Enhancing the Aquatic Gap Analysis for Arkansas: Developing Reach-Scale Species Distribution Maps for Crayfish and Mussel Species in Arkansas

Project Summary

Project Leader
Ethan Inlander, Ozark Rivers Program Director
Ozark Highlands Office, The Nature Conservancy
einlander@tnc.org
675 North Lollar Lane
Fayetteville, Arkansas 72701
(479) 973-9110 (office)
(479) 973-9135 (fax)

Project Partner
Scott Sowa
Assistant Director
Missouri Resource Assessment Partnership (MoRAP)
sowasp@missouri.edu
4200 New Haven Road
Colombia, MO 65201
(573) 441-2791

The total cost of this project will be $81,840.

The Nature Conservancy respectfully requests $40,920 to complete this project and will provide the remaining $40,920 as match (50%).

INTRODUCTION

In July of 2007, the USGS National Aquatic Gap Analysis Program awarded a grant to the Missouri Resources Assessment Partnership (MoRAP) and The Nature Conservancy of Arkansas (TNC) to create stream reach-scale distribution models for all fish species known to occur in the state of Arkansas. This is the first major step toward completing an aquatic gap analysis for Arkansas, which will provide significant information, and recommendations for future revisions of the Arkansas Wildlife Action Plan (AWAP).

BACKGROUND

Gap analysis is a conservation assessment methodology that compares the distribution of species or natural communities to the distribution of lands and waters that are managed for native species and natural ecosystem processes. To accomplish this task, it is necessary to develop distribution maps of individual animal species for comparison with maps of land stewardship and management status. With this information, more accurate estimates can be made about the amount of available habitat for each species, how much habitat has been lost, how much is currently represented within the existing matrix of protected lands, and where are the best management options for conserving a particular species or community. A terrestrial gap analysis was completed for the state of Arkansas in the 1990s.

Major tasks to be completed for aquatic gap analysis are aquatic ecological classification, stewardship analysis, human threats analysis, and species modeling. The project currently funded by USGS will complete two pieces of information critical to an aquatic gap analysis including the creation of a classified stream network and predictive distribution models of fish species throughout Arkansas.

This project requires compiling biological sampling data from various sources in Arkansas and characterizing the local and watershed conditions of every stream segment. These data will then be utilized to model fish species affinities to particular local and watershed characteristics (habitats) within the species' range. Maps and a GIS database will be generated to identify all the suitable habitats for these species throughout their professionally reviewed range in Arkansas.

Deliverables will include a relational database of fish community sampling data, modified 1:100,000 NHD stream networks with local and watershed metrics, GIS datasets of all the species model results, and maps of the resulting species models. MoRAP will handle most tasks pertaining to stream classification and attribution while TNC will undertake most of the work pertaining to assembling the species collection information and development of the species models with limited assistance from MoRAP.

Specific Objectives of Funded Project

By September of 2010 we will develop a classified stream network and stream reach-scale models for 207 native and 19 introduced fish species for a total of 226 models across the state of Arkansas.The specific objectives of this project are to:

1. Classify stream segments contained within the 1:100,000 National Hydrography Dataset (NHD) into distinct valley segment types according to distinct combinations of factors known to individually and collectively influence local biophysical conditions.
2. Generate stream reach specific drainage area polygons (catchments/segmentsheds) for each primary channel stream segment in the 1:100,000 scale NHD.
3. Generate local and watershed statistics pertaining to soils, geology, and landform for every stream segment in the NHD.
4. With a reasonable level of effort, compile existing fish sampling data collected from streams across Arkansas.
5. Enter or transfer compiled data into a Microsoft Access relational database
6. Link each sampling record to the USGS/EPA National Hydrography Dataset and the USGS/NRCS Hydrologic Unit Coverage.
7. Develop professionally reviewed range maps for all fish species by 8, 10 or 12-digit HU.
8. Develop predictive distribution models and maps showing the stream segments each fish species is likely to occur in under relatively natural conditions in Arkansas.
9. Write final report.

SEE ***ATTACHMENT A*** FOR THE COMPLETE PROJECT PROPOSAL FUNDED BY USGS.

PROJECT PROPOSAL

TNC proposes to enhance the project currently funded by USGS (modeling for fish species) by developing stream reach-scale distribution models for crayfishes and mussels in Arkansas. Objectives 1, 2, and 3 above will be completed entirely through the USGS project. The proposed project would fund Objectives 4 to 9 for crayfish and mussels over a three-year period. These new objectives are referred to below as Objectives 4A to 9A.

YEAR 1:

Objective 4A:

With a reasonable level of effort, compile existing crayfish and mussel sampling data collected from streams across Arkansas. (Coordinating this effort with the data collection effort for fishes will reduce redundancy so that collections holders will only need to provide data one for fishes, crayfish and mussels, instead of providing fish data in 2008, and providing crayfish and mussel data in the future.)

Deliverable: Paper and digital record of species and community collections.

YEAR 2:

Objective 5A:

Enter or transfer compiled data into a Microsoft Access relational database.

Deliverable:

Attributed Microsoft Access relational database

Objective 6A:
Link each sampling record to the USGS/EPA National Hydrography Dataset and the USGS/NRCS Hydrologic Unit Coverage.
Deliverable:
Georeferenced species occurrences, referenced to both stream reaches and HUCs.

<u>YEAR 3:</u>
Objective 7A:
Develop professionally reviewed range maps for all fish species by 8, 10 or 12-digit HU.
Deliverable: Professionally reviewed range maps for each species.

Objective 8A:
Develop predictive distribution models and maps showing the stream segments each fish species is likely to occur in under relatively natural conditions in Arkansas.
Deliverable: Predicted distribution model and map for each species.

