Fish–Habitat Relations across Spatial Scales in Prairie Streams

Keith B. Gido*, Jeffrey A. Falke
Division of Biology, Kansas State University
Manhattan, Kansas 66506, USA

Robert M. Oakes
University of Pennsylvania Law School
Philadelphia, Pennsylvania 19104, USA

Kristen J. Hase
Tallgrass Prairie National Preserve
Cottonwood Falls, Kansas 66845, USA

Abstract.—Habitat data collected at three spatial scales (catchments, reaches, and sites) were used to predict individual fish species occurrences and assemblage structure at 150 sites in the Kansas River basin. Habitat measurements for the catchments and reaches of each sample site were derived from available geographic information system (GIS) data layers. Habitat measurements at the sample sites were collected at the time of fish sampling. Because habitat measurements are typically more difficult to collect as the spatial scale of sampling decreases (i.e., field measurement versus a GIS analysis), our objective was to quantify the relative increase in predictive ability as we added habitat measurements from increasingly finer spatial scales. Although the addition of site-scale habitat variables increased the predictive performance of models, the relative magnitude of these increases was small. This was largely due to the general association of species occurrences with measurements of catchment area and soil factors, both of which could be quantified with a GIS. Habitat measurements taken at different spatial scales were often correlated; however, a partial canonical correspondence analysis showed that catchment-scale habitat measurements accounted for a slightly higher percent of the variation in fish-assemblage structure across the 150 sample sites than reach- or site-scale habitat measurements. We concluded that field habitat measurements were less informative for predicting species occurrences within the Kansas River basin than catchment data. However, because of the hierarchical nature of the geomorphological processes that form stream habitats, a refined understanding of the relationship between catchment-, reach- and site-scale habitats provides a mechanistic understanding of fish–habitat relations across spatial scales.

INTRODUCTION

Understanding species–environment relations is an important step toward the conservation of aquatic communities. This is particularly important in regions such as the Great Plains, which have a highly endangered fauna as a result of extensive human alterations to aquatic systems (Cross et al. 1985; Cross and Moss 1987; Fausch and Bestgen 1997; Dodds et al. 2004). However, identifying the appropriate scale at which to measure these relations is complicated. Both natural and human factors work across multiple spatial scales to constrain species distributions. At broad spatial scales, natural factors such as basin geology and stream network configurations influence...
fish assemblage structure (e.g., Benda et al. 2004) and food web interactions (Power and Dietrich 2002). At finer spatial scales, fish species partition habitats based on stream substrate size, depth, and current velocity (e.g., Gorman 1988; Aadland 1993; Gido and Propst 1999). Human impacts at broad spatial scales, such as groundwater mining or impoundment, have reduced and fragmented Great Plains river systems. Although these factors have left many local habitats intact, some species have clearly been affected by these disturbances (Cross and Moss 1987; Fausch and Bestgen 1997). In addition, local-scale alterations to habitat (e.g., gravel mining or bridge construction) can also affect local fish-assemblage structure (Burns 1972; Kondolf 1997).

Understanding of aquatic species distributions and ecosystem processes vary widely with the spatial scale of analysis (Allan et al. 1997). Several studies have previously reported species–habitat relationships across spatial scales, but these findings suggest reach and site habitats are primarily important in studies with a limited spatial extent (i.e., 100–10,000 km²). For example, Bond and Lake (2003; spatial extent 225 km²) evaluated habitat associations across three spatial scales (among streams, among sites, and within sites). They found that three of four species showed significant habitat associations at local spatial scales and that differences in habitat availability among sites and streams were less influential. Walters et al. (2003; spatial extent 300–400 km²) found that species composition in Piedmont streams was best predicted by reach-level geomorphic variables, including stream slope, bed texture, bed mobility, and tractive forces. However, species richness and density were best predicted by stream size. Joy and Death (2003; spatial extent 5,000 km²) successfully predicted the occurrences of 12 fish and invertebrate species in New Zealand streams based on reach-scale habitat features and two spatial variables (elevation and distance from the coast). Rich et al. (2003; spatial extent 7,900 km²), found that predictive models with both local-habitat and biotic variables (proximity to source populations) best explained the occurrence of bull trout Salvelinus fontinalis in 112 Montana streams.

Across broader spatial scales, quantifying catchment habitats may become more important in predicting local fish-assemblage structure. Because of the hierarchical structuring of streams, catchment landscapes can be closely linked to a stream’s hydrologic regime, chemistry, and physical structure (Ritter et al. 1995). In addition, interactions between the catchment landscapes and stream organisms can change predictably with longitudinal position (Power and Dietrich 2002). Thus, because fish assemblages are tightly linked to hydrologic variability (e.g., Poff and Allan 1995), catchment area and geology are predicted to be closely associated with fish-assemblage structure. Santoul et al. (2004; spatial extent 57,000 km²) provide an example of this association, as they were able to explain 87.2% of the variation in fish species composition across 329 sites in southwestern France with the variables altitude, slope, and catchment area. The studies listed above indicate that stream fish assemblages are influenced by factors that occur at multiple spatial scales, but catchment or stream segment habitat measurements may be the best predictors of assemblage structure across broad spatial scales.

Broad-scale patterns of land use have been linked to aquatic community structure, but these effects also can be scale dependent. For example, Rashleigh (2004) found that land use associated with agriculture and urban development was associated with fish-assemble structure, primarily by shifting trophic composition. Wang et al. (2001) found that the amount of connected impervious surface in a catchment influenced stream fish assemblages, instream habitat, and base flow in 47 small Wisconsin streams. Roth et al. (1996) found that land use quantified at the catchment scale was more reflective of stream biotic integrity (index of biotic integrity) than local scale habitat measurements. These results also are consistent with studies of macroinvertebrate assemblage structure, which has been associated with differences in land use (e.g.,
native versus nonnative land cover; Townsend et al. 2004).

