gap

KEEPING COMMON SPECIES COMMON

Gap Analysis BULLETIN No. 12

A Geographic Approach to Planning for Biological Diversity

U.S. Department of the Interior
U.S. Geological Survey

2003/2004

USGS
science for a changing world

CONTENTS

A Comprehensive Biological Inventory Database for the Iowa Aquatic GAP Project

ANNA LOAN-WILSEY[1], ROBIN L. MCNEELY[2], PATRICK D. BROWN[2], KEVIN L. KANE[2], AND CLAY L. PIERCE[3]

[1]Department of Natural Resource Ecology and Management, Iowa State University, Ames
[2]Geographic Information Systems Support and Research Facility, Iowa State University, Ames
[3]USGS, Iowa Cooperative Fish and Wildlife Research Unit, Iowa State University, Ames

Before the implementation of the Iowa Aquatic Gap Analysis, project coordinators had no sense of the breadth of biological sampling data available for fish. However, it was considered important to have the most extensive biological data set possible. We were able to systematically compile a fish inventory database that we believe satisfies this objective. Other Aquatic GAP projects may find themselves in a similar situation and thus benefit from our approach to compiling a comprehensive biological inventory database.

Database Design

Before compiling any data set, it is essential to determine what types of information are to be included. First, we modified the Microsoft Access relational database originally designed by the Missouri Aquatic GAP Project by expanding it to reflect the additional information we wished to capture for Iowa, including additional tables for source, collector, collector samples, gear type, and negative data (where taxa were sampled for and not found). Elaborating on the original source field found in the samples table, the new collector tables included fields for collectors' names and associated samples, whereas the source table included the name of the associated institution, the citation or description of the source, and location of the original data. Unlike the sampled species table, which indicates the presence of a species in a sample, the new table for negative data indicated the absence of a species in a sample when an explicit search for that species had been made. In addition to adding tables, we expanded the number of fields in preexisting tables. Additional fields include (a) information about abundance, (b) sample type (community versus target), (c) descriptive location details, (d) descriptive method details, (e) individual specimen details, (f) a flag field for records not used in the professionally reviewed copy of the database, (g) a flag field to indicate that the sample has a corresponding feature in a GIS shapefile, and (h) a field for the Index of Biological Integrity (a widely used index of stream health).

Data Acquisition

Once the database was designed, the next step was to acquire the raw data. We first compiled a detailed list of all possible and known sources of data including historic and recent, print and electronic, and published and unpublished sources. We then compiled a detailed list of possible data acquisition strategies. We proceeded to match appropriate strategies with possible sources and pursued those sources. For example, museum collections are a possible source for historic data. Possible strategies for retrieving museum records could be to search their on-line database and/or contact individual museum curators. We identified possible museums, both public and private institutions, at the local, state, or national level. After performing a comprehensive Internet search to identify all museums that might have fish collections, we either searched their on-line database for Iowa records or contacted the curator.

Through this process we identified seven categories of source data:

- Published literature: monographs, theses, dissertations, and journal articles
- Federal reports: EPA, U.S. FWS, Army Corps of Engineers
- Museum collections
- Iowa Department of Natural Resources (IDNR) reports
- IDNR field notes
- Statewide biological inventory databases
- Individual researchers' unpublished field notes

We grouped all data acquisition strategies into four categories: literature searches, IDNR field trips, museum collection inquiries, and individual contacts. Although searching Internet access databases, such as FishBase (Froese and Pauly 2003), as a strategy was initially pursued, we discovered little Iowa community data that was not already available in primary sources.

Literature Searches

To compile fish data from published literature, we conducted literature searches using several different methods. We used bibliographies of known published sources of data or from appropriate secondary sources in order to trace back to historically published data in the same way one would use a citation index. This was useful for including journal articles and published reports that are not indexed elsewhere. For both historic and recent journal articles, we searched both print and electronic forms of subject indexes and abstracts. To ensure that the searches were comprehensive, Boolean keyword searching, field-limited searches, as well as controlled vocabulary were used. To find published reports, monographs, theses, and dissertations, we searched library catalogs at the state and national level as well as the WorldCat database, an on-line union catalog of 23,000 libraries in 63 countries. Thirty-three sources were found through this strategy.

Iowa Department of Natural Resources Field Trips

No centralized depository for stream fish community data existed in Iowa before this project. We gathered fish sampling data during visits to all 15 IDNR regional fisheries stations as well as the headquarters. During these station visits, we met with IDNR fisheries biologists and technicians to explain and promote the Aquatic GAP project. We also acquired all of the riverine fish data located at each station. Almost half of all sources used for the database were obtained during these visits, including management and research reports not available elsewhere. As an example, over 1,700 fish community samples from 1941 to 2003 were obtained just from field notes stored in filing cabinets.

Museum Collections

During early explorations of Internet sources, we discovered the most useful source of such data came from museum collection's on-line databases. After eliminating museum databases that did not include fish collections, we conducted searches on each database for Iowa-specific records. However, we also came across museum fish collections that were not available electronically. For those museums, we acquired Iowa-specific records by contacting the curator directly through e-mail. We identified over 40 museums with Iowa fish collection records. For the purposes of the Iowa Aquatic GAP Project, we were able to use the records of nine museum collections totaling 261 historic fish community samples ranging in date from 1854 to 2000.

Individual Contacts

Through an extensive network of cooperators, both at Iowa State University and the IDNR, we were directed to individuals who had collected fish community samples in Iowa. We contacted most of these individuals by e-mail. Individuals contacted ranged from retired faculty of liberal arts colleges in Iowa to out-of-state fisheries biologists who had visited the state only once. The majority of the resulting data was in the form of unpublished, hand-written field notes ranging from 1932-2000. The data uncovered in this fashion were extensive, resulting in over 2,400 fish community samples covering all geographical regions of the state.

Data Organization

For verification purposes, it is important to ensure a direct relationship back to the original data. Therefore, we also organized the raw data for easy retrieval. As we had a tremendous amount of print material, we labeled each print sample with its unique sample identifier and each print source with its unique source identifier. These materials were categorized and their locations indicated in the database using a field in the source table, e.g., "File Folder: Reports, Government- Mississippi River" or "Dissertation: contact ISU Parks Library Call No. SH156wa." For electronic data, we made use of the cross-reference tables designed by the Missouri Aquatic GAP Project, which essentially provided the same ability to go from the biological inventory database back to a specific source or sample. We also used the source table field in the database to indicate the name and location of each electronic source file, e.g., c:\\...\Manchester\2004_season.xls. This level of organizational detail aids in the data entry and error checking process and makes it easier to access the data for future use.

Database Summary

This database is available on the Internet at http://maps.gis.iastate.edu/iris/. It contains 11,683 fish community samples taken from 1884-2003. It contains 98,206 sampled species records including 142 native and 13 exotic species. It has samples from every county, every 8-digit, and almost every 10-digit hydrological unit in Iowa (see Table 1).

Table 1. Iowa Aquatic GAP database summary

Number of fish community samples	11,683
Number of species occurrences	98,206
Number of fish species sampled	142 native, 13 exotic
Sampling date range	1884-2003
Number of individual sources of data	202
Number of Iowa counties sampled (99 total)	99
Number of unique stream reaches sampled	3224
Percent of all 8-digit HUCs sampled	100
Percent of all 10-digit HUCs sampled	92.4
Percent of all 12-digit HUCs sampled	73.5

Literature Cited

Froese, R., and D. Pauly, editors. 2003. FishBase. World Wide Web electronic publication. URL: www.fishbase.org, version 10. March 2004.

Surveys to Evaluate Fish Distribution Models for the Upper Missouri River Basin Aquatic GAP Project

Steve E. Freeling[1], Charles R. Berry, Jr.[2], Ryan M. Sylvester[1], Steven S. Wall[1], and Jonathan A. Jenks[1]
[1]Department of Wildlife and Fisheries Sciences, South Dakota State University, Brookings
[2]U.S. Geological Survey, South Dakota Cooperative Fish and Wildlife Research Unit, South Dakota State University, Brookings

Introduction

For terrestrial vertebrates, the Gap Analysis Program has generated what Scott et al. (1993) called “the necessary ingredients for anticipation of endangerment of species with the ultimate goal of predicting areas of high biodiversity.” The necessary ingredients include maps of land cover, terrestrial vertebrate distributions, and land stewardship. With the aquatic component of Gap Analysis, analyses are done within watershed boundaries using valley segments as the finest resolution (Wall et al. 2004). We report here on surveys used to evaluate fish species distribution models for the aquatic GAP project of the huge Missouri River Basin.

The longest river in North America, the Missouri flows through the northern Great Plains for 3,768 km to its confluence with the Mississippi River. The river has been greatly altered in the past century for flood protection, navigation, irrigation, and power production. Twenty-five families, containing 136 species, compose its ichthyofauna. Populations of 24 species are known to be declining. Eleven fishes are listed as imperiled by two or more of the seven main-stem states (Galat et al. 2004). Plans for conserving these species and areas of high species diversity might be assisted by the Gap Analysis data provided by our project.

The Missouri River Gap Analysis Project is a partnership between South Dakota State University, working in the upper basin, the Missouri Resource Assessment Program, working in the lower basin, and the U.S. Geological Survey. Models are being developed to predict the distribution of fish species. Our purpose is to report on the initial fieldwork done in the upper basin to test the accuracy of the fish distribution models.

Site Selection

One watershed was selected from each U.S. state and one Canadian province in the upper Missouri River basin. We met with each state and provincial game and fish agency to inform them about the GAP program and select watersheds for sampling. We tended to choose watersheds that lacked fish community data and were the right size for our planned effort. The selected watersheds (Figure 1) were the Beaver River (North Dakota), Elm River (North Dakota and South Dakota), Frenchman River (Saskatchewan and Montana), Nowood River (Wyoming), and Sweetgrass River (Montana).

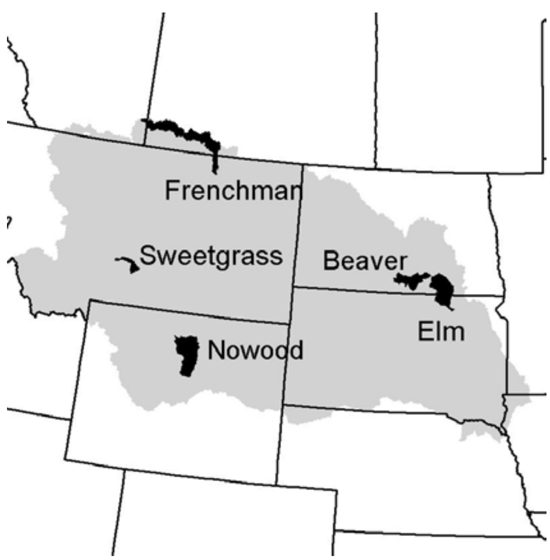


Figure 1. Five watersheds sampled (black polygons) in the Upper Missouri River Basin (gray polygon).

Streams in each watershed were stratified into three stream types: headwaters, creeks, and small rivers, which were determined from the shreve order (Shreve 1967). Shreve orders were < 9 for headwaters, 10-75 for creeks, and 76-1500 for small rivers (Wall et al. 2004). Eight general sites were selected for each stream type in each watershed, with the goal of sampling six reaches in each of the three stream types. A reach was a stream segment approximately 39 times the mean stream width (Patton et al. 2000) (50 to 200 m) and included at least one riffle, pool, and run. Our goal of sampling 18 reaches was not met in three watersheds (Table 1) for a variety of reasons, including few tributaries to choose from, lack of access permission, and lack of flow.

Table 1. Number of sample reaches for each stream type (headwater, creek, and small river) for five selected watersheds in the Upper Missouri River Basin.

Watershed	Headwater	Creek	Small River
Nowood	6	6	6
Frenchman	5	7	6
Sweetgrass	1	3	5
Beaver		1	6
Elm			6

Fish Sampling

Fish were collected with a battery-powered electrofisher (Smith-Root model LR-24) and a bag seine (9.1 m x 1.2 m with 5-mm delta mesh). The electrofisher was inefficient in wide streams, in deep pools, or in turbid water, thus seining was also used. Sampling started at the downstream end of a reach and progressed upstream in a zigzag pattern; no block nets were used (Simonson and Lyons 1995). Fish were held in 11.4-L plastic pails before being identified to species, counted, and released at the downstream end of the reach. Seining was done after electrofishing was completed.

Ancillary studies were planned to augment the basic project to access GAP fish distribution models. Habitat measurements included water chemistry, channel morphology, and riparian vegetation. These data may be valuable in future fisheries studies. Macroinvertebrates were collected with 15 sweeps from a D-frame dip net and three sediment core samples. These data are some of the first collected in these watersheds. White sucker were collected to determine population metrics. White suckers were chosen because of their occurrence in all watersheds, thus leading to the possibility of analysis of growth over a large spatial scale.

Results

A total of 41 species were identified among the 19,556 fish collected in the five watersheds. Fathead minnow and white sucker were most abundant (50% of all fish sampled) and were found in all five watersheds (Table 2). The Beaver River watershed had the highest species richness (21 species) and contained six species not found elsewhere (i.e., emerald shiner, red shiner, white bass, spottail shiner, yellow perch, and gizzard shad). The first three species inhabit open channels of large, permanently flowing rivers with low gradient (Pflieger 1997). The high species richness and presence of unique species probably occurred because this was the only direct tributary to the Missouri River. Northern redbelly dace, a cool water species (Brown 1971), were recorded for the first time in ten years in the Beaver watershed. The Elm River contained 17 species and had the greatest number of fish per site (971). Three lentic predatory fish (bluegill, black crappie, and largemouth bass) were only found in the Elm watershed. The presence of these lentic species may be due to stocking in the numerous impoundments in the watershed and also may have localized effects on riverine species richness and abundance.

Table 2. Fish species richness and abundance in five watersheds of the Upper Missouri River Basin.

Species	Beaver	Elm	Frenchman	Nowood	Sweetgrass
Red shiner *Cyprinella lutrensis*	111				
Emerald shiner *Notropis atherinoides*	84				
Yellow perch *Perca flavescens*	6				
White bass *Morone chrysops*	2				
Spottail shiner *Notropis hudsonius*	2				
Gizzard shad *Dorosoma cepedianum*	1				
Black bullhead *Ameiurus melas*	888	241			
Sand shiner *Notropis ludibundus*	198	813			
Orangespotted sunfish *Lepomis humilis*	66	667			
Johnny darter *Etheostoma nigrum*	37	120			
Channel catfish *Ictalurus punctatus*	5	11			
Bluegill *Lepomis macrochirus*		2			
Bigmouth buffalo *Ictiobus cyprinellus*		3			
Black crappie *Pomoxis nigromaculatus*		5			
River carpsucker *Carpiodes carpio*		8			
Green sunfish *Lepomis cyanellus*		25			
Largemouth bass *Micropterus salmoides*		26			
Creek chub *Semotilus atromaculatus*		48			
Northern pike *Esox lucius*	12		3		
Brassy minnow *Hybognathus hankinsoni*	11		27		
Northern redbelly dace *Phoxinus eos*	12		53		
Iowa darter *Etheostoma exile*	58	15	38		
Walleye *Stizostedion vitreum*		1	1		
Common carp *Cyprinus carpio*	330	157	11	59	
Fathead minnow *Pimephales promelas*	328	3657	4314	29	15
White sucker *Catostomus commersoni*	83	28	592	201	92
Brook stickleback *Culaea inconstans*	33		254		10

Stonecat *Noturus flavus*	19		3	19	13
Shorthead redhorse *Moxostoma macrolepidotum*	10		6	5	2
Plains minnow *Hybognathus placitus*			8		
Pearl dace *Margariscus margarita*			333		
Hybognathus *spp.*			1887		
Flathead chub *Platygobio gracilis*			186	19	
Mountain sucker *Catostomus platyrhynchus*			30	74	103
Brook trout *Salvelinus fontinalis*			67	53	85
Lake chub *Couesius plumbeus*			605	13	113
Longnose dace *Rhinichthys cataractae*			681	681	276
Rainbow trout *Oncorhynchus mykiss*				60	
Longnose sucker *Catostomus catostomus*				28	3
Brown trout *Salmo trutta*				336	30
Mountain whitefish *Prosopium williamsoni*					4
Mottled sculpin *Cottus bairdi*					11
Totals	2296	5827	9099	1577	757

The Frenchman River watershed contained two unique species (plains minnow and pearl dace). This is the first documented occurrence of the plains minnow in Canada. The Nowood River and Sweetgrass River watersheds had very similar species assemblages with 11 of the same species, mainly trout and sucker species (Table 2); this is because both are cold-water, mountainous watersheds. Fish data and habitat measurements have been provided to the state or provincial agencies for use in their future management decisions and reports.

Future Plans

Fish habitat models are being developed using Chi-squared Automatic Interaction Detector (CHAID; SPSS 2001), which is a decision tree derived from algorithms. The Lower Missouri River Aquatic GAP project team is using similar methods. Accuracy will be assessed using data splitting, jackknifing, resubstitution, and an independent data set (Fielding and Bell 1997). Cohen's Kappa will be used to assess chance corrected accuracy of the model (Titus et al. 1984). Macroinvertebrate assemblage will be used to determine if relationships exist between fish presence and macroinvertebrate presence (Lammert and Allan 1999). Macroinvertebrate assemblage may also be used in the discussion of commission and omission errors in the model.

In summary, the Missouri River Aquatic GAP Project is on schedule. The fieldwork has added information to fisheries databases managed by states and the province of Saskatchewan. Our experiences will be useful for Aquatic GAP projects that follow.

Literature Cited

Brown, C.D. 1971. Fishes of Montana. Big Sky Books, Bozeman, Montana. 207 pp.

Fielding, A.H., and J.F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24:38-49.

Galat, D., C. Berry, W. Gardner, J. Hendrickson, G. Mestl, G. Power, C. Stone, and M. Winston. 2004. Spatiotemporal patterns and changes in Missouri River fishes. In J. Rine, R. Hughes, and R. Calamusso, editors. Historical changes in fish assemblages of large American Rivers. American Fisheries Society Symposium.