Objective 9A:
Write final report.
Deliverable: Final Report

NEW BUDGET

Year 1	SWG	TNC	Sub-Total	% Match
Personnel / Fringe	$ 3,252	$ 3,252	$ 6,504	50%
Supplies	$ 1,355	$ 1,355	$ 2,710	50%
Travel	$ 542	$ 542	$ 1,084	50%
Sub-Total	$ 5,149	$ 5,149	$ 10,298	50%
Indirect (23%)*	$ 1,184	$ 1,184	$ 2,369	50%
Sub-Total	**$ 6,333**	**$ 6,333**	**$ 12,667**	**50%**

Year 2	SWG	TNC	Sub-Total	% Match
Personnel / Fringe	$ 5,775	$ 5,775	$ 11,550	50%
Supplies	$ 1,000	$ 1,000	$ 2,000	50%
Travel	$ 667	$ 667	$ 1,333	50%
Sub-Total	$ 7,442	$ 7,442	$ 14,883	50%
Indirect (23%)*	$ 1,712	$ 1,712	$ 3,423	50%
Sub-Total	**$ 9,153**	**$ 9,153**	**$ 18,307**	**50%**

Year 3	SWG	TNC	Sub-Total	% Match
Personnel / Fringe	$ 19,011	$ 19,011	$ 38,022	50%
Supplies	$ 1,000	$ 1,000	$ 2,000	50%
Travel	$ 667	$ 667	$ 1,333	50%
Sub-Total	$ 20,678	$ 20,678	$ 41,355	50%
Indirect (23%)*	$ 4,756	$ 4,756	$ 9,512	50%
Sub-Total	**$ 25,433**	**$ 25,433**	**$ 50,867**	**50%**

Project Total	SWG	TNC	Total	% Match
Personnel / Fringe	$ 28,038	$ 28,038	$ 56,076	50%
Supplies	$ 3,355	$ 3,355	$ 6,710	50%
Travel	$ 1,875	$ 1,875	$ 3,751	50%
Sub-Total	$ 33,268	$ 33,268	$ 66,536	50%
Indirect (23%)*	$ 7,652	$ 7,652	$ 15,303	50%
Total	**$ 40,920**	**$ 40,920**	**$ 81,840**	**50%**

QUALIFICATIONS

Ethan Inlander has been applying geospatial technologies and physical sciences to conservation issues for over 12 years. He received his undergraduate and master's degrees from the Department of Geography at University of California Santa Barbara, the #1 geography program in the US (NRC, phds.org). His thesis topic was "An Integrated Methodology for the Mapping and Inventory of Riparian Areas in the Upper Santa Ynez Watershed, California ". Before joining The Nature Conservancy, Ethan applied geographical information systems technology to address multiple scale conservation problems in riparian and costal habitats of California. Since joining The Nature Conservancy, Ethan has applied these same techniques to identify and reduce impacts and habitat degradation to freshwater stream ecosystems, conduct local, watershed, and regional threat assessments of subterranean environments, and prioritize and implement karst and riverine conservation actions at multiple scales.

Scott Sowa is the Assistant Director for the Missouri Resource Assessment Partnership (MoRAP). His duties include assisting the Director with developing, planning and coordinating the progress of existing and future projects. He also assists with grant procurement and management, and the recruitment, hiring and oversight of staff. Scott serves as the principal investigator for numerous projects, particular those pertaining to aquatic ecosystems. Scott holds a Ph.D. in Fisheries and Wildlife from the University of Missouri. His dissertation topic was; "A Multiscale Investigation of Factors Influencing Local Biophysical Conditions in Headwater Streams of the Missouri Ozarks". Scott also holds a M.S. in Fisheries and Wildlife from the University of Missouri and a B.S. in Fisheries and Wildlife from Michigan State University. His areas of expertise include stream ecology, riparian ecology, fluvial geomorphology, aquatic ecological classification, and GIS.

Project title:
Developing Stream Reach-Scale Predicted Distribution Models for Fish Species in Arkansas

Principal Investigators:

Scott P. Sowa, PhD
4200 New Haven Road
Columbia, Missouri 65201
Ph: (573) 441-2791
FAX: (573) 876-1863
Email: Sowasp@missouri.edu
Missouri Resource Assessment Partnership
University of Missouri

Ethan Inlander
675 North Lollar Lane
Fayetteville, Arkansas 72701
(479) 973-9110
(479) 973-9135
einlander@tnc.org
The Nature Conservancy

Project duration: 3 Years

Proposed project start date: September 2007

Total Funds Requested: $280,000

SYNOPSIS OF PROJECT

The Missouri Resource Assessment Partnership (MoRAP) and The Nature Conservancy (TNC) propose to create stream reach-scale distribution models for all fish species known to occur in the state of Arkansas. This project requires compiling biological sampling data from various sources in Arkansas and characterizing the local and watershed conditions of every stream segment in the 1:100,000 scale National Hydrography Dataset (NHD). These data will then be utilized to model fish species affinities to particular local and watershed characteristics (habitats) within the species' range. Maps and a GIS database will be generated to identify all the suitable habitats for these species throughout their professionally reviewed range in Arkansas. Deliverables will include a relational database of fish community sampling data, modified 1:100,000 NHD stream networks with local and watershed metrics, GIS datasets of all the species model results, and maps of the resulting species models. The resulting stream network and species models are two critical components for conducting an aquatic gap analysis. This project will be a joint effort with MoRAP taking on most tasks pertaining to stream classification and attribution while TNC will undertake most of the work pertaining to assembling the species collection information and development of the species models with limited assistance from MoRAP.

OBJECTIVES

By September of 2010 we will develop a classified stream network and stream reach-scale models for 207 native and 19 introduced fish species for a total of 226 models across the state of Arkansas. Major tasks that must be completed in order to conduct a gap analysis are aquatic ecological classification, stewardship analysis, human threats analysis, and species modeling. This project will complete two pieces of information critical to an aquatic gap analysis including the creation of a classified stream network and predictive distribution models of fish species throughout Arkansas. Arkansas comprises a portion of the Mississippi River Basin, an area of interest to the Gap Analysis Program. The specific objectives of this project are to:

1. Classify stream segments contained within the 1:100,000 National Hydrography Dataset (NDH) into distinct valley segment types according to distinct combinations of factors known to individually and collectively influence local biophysical conditions.
2. Generate stream reach specific drainage area polygons (catchments/segmentsheds) for each primary channel stream segment in the 1:100,000 scale NHD.
3. Generate local and watershed statistics pertaining to soils, geology, and landform for every stream segment in the NHD.
4. With a reasonable level of effort, compile existing fish sampling data collected from streams across Arkansas.
5. Enter or transfer compiled data into a Microsoft Access relational database
6. Link each sampling record to the USGS/EPA National Hydrography Dataset and the USGS/NRCS Hydrologic Unit Coverage.
7. Develop professionally reviewed range maps for all fish species by 8 or 11-digit HU.
8. Develop predictive distribution models and maps showing the stream segments each fish species is likely to occur in under relatively natural conditions in Arkansas.

