

ARTICLE

Habitat Associations of Fish Species of Greatest Conservation Need at Multiple Spatial Scales in Wadeable Iowa Streams

Anthony R. Sindt*¹

Department of Natural Resource Ecology and Management, Iowa State University, 339 Science II, Ames, Iowa 50011, USA

Michael C. Quist

U.S. Geological Survey, Idaho Cooperative Fish and Wildlife Research Unit, Department of Fish and Wildlife Sciences, University of Idaho, Box 441141, Moscow, Idaho 83844, USA

Clay L. Pierce

U.S. Geological Survey, Iowa Cooperative Fish and Wildlife Research Unit, Department of Natural Resource Ecology and Management, Iowa State University, 339 Science II, Ames, Iowa 50011, USA

Abstract

Fish and habitat data were collected from 84 wadeable stream reaches in the Mississippi River drainage of Iowa to predict the occurrences of seven fish species of greatest conservation need and to identify the relative importance of habitat variables measured at small (e.g., depth, velocity, and substrate) and large (e.g., stream order, elevation, and gradient) scales in terms of their influence on species occurrences. Multiple logistic regression analysis was used to predict fish species occurrences, starting with all possible combinations of variables (5 large-scale variables, 13 small-scale variables, and all 18 variables) but limiting the final models to a maximum of five variables. Akaike's information criterion was used to rank candidate models, weight model parameters, and calculate model-averaged predictions. On average, the correct classification rate (CCR = 80%) and Cohen's kappa ($\kappa = 0.59$) were greatest for multiple-scale models (i.e., those including both large-scale and small-scale variables), intermediate for small-scale models (CCR = 75%; $\kappa = 0.49$), and lowest for large-scale models (CCR = 73%; $\kappa = 0.44$). The occurrence of each species was associated with a unique combination of large-scale and small-scale variables. Our results support the necessity of understanding factors that constrain the distribution of fishes across spatial scales to ensure that management decisions and actions occur at the appropriate scale.

Conservation of freshwater ecosystems is an important goal of resource managers and an understanding of species distributions and habitat requirements greatly increases the probability of successful ecosystem restoration and preservation. Furthermore, an understanding of species–habitat relationships can provide insight into the effects of land use practices, habitat alterations, and climate change on species distributions (Wang

et al. 2003; Wall et al. 2004; Lyons et al. 2010). Modeling of species distributions is an important tool for addressing many issues in conservation (Guisan and Thuiller 2005), and the use of predictive occurrence models to further the understanding of fish species' relationships with habitat features in freshwater systems is common (e.g., Olden and Jackson 2001; Rich et al. 2003; Steen et al. 2008). As habitat loss and degradation

*Corresponding author: anthony.sindt@dnr.state.oh.us

¹Present address: Ohio Department of Natural Resources, 10517 Canal Road Southeast, Hebron, Ohio 43025, USA.

Received January 29, 2012; accepted July 18, 2012

continue to threaten fish biodiversity in North America (Miller et al. 1989; Richter et al. 1997; Jelks et al. 2008), species distribution models are playing an increasingly important role in conservation. However, predicting the distribution of fish species is challenging because fish species occurrences are influenced by a combination of abiotic and biotic processes acting across multiple spatial and temporal scales (Poff 1997; Marsh-Matthews and Matthews 2000; Jackson et al. 2001). A further challenge is that the most influential processes vary by species and system (Pont et al. 2005; Monti and Legendre 2009).

The link between landscapes and aquatic ecosystems has been appreciated for decades (Hynes 1975; Vannote et al. 1980; Allan 2004), and the relationships between processes acting at larger scales and habitats and biota at smaller scales are of great interest to aquatic ecologists (Richards et al. 1996; Wang et al. 2003; Hughes et al. 2006). Habitat variables measured at large scales are often able to explain substantial variation in fish assemblages, particularly in highly variable landscapes (e.g., Rocky Mountain streams; Rahel and Hubert 1991) or across large spatial scales (e.g., entire midwestern USA; Marsh-Matthews and Matthews 2000). Additionally, the modeling of species occurrences with large-scale habitat variables is economical because such variables are easily measured with a GIS, thereby reducing the need for costly field sampling (Oakes et al. 2005; Steen et al. 2008). However, in other regions, large-scale habitat variables such as land use may not effectively explain variation in fish assemblage structure (Rich et al. 2003; Heitke et al. 2006) or may only do so indirectly (Rowe et al. 2009a).

Relationships between fish assemblages and environmental features measured at small scales are also well documented (Gorman and Karr 1978; Schlosser 1982; Lobb and Orth 1991). The influence of instream physical habitat features (e.g., depth and substrate composition) on food, refuge, and spawning habitat availability is easy to conceptualize, and their influence on fish assemblages is well studied (Fischer and Paukert 2008; Rowe et al. 2009b). Thus, many studies have used instream habitat variables to explain the distribution or abundance of fish species (e.g., Hubert and Rahel 1989; Quist et al. 2005; Rashleigh et al. 2005). Furthermore, an understanding of the constraints on fish species occurrences measured at small scales (e.g., canopy cover) may be most useful for conservation efforts because small-scale characteristics can be more easily manipulated for restoration than landscape-level features (e.g., elevation).

Identification of the factors and spatial scales influencing fish species occurrences remains a central focus of fisheries ecologists (e.g., Porter et al. 2000; Rich et al. 2003; Pont et al. 2005). Biotic communities are influenced by a hierarchical system of constraints in which large-scale processes constrain smaller-scale processes (Tonn 1990; Poff 1997; Quist et al. 2005). This is especially evident for stream habitats, which are hierarchically organized within catchments, segments, reaches, macrohabitats, and microhabitats (Frissel et al. 1986). Thus, species occurrence models that account for effects measured at multiple spatial scales may provide more predictive power and

transferability than models that incorporate variables measured at a single spatial scale (Leftwich et al. 1997; Rich et al. 2003; Pont et al. 2005).

In Iowa, 68 of approximately 144 extant fish species are classified as species of greatest conservation need (SGCN), and the protection and enhancement of habitats to improve their status are of high priority (Zohrer 2005). Fifty percent of the fish SGCN examined by Sindt et al. (2012) were found to have declining distributions. Despite concern regarding their conservation status, the habitat associations of SGCN and most other Iowa stream fish species are poorly understood. Therefore, an understanding of factors influencing the occurrences of stream fish SGCN at multiple spatial scales is important for guiding the conservation of these species (Lewis et al. 1996; Rabeni and Sowa 1996; Durance et al. 2006; Dunham and Peterson 2010).

Our objectives were to (1) identify important large-scale and small-scale habitat features that influence the occurrence of each fish SGCN; (2) predict the occurrences of fish SGCN by using habitat variables measured at a large scale, a small scale, and both scales (i.e., multiple-scale variables); and (3) evaluate the relative influence of large-scale and small-scale habitat variables on fish species occurrences. Furthermore, because inaccurate species distribution models may be detrimental to conservation efforts, models were validated by using an independent data set to gauge confidence in model predictions and test model generality. We expected that (1) the habitat features with the greatest influence on species occurrence would vary among species due to species-specific habitat associations and (2) species occurrence models that included habitat variables measured at multiple spatial scales would explain the most variability in species occurrence and would exhibit the greatest model generality.

METHODS

Study area and field sampling.—Fish assemblages and small-scale physical habitat characteristics were sampled from 84 wadeable (i.e., second through fifth order) Iowa stream reaches during the spring and summer (May–August) of 2009 and 2010 (Figure 1). Rather than using a simple random sampling design, stream reaches were selected to maximize the number of reaches sampled for each SGCN in three stream segment categories for a concurrent study (Sindt et al. 2012). The three stream segment categories within each species' respective distribution included (1) previously sampled stream segments where the species was documented to be present within the last 50 years (i.e., since 1958); (2) stream segments that had not been previously sampled but where the species was predicted to be present by an existing species distribution model (Iowa Aquatic Gap Analysis Project; Loan-Wilsey et al. 2005); and (3) stream segments that had not been previously sampled and where the species was not predicted to be present. Sixty-eight Iowa fish species are classified as SGCN, and knowledge about habitat associations for all SGCN is needed to guide conservation efforts (Zohrer 2005). However, only seven species were present in a sufficient number of the 84 sampled stream reaches to allow for modeling:

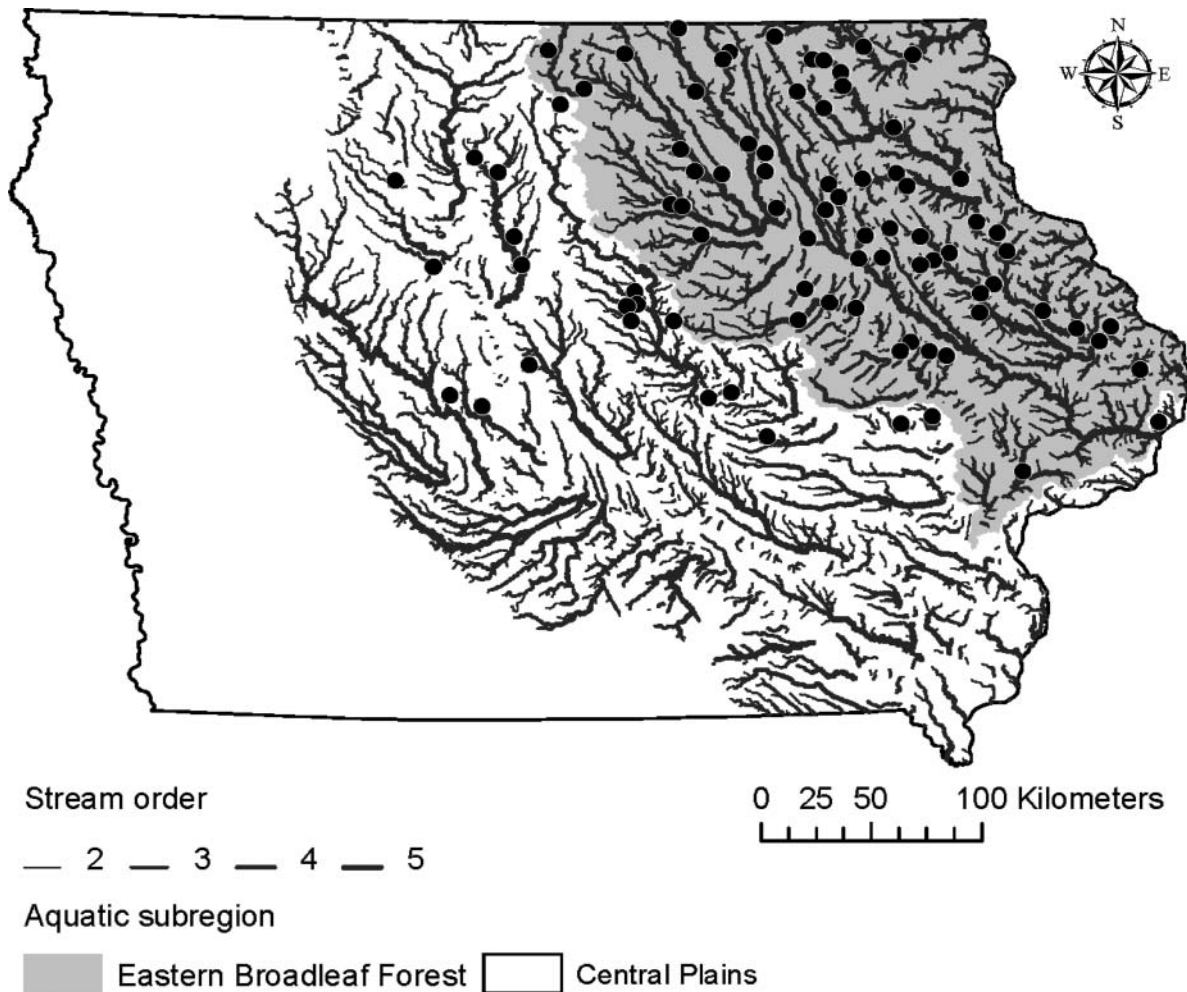


FIGURE 1. Locations of 84 wadeable (second through fifth order) stream reaches sampled in the Mississippi River drainage of Iowa during spring and summer (May–August) 2009 and 2010.

the banded darter *Etheostoma zonale*, American brook lamprey *Lampetra appendix*, Ozark minnow *Notropis nubilus*, blackside darter *Percina maculata*, southern redbelly dace *Phoxinus erythrogaster*, longnose dace *Rhinichthys cataractae*, and central mudminnow *Umbra limi*.

All sampled stream reaches were located in the Mississippi River drainage of Iowa, which contains 34 eight-digit hydrologic unit code (HUC) watersheds (HUCs designated by the U.S. Geological Survey) and portions of seven ecoregions (Des Moines Lobe, Loess Hills and Steeply Rolling Prairies, Southern Iowa Rolling Loess Prairies, Central Irregular Plains, Iowan Surface, Paleozoic Plateau, and Interior River Lowland; Griffith et al. 1994). The 84 stream reaches were sampled from 18 watersheds and included reaches from the Iowan Surface ($n = 47$ reaches), Des Moines Lobe ($n = 15$), Southern Iowa Rolling Loess Prairies ($n = 14$), and Paleozoic Plateau ($n = 8$) ecoregions. Stream reaches were also sampled from both of Iowa's aquatic subregions (i.e., Eastern Broadleaf Forest and Central Plains), which have unique physiographic characteristics and

riverine assemblages (Sowa et al. 2004). Aquatic subregions are similar to the ecoregion provinces defined by Bailey (1995), but the boundaries were delineated by the Missouri Resource Assessment Partnership (University of Missouri, Columbia) to align with drainage divides. Twenty-one of the sampled stream reaches were in the Central Plains aquatic subregion, which is characterized by thick loess deposits over flat to gently sloping terrain with wide stream valleys. Many Central Plains streams were historically dominated by fine silt and sand substrates, and sediment input has been exacerbated by the conversion of native prairie to cropland. The other 63 stream reaches were sampled in the Eastern Broadleaf Forest aquatic subregion of northeastern Iowa; this aquatic subregion is also highly altered from agricultural practices but is geologically more diverse than the Central Plains subregion. The Eastern Broadleaf Forest contains the Paleozoic Plateau ecoregion, which was less impacted by glaciation and is characterized by high topographic relief, rocky outcroppings, and dense forests (Griffith et al. 1994). Overall, Iowa's landscape is relatively homogeneous, with

elevation only varying from 146 to 509 m, and 72% of the land has been converted to row crop agriculture (U.S. Department of Agriculture 2007). Previous studies have found increasing aquatic ecological integrity (e.g., index of biotic integrity) from southwest Iowa to northeast Iowa (Wilton 2004; Heitke et al. 2006). Greater ecological integrity in northeast Iowa is likely a reflection of relatively high-quality habitat due to the greater topographic relief, less-disturbed riparian habitats, and coarser substrates than in other regions of Iowa.

Sample reaches were 300–400 m in length and were divided into macrohabitat units (riffles, pools, runs, or off-channel units). When feasible, sample reaches were selected to encompass as many different macrohabitat units as possible and to be more than 100 m from an artificial structure (e.g., a bridge or low-head dam). All sampling occurred during base flow conditions to minimize sampling inefficiencies associated with high flow, depth, and turbidity. Fish assemblages were sampled in each macrohabitat separately by using single-pass upstream electrofishing with a pulsed-DC electrofishing unit (Simonson and Lyons 1995). When feasible, a generator-powered, barge-mounted electrofishing unit with three anodes (Model VVP-15B; Smith-Root, Inc., Vancouver, Washington) was used. However, if streams were too shallow or were inaccessible, a battery-powered backpack electrofishing unit (Smith-Root LR-20) was used. For both backpack and barge-mounted electrofishing, three netters used 6.34-mm-mesh dip nets to collect fish. An effort was made to sample all accessible habitat types in each macrohabitat. Voltage output was adjusted to maximize efficiency and reduce incidental mortality in each sample reach. Fish that were collected in each macrohabitat unit were examined for external abnormalities, identified, counted, and released. Up to five voucher specimens of each SGCN were preserved in a 10% solution of formalin. Fish that could not be identified in the field were preserved and transported to the laboratory for identification.

Habitat characteristics were quantified by measuring physical habitat features in each macrohabitat unit. The length of each macrohabitat unit was measured along the thalweg, and the width of each macrohabitat unit was measured along a transect perpendicular to the thalweg at 25, 50, and 75% of the macrohabitat length. Depth, current velocity, and substrate size were measured at 20, 40, 50, 60, and 80% of the length of each transect. Current velocity was measured with a portable velocity meter (Flo-Mate Model 2000; Marsh-McBirney, Inc., Frederick, Maryland) at 60% of the depth when depth was less than 0.75 m; at depths greater than 0.75 m, velocities were measured at 20% and 80% of the depth and the two values were averaged (Buchanan and Somers 1969). Substrate was classified as artificial (e.g., tire), soil, wood (e.g., submerged tree), detritus, hardpan, clay (<0.004 mm), silt (0.004–0.063 mm), sand (0.062–2.000 mm), gravel (2–16 mm), coarse gravel (16–64 mm), cobble (64–256 mm), boulder (>256 mm), or bedrock (i.e., modified Wentworth scale). Canopy cover was measured along each transect by using a spherical densiometer (1) at

and facing each streambank and (2) facing upstream and downstream at the midpoint of the transect. Bank cover characteristics (percent coverage of woody vegetation, nonwoody vegetation, roots, boulders, eroding ground, and bare ground) and distance to anthropogenic disturbance (on the bank, ≤ 10 m from the bank, > 10 m from the bank, or absent) were visually estimated for the left and right banks of each transect. One length measurement, three width measurements, and three depth measurements were recorded for each unit of instream cover (artificial structure, boulder, rip-rap, filamentous algae, aquatic macrophytes, terrestrial vegetation, overhanging vegetation, undercut bank, island, small brush, tree root, and large woody debris) at least 0.3 m in length within each macrohabitat.

The area of each macrohabitat unit was estimated by multiplying the thalweg length by the mean width. Mean depth, width, current velocity, canopy cover, and bank coverage percentages were calculated for each macrohabitat unit. Additionally, the CV in depth, width, current velocity, and canopy cover was calculated ($CV = 100 \times [SD/mean]$). The proportions of each substrate type and distance to disturbance category were also quantified for each macrohabitat unit. All habitat characteristics (except instream cover, which was censused and not estimated) were averaged for each macrohabitat category (riffle, pool, run, and off-channel) within each stream reach. Furthermore, averaged values were weighted by the proportion of the total stream reach area that was represented by the respective macrohabitat category. Weighted values were summed to quantify habitat characteristics for the entire stream reach. Aerial coverage of each instream cover type was quantified as the proportion of reach area. Additional composite variables were created by summing two or more habitat variables (e.g., proportion of coarse substrates).

Independent variables.—Large-scale and small-scale habitat variables characterizing sampled stream reaches were selected as independent variables for SGCN occurrence models. Sixteen GIS-measured variables that were used to develop fish distribution models for the Iowa Aquatic Gap Analysis Project (Loan-Wilsey et al. 2005) were obtained using ArcMap version 9.3 (Environmental Systems Research Institute, Redlands, California) and were considered candidate large-scale variables. The 16 variables included one or more measures of flow (intermittent or permanent), stream order (Strahler 1957), channel gradient, elevation, soil type, aquatic subregion (Eastern Broadleaf Forest or Central Plains), temperature (coldwater or warmwater), and size discrepancy with downstream segment. The temperature variable was derived from the Iowa Department of Natural Resources' coldwater streams designation, which is based on biotic communities and maximum summer temperatures. Candidate small-scale habitat variables characterizing channel morphology (e.g., width, depth, and macrohabitat composition), current velocity, canopy cover, bank cover, substrate composition, instream cover (e.g., woody debris and boulders), and distance to disturbance (e.g., row crop agriculture, pasture, and road) were obtained from field sampling.

TABLE 1. Large-scale and small-scale habitat variables for 84 Wadeable Iowa stream reaches that were sampled during spring and summer 2009 and 2010; these variables were used as independent variables in fish species occurrence models (min = minimum; max = maximum).

