water-Resources Investigations Report 98-4052, 77 p.
26. Moulton SR, Kennen JG, Goldstein RM, Hambrook JA. 2002 Revised protocols for sampling algal, invertebrate, and fish communities as part of the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 02-150, 75 p.
27. Page LM, Burr BM. 2011 *A field guide to freshwater fishes of North America north of Mexico*, 2nd edn. New York, NY: Houghton Mifflin Harcourt.
28. Gotelli NJ, Colwell RK. 2001 Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.* **4**, 379–391. (doi:10.1046/j.1461-0248.2001.00230.x)
29. Gotelli NJ, Entsminger GL. 2008 *EcoSim: null models software for ecology*. Version 7. Jericho, VT: Acquired Intelligence Inc. & Kesey-Bear.
30. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005 Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978. (doi:10.1002/joc.1276)
31. Cade BS, Noon BR. 2003 A gentle introduction to quantile regression for ecologists. *Front. Ecol. Environ.* **1**, 412–420. (doi:10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2)
32. Cade BS, Terrell JW, Schroeder RL. 1999 Estimating effects of limiting factors with regression quantiles. *Ecology* **80**, 311–323. (doi:10.1890/0012-9658(1999)080[0311:EEOLFJ]2.0.CO;2)
33. Dunham JB, Cade BS, Terrell JW. 2002 Influences of spatial and temporal variation on fish-habitat relationships defined by regression quantiles. *Trans. Am. Fish. Soc.* **131**, 86–98. (doi:10.1577/1548-8659(2002)131<0086:IOSATV>2.0.CO;2)
34. Burnham KP, Anderson DR. 2004 Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Method. Res.* **33**, 261–304. (doi:10.1177/0049124104268644)
35. Kleinbaum DG, Kupper LL, Nizam A, Muller KE. 2008 *Applied regression analysis and other multivariate methods*. Belmont, CA: Duxbury Press.
36. He X, Hu F. 2002 Markov chain marginal bootstrap. *J. Am. Stat. Assoc.* **97**, 783–795. (doi:10.1198/016214502388618591)
37. Chen C. 2004 An introduction to quantile regression with the QUANTREG Procedure. In *Proc. of the 30th annual meeting of SAS Users Group International (Paper 213-30)*. Cary, NC: SAS Institute.
38. Koenker R, Mochado JFA. 1999 Goodness of fit and related inference processes for quantile regression. *J. Am. Stat. Assoc.* **94**, 1296–1310. (doi:10.1080/01621459.1999.10473882)
39. Koenker R. 2013 quantreg: Quantile Regression. R package version 4.98. See <http://CRAN.R-project.org/package=quantreg>.
40. R Core Team. 2013 *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See <http://www.R-project.org/>.
41. Taylor CM, Winston MR, Matthews WJ. 1993 Fish species-environment and abundance relationships in a Great-Plains river system. *Ecography* **16**, 16–23. (doi:10.1111/j.1600-0587.1993.tb00054.x)
42. Taylor CM, Lienesch PW. 1995 Environmental correlates of distribution and abundance for *Lythrurus snelsoni*: a range-wide analysis of an endemic fish species. *Southwest Nat.* **40**, 373–378.
43. Holmgren K, Appelberg M. 2000 Size structure of benthic freshwater fish communities in relation to environmental gradients. *J. Fish Biol.* **57**, 1312–1330. (doi:10.1111/j.1095-8649.2000.tb00489.x)
44. Carey MP, Mather ME. 2009 Identifying across-system sources of variation in a generalist freshwater fish: correlates of total and size-specific abundance of yellow perch. *Ecol. Freshw. Fish* **18**, 145–155. (doi:10.1111/j.1600-0633.2008.00335.x)
45. Wiens JJ, Donoghue MJ. 2004 Historical biogeography, ecology and species richness. *Trends Ecol. Evol.* **19**, 639–644. (doi:10.1016/j.tree.2004.09.011)
46. Gause GF. 1931 Influence of ecological factors on the size of populations. *Am. Nat.* **65**, 70–76. (doi:10.1086/280347)
47. Niu SQ, Franczyk M, Knouft JH. 2012 Regional species richness, hydrological characteristics and the local species richness of assemblages of North American stream fishes. *Freshw. Biol.* **57**, 2367–2377. (doi:10.1111/fwb.12016)
48. Ficklin DL, Barnhart BL, Knouft JH, Stewart IT, Maurer EP, Letsinger SL, Whittaker GW. 2014 Climate change and stream temperature projections in the Columbia River basin: habitat implications of spatial variation in hydrologic drivers. *Hydrol. Earth. Syst. Sci.* **18**, 4897–4912. (doi:10.5194/hess-18-4897-2014)
49. Knouft JH, Chu ML. 2015 Using watershed-scale hydrological models to predict the impacts of increasing urbanization on freshwater fish assemblages. *Ecology* **8**, 273–285. (doi:10.1002/eco.1506)
50. Stefan HG, Preudhomme EB. 1993 Stream temperature estimation from air-temperature. *Water Resour. Bull.* **29**, 27–45. (doi:10.1111/j.1752-1688.1993.tb01502.x)
51. Mohseni O, Stefan HG. 1999 Stream temperature air temperature relationship: a physical interpretation. *J. Hydrol.* **218**, 128–141. (doi:10.1016/S0022-1694(99)00034-7)