BACKGROUND AND JUSTIFICATION
Gap analysis is a conservation assessment methodology that compares the distribution of several elements of biological diversity to the distribution of lands and waters that have been set aside and are primarily managed for native species and natural ecosystem processes (Scott et al. 1993). To accomplish this task, it is necessary that GAP develop relatively high-confidence distribution maps of individual animal species for comparison with maps of land stewardship and management status. The purpose of these predicted distribution maps is to provide more precise information about the distribution of individual native and nonnative species. With this information, more accurate estimates can be made about the amount of available habitat for each species, how much has been lost, how much is currently represented within the existing matrix of public lands, and where are the best management options for conserving a particular species or community. The goal of this project is to develop statewide predicted distribution models and maps for all fish species in Arkansas.

For each species we will develop a probability of occurrence model and a simple binary presence model. The models will be applied to the statewide 1:100,000 modified national hydrography dataset (NHD) and clipped to each species' professionally-reviewed range. All of the individual model results will be merged into a "hyperdistribution" format, which allows users to simultaneously view all of the species predicted to occur in each stream segment in the state and their associated probabilities of occurrence.

Expected Benefits

The refined species range and distribution maps will display individual stream reaches that are likely to provide habitat for each fish species under relatively natural conditions. These reach-scale maps can be used in future State Wildlife Grant (SWG) implementation projects, or other conservation efforts, to identify specific sub-watersheds, stream reaches, or even individual properties where habitat protection or restoration will be most effective. These refined species maps will update the Arkansas Wildlife Action Plan, and will increase the efficiency and effectiveness of future on the ground conservation actions.

The maps will likely reveal previously unknown populations of modeled species, as was evidenced by work done by MoRAP for the Missouri Aquatic Gap Project (Sowa et al. 2005). A significant example from the MoRAP project was demonstrated for the Plains Topminnow, which had three known historic populations, only one of which was thought to still exist. Habitat affinity modeling identified 15 additional potential areas for the species. Eleven of these new areas were sampled, and four of the areas had living populations of the species. Results like these could be used in Arkansas to modify the priority scores for individual species of greatest conservation need (SGCN) if a species has greater populations than previously known.

WORK PLAN/APPROACH/METHODS

Task 1: Valley Segment Type Classification

We will classify stream segments contained within the 1:100,000 National Hydrography Dataset into distinct valley segment types according to distinct combinations of factors known to individually and collectively influence local biophysical conditions. For this objective we will follow the protocols developed by MoRAP for the Missouri Aquatic Gap Pilot Project (Annis et al. 2002; Sowa et al. 2005).

Valley Segment Types (VSTs) are defined and mapped to account for longitudinal and other linear variation in ecosystem structure and function that is so prevalent in lotic environments. VSTs represent hydrogeomorphic units defined by local physical factors and their position in the stream network. They stratify stream networks into major functional components that define broad similarities in fluvial processes, sediment transport, riparian conditions, and thermal regimes. Each individual valley segment is a spatially distinct habitat, but valley segments of the same size, temperature, flow, gradient, etc. all fall under the same VST.

Stream segments within the 1:100,000 USGS/EPA National Hydrography Dataset (NHD) will be attributed according to various categories of stream size, flow, gradient, temperature, and geology through which they flow, and also the position of the segment within the larger drainage network. These variables have been consistently shown to be associated with geographic variation in assemblage composition (Moyle and Cech 1988; Pflieger 1989, Osborne and Wiley 1992; Allan 1995; Seelbach et al. 1997; Matthews 1998). Each distinct combination of variable attributes represents a distinct VST. Stream size classes (i.e., headwater, creek, small river, large river, and great river) will be based on those of Pflieger (1989), which were empirically derived with multivariate analyses and prevalence indices.

I. *Preprocessing*

1. Gather and assemble the necessary GIS data layers. Necessary data layers include:
   - 1:100,000 National Hydrography Dataset (NHD) stream networks for Arkansas. <http://nhd.usgs.gov/>
   - Digital elevation model (DEM) 30 meter resolution, from the National Elevation Dataset (NED).
   - Digital Geologic Map of Arkansas (1:500,000). Version 1.0. Published in 2000. <http://pubs.usgs.gov/sm/arkansas/download>
   - STATSGO soils (1:250,000).
   - Coldwater stream segments (pending availability)
2. Code primary and secondary channels. To run stream ordering programs on the networks it becomes necessary to code and temporarily remove the secondary channels (loops and braids) from the primary channels (main flow paths).

II. *Attribute and Classify Stream Networks*

1. Assign stream order and downstream order (both Strahler and Shreve link number)
2. Assign stream size categories using ranges of Shreve link number. Size categories will consist of Headwater, Creek, Small River, Large River, and Great River.

3. Assign stream size discrepancy categories. Connectivity is an important variable influencing aquatic communities. Coding streams according to stream size discrepancy can be accomplished by finding the link number (and consequently the stream size class) of the next downstream reach for every stream reach.
4. Attribute according to dominant geologic component each stream segment flows through. Stream segments are assigned a geologic type code based on what the majority of the segment is flowing through. This approach is used to avoid having to break a stream segment into numerous small segments every time it crosses a geologic boundary.
5. Attribute each stream segment with an upstream and downstream elevation in meters from the DEM.
6. Compute stream gradient in meters per kilometer for each stream segment (also between major confluences of the larger rivers)
7. Assign relative gradient categories. Gradient is made relative with respect to both stream size class and Aquatic Subregion.
8. Create and assign stream valley segment type (VST) codes by concatenating the individual classification variables into one longer numeric code. Because each position in the code has meaning (represents a specific variable) each group of stream segments with the same code are hydrogeomorphically similar.

Task 2: Characterize the Watershed of Each Stream Segment

Characterize the physiographic condition of the watershed of all primary channel stream segments in Arkansas, but excluding the Mississippi River.