Variable	Description	Mean	SD	Min	Max
Large-scale variables					
Temperature	Factor (1 = coldwater; 2 = warmwater)	1.93	0.26	1.00	2.00
Subregion	Factor (1 = Central Plains; 2 = Eastern Broadleaf Forest)	1.75	0.44	1.00	2.00
Order	Strahler stream order	3.11	0.76	2.00	5.00
Elevation	Elevation (m) of the upstream end of the stream segment	295.27	45.43	189.00	387.00
Gradient	Measure of reach gradient (1 = 0.0–0.4 m/km; 2 = 0.5–1.0 m/km; 3 = 1.1–2.0 m/km; 4 = 2.1–3.0 m/km; 5 = 3.1–5.0 m/km; 6 = 5.1–7.0 m/km; 7 = 7.1–10.0 m/km; 8 = 10.1–13.0 m/km; 9 = 13.1–17.0 m/km; 10 = ≥ 17.1 m/km)	3.18	1.46	1.00	8.00
Small-scale variables					
Width	Mean wetted channel width (m)	10.84	7.07	2.28	40.95
CVWidth	Mean CV of wetted channel width	14.93	5.91	4.69	32.29
Depth	Mean depth (m)	0.38	0.16	0.08	0.85
CVDepth	Mean CV of depth	39.87	12.50	10.97	68.37
Velocity	Mean current velocity (m/s)	0.30	0.13	0.09	0.65
Pool	Percentage of reach area as pool	5.43	10.14	0.00	53.31
Riffle	Percentage of reach area as riffle	17.24	17.98	0.00	74.21
Canopy	Mean canopy cover (%)	39.69	24.21	0.03	87.30
PctBankCover	Percentage of total bank covered (woody, nonwoody, boulder or rip-rap, and roots)	50.61	16.38	18.33	100.00
Coarse	Percentage of substrate that is coarse (coarse gravel, cobble, and boulder)	21.15	25.58	0.00	88.51
Dist10m	Percentage of banks with disturbance on the bank or within 10 m of the bank	25.69	38.89	0.00	100.00
WoodyCover	Percentage of reach area with woody debris, tree roots, or small brush as cover	5.75	5.86	0.00	23.04
VegCover	Percentage of reach area with filamentous algae, aquatic macrophytes, overhanging vegetation, or terrestrial vegetation as cover	5.96	10.33	0.00	50.06

Many of the large-scale variables were redundant (e.g., three variables described stream size). To avoid multicollinearity, the most ecologically relevant and interpretable variable from redundant groups was retained. The variable describing flow was excluded because streams with intermittent flow were not sampled. Pearson's product-moment correlations were calculated for all pairs of large-scale variables to ensure that no highly correlated (Pearson's $r \geq |0.70|$) variables were retained.

Small-scale habitat characteristics have been found to be associated with fish assemblage characteristics in Iowa (Rowe et al. 2009b). Thus, 62 small-scale variables were created to describe local physical habitat features. Reducing the number of variables was essential for removing multicollinearity between variables and improving model interpretability. Ecologically relevant small-scale variables were selected as potential explanatory variables. Pearson's product-moment correlations were calculated for all pairs of small-scale variables. When two

or more ecologically relevant variables were highly correlated, the most relevant and interpretable variable was retained.

Habitat variables that were used to develop species occurrence models included 5 large-scale variables and 13 small-scale variables (Table 1). The large-scale variables included temperature, subregion, stream order, elevation, and gradient. Temperature and subregion were categorical variables and therefore were treated as factors rather than as continuous variables in model development. The 13 small-scale variables included measures of channel morphology, current velocity, canopy cover, bank cover, substrate composition, instream cover, and distance to disturbance.

Species occurrence models.—We used an information theoretic approach (Burnham and Anderson 2002) to evaluate competing multiple logistic regression models and to identify habitat variables measured at multiple spatial scales that influence the occurrences of the seven fish SGCN. Multiple logistic

TABLE 2. Model performance measures for large-scale, small-scale, and multiple-scale models of occurrence for seven Iowa fish species of greatest conservation need (n = number of stream reaches sampled in each species' historical distribution; CCR = correct classification rate expressed as a proportion; κ = Cohen's kappa value; P is associated with the null hypothesis that $\kappa = 0$).

Species	n	Prevalence	Large-scale models			Small-scale models			Multiple-scale models		
			CCR	κ	P	CCR	κ	P	CCR	κ	P
Banded darter	80	0.40	0.78	0.54	<0.001	0.80	0.60	<0.001	0.83	0.64	<0.001
American brook lamprey	53	0.32	0.51	0.14	0.118	0.55	0.21	0.041	0.58	0.27	0.011
Ozark minnow	63	0.25	0.59	0.27	0.007	0.70	0.39	0.001	0.79	0.51	<0.001
Blackside darter	83	0.27	0.72	0.37	0.001	0.64	0.33	<0.001	0.80	0.54	<0.001
Southern redbelly dace	77	0.55	0.77	0.51	<0.001	0.71	0.40	<0.001	0.79	0.58	<0.001
Longnose dace	50	0.20	0.92	0.75	<0.001	0.92	0.77	<0.001	0.96	0.88	<0.001
Central mudminnow	34	0.21	0.85	0.47	0.047	0.91	0.74	<0.001	0.88	0.70	<0.001

regression analysis is a common multivariate approach for predicting the binary response of fish species presence or absence (e.g., Porter et al. 2000; Rich et al. 2003; Rashleigh et al. 2005). Using retained variables (i.e., the 5 large-scale variables and 13 small-scale variables), candidate multiple logistic regression models were created for all possible combinations of variables measured at a large scale, a small scale, and multiple scales (i.e., both large-scale and small-scale variables). Because large-scale models could include a maximum of five variables, up to five variables were allowed in all candidate models to eliminate the possibility of small-scale and multiple-scale models explaining more variation than large-scale models simply due to the inclusion of more variables. Furthermore, to prevent model overfitting and spurious results, the maximum number of variables allowed in a candidate model was 10% of the number of stream reaches that were sampled within the species' respective distribution (Table 2). For example, we sampled 34 stream reaches within the historical distribution of the central mudminnow, and thus we only allowed up to three variables to be included in models predicting the occurrence of central mudminnow.

Confidence model sets were selected from among all candidate models for each model type (large-scale, small-scale, and multiple-scale models) based on Akaike's information criterion (AIC), and these confidence model sets were used to create model-averaged models. The AIC reflects model parsimony by measuring the goodness of fit while penalizing for the number of parameters (Burnham and Anderson 2002). Recently, the use of AIC corrected for small sample size (AIC_c) has been common; however, Richards (2005) found that AIC_c did not increase the likelihood of selecting the best-approximating model. Thus, we used AIC rather than AIC_c . Richards (2005) also suggested that an AIC difference (ΔAIC) between 4 and 7 should be used as a selection criterion for 95% confidence that the best-approximating model is included in the confidence model set. Thus, all candidate models with a ΔAIC of 6 or less were included in a given confidence model set. Furthermore, because more than one combination of variables was likely to have evidence of being the best-approximating model, all mod-

els that were included in confidence model sets were averaged to account for model selection uncertainty (Burnham and Anderson 2002). Model-averaged coefficients were calculated by weighting the coefficient values for each model in the confidence model set by the models' respective Akaike weights. The relative importance of habitat variables in confidence model sets was assessed by summing the Akaike weights for all models within the confidence model set in which the variable of interest was included. Variables with summed Akaike weights of 0.60 or greater were interpreted as important. Summed relative weights must be interpreted carefully because they reflect the importance of the habitat variable but not the strength of the relationship (i.e., they do not account for coefficient values). All multiple logistic regression analyses were performed with R software (R Development Core Team 2009).

Logistic regression models predict the probability of species presence as a value ranging from 0 to 1. Thresholds of 0.5 are often used in species presence-absence models, with values greater than 0.5 being interpreted as presence and values less than 0.5 being interpreted as absence (e.g., Porter et al. 2000; Pont et al. 2005; Ruiz and Peterson 2007). However, 0.5 is not always the most appropriate threshold because species prevalence can bias logistic regression scores toward the larger group (i.e., present or absent; Fielding and Bell 1997). We used an approach similar to that described by Olden and Jackson (2001) to select optimum thresholds for maximizing the area under the curve of the receiver operating characteristic plot for the training data (i.e., the same data that were used to create the model). The receiver operating characteristic plot is a graph of model sensitivity versus $1 - \text{specificity}$, where sensitivity is the proportion of observed presences that are correctly predicted and specificity is the proportion of observed absences that are correctly predicted. Optimal threshold values were chosen by assuming that the costs of misclassifying a species as absent or present were equal. The optimal threshold value and model-averaged predictions for each species occurrence model were used to calculate confusion matrix values and performance measures to self-evaluate model performance after re-substituting the same

data that were used to parameterize the model (Fielding and Bell 1997). The correct classification rate (CCR) is a simple measure of the percentage of cases that are correctly predicted by a model, but its interpretation can be biased by species prevalence (Olden et al. 2002). Therefore, the primary statistic used to self-evaluate model performance was Cohen's kappa (κ), an index used to assess the correct classification of events (presence or absence) relative to that expected by random chance (Cohen 1960). Kappa values less than or equal to zero indicate that model performance is no better than random chance, whereas a κ value of 1 indicates perfect model performance. Landis and Koch (1977) arbitrarily characterized κ values of less than 0.40 as indicating poor to fair model performance and κ values greater than 0.40 as indicating moderate to near-perfect model performance. Cohen's κ is commonly used to evaluate fish species presence-absence models (e.g., Olden and Jackson 2001; Rashleigh et al. 2005; Hayer et al. 2008) and is also used in the medical and remote sensing fields (Congalton 1991; Manel et al. 2001). Kappa has limitations and may be overly sensitive to species prevalence (McPherson et al. 2004; Vaughan and Omerod 2005); however, Manel et al. (2001) found that κ was a robust indicator of model performance and was negligibly influenced by species prevalence. The SE, significance, and 95% confidence interval of κ can be calculated and used to test the null hypothesis that κ is equal to zero (Titus et al. 1984). The κ values were calculated and significance tests ($\alpha = 0.05$) were performed using the `fmsb` package in R (R Development Core Team 2009). The CCRs and κ values were used to assess the performance of models and to evaluate the relative influence of habitat variables measured at each spatial scale on species occurrences.

Model validation.—The most robust model validation techniques use data that are independent from those used to estimate the model parameters. Therefore, independent data were used to validate and test the generality of species occurrence models. Data were obtained from a previous study by Rowe et al. (2009b) in which fish assemblages were sampled via the Iowa Department of Natural Resources' wadeable stream bioassessment protocol (Wilton 2004) and physical habitats were sampled by following the wadeable streams physical habitat protocol of the U.S. Environmental Protection Agency's Environmental Monitoring and Assessment Program (EMAP; Peck et al. 2006). Rowe et al. (2009b) analyzed data from 93 randomly selected wadeable streams across Iowa. However, only 65 of the wadeable streams were sampled from the Mississippi River drainage to which our models were constrained. Similar methods were used to sample fish, but the methods used to quantify habitat characteristics were slightly different from those used to collect data for our study; however, in most applications it is unlikely that data will always be collected with the exact same protocol. Thus, our use of data collected with differing methodologies offers an opportunity to validate the models with an independent data set and to test the models' generality for use with data collected by different

methods, thereby presenting a "conservative" assessment of model accuracy.

