GIS-BASED EXPERT SYSTEMS MODEL FOR PREDICTING HABITAT
SUITABILITY OF BLACKSIDE DACE IN SOUTHEASTERN KENTUCKY

Benjamin Blandford, Ph.D. candidate, Research Associate
and
John Ripy, Systems Engineer
and
Ted Grossardt, Ph.D., Research Program Manager
Kentucky Transportation Center
176 Raymond Building
University of Kentucky
Lexington KY 40506
Tel: 859 257 7504
Fax: 859 323 3734
E-mail: Benjamin.Blandford@uky.edu;
johnripy@adelphia.net;
tgrossardt@uky.edu;

A paper prepared for the 2013 TRB Annual meeting
Submitted August 01, 2012
Number of words in text: 5781
Number of figures and tables: 5
Total Word count: 7031
ABSTRACT
This study presents a GIS-based predictive habitat suitability model for the blackside dace, a federally-listed threatened species of the Upper Cumberland River basin in southeastern Kentucky. The model is a rules-based system which incorporates expert knowledge about habitat preferences for the species. The five habitat factors identified by experts and included in this model are stream gradient, canopy coverage, riparian vegetation type, riparian zone width, and stream order. Using GIS, the five habitat parameters were parameterized and combined across the entire stream network. Combinations were evaluated by blackside dace experts in terms of habitat suitability. The resulting model was tested against known blackside dace occurrences using locational modeling statistics. This analysis demonstrates success at identifying stream areas of both high and low likelihood of occurrences. Model results could be of particular usefulness to transportation planners in identifying sensitive areas in the landscape that may impact transportation planning.
As part of the National Environmental Policy Act (NEPA) process, transportation planners must be responsive to both state and federal Departments of Fish and Wildlife when making determinations about potential impacts on Threatened and Endangered Species. Since much of this determination hinges on habitat impacts associated with particular species, better understanding of the habitat combinations and locations will assist transportation planners in avoiding obvious problems and subsequently be more practical in planning decisions.

Using insights gained from the predictive modeling of pre-historic archeological sites in Kentucky, the team developed a rules-based Geographic Information Systems (GIS) model that uses existing or available data coverages, combined with expert knowledge about optimum habitat, to build a more comprehensive and accurate predictive model for the habitat conducive to certain Threatened and Endangered species in Kentucky. As a prototype, the team used blackside dace habitat as a modeling goal. This type of minnow is found in certain types of small streams, especially in Southeastern Kentucky, and is especially sensitive to changes in stream quality and type. Thus it is the type of species that would constitute a serious challenge to the Kentucky Transportation Cabinet. Further, the habitat characteristics for stream-dwelling species include landside descriptors as well as stream descriptors, so it encourages work in both types of GIS data.

For conservation purposes, general guidelines are available for identifying areas with potential blackside dace populations in accordance to its ecological preferences. However, explicit spatial modeling of blackside dace distribution among the subwatersheds within the species’ nature range has not yet been done in the past. More accurately delineating potential areas with quality habitat for the blackside dace at the regional scale is important for better conserving this species and managing the riparian zones that its subsistence depends upon. Such knowledge is also beneficial for transportation planners, who, having such knowledge from the outset, may either avoid or account for transportation planning projects that disturb critical habitats. This saves both time and money for the DOT.

EXPERT SYSTEMS

A common approach to current habitat modeling is to construct a regression analysis using many possible environmental factors and existing GIS data. The existing data may be only a crude representation of the imagined critical factor, and when combined with many other similarly lightly-theorized data strings in a model, can lead to a misunderstanding of relationships.

Aware of the potential problems of this approach, the project team instead adopted a ‘rules-based’ philosophy, relying on existing expert knowledge and research to help carefully describe the most relevant habitat factors related to the survival of the blackside dace. To pursue this, the team contracted with the Kentucky State Nature Preserves, as holders of both data and knowledge regarding the blackside dace. Through a series of interviews and literature reviews, the research team developed a list of most critical habitat factors. This list drove the team’s extensive GIS development process, as the team worked to most closely match, with data, the habitat conditions identified verbally or in the literature.