I. *Preprocessing*

1. Gather and assemble the necessary GIS data layers. Necessary data layers include:
   - 1:100,000 Valley Segment Type stream network as developed for Objective 1 above.
   - Digital elevation model (DEM) 30 meter resolution, from the National Elevation Dataset (NED).
   - Digital Geologic Map of Arkansas (1:500,000). Version 1.0. Published in 2000. <http://pubs.usgs.gov/sm/arkansas/download>
   - STATSGO soils (1:250,000).
2. Use ESRI's Arc Hydro Tools to create small stream segment specific (or reach specific) drainage area polygons for all of the primary channel streams contained in the 1:100,000 stream valley segment type coverage for Arkansas using a 30-meter resolution digital elevation model (DEM). These polygons are often referred to as simply "catchments" or "segmentsheds" (Figure 1).

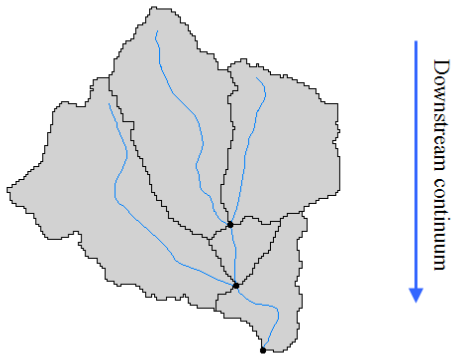

Figure 1. Example stream reach specific drainage area polygons and corresponding stream segments. These polygons are also referred to as "catchments" or "segmentsheds".

II. *Attribute Catchment Polygons and Characterize the Watersheds of Each Stream Segment*

1. Tabulate the area of each catchment polygon contained in each soil surface texture class, soils hydrologic group class, and each geologic class.
2. Attach the tabulations from step one to the stream segments via the common identifier.
3. Develop and run programs to quantify the percent of the drainage area above every stream segment contained in each soil, geologic, and gradient class from step one (Figure 2).

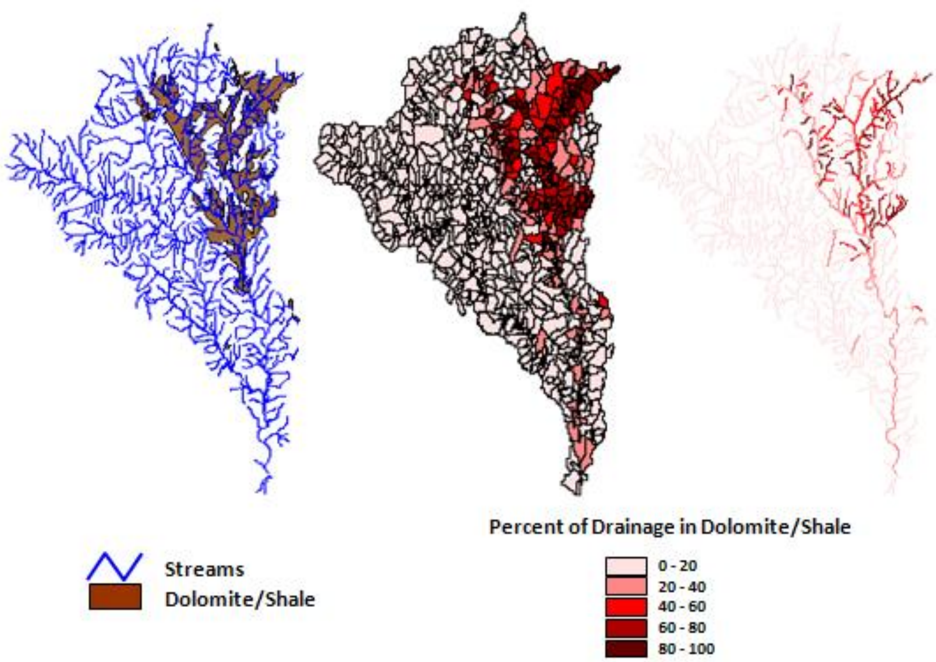

Figure 2. Map example showing results of generating overall watershed percentages for each individual stream segment. This example shows the percent of dolomite/shale geology within each stream segment's watershed.

Task 3: Compile Fish Community Collection Records

We will acquire community collection records for all fish throughout Arkansas. We will contact state and federal agencies, academic institutions and museums to locate existing fish sampling data in Nebraska. Sources will include the Arkansas Natural Heritage Inventory, other state agency collections (Arkansas Game and Fish Commission, Arkansas Department of Environmental Quality, others), and universities.

Task 4: Create a Relational Species Collections Database

All available species community collection records will be integrated and transferred into a Microsoft Access relational database developed for this project. Database fields will include species common and scientific names, ITIS codes, collection date, global and state ranks.

Task 5: Geo-reference Species Occurrence Data

All available species community collection records will be geo-referenced into a GIS database using standard GIS tools and protocols. Each collection record will be given an identifier to attach in to the stream segment from which the collection was taken.

Task 6: Generate Professionally Reviewed Species Range Maps

Based on the geo-referenced species collection information we will create digital range maps by 8-digit or 11-digit hydrologic units. These range maps will then be presented to a committee for professional review. The committee will adjust the ranges for each species by adding or removing any hydrologic units that represent data errors or anomalies (Figure 3). The resulting final species range maps will be stored in an ArcGIS database.

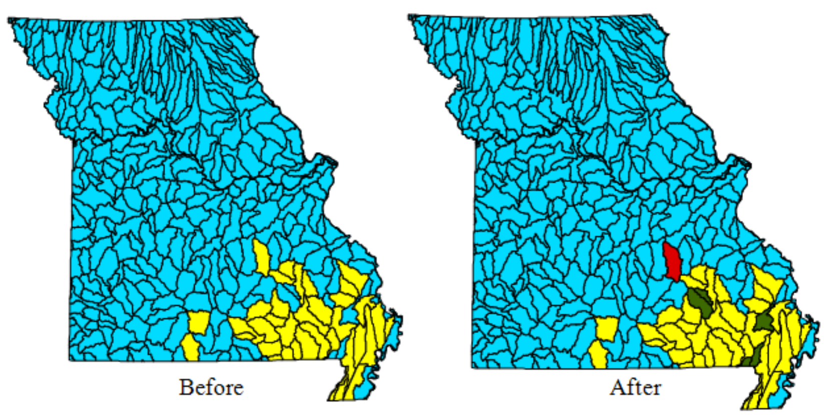

Figure 3. The map on the left shows the range of the Redspotted sunfish in Missouri based on actual collections. The map on the right shows the results of the professional review process where the hydrologic unit in red was removed from the species range because it was considered an error by the professional reviewers, while the hydrologic units in green were added to the range because the units had been undersampled (5 or fewer samples).