Correlations between fish-assemblage structure and broad-scale habitat measurements such as land use and geology are due to the response of individuals to local conditions that are partly influenced by broader-scale factors. That is, major driving factors that structure assemblages may be correlated with measurements taken at different spatial scales due to the hierarchical organization of lotic systems (Frissell et al. 1986). For example, Richards et al. (1996) found a strong association between land-use and geologic variables, both of which were associated with macroinvertebrate assemblage structure. Despite the potential influence of broad-scale factors, these coarse landscape filters typically cannot explain all levels of variation in species distributions because they only represent an average of variability in finer scale habitat features that are relevant to the biota (Poff 1997). Thus, we would predict that fine-scale characterization of habitat would increase our ability to predict assemblage structure at fine to intermediate spatial scales because they more adequately reflect macro- and microhabitats used by fishes.

Given that fish assemblages are structured by factors that operate across multiple spatial scales, an understanding of these species-habitat relationships should provide information to help manage and predict the consequences of habitat changes that occur at different spatial scales. In addition, identifying the scales at which species respond to habitat will allow managers to estimate the scale at which to apply conservation efforts. For example, the scale of conservation is likely dependent on dispersal ability of organisms, and reach- or segment-level conservation may only be appropriate for those organisms with poor dispersal ability (Fausch et al. 2002; Wishart and Davies 2003). Species that require moderate to high levels of dispersal may require larger, regional conservation efforts.

Our objective was to evaluate the association between local fish-assemblage structure and habitat characteristics of 150 sample sites quantified at three spatial scales. Habitat measurements from field sites (~0.1 km) were taken in conjunction with fish assemblage sampling. Reach (~1.0 km) and catchment (>5 km²) habitat characteristics were measured using a GIS. This partitioning of our habitat measurements allowed us to compare the relative predictive ability of field versus GIS methods of quantifying habitat. In addition, our approach was to first explain variability of assemblage structure with catchment variables and then quantify how much additional variation could be explained by adding habitat variables from finer spatial scales. This nested sequence of analyses was used because of the hierarchical structuring of stream habitats and the relative effort required in gathering habitat information as opposed to quantifying landscape-level factors. In addition, important processes in lotic systems typically occur in a downstream direction (i.e., from catchment to sites). Based on the broad spatial extent of our study (67,000 km²), we predicted that catchment-scale habitats that can be quantified by GIS are the best predictors of fish-assemblage structure, and little additional variability will be explained by adding habitat information taken at finer spatial scales (i.e., field measurements).

**METHODS AND MATERIALS**

**Study Area**

Fishes were collected at 150 sites in the lower Kansas River basin (Figure 1). This area falls within four EPA level III ecoregions: Central Irregular Plains (13.0%), Flint Hills (17.0%), Western Corn Belt Plains (19.5%), and Central Great Plains (50.5%; Omernik 1987). Primary land uses within the lower Kansas River basin were row crop agriculture (53%), seminatural herbaceous vegetation (20%), and pasture/hay (13%). Urban and wetland land uses combined composed ~3% of land use in the basin (USGS 1992). Stream order at the collection sites ranged from first to fifth order (Strahler 1957) and catchment...
area of study reaches ranged from 5 to 33,775 km². Three major impoundments (4,734–6,257-ha surface area) and numerous small impoundments are spread throughout this region.

Fish Data

Our analysis was based on 150 fish collections taken by the Kansas Department of Wildlife and Parks (KDWP) between September 1995 and August 2000. Fifty-three fish species were captured, 41 of which occurred at six or more sites. The frequency of occurrence across sites for these species ranged from less than 1.0% to 96% (mean 36%; Table 1). Included in these collections were four species that have been identified as being of special conservation concern in Kansas: Topeka shiner Notropis topeka, southern redbelly dace Phoxinus erythrogaster, common shiner Luxilus cornutus, and Johnny darter Etheostoma nigrum (Haslouer et al. 2005). The KDWP sampling followed the protocol of Lazorchak et al. (1998); at each site a reach that was 40 times the average wetted width of the stream (minimum 150 m, maximum 300 m) was sampled using a combination of straight and bag seines (4.7-mm mesh) and DC-pulsed backpack electrofishing.

Environmental Variables

Predictive modeling was conducted using habitat measurements taken at three spatial scales (sites, reaches, and catchments; Table 2). Habitat variables included in our analysis were those that were available for the entire region and that we considered to potentially influence stream fish assemblages, based on a review of pertinent literature and personal observations. Habitat measured at field sites included the length of stream from which fish were sampled (i.e., between...
150 m and 300 m). Channel width, bank angle, and canopy cover (based on densiometer reading) were measured at 11 equally spaced transects along the site. Depth and substrate size were measured at 5 points along each of these transects. All values were averaged for each site. In addition, we measured discharge, specific conductance, turbidity, total dissolved solids, nitrate, ammonia, and phosphorus.

A stream reach was defined as a stream course from its upstream confluence to its downstream confluence with other tributaries (mean length = 1.6 km). Landscape-scale habitat variables were quantified for each reach and were chosen to

Table 1. Common and scientific names of fishes, species codes, and number of sample sites occupied by each species in the Kansas River basin.