Lammert, M., and J.D. Allan. 1999. Assessing biotic integrity of streams: Effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. *Environmental Management* 23:257-270.

Patton, T.M., W.A. Hubert, F.J. Rahel, and K.G. Gerow. 2000. Effort needed to estimate species richness in small streams on the Great Plains in Wyoming. *North American Journal of Fisheries Management* 20:394-398.

Pflieger, W.L. 1997. The Fishes of Missouri. Conservation Commission of the State of Missouri, Jefferson City, Missouri. 372 pp.

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 R.G. Wright. 1993. Gap Analysis: A geographic approach to protection of biological diversity. *Wildlife Monographs* 123:1-41.

Shreve, R.L. 1967. Infinite topologically random channel networks. *Journal of Geology* 75:178-186.

Simonson, T.D., and J. Lyons. 1995. Comparison of catch per effort and removal procedures for sampling stream fish assemblages. *North American Journal of Fisheries Management* 15:419-427.

SPSS Inc. 2001. AnswerTree 3.0 User's Guide. SPSS Inc., Chicago, Illinois. 226 pp.

Titus, K., J.A. Mosher, and B.K. Williams. 1984. Chance-corrected classification for use in discriminant analysis: Ecological applications. *The American Midland Naturalist* 111:1-7.

Wall, S.S., C.R. Berry, Jr., C.M. Blausey, J.A. Jenks, and C.J. Kopplin. 2004. Fish-habitat modeling for gap analysis to conserve the endangered Topeka shiner (*Notropis topeka*). *Canadian Journal of Fisheries and Aquatic Sciences* 61:954-973.

An Overview of the Data Developed for the Missouri Aquatic GAP Project and an Example of How it Is Being Used for Conservation Planning

Scott P. Sowa[1], Gust M. Annis[1], David D. Diamond[1], Dennis Figg[2], Michael E. Morey[1], and Timothy Nigh[2]

[1]Missouri Resource Assessment Partnership, University of Missouri, Columbia

[2]Missouri Department of Conservation, Jefferson City, Missouri

At the beginning of the Missouri Aquatic GAP Project, my coworkers and I at the Missouri Resources Assessment Partnership (MoRAP) expected that conservation gaps would be the norm and not the exception. Consequently, from the start we focused on compiling and producing data that would assist planners and managers with developing conservation plans for filling those gaps. These ambitions have recently become a reality when the Missouri Department of Conservation began using our data as the core decision support system for developing a statewide conservation plan for conserving freshwater biodiversity.

Before discussing the specific data we compiled or developed for the Missouri Aquatic GAP Project, we believe it necessary to provide an overview of conservation planning. This overview will provide a general context that will more clearly illustrate why we developed each geospatial data layer. Margules and Pressey (2000) and Groves (2003) both provide excellent overviews of conservation planning, and we essentially cover the most basic elements discussed by these authors in our review of the topic.

The first step in conservation planning is to establish a goal expressing the focus of the effort. This should not be confused with the quantitative conservation goals that are established when devising a specific conservation strategy (see below). Goals pertaining to biodiversity conservation have been variously described, but all have in common the conservation and restoration of the processes that generate or sustain biodiversity.

Once a goal has been established, the fundamental principles, theories, and assumptions that must be considered in order to achieve this goal must be identified. These generally pertain to basic ecological or conservation principles and theories that will be used to guide the development of a conservation strategy for achieving the overall goal.

Because conservation planning is a geographical exercise, the next step in the process involves selecting a suitable geographic framework. More specifically, this involves selecting, defining, and mapping *planning regions* and *assessment units*. A planning region refers to the area for which the conservation plan will be developed. It defines the spatial extent of the planning effort(s). Assessment units are geographic subunits of the planning region. These units define the spatial grain of analysis and represent those units among which relative quantitative or qualitative comparisons will be made in order to select specific geographic locations as priorities for conservation. Planning regions and assessment units can be variously defined and should be hierarchical in nature to allow for multiscale assessment and planning (Wiens 1989). Boundaries could be based on sociopolitical boundaries (e.g., nations, states, counties, townships), regular grids (e.g., UTM zones or EPA EMAP hexagons), or ecologically defined units (e.g., watersheds or ecoregions). Since biodiversity does not follow sociopolitical boundaries or regular grids, whenever possible planning regions and assessment units should be based on ecologically defined boundaries, since these boundaries provide a more informative ecological context (Bailey 1995, Omernik 1995, Leslie et al. 1996, Higgins 2003).

Next, because it is impossible to directly measure or map biodiversity, surrogate targets for conservation must be identified and mapped (Margules and Pressey 2000, Noss 2004). For the terrestrial component of GAP these surrogates generally include plant communities or vegetation types and vertebrate species (Scott et al. 1991). The assumption here is that by taking measures to conserve these surrogates we are in fact taking measures to also conserve those unmapped or unmappable elements of biodiversity. Because different targets often lead to different answers on which locations should be a priority for conservation, it is generally more effective to use a variety of targets (Kirkpatrick and Brown 1994, Noss 2004, Diamond et al. *in press*). Also, because biological survey data are often incomplete, biased, or completely lacking, abiotic targets (e.g., ecosystems, landscapes, or habitats), which are usually easier to map, are often considered as targets (Belbin 1993, Nicholls et al. 1998, Noss et al. 2002, Noss 2004). Angermeier and Schlosser (1995) and Noss (2004) provide excellent discussions on the reasons for using both biotic and abiotic surrogates. Also, a study by Kirkpatrick and Brown (1994) revealed that using both biotic and abiotic targets would likely be the most successful approach to representing the range of biodiversity within a planning region.

Once planning regions, assessment units, and conservation targets have been identified and mapped, an overall conservation strategy for selecting priority areas within the planning region must be established. Unfortunately, there are no detailed guidelines, and even when there is some guidance (e.g., biogeography theory, population viability analysis, or metapopulation theory) the data needed for these more detailed evaluations are usually lacking (Margules and Pressey 2000, Groves 2003). Expert opinion will therefore often play a major role in developing the overall conservation strategy.

In addition to establishing a general conservation strategy, quantitative and/or qualitative assessment criteria that will be used to make relative comparisons among assessment units must also be established. These criteria include measures of relative significance or irreplaceability, condition, future threats, costs, and opportunities, which guide the selection of one particular assessment unit over another (Groves 2003). These criteria should also be based upon the previously established fundamental principles, theories, and assumptions.

Examples include

Significance/irreplaceability: species richness, number or percent of endemic species, diversity of habitats, presence of unique habitats, species, communities, or processes

Condition: percent urban or agriculture, road density, degree of fragmentation, extent of channelization, degree of hydrologic modification, mine density, etc.

Future threat: recent or projected population trends, potential for future extractive uses

Costs: acquisition cost, restoration cost, loss of socioeconomic benefits

Opportunities: leveraging of funds or cooperation among stakeholders, local interest or involvement, ability to receive federal, state, or local funding

After addressing the issues discussed above, the next step involves selecting priority locations within the planning region(s).

Since conservation planning is a geographical exercise, it is no surprise that Geographical Information Systems (GIS) are an invaluable tool. However, because not all of the essential data are in a geospatial format, and because much of the available data often lack the necessary detail, expert knowledge must often be incorporated into the planning process. The GIS data provide a more objective, spatially explicit, and comprehensive view of the planning region, while the experts may provide additional and more detailed information for certain locations.

Conservation planning is also a logistical exercise, and once priority areas have been identified, much work remains to be done. Many questions have to be addressed, such as: Who owns the land within and around each priority area? What are the critical structural features, functional processes, and species or communities of concern within each priority area? How are we going to eliminate or minimize threats? When should conservation actions be taken, immediately or is there time? Why was each priority area selected, and why is one more "important" than another? Answering these questions is often more difficult than building the geospatial data sets and associated tools used to select priority areas. However, not addressing these important questions could lead to failure in our efforts to conserve biodiversity (Margules and Pressey 2000). Once these logistical questions have been answered, then on-the-ground conservation actions can be taken. Monitoring programs must also be established to ensure that conservation efforts are successful and to signal when and possibly how management actions should be modified. Because of the complexity and dynamic nature of ecosystems, adaptive management will be key to long-term conservation of biodiversity (Leslie et al. 1996).

So, what does this abbreviated overview of conservation planning have to do with the Missouri Aquatic GAP Project? Well, in order to adequately assess gaps in biodiversity conservation we must first identify what constitutes a gap and the only way to do this is to develop criteria for what constitutes "effective" conservation. These very criteria are established in the conservation planning process. Building on the solid foundation of the terrestrial component of GAP and going through the above process were the two most influential factors that guided the decisions we faced about the data to be compiled or developed as well as the overall approach to the Missouri Aquatic GAP Project.

The Data

The following overview of the geospatial data developed for the Missouri Aquatic GAP Project explains why and how these data were developed as a precursor to the conservation planning case study that comes later. The process for data development has four steps that are described in detail in the following sections:

1. Classify and map relatively distinct riverine ecosystems at multiple spatial scales.
2. Develop predictive distribution maps for each of the fish, mussel, and crayfish species of Missouri.
3. Develop local, watershed, and upstream riparian stewardship statistics for each stream segment within Missouri.
4. Develop or assemble geospatial data on anthropogenic threats or stressors necessary to quantitatively or qualitatively account for the current conservation status of each ecosystem unit.

Step 1: Classifying riverine ecosystems

Purpose:

- Provide the ecological and evolutionary context necessary for making truly relative comparisons among two or more locations.
- Provide an ecologically meaningful geographic framework for conservation planning (i.e., planning regions and assessment units).
- Provide surrogate abiotic conservation targets to complement biotic targets.
- Account for broader ecosystem or evolutionary processes that are often not considered with the use of species data alone.
- Account for poorly known or unknown ecosystem processes, aquatic assemblages, and organisms.
- Provide a geographic template and predictor variables for developing predictive species distribution models and maps.
- Provide the necessary reductionist tool for generating inventory statistics, conducting conservation assessments, and developing conservation plans.
- Enhance our understanding of the number and spatial distribution of distinct ecosystem types and riverine assemblages.
- Enhance communication among resource professionals, legislators, and the public.

It is widely accepted that to conserve biodiversity we must conserve ecosystems (Franklin 1993, Grumbine 1994). It is also widely accepted that ecosystems can be defined at multiple spatial scales (Noss 1990, Orians 1993). Consequently, a key objective was to define and map distinct riverine ecosystems (often termed ecological units) at multiple levels. Yet, before distinct riverine ecosystems could be classified and mapped, the question "What factors make an ecosystem distinct?" needed to be answered. Ecosystems can be distinct with regard to their structure, function, or composition (Noss 1990). Structural features in riverine ecosystems include factors such as depth, velocity, substrate, or the presence and relative abundance of habitat types. Functional properties include factors such as flow regime, thermal regime, sediment budgets, energy sources, and energy budgets. Composition can refer to either abiotic (e.g., habitat types) or biotic factors (e.g., species). While both are important, our focus here will be on biological composition, which can be further subdivided into ecological composition (e.g., physiological tolerances, reproductive strategies, foraging strategies, etc.) or taxonomic composition (e.g., distinct species or phylogenies) (Angermeier and Schlosser 1995). Geographic variation in ecological composition is generally closely associated with geographic variation in ecosystem structure and function. For instance, fish species found in streams draining the Central Plains of northern Missouri generally have higher physiological tolerances for low dissolved oxygen and high temperatures than species restricted to the Ozarks, which corresponds to the prevalence of such conditions within the Central Plains (Pflieger 1971, Matthews 1987, Smale and Rabeni 1995a, 1995b). Differences in taxonomic composition, not related to differences in ecological composition, are typically the result of differences in evolutionary history between locations (Mayr 1963). For instance, differences among biological assemblages are found on islands despite the physiographic similarity of the islands.

Considering the above, a more specific objective was to identify and map riverine ecosystems that are relatively distinct with regard to ecosystem structure, function, and evolutionary history (i.e., biological composition) at multiple levels. To accomplish this, an eight-level classification hierarchy was developed in conjunction with The Nature Conservancy's Freshwater Initiative (Higgins 2003) (Figure 1). These eight geographically dependent and hierarchically nested levels (described next) were either empirically delineated using biological data or delineated in a top-down fashion using landscape and stream features (e.g., drainage boundaries, geology, soils, landform, stream size, gradient, etc.). These features have consistently been shown to be associated with or ultimately control structural, functional, and compositional variation in riverine ecosystems (Hynes 1975, Dunne and Leopold 1978, Matthews 1998). More specifically, levels 1-3 and 5 account for geographic variation in *taxonomic or genetic-level composition* resulting from distinct evolutionary histories, while levels 4 and 6-8 account for geographic variation in ecosystem structure, function, and *ecological composition* of riverine assemblages. The most succinct way to think about the hierarchy is that it represents a merger between the different approaches taken by biogeographers and physical scientists for tesselating the landscape into distinct geographic units.

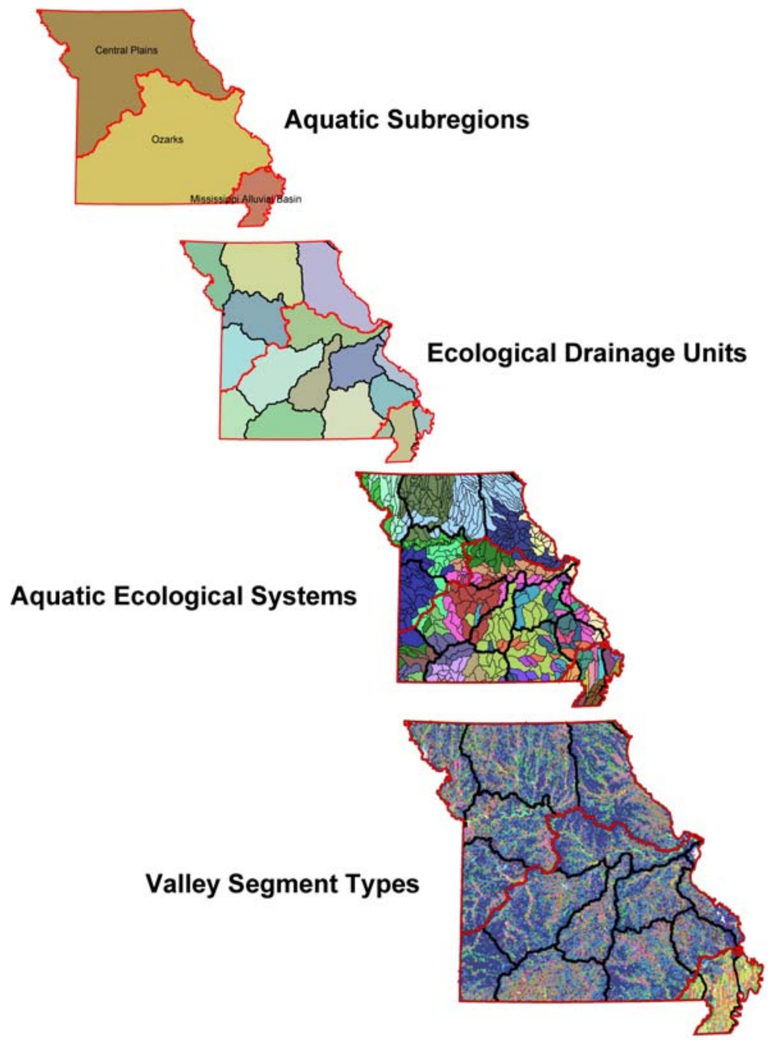


Figure 1. Maps of Missouri showing four of the eight levels of the MoRAP aquatic ecological classification hierarchy. Maps of the upper three levels (Zone, Subzone, and Region) of the hierarchy are provided in Maxwell et al. (1995). Level 8 of the hierarchy is also not shown since the distinct units within this level (e.g., riffles, pools, glides) cannot be mapped within a GIS at a scale of 1:100,000.

Levels 1 – 3: Zone, Subzone, and Region

The upper three levels of the hierarchy are largely zoogeographic strata representing geographic variation in taxonomic (family- and species-level) composition of aquatic assemblages across the landscape resulting from distinct evolutionary histories (e.g., Pacific versus Atlantic drainages). For these three levels we adopted the ecological units delineated by

Maxwell et al. (1995) who used existing literature and data, expert opinion, and maps of North American aquatic zoogeography (primarily broad family-level patterns for fish and also unique aquatic communities) to delineate each of the geographic units in their hierarchy. More recent quantitative analyses of family-level faunal similarities for fishes conducted by Matthews (1998) provide additional empirical support for the upper levels of the Maxwell et al. (1995) hierarchy. The ecological context provided by these first three levels may seem of little value; however, such global or subcontinental perspectives are critically important for research and conservation (see pp. 261-262 in Matthews 1998). For instance, the physiographic similarities along the boundary of the Mississippi and Atlantic drainages often produce ecologically similar (i.e., functional composition) riverine assemblages within the smaller streams draining either side of this boundary, as Angermeier and Winston (1998) and Angermeier et al. (2000) found in Virginia. However, from a species composition or phylogenetic standpoint, these ecologically similar assemblages are quite different as a result of their distinct evolutionary histories (Angermeier and Winston 1998, Angermeier et al. 2000). Such information is especially important for those states that straddle these two drainages, such as Georgia, Maryland, New York, North Carolina, Pennsylvania, Tennessee, Virginia, and West Virginia, since simple richness or diversity measures not placed within this broad ecological context would fail to identify, separate, and thus conserve distinctive components of biodiversity. The importance of this broader context also holds for those states that straddle the continental divide or any of the major drainage systems of the United States (e.g., Mississippi Drainage vs. Great Lakes or Rio Grande Drainage).