Task 7: Develop Predictive Distribution Models and Maps for Fish

For the modeling process we will follow the methods of Sowa et al. (2006). We will merge the species collection database with the Valley Segment data layer using the common identifier seg_id as the link between the two datasets. We will then use Decision Tree Analyses to generate predictive distribution models for each species. Nonlinear classification tree statistics will be utilized to model species affinities to particular local and watershed characteristics (habitats) within the species' range. Once an affinity is defined for a particular species, other similar stream reaches will be characterized for the likelihood the species will occur there. The resulting GIS dataset will identify stream reaches with suitable habitat for the species.

Predictor Variables

Riverine fishes are influenced by numerous landscape and in-channel factors and processes operating at multiple spatial and temporal scales. Of particular interest are those landscape factors operating within the overall watershed and immediate drainage of a particular stream segment and the local in-channel factors associated with that specific stream segment.
All of the watershed variables will be converted into a 10 category, equal interval, variable prior to modeling. Since the range of values for each watershed variable ranges from 0 to 100%, each category represents a 10% range in values. An example list of the watershed variables that were used in completing this process for the state of Nebraska is provided in Table 1. We will develop a specific list of watershed variables for Arkansas, but like Nebraska they will generally pertain to soils, geology, and landform.

Table 1. Descriptions for the 23 local and watershed predictor variables used for the Nebraska species modeling.

Local variable	**Description**
Flow	Binary variable that differentiates perennial and intermittent flow
Temp	Binary variable that differentiates cold and warm water streams
Linkr10	A ten category description of stream size based on Shreve link magnitude (Shreve 1966)
Sdiscr_2c	Binary variable that differentiates stream segments that flow into either the same size stream or larger stream
Grdseg10	A ten category designation of stream gradient (m/km)
Neb_geol	A 14 category variable designating the surficial geology through which each stream segment flows
Stxt4cat	A 4 category variable designating the general soil texture class through which each stream segment flows
Drn_grp	A 5 category variable designating the major drainage group in which a given stream segment occurs
Watershed Variable	
Avegrd10	Average gradient of all stream segments in the watershed
Hyda_p	Percent of watershed containing Hydrologic Soil Group A placed into ten categories
Hydb_p	Percent of watershed containing Hydrologic Soil Group B placed into ten categories
Hydc_p	Percent of watershed containing Hydrologic Soil Group C placed into ten categories
Hydd_p	Percent of watershed containing Hydrologic Soil Group D placed into ten categories
Hydbc_p	Percent of watershed containing Hydrologic Soil Group B/C placed into ten categories
Stxt01_p	Percent of watershed containing Soil Surface Texture Class 1 (Sand) placed into ten categories
Stxt02_p	Percent of watershed containing Soil Surface Texture Class 2 (Loamy Sand) placed into ten categories
Stxt03_p	Percent of watershed containing Soil Surface Texture Class 3 (Sandy loam) placed into ten categories
Stxt04_p	Percent of watershed containing Soil Surface Texture Class 4 (Silt loam) placed into ten categories
Stxt06_p	Percent of watershed containing Soil Surface Texture Class 6 (Loam) placed into ten categories
Stxt08_p	Percent of watershed containing Soil Surface Texture Class 8 (Silty clay loam) placed into ten categories
Stxt09_p	Percent of watershed containing Soil Surface Texture Class 9 (Clay loam) placed into ten categories
Stxt11_p	Percent of watershed containing Soil Surface Texture Class 11 (Silty clay) placed into ten categories
Stxt12_p	Percent of watershed containing Soil Surface Texture Class 12 (Clay) placed into ten categories

Statistical Methods

We will use the Classification Tree add-on of SPSS version 14.0 for modeling our species distributions. Classification tree analyses are nonlinear/nonparametric modeling techniques that typically employ a recursive-partitioning algorithm which repeatedly partitions the input data set into a nested series of mutually exclusive groups, each of which is as homogeneous as possible with respect to the response variable (Olden and Jackson 2002). The resulting tree-shape output represents sets of decisions or rules for the classification of a particular dataset. These rules can then be applied to a new unclassified dataset to predict which records or, in our case, location will have a given outcome.

The specific modeling algorithm we will use is Exhaustive CHAID, which is a modification of CHAID developed by Biggs et al. (1991). It was developed to address some weaknesses of the CHAID method. In some instances CHAID may not find the optimal split for a variable since it stops merging categories as soon as it finds that all remaining categories are statistically different (AnswerTree® User's Guide 2001). Exhaustive CHAID remedies this problem by continuing to merge categories of the predictor variable until only two "supercategories" are left and then examines the series of merges for the predictor and finds the set of categories that gives the strongest association with the target variable and computes an adjusted-*p* value for that association. Consequently, exhaustive CHAID can find the best split for each individual predictor and then choose which of these predictors to split on at each level in the tree by comparing the adjusted-*p* values.

Model Outputs

Probability of Occurrence

Each terminal node, in the classification tree models, represents a combination of watershed and local conditions and provides a corresponding probability of occurrence for a given species under that set of conditions. These probabilities can be applied to an independent dataset using a series of if/then model statements that are generated by SPSS. Once a model is completed, for a given species, we will apply the resulting if/then statement model to the attribute table of the statewide 1:100,000 modified NHD. This process produces a column for that particular species which provides the probability of occurrence for each of the stream segments in the state. However, since potential distributions are often constrained by isolation mechanisms and by biotic interactions, particularly predation (Jackson et al. 2001) and competition (Winston 1995), species rarely, if ever, occupy all suitable locations (Hutchinson 1957). Therefore, for all stream segments falling outside of the professionally-reviewed geographic range, we will convert the probabilities generated by the model to zero. Results of this process can be seen in the maps provided in Figure 4.