<table>
<thead>
<tr>
<th>Species</th>
<th>Scientific name</th>
<th>Species code</th>
<th>Number of sites occupied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black bullhead</td>
<td>Ameiurus melas</td>
<td>AMEMEL</td>
<td>54</td>
</tr>
<tr>
<td>Yellow bullhead</td>
<td>A. natalis</td>
<td>AMENAT</td>
<td>78</td>
</tr>
<tr>
<td>Freshwater drum</td>
<td>Aplodinotus grunniens</td>
<td>APLGRU</td>
<td>17</td>
</tr>
<tr>
<td>Central stoneroller</td>
<td>Catostomus commersonii</td>
<td>CATCOM</td>
<td>61</td>
</tr>
<tr>
<td>Red shiner</td>
<td>Cyprinella lutrensis</td>
<td>CYPLUT</td>
<td>144</td>
</tr>
<tr>
<td>Common carp</td>
<td>Cyprinus carpio</td>
<td>CPCAR</td>
<td>54</td>
</tr>
<tr>
<td>Gizzard shad</td>
<td>Dorosoma cepedianum</td>
<td>DORCEP</td>
<td>22</td>
</tr>
<tr>
<td>Johnny darter</td>
<td>Etheostoma nigrum</td>
<td>ETHNIG</td>
<td>32</td>
</tr>
<tr>
<td>Orangethroat darter</td>
<td>E. spectabile</td>
<td>ETHSPE</td>
<td>91</td>
</tr>
<tr>
<td>Western mosquitofish</td>
<td>Gambusia affinis</td>
<td>GAMAFF</td>
<td>10</td>
</tr>
<tr>
<td>Channel catfish</td>
<td>Ictalurus punctatus</td>
<td>ICTPUN</td>
<td>88</td>
</tr>
<tr>
<td>Smallmouth buffalo</td>
<td>Ictiobus bubalus</td>
<td>ICTBUB</td>
<td>8</td>
</tr>
<tr>
<td>Longnose gar</td>
<td>Leptosostus osseus</td>
<td>LEPOSS</td>
<td>21</td>
</tr>
<tr>
<td>Green sunfish</td>
<td>Lepomis cyanellus</td>
<td>LEPCYA</td>
<td>138</td>
</tr>
<tr>
<td>Orangespotted sunfish</td>
<td>L. humilis</td>
<td>LEPHUM</td>
<td>48</td>
</tr>
<tr>
<td>Bluegill</td>
<td>L. macrochirus</td>
<td>LEMAC</td>
<td>67</td>
</tr>
<tr>
<td>Longear sunfish</td>
<td>L. megalotis</td>
<td>LEMPMEG</td>
<td>34</td>
</tr>
<tr>
<td>Common shiner</td>
<td>Luxilus cornutus</td>
<td>LUXCOR</td>
<td>50</td>
</tr>
<tr>
<td>Redfin shiner</td>
<td>Lythrurus unbrathilis</td>
<td>LYTUMB</td>
<td>38</td>
</tr>
<tr>
<td>Largemouth bass</td>
<td>Micropterus salmoides</td>
<td>MICSAL</td>
<td>86</td>
</tr>
<tr>
<td>White bass</td>
<td>Morone chrysops</td>
<td>MORCHR</td>
<td>6</td>
</tr>
<tr>
<td>Golden redhorse</td>
<td>Moxostoma erythrurum</td>
<td>MOXERY</td>
<td>12</td>
</tr>
<tr>
<td>Shortnose redhorse</td>
<td>M. macrolepidotum</td>
<td>MOXMAC</td>
<td>28</td>
</tr>
<tr>
<td>Golden shiner</td>
<td>Notemigonus crysoleucas</td>
<td>NOTCARY</td>
<td>19</td>
</tr>
<tr>
<td>Bigmouth shiner</td>
<td>Notropis dorsalis</td>
<td>NOTDOR</td>
<td>15</td>
</tr>
<tr>
<td>Rosyface shiner</td>
<td>N. rubellus</td>
<td>NOTRUB</td>
<td>16</td>
</tr>
<tr>
<td>Sand shiner</td>
<td>N. stramineus</td>
<td>NOTSTR</td>
<td>110</td>
</tr>
<tr>
<td>Topeka shiner</td>
<td>N. topeka</td>
<td>NOTTOP</td>
<td>8</td>
</tr>
<tr>
<td>Slender madtom</td>
<td>Noturus exilis</td>
<td>NOTEXI</td>
<td>46</td>
</tr>
<tr>
<td>Stonecat</td>
<td>N. flavus</td>
<td>NOTFLA</td>
<td>57</td>
</tr>
<tr>
<td>Logperch</td>
<td>Percina caprodes</td>
<td>PERCAP</td>
<td>27</td>
</tr>
<tr>
<td>Suckermouth minnow</td>
<td>Phenacobius mirabilis</td>
<td>PHEMIR</td>
<td>111</td>
</tr>
<tr>
<td>Southern redbelly dace</td>
<td>Phoxinus erythrogaster</td>
<td>PHOERY</td>
<td>14</td>
</tr>
<tr>
<td>Bluntnose minnow</td>
<td>Pimephales notatus</td>
<td>PIMNOS</td>
<td>120</td>
</tr>
<tr>
<td>Fathead minnow</td>
<td>P. promelas</td>
<td>PIMPRD</td>
<td>105</td>
</tr>
<tr>
<td>Bullhead minnow</td>
<td>P. vigilax</td>
<td>PIMVIG</td>
<td>13</td>
</tr>
<tr>
<td>White crappie</td>
<td>Pomoxis annularis</td>
<td>POMANN</td>
<td>27</td>
</tr>
<tr>
<td>Flathead catfish</td>
<td>Pylodictis olivaris</td>
<td>PYLOLI</td>
<td>30</td>
</tr>
<tr>
<td>Creek chub</td>
<td>Semotilus atramoculatus</td>
<td>SEMATR</td>
<td>119</td>
</tr>
</tbody>
</table>
Table 2. Reference codes and descriptions of habitat variables measured at three different spatial scales (catchment, reach, and site) in the Kansas River basin. Variable loadings and percent variance explained is given for the first four axes of principal component analyses to summarize variation in measurements across sites.