Level 4: Aquatic Subregions

Aquatic Subregions are physiographic or ecoregional substrata of Regions and thus account for differences in the ecological composition of riverine assemblages resulting from geographic variation in ecosystem structure and function. However, the boundaries between Subregions follow major drainage divides to account for drainage-specific evolutionary histories in subsequent levels of the hierarchy. The three Aquatic Subregions that cover Missouri (i.e., Central Plains, Ozarks, and Mississippi Alluvial Basin) largely correspond to the three major aquatic faunal regions of Missouri described by Pflieger (1989). Pflieger (1989) used a species distributional limit analysis and multivariate analyses of fish community data to empirically define these three major faunal regions. Subsequent studies examining macroinvertebrate assemblages have provided additional empirical evidence that these Subregions are necessary strata to account for biophysical variation in Missouri's riverine ecosystems (Pflieger 1996, Rabeni et al. 1997, Rabeni and Doisy 2000). Each Subregion contains streams with relatively distinct structural features, functional processes, and aquatic assemblages in terms of both taxonomic and ecological composition.

Level 5: Ecological Drainage Units

Level 5 of the hierarchy, Ecological Drainage Units (EDUs), accounts for differences in taxonomic composition (Figure 2). An initial set of EDUs was empirically defined by grouping USGS 8-digit hydrologic units (HUs) with relatively similar fish assemblages, based on the results of multivariate analyses of fish community data (Nonmetric Multidimensional Scaling, Principal Components Analysis, and Cluster Analysis). We then used collection records for three other taxa (crayfish, mussels, and snails) to further examine faunal similarities among the major drainages within each Subregion and refined the boundaries of this draft set of EDUs when necessary. Spatial biases and other problems with the data prohibited including these taxa in the multivariate analyses. In only one instance were the draft boundaries altered. Within the Ozark Aquatic Subregion the subdrainages of the Osage and Gasconade basins consistently grouped together using the methods described above. However, a more general assessment using Jacaard similarity coefficients suggested the need to separate these two drainages. Using just fish community data, the Jacaard similarity coefficient among these two drainages is 86, while when using combined data for crayfish, mussels, and snails the similarity coefficient drops to only 56.

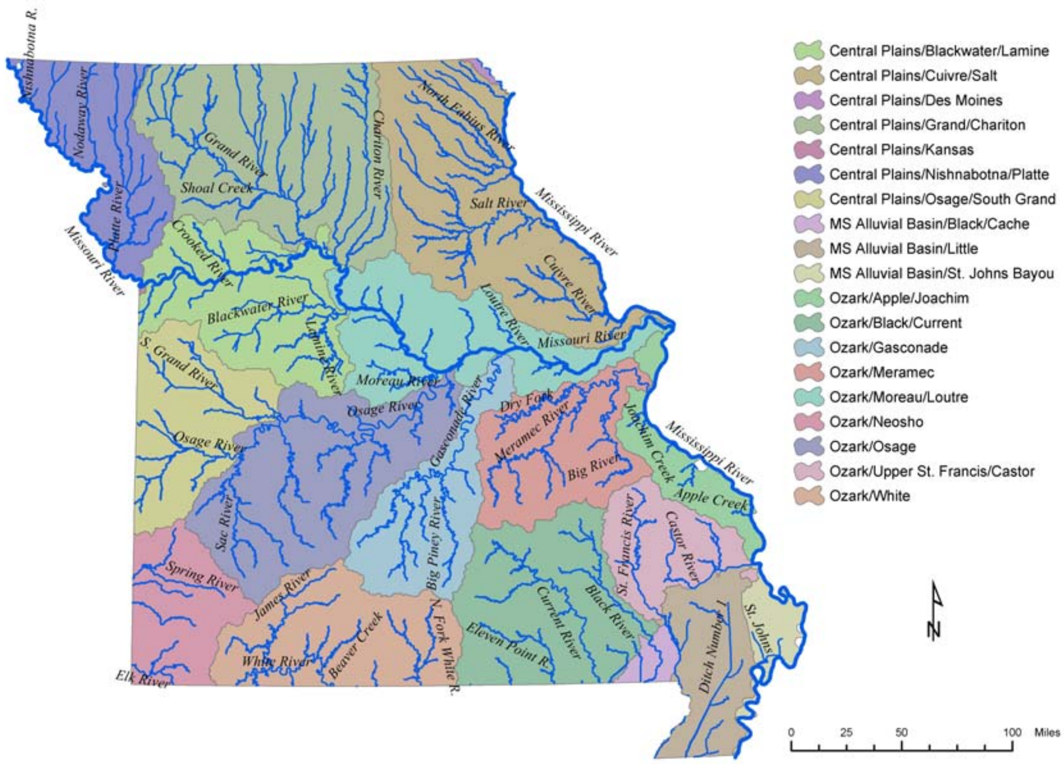


Figure 2. Map of the Ecological Drainage Units (EDUs) of Missouri.

Level 6: Aquatic Ecological System Types

To account for finer-resolution variation in ecological composition we used multivariate cluster analysis of quantitative landscape data to group small- and large-river watersheds into distinct Aquatic Ecological System Types (AES-Types). AES-Types represent watersheds or subdrainages that are approximately 100 to 600 mi² with relatively distinct (local and overall watershed) combinations of geology, soils, landform, and groundwater influence (Figure 3). We determined the number of distinct types by examining relativized overlay plots of the cubic clustering criterion, pseudo F-statistic, and the overall R-square as the number of clusters was increased (Calinski and Harabasz 1974, Sarle 1983). Plotting these criteria against the number of clusters and then determining where these three criteria are simultaneously maximized provides a good indication of the number of distinct clusters within the overall data set (Calinski and Harabasz 1974, Sarle 1983, Milligan and Cooper 1985, SAS 1990, Salvador and Chan 2003). Thirty-eight AES-Types were identified for Missouri with this method.

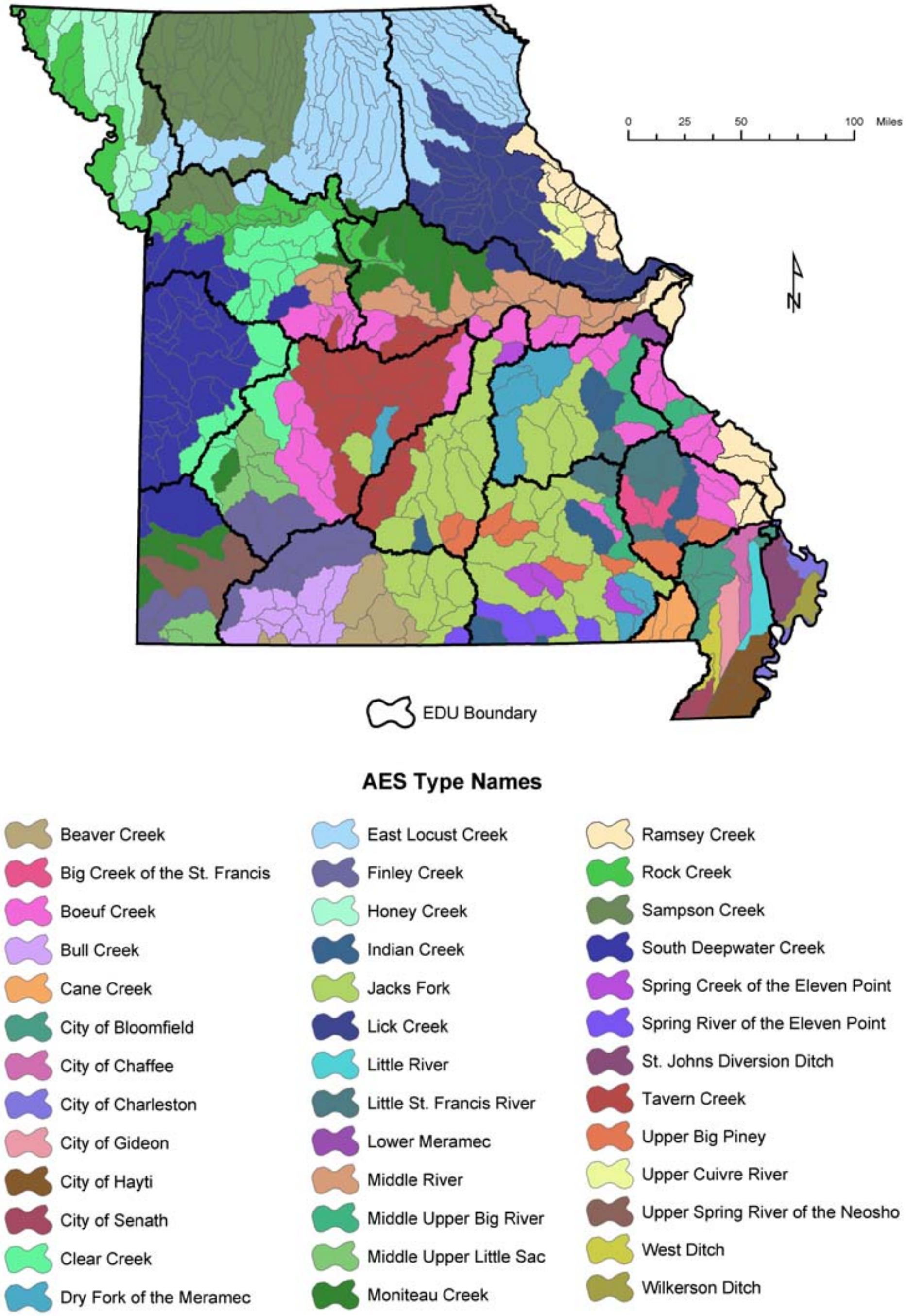


Figure 3. Map of the Aquatic Ecological Systems (AESs) and Types (AES-Types) for Missouri.

AES-Types often initially generate confusion simply because the words or acronym used to name them are unfamiliar. AES-Types are just "habitat types" at a much broader scale than most aquatic ecologists are familiar with. For example, a riffle is a habitat type, yet there are literally millions of individual riffles that occupy the landscape. Each riffle is a spatially distinct habitat; however, they all fall under the same habitat type with relatively similar structural features, functional processes, and ecologically defined assemblages. The same holds true for AES-Types. Each individual AES is a spatially distinct macrohabitat, however, all individual AESs that are structurally and functionally similar fall under the same AES-Type.

Level 7: Valley Segment Types

In Level 7 of the hierarchy 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 (Figure 4). Stream segments within the 1:100,000 USGS/EPA National Hydrography Dataset were 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) are based on those of Pflieger (1989), which were empirically derived with multivariate analyses and prevalence indices. As in the level 6 AESs, VSTs may seem foreign to some, yet if they are simply viewed as habitat types the confusion is removed. 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.

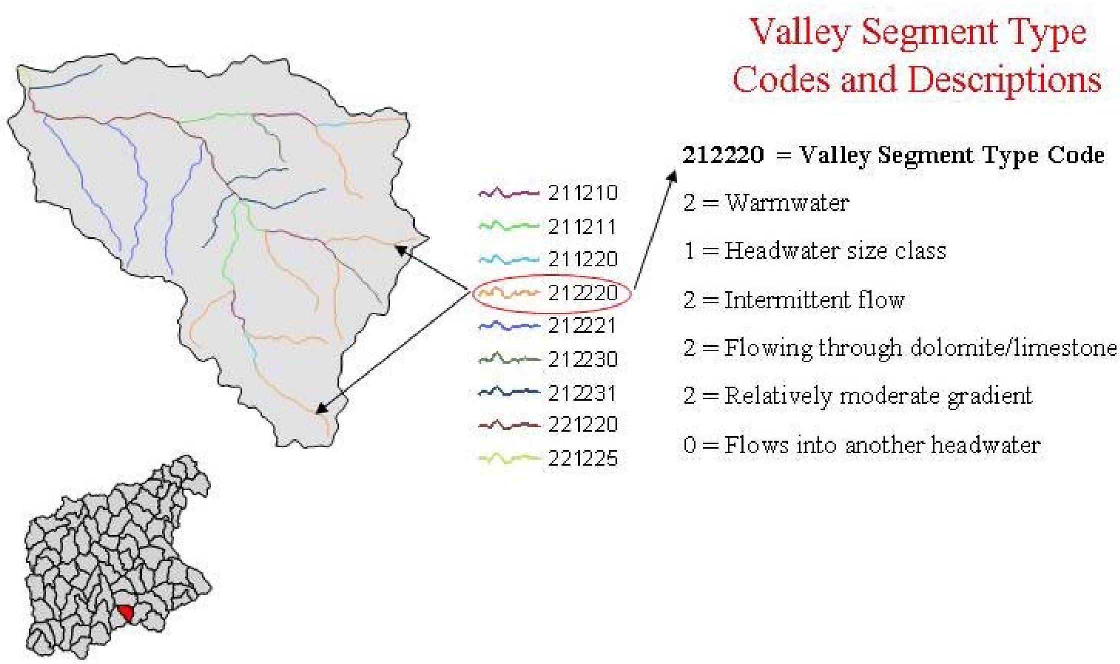


Figure 4. Map showing examples of several different Valley Segment Types (VSTs) within a small watershed of the Meramec EDU.

Level 8: Habitat Types

Units of the final level of the hierarchy, Habitat Types (e.g., high-gradient riffle, lateral scour pool), are simply too small and temporally dynamic to map within a GIS across broad regions or at a scale of 1:100,000. However, we believe it is important to recognize this level of the hierarchy, since it is a widely recognized component of natural variation in riverine assemblages (Bisson et al. 1982, Frissell et al. 1986, Peterson 1996, Peterson and Rabeni 2001).

<u>Step 2: Develop predictive distribution maps for fish, mussels, and crayfish</u>

Purpose:

- Only 0.03% of the stream miles in Missouri have been sampled, and much of this data is spatially and temporally biased. Predicted distribution maps provide us with spatially comprehensive biological data at the finest level of our gap analysis (individual stream segment), which is a resolution that managers can comprehend and at which conservation action typically takes place.
- Since we cannot directly measure or map biodiversity, species within those taxa for which adequate sampling data is available and the associated assemblages must serve as surrogate biotic targets for biodiversity conservation, which complement the abiotic targets.

- Conservation values of society are largely biologically based. The public, legislators, and even scientists can more readily comprehend and relate to biologically based assessments than other measures of biodiversity (e.g., habitat or processes).

To construct our predictive distribution models we compiled nearly 7,000 collection records for fish, mussels, and crayfish and spatially linked these records to the 12-digit USGS/NRCS Hydrologic Unit coverage for Missouri and also to the Valley Segment GIS coverage. Range maps were produced for each of the 315 species, sent out for professional review, and modified as needed. Then we used Decision Tree Analyses to construct predictive distribution models for each species. Ultimately, a total of 571 models were developed to construct reach-specific predictive distribution maps for the 315 species. The resulting maps were merged into a single hyperdistribution (Figure 5), which is related to a database containing information on the conservation status, ecological character, and endemism level of each species.

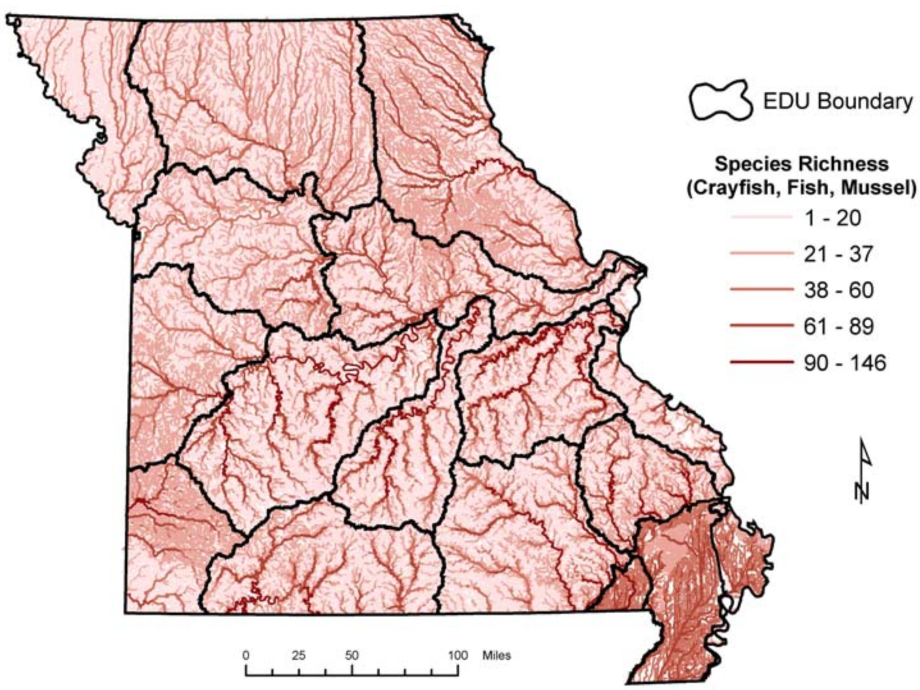


Figure 5. Map of predicted species richness for fish, mussels, and crayfish. This map reflects resource potential and not present-day richness since human disturbances were not included in the models.

Users can select an individual stream segment within the Valley Segment coverage and generate a list of those species (and associated information) predicted to occur in that segment under relatively undisturbed conditions (anthropogenic stressors were not or could not be accounted for). An accuracy assessment was conducted for each taxonomic group using independent data. Commission errors, averaged across all three taxa, were relatively high (55%), while omission errors were relatively low (9%). We believe these accuracy statistics can be improved by incorporating watershed variables as predictors as well as by getting more detailed temperature data for valley segments. However, it must be pointed out that this accuracy assessment is fraught with problems mainly related to the inadequacy of the independent data used to evaluate the accuracy of our models (e.g., insufficient length of stream sampled, only a single sample at a single point in time, inefficient gear, and many of the sampling sites were degraded to some degree while our models predict composition under relatively undisturbed conditions). An assessment of a handful of relatively high-quality, intensively sampled streams revealed a much lower commission error rate (35%) but also a higher omission error rate (18%).

Step 3: Develop local, watershed, and upstream riparian stewardship statistics for each stream segment

Purpose:

- Assess representation of biotic and abiotic targets within the existing matrix of public lands.
- Assist with conservation planning by providing decision makers with information on which to base the selection of focus areas for conservation. For instance, a deciding factor between two locations might be the percentage of the watershed in public ownership (e.g., 10% vs. 50%).
- Assist with conservation planning by providing decision makers with information on who owns the stream segment(s) under consideration as well as the percentage of watershed or upstream riparian ownership by each agency or organization.