As detailed later in this paper, much of the team’s time and effort was devoted to locating, acquiring, translating, and transforming data from a variety of sources into GIS data layers that could then be used for analytic purposes. At least one factor identified as critical by the wildlife professionals could not be adequately captured through any available data sources for the geographic extent of the research area, nor could it be derived through extrapolation methods. Other factors were represented with greater or lesser faithfulness. Regardless, the
careful attention to this step was important, as the same base data layers were then used in two ways to attempt to predict the likelihood of species occurrence.

A ‘rules-based’ system basically uses expert opinion and published research to specify most likely relationships between the factors modeled spatially, and the occurrence of the species of interest. An approach like this allows for greater complexity in the interaction of the factors, if in fact, such complexity is understood. Because it is based on general and broad-based knowledge about the species of interest, it may lack some local specification or factor that is particularly relevant. Conversely, the model, once well developed, may be more broadly applicable across the landscape.

Additionally, a system like this does not rely on the existence of an observed data set to develop a spatial model, as in the case of statistical models. In cases where data is sparse, missing, or unreliable, it may be the only option for developing a model. In the case of the blackside dace, a certain amount of observation data was available from the Kentucky State Nature Preserves GIS files.

Because of the existence of some limited observation data regarding the blackside dace, and because of the detailed derivation of GIS data layers to represent identified critical factors, it was also possible to derive a model based on the spatial relationship between observed instances of the species and the associated landscape properties. This modeling procedure followed parameter estimation methods already developed in the literature, but used more advanced GIS data developed for the rules-based system.

Given exhaustive surveying of species occurrence at the regional scale and above is very difficult, simulation techniques known as species distribution modeling, habitat modeling, or ecological niche modeling (ENM) have been very useful to construct distribution maps of species that approximates reality. Such modeling often first utilizes limited field data and environmental variables to establish relationships between species occurrence and habitat attributes. The established model is then projected to an area of interest with continuous environmental data layers using GIS. In particular, modeling practices often employ influencing abiotic factors (i.e. scenopoetic factors), such as climate, landforms, and vegetation (which are more available in continuous formats) as predictors. Consideration of only physical variables generates a fundamental niche (environmental space). When limited by local factors in a given region, all areas that meet the requirements of a fundamental niche then forms a potential niche of a species. More in-depth modeling should also incorporate biotic interactions (i.e. bionomic factors) and species/population movement patterns (i.e. movement-related factors). The latter two factors being further determinants of the realized niche, in many cases, are hard to measure. Therefore, most studies employing ENM have focused on predicting the potential niche of species relying on corresponding environmental constraints in study areas. Anthropogenic impacts, in the case of blackside dace, as well as many other species, are indirect forces that alter physical conditions of habitats; therefore their consequences are reflected in environmental changes. However, when direct measurements of environmental quality (such as water turbidity) are not available, continuous land use maps can be used to incorporate human factors into prediction models. For example, proximity to coal mines may serve as a good proxy of disturbance to stream silt content, which in turn affects livelihood of aquatic species (citation). Since human impact commonly functions through changing abiotic factors of habitats, results from prediction using both natural and societal variables may be considered within the scope of potential niche.
The blackside dace was first identified scientifically and described by Starnes and Starnes (1978). The minnow is endemic to streams of the Upper Cumberland drainage, particularly in Southeastern Kentucky but also in Northeastern Tennessee. It resides primarily in small upland streams characterized by moderate flow, full overhead canopy, and an abundance of rocky substrates.

Southeastern Kentucky has for many decades been a site of significant natural resource extraction, particularly in the form of logging and mining. By the time of the blackside dace’s identification, significant extirpation was evident and largely attributed to manmade environmental disturbances. A survey by Starnes and Starnes (1980) of 168 streams in the Upper Cumberland drainage found the blackside dace present in only 27 of the streams. They estimated the minnow had been extirpated from 60 – 70 percent of its historical range.

According to the U.S. Fish & Wildlife, there are believed to be approximately 70 extant blackside dace populations in the Upper Cumberland drainage, many of which are small and disconnected remnant populations. Though many streams in the region exist that are considered habitable, blackside dace only inhabit a small number of these streams. For these reasons, the blackside dace was officially listed as a Threatened Species in 1987 by the U.S. Department of Fish and Wildlife.