<table>
<thead>
<tr>
<th>Reference code</th>
<th>Variable description</th>
<th>PCA1</th>
<th>PCA2</th>
<th>PCA3</th>
<th>PCA4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Catchment</strong></td>
<td></td>
<td>28.4%</td>
<td>22.5%</td>
<td>12.6%</td>
<td>9.1%</td>
</tr>
<tr>
<td>CmtAREA</td>
<td>Catchment area (km²)</td>
<td>0.026</td>
<td>-0.127</td>
<td>0.351</td>
<td>0.367</td>
</tr>
<tr>
<td>Density</td>
<td>Stream density (km/km²)</td>
<td>0.072</td>
<td>-0.321</td>
<td>-0.187</td>
<td>-0.004</td>
</tr>
<tr>
<td>Slope</td>
<td>Field slope (%)</td>
<td>-0.375</td>
<td>0.138</td>
<td>0.008</td>
<td>0.127</td>
</tr>
<tr>
<td>WTDEP</td>
<td>Water table depth (m)</td>
<td>-0.339</td>
<td>-0.215</td>
<td>0.277</td>
<td>0.161</td>
</tr>
<tr>
<td>KFACT</td>
<td>Soil erodibility factor (tons/unit of rainfall erosion index)</td>
<td>0.363</td>
<td>-0.156</td>
<td>-0.062</td>
<td>-0.121</td>
</tr>
<tr>
<td>PERM</td>
<td>Soil permeability (cm/h)</td>
<td>-0.032</td>
<td>-0.209</td>
<td>0.584</td>
<td>0.162</td>
</tr>
<tr>
<td>BD</td>
<td>Bulk density of soils (g/cm³)</td>
<td>0.072</td>
<td>0.388</td>
<td>-0.299</td>
<td>0.260</td>
</tr>
<tr>
<td>OM</td>
<td>Organic matter content of soils (% by weight)</td>
<td>-0.291</td>
<td>0.252</td>
<td>-0.173</td>
<td>0.284</td>
</tr>
<tr>
<td>TFACT</td>
<td>Soil loss tolerance factor (tons/acre/year)</td>
<td>0.370</td>
<td>-0.056</td>
<td>0.155</td>
<td>0.256</td>
</tr>
<tr>
<td>WEG</td>
<td>Wind erosion group</td>
<td>-0.164</td>
<td>-0.060</td>
<td>0.056</td>
<td>-0.654</td>
</tr>
<tr>
<td>Urban</td>
<td>Urban land</td>
<td>0.132</td>
<td>0.132</td>
<td>0.275</td>
<td>-0.210</td>
</tr>
<tr>
<td>Forest_u</td>
<td>Forested land</td>
<td>0.104</td>
<td>0.389</td>
<td>0.298</td>
<td>-0.124</td>
</tr>
<tr>
<td>Grasslan</td>
<td>Grassland</td>
<td>-0.435</td>
<td>-0.085</td>
<td>-0.070</td>
<td>0.107</td>
</tr>
<tr>
<td>Pasture</td>
<td>Pasture</td>
<td>0.198</td>
<td>0.358</td>
<td>0.050</td>
<td>0.128</td>
</tr>
<tr>
<td>Sm_grain</td>
<td>Small grain</td>
<td>0.302</td>
<td>-0.271</td>
<td>-0.198</td>
<td>0.228</td>
</tr>
<tr>
<td>Wetlands</td>
<td>Wetlands</td>
<td>0.069</td>
<td>0.386</td>
<td>0.251</td>
<td>-0.050</td>
</tr>
<tr>
<td><strong>Reach</strong></td>
<td></td>
<td>21.1%</td>
<td>14.3%</td>
<td>11.0%</td>
<td>9.9%</td>
</tr>
<tr>
<td>Res_dist</td>
<td>Downstream distance to the nearest reservoir (km)</td>
<td>0.052</td>
<td>0.244</td>
<td>-0.260</td>
<td>0.341</td>
</tr>
<tr>
<td>STRAHLER</td>
<td>Strahler order of stream segment</td>
<td>0.401</td>
<td>0.129</td>
<td>0.168</td>
<td>0.007</td>
</tr>
<tr>
<td>DOWNORD</td>
<td>Strahler order of downstream segment</td>
<td>0.384</td>
<td>0.224</td>
<td>-0.030</td>
<td>-0.212</td>
</tr>
<tr>
<td>MAX_ELEV</td>
<td>Maximum elevation of the stream segment (m)</td>
<td>-0.132</td>
<td>0.419</td>
<td>-0.049</td>
<td>0.361</td>
</tr>
<tr>
<td>Clay_loa</td>
<td>NRCS soil texture class clay loam in the stream segment</td>
<td>-0.241</td>
<td>0.218</td>
<td>-0.318</td>
<td>-0.234</td>
</tr>
<tr>
<td>Silty_cl</td>
<td>NRCS soil texture class silty clay in the stream segment</td>
<td>0.162</td>
<td>-0.483</td>
<td>-0.222</td>
<td>0.088</td>
</tr>
<tr>
<td>Siltycl</td>