The GAP stewardship coverage for Missouri was used in conjunction with the Valley Segment coverage to identify stream segments flowing through public lands. A special Arc Macro Language (AML) program was used to identify only those segments that have the majority of their length ($\geq$ 51%) within public lands (Figure 6). Each segment flowing through public land is further classified according to the GAP stewardship categories (1-4) and the specific owner. Another AML was used to calculate the percentage of each segment's watershed and upstream riparian area in public ownership by GAP stewardship category and owner (Figure 6). Because the watersheds for many of the stream segments within Missouri extend beyond the state, the stewardship coverages for the neighboring states of Iowa, Kansas, and Nebraska were merged with that of Missouri. With these attributes users can now select any of the more than 170,000 individual stream segments within Missouri and see which segments are flowing through public lands, who owns which segments, and what percentage of the overall watershed and upstream riparian area is within public ownership, by either GAP stewardship category or owner.

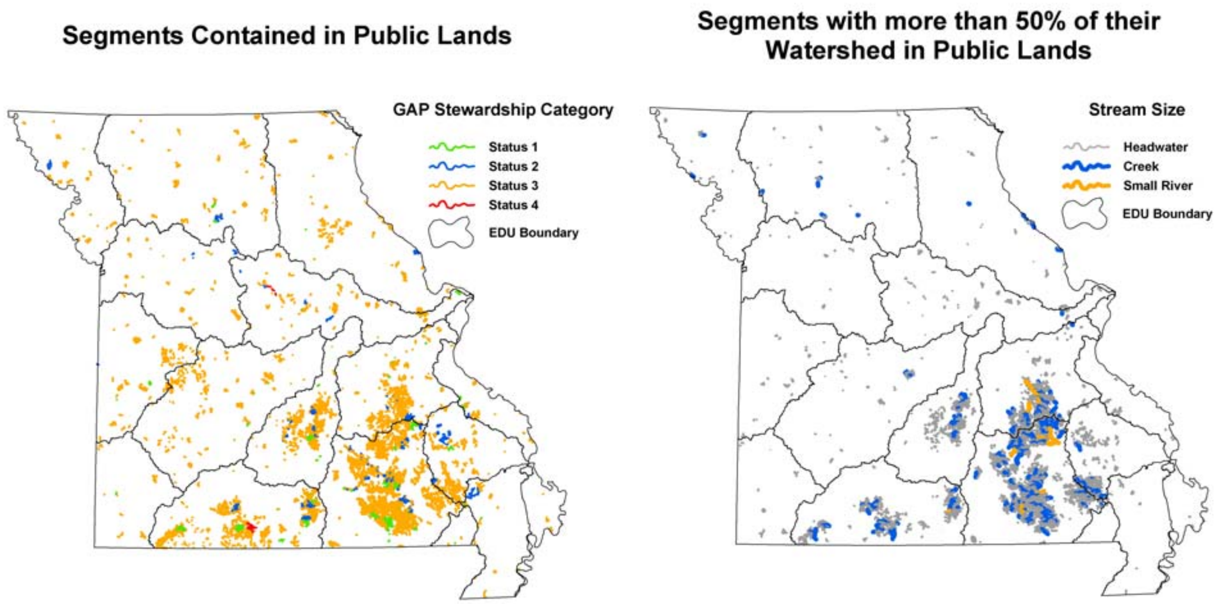


Figure 6. Maps showing, a) stream segments with most of their length in public land, classified by GAP Stewardship categories 1-4, and b) stream segments with $\geq$ 50% of their upstream watershed within existing publics lands.

Step 4: Develop and assemble geospatial data on threats or human stressors

Purpose:

- Because ownership does not ensure effective long-term conservation, measures must be taken to account for human stressors that might significantly impair the ecological integrity of those segments currently within public ownership.
- Assist with conservation planning by providing decision makers with quantitative and qualitative information that can be used to identify relatively high-quality locations in order to conserve a given conservation target.
- Assist with conservation planning by providing decision makers with quantitative and qualitative information that can be used to identify what factors threaten the ecological integrity of a particular priority location, and which can then be used to prioritize management objectives.
- Provide spatially explicit information on human stressors to allow resource managers to pinpoint the specific location of the stressor(s) within the drainage network or watershed.

There are a multitude of stressors that negatively affect the ecological integrity of riverine ecosystems (Allan and Flecker 1993, Richter et al. 1997). The first step in any effort to account for anthropogenic stressors is to develop a list of candidate causes (U.S. EPA 2000). Working in consultation with a team of aquatic resource professionals, a list of the principal human activities known to affect the ecological integrity of streams in Missouri was generated. Then the best available (i.e., highest resolution and most recent) geospatial data that could be found for each of these stressors were assembled (Table 1). Fortunately, and somewhat surprisingly, data were available for most stressors. However, for some, such as channelized stream segments, there were no available geospatial data, and efforts to develop a coverage of such segments using a sinuosity index proved ineffective. Most of the geospatial data were acquired from U.S. EPA and the Missouri Departments of Conservation and Natural Resources.

Table 1. List of the GIS coverages, and their sources, that are used to assess the current conservation status and threats during the conservation planning process for the Missouri CWCS.

Data layer	**Source**
303d Listed Streams	Missouri Department of Natural Resources (MoDNR)
Cafos	MoDNR
Dam Locations	U.S. Army Corps of Engineers (1996)
Drinking Water Supply (DWS) Sites	U.S. Environmental Protection Agency (USEPA)
High Pool Reservoir Boundaries	Elevations from U.S. Army Corps of Engineers
Industrial Facilities Discharge (IFD) Sites	USEPA
Landcover	1992 Missouri Resource Assessment Partnership (MoRAP) Land Cover Classification
Landfills	Missouri Department of Natural Resources, Air and Land Protection Division, Solid Waste Management Program
Mines - Coal	U.S. Bureau of Mines
Mines - Instream Gravel	Missouri Department of Conservation (MDC)
Mines - Lead	U.S. Bureau of Mines
Mines (other/all)	U.S. Bureau of Mines
Nonnative Species	Missouri Aquatic GAP Project - Predicted Species Distributions (MoRAP)
Permit Compliance System (PCS) Sites	USEPA; Ref: http://www/epa.gov/enviro
Resource Conservation and Recovery Information System (RCRIS) Sites	USEPA; Ref: http://www.epa.gov/enviro
Riparian Land Cover	MDC
Superfund National Priority List Sites	USEPA; Ref: http://www.epa.gov/enviro
TIGER Road Files	United States Department of Commerce, Bureau of the Census
Toxic Release Inventory (TRI) Sites	USEPA; Ref: http://www.epa.gov/enviro

Using the Missouri Aquatic GAP Data for Biodiversity Conservation Planning

In fall 2001, federal legislation established a new State Wildlife Grants (SWG) program, which provides funds to state wildlife agencies for conservation of fish and wildlife species, including nongame species. In order to continue receiving federal funds through the SWG program, Congress charged each state and territory with developing a statewide Comprehensive Wildlife

Conservation Strategy (CWCS). In Missouri, the Conservation Department (MDC) is responsible for developing the CWCS. The MDC contacted MoRAP and provided funds to develop customized GIS projects that would assist in the development of a statewide plan for conserving aquatic biodiversity. These customized GIS projects include all of the data compiled or created for the Missouri Aquatic GAP Project, as well as other pertinent geospatial data. At the same time, the MDC developed customized GIS projects for developing a statewide plan for conserving terrestrial biodiversity. Interim results of these two plans will be merged into a single CWCS for the state.

After the customized GIS projects were developed, a team of aquatic resource professionals from around Missouri was assembled. The objective of this team was to address each of the basic components of conservation planning discussed above. The team formulated the following goal: *Ensure the long-term persistence of native aquatic plant and animal communities, by conserving the conditions and processes that sustain them, so people may benefit from their values in the future.*

The team then identified a list of principles, theories, and assumptions that must be considered in order to achieve this goal. Many were similar to those presented above and related mainly to basic principles of stream ecology, landscape ecology, and conservation biology. However, some reflected the personal experiences of team members and the challenges they face when conserving natural resources in regions with limited public land holdings. For instance, one of the assumptions identified by the team was: "Success will often hinge upon the participation of local stakeholders, which will often be private landowners." In fact, the importance of private lands management for aquatic biodiversity conservation was a topic that permeated throughout the initial meetings of the team.

The MoRAP aquatic ecological classification hierarchy was adopted as the geographic framework (i.e., Planning Regions and Assessment Units) for developing the conservation plan. From this classification hierarchy the team selected AES-Types and VSTs as abiotic conservation targets. They also agreed that, in order to fully address biotic targets, a list of target species (fish, mussel, and crayfish) should be developed for each EDU. These lists were developed, and they represent species of conservation concern (i.e., global ranks: G1-G3 and state ranks: S1-S3), endemic species, and focal or characteristic species (e.g., top predators, dominant prey species, unique ecological role, etc.).

Next the team crafted a general conservation strategy. The reasoning behind each component of this strategy is best illustrated by discussing what conservation objectives the team hoped to achieve with each component. These reasons are provided in Box 1.

The conservation strategy

- must develop separate conservation plans for each EDU (Primary Planning Regions);

- whenever possible, represent two distinct spatial occurrences/populations of each target species;
- represent at least one example of each AES-Type within each EDU;
- within each selected AES, represent at least 1 km of the dominant VSTs for each size class (headwater, creek, small river, and large river) as an interconnected complex; and
- represent a least three separate headwater VSTs.

The team then established quantitative and qualitative assessment criteria for making relative comparisons among the assessment units. Since the assessment was conducted at two spatial grains (AES and VST), there exist two different assessment units with assessment criteria developed separately for each.

<u>AES level criteria (listed in order of importance)</u>

- Highest predicted richness of target species
- Lowest Human Stressor Index value (also qualitatively examine individual stressors)
- Highest percentage of public ownership
- Overlap with existing conservation initiatives
- Ability to achieve connectivity among dominant VSTs across size classes
- When necessary, incorporate professional knowledge of opportunities, constraints, or human stressors not captured within the GIS projects to guide the above decisions.

<u>VST level (listed in order of importance)</u>

- If possible, select a complex that contains known viable populations of species of special concern.
- If possible, select the highest-quality VST complex by qualitatively evaluating the relative local and watershed condition using the full breadth of available human stressor data.
- If possible, select a VST complex that is already within the existing matrix of public lands.

If possible, select a VST complex that overlaps with existing conservation initiatives or where local support for conservation is high.

- When necessary, incorporate professional knowledge of opportunities, constraints, or human stressors not captured within the GIS projects to guide above decisions.

The conservation strategy and assessment boils down to a five-step process:

1) Use the AES selection criteria to identify one priority AES for each AES-Type within the EDU.
2) Within each priority AES, use the VST selection criteria to identify a priority complex of the dominant VSTs.
3) For each complex of VSTs create a map of the localized subdrainage (termed “focus area”) that specifically contains the entire interconnected complex.
4) Evaluate the capture of target species.
5) If necessary, select additional focus areas to capture underrepresented target species.

The team then used the conservation strategy and assessment process to develop a conservation plan for the Meramec EDU. By using the above process, all of the objectives of the conservation strategy were met with 11 focus areas (Figure 7). With the initial assessment process and selection criteria, which focus on abiotic targets (AESs and VSTs), 10 separate focus areas were selected. These 10 areas represent the broad diversity of watershed and stream types that occur throughout the Meramec EDU. Within this initial set of 10 focus areas all but five of the 103 target species were captured (Table 2). The distribution of all five of these species overlapped within the same general area of the EDU, near the confluence of the Meramec and Dry Fork Rivers. Consequently, all five of these species were captured by adding a single focus area (the Dry Fork/Upper Meramec focus area, see Figure 7).

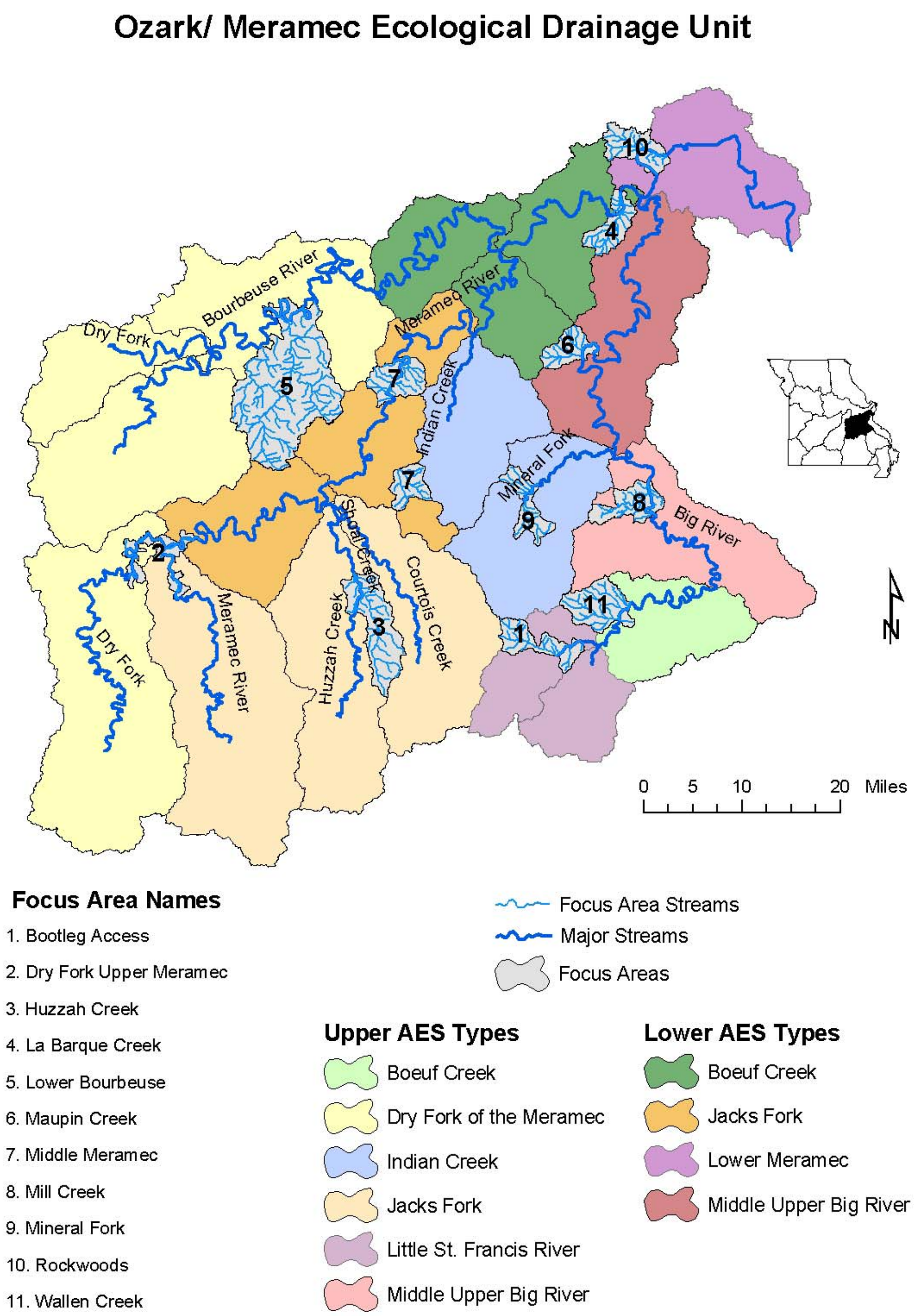


Figure 7. Map showing the 11 Focus Areas selected for the Meramec EDU as part of the aquatic component of the Missouri Comprehensive Wildlife Conservation Strategy. The stream segments within Focus Area number 2 (Dry Fork Upper Meramec) were selected in order to capture those target species not captured in the 10 Focus Areas selected using the initial assessment and selection criteria, which focus on abiotic targets.

Table 2. Target species not captured by the initial conservation planning effort in the Ozark/Meramec EDU, but captured by adding a single Focus Area – Dry Fork/Upper Meramec

Taxon	Common	Scientific	Grank	Srank	Endemism
Fish	blacknose shiner	Notropis heterolepis	G4	S2	Subzone
	lake chubsucker	Erimyzon sucetta	G5	S2	Subzone
	plains topminnow	Fundulus sciadicus	G4	S3	Region
	southern cavefish	Typhlichthys subterraneus	G4	S2S3	Subzone
Crayfish	Salem cave crayfish	Cambarus hubrichti	G2	S3	Subregion

The final set of priority valley segments, within the 11 focus areas, constitutes 186 miles of stream. This represents 2.8% of the total stream miles within the Meramec EDU. The focus areas themselves represent an overall area of 213 mi^2, which is 5% of the nearly 4,000 mi^2 contained within the EDU. Obviously, efforts to conserve the overall ecological integrity of the Meramec EDU cannot be strictly limited to the land area and stream segments within these focus areas. In some instances the most important initial conservation action will have to occur outside of a given focus area, yet the intent of those actions will be to conserve the integrity of the particular focus area. Specific attention to, and more intensive conservation efforts within, these 11 focus areas provides an efficient and effective strategy for the long-term maintenance of relatively high-quality examples of the various ecosystem and community types that exist within this EDU.

In addition to selecting focus areas, the team provides information that can assist with the remaining logistical tasks. This information is captured within a database that can be spatially related to the resulting GIS coverage of the focus areas. Specifically, each focus area is given a name that generally corresponds with the name of the largest tributary stream, then each of the following items is documented:

- all of the agencies or organizations that own stream segments within the focus area and own portions of the overall watershed or upstream riparian area,
- the specific details of why each AES and VST complex was selected,
- any uncertainties pertaining to the selection of the AES or VST complex and if there are any alternative selections that should be further investigated,
- how these uncertainties might be overcome, such as conducting field sampling to evaluate the accuracy of the predictive models or doing site visits to determine the relative influence of a particular stressor,
- all of the management concerns within each focus area and the overall watershed,
- any critical structural features, functional processes, or natural disturbances,
- what fish, mussel, and crayfish species exist within the focus area for each stream size class, and
- any potential opportunities for cooperative management or working in conjunction with existing conservation efforts.