Recent research has yielded significant findings regarding the habitat characteristics of the blackside dace. Dace streams are generally described as containing small to moderate populations that are relatively isolated, have restricted mobility, and are suffering from varying degrees of environmental impacts. Habitat modeling by Jones (2005) suggested that the following environmental characteristics were of significance to blackside dace populations: gradient, turbidity, dissolved oxygen content, water temperature, conductivity, percent riffle, and link magnitude. Further habitat modeling by Black (2007) suggested that conductivity and
water temperature were the most statistically relevant factors involved in presence/absence of blackside dace in the sample streams.

HABITAT FACTORS

The existing knowledge of blackside dace habitat suitability characteristics found in the literature provided a starting point for identifying which habitat indicators would be included in this project. Additionally, habitat factors of the blackside dace were discussed at length with experts at the Kentucky State Nature Preserves. Based on their knowledge and experience in the field surveying dace streams, a number of other factors were suggested for consideration, including riparian vegetation type and width, presence/absence of predator fish, density of bridges and culverts over the stream network, riffle/pool ratio, and the density of oil and gas wells in the watershed.

In determining which habitat factors to include for modeling, several considerations were involved. The primary consideration was the availability of relevant data for creating the model. A number of the suggested habitat factors from the literature and from the expert knowledge of the Nature Preserves biologists were either not contained in any available data sets or were inadequately described in available data sets. A second consideration was the relative likelihood of the habitat factor to impact the presence/absence of blackside dace in the streams. Finally, the total number of habitat factors needed to be limited for modeling purposes. After thorough deliberation between the project team and the Nature Preserves over these considerations, the following habitat factors were selected for modeling: gradient, canopy, riparian vegetation type, riparian zone width, and stream order.

Gradient

Gradient was chosen as habitat factor for analysis because of its correlation to the presence/absence of blackside dace in Upper Cumberland streams. Because the terrain of the Upper Cumberland drainage is so mountainous, gradient is believed to significantly influence the distribution of blackside dace. Streams characterized by moderate to high gradient are not believed to be conducive to blackside dace habitation, as are streams of unusually low gradient. According to statistical modeling of stream sampling conducted by Jones (2005), dace are most likely to inhabit streams of a gradient of one to six percent at the stream scale. Mattingly (2005) similarly found that blackside dace were four times more likely to occur in streams of a crude gradient of one to six percent.

For this model, data for stream gradient was acquired from the Kentucky Geospatial Data Clearinghouse; the data was derived from ten meter Digital Elevation Models (DEM) and mosaiced together to form a single image. Calculating gradient at the stream level entailed calculating a single average gradient value for each stream segment based on the change in elevation from the beginning of the segment to the end.

Canopy

Although canopy itself is seldom mentioned in the literature as directly correlating to the presence/absence of blackside dace, canopy is known to greatly impact the water temperature of Upper Cumberland streams. Temperature is described in the literature as being a critical component of dace habitats. Indeed, Black (2007) found temperature to be one of the most
significant predictive factors for blackside dace presence/absence. Unfortunately, existing and sufficient water temperature data was not available for all streams of the Upper Cumberland. However, because of the close relationship between canopy and water temperature, the team included canopy as a potential indicator of habitat suitability.

Several data sources were used to calculate percent canopy across the region: USGS Gap Analysis Vegetation Cover, and Kentucky's Statewide 2' Aerial Imagery. The imagery was split into three bands using the Composite Bands (Data Management) tool in ArcGIS. An “if” statement was created to locate cells where the blue band was greater than both the green and red, and the green band was greater than the red. The Gap vegetation cover (using Level III) was used as an overlay to identify false positive locations, such as ponds and lakes. Non-vegetation cells were clipped from the layer above. The percentage coverage was calculated between 2, the lowest possible blue value, and 255.

In order to derive fractional canopy cover along the streams, both aerial imagery and land cover map were used. Blue band from the visible color aerial imagery was empirically found to be most predictive in determining canopy. A band comparison criterion, blue > red and blue > green and green > red, was applied on aerial imagery to extract canopy shaded areas. Ponds, lakes, and non-vegetation were masked out using GAP analysis level III land cover map. Then for each remaining pixel in the area, a fractional coverage estimate was calculated using the ratio of the blue band digital number (8bit, 0-255) to the maximum value (255) present in the area.

**Riparian Vegetation**

Riparian corridors are known to play an essential role in water quality and habitat maintenance within an ecosystem. For streams of the Upper Cumberland drainage, effective riparian corridors filter out pollutants and other sediments, minimize flood events and regulate water temperatures. Damage to the riparian corridor is most often caused by changes to the land use, such as through logging, mining or agricultural practices. Such changes result in elevated stream temperatures and decreases in dissolved oxygen content of the water, both of which cause elevated stress levels and decreased competitive ability of blackside dace. These factors stress the importance of healthy riparian corridors to the habitability of streams, especially for blackside dace.