<td>NRCS soil texture class silty clay loam in the stream segment</td>
<td>-0.145</td>
<td>0.113</td>
<td>0.502</td>
<td>0.199</td>
</tr>
<tr>
<td>Silt_loa</td>
<td>NRCS soil texture class silty loam in the stream segment</td>
<td>0.232</td>
<td>0.255</td>
<td>-0.060</td>
<td>0.129</td>
</tr>
<tr>
<td>Sandy_lo</td>
<td>NRCS soil texture class sandy loam in the stream segment</td>
<td>0.072</td>
<td>-0.012</td>
<td>-0.015</td>
<td>-0.402</td>
</tr>
<tr>
<td>Loom</td>
<td>NRCS soil texture class loam in the stream segment</td>
<td>-0.041</td>
<td>-0.061</td>
<td>0.108</td>
<td>-0.308</td>
</tr>
<tr>
<td>Loamy_up</td>
<td>NRCS range site information, loamy upland</td>
<td>-0.113</td>
<td>0.124</td>
<td>0.274</td>
<td>-0.329</td>
</tr>
<tr>
<td>Loamy_lo</td>
<td>NRCS range site information, loamy lowland</td>
<td>0.308</td>
<td>-0.302</td>
<td>-0.205</td>
<td>0.300</td>
</tr>
<tr>
<td>Clay_up</td>
<td>NRCS range site information, clay upland</td>
<td>-0.243</td>
<td>0.254</td>
<td>-0.331</td>
<td>-0.016</td>
</tr>
<tr>
<td>Limy_up</td>
<td>NRCS range site information, limy upland</td>
<td>-0.099</td>
<td>0.001</td>
<td>0.381</td>
<td>0.158</td>
</tr>
<tr>
<td>Clay_lo</td>
<td>NRCS range site information, clay lowland</td>
<td>0.036</td>
<td>0.061</td>
<td>0.105</td>
<td>-0.025</td>
</tr>
<tr>
<td>Cenozoic</td>
<td>Presence of Cenozoic soils in the segment</td>
<td>0.150</td>
<td>0.339</td>
<td>-0.045</td>
<td>0.035</td>
</tr>
<tr>
<td>Sinuosity</td>
<td>Sinuosity (km/km)</td>
<td>0.047</td>
<td>-0.066</td>
<td>0.271</td>
<td>0.172</td>
</tr>
<tr>
<td><strong>Site</strong></td>
<td></td>
<td>20.1%</td>
<td>18.2%</td>
<td>12.7%</td>
<td>9.7%</td>
</tr>
<tr>
<td>Mean_sub</td>
<td>Mean substrate size</td>
<td>-0.342</td>
<td>0.161</td>
<td>0.021</td>
<td>0.339</td>
</tr>
<tr>
<td>Bnk_angl</td>
<td>Bank angle (degrees)</td>
<td>-0.011</td>
<td>0.306</td>
<td>0.130</td>
<td>0.312</td>
</tr>
<tr>
<td>Canopy</td>
<td>Percent canopy cover (%)</td>
<td>-0.178</td>
<td>0.215</td>
<td>-0.063</td>
<td>0.591</td>
</tr>
<tr>
<td>Mean_dep</td>
<td>Mean depth (m)</td>
<td>-0.074</td>
<td>0.520</td>
<td>0.084</td>
<td>-0.047</td>
</tr>
<tr>
<td>Mean_wid</td>
<td>Mean width (m)</td>
<td>-0.041</td>
<td>0.534</td>
<td>-0.068</td>
<td>-0.272</td>
</tr>
<tr>
<td>Discharge</td>
<td>Discharge (L/s)</td>
<td>0.160</td>
<td>0.470</td>
<td>0.084</td>
<td>-0.306</td>
</tr>
<tr>
<td>Cond</td>
<td>Conductivity (siemens/cm)</td>
<td>-0.146</td>
<td>0.043</td>
<td>0.636</td>
<td>-0.064</td>
</tr>
<tr>
<td>Turb</td>
<td>Turbidity (NTU)</td>
<td>0.473</td>
<td>0.144</td>
<td>-0.105</td>
<td>0.157</td>
</tr>
<tr>
<td>Alkil</td>
<td>Alkalinity (mg/L)</td>
<td>-0.194</td>
<td>-0.084</td>
<td>0.550</td>
<td>0.108</td>
</tr>
<tr>
<td>Chlor</td>
<td>Chlorides (mg/L)</td>
<td>0.332</td>
<td>0.060</td>
<td>0.105</td>
<td>-0.117</td>
</tr>
<tr>
<td>Ammonia</td>
<td>Ammonia (mg/L)</td>
<td>0.391</td>
<td>-0.113</td>
<td>0.227</td>
<td>0.102</td>
</tr>
<tr>
<td>Nitrate</td>
<td>Nitrate (mg/L)</td>
<td>0.368</td>
<td>0.106</td>
<td>-0.107</td>
<td>0.444</td>
</tr>
<tr>
<td>Phos</td>
<td>Phosphorous (mg/L)</td>
<td>0.371</td>
<td>-0.030</td>
<td>0.411</td>
<td>0.085</td>
</tr>
</tbody>
</table>