All of this information is critical to the remaining logistical aspects of conservation planning that must be addressed once geographic priorities have been established. Also, since work cannot be immediately initiated within all of the focus areas, there must be priorities established among the focus areas in order to develop a schedule of conservation action (Margules and Pressey 2000). For Missouri, this will initially take place within each EDU and then again from a statewide perspective. An important aspect of generating a "comprehensive" plan is that conservation is often driven by opportunity; by identifying a portfolio of priority locations, quick action can be taken when opportunities arise (Noss et al. 2002).

At present, the selection of focus areas has been completed for 13 of the 17 EDUs. The remaining EDUs will be completed by August 2004. Some of the most important things learned from this process include:

- Local experts are often humbled by the GIS data. Often, what appear to be the best places to conserve are those places that the local managers know little or nothing about. This exemplifies that the world is a big place, and we cannot expect a handful of experts to know every square inch of 4,000+ mi^2.
- The GIS data are often insufficient and, if solely relied upon, would often lead to poor decisions. There have been several instances where the GIS data point us to a particular location, while the local experts quickly point out that, for example, the sewage treatment facility just upstream has one of the worst spill records in the state, and fish kills occur almost on an annual basis. While the GIS data show the location of the sewage treatment facility, they do not contain this more detailed information.
- Even in the most highly altered and severely degraded landscapes there almost always exist "hidden jewels" that have somehow escaped the massive landscape transformations and other insults in neighboring watersheds. This experience has really revealed the social aspects of land use patterns described by Meyer (1995).
- Ninety-five to 100% of the biotic targets are captured by initially only focusing on abiotic targets (AES-Types and VSTs). This is especially surprising in the Ozark Aquatic Subregion, which contains numerous local endemics with very restricted and patchy distributions. This suggests that these classification units do a good job of capturing the range of variation in stream characteristics that are partly responsible for the patchy distribution of these species.
- All of the abiotic and biotic targets can be captured within a relatively small fraction of the overall resource base. Unfortunately, the area of interest for managing these focus areas is often substantially larger and much more daunting. However, the reason priority locations were termed "focus areas" was that the streams and assemblages within each priority location are the ultimate focus of conservation action. Even when work is being conducted outside of a focus area, it should be directed at maintaining or restoring conditions within a particular focus area.
- If possible, priorities should be established at a scale that managers can understand and use (e.g., individual stream segments) in order to apply spatially explicit conservation

actions. Each team of local experts has found the process much more useful than previous planning efforts that have identified relatively large areas as priorities for conservation. The managers have stated that, because we are selecting localized complexes of specific stream segments, much of the guesswork on where conservation action should be focused has been taken "out of the equation," which will expedite conservation action.

Identifying Gaps in the Existing Matrix of Public Lands

Going through the above conservation planning exercise allowed us to more specifically quantify what constitutes a "gap." Arguments about the validity of the specific criteria aside (e.g., why not three occurrences of each target species?), the value of this exercise must be viewed in a broad sense. The criteria embedded within the general conservation strategy are a significant improvement over basic species- or habitat-specific stewardship statistics (e.g., percent of each species range within GAP 1 or 2 lands), which are insufficient for quantifying the true extent of the problem since these statistics lack other important contextual information (e.g., connectivity, number of distinct populations, environmental quality).

What are the results if the criteria used to identify focus areas for the Missouri CWCS are used to assess gaps in the existing conservation network? (see Figure 8). *Note: these statistics pertain to all public lands, not just those meeting criteria for GAP stewardship categories 1 and 2.*

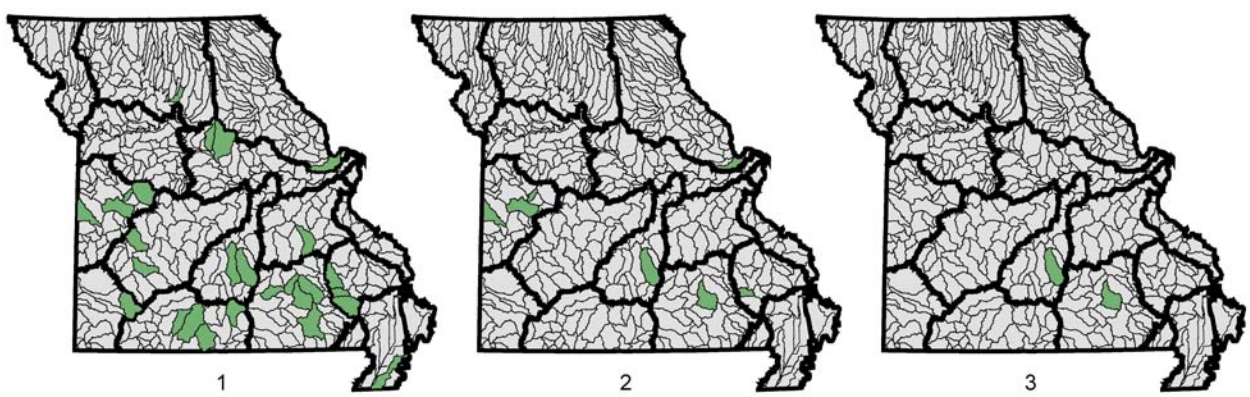


Figure 8. Maps showing,1) those individual AESs that have a least 1 km of the dominant VSTs (for all size classes) currently captured in public lands, 2) those from 1, where the dominant VSTs are captured as an interconnected complex, and 3) those from 2, that can be considered relatively viable options for long-term conservation.

How many individual AESs have at least 1 km of the dominant VSTs (for each size class) captured in existing public lands? 28

How many of these 28 have the dominant VSTs captured as an interconnected complex? 7

How many of these 7 can be considered viable (relatively undisturbed) ecosystems? 2

It is apparent from these results and Figure 8 that none of the EDUs have their full range of watershed or stream types currently captured within the existing matrix of public lands. Furthermore, none of the EDUs even come close to having two occurrences of all target species captured. From a conservation reserve standpoint, these results paint a bleak picture. However, these results should not come as a surprise, considering the fact that conservation of biodiversity, especially riverine biodiversity, has rarely been considered in the acquisition of public lands.

Currently, 7% of the total stream miles in Missouri are in public ownership, yet only a handful of watersheds meet the basic elements of our conservation strategy. Results, thus far, from the statewide conservation planning effort suggest that a reserve network using the outlined conservation strategy would encompass approximately 5-6% of the total stream miles in the state. Consequently, there are more stream miles currently in public ownership than what the conservation planning results suggest is minimally required to represent the "full range" of variation in stream ecosystem types and multiple populations of all fish, mussel, and crayfish species that occur within the state. This irony illustrates the importance of location and spatial arrangement for conserving riverine biodiversity, which heretofore has not been considered in the acquisition of conservation lands. Fortunately, the focus areas presently being identified for the Missouri CWCS serve as an important conservation blueprint to help fill the many voids within the existing conservation network.

Conclusions

The foundation provided by the terrestrial component of GAP in conjunction with an understanding of the basic elements of conservation planning were the key elements that have driven the approach taken in the Missouri Aquatic GAP Project. The data developed for the project are currently being used as the core information in a decision support system for developing a statewide freshwater biodiversity conservation plan. Going through the conservation planning process enabled those involved to more specifically define what constitutes effective conservation for a particular ecosystem and thus better define what constitutes a conservation gap. The gap analysis results are not encouraging. However, the results from the conservation planning efforts provide hope that relatively intact ecosystems still exist even in highly degraded landscapes. Results also suggest that a wide spectrum of the abiotic and biotic diversity can be represented within a relatively small portion of the total resource base, with the understanding that for riverine ecosystems the area of conservation concern is often substantially larger than the focus areas.

Selecting focus areas for conservation is the first step toward effective biodiversity conservation, and the Gap Analysis Program is providing data critical to this task. Yet, establishing geographic priorities is only one of the many steps in the overall process of achieving real conservation. Achieving the ultimate goal of conserving biodiversity will require vigilance on the part of all responsible parties, with particular attention to addressing and coordinating the remaining logistical exercises.

Box 1: What We Are Trying to Achieve with Each Component of the General Conservation Strategy Established for the Missouri CWCS

By attempting to conserve every EDU:

- Provide a holistic ecosystem approach to biodiversity conservation, since each EDU represents an interacting biophysical system.
- Represent all of the characteristic species and species of concern within the broader Aquatic Subregion and the entire state, since no single EDU contains the full range of species found within the upper levels of the classification hierarchy.
- Represent multiple distinct spatial occurrences ("populations") or phylogenies for large-river or wide-ranging species (e.g., sturgeon, catfish, paddlefish), which, from a population standpoint, can only be captured once in any given EDU.

By attempting to conserve an individual example of each AES-Type within each EDU:

- Represent a wide spectrum of the diversity of macrohabitats (distinct watershed types) within each EDU.
- Account for successional pathways and safeguard against long-term changes in environmental conditions caused by factors like Global Climate Change. For instance, gross climatic or land use changes may make conditions in one AES-Type unsuitable for a certain species but at the same time make conditions in another AES-Type more favorable for that species.
- Represent multiple distinct spatial occurrences ("populations") for species with moderate (e.g., bass or sucker species) and limited dispersal capabilities (e.g., darters, sculpins, certain minnow species, most crayfish and mussels).
- Account for metapopulation dynamics (source/sink dynamics).

By attempting to conserve the dominant VSTs for each size class within a single AES:

- Represent the dominant physicochemical conditions within each AES, which we assume represent the environmental conditions to which most species in the assemblage have evolved adaptations for maximizing growth, reproduction, and survival (*sensu* Southwood 1977).

- Represent a wide spectrum of the diversity of mesohabitats (i.e., stream types) within each EDU, since the dominant stream types vary among AES-Types.
- Promote an ecosystem approach to biodiversity conservation by representing VSTs within a single watershed.
- Account for metapopulation dynamics (source/sink dynamics).

By attempting to conserve an interconnected complex of dominant VSTs:

- Account for seasonal and ontogenetic changes in habitat use or changes in habitat use brought about by disturbance (floods and droughts).
    - For instance, during periods of severe drought many headwater species may have to seek refuge in larger streams in order to find any suitable habitat due to the lack of water or flow in the headwaters.
- Account for metapopulation dynamics (source/sink dynamics).
- Further promote an ecosystem approach to conservation by conserving an interconnected/interacting system.

By attempting to conserve at least 3 headwater VSTs within each Focus Area:

- Represent multiple distinct spatial occurrences ("populations") for species with limited dispersal capabilities (e.g., darters, sculpins, certain minnow species, most crayfish and mussels).
- Represent multiple high-quality examples of key reproductive or nursery habitats for many species.

By attempting to conserve at least a 1 km of each priority VST:

- Represent a wide spectrum of the diversity of Habitat Types (e.g., riffles, pools, runs, backwaters, etc.) within each VST and ensure connectivity of these habitats.
- Account for seasonal and ontogenetic changes in local habitat use or changes in habitat use brought about by disturbance (e.g., floods and droughts).
    - For instance, many species require different habitats for foraging (deep habitats with high amounts of cover), reproduction (high-gradient riffles), overwintering (extremely deep habitats with flow refugia or thermally stable habitats like spring branches), or disturbance avoidance (deep or shallow habitats with flow refugia).
- Account for metapopulation dynamics (source/sink dynamics).
- Again, further promote an ecosystem approach to biodiversity conservation by representing an interacting system of Habitat Types.

Literature Cited

Allan, J.D., and A.S. Flecker. 1993. Biodiversity conservation in running waters: Identifying the major factors that threaten destruction of riverine species and ecosystems. *Bioscience* 43:32-43.

Angermeier, P.L., and I.J. Schlosser. 1995. Conserving aquatic biodiversity: Beyond species and populations. Pages 402-414 in J.L. Nielsen, editor. Evolution and the aquatic ecosystem:

Defining unique units in population conservation. American Fisheries Society Symposium 17, Bethesda, Maryland.

Angermeier, P.L., and M.R. Winston. 1998. Local vs. regional influences on local diversity in stream fish communities of Virginia. *Ecology* 79:911-927.

Angermeier, P.L., R.A. Smogor, and J.R. Stauffer. 2000. Regional frameworks and candidate metrics for assessing biotic integrity in mid-Atlantic highland streams. *Transactions of the American Fisheries Society* 129:962-981.

Bailey, R.G. 1995. Description of the ecoregions of the United States. Second edition. U.S. Forest Service, Washington, DC. Miscellaneous Publication No. 1391.

Belbin, L. 1993. Environmental representativeness: Regional partitioning and reserve selection. *Biological Conservation* 66:223-230.

Bisson, P.A., J.L. Nielsen, R.A. Palmason, and L.E. Grove. 1981. A system of naming habitat in small streams, with examples of habitat utilization by salmonids during low streamflow. Pages 62-73 in N.B. Armantrout, editor. Acquisition and utilization of aquatic habitat inventory information. Proceedings of a symposium, Oct. 28-30, 1981, Portland, Oregon. Hagen Publishing Co., Billings, Montana.

Calinski, T., and J. Harabasz. 1974. A dendrite method for cluster analysis. *Communications in Statistics* 3:1-27.

Diamond, D.D., C.D. True, T.M. Gordon, S.P. Sowa, W.E. Foster, and K.B. Jones. In press. Influence of targets and area of assessment on perceived conservation priorities. *Environmental Management.*

Dunne, T., and L. Leopold. 1978. Water in Environmental Planning. Freeman and Company, New York.

Franklin, J.F. 1993. Preserving biodiversity: Species, ecosystems, or landscapes. *Ecological Applications* 3:202-205.

Frissel, C.A., W.J. Liss, C.E. Warren, and M.D. Hurley. 1986. A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. *Environmental Management* 10:199-214.

Groves, C. 2003. Drafting a conservation blueprint: A practitioner's guide to planning for biodiversity. Island Press, Washington, DC. 457 pp.

Grumbine, R.E. 1994. What is ecosystem management? *Conservation Biology* 8:27-38.

Higgins, J.V. 2003. Maintaining the ebbs and flows of the landscape: Conservation planning for freshwater ecosystems. Pages 291-318 in Groves, C., editor. Drafting a conservation blueprint: A practitioner's guide to planning for biodiversity. Island Press, Washington, DC. 457 pp.

Hynes, H.B.N. 1975. The stream and its valley. *Verh. Int. Theor. Ang. Limnol.* 19:1-15.

Kirkpatrick, J.B., and M.J. Brown. 1994. A comparison of direct and environmental domain approaches to planning reservation of forest higher plant communities and species in Tasmania. *Conservation Biology* 8:217-224.

Leslie, M., G.K. Meffe, J.L. Hardesty, and D.L. Adams. 1996. Conserving biodiversity on military lands: A handbook for natural resources managers. The Nature Conservancy, Arlington, Virginia.

Margules, C.R., and R.L. Pressey. 2000. Systematic conservation planning. *Nature* 405:243-253.

Matthews, W.J. 1987. Physicochemical tolerance and selectivity of stream fishes as related to their geographic ranges and local distributions. Pages 111-120 in W.J. Matthews and D.C. Heins, editors. Community and evolutionary ecology of North American stream fishes. University of Oklahoma Press, Norman, Oklahoma.

Matthews, W.J. 1998. Patterns in freshwater fish ecology. Chapman and Hall, New York.

Maxwell, J.R., C.J. Edwards, M.E. Jensen, S.J. Paustian, H. Parrott, and D.M. Hill. 1995. A hierarchical framework of aquatic ecological units in North America (Nearctic Zone). U.S. Forest Service, North Central Forest Experiment Station, St. Paul, Minnesota. General Technical Report NC-176.

Mayr, E. 1963. Animal species and evolution. Belknap Press of Harvard University Press. Cambridge, Massachusetts. 797 pp.

Meyer, W.B. 1995. Past and present land use and land cover in the USA. *Consequences* 1(1): online journal available at http://www.gcrio.org/CONSEQUENCES/.

Milligan, G.W., and M.C. Cooper. 1985. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 50:159-179.

Moyle, P.B., and J.J. Cech, Jr. 1988. Fishes: An introduction to ichthyology, 2nd edition. Prentice-Hall, Englewood Cliffs, New Jersey.

Nicholls, W.F., K.T. Kinningbeck, and P.V. August. 1998. The influences of geomorphological heterogeneity on biodiversity: A landscape perspective. *Conservation Biology* 12:371-379.

Noss, R.F. 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4:355-364.

Noss, R.F., C. Carroll, K. Vance-Borland, and G. Wuerthner. 2002. A multicriteria assessment of the irreplaceability and vulnerability of sites in the Greater Yellowstone Ecosystem. *Conservation Biology* 16:895-908.

Noss, R.F. 2004. Conservation targets and information needs for regional conservation planning. *Natural Areas Journal* 24: 223-231.

Omernik, J.M. 1995. Ecoregions: A spatial framework for environmental management. Pages 49-62 in W. Davis and T. Simon, editors. Biological assessment and criteria: Tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Orians, G.H. 1993. Endangered at what level? *Ecological Applications* 20:206-208.

Osborne, 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.

Peterson, J.T. 1996. The evaluation of a hydraulic unit-based habitat system. Ph.D. dissertation. University of Missouri, Columbia, Missouri. 397 pp.

Peterson, J.T., and C.F. Rabeni. 2001. The relation of fish assemblages to channel units in an Ozark stream. *Transactions of the American Fisheries Society* 130:911-926.

Pflieger, W.L. 1971. A distributional study of Missouri fishes. University of Kansas Publications, Lawrence, Kansas. Museum of Natural History Volume 20:225-570.

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

Pflieger, W.L. 1996. The crayfishes of Missouri. Missouri Department of Conservation, Jefferson City, Missouri.