The Riparian Vegetation habitat factor did not entail significant data processing. The data acquired from USGS Gap Analysis Vegetation Cover contained assigned values for each cell pertaining to specific type of land cover. For example, a value of 1204 indicated developed land of high intensity; a value of 1402 indicated cultivate cropland; a value of 4126 indicated deciduous dominated forest and woodland; and so forth. The Riparian Vegetation habitat was calculated by considering the corresponding land use type adjacent to the respective streams along the network.

**Riparian Zone Width**

A riparian zone is characterized not only by the vegetation type but other factors as well including width, height and bank slope. Riparian zone width, in particular, is important in that it is a function of the effectiveness of the riparian zone at positively impacting waterways by filtering out pollutants and sediments and minimizing flood events. Sufficiently wide riparian
zones are likely to have a positive impact on suitability of streams. This model included riparian zone width as a potential indicator of habitat suitability for blackside dace.

Using a similar method described by Inlander (2004)\textsuperscript{14}, the riparian width for the study area was calculated utilizing spatial data obtained from several sources: USGS Gap Analysis Vegetation Cover, Kentucky's Statewide 2' Aerial Imagery, and Kentucky Single Zone 30ft DEM (USGS). Using the 30ft DEM, a boundary was delineated using topographic variables:
1) Slope (degrees) Less than/equal to 7
2) 3x3 Focal Variety Less than/equal to 5
3) Path Distance (slope as cost) Less than/equal to 225 to stream
The 2ft canopy cover layer was then clipped by the above boundary. From this, a Euclidean distance was calculated from the streams layer and constrained by the canopy cover. The maximum value was used as the riparian zone width.

**Stream Order (Strahler)**

Blackside dace are known to prefer small to moderate streams.\textsuperscript{5,6} For modeling purposes, Strahler stream order provides a useful way of capturing the relative volume and accumulation of a stream segment within a network. Under this hierarchical system, each stream segment is treated as a node. Accordingly, when two first-order streams, or source streams, join together, they form a second-order stream. When two second-order streams join together, they form a third-order stream, and so on. This hierarchical system provides a useful way of understanding the water volume and accumulation within a given stream network. Stream order (also known as link magnitude) has been found to be positively correlated with blackside dace presence/absence. Habitat modeling by Jones (2005) found stream order – when combined with connectivity – to be the strongest predictive habitat model of those tested.\textsuperscript{8}

The stream order data was created using the Hydrology Tools in ArcGIS. The Hydrology Tools operate by creating a predicted stream network based on the provided DEM for the area. Once the stream raster network was created with these tools, Stream Order was calculated using the Strahler ordering method.

**Factors Not Included**

A number of other factors identified either in the literature or by biologists at the Kentucky State Nature Preserves were not included in the model for a variety of reasons. One is the specific conductivity of streams. In Southeastern Kentucky, resource extraction, in particular strip mining, is argued to be a major cause of habitat degradation and decreasing populations of blackside dace.\textsuperscript{3,7,10,11} Among the potential impacts of strip mining on stream quality is elevated water conductivity levels. Specific conductivity is, in effect, a function of the dissolved solids within the water, and it provides a useful measurement for understanding water quality.\textsuperscript{15} High conductivity levels are known to negatively and severely impact the habitat of blackside dace.\textsuperscript{11} Unfortunately for this project, sufficient water conductivity data was not available across the entire study area. Several proxy measures were attempted to try and account for conductivity levels, including several different measures of mine density and oil and gas well density at various geographic scales. Ultimately, lacking any direct causal link between these proxy measures and the water conductivity levels for any and every stream link, this factor was omitted from the model.
Another factor not included in this model was a comprehensive measure of stream network connectivity. For blackside dace, connectivity may be limited in the stream network due to disturbances, such as manmade obstacles (e.g. dams or poorly aligned culverts) or naturally occurring obstacles (e.g. beaver dams), or because of significantly degraded or polluted streams. Limited stream network connectivity is considered to be a major problem facing blackside dace.\textsuperscript{5,8} Data describing such disturbances was largely not available for modeling purposes. For example, naturally occurring dams are largely temporal and may fluctuate significantly; GIS data simply does not exist for such occurrences.