Even though the habitat sampling methodologies of this study were not identical to those used to collect the independent data, many of the same small-scale habitat features were quantified, and the large-scale variables could be easily obtained with a GIS (ArcMap version 9.3). Rather than measuring depths across the width of the stream reach, the EMAP protocol only measures depths along the stream thalweg. To adjust the mean thalweg depth to better represent mean depth, linear regression analysis was used to estimate mean macrohabitat depth from the mean of the maximum macrohabitat transect depths by using data collected from all 84 stream reaches in our study. The linear regression explained a large amount of the variation in mean depth, with an r^2 of 0.91 ($P < 0.0001$). Therefore, mean depth was estimated using the equation

$$\text{mean depth} = (0.773 \times \text{mean thalweg depth}) - 0.0224.$$

Similarly, the EMAP protocol does not include measurement of current velocity. Therefore, we estimated mean current velocity by using a linear regression equation developed with data collected during our study:

$$\begin{aligned} \text{mean current velocity} = & (0.0073 \times \text{mean channel width}) \\ & + 0.2209. \end{aligned}$$

Mean channel width explained a small amount of the variation in mean current velocity ($r^2 = 0.17$; $P < 0.0001$), but this was the strongest relationship between it and any of the other habitat variables. We argue that using this equation to estimate mean current velocity was more appropriate than assuming a constant mean current velocity across sites. Among the small-scale variables that were used to develop our species occurrence models, the only one that could not be obtained from the EMAP protocol for the independent data set was the proportion of banks with disturbance on the bank or within 10 m of the bank. Because missing values resulting from different sampling protocols is a real-world scenario, all values for the distance to disturbance variable were set to zero to simulate missing data. We hypothesize that having disturbance within 10 m of the bank is likely to bias the models toward predicting the species as absent, and thus we will obtain a conservative estimate of model performance.

Model-averaged multiple logistic regression models developed with large-scale, small-scale, and multiple-scale data were used to predict the presence or absence of the seven SGCN by using the habitat variables from the independent data set. Optimal thresholds were used to determine whether each species was predicted to be present or absent in stream reaches that were sampled within the species' respective distribution. Model-predicted presences and absences were compared with the actual presence or absence of each species and were used to calculate

TABLE 3. Model-averaged parameters and optimal threshold values for large-scale (LS), small-scale (SS), and multiple-scale (MS) multiple logistic regression models of species occurrence. Multiple logistic regression equations are in the form of $P = e^{g(x)} / [1 + e^{g(x)}]$, where P = probability of occurrence and $g(x) = b_0 + b_1x_1 + b_2x_2 + b_kx_k$ (b_0 is the intercept, b_1 to b_k are coefficients, and x_1 to x_k are the variables). If P is greater than the optimal threshold value, the species is predicted to be present. See text for definition of optimal threshold; variables are defined in Table 1.

Variable	Banded darter		American brook lamprey		Ozark minnow		Blackside darter		Southern redbelly dace		Longnose dace		Central mudminnow	
	LS or SS	MS	LS or SS	MS	LS or SS	MS	LS or SS	MS	LS or SS	MS	LS or SS	MS	LS or SS	MS
LS or MS optimal threshold	0.378	0.467	0.145	0.056	0.248	0.303	0.314	0.278	0.442	0.578	0.336	0.248	0.375	0.273
LS or MS intercept	-9.952	-4.896	-14.918	-17.848	-16.699	-12.588	-11.007	-4.918	8.862	4.039	-12.669	-10.391	-3.643	0.540
SS optimal threshold	0.316		0.206		0.206		0.212		0.417		0.282		0.352	
SS intercept	-2.356		-3.698		-4.676		3.484		-0.417		-4.478		-4.795	
LS variables														
Temperature 2 (warmwater)	0.205	0.201	0.331	-0.065	-0.030	-0.237	10.644	5.419	-9.355	-5.462	-0.203	-1.435	3.567	0.243
Subregion 2 (Eastern Broadleaf Forest)	0.225	0.117	0.642	0.209	13.778	4.904	-0.120	-0.001	4.152	4.103	8.802	1.971	0.074	0.000
Order	1.542	0.280	1.023	0.018	-0.045	-0.043	-0.498	-0.638	-1.055	-0.496	0.521	0.586	0.028	0.000
Elevation	0.014	0.008	0.034	0.053	0.007	0.010	0.007	0.016	0.001	0.004	-0.012	-0.003	0.000	-0.001
Gradient	0.054	-0.014	-0.066	-0.016	0.013	0.000	-0.369	-0.135	0.077	0.030	1.206	1.339	-0.522	-1.168
SS variables														
Width	0.205	0.179	0.107	0.128	0.001	0.000	0.003	0.002	-0.084	-0.082	0.004	0.002	0.026	0.015
CVWidth	0.021	0.012	0.025	0.013	0.012	0.005	-0.003	-0.001	-0.009	-0.003	-0.050	-0.127	0.005	0.001
Depth	0.655	0.301	0.010	0.275	0.128	0.042	1.258	3.044	-0.040	-0.100	0.241	0.122	0.071	0.005
CVDepth	0.027	0.020	0.048	0.010	0.110	0.115	-0.071	-0.059	0.036	0.006	0.013	0.025	0.043	0.018
Velocity	0.014	0.110	-0.454	-0.142	-3.238	-2.036	-3.232	-1.855	-0.346	-0.123	0.293	5.016	-5.712	-7.326
Pool	0.001	0.000	0.016	0.008	-0.049	-0.048	0.001	0.004	0.007	0.005	0.016	0.004	0.033	0.013
Riffle	0.000	0.001	-0.020	-0.001	-0.001	0.000	-0.002	-0.001	0.001	0.002	0.132	0.087	0.000	0.005
Canopy	0.000	0.000	0.000	0.000	-0.002	-0.001	0.005	0.004	-0.002	-0.002	-0.006	-0.003	0.015	0.010
PctBankCover	-0.055	-0.059	-0.001	-0.005	0.007	0.003	-0.033	-0.057	0.009	0.003	-0.004	-0.001	0.035	0.013
Coarse	0.030	0.031	-0.003	-0.003	-0.002	-0.001	0.002	0.001	0.003	0.001	0.002	0.000	0.006	0.028
Dist10m	0.000	0.000	-0.019	-0.002	-0.039	-0.039	-0.001	-0.003	0.002	0.002	0.001	0.000	-0.042	-0.028
WoodyCover	-0.006	-0.005	-0.006	-0.010	0.002	0.009	0.019	0.005	0.003	0.006	-0.053	-0.022	0.004	0.001
VegCover	-0.001	-0.002	-0.038	-0.289	0.003	0.001	-0.004	-0.001	-0.025	-0.008	-0.136	-0.418	-0.002	-0.002

the same model performance measures (CCR and κ value) that were examined for the self-evaluation of model performance.

RESULTS

The number of stream reaches that were sampled within the historical distribution of the seven SGCN varied from 34 for the central mudminnow to 83 for the blackside darter, and frequency of occurrence varied from 0.20 for the longnose dace to 0.55 for the southern redbelly dace (Table 2). Large-scale, small-scale, and multiple-scale model-averaged models were created for all seven species; as expected, the optimal threshold values differed from the commonly used value of 0.5 (Table 3). Multiple-scale models had greater κ values than most of the large- or small-scale models, and small-scale models tended to have greater κ values than large-scale models.

Species occurrence models that were developed with large-scale habitat variables predicted the occurrences of six of the seven species significantly better than chance ($\kappa > 0.0$, $P \leq 0.05$), and κ values varied from 0.14 for the American brook lamprey to 0.75 for the longnose dace (mean $\kappa = 0.44$, SE = 0.07; Table 2). Large-scale models correctly predicted 51–92% of species presences and absences, with a mean CCR of 73% (SE = 5). The sum of Akaike weights for all models in confidence model sets in which a variable occurred indicated that all five large-scale variables were important in at least one of the seven species models (Table 4). Stream order was an important variable in four species models, elevation and gradient were important in three models, subregion was important in two models, and temperature was important in one model. Four large-scale variables were important in the blackside darter model, and only one variable was important in Ozark minnow and central mudminnow models.

TABLE 4. Relative importance and the direction of influence for independent variables in large-scale (LS), small-scale (SS), and multiple-scale (MS) species occurrence models for seven Iowa fish species of greatest conservation need. Relative importance for a given variable is the sum of Akaike weights for all models in the confidence model set (i.e., with Akaike's information criterion difference [ΔAIC] ≤ 6) that included the variable of interest. Values greater than 0.60 (shown in bold italics) provide substantial evidence that the variable is important for predicting the occurrence of the species. Variables are defined in Table 1.