Individual models will be merged into two similar, but distinct DBASE (hyperdistribution) files that can be spatially related to the modified 1:100,000 NHD, in ArcView or ArcGIS via the segment id code (named: seg_id). One of these DBASE files will be in a “flat file” format that contains a single row for each stream segment in Arkansas and a column for each species, which is denoted by a species code (e.g., F###). A column for each species will provide the probability of occurrence for that species for each stream segment. This database is suited to addressing research or management questions pertaining to a single species. The other DBASE file will provide this same information in a relational database, or list, format. In this database the segment id code will repeat as many times as there are species predicted to have a greater than zero probability of occurring in that segment. There will also be columns that provide the species common name and the probability of occurrence. This database will be best suited to addressing research or management questions pertaining to multiple species since it will display a list of species and associated occurrence probabilities for each stream segment.

Presence

While the probability of occurrence databases provide a wealth of distributional information for each species, they are not particularly suited to multi-species conservation assessments like a gap analysis, which often rely on richness or diversity measures. Calculating richness or diversity measures require explicit yes or no statements about species presence, which are not provided with a continuous probability of occurrence. In many instances, modelers deem a species as being present at locations where it has a greater than 50% probability of occurrence. However, because sampling methods are often inefficient or insufficient, species that have low detection probabilities rarely have probability of occurrences that exceed 50% and would therefore never be predicted as “present” across their entire range, even within optimal habitat.

To overcome this problem we will use the “relative-50%” rule developed by Sowa et al. (2005) in order to generate a pure presence model for each species. Specifically, for each model we will first identify the terminal node having the highest occurrence percentage that also contained at least 50 collection records. Selecting the highest occurrence percentage only from those terminal nodes with 50 or more collection records will account for the fact that terminal nodes with only a handful of samples generally provide somewhat inflated or deflated occurrence percentages. We will then divide the highest occurrence percentage by 2 and select all nodes having occurrence percentages greater than or equal to this percentage. For example, if the highest occurrence percentage was 80% we would select all nodes with occurrence percentages greater than or equal to 40% and if the highest occurrence percentage was 20% we would select all nodes with occurrence percentages greater than or equal to 10%. Within the statewide probability of occurrence database, described above, we will query the appropriate probability of occurrence column to select all stream segments having a probability of occurrence greater than or equal to the specific relative 50% occurrence identified for a given species. Then, we will create a new column for that species and attribute all of the selected segments with a value of 1 to denote presence, while all other segments will be attributed with a 0. Results of this process can be seen in the maps provided in Figure 4.

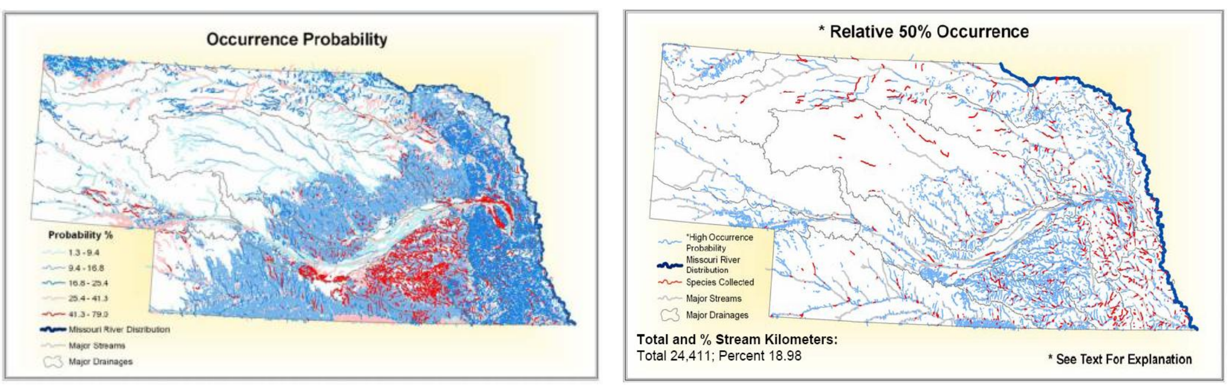

Figure 4. Resulting models for the Black Bullhead in Nebraska.

The above statistical methods will be used to generate predictive distribution models for most of the fish species. However, for some species were there may be too few occurrence records (i.e., less than 10) to develop a statistical model. In addition, species that are only found in the Mississippi River will not be modeled statistically. In these instances the models will be developed subjectively with help from a committee of professional reviewers.

Task 8: Validate Species Models

Species distribution predictions will be validated by withholding some species occurrence records will from the initial modeling. These data will be used to identify the accuracy of the models.

Task 9: Write Report

A report will be written that describes the concept, methodology and results of the project.

PRELIMINARY RESULTS AND PRIOR WORK

As part of the Missouri Aquatic GAP Pilot Project, The Missouri Resource Assessment Partnership (MoRAP) has developed a methodology for conducting gap analyses to identify and prioritize biodiversity conservation needs for stream ecosystems. MoRAP has been involved in ecological classification and Aquatic Gap Analysis for the Lower Missouri River Basin and has developed aquatic species models for Missouri (fish, mussels, and crayfish) and Nebraska (fish). For more information about some of our past work see *A Gap Analysis for Riverine Ecosystems of Missouri* (Sowa et al. 2005) and *Developing Predictive Distribution Models for Fish Species in Nebraska* (Sowa et al. 2006).

PLANNED PRODUCTS

1. GIS layer of 1:100,000 scale valley segment types covering Arkansas
2. GIS related table of fish occurrence records
3. Relational (Microsoft Access) database holding all of the fish sampling data compiled in Arkansas
4. Species range maps by 11 or 8-digit Hydrologic Unit for all fish species in Arkansas
5. Stream reach scale predicted distribution maps for all fish species in Arkansas
6. Related Dbase file consisting of watershed statistics for each stream reach pertaining to geology, soils, and relief.
7. Final Report

DISSEMINATION OF PROJECT RESULTS

All project products will be delivered to the USGS National GAP Analysis Program, the Arkansas Fish and Game Commission, Arkansas Natural Resources Commission, Arkansas Department of Environmental Quality, the University of Arkansas, United States Fish and Wildlife Service, National Park Service, and will be distributed to GeoStor, Arkansas' Official GeoData Clearinghouse. A workshop for end-users describing the methodology and appropriate uses of the data will be hosted at the end of the project.