On a smaller scale, a habitat factor frequently mentioned in the literature but not included in this model was the stream segments’ riffle/pool ratio. Starnes and Starnes (1981) describe the minnow as preferring riffle/pool rations near or below 60:40.\textsuperscript{4} Habitat modeling by Jones (2005) suggested that percent riffle was among the strongest single predictors of blackside dace presence.\textsuperscript{8} Riffle/pool ratio was not included in this model because of the lack of sufficient data to cover the entire Upper Cumberland drainage. Additionally, presence/absence of predator fish, such as largemouth bass and redbreast sunfish, was not included due to lack of sufficient data.

Despite these limitations, the project team felt the habitat factors that were included in the model – canopy, gradient, riparian vegetation, riparian width, and stream order – provided sufficient descriptive information toward the evaluation of potential habitat suitability for blackside dace. These habitat factors formed the basis for habitat model generation.

**RULES-BASED SURFACE GENERATION**

Parameterizing the Habitat Factors

To create an expert-based system, the habitat factors described above need to be categorized, or parameterized, in terms of likely impact on habitat suitability for blackside dace. This entails dividing each habitat factor individually into classes with distinct break off points indicating likely changes in suitability. For this model, the habitat factors were divided into four classes (Table 1), with a value of 1 indicating least suitability and a value of 4 indicating highest suitability. This classification scheme was created through consultation with Kentucky State Nature Preserves and previous habitat modeling research from the literature review.

**TABLE 1  Habitat factors and their parameterization in terms of habitat suitability for the Blackside dace.**

<table>
<thead>
<tr>
<th>Habitat Factor (Low)</th>
<th>(High)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gradient (stream level)</td>
<td>&lt;6 percent</td>
</tr>
<tr>
<td>Canopy (percent coverage)</td>
<td>0 – 50 percent</td>
</tr>
<tr>
<td>Riparian Vegetation</td>
<td>Cultivated, Developed, Barren</td>
</tr>
<tr>
<td>Riparian Zone Width</td>
<td>&lt;6 meter</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TRB 2013 Annual Meeting Paper revised from original submittal.
For each habitat factor, a stream raster was created and classified according to the criteria in Table 1. Raster cells for the streams were created at 10 meters. This resulted in five stream raster networks (one for each habitat factor) with encoded habitat factor data for every 10 meter segment along the stream network. Thus, for example, every 10 meter raster cell for riparian vegetation contains a value of one through four, indicating the type of riparian vegetation for that respective 10 meter stream segment (as classified in Table 1). This was produced across the entire stream network at 10 meter intervals for all five habitat factors.

Once each of the habitat factors were rasterized and parameterized, they were combined into a single raster representing the stream network of the Upper Cumberland drainage (Figure 2). For each 10 meter raster cell along the stream network, a unique value was produced which combined the data from each of the respective habitat factors. In effect, this created 10 meter snapshots of the stream segments containing highly detailed blackside dace related habitat information. For example, a cell may be coded 43232: each successive digit reflecting the parameterization of each individual habitat factor for that 10 meter cell. In this case, the gradient (4) is <2 percent; the canopy (3) is 70 – 90 percent; the riparian vegetation (2) is grass, herbaceous, or pasture; the riparian zone width (3) is 12 – 18 meter; and the stream order (2) is an order of 1. This information was computed for every 10 meter segment of the stream network.

<table>
<thead>
<tr>
<th>Stream Order (Strahler)</th>
<th>6 – 7</th>
<th>1</th>
<th>4 – 5</th>
<th>2 - 3</th>
</tr>
</thead>
</table>

**FIGURE 2** Five GIS layers are combined into one raster dataset representing combinations of habitat factors.
Expert Evaluation

With each of the habitat factors clearly identified, parameterized, and mapped in GIS, the next step was to evaluate the possible habitat factor combinations (scenarios) in terms of suitability to blackside dace. Again, the project team solicited expert input from the Kentucky State Nature Preserves. Evaluations were conducted by considering the combined influence of each of the habitat factors for a given scenario, and then rating the scenario on a scale of one to four (one being not suitable, four being highly suitable). With such an expert-driven system, habitat suitability can be ascertained across the entire stream network of the Upper Cumberland drainage.