Rabeni, C.F., R.J. Sarver, N. Wang, G.S. Wallace, M. Weiland, and J.T. Peterson. 1997. Development of regionally based biological criteria for streams of Missouri. A report to the Missouri Department of Natural Resources from the Missouri Cooperative Fish and Wildlife Research Unit, University of Missouri, Columbia, Missouri.

Rabeni, C.F., and K.E. Doisy. 2000. The correspondence of stream benthic invertebrate communities to regional classification schemes in Missouri. *Journal of the North American Benthological Society* 19:419-428.

Richter, B.D, D.P. Braun, M.A. Mendelson, and L.L. Master. 1997. Threats to imperiled freshwater fauna. *Conservation Biology* 11:1081-1093.

Salvador, S., and P. Chan. 2003. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. Department of Computer Sciences Technical Report CS-2003-18, Florida Institute of Technology. Melbourne, Florida.

Sarle, W.S. 1983. The Cubic Clustering Criterion. SAS Technical Report A-108. SAS Institute Inc. Cary, North Carolina. ....

SAS Institute. 2001. SAS user's guide: Statistics. Version 8.2. SAS Institute Inc., Cary, North Carolina.

Scott, J.M., B. Csuti, K. Smith, J.E. Estes, and S. Caicco. 1991. Gap Analysis of species richness and vegetation cover: An integrated biodiversity conservation strategy. Pages 282-297 in K. Kohm, editor. Balancing on the brink of extinction. Island Press, Washington, DC.

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, Michigan. 51.pp.

Shaffer, M.L. 1990. Population viability analysis. *Conservation Biology* 4:39-40.

Shaffer, M.L., and B.A. Stein. 2000. Safeguarding our precious heritage. Pages 301-321 in B.A. Stein, L.S. Kutner, and J.S. Adams, editors. Precious heritage: The status of biodiversity in the United States. The Nature Conservancy, Oxford University Press, New York.

Smale, M.A., and C.F. Rabeni. 1995a. Hypoxia and hypothermia tolerances of headwater stream fishes. *Transactions of the American Fisheries Society* 124:698-710.

Smale, M.A., and C.F. Rabeni. 1995b. Influences of hypoxia and hyperthermia on fish species composition in headwater streams. *Transactions of the American Fisheries Society* 124:711-725.

Southwood, T.R.E. 1977. Habitat, the template for ecological strategies? *Journal of Animal Ecology* 46:337-365.

U.S. EPA. 2000. Stressor identification guidance document. U.S. Environmental Protection Agency, Office of Water, Office of Research and Development, Washington, DC. EPA 822-B-00-025.

Wiens, J.A. 1989. Spatial scaling in ecology. *Functional Ecology* 3:385-397.

APPLICATIONS

A Framework to Extend Gap Analysis to Multi-Objective Conservation Planning

DAVID STOMS, FRANK DAVIS, CHRIS COSTELLO, ELIA MACHADO, AND JOSH METZ
University of California, Santa Barbara

Introduction

While we (FD and DS) were developing the California Gap Analysis Project in the mid-1990s, we naturally became interested in using our GAP data to design reserve networks to fill the gaps we were identifying. We began collaborating on various reserve selection methods for choosing sets of sites that would achieve conservation targets efficiently. These efforts were satisfying intellectually as they became more sophisticated, but we were somewhat frustrated that this kind of systematic approach to conservation planning was not being adopted widely in public land use planning. An opportunity to rethink the conservation planning problem arose after California's legislative watchdog agency had been critical of the state's conservation program. They cited a lack of coordination among agencies with different agendas and an inability to formally evaluate properties when they were offered for acquisition. Was the state moving cost-effectively toward some desired endpoint? As a result, the California Resources Agency contracted with the National Center for Ecological Analysis and Synthesis (NCEAS) to convene a working group that would bring systematic conservation planning theory and methods to bear on the design and implementation of the state's conservation programs through their California Legacy Project (CLP, http://www.legacy.ca.gov). Because California's conservation programs (like many others) act on voluntary offers of private lands to be acquired from a fixed budget (e.g., proceeds from a bond initiative), they did not need or want a process to develop a long-range plan that may take decades to implement. Rather they needed a process to evaluate and prioritize the set of properties that are currently available for conservation. In this paper we provide a brief overview of a planning framework produced by the NCEAS working group. A detailed technical description of the framework can be found online at http://www.nceas.ucsb.edu/nceas-web/projects/4040/TerrBiod_framework-report.pdf.

Prioritizing Places for Conservation Investments

What makes this framework different from earlier examples? It differs primarily in three aspects: the overall focus of systematic conservation planning, multiple rather than single objectives, and the measure of site conservation value.

Most examples of conservation planning tools follow a reserve selection paradigm that either meets conservation goals for protection at minimum cost (or area) or maximizes biodiversity protected for a limited budget. The performance measure is how much biodiversity is contained (represented) in the reserve network. Biodiversity outside of reserves is not credited. In our framework we shifted the focus to maximizing how much biodiversity is expected to remain in the future (whether in reserves or not). Adding new reserves becomes the means to that end rather than the end in itself.

The planning framework is organized into a hierarchy of five conservation objectives for terrestrial biodiversity currently used in conservation practice:

1) Protect hotspots of rare and endangered species;
2) Protect underrepresented species and communities (the GAP perspective used in most reserve selection models);
3) Maintain ecological and evolutionary processes in landscapes;
4) Protect wildlands for large carnivores and area-dependent species;
5) Expand existing reserves.

Each objective represents a different policy for prioritizing conservation investments, and each invokes a somewhat distinctive set of ecological and spatial criteria.

Arguably the most important feature of any planning approach is the specification of the performance measure by which sites can be prioritized. The performance measure may also be seen as a "marginal utility function." In traditional reserve selection approaches, this measure is often some form of complementarity. Because our framework is concerned with maximizing the amount of biodiversity that remains at some future time, it requires a change in how we measure the value of individual sites. Instead of measuring how much biodiversity a site has today, we project how much biodiversity would be lost if it were not protected. Or more concisely, what difference would conservation make? If a site is not threatened, its biodiversity is likely to remain even without formal protection, and so our framework would assign it a low priority. Therefore, the framework requires a spatially explicit scenario of future land use that is used to estimate the potential loss of biodiversity in the site and the region. Specifically, we compute the difference between the area of each element with and without conservation as an index of threat.

The framework establishes a relationship between the level or amount of a resource (e.g., the area of a particular habitat type) and the "utility" associated with it. Economists recognize that the utility of the next unit of some resource depends on how much you already have. If there were a million hectares of habitat for a species, protecting the first hectare probably has more social utility than protecting the millionth. In a gap analysis context, planners set goals of how much habitat to protect, which is usually an estimate of the minimum area needed (with some degree of confidence) for the species or community to persist. The implicit relationship between resource amount and marginal utility is a step function, where marginal utility of remaining unprotected sites becomes zero once minimum conservation goals are achieved. That implies that society would be satisfied with a set of reserves surrounded by intense land use, which we believe grossly oversimplifies the social demand for conservation. This is rather like being satisfied with a subsistence diet instead of recognizing that as a bare minimum. To reserve selection algorithms, the goal is treated as the ceiling (this much and no more) whereas to us, the goal is like a floor (at least this much for persistence but the more the better). The step function is a special case of our more general diminishing marginal utility function.

To measure a site's overall value for conserving terrestrial biodiversity, we estimate the site's marginal utility for each of the five conservation objectives listed above. These are combined by weighting each objective (according to stakeholders' values) and summing the weighted values for each site. The final step in our framework is a budget allocation model. Our approach to measuring conservation value is based on cost-effectiveness, with the cost for whatever action is deemed necessary to remove threat. The problem the allocation model solves is to invest a fixed budget in a set of sites that, if conserved, would minimize the loss of terrestrial biodiversity during the planning period.

Summary and Future Directions

The framework we have developed for the California Legacy Project has not been fully vetted with the relevant state agencies or other stakeholder groups, so it remains to be seen whether the ideas and methods will prove useful in real planning efforts. We believe the strengths of the framework are its generality, explicitness, modest data requirements (such as GAP), flexibility for exploring alternatives, formal consideration of threats and costs, and—perhaps most importantly—its ability to help in choosing among competing projects. For many organizations, this may be more useful than optimizing grand conservation plans that are often out of date the moment they are adopted. We believe the framework could be adapted to other regions, scales, and ownerships (e.g., prioritizing acquisitions for National Wildlife Refuges, National Forest or BLM land management planning, land exchanges), restoration projects, and aquatic biodiversity. We are eager to explore these opportunities further with you (contact stoms@bren.ucsb.edu).

The framework has only been implemented sufficiently to demonstrate the concepts and therefore is currently rather cumbersome. NatureServe (http://www.natureserve.org/) is developing a planning support system, and we are working with them to codify our process.

LAND COVER

Land Cover Mapping Using StatMod: An ArcView® 3.X Extension for Classification Trees Using S-PLUS®

JOHN LOWRY, CHRISTINE GARRARD, AND DOUG RAMSEY
Remote Sensing/GIS Laboratory, College of Natural Resources, Utah State University, Logan

Introduction

The use of classification trees (CT) for land cover mapping is becoming increasingly common (Hansen et al. 1996, Lawrence and Wright 2001, Pal and Mather 2003, Brown de Colstoun 2003). Classification trees, sometimes called decision trees, or CART (Classification and Regression Trees) offer several advantages over classification algorithms traditionally used for land cover mapping. One advantage is the ability to effectively use both categorical and continuous predictor data sets with different measurement scales. Other advantages include the ability to handle nonparametric training and predictor data, good computational efficiency, and an intuitive hierarchical representation of discrimination rules. Classification trees use multiple explanatory variables to predict a single response variable.

A major challenge with using classification trees for land cover mapping lies in spatially applying the rules generated from the CT software within a geographic information system (GIS). StatMod for ArcView 3.X was developed by Christine Garrard at Utah State University with the purpose of interfacing the CT tools available in S-PLUS with ArcView GIS (Garrard 2002). StatMod is available free and can be downloaded, with an accompanying user's guide, from http://www.gis.usu.edu/~chrisg/avext/.

StatMod ArcView 3.X Extension

StatMod provides the option to automatically submit jobs to S-PLUS, in which case all interactions with S-PLUS are through an ArcView dialog box. Alternatively StatMod allows manual creation of tree models in S-PLUS, which can thereafter be spatially applied in ArcView. This flexibility extends the functionality of StatMod to a range of users—from those with little experience using S-PLUS to experienced S-PLUS users. A basic knowledge of ArcView is, however, necessary to successfully use StatMod.

<u>Response and Explanatory Variables</u>

The response variable, or training theme, is represented by training sites distributed throughout the study area and may be either a point or polygon theme. The training theme must have an attribute field containing codes or descriptions for the land cover classes to be modeled using the CT. Explanatory variables are spatial data layers from which CT rules will be generated to predict the spatial distribution of land cover. Examples of explanatory variables include individual satellite image bands, band transformations such as NDVI (Normalized Difference Vegetation Index), a digital elevation model, and topographic aspect or geology GIS data sets. When the training variable is a point theme, explanatory variables can be either polygon or grid themes. When the training variable is a polygon theme, explanatory variables must be grid themes.

Once a training theme and multiple explanatory themes are added to the View, the associated value for each training site is obtained by intersecting the training theme through the explanatory themes. When the training theme is a point theme, the value in each training site is the intersected value taken from the explanatory themes. When the training theme is a polygon theme, StatMod provides a choice of statistics such as mean, maximum, or majority value to characterize each training site. Refer to Figure 1 for a graphic depiction of how the response (dependent) variable and explanatory (independent) variables are identified in the StatMod dialog box.

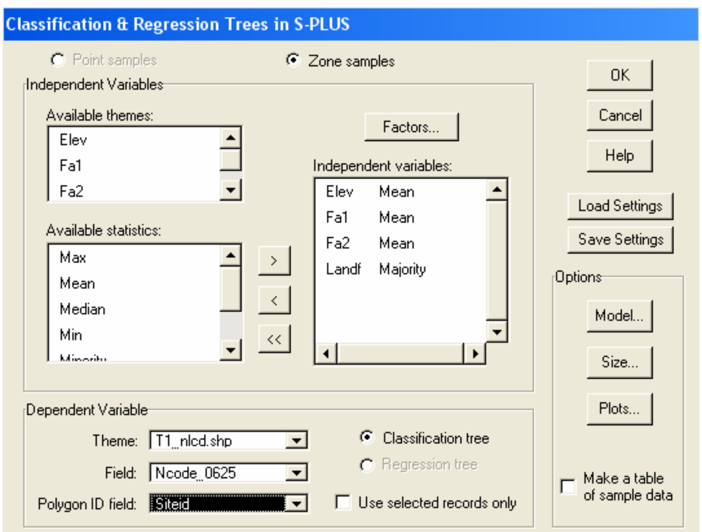


Figure 1. StatMod dialog box for Classification and Regression Trees.

<u>Building a Classification Tree</u>

The CT algorithm determines the appropriate characteristics of the response variable by recursively splitting the explanatory data into increasingly more homogeneous groups (Figure 2), producing a hierarchical tree composed of “rules” defining the characteristics of each response category (Figure 3). Commonly CT models are overfitted to the training data, that is, the CT algorithm recursively splits the data until rules are generated for specific training sites rather than entire response categories. Once an overfitted tree is generated, it can be reduced in size to create a tree that is neither precisely fitted to the training data nor so general that it is not meaningful. S-PLUS offers two methods for reducing tree size: “pruning” and “shrinking.”

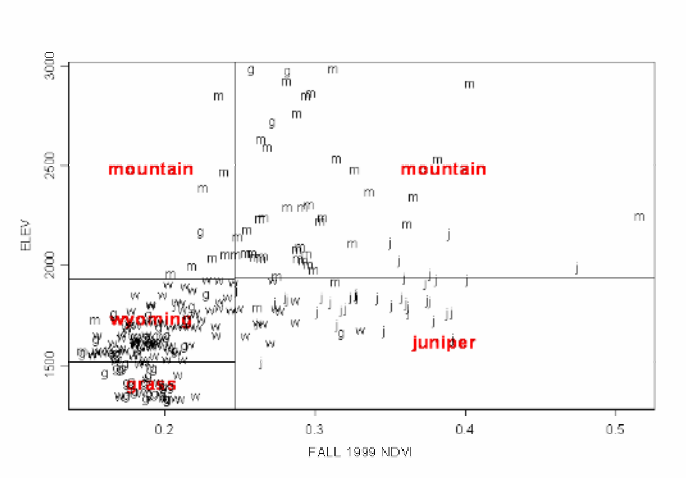


Figure 2. Example of four cover types discriminated by elevation and Fall NDVI.

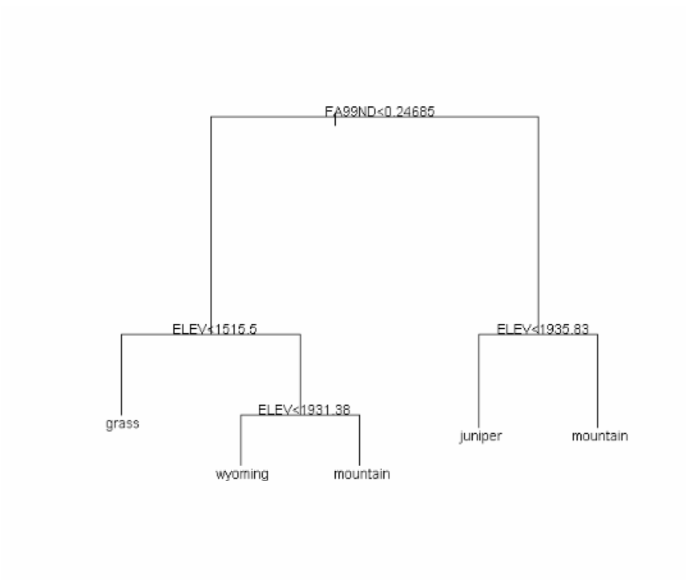


Figure 3. Discrimination rules from Figure 2 presented as a classification tree.

Choosing the best tree reduction method is typically achieved through iteratively growing and reducing tree models, with subsequent evaluation of deviation or misclassification error rates and testing different predictor variables and pruning or shrinking criteria. StatMod provides a convenient interface allowing the user to choose one of several methods of controlling tree size (Figure 4). These include a one standard error rule, Akaike's information criterion (AIC), the size or number of tree nodes, and a cost complexity parameter. For more detailed information on options for controlling tree size refer to the StatMod user's guide (Garrard 2002b).

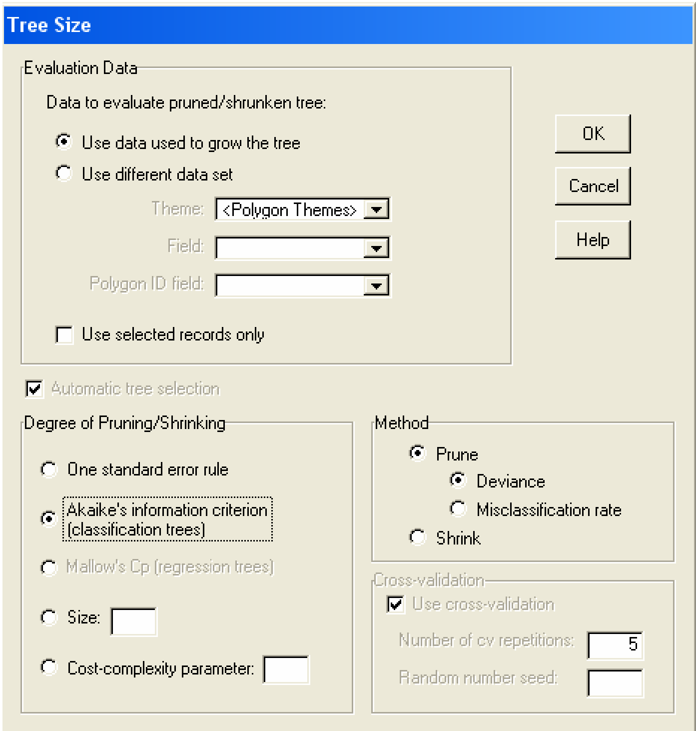


Figure 4. StatMod dialog box used to control tree size.