Variable	Banded darter		American brook lamprey		Ozark minnow		Blackside darter		Southern redbelly dace		Longnose dace		Central mudminnow	
	LS or SS	MS	LS or SS	MS	LS or SS	MS	LS or SS	MS	LS or SS	MS	LS or SS	MS	LS or SS	MS
LS variables														
Temperature 2 (warmwater)	0.29	0.13	0.31	-0.08	-0.25	-0.14	0.65	0.32	-0.58	-0.34	-0.30	-0.34	0.25	0.02
Subregion 2 (Eastern Broadleaf Forest)	0.35	0.11	0.44	0.15	0.82	0.29	-0.30	-0.04	1.00	1.00	0.51	0.12	0.24	-0.02
Order	1.00	0.25	0.84	0.11	-0.27	-0.10	-0.67	-0.48	-0.96	-0.50	0.48	0.35	0.24	-0.02
Elevation	0.89	0.49	1.00	1.00	0.59	0.56	0.61	0.73	0.27	0.11	-0.65	-0.17	-0.24	-0.04
Gradient	0.33	-0.08	-0.32	-0.09	0.25	-0.04	-0.80	-0.28	0.37	0.15	1.00	0.67	-0.69	-0.76
SS variables														
Width	0.99	0.87	0.85	0.87	0.12	-0.05	0.13	0.05	-0.92	-0.68	0.20	0.07	0.16	0.09
CVWidth	0.27	0.16	0.27	0.15	0.18	0.08	-0.10	-0.04	-0.19	-0.11	-0.37	-0.57	0.06	0.03
Depth	0.22	0.10	0.15	0.12	0.11	0.05	0.37	0.45	-0.19	-0.12	0.19	0.05	0.04	0.02
CVDepth	0.44	0.33	0.60	0.18	0.99	1.00	-0.94	-0.62	0.72	0.20	0.26	0.21	0.30	0.13
Velocity	0.11	0.07	-0.16	-0.08	-0.48	-0.31	-0.62	-0.31	-0.19	-0.11	0.19	0.39	-0.50	-0.53
Pool	0.11	0.05	0.33	0.16	-0.61	-0.58	0.11	0.10	0.24	0.16	0.31	0.08	0.30	0.12
Riffle	-0.12	0.08	-0.50	-0.09	-0.11	-0.05	-0.14	-0.05	0.14	0.13	1.00	0.68	-0.02	0.07
Canopy	0.10	-0.05	0.10	-0.09	-0.15	-0.07	0.29	0.18	-0.19	-0.15	-0.25	-0.09	0.29	0.19
PctBankCover	-0.92	-0.91	-0.10	-0.16	0.24	0.11	-0.74	-0.91	0.33	0.16	-0.20	-0.05	0.36	0.15
Coarse	0.94	0.91	-0.18	-0.14	-0.14	-0.08	0.18	0.09	0.24	0.11	0.19	0.05	0.14	0.36
Dist10m	-0.10	0.05	-0.84	-0.16	-1.00	-1.00	-0.13	-0.05	0.11	0.18	0.19	0.05	-0.51	-0.29
WoodyCover	-0.15	-0.09	-0.11	-0.14	0.10	0.04	0.27	0.09	0.14	0.14	-0.32	-0.12	0.07	0.02
VegCover	-0.11	-0.06	-0.26	-0.87	0.14	0.06	-0.16	-0.06	-0.52	-0.21	-0.39	-0.42	-0.05	-0.04

Species occurrence models that were developed with small-scale variables predicted the occurrences of all seven species significantly better than chance, with κ values varying from 0.21 for the American brook lamprey to 0.77 for the longnose dace (mean $\kappa = 0.49$, SE = 0.08). Small-scale models correctly classified 55–92% of species presences and absences, with a mean CCR of 75% (SE = 5). Of the 13 small-scale variables that were hypothesized to influence fish species distributions, 9 of the variables were identified as important in small-scale models (Table 4). Mean stream width and mean CV of depth were the variables that were most commonly designated as important; they were identified as important in the small-scale models for three species. The number of important small-scale variables varied from zero in the central mudminnow model to three in the banded darter, Ozark minnow, and blackside darter models. Although the small-scale model performed better than the large-scale model for central mudminnow, no small-scale variables were identified as important; this is likely attributable to the small number of stream segments (i.e., 34) that were sampled within the central mudminnow's distribution, and spurious correlations probably accounted for the explained variation in occurrence.

For six of the seven species, multiple-scale models had a higher κ value and CCR than either the large- or small-scale models (Table 2). The κ values were significantly greater than zero for all multiple-scale models, varying from 0.27 for the American brook lamprey to 0.88 for the longnose dace (mean $\kappa = 0.59$, SE = 0.07). Six of the multiple-scale models had κ values greater than 0.40, suggesting moderate to substantial performance. The CCRs for most of the multiple-scale models were relatively high (79–96%), but the percentage of correctly predicted presences and absences was low (CCR = 58%) for the American brook lamprey (Table 2). As expected, the habitat variables identified as important in multiple-scale models were similar to those identified for the large-scale and small-scale models; however, only three of the large-scale variables were important in at least one multiple-scale model (Table 4). Although stream order was important in the large-scale models for four species, stream width (a small-scale variable) was selected as a more appropriate measure of stream size in the multiple-scale models. Six of the variables that were important in small-scale models were also important in multiple-scale models, and the variable representing the availability of instream vegetation cover (e.g., aquatic macrophytes and overhanging vegetation)

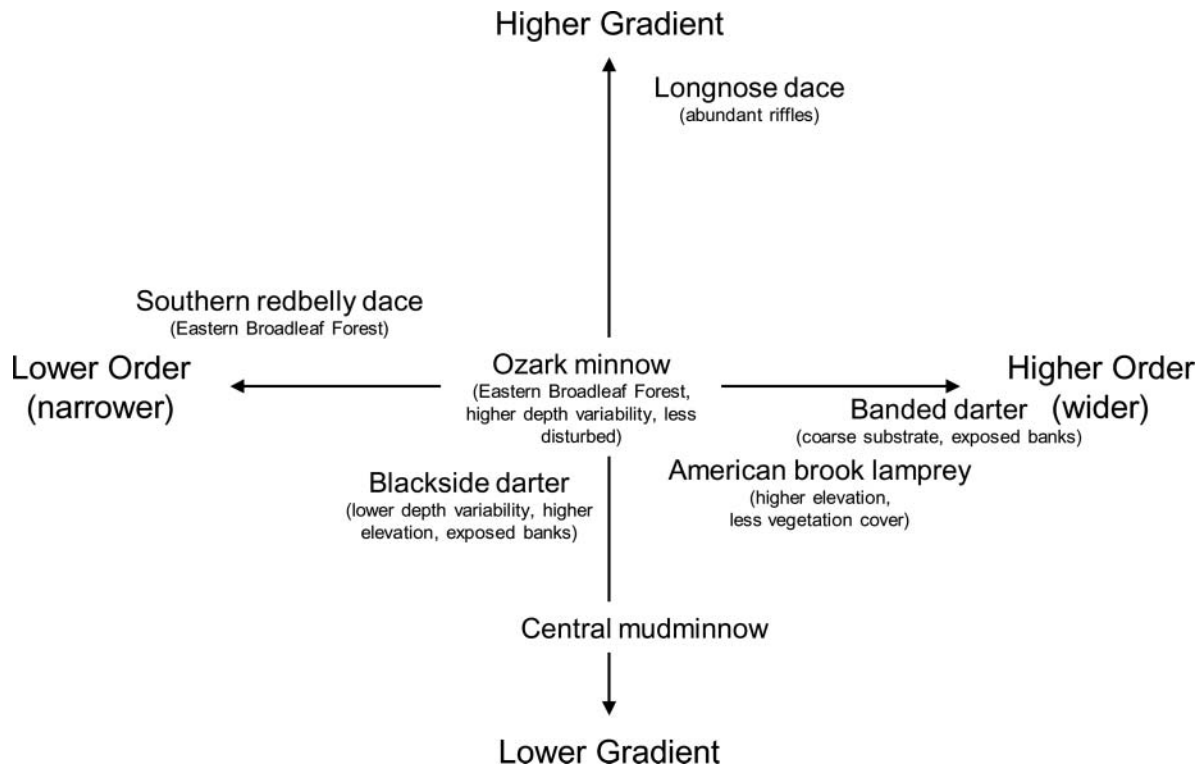


FIGURE 2. Qualitative associations of seven fish species of greatest conservation need with large-scale and small-scale habitat variables measured in Wadeable Iowa streams. Quantitative relationships supporting these associations were primarily obtained from the multiple-scale model summary statistics shown in Tables 4 and 5.

was identified as important in the American brook lamprey's multiple-scale model even though it was not important in the small-scale model.

Multiple-scale models predicted the occurrence of most fish SGCN with the greatest accuracy; thus, the habitat variables that were identified as important in each species' multiple-scale model were weighted the most for interpretation. Stream size (stream order and stream width) or gradient variables were important in multiple-scale models for five of the seven species (Table 4). Thus, the importance of these variables in species occurrence models was used in combination with model-averaged coefficient values to conceptualize the relative associations with the occurrence of each species (Figure 2). For example, the occurrence of southern redbelly dace was associated with narrow streams, whereas the occurrence of banded darters and American brook lampreys was associated with larger streams. Similarly, the occurrence of longnose dace was associated with high-gradient streams, while central mudminnow occurrence was associated with low-gradient streams. Neither stream size nor gradient was important in the multiple-scale models for the blackside darter and Ozark minnow, but the large-scale model showed that blackside darter occurrence was associated with smaller, low-gradient streams. Furthermore, up to three other variables were identified as important in each species' multiple-scale model.

Stream reaches in the independent data set were sampled from 28 of the 34 eight-digit HUC watersheds in the Mississippi River drainage of Iowa. From 15 to 54 stream reaches were sampled within the selected species' historical distributions and each species was collected in 1 (Ozark minnow and central mudminnow) to 16 (southern redbelly dace) stream reaches (Table 5). When applied to the independent data set, large-scale models were able to predict the occurrences of banded darters, American brook lampreys, southern redbelly dace, and longnose dace significantly better than random chance, but they could not predict the occurrence of Ozark minnow, blackside darters, or central mudminnow better than random chance (Table 5). Although small-scale models performed better than large-scale models for the model development data set, when the small-scale models were applied to the independent data set, they only predicted the occurrences of banded darters and longnose dace better than would be expected by random chance. When variables measured at both the large and small scales were included in the multiple-scale models, the presence or absence of Ozark minnow, blackside darters, and central mudminnow was still not predicted more accurately than would be expected by chance. However, multiple-scale models predicted the occurrences of banded darters, southern redbelly dace, and longnose dace better than large- or small-scale models.

TABLE 5. Model performance measures for large-scale, small-scale, and multiple-scale species occurrence models when validated against an independent data set collected from 65 wadeable stream reaches in Iowa (n = number of stream reaches sampled in each species' historical distribution; CCR = correct classification rate expressed as a proportion; κ = Cohen's kappa value; P is associated with the null hypothesis that $\kappa = 0$).

Species	n	Prevalence	Large-scale models			Small-scale models			Multiple-scale models		
			CCR	κ	P	CCR	κ	P	CCR	κ	P
Banded darter	40	0.20	0.68	0.33	0.021	0.83	0.48	0.016	0.90	0.66	0.004
American brook lamprey	26	0.23	0.77	0.52	0.006	0.42	0.08	0.309	0.73	0.41	0.026
Ozark minnow	28	0.04	0.50	0.06	0.363	0.54	0.07	0.353	0.82	0.24	0.242
Blackside darter	54	0.26	0.63	0.19	0.105	0.67	0.17	0.152	0.72	0.18	0.169
Southern redbelly dace	40	0.41	0.66	0.35	0.009	0.51	0.10	0.236	0.71	0.44	0.002
Longnose dace	35	0.23	0.74	0.40	0.019	0.86	0.53	0.018	0.89	0.68	0.002
Central mudminnow	15	0.07	0.87	0.44	0.167	0.27	-0.14	0.764	0.73	0.25	0.236

DISCUSSION

An understanding of factors that influence the distribution of fish species is vital for conservation success, and models that predict species occurrences may serve as a useful tool for fisheries managers. We developed models that accurately predicted the occurrences of all seven Iowa fish SGCN and identified important associations with both large-scale and small-scale habitat features. Most species occurrences were effectively predicted by using either large- or small-scale variables, and the most influential habitat features and spatial scale varied by species. As hypothesized, the inclusion of variables that were measured at both spatial scales predicted the occurrences of fish SGCN with the greatest accuracy.