Because the total number of potential combinations of habitat factors (over a thousand) is too large for each to be individually evaluated by experts, a weighted system was created. This system began with a set of rules identified by experts. For example, all cells within the Stream Order parameter 1 (indicating a Strahler order of 6-7, per Table 1) were automatically given an overall score of 1 in terms of habitat suitability. The experts deemed all large streams as highly unsuitable for blackside dace, regardless of the other habitat factors. Additionally, any cells within the Gradient parameter 1 (greater than 6 percent) and Stream Order parameter 2 (Strahler order of 1) were also automatically given an overall score of 1 in terms of habitat suitability. In this case, shallow streams along steep slopes were also deemed highly unsuitable for blackside dace. Once these initial rules were accounted for, the following weighted formula was applied across all remaining cells.

\[ S = \sum Wi \times Xi \]

Where:
- \( S \) = surface of total probability score
- \( Wi \) = influence or weight factor of the ith factor
- \( Xi \) = Criteria score for the ith parameter

So,

\[ S = (GRw \times GRx + CAw \times CAx + VGw \times VGx + RWw \times RWx + SOW \times SOx) \]

Where:
- \( GR \) = Gradient
- \( CA \) = Canopy
- \( VG \) = Riparian Vegetation
- \( RW \) = Riparian Zone Width
- \( SO \) = Stream Order

In this formula, Gradient and Stream Order were given the most weight, at 30 percent each, Canopy was given 20 percent weight, and Riparian Vegetation and Riparian Width were each given 10 percent. Figure 3 shows the stream raster data set according to the resulting weighted score.
FIGURE 3 Stream raster dataset is evaluated by experts, and combinations are rated for habitat suitability.

OBSERVATIONS

To test the precision of the resulting habitat suitability model, blackside dace observation data was obtained from the Kentucky State Nature Preserves. This GIS data was in the form of spatially referenced points depicting where blackside dace had been previously observed by state biologists. There were a total of 213 observance points distributed across the upper Cumberland watershed. This compared to the 620,940 unique 10 meter raster cells created through the rules-based expert system model. Put together, blackside dace occurrences have been observed at only .034 percent of the potential 10 meter segments along the entire stream network (213/620,940). This ratio is the basis for testing the relative strengths and weaknesses of the model created for this project. For example, if a random circle was drawn over a set of the total possible sites, one would expect to catch, on average, the same ratio of positive sites to total sites. Thus a random model would yield 3.4 hits for every 10,000 sites. Numbers below that are categories good at predicting absence, and numbers above that are categories good at predicting presence.

Blackside dace observation data obtained from the Kentucky State Nature Observes was used to test the rules-based model after it was built. The model was analyzed using locational modeling statistics as described by Kvamme (2006). This data set is the 'test bed' against which the categorized overlays are tested. The rules-based model predicts one of four categories of likelihood ranging from lowest (1) to highest (4) or simply categories 1-4 (Figure 4).

- Habitat Suitability Rank 1 (low suitability) covers 28.49 percent of the total applicable stream bed. Only two dace presence observations occurred at stream segments with this ranking, resulting in a model accuracy of 0.94 percent. At any location within this category, there is a 0.00113 percent chance of hitting a blackside dace site. Within this category, it is 0.02 times as likely to have dace occurrences as all other classified locations in the model.

- Habitat Suitability Rank 2 (below average suitability) covers another 22.28 percent of this total applicable stream bed. 14 dace presence observations occurred at stream segments with this ranking, resulting in a model accuracy of 6.57 percent. At any location within this category, there is a 0.01012 percent chance of hitting a blackside dace site. Within this category, it is 0.25 times as likely to have dace occurrences as all other classified locations in the model.
- Habitat Suitability Rank 3 (average suitability) covers another 26.46 percent of the total applicable stream beds. 52 dace presence observations occurred at stream segments with this ranking, resulting in a model accuracy of 24.41 percent. At any location within this category, there is a 0.03165 percent chance of hitting a blackside dace site. Within this category, it is 0.90 times as likely to have dace occurrences as all other classified locations in the model.

- Habitat Suitability Rank 4 (high suitability) covers the last 22.78 percent of the total applicable stream beds. 145 dace presence observations occurred at stream segments with this ranking, resulting in a model accuracy of 68.08 percent. At any location within this category, there is a 0.10253 percent chance of hitting a blackside dace site. Within this category, it is 7.23 times more likely to have dace occurrences than all other classified locations in the model.