Case Study Using StatMod

Objectives and Methods

The mapping area comprises 5 million acres in Utah's High Plateau region situated on the western edge of the Rocky Mountains. Vegetation cover includes basin big sagebrush at lower elevations, with expanses of pinyon-juniper communities at mid-elevations. Upper montane communities include Douglas-fir, aspen, and ponderosa pine, and at higher elevations spruce/fir mixes, aspen, and tundra dominate. Barren areas are present in the southeastern edge of the mapping area, which borders Utah's slickrock country.

Approximately 3,800 training samples were available for the mapping area. All training samples were labeled with one of seven NLCD (National Land Cover Database) class codes. These correspond to Barren Lands, Deciduous Forest, Evergreen Forest, Mixed Forest, Shrub/Scrub, Grassland/Herbaceous, and Woody Wetlands. Twenty percent of the sample sites were randomly selected and withheld for accuracy assessment.

Predictor layers used for the classification tree included a digital elevation model, a raster landform model, and Enhanced Thematic Mapper (ETM) bands 1-5 and 7 (converted to grids) for a summer and fall date. Using StatMod, a classification tree was created using default S-PLUS model parameters. The tree was pruned to optimal size using Akaike's information criterion (AIC).

Results

StatMod produced the predicted land cover map and a text file with an .smg extension. The .smg file reports the predictor variables used in the construction of the tree, the number of terminal nodes produced, and the misclassification error rate. It also contains a textual presentation of the rules that comprise the classification tree. For this study, all predictor variables were used, and the tree was comprised of 70 terminal nodes. The tree had a misclassification rate of 0.19, meaning that 81% of the training data could be predicted by the classification tree.

The predicted map was produced as a grid and was displayed in the active View. Attribute information stored in separate fields in the .vat of the grid include the predicted land cover class, the probability of correct classification (inverse of misclassification), and calculated deviance for each grid cell. Figure 5 shows the probability values associated with each cell for the predicted grid. Low (black) to high (white) probabilities of correct classification are displayed using a graduated color ramp. It should be noted that “probability” is based on misclassification rates determined by model fit and not from an independent data source.

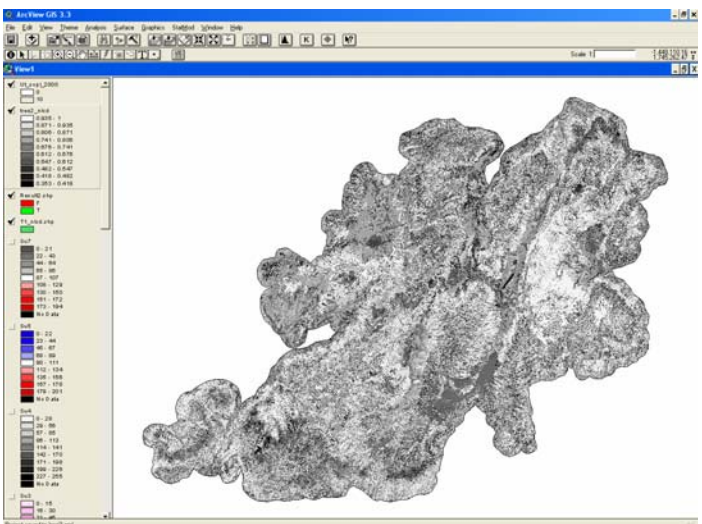


Figure 5. Probability of correct classification ranging from low (black) to high (white).

StatMod also provides a convenient tool for assessing accuracy with a traditional error matrix and kappa calculation. Using the withheld 20% of the sample data, the Kappa tool in StatMod was used to intersect 712 withheld sample sites through the predicted land cover map. When polygon sample data are used, the tool assumes a correct classification when the majority of cells in the predicted map agree with the sample polygon. Overall accuracy was 75% with a kappa statistic of .67. User’s accuracies were as follows: Barren Lands (72%), Deciduous Forest (81%), Evergreen Forest (79%), Mixed Forest (55%), Shrub/Scrub (64%), Grassland/Herbaceous (76%), and Woody Wetlands (47%).

Summary

StatMod provides an easy-to-use and inexpensive tool for spatially applying the classification rules generated from the CT algorithm in S-PLUS. While the focus of this article was to use StatMod for classification trees, StatMod functions in a similar manner for regression trees. Classification trees are appropriate for discriminating distinct classes such as land cover. In a regression tree, the response variable is a continuous numeric field such as percent canopy cover. In addition to interfacing with S-PLUS for classification and regression trees, StatMod can be used to interface with SAS® to create and spatially apply logistic regression models.

Literature Cited

Brown de Colstoun, E.C., M.H. Story, C. Thompson, K. Commisso, T.G. Smith, and J.R. Irons. 2003. National Park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier. *Remote Sensing of Environment* 85:316-327.

Garrard, C.M. 2002. StatMod: A tool for interfacing ArcView GIS with statistical software to facilitate predictive ecological modeling. Master of Science Thesis. Utah State University.

Garrard, C.M. 2002b. StatMod Zone user's guide. WWW URL: http://www.gis.usu.edu/~chrisg/avext/

Lawrence, R.L., and A. Wright. 2001. Rule-based classification systems using Classification and Regression Tree (CART) Analysis. *Photogrammetric Engineering and Remote Sensing* 67:1137-1142.

Pal, M., and P.M. Mather. 2003. An assessment of effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment* 86:554-565.

The Wide Dynamic Range Vegetation Index and its Potential Utility for Gap Analysis

GEOFFREY M. HENEBRY, ANDRÉS VIÑA, AND ANATOLY A. GITELSON
Center for Advanced Land Management Information Technologies (CALMIT), School of Natural Resources, University of Nebraska-Lincoln

Introduction

In landscapes with moderate to high densities of green biomass, the widely used Normalized Difference Vegetation Index (NDVI) has long been known to exhibit reduced sensitivity to moderate-to-high vegetation density. This loss of sensitivity diminishes the utility of the NDVI to discriminate among land cover types or land cover quality. A straightforward modification of the NDVI, the Wide Dynamic Range Vegetation Index (WDRVI), was recently developed (Gitelson 2004) and has been shown to be effective in tracking spatio-temporal variation in diverse ecoregions throughout the conterminous United States (Viña et al. 2004). In this brief note, we illustrate the prevalence of reduced sensitivity of the NDVI, introduce the WDRVI, and illustrate the advantages of the WDVRI over the NDVI using Landsat ETM+ data that spans a range of canopy densities.

Limitations of the NDVI

The Normalized Difference Vegetation Index is calculated as the ratio of the difference between near infrared (ρ_{NIR}) and red (ρ_{red}) reflectance divided by their sum: ($\rho_{NIR}-\rho_{red}$) / ($\rho_{NIR}+\rho_{red}$). Values range from -1 to +1. The specific value of NDVI for a scene depends on the wavelengths used to represent ρ_{NIR} and ρ_{red}, the radiometric and spatial resolutions of the sensor, the illumination and atmospheric conditions, the sun-target-sensor geometry, and the distribution and types of objects within a scene. The proper biogeophysical interpretation of the NDVI is the fraction of absorbed photosynthetically active radiation (fAPAR). The NDVI loses sensitivity when the leaf area index (LAI) exceeds about 2. Reduction in its dynamic range means fewer distinct levels of NDVI are observable. When the LAI is much larger than 2, even a large change in the LAI may be undetectable using the NDVI. This has implications for land cover/land use change studies but land cover classification as well. A limited dynamic range may distort and obscure interesting spectral features that could aid classification.

During a significant portion of the temperate growing season, it is as if a green veil obscures changes across the vegetated land surface. We can visualize the duration and extent of the green veil using the biweekly composites of maximum NDVI as observed by the NOAA AVHRR sensors. Here we simply count the number of times during the growing season that a pixel exceeds a specific NDVI threshold associated with the transition to reduced sensitivity. In Figure 1, for example, there are some dark areas of the region that never experience reductions in NDVI sensitivity (e.g., lakes, reservoirs, badlands) and others that are in the zone throughout the growing season (e.g., coniferous forests, deciduous forests in eastern Kansas, integrated agribusiness complex near Garden City, KS). Note the distinct bright triangle in Nebraska south of the Platte River (see arrow); we will zoom into this area in an example below.

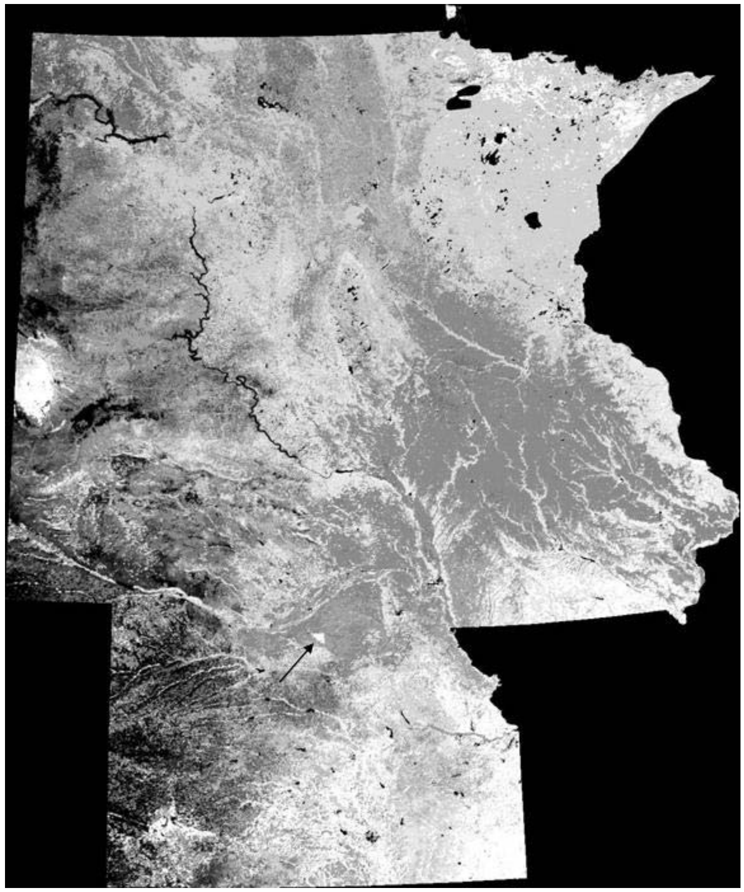


Figure 1. Persistence of reduced NDVI sensitivity over the GAP Great Plains region (ND, SD, NE, KS, MN, IA) using AVHRR composites from 2000. Brighter pixels spent more time during the 15 biweekly compositing periods of the growing season in the zone of reduced NDVI sensitivity.

Lifting the green veil with the WDRVI

Gitelson (2004) introduced the WDRVI as a way to enhance the dynamic range of the NDVI by applying a weighting parameter α to the near infrared reflectance:

$$WDRVI = (\alpha * \rho_{NIR} - \rho_{red}) / (\alpha * \rho_{NIR} + \rho_{red}). \quad [1]$$

If α equals 1, then the WDRVI is equivalent to the NDVI. If α equals $(\rho_{red} / \rho_{NIR})$, then the WDRVI equals zero. Think of α as a tuning knob that adjusts the gain on the index. Selection of the coefficient for the α parameter requires some forethought, so we will illustrate the effect of different coefficient values on the WDRVI.

Example with Landsat 7 data

We have chosen a small piece of an ETM+ scene acquired on August 4, 2001 (Path 29, Row 32) with a nominal spatial resolution of 28.5 m. Figure 2 shows the NDVI calculated from sensor reflectances without any atmospheric correction. For this same image, we also calculated the

WDRVI at different levels of α (0.20, 0.10, 0.05) that had been used by Gitelson (2004). We also calculated a coefficient value adjusted to scene characteristics using the heuristic:

$$\alpha_{est} = 2 * (\text{average } \rho_{red}) / (\text{maximum } \rho_{NIR}) \quad [2]$$

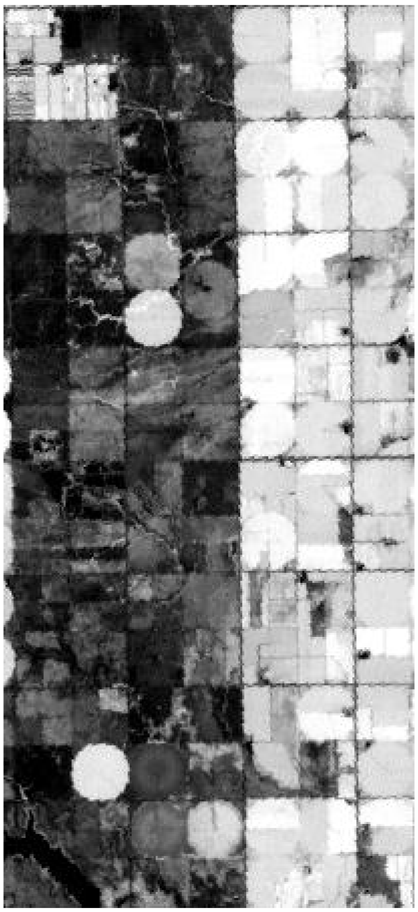
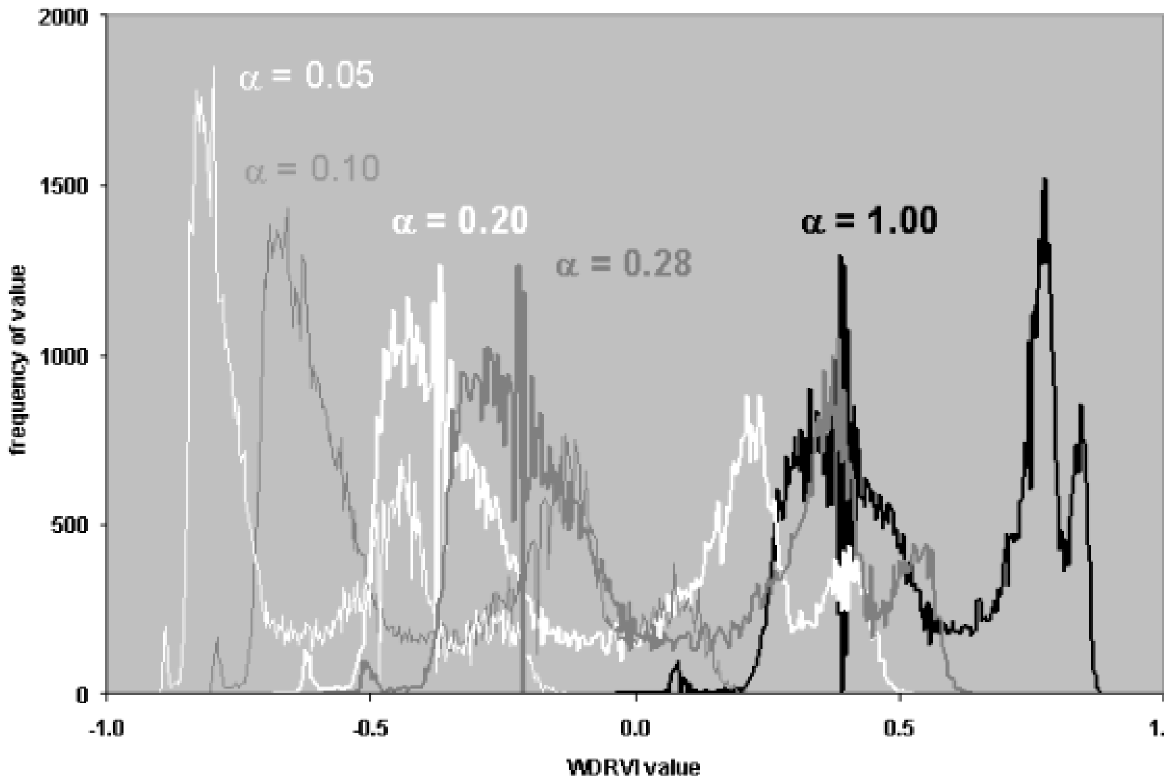


Figure 2. NDVI calculated from a Landsat ETM+ image (P29, R32) acquired August 4, 2001. Location is at the edge of a research farm near Hastings, NE. Brighter tones indicate higher fAPAR. Circles are quarter-section (160 acre; 65 ha) fields irrigated by center pivot.

Our scene had an average red reflectance of 7.7% and maximum near infrared reflectance of 54.9%, thus the α_{est} equaled 0.28. Figure 3 shows the histograms that result from calculating the WDRVI with different α values, and Table 1 provides a statistical summary of these distributions.