represent factors linked to fish-assemblage structure, including stream size and position in the watershed (Sheldon 1968; Osborne and Wiley 1992; Zorn et al. 2002) and surface geology of the reach (Matthews and Robison 1988; Nelson et al. 1992; Mandrak 1995). Measures of stream size and position were calculated from a modified version of the national hydrography data set (USGS 1997). Surficial geology of the reach was based on soil measurements obtained from the STATSGO database (NRCS 1994). Environmental variables were classified at the stream-reach scale for use in modeling with ArcMap version 8.2 software and methods described in Maidment (2002).

Catchment area was calculated using a GIS. Catchment-scale soil factors, land use, stream density, and basin slope were extracted from available data layers for the region. As above, soil variables were obtained from the STATSGO database (NRCS 1994) and land use was calculated from the national land cover data set (USGS 1992), which contains 21 categories at a resolution of 30 m².

Statistical Analyses

Concordance of habitat measurements across spatial scales.—Prior to analysis, all habitat measurements were evaluated for normality, and appropriate transformations were applied to reduce heterogeneity of variances and the effects of outliers. Associations among variables were evaluated prior to modeling and those exhibiting a high degree of concordance (i.e., Pearson correlation, r > 0.70) were eliminated to reduce multicollinearity in the data set. Of the highly correlated variables, we retained those expected to be most closely associated with assemblage structure based on previous research in this system (Oakes et al. 2005). To evaluate correlations of the remaining habitat measurements across spatial scales we first conducted a principal components analyses (PCA) to summarize variation in habitat measurements across sites for each of the three spatial scales. For these analyses, all variables were centered and standardized to a standard deviation of one and mean of zero. Next, to characterize the association of environmental variables across spatial scales, we used a Procrustean analysis (PROTEST) to test for congruence among PCA scores from the analysis of habitat measurements from different scales. PROTEST is similar to the Mantel test and is equal in power for comparing matrices with no correlation between or within them and more powerful in comparing correlated matrices (Peres-Neto and Jackson 2001). The PROTEST analysis provided a graphical relationship of two matrices and a residual error value between paired observations. The matrices were matched by translating both matrices to a common centroid and then rotating and dilating one configuration to match the other configuration and minimize the sum of the squared deviations between landmarks (the deviations between landmarks were vector residuals). A small vector residual indicates a close agreement between the corresponding landmarks. The measure of fit ($m^2$) was based on the sum of the squared deviations (Gower 1971). To test if the original $m^2$ was smaller than expected due to chance, the observed $m^2$ was compared to 9,999 permutations using PROTEST. To test for correlations in habitat measurements across spatial scales, we used PROTEST to perform pair-wise comparisons of PCA scores for habitat measurements taken at the scale of catchments, reaches, and field sites. All comparisons were based on the first four axes of the PCA.

Predicting individual species occurrences.—We used discriminant function analysis (DFA), which uses linear combinations of predictor variables to maximize the separation between groups (i.e., presence and absence) to test the association between habitat variables and species occurrences. Discriminant function analysis is expected to perform similar to logistic regression when parametric assumptions are met (Efron 1975) and therefore is representative of traditional, parametric approaches in general. Discriminant function analysis has been used in...
ecological studies of fishes (e.g., Joy and Death 2002) but is constrained by the assumptions of linear relationships among variables (e.g., Olden and Jackson 2002). Nevertheless, previous comparisons of several modeling techniques, including DFA, classification trees, and artificial neural networks, revealed only slight differences among these approaches (Olden and Jackson 2002; K.B. Gido, Kansas State University, unpublished data). Thus, we chose DFA because of the availability of a stepwise selection procedure to eliminate variables that were redundant or explained a small amount of variation in species occurrences. For these analyses, variables with partial correlation coefficients with \( P \)-values greater than 0.05 were entered and those with \( P \) less than 0.10 were removed. Because this stepwise procedure requires multiple tests, it is likely to include superfluous variables in the model (Legendre and Legendre 1998). However, because we were primarily interested in forecasting species occurrences, and secondarily interested in variable importance, we did not correct for this error (e.g., with a Bonferroni correction) in order to maximize the predictive ability of our models. Within-group covariance matrices were used and prior probabilities were computed from group sizes (i.e., species prevalence). Model effectiveness was assessed by our ability to correctly classify species presence or absence across the 150 sites. All error rates were expressed as percentages. For all model predictions a decision threshold of 0.5 was used to classify a species as present or absent. All models were generated using SPSS, version 11.0. To evaluate the predictive power of independent variables, we tallied the number of individual species models that included each habitat variable.

Individual species models were evaluated using a jackknife procedure, in which one site was excluded, a model was constructed using \( n - 1 \) sites, and the excluded site was predicted using this model. This procedure was repeated for each of the 150 sites to give a prediction matrix that was compared to the actual occurrences of each species to give error rates across sites.

**Predicting assemblages.**—Fish-assemblage structure was characterized by both indirect and direct gradient analysis. Correspondence analysis (CA) is an indirect gradient analysis that was used to characterize the variation among fish samples without the constraint of measured environmental gradients. Canonical correspondence analysis (CCA) also ordinates samples and species, but under the constraint of the measured habitat features at the different spatial scales. These methods were chosen over linear-based approaches (e.g., principle components analysis and redundancy analysis) because of the large environmental gradients across our study area and predicted monotonic relationships between species occurrences and habitat measurements (Legendre and Legendre 1998). Because these two methods use the same algorithm to ordinate samples, the total inertia (variance) from both approaches can be directly compared to evaluate how well our environmental variables explained variability in the fish assemblage across sample sites. In addition, we were able to partition variation based on habitat measures at the three spatial scales by including different subsets of our data as covariables and evaluating the residual variation explained by the remaining data (Borcard et al. 1992). Thus, we estimated the percent variability explained by habitat variables taken at each of the three spatial scales and the interactions among these variables following the method of Anderson and Gribble (1998). Eigenvalues and site scores for CCA were calculated using CANOCO software, version 4.5 (ter Braak and Smilauer 2002). A forward selection procedure was used to select a subset of variables that significantly contributed to the variation in the species data for each scale of analysis. In sequential steps, each variable that explained the greatest amount of added variation to the model was included in a Monte Carlo permutation test (1,000 iterations) to evaluate if that variable explained a significant proportion of the species data. To correct for the inflated type I error rate associated with this multiple selection procedure, a Bonferroni adjusted alpha was used to select
variables for inclusion in the CCA. Once a variable failed to explain a significant proportion of the variation, no additional variables were added to the model.

RESULTS

Concordance of Habitat Measurements across Spatial Scales

Several highly correlated ($r > 0.70$) variables were removed from the analysis; 16 of 23, 17 of 21, and 13 of 13 variables were retained for catchment-, reach-, and site-scale habitat measurements. Principal component analyses of the remaining habitat measurements at the three spatial scales summarized the majority of variation (between 56% and 73%) across sites on the first four axes (Table 2). Variation in habitat measurements among sites at the catchment scale was primarily attributed to soil variables (soil loss tolerance factor [TFACT], soil erodibility factor [KFACT], and catchment slope), and secondarily to land use and stream density. Variation in habitat measurements across sites at the reach scale was primarily due to variables associated with stream size (stream order) and secondarily by soil factors and elevation. Of the variables measured at the field sites, water quality parameters (turbidity, ammonia, and phosphorus) were negatively associated with mean substrate size, and a secondary axis was weighted by measures of stream size (mean width, depth, and discharge).