<table>
<thead>
<tr>
<th>Habitat Suitability Rank</th>
<th>M</th>
<th>M'</th>
<th>S</th>
<th>P(M): Base rate or chance probability that a model will indicate a site; proportion of study region mapped to M</th>
<th>P(M/S): Model accuracy; probability that a model will correctly indicate a site</th>
<th>P(S/M): Probability of Blackside dace presence when model specifies an occurrence</th>
<th>P(S/M)/p(S/M'): Indicates how many times more likely an occurrence is in M versus M'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>176879</td>
<td>176877</td>
<td>2</td>
<td>28.49%</td>
<td>0.94%</td>
<td>0.0013%</td>
<td>0.02</td>
</tr>
<tr>
<td>2</td>
<td>138361</td>
<td>138347</td>
<td>14</td>
<td>22.28%</td>
<td>6.57%</td>
<td>0.01012%</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>164272</td>
<td>164230</td>
<td>52</td>
<td>26.46%</td>
<td>24.41%</td>
<td>0.03165%</td>
<td>0.90</td>
</tr>
<tr>
<td>4</td>
<td>141428</td>
<td>141283</td>
<td>145</td>
<td>22.78%</td>
<td>68.08%</td>
<td>0.10253%</td>
<td>7.23</td>
</tr>
</tbody>
</table>

**FIGURE 4 Model Results Using Locational Model Statistics.**

The results from this model are promising. The low habitat suitability categories of 1 and 2 indeed result in a very low likelihood of dace occurrences. The average habitat suitability category of 3 results in a nearly statistically even compared with random chance (0.90) of having a dace occurrence. The high suitability category of 4 has a much greater than random chance of predicting dace occurrences. These indicators demonstrate considerable precision of the model in identifying both highly suitable areas and highly unsuitable areas for blackside dace.

This model does have room for improvement as well. The two blackside dace occurrences in the lowest category may be too many for the comfort of transportation planners. Closer analysis of the GIS model reveals that the model’s shortcomings here are likely related to the gradient calculations. In both cases, dace presence was recorded near the mouth of small streams of overall high gradient – however the stream gradient leveled out somewhat near the mouth of the streams where the dace presences were recorded. In one of the instances, the dace occurrence was less than one hundred meters away from the stream’s conflux with a highly rated stream (category 4). Higher precision in gradient calculations could improve the model’s performance.

Based on the literature and expert testimony, this model could also be improved by the inclusion of stream specific conductivity measures. This particular habitat factor was heavily stressed by experts as being correlated to blackside dace occurrences in a limiting manner. In other words, where specific conductivity readings exceeded a particular level, streams were deemed highly unsuitable for blackside dace, regardless of any other habitat characteristics. Sufficient and comprehensive conductivity data across the entire stream network however is not
Currently available. Until or unless such exhaustive water sampling is recorded, this habitat factor will remain omitted from future similar modeling endeavors.

**CONCLUSIONS**

This paper demonstrates how a GIS-based expert systems model can be effectively deployed to identify habitat suitability for blackside dace across the Upper Cumberland watershed in southeastern Kentucky. The model successfully identified areas of both high and low likelihood of blackside dace occurrences. Further fine tuning of the model could potentially yield even more precision in terms of comparing the model results against known blackside dace occurrences.

These results suggest the potential viability of such GIS-based predictive habitat models for other known threatened or endangered species of interest to environmental professionals and transportation planners. Future research could be directed toward aquatic and terrestrial habitats alike. This type of model is most successful in situations where sufficient expert knowledge about the species and its preferred habitat already exists prior to modeling. In situations of high uncertainty regarding species habitat preference, this type of model may require many iterations to successfully identify the most correlative habitat factors for inclusion. However, this type of model is particularly useful in situations where the exact range of a species has not yet been determined. Incorporating expert knowledge as the input for predictive habitat modeling, rather than already known presence/absence data, is demonstrated to successfully identify suitable habitats with considerable precision.
ACKNOWLEDGMENTS

This project benefitted greatly through collaboration with experts at the Kentucky State Nature Preserves, in particular Ryan Evans and Sara Hines. Their knowledge, experience, and data regarding the habitat preferences for blackside dace helped inform much of the project execution.
REFERENCES