Physical and biological changes along the longitudinal gradients of streams are well documented (Vannote et al. 1980), and large-scale environmental gradients have been shown to influence the longitudinal distribution of fish species (Rahel and Hubert 1991; Quist et al. 2004). Consistent with the results of others, we found that large-scale measures of stream order (Paller 1994), elevation (Quist et al. 2004), channel gradient (Pont et al. 2005), temperature (Buisson et al. 2008; Lyons et al. 2010), and aquatic subregion each had an important influence on the distribution of at least one of the seven fish species. Stream order, elevation, and gradient were among the most important variables in large-scale models, suggesting that Iowa stream fish assemblages may be strongly influenced by longitudinal gradients and stream size thresholds. Similarly, Quist et al. (2004) showed that fishes persisted within specific elevation and stream width boundaries in Wyoming streams. In the present study, we found that southern redbelly dace were common in small second- and third-order streams, but banded darters and American brook lampreys were never collected from a stream that was less than third order. Paller (1994) reported similar associations between fish assemblages and stream order in following patterns of species additions and replacements. Models for three Iowa fish SGCN identified important associations with stream gradient. Longnose dace were generally sampled from higher-gradient streams, including the two highest gradient streams

sampled (i.e., 7.0–13.0 m/km). In contrast, blackside darters and central mudminnow were only sampled in low-gradient streams for which gradient did not exceed 3.0 m/km. The positive association between longnose dace occurrence and gradient likely reflects this species' similar positive association with riffle habitat availability. This result is consistent with what others have reported as common longnose dace habitat in Iowa (Harlan and Speaker 1969). Central mudminnow are suggested to occupy habitats with little to no current, silty substrates, and usually the presence of vegetation (Harlan and Speaker 1969; Martin-Bergmann and Gee 1985); thus, high-gradient streams would not be expected to provide suitable habitat conditions for central mudminnow.

Iowa topography is relatively gentle, and the elevation of streams sampled for this study only varied from 189 to 387 m. Thus, the importance of elevation in the occurrence models for four of the species was unexpected. Quist et al. (2004) suggested that elevation acts as a proxy for temperature, which influences the survival, growth, reproduction, and distribution of fish (Buisson et al. 2008; Lyons et al. 2010; Robinson et al. 2010). Although elevation probably does not act as a surrogate measure for temperature in Iowa, differences in elevation likely correspond to differences in underlying geology and climate across the state. The subregion variable similarly differentiates between two Iowa regions that have different underlying physiographic features. Heitke et al. (2006) showed that fish assemblages and physical habitat conditions vary significantly across Iowa's ecoregions. Subregions may also represent regions with unique aquatic biodiversity as a result of processes such as speciation, colonization, and extinction. For example, there is a unique group of fishes—including the Ozark minnow—that exhibits two disjunct distributions: one distribution extends into the Eastern Broadleaf Forest subregion of Iowa, and the other is found in the northern Ozarks of southern Missouri and surrounding states. Berendzen et al. (2010) showed that this unique, disjunct distribution is likely attributable to the expansion of fishes from the northern Ozarks into the Paleozoic Plateau of Iowa after glacial periods, followed by isolation of the two populations

as suitable habitats between the two regions were lost due to glacial deposition of loess and till. Our results were consistent with this hypothesis and showed that the occurrences of Ozark minnow and southern redbelly dace were positively associated with the Eastern Broadleaf Forest subregion. Our results also suggest that southern redbelly dace are not constrained by specific small-scale habitat features within stream reaches where large-scale habitat conditions are suitable. These results are not surprising, as southern redbelly dace are recognized as being prevalent in headwater streams of the Eastern Broadleaf Forest subregion in Iowa (Harlan and Speaker 1969). Thus, the sole use of GIS-measured variables representing stream order and subregion can accurately predict southern redbelly dace occurrence in the Mississippi River drainage of Iowa. In contrast, even though the distribution of Ozark minnow is constrained to northeast Iowa (Harlan and Speaker 1969), small-scale variables predicted the occurrence of Ozark minnow better than the subregion variable or any combination of large-scale variables. Ultimately, the multiple-scale model was able to predict the occurrence of Ozark minnow with the greatest accuracy and indicated that Ozark minnow occurrence was positively associated with depth variability and negatively associated with proximity to anthropogenic disturbance. This is consistent with the work of Pflieger (1997) who reported that Ozark minnow inhabit areas of slow current in fast-flowing streams, which is indicative of high depth variability.

Similar to large-scale variables, we found that many small-scale variables were important for predicting the occurrence of fish SGCN. Understanding the small-scale physical habitat requirements of stream fish species has long been a central theme of fish ecologists (Schlosser 1982; Hubert and Rahel 1989). Early work by Gorman and Karr (1978) related fish species to specific instream physical habitat measures (e.g., depth, substrate, and velocity) and showed a positive relationship between species diversity and habitat diversity. Later, Leonard and Orth (1988), among others (e.g., Schlosser 1982; Lobb and Orth 1991), identified guilds of species that were associated with similar instream physical habitat features. Our results were similar in that a variety of small-scale habitat variables (channel width, depth variability, current velocity, coarse substrate, pool and riffle habitat availability, bank characteristics, vegetation, and distance to disturbance) were important in at least one species' occurrence model. The CVs in depth and width variables were assumed to represent habitat complexity and thus the ability to support greater biodiversity (Gorman and Karr 1978). The occurrences of southern redbelly dace and Ozark minnow were positively associated with depth variability; however, blackside darter occurrence was negatively associated with depth variability. The occurrence of blackside darters was also negatively associated with covered banks. The negative association with depth variability and covered banks likely reflects the common occurrences of blackside darters in channelized streams of the Central Plains subregion. As a result of stream straightening, bank erosion generally increases, pool and riffle habitats

become filled and covered with fine sediments, and the variability in water depth and current velocity decreases (Bulkley 1975; Zimmer and Bachmann 1978).

Coarse substrates can provide essential fish spawning habitat and refuge from current and predators. Additionally, coarse substrates provide excellent habitat for macroinvertebrate prey and are often associated with greater ecological integrity and fish biodiversity (Heitke et al. 2006; Rowe et al. 2009b). In our study, the banded darter was the only species for which occurrence was positively associated with the proportion of coarse substrate. Banded darter occurrence was also positively associated with stream width and was negatively related to covered banks. Anecdotal reports are similar to our results and suggest that banded darters occur in deep riffles over rocky substrate (Harlan and Speaker 1969; Cross and Collins 1995).

The spatial scale that is most influential on stream fauna is context dependent and varies by species and system. Monti and Legendre (2009) showed that environmental factors were important in structuring biotic communities in high-flow lotic systems but that biotic interactions were more influential in low-flow lotic systems. Pont et al. (2005) showed that the influences of regional-scale and small-scale processes on species occurrence were species specific rather than system specific. Similarly, we found differences in the relative influence of large-scale and small-scale habitat features on each fish species. Thus, it is important to determine the appropriate scale for management efforts on a case-by-case basis because models that are developed in one region or for one species are unlikely to exhibit generality to another region or species. Many studies have focused on determining the spatial scale that is most influential on stream biotic communities, but the results are inconsistent. For example, Creque et al. (2005) developed models to predict the density of five sport fish species in Michigan rivers by using only GIS-measured landscape variables and only small-scale habitat variables; models that were developed with landscape variables explained more variation than models that were developed with small-scale variables. In contrast, Rich et al. (2003) found that variation in the occurrence of bull trout *Salvelinus confluentus* in a Montana watershed was explained to a greater extent by small-scale abiotic and biotic variables than by large-scale variables. Although either large- or small-scale variables may accurately predict the distribution of fish species, Leftwich et al. (1997) suggested that incorporating variables from multiple spatial scales is likely to improve the generality of models and to provide the greatest interpretive value. Our results support this conclusion in that models incorporating both large-scale and small-scale variables predicted fish species occurrences with the greatest accuracy. Specifically, our multiple-scale models correctly predicted the occurrences of the seven species 58–96% of the time, with κ values varying from 0.27 to 0.88. Pont et al. (2005) similarly used multiple logistic regression models with variables measured at multiple spatial scales to predict the occurrences of 13 common fish species in France; they reported CCRs of 71–92% and κ values from 0.10 to 0.61. When applied

to an independent data set, our multiple-scale models generally exhibited better performance than the large- or small-scale models. Specifically, the occurrences of banded darters, southern redbelly dace, and longnose dace were predicted considerably better by multiple-scale models than by single-scale models. These results emphasize the importance of understanding the influence of factors acting at multiple spatial scales before making important conservation decisions. The largest spatial scale at which factors are constraining species occurrences should be treated as a limiting factor; unless management efforts address issues at that scale, the success of such efforts will be limited. For example, increasing fish species diversity with small-scale habitat manipulations may be unsuccessful if water quality or connectivity issues are the limiting factor (Tonn 1990; Pretty et al. 2003). Results from this study suggest that habitat restoration efforts to create riffle habitats for longnose dace could be unsuccessful if implemented in low-gradient streams.

Despite the large number of habitat variables considered in our species occurrence models, none of the variables provided measures of temporal variability, which may have the ability to increase predictive power. For example, the timing of flooding events may greatly influence fish assemblages (Harvey 1987). In Oregon, Pearson et al. (1992) documented changes in fish assemblages before and after flood events and linked assemblage resilience to flood events with hydraulic complexity. As a result of agricultural practices (e.g., drainage tiling and channelization) in Iowa, the hydrologic regimes of stream systems are highly altered, and the intensity, timing, and frequency of flood events are likely to influence aquatic communities (Meyers et al. 2010). Similarly, the location and magnitude of fish barriers (e.g., dams, culverts, and grade-control structures) can fragment populations and prevent source populations from re-colonizing suitable habitats (Compton et al. 2008; Litvan et al. 2008). Variables that characterize water quality and biotic interactions may also explain further variation in species occurrences. For example, predation and competition have been shown to influence fish assemblage structure (Fausch and White 1981; Power et al. 1985), and variables characterizing the presence of competitors or the abundance of predators have been used to improve predictive power in recent species distribution models (Rich et al. 2003; Quist et al. 2005). The number of factors that potentially influence fish assemblages and species distributions is immeasurable; therefore, understanding the factors that explain the most variation in species occurrences and the linkages between those factors across spatial scales will improve the efficiency of conservation efforts.