Procrustean analysis revealed highly significant concordance in the relative position of sample scores on the first four PCA axes based on habitat measurements taken at the three different scales. The strongest association occurred between ordinations of reach and site habitat variables ($m^2 = 0.803, P < 0.001$) and the weakest between ordination of catchment and reach variables ($m^2 = 0.634, P < 0.001$). The association between ordinations based on catchment and site habitat measurements also was quite high ($m^2 = 0.799, P < 0.001$).

Predicting Individual Species Occurrences

As expected, total classification error across species generally declined as habitat measurements from finer spatial scales were introduced into DFA models (Figure 2). However, these differences were minor, as classification error only decreased, on average, approximately 1% for each added group of habitat variables; 26.6%, 25.1%, and 24.0% for catchment, catchment + reach, and catchment + reach + site models, respectively.

An analysis of variable importance from DFA models showed that variables representing geology and stream size were generally important in predicting the occurrences of fish species in the Kansas River basin. KFACT was selected in 13 of the 41 species models, regardless of whether reach or site variables were included in individual DFA models (Figure 3). Catchment area, however, was notably more important when only catchment variables were included in individual species DFA models (included in 25 of 41 models) than when combined with reach (13 of 41 models) or reach and site (6 of 41 models) variables. This was likely because habitat variables at finer spatial scales (e.g., stream order and mean width) were correlated with catchment area and were selected to replace catchment area in the stepwise selection. No single variable was selected for more than one-third of the species models, when habitat information from all three spatial scales was included. Of the reach habitat variables, stream order (of the downstream reach), sinuosity, and geology (percent sandy loam and Cenozoic) were most frequently included in the species models. Mean width was the most commonly selected site variable but was only selected for 10 of 41 species models.

Predicting Assemblages

Correspondence analysis explained approximately 40% of the variation in species composition on the first four axes. The first axis, which explained 16% of the variation in fish-assemblage structure across sites contrasted sites with
species typical of small, clear-water streams (e.g., southern redbelly dace) with those typical of mid-sized streams to large rivers (Figure 4). The second axis appeared to separate species that occur in mid-sized streams (bullheads *Amieturus* spp.) with those that typically occur in large, sand-bottom rivers (e.g., white bass *Morone chrysops* and longnose gar *Lepisosteus osseus*).

Canonical correspondence analysis, constrained by environmental variables, explained over half the variability in fish-assemblage structure that was explained by the unconstrained ordination (CA). Percent variation explained by the first four axes of the CCA increased as variables from finer spatial scales were included in the analysis. The CCA explained 27.4%, 29.5%, and 31.6% of the variation in fish-assemblage structure based on the inclusion of catchment, catchment + reach, and catchment + reach + site variables, respectively. Thus, less than 3% of the variation in fish-assemblage structure was explained with the addition of habitat measurements from field sites. Fish-assemblage structure along the first axis of the CCA, which only included catchment-scale habitat measurements, was primarily associated with catchment geology (KFACT) and organic content of soils (OM) (Figure 5), whereas the second axis was related to catchment area and catchment geology (TFACT). When habitat measurements from reaches were included in the analysis, several soil factors (Cenozoic and Clay upland) and
Figure 3. Number of species models for which each habitat variable was selected in a stepwise discriminant function analysis to predict its occurrence across the 150 sample locations in the Kansas River basin. Reference codes as in Table 2.
sinuosity were included in the ordination (Figure 6) without the loss of major variables from the ordination of catchment-only variables. When habitat measurements from all spatial scales were added, the importance of water table depth along the first axis declined, and mean width and depth were included with catchment area as an important gradient along the second axis (Figure 7).

A partial CCA that evaluated residual variation attributed to “pure” effects of variables from the three spatial scales illustrated the strong interaction among habitats measured at different scales (Figure 8). When using reach- and site-scale habitat variables as covariates, catchment variables explained 8.8% and reach-scale variables explained 6.4% of the variability in the CCA.
that included habitat variables from all spatial scales. Variables measured at field sites explained 6.2% of the variation in fish-assemblage structure explained by the full CCA. The percent variation attributed to the interaction of all three variables (10.2%) was also quite large because of the high degree of correlation among habitat measurements taken at different spatial scales.

**DISCUSSION**

Most of the variation in fish-assemblage structure and in individual species occurrences could be explained by environmental factors that were quantified using GIS-derived habitat measurements at the scale of catchments and reaches. However, because of the hierarchical nature of...
stream habitats (Frissell et al. 1986), it is hard to separate the relative importance of habitat measured across different spatial scales. Indeed, we found a high degree of correlation in habitat measurements across scales, and it is possible that habitats of catchments simply reflect local habitat conditions. Nevertheless, stream sites that were sampled in our study may not encompass all habitats used by some fishes throughout their different life stages (Schlosser and Angermeier 1995; Fausch et al. 2002), and these populations may respond to habitats defined at broad spatial scales. In particular, fish populations may be dependent on the spatial connectivity of habitats in reaches or stream segments (e.g., Fausch et al. 2002; Rich et al. 2003). For example, if a fish population is highly mobile and uses a variety of habitats throughout its life, the average habitat measurements from a catchment might best reflect the amount of suitable habitat for that population. In contrast, some stream fishes typically move relatively...
little during most life stages (Skalski and Gilliam 2000) and thus should respond to factors operating at finer (e.g., reach and field site) spatial scales. Yet, habitat measurements at fine spatial scales may be confounded by biotic interactions, thus limiting the statistical relationship with species occurrences and habitat features. Regardless of the habitat scale at which fish populations respond, it seems clear that GIS measurements of catchment and reach habitats were able to explain the majority of variation in species occurrences that was explained by variables measured at field sites.