REFERENCES CITED

Allan, J. D. 1995. *Stream ecology: structure and function of running waters*. Chapman and Hall, New York, NY.

Annis G. M., S. P. Sowa and D. D. Diamond. 2002. Detailed procedures for classifying stream valley segment types using the National Hydrography Dataset. Missouri Resource Assessment Partnership Technical Bulletin, Number 2002-01.

AnswerTree 3.0 User's Guide. 2001. SPSS Inc. Chicago, IL.

Biggs, D., B. de ville, and E. Suen. 1991. A method of choosing multiway partitions for classification and decision trees. Journal of Applied Statistics 18: 49-62.

Hutchinson, G. E. 1957. A Treatise on Limnology, vol 4. The Zoobenthos. John Wiley and Sons, New York, NY.

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.

Matthews, W. J. 1998. *Patterns in Freshwater Fish Ecology*. Chapman and Hall, New York, NY.

Moyle, P. B. and J. J. Cech, Jr. 1988. *Fishes: An Introduction to Ichthyology*, 2nd edition. Prentice-Hall, Englewood Cliffs, NJ.

Olden, J. D. and D. A. Jackson. 2002. A comparison of statistical approaches for modelling fish species distributions, Freshwater Biology 47: 1976-1995.

Osborne, L. L. and M. J. Wiley. 1992. Influence of tributary spatial position on the structure of warmwater fish communities. Canadian Journal of Fisheries and Aquatic Sciences 49: 671-681.

Pflieger, W. L. 1989. *Aquatic community classification system for Missouri*. Jefferson City, MO: Missouri Department of Conservation, Aquatic Series No. 19.

Scott, J.M., F. Davis, B. Csuti, R. Noss, B. Butterfield, C. Groves, H. Anderson, S. Caicco, F. D'Erchia, T. C. Edwards, Jr., J. Ulliman, and G. Wright. 1993. Gap analysis: A geographic approach to protection of biological diversity. Wildlife Monographs 123.

Seelbach, P.W., M.J. Wiley, J.C. Kotanchik and M.E. Baker.1997. *A Landscape- based ecological classification system for river valley segments in Lower Michigan*. Fisheries Research Report No. 2036. Michigan Department of Natural Resources, Ann Arbor, MI. 51.pp.

Shreve, R. L. 1966. Statistical law of stream numbers. Journal of Geology 74: 1737.

Sowa, S. P., G. M. Annis, M. E. Morey, and A. J. Garringer. 2006. Developing Predictive Distribution Models for Fish Species in Nebraska. Final Report, submitted to the USGS National Gap Analysis Program, Moscow, ID. 494 pp.

Sowa, S. P., D. D. Diamond, R. Abbitt, G. Annis, T. Gordon, M. E. Morey, G. R. Sorensen, and D. True. 2005. A Gap Analysis for Riverine Ecosystems of Missouri. Final Report, submitted to the USGS National Gap Analysis Program, Moscow, ID. 1675 pp.

Winston, M. R. 1995. Co-occurrence of morphological similar species of stream fishes. American Naturalist 145: 527-545.

PROJECT PERSONNEL

Scott P. Sowa, Ph.D. (PI): Scott obtained a M.S. and Ph.D. in Fisheries and Wildlife from the University of Missouri and a B.S. in Fisheries and Wildlife from Michigan State University. His areas of expertise include freshwater ecology, aquatic ecological classification, biodiversity conservation, and GIS. Scott has been working on Aquatic GAP projects for the last ten years. He served as the PI on the Missouri Aquatic GAP Project, Nebraska Fish Modeling Project, and as one of the CO-PIs on the Lower Missouri River Basin Aquatic GAP Data Development project.

Ethan Inlander, (PI): Ethan received his undergraduate and master's degrees from the Department of Geography at University of California Santa Barbara. His thesis topic was "An Integrated Methodology for the Mapping and Inventory of Riparian Areas in the Upper Santa Ynez Watershed, California". Before joining The Nature Conservancy, Ethan applied geographical information systems technology to address multiple scale conservation problems in riparian and costal habitats of California. Since joining The Nature Conservancy, Ethan has applied these same techniques to identify and reduce impacts and habitat degradation to freshwater stream ecosystems, conduct local, watershed, and regional threat assessments of subterranean environments, and prioritize and implement karst and riverine conservation actions at multiple scales.

Gust Annis, (CO-PI): Gust obtained a M.A. in Geography from Western Illinois University and a B.S. in Education as well as the post baccalaureate equivalent of a Geography major from Northern Michigan University. His areas of expertise include freshwater ecosystem management and GIS. He has served as the lead GIS specialist for the Missouri Aquatic GAP Project, Nebraska Fish Modeling Project, and the Lower Missouri River Basin Aquatic GAP Data Development project.

David D. Diamond, Ph.D.: David obtained a B.S. in Biology from Eastern Montana College and M.S. in Botany and Ph.D. in Range Science from Texas A&M University. He has served as the Director of MoRAP since its inception in 1996 and is responsible for the overall management of MoRAP staff and projects.

Michael E. Morey: Mike obtained a B.S. in Journalism from the University of Kansas and a B.A. in Computer Science from the University of Missouri-Kansas City. His areas of expertise include database management and GIS. He has served as the lead database manager for the Missouri Aquatic GAP Project, Nebraska Fish Modeling Project, and the Lower Missouri River Basin Aquatic GAP Data Development project.

Aaron Garringer: Aaron obtained a B.A. degree in Geography from the University of Missouri. He has considerable experience in compiling geospatial data on human stressors and quantifying their effects on freshwater ecosystems.

Dyanna Pursell: Dyanna obtained a B.A. degree in History and an M.A. degree in Geography from the University of Missouri. She has been working for MoRAP for the last three years and has served as a GIS specialist on a wide array of projects.