Our study is one of the few studies that have used an independent data set to validate fish species occurrence models and to quantify confidence in model predictions (e.g., Leftwich et al. 1997; Steen et al. 2008). Most often, species occurrence models fail to accurately predict species distributions outside of the region for which the models were developed. For instance, Porter et al. (2000) developed models that predicted the occurrence of 13 fish species in the Blackwater River drainage, British

Columbia, with CCRs of 73–90%, but the accuracy of the models greatly decreased when they were applied to data collected in a distant drainage. Results of our model validation with an independent data set showed that large-scale and multiple-scale models were able to predict the occurrences of banded darters, American brook lampreys, southern redbelly dace, and longnose dace better than random chance. In a concurrent study (Sindt et al. 2012), we also found that models created with large-scale explanatory variables (Iowa Aquatic Gap Analysis Project models; Loan-Wilsey et al. 2005) predicted the occurrences of banded darters, southern redbelly dace, and longnose dace better than would be expected by random chance, but this was not the case for nine other species, including the American brook lamprey, Ozark minnow, and blackside darter. Thus, in Iowa, the most effective variables for predicting the occurrences of many fish SGCN, including the Ozark minnow and blackside darter, are still undetermined. The combined results of the present study and our previous study (Sindt et al. 2012) further emphasize the importance of understanding the most appropriate scale for management and conservation efforts on a species-by-species and system-by-system basis. Therefore, we caution that even though several of our models exhibited generality when validated with an independent data set, they might not be effective outside of the Mississippi River drainage in Iowa.

An understanding of factors that shape the distribution of fish species across spatial scales is important for resource managers and is vital to conservation success. For this study, we adopted an exploratory approach to identify abiotic variables measured at two spatial scales influencing the occurrences of fish SGCN in wadeable Iowa streams and to provide insight into species ecology. Our results suggest that managing for stream fish biodiversity requires the protection and restoration of habitat complexity across a broad spectrum of large-scale habitat conditions. The performance of our models complements other studies in which large-scale and small-scale habitat features were found to explain variability in species occurrences. However, we have also shown that the most influential spatial scale and habitat variables are species specific and that the incorporation of variables measured at multiple spatial scales provides the greatest model generality.

ACKNOWLEDGMENTS

We thank Chris Smith, Rebecca Krogman, Josh Bruegge, Collin Hinz, Nick Johnson, and Maria Dzul for field assistance. Comments and suggestions provided by Michael Colvin, Jesse Fischer, Vivekananda Roy, Jennifer Biederman, Christopher Chizinski, and three anonymous reviewers helped to improve this manuscript. We also thank all those involved with creating and contributing to the Iowa Aquatic Gap Analysis Project, especially Anna Loan-Wilsey and Kevin Kane. This project was supported in part by the Department of Natural Resource Ecology and Management at Iowa State University, the

Iowa Cooperative Fish and Wildlife Research Unit, and the Iowa Department of Natural Resources through State Wildlife Grant T-1-R-19. Reference to trade names does not imply endorsement by the U.S. Government.

REFERENCES

- Allan, J. D. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. *Annual Review of Ecology, Evolution, and Systematics* 35:257–284.
- Baily, R. G. 1995. Description of the ecoregions of the United States, 2nd edition revised. U.S. Forest Service, Miscellaneous Publication 1391, Washington, D.C.
- Berendzen, P. B., J. F. Dugan, and T. Gamble. 2010. Post-glacial expansion into the Paleozoic Plateau: evidence of an Ozarkian refugium for the Ozark minnow *Notropis nubilus* (Teleostei: Cypriniformes). *Journal of Fish Biology* 77:1114–1136.
- Buchanan, T. J., and W. P. Somers. 1969. Discharge measurements at gaging stations. Techniques of water-resource investigation, book 3. U.S. Geological Survey, Washington, D.C.
- Buisson, L., L. Blanc, and G. Grenouillet. 2008. Modelling stream fish species distribution in a river network: the relative effects of temperature versus physical factors. *Ecology of Freshwater Fish* 17:244–257.
- Bulkley, R. V. 1975. Inventory of major stream alterations in Iowa. Completion report: a study of the effects of stream channelization and bank stabilization on warm water sport fish in Iowa. U.S. Fish and Wildlife Service, Subproject 1, Contract 14-16-008-745, Ames, Iowa.
- Burnham, K. P., and D. R. Anderson. 2002. Model selection and multi-model inference: a practical information theoretic approach, 2nd edition. Springer, New York.
- Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurements* 20:37–46.
- Compton, R. I., W. A. Hubert, F. J. Rahel, M. C. Quist, and M. R. Bower. 2008. Influences of fragmentation on three species of native warmwater fishes in a Colorado River basin headwater stream system, Wyoming. *North American Journal of Fisheries Management* 28:1733–1743.
- Congalton, R. G. 1991. A review of assessing the accuracy of classification of remotely sensed data. *Remote Sensing of Environment* 37:35–46.
- Creque, S. M., E. S. Rutherford, and T. G. Zorn. 2005. Use of GIS-derived landscape-scale habitat features to explain spatial patterns of fish density in Michigan rivers. *North American Journal of Fisheries Management* 25:1411–1425.
- Cross, F. B., and J. T. Collins. 1995. Fishes in Kansas, 2nd edition. University of Kansas, Natural History Museum, Lawrence.
- Dunham, J., and J. T. Peterson. 2010. Scale and fisheries management. Pages 81–105 in W. A. Hubert and M. C. Quist, editors. *Inland fisheries management in North America*, 3rd edition. American Fisheries Society, Bethesda, Maryland.
- Durance, I., C. Lepichon, and S. J. Omerod. 2006. Recognizing the importance of scale in the ecology and management of riverine fish. *River Research and Applications* 22:1143–1152.
- Fausch, K. D., and R. J. White. 1981. Competition between brook trout (*Salvelinus fontinalis*) and brown trout (*Salmo trutta*) for positions in a Michigan stream. *Canadian Journal of Fisheries and Aquatic Sciences* 38:1220–1227.
- Fielding, A. H., and J. F. Bell. 1997. A review of methods for assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24:38–49.
- Fischer, J. R., and C. P. Paukert. 2008. Habitat relationships with fish assemblages in minimally disturbed Great Plains regions. *Ecology of Freshwater Fish* 17:597–609.
- Frissel, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. *Environmental Management* 10:199–214.
- Gorman, O. T., and J. R. Karr. 1978. Habitat structure and stream fish communities. *Ecology* 59:507–515.
- Griffith, G. E., J. M. Omernik, T. F. Wilton, and S. M. Pierson. 1994. Ecoregions and subregions of Iowa: a framework for water quality assessment and management. *Journal of the Iowa Academy of Science* 101:5–13.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8:993–1009.
- Harlan, J. R., and E. B. Speaker. 1969. Iowa fish and fishing. Iowa Conservation Commission, Des Moines.
- Harvey, B. C. 1987. Susceptibility of young-of-the-year fishes to downstream displacement by flooding. *Transactions of the American Fisheries Society* 116:851–855.
- Hayer, C., S. S. Wall, and C. R. Berry Jr. 2008. Evaluation of predicted fish distribution models for rare fish species in South Dakota. *North American Journal of Fisheries Management* 28:1259–1269.
- Heitke, J. D., C. L. Pierce, G. T. Gelwicks, G. A. Simmons, and G. L. Siegwirth. 2006. Habitat, land use, and fish assemblage relationships in Iowa streams: preliminary assessment in an agricultural landscape. Pages 287–303 in R. M. Hughes, L. Wang, and P. Seelbach, editors. *Landscape influences on stream habitat and biological assemblages*. American Fisheries Society, Symposium 48, Bethesda, Maryland.
- Hubert, W. A., and F. J. Rahel. 1989. Relations of physical habitat to abundance of four nongame fishes in high-plains streams: a test of habitat suitability index models. *North American Journal of Fisheries Management* 9:332–340.
- Hughes, R. M., L. Wang, and P. Seelbach, editors. 2006. *Landscape influences on stream habitat and biological assemblages*. American Fisheries Society, Symposium 48, Bethesda, Maryland.
- Hynes, H. B. N. 1975. The stream and its valley. *Verhandlungen Internationale Vereinigung Limnologie* 19:1–15.
- Jackson, D. A., P. R. Peres-Neto, and J. D. Olden. 2001. What controls who is where in freshwater fish communities—the roles of biotic, abiotic, and spatial factors. *Canadian Journal of Fisheries and Aquatic Sciences* 58:157–170.
- Jelks, H. L., S. J. Walsh, N. M. Burkhead, S. Contreras-Balderas, E. Díaz-Pardo, D. A. Hendrickson, J. Lyons, N. E. Mandrak, F. McCormick, J. S. Nelson, S. P. Platania, B. A. Porter, C. B. Renaud, J. J. Schmitter-Soto, E. B. Taylor, and M. L. Warren Jr. 2008. Conservation status of imperiled North American freshwater and diadromous fishes. *Fisheries* 33:372–386.
- Landis, J. R., and G. G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33:159–174.
- Leftwich, K. N., P. L. Angermeier, and C. A. Dolloff. 1997. Factors influencing behavior and transferability of habitat models for a benthic stream fish. *Transactions of the American Fisheries Society* 126:725–734.
- Leonard, P. M., and D. J. Orth. 1988. Use of habitat guilds of fishes to determine instream flow requirements. *North American Journal of Fisheries Management* 8:399–409.
- Lewis, C. A., N. P. Lester, A. D. Bradshaw, J. E. Fitzgibbon, K. Fuller, L. Hakanson, and C. Richards. 1996. Considerations of scale in habitat conservation and restoration. *Canadian Journal of Fisheries and Aquatic Sciences* 53:440–445.
- Litvan, M. E., C. L. Pierce, T. W. Stewart, and C. J. Larson. 2008. Fish passage in a western Iowa stream modified by grade control structures. *North American Journal of Fisheries Management* 28:1384–1397.
- Loan-Wilsey, A. K., C. L. Pierce, K. L. Kane, P. D. Brown and R. L. McNeely. 2005. The Iowa aquatic gap analysis project, final report. Iowa Cooperative Fish and Wildlife Research Unit, Iowa State University, Ames.
- Lobb, M. D. III, and D. J. Orth. 1991. Habitat use by an assemblage of fish in a large warmwater stream. *Transactions of the American Fisheries Society* 120:65–78.
- Lyons, J., J. S. Stewart, and M. Mitro. 2010. Predicted effects of climate warming on the distribution of 50 stream fishes in Wisconsin, U.S.A. *Journal of Fish Biology* 77:1867–1898.
- Manel, S., H. Ceri Williams, and S. J. Omerod. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* 38:921–931.