Variables associated with stream size were important predictors of individual species as well as assemblage structure. This was not surprising given the tendency for fish assemblages (Huet 1959; Schlosser 1987; Edds 1993) and ecosystem processes (Vannote et al. 1980) to change longitudinally in river systems. Also, this study included a wide range of stream sizes across sample sites. Because habitat variables...
that were representative of stream size occurred at all three spatial scales (e.g., catchment area, stream order, and stream width), it was hard to evaluate the relative predictive ability of the different measurements. Whereas large rivers have unique reach-scale habitats, such as braided channels, turbid water, fine substrates, and reduced variability in flows, catchment-scale habitats of large rivers also contain complex stream networks and a greater degree of habitat heterogeneity than small catchments. Reach habitats in large rivers should favor generalist, water column, and predatory species (e.g., red shiner and channel catfish) that may require a diversity of habitats throughout their life cycles. Alternatively, smaller tributary streams with flashy flows and reduced turbidity were dominated by benthic-feeding herbivorous and insectivorous fish such as central stoneroller *Campostoma anomalum* and orangethroat darter *Etheostoma spectabile*. An evaluation of how individual species with different life history traits respond to habitat measurements taken at different spatial scales would help partition the relative influence of stream size in determining species distributions.

Although stream size was an important predictor of fish-assemblage structure, soil factors were of equal or greater importance for many species. For example, KFACT, which is a measure of...
soil erodeability, averaged over each catchment was of considerable importance for both individual species and assemblage modeling. In our study area, such soil variables separated high gradient streams flowing through terraced limestone and shale layers from streams with low gradient, high turbidity that flowed through highly erodible soils. This is consistent with previous studies in the Big Blue River basin that show fish assemblages are separated based on three dominant substrate types (mud, gravel, and sand; Minckley 1959; Gido et al. 2002). In addition, catchments with porous soils and steep slopes should also have a higher degree of groundwater input and less variable flow regimes (e.g., Baker et al. 2003). Thus, species that typically occupy spring-fed stream reaches with stable hydrologic regimes should respond to these geologic factors. This is illustrated by the inverse relationship between KFACT and the occurrences of southern redbelly dace, which typically occupy headwater springs (Cross 1967; Figures 6 and 7).

It is informative to learn that adding site habitat measurements to models with habitat measurements from broader spatial scales adds little to our ability to predict species occurrences, particularly because broad-scale habitat measurements are relatively easy to capture using data layers from a GIS. However, because landscape features may be surrogates for local-scale phenomena (Wall et al. 2004), it is often necessary to identify local-scale habitat conditions to understand the mechanistic response of fish assemblages along these gradients. For example, feeding and reproduction of two darter species evaluated in this study (orangethroat darter and Johnny darter) are closely linked to coarse substrates (Cross 1967). However, in our DFA models, geologic (KFACT) and landcover (percent grassland) variables were selected as important predictors of their occurrences, rather than substrate size. Thus, even though coarse substrates are required for the survival of this species, other habitat variables that may covary with substrate size or that reflect other important habitat features appear to be equally important.

Alternatively, habitat data averaged across a field site may be poor predictors of fish-assemblage composition. In particular, averaging habitat measurements such as substrate size and stream depth may not adequately reflect important meso- or microhabitats required by particular fish species. Rather, fishes may respond to smaller habitat patches, such as deep pools or riffle characteristics. There are also a number of habitat features that were not quantified in our sampling such as the occurrence of large woody debris, which is known to influences the abundance of stream fishes (Angermeier and Karr 1984). Perhaps a more thorough quantification of mesohabitats would have increased our predictive ability. Nevertheless, to identify the utility of habitat measurements at finer spatial scales, it would be necessary to divide field sites into smaller spatial units (i.e., local habitat units) in order to partition variation attributed to variables measured at these fine spatial scales.

Spatial extent of sampling influences the relative importance of variables measured at different spatial scales (Roth et al. 1996). Because of the large spatial extent of this study, it is possible the inclusion of soil factors simply reflected zoogeographic barriers among catchments. For example, bigmouth shiner *Notropis dorsalis* was primarily captured in the Delaware River basin, and results from both the CCA and DFA indicated that soil erodibility (KFACT) was an important predictor of its occurrence. Although there is potentially unique habitat in the Delaware River basin, bigmouth shiner may simply be isolated in this system and have not dispersed to other regions within our study area. Thus, the relatively high importance of catchment-scale habitat measurements could be due to the characterization of zoogeographic boundaries.

In conclusion, our ability to predict fish-assemblage structure in the Kansas River basin did not notably improve with the addition of sitescale habitat measurements. These findings are consistent with recent studies (e.g., Roth et al.}
282 Gido et al.

1996; Poff 1997; Wang et al. 2001; Joy and Death 2004) that indicate landscape features exhibit important constraints on the distribution of fishes. Moreover, predictive modeling for conservation of aquatic systems, such as in Gap Analysis (e.g., Sowa et al. 2004; Wall et al. 2004), may not require intensive site-scale habitat quantification. Nevertheless, field habitat measurements provide a mechanistic understanding of species–habitat relationships that contribute to a better understanding of large-scale problems (Poff 1997; Park et al. 2003). We suggest that further research on the linkages between habitats across spatial scales will likely improve our understanding of species–environment relationships across spatial scales.

ACKNOWLEDGMENTS

We thank the Kansas Department of Wildlife and Parks for providing data used in this study. Scott Sowa and Gust Annis of the Missouri Resource Assessment Partnership provided assistance in developing GIS variables. Comments on an earlier version of this manuscript by Shannon Brewer, Charles Rabeni, Paul Seelbach, Scott Sowa and an anonymous reviewer were very much appreciated. This work was supported by grants from the U.S. Geological Survey National Gap Analysis Program and the Kansas Department of Wildlife and Parks.

REFERENCES