Tammy Martin: Tammy obtained a B.A. degree in in Mass Communications with an emphasis in Journalism from Truman State University. She has been the Administrative Assistant for MoRAP since 2001 and helps with the fiscal and operational procedures associated with all MoRAP projects.

Eylem Mutlu: Eylem received her B.S. as honor student in Environmental Engineering from Ondokuz Mayis University, Samsun, Turkey, her M.S. in Environmental Engineering from the University of Missouri-Rolla and her PhD. Biological & Agricultural Engineering Department from University of Arkansas-Fayetteville in 2007. Her PhD. thesis topic was "*Development* of Artificial Neural Network Models and Watershed Models for Hydrologic Prediction in an Agricultural watershed". During her PhD degree, she applied application of geographic information systems and remote sensing in developing decision support system for ecosystem management, identifying and reducing impacts and habitat degradation to freshwater stream ecosystems, conduct local, watershed, and regional threat assessments of subterranean environments.

PROJECT TIMELINE AND BUDGET

Exhibit A

Table 2 provides a task-related budget and production timeline for the proposed project. Table 3 provides an itemized budget for the proposed project.

Table 2. Task-related budget and production timeline.

Task	Time Estimate (Months from award)			Total Cost	MoRAP Yr 1 Effort	AR TNC Yr 1 Effort	MoRAP Yr 2 Effort	AR TNC Yr 2 Effort	MoRAP Yr 3 Effort	AR TNC Yr 3 Effort
	1 3 6 9 12	13 15 18 21 24	25 27 30 33 36							
Develop VST Coverage for Arkansas										
Compile NHD files	*			**$500**	$500	$0	$0	$0	$0	$0
Preprocess NHD linework	***************	***		**$43,000**	$22,000	$0	$21,000	$0	$0	$0
Attribute NHD linework		*******		**$33,000**	$0	$0	$33,000	$0	$0	$0
QA/QC VST coverage		**		**$1,500**	$0	$0	$1,500	$0	$0	$0
Generate segmentsheds and watershed statistics for each stream segment										
Generate catchments/segmentsheds		***		**$5,000**	$0	$0	$5,000	$0	$0	$0
Generate local and watershed statistics		************	*******	**$32,000**	$0	$0	$18,000	$0	$14,000	$0
Transfer statistics to VST layer for modeling			*	**$2,000**	$0	$0	$0	$0	$2,000	$0
QA/QC statistics			*	**$2,500**	$0	$0	$0	$0	$2,500	$0
Generate distribution models for fish										
Compile community sampling data	******************			**$15,000**	$3,000	$12,000	$0	$0	$0	$0
Create relational community sampling database		******************		**$27,000**	$0	$0	$5,500	$21,500	$0	$0
Geo-reference community sampling data to VST layer			*****	**$23,000**	$0	$0	$0	$0	$4,600	$18,400
Create professionally reviewed species range maps by Hydrologic Unit			****	**$15,000**	$0	$0	$0	$0	$3,000	$12,000
Model fish			*****	**$28,000**	$0	$0	$0	$0	$5,600	$22,400
QA/QC species models			**	**$5,000**	$0	$0	$0	$0	$1,000	$4,000
Final Report			***	**$16,000**	$0	$0	$0	$0	$0	$16,000
Supplies	NA	NA	NA	**$25,000**	$5,000	$5,000	$5,000	$2,500	$5,000	$2,500
Travel Costs	NA	NA	NA	**$6,500**	$500	$2,000	$1,000	$1,000	$1,000	$1,000
				$280,000	$31,000	$19,000	$90,000	$25,000	$38,700	$76,300
					Total Year 1 =	**$50,000**	**Total Year 2 =**	**$115,000**	**Total Year 3 =**	**$115,000**

MoRAP Total Project	**$159,700**
TNC Total Project	**$120,300**
Total Project Cost	**$280,000**

Table 3. Itemized budget.

Personnel	Functional Title	Annual Salary	Year 1 Effort	Year 1 Cost	Year 2 Effort	Year 2 Cost	Year 3 Effort	Year 3 Cost
Annis, Gust	GIS Lead	$ 42,715	10.00%	$ 4,272	20.00%	$ 8,543	10.00%	$ 4,272
Diamond, David	MoRAP Administrator	$ 82,411	0.00%	$ -	5.00%	$ 4,121	0.00%	$ -
Garringer, Aaron	GIS Technician	$ 33,400	5.00%	$ 1,670	30.00%	$ 10,020	15.00%	$ 5,010
Morey, Mike	Database Manager	$ 40,856	5.00%	$ 2,043	30.00%	$ 12,257	15.00%	$ 6,128
Martin, Tammy	Administrative Assistant	$ 29,803	5.868%	$ 1,749	10.561%	$ 3,147	5.282%	$ 1,574
Pursell, Dyan	GIS Technician	$ 32,427	0.00%	$ -	35.00%	$ 11,349	0.00%	$ -
Sowa, Scott	Project Coordinator	$ 57,098	10.00%	$ 5,710	10.00%	$ 5,710	5.00%	$ 2,855
Subtotal				$ 15,443		$ 55,147		$ 19,839
Benefits @ 29.15%				$ 4,502		$ 16,075		$ 5,783
Total Salary plus Benefits				**$ 19,945**		**$ 71,222**		**$ 25,622**
Travel				**$ 500**		**$ 1,000**		**$ 1,000**
Computing Supplies				**$ 5,000**		**$ 5,000**		**$ 5,000**
MoRAP Direct Costs (exclusive of Contract)				**$ 25,445**		**$ 77,222**		**$ 31,622**
Services / Contract - Arkansas Nature Conservancy				**$ 19,000**		**$ 25,000**		**$ 76,300**
Total Direct Costs				**$ 44,445**		**$ 102,222**		**$ 107,922**
Indirect @ 12.5% (applied to first $25,000 of Contract)				**$ 5,556**		**$ 12,778**		**$ 7,078**
MoRAP Total (MoRAP Direct + Indirect)				**$ 31,000**		**$ 90,000**		**$ 38,700**
Total Grant Cost by Year (MoRAP Total + Contract)				**$ 50,000**		**$ 115,000**		**$ 115,000**

Grand Total **$280,000**