- Marsh-Matthews, E., and W. J. Matthews. 2000. Geographic, terrestrial and aquatic factors: which most influence the structure of stream fish assemblages in the midwestern United States? *Ecology of Freshwater Fish* 9:9–21.
- Martin-Bergmann, K. A., and J. H. Gee. 1985. The central mudminnow, *Umbra limi* (Kirtland), a habitat specialist and resource generalist. *Canadian Journal of Zoology* 63:1753–1764.
- McPherson, J. M., W. Jetz, and D. J. Rogers. 2004. The effects of species' range size on the accuracy of distribution models: ecology phenomenon or statistical artifact? *Journal of Applied Ecology* 41:811–823.
- Meyers, E. M., B. Dobrowski, and C. L. Tague. 2010. Climate change impacts on flood frequency, intensity, and timing may affect trout species in Sagehen Creek, California. *Transactions of the American Fisheries Society* 139:1657–1664.
- Miller, R. R., J. D. Williams, and J. E. Williams. 1989. Extinctions of North American fishes during the past century. *Fisheries* 14(6):22–38.
- Monti, D., and P. Legendre. 2009. Shifts between biotic and physical driving forces of species organization under natural disturbance regimes. *Canadian Journal of Fisheries and Aquatic Sciences* 66:1282–1293.
- Oakes, R. M., K. B. Gido, J. A. Falke, J. D. Olden, and B. L. Brock. 2005. Modelling of stream fishes in the Great Plains, USA. *Ecology of Freshwater Fish* 14:361–374.
- Olden, J. D., and D. A. Jackson. 2001. Fish-habitat relationships in lakes: gaining predictive and explanatory insight by using artificial neural networks. *Transactions of the American Fisheries Society* 130:878–897.
- Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2002. Predictive models of fish species distribution: a note on proper model validation and chance predictions. *Transactions of the American Fisheries Society* 131:329–336.
- Paller, M. H. 1994. Relationships between fish assemblage structure and stream order in South Carolina coastal plains streams. *Transactions of the American Fisheries Society* 123:150–161.
- Pearson, T. N., H. W. Li, and G. A. Lamberti. 1992. Influence of habitat complexity on resistance to flooding and resilience of stream fish assemblages. *Transactions of the American Fisheries Society* 121:427–436.
- Peck, D. V., A. T. Herlihy, B. H. Hill, R. M. Hughes, P. R. Kaufmann, D. J. Klemm, J. M. Lazorchak, F. M. McCormick, S. A. Peterson, P. L. Ringold, T. Magee, and M. R. Cappaert. 2006. Environmental monitoring and assessment program-surface waters western pilot study: field operations manual for Wadeable streams. U.S. Environmental Protection Agency, EPA 600/R-06/003, Washington, D.C.
- Pflieger, W. L. 1997. *The Fishes of Missouri*, revised edition. Missouri Department of Conservation, Jefferson City.
- Poff, N. L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. *Journal of North American Benthological Society* 16:391–409.
- Pont, D., B. Hugueny, and T. Oberdorff. 2005. Modelling habitat requirement of European fishes: do species have similar responses to small and regional environmental constraints? *Canadian Journal of Fisheries and Aquatic Sciences* 62:163–173.
- Porter, M. S., J. Rosenfeld, and E. A. Parkinson. 2000. Predictive models of fish species distribution in the Blackwater Drainage, British Columbia. *North American Journal of Fisheries Management* 20:349–359.
- Power, M. E., W. J. Matthews, and A. J. Stewart. 1985. Grazing minnows, piscivorous bass, and stream algae—dynamics of a strong interaction. *Ecology* 66:1448–1456.
- Pretty, J. L., S. S. C. Harrison, D. J. Shepherd, C. Smith, A. G. Hildrew, and R. D. Hey. 2003. River rehabilitation and fish populations: assessing the benefit of instream structures. *Journal of Applied Ecology* 40:251–265.
- Quist, M. C., W. A. Hubert, and F. J. Rahel. 2004. Elevation and stream-size thresholds affect distributions of native and exotic warmwater fishes in Wyoming. *Journal of Freshwater Ecology* 19:227–235.
- Quist, M. C., F. J. Rahel, and W. A. Hubert. 2005. Hierarchical faunal filters: an approach to assessing effects of habitat and nonnative species on native fishes. *Ecology of Freshwater Fish* 14:24–39.
- R Development Core Team. 2009. R version 2.10.0. Available: r-project.org. (August 2010).
- Rabeni, C. F., and S. P. Sowa. 1996. Integrating biological realism into habitat restoration and conservation strategies for small streams. *Canadian Journal of Fisheries and Aquatic Sciences* 53:252–259.
- Rahel, F. J., and W. A. Hubert. 1991. Fish assemblages and habitat gradients in a Rocky Mountain–Great Plains stream: biotic zonation and additive patterns of community change. *Transactions of the American Fisheries Society* 120:319–332.
- Rashleigh, B., R. Parmar, J. M. Johnston, and M. C. Barber. 2005. Predictive habitat models for the occurrence of stream fishes in the Mid-Atlantic highlands. *North American Journal of Fisheries Management* 25:1353–1366.
- Rich, C. F., T. E. McMahon, B. E. Rieman, and W. L. Thompson. 2003. Small-habitat, watershed, and biotic features associated with bull trout occurrence in Montana streams. *Transactions of the American Fisheries Society* 132:1053–1064.
- Richards, C., L. B. Johnson, and G. E. Host. 1996. Landscape-scale influences on stream habitats and biota. *Canadian Journal of Fisheries and Aquatic Sciences* 53:295–311.
- Richards, S. A. 2005. Testing ecological theory using the information-theoretic approach: examples and cautionary results. *Ecology* 86:2805–2814.
- Richter, B. D., D. P. Braun, M. A. Mendelson, and L. L. Master. 1997. Threats to imperiled freshwater fauna. *Conservation Biology* 11:1081–1093.
- Robinson, J. B., D. C. Josephson, B. C. Weidel, and C. E. Kraft. 2010. Influence of variable interannual summer water temperatures on brook trout growth, consumption, reproduction, and mortality in an unstratified Adirondack lake. *Transactions of the American Fisheries Society* 139:685–699.
- Rowe, D. C., C. L. Pierce, and T. F. Wilton. 2009a. Physical habitat and fish assemblage relationships with landscape variables at multiple spatial scales in Wadeable Iowa streams. *North American Journal of Fisheries Management* 29:1333–1351.
- Rowe, D. C., C. L. Pierce, and T. F. Wilton. 2009b. Fish assemblage relationships with physical habitat in Wadeable Iowa streams. *North American Journal of Fisheries Management* 29:1314–1332.
- Ruiz, J. C., and J. T. Peterson. 2007. An evaluation of the relative influence of spatial, statistical, and biological factors on the accuracy of stream fish species presence models. *Transactions of the American Fisheries Society* 136:1640–1653.
- Schlosser, I. J. 1982. Fish community structure and function along two habitat gradients in a headwater stream. *Ecological Monographs* 52:395–414.
- Simonson, T. D., and J. Lyons. 1995. Comparison of catch per unit effort and removal procedures for sampling stream fish assemblages. *North American Journal of Fisheries Management* 15:419–427.
- Sindt, A. R., C. L. Pierce, and M. C. Quist. 2012. Fish species of greatest conservation need in Wadeable Iowa streams: current status and effectiveness of aquatic Gap Program distribution models. *North American Journal of Fisheries Management* 32:135–146.
- Sowa, S. P., G. M. Annis, D. D. Diamond, D. Figg, M. E. Morey, and T. Nigh. 2004. An overview of the data developed for the Missouri Aquatic GAP Project and an example of how it is being used for conservation planning. *GAP Analysis Bulletin* 12:7–19.
- Steen, P. J., T. G. Zorn, P. W. Seelbach, and J. S. Schaeffer. 2008. Classification tree models for predicting distributions of Michigan stream fish from landscape variables. *Transactions of the American Fisheries Society* 137:976–996.
- Strahler, A. N. 1957. Quantitative analysis of watershed geomorphology. *Transactions, American Geophysical Union* 38:913–920.
- Titus, K., J. A. Mosher, and B. K. Williams. 1984. Chance-corrected classification for use in discriminant analysis: ecological applications. *American Midland Naturalist* 111:1–7.
- Tonn, W. M. 1990. Climate change and fish communities: a conceptual framework. *Transactions of the American Fisheries Society* 119:337–352.
- U.S. Department of Agriculture. 2007. Summary report: 2007 national resources inventory. Natural Resource Conservation Service, Washington, D.C., and Statistical Laboratory, Iowa State University, Ames.

- Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences* 37:130–137.
- Vaughan, I. P., and S. J. Omerod. 2005. The continuing challenges of testing species distribution models. *Journal of Applied Ecology* 42:720–730.
- Wall, S. S., C. R. Berry Jr., C. M. Blausey, J. A. Jenks, and C. J. Kopplin. 2004. Fish-habitat modeling for gap analysis to conserve the endangered Topeka shiner (*Notropis topeka*). *Canadian Journal of Fisheries and Aquatic Sciences* 61:954–973.
- Wang, L., J. Lyons, P. Rasmussen, P. Seelbach, T. Simon, M. Wiley, P. Kanehl, E. Baker, S. Niemela, and P. M. Stewart. 2003. Watershed, reach, and riparian influence on stream fish assemblages in the northern lakes and forest ecoregion, U.S.A. *Canadian Journal of Fisheries and Aquatic Sciences* 60:491–505.
- Wilton, T. F. 2004. Biological assessment of Iowa's wadeable streams. Iowa Department of Natural Resources, Des Moines.
- Zimmer, D. W., and R. W. Bachmann. 1978. Channelization and invertebrate drift in some Iowa streams. *Journal of the American Water Resources Association* 14:868–883.
- Zohrer, J. J. 2005. Securing a future for fish and wildlife: a conservation legacy for Iowans. Iowa wildlife action plan report. Iowa Department of Natural Resources, Des Moines.