α = 0.05
α = 0.10
α = 0.20
α = 0.28
α = 1.00
frequency of value
2000
1500
1000
500
0
-1.0
-0.5
0.0
0.5
1.0
WDRVI value

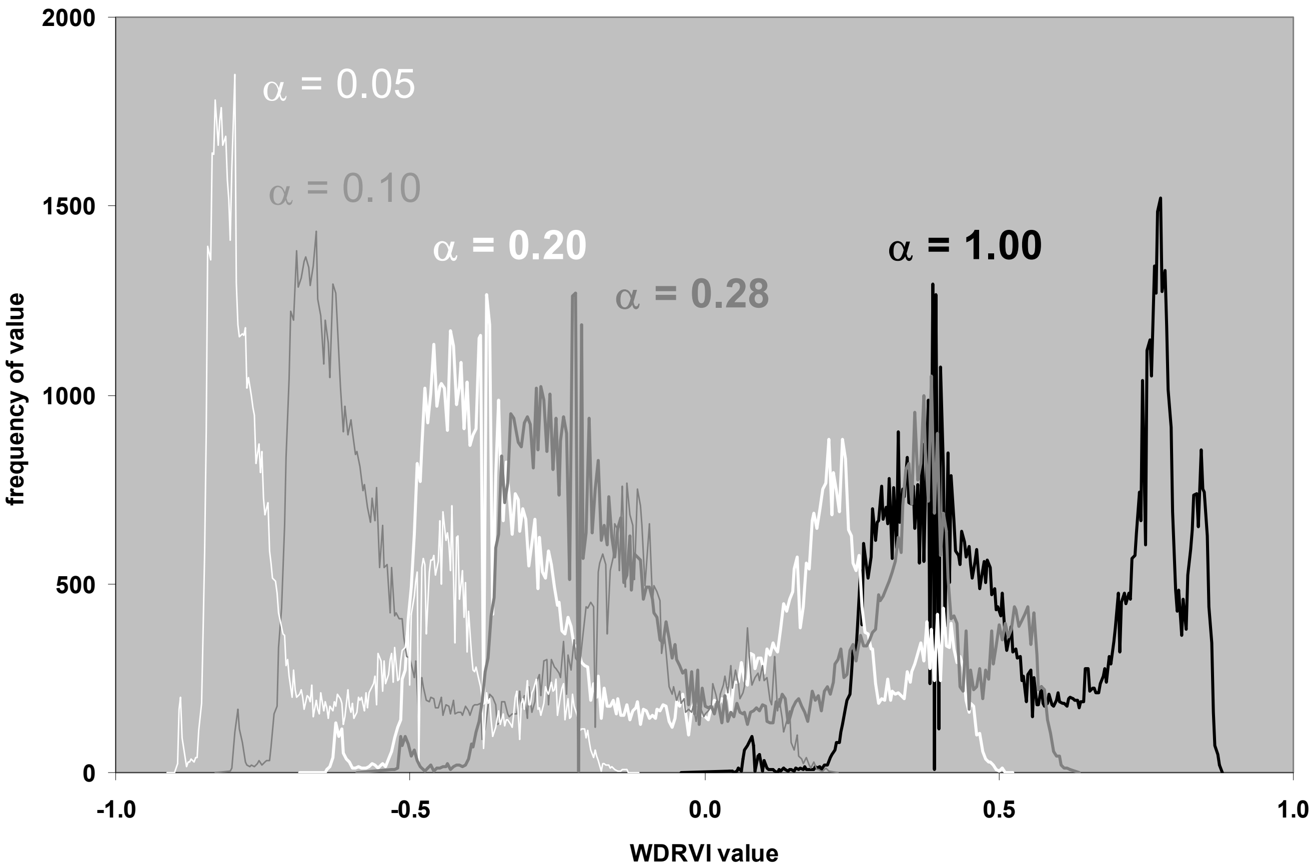


Figure 3. Histograms of WDRVI obtained for different values of α.

Table 1. Summary statistics for WDRVI calculated with various values for α.

Coefficient Value	Mean	Maximum	Minimum	Range	Change in Range over the NDVI
α = 1.00	0.553	0.883	−0.040	0.923	--
α = 0.28	0.037	0.635	−0.589	1.224	+33%
α = 0.20	−0.115	0.524	−0.688	1.212	+31%
α = 0.10	−0.406	0.231	−0.831	1.062	+15%
α = 0.05	−0.636	−0.110	−0.912	0.802	-13%

Discussion

It can be seen that the shape of the distributions changes significantly with change in α; in particular, the two modes at high NDVI values spread out as α decreases (Figure 3). However, the cost of enhanced dynamic range at the high end is some loss of sensitivity at the low end. Notice the contraction of the small mode at the low end as α decreases (Figure 3). Attenuation of the near infrared reflectance can increase the dynamic range of the WDRVI over the NDVI: α =

0.20 yields more than 30% increase in dynamic range (Table 1). Notice that $\alpha_{est} = 0.28$ gives a slight improvement in dynamic range over $\alpha = 0.20$, but tuning the coefficient value to particular scene characteristics could impair scene mosaicking and temporal comparisons. We suggest that since $\alpha = 0.20$ has been shown to be effective with proximal sensors (Gitelson 2004) as well as with AVHRR (Viña et al. 2004) and Landsat ETM+ (this note) imagery in the absence of atmospheric correction, it is a good initial value from which to explore the potential of the WDRVI in revealing more variation in settings with moderate to high green LAI.

Conclusions

1. The WDRVI offers a simple way to enhance dynamic range that is limited by the NDVI under conditions of moderate to high biomass (LAI > 2).
2. Tuning the weighting parameter α to different values changes histogram shape.
3. A coefficient value of 0.20 for α appears to be generally effective.
4. For low biomass settings (LAI < 1), the NDVI still works best for distinguishing vegetation.

Acknowledgements

This work was supported in part through the GAP Research Project *Regionalization of disparate land cover maps using image time series* as well as by a NSF Biodiversity and Ecosystem Informatics (BDEI) grant to Henebry and a NASA Land Cover Land Use Change (LCLUC) grant to Gitelson and Henebry.

Literature Cited

Gitelson, A.A. 2004. Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. *Journal of Plant Physiology* 161:165-173.

Viña, A., G.M. Henebry, and A.A. Gitelson. 2004. Satellite monitoring of vegetation dynamics: Sensitivity enhancement by the Wide Dynamic Range Vegetation Index. *Geophysical Research Letters* 31 (4) L04503. doi:10.1029/2003GL019034.

Digital Aerial Photograph Interpretation: Examples and Techniques from Arizona and the Southwest Regional Gap Analysis Program

KEITH POHS AND KATHRYN THOMAS

U. S. Geological Survey, Southwest Biological Science Center, Flagstaff, Arizona

Introduction

The Southwest Regional Gap Analysis Program (SWReGAP) is developing land cover maps using a biophysical modeling procedure that incorporates satellite imagery, maps of environmental variables, and extensive reference observations of vegetation types as model input data. Field crews have collected these reference observations throughout the five-state SWReGAP region—Arizona, Colorado, Nevada, New Mexico, and Utah. However, field investigation sometimes is not possible or is prohibitively costly. The former is due largely to limited access to certain lands. The latter occurs in large roadless areas where access is mainly by foot, often in extremely rugged terrain, and field points are less efficiently obtained. Since the biophysical modeling approach is most effective where there are sufficient observations for each vegetation type and adequate geographical representation across its occurrence, in Arizona we have used a method of digital aerial photograph interpretation to collect additional observation points. Using Digital Ortho Quarter Quadrangles (DOQQs) downloaded from the Arizona Regional Image Archive (ARIA, http://aria.arizona.edu/index.html) Web site or Digital Ortho Quadrangles (DOQs) interpreted on the TerraServer Web site (terraserver-usa.com), the Arizona team was able to delineate over a thousand additional vegetation observation points and polygons, classified to ecological systems as developed by NatureServe. In this paper we present the methods used to obtain the vegetation observations using these digital sources and the limitations and advantages of this methodology for regional mapping of land cover.

Methods

Two methods involving the use of DOQQs or DOQs were developed, depending on whether the imagery was accessed from the ARIA Web site or the TerraServer Web site. Initially DOQQs were downloaded from ARIA for various parts of Arizona. Images from 1992, 1996, and 1997 were selected by quad sheet name, based upon a GIS grid layer of quads for Arizona, and then systematically added as a layer into ArcMap. A land cover analyst was able to distinguish a number of ecological systems and delineate them as polygons. Initially we needed additional ground reference observations for a large area of the Sonoran Desert where field sampling was not possible due to access restrictions. Several ecological systems were discernible, including Sonoran-Mojave Creosote-White Bursage Desert Scrub and Sonoran Palo Verde-Mixed Cacti Desert Scrub.

Some ecological systems seemed underrepresented in our field observations. This was the case for a number of different reasons, but one was the placement of roads (our primary field sampling corridors) in areas of gentler topography relative to the surrounding landscape. This was apparent on the Colorado Plateau where travel routes simply avoid the steep slopes and bedrock expanses of the Colorado Plateau Mixed Bedrock Canyon and Tableland ecological system. These predominantly barren features, however, are readily discernible in aerial photography, and we obtained adequate representation of this type through interpretation of DOQQs. We also obtained reference observations for dunes, playas, cinder cones, and lava flows in this region and throughout the state. All of these features and their associated sparse vegetation are classified into a described NatureServe ecological system. Other ecological systems were digitized based primarily on vegetative cover and included Rocky Mountain Aspen Forest and Woodland and Rocky Mountain Gambel Oak-Mixed Montane Shrubland.

The ARIA site had considerable DOQQ coverage for much of Arizona except for the central part of the state. An alternative source of remotely sensed data was sought for these areas as well as areas in adjacent states for which the Arizona team has mapping responsibilities. This led to the development of a second methodology using imagery available on the TerraServer Web site.

DOQs for all of Arizona and adjacent study areas of Utah and New Mexico are available on the TerraServer Web site. The site hosts USGS aerial imagery from 1997 and scanned images of USGS topographic maps from various years. The team's land cover analyst obtained reference observations by navigating to a particular region or feature on the topographic maps and then switching to the aerial photograph for that site using the built-in features of TerraServer (Figure 1a).

Figure 1a. Interpretation of the Invasive Southwest Riparian Woodland and Shrubland Ecological System along the Little Colorado River, Arizona, using TerraServer. A USGS digital topographic map of this same view is a related link on the toolbar to the left of the image.

Alternatively, the analyst opportunistically scanned areas for discernable systems using the DOQs and then found the location of the type on the Web site's digital topographic maps. TerraServer does not readily support systematic downloading of DOQ images and thus digitization in ArcMap nor does it provide sufficiently accurate geographic coordinates for locations of interest. We determined the geographic coordinates for identified ecological systems by determining the center of the system on the Web site's digital topographic map and then finding the same location on a state National Geographic Digital USGS Topographic Map™ (National Geographic Maps, San Francisco). Using the navigational TOPO! Software™ and the compass tool that is part of this product, the UTM coordinates for the interpreted sites were obtained (Figure 1b).

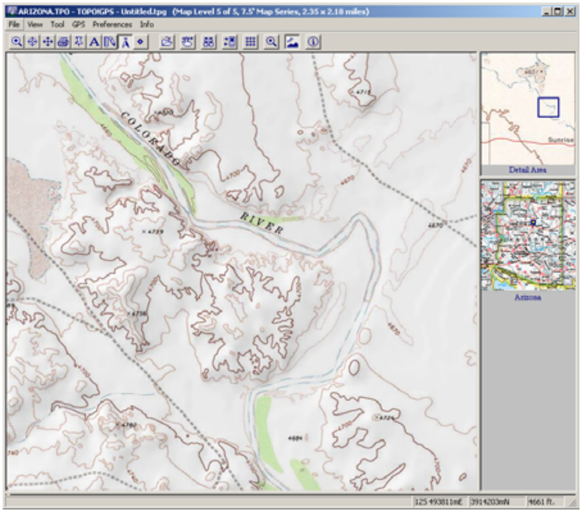


Figure 1b. National Geographic digital topographic map of same site in TOPO! view. Using the available compass tool, a UTM coordinate can be obtained.

We were able to develop several hundred ecological system points using this methodology and state National Geographic Digital Topographic Maps for Arizona, Utah, and New Mexico. In Arizona, data was obtained for a number of higher elevation areas that had not been sampled on the ground. This technique was also used in the roadless Gila and Aldo Leopold Wilderness areas in southwestern New Mexico. The most common ecological systems delineated in these regions included aspen forests and Gambel oak shrublands, as the dominant deciduous species were readily discernible from nearby coniferous systems.

The preponderance of private land along stream and river corridors, the occasionally steep topographic relief of canyon environments inhibiting foot and automobile travel, and in some areas wilderness designation can make the efficient collection of field points in riparian areas exceedingly difficult. The TerraServer remote sensing methodology again allows for some interpretation of vegetation communities in these otherwise inaccessible but ecologically important areas. Several NatureServe ecological systems are discernible, including Invasive Southwest Riparian Woodland and Shrubland (Figure 1a) and Rocky Mountain Lower Montane Riparian Woodland and Shrubland.

In some instances more points of a particular ecological system in a mapping area are needed for modeling than were obtained in the field. One working unit comprising much of northeastern Arizona and adjacent New Mexico, aspen forests, though locally common above 8,500 feet along the length of the Chuska Mountains did not appear to be sufficiently well represented by field observations to be mapped. Using the TerraServer methodology, additional points were obtained, and this important system may now be mapped.

Discussion

Use of the Web-based imagery allowed the Arizona SWReGAP team to acquire reference data that otherwise would simply not be available or available at a far greater cost than feasible for the project. It allowed the collection of data in areas where access was not possible on the ground. It also allowed us to increase the geographic representation of our reference observations, augmenting those collected in the field. The imagery used for the work is available free of charge on the Internet, and the only material cost was the purchase of commercial digital topographic maps on CD-ROM for two states, Arizona and New Mexico. The major limitation of both methodologies is that only some ecological systems are discernible given the resolution of the imagery used. Most of the higher elevation coniferous systems, such as spruce-fir and limber-bristlecone pine, cannot be easily interpreted. Ponderosa pine and pinyon-juniper systems can be interpreted but were not a priority for this work, as a considerable number of field observations had already been obtained. The technique does require familiarity with the ecological systems being mapped and their expression on the landscape. In our case, the land cover analyst doing the photointerpretation had spent extensive time on the ground acquiring reference data and was able to directly use this acquired knowledge in the interpretive work. In addition, the analyst identified known field reference data observations of ecological systems that were targeted for photointerpretation on the digital imagery to verify his interpretation.

The digitized polygons of this work can be overlaid on the developing map, or point locations can be extracted from within the polygons for inclusion in the training data, along with other photo-interpreted points. These and other training data are then used to produce the new GAP land cover map for Arizona and portions of adjacent states.

Hierarchical Land Cover Classification for Hawaii

STEVEN HOCHART[1,2], DAN DORFMAN[1,3], SAMUEL GON III[4], AND DWIGHT MATSUWAKI[1]
[1]Hawaii Gap Analysis Program, University of Hawaii, Honolulu
[2]USInfrastructure-Hawaii, Honolulu
[3]The Nature Conservancy, Marine Initiative, Santa Cruz, California
[4]The Nature Conservancy, Honolulu, Hawaii

Introduction

The state Gap Analysis projects each have developed approaches to image interpretation. In Hawaii, most of the land area has been surveyed at some time, and detailed vegetation classifications are available for more than a dozen significant areas covering portions of each island.

GAP led the development of the Multi-Resolution Land Cover Consortium (MRLC), and most GAP state projects enjoy access to the MRLC archive (Hegge et al. 2001). In most cases this includes Landsat Thematic Mapper 7 images available for each path/row representing three seasons. Under the MRLC these images are preprocessed to standardize geographic location as well as correct for terrain displacement and atmospheric reflection. Additionally, the EROS Data Center now makes MRCL images available that have been corrected to "at-satellite radiance values" (Homer and Hegge, EROS Data Center/Raytheon). The process also employs the Sun-Earth and Earth-radiometer distances at the time the image was taken to compensate for the radiometric distortion effects of the Earth's atmosphere, making images taken on different revolutions more comparable. The Hawaiian entry for MRLC "at-satellite radiance" images had not been populated prior to the HI-GAP effort, and we were able to partner with EROS Data Center to select and process scenes for each path/row representing seasonality as well as completing a cloud-free mosaic using Landsat TM images from 12/99-12/02. These scenes represent a consistent data set on which the HI-GAP spectral decision tree classification was implemented.

Classification Research

Several image interpretation methods were tested for the HI-GAP application. Classification and Regression Tree Analysis was considered for its objectivity and statistical strength (De'ath and Fabricius 2000, Hansen et al. 1996, Lawrence and Wright 2001), but this approach requires a significant investment in field data collection, and the majority of the land area in Hawaii has been previously surveyed. Hawaiian vegetation systems are relatively well-studied and have been mapped and classified several times previously. Detailed vegetation maps are available for significant portions of many of the Hawaiian Islands. Additionally, previous and concurrent land cover research has led to a significant spectral signature library for Hawaiian native and invasive vegetation types. Research at The University of California, Santa Barbara has focused on employing aspects of spectral signatures to perform classifications on AVRIS hyperspectral imagery (Roberts et al. 1998, Serrano et al. 2000). But AVRIS imagery is not available statewide, and Landsat 7 TM data does not have sufficient spectral resolution to enable a library/signature-based approach under the conditions and needs of HI-GAP. A more ecologically driven classification approach has been developed at Duke University's Nicholas School of the Environment, where radiometric enhancements are employed to enable classification based on "natural" variables such as level of vegetation or soil exposure (Khorram et al. 1992). Also, three recent case studies developed for mapping impervious surfaces from Landsat 7 ETM were consulted for their possible applicability in land cover mapping approaches for HI-GAP (Yang et al., in review).

Classification Methodology

After extensive research on different methods for land cover classification in Hawaii, the HI-GAP team chose to employ an ecologically driven spectral decision tree approach to land cover classification. The approach is based on the application of ERDAS Imagine's Knowledge Engineer software platform. Knowledge Engineer was selected because it provided the hierarchical structure to perform image classification and offered a good platform for storing and analyzing spectral properties. Knowledge Engineer files were developed for each image and then integrated into a central Knowledge Engineer to produce the final classification...........

The first stage in this process is removing the ocean and clouds by masking the image. The remaining areas of water are the first branch of our decision tree classification. These areas have strong absorption of near infrared light and therefore very low values in Landsat band 4. We are able to use this “natural” or “ecological” property to form the basis of a decision. If the values recorded for band 4 are below a defined cutoff point, then we expect those cells to represent standing water and classify those areas accordingly. In addition we are able to clearly identify areas of industrial or urban land cover as having very high reflectance in certain raw bands and can build this principle into a spectral decision tree classification.

Many of the vegetation types in Hawaiian forests cannot be distinguished clearly from the information available in raw TM bands for a variety of reasons, ranging from complex topography to small-scale mosaics of adjacent vegetation types within a limited geographic area. We employ two techniques to address this natural complexity. First, vegetation is known to have a low reflectance in Landsat band 3 (0.63-0.69 nm) and a high reflectance in Landsat band 4 (0.76-0.90 nm). As a result, vegetation indices have been designed to isolate this spectral feature and distinguish the amount of vegetation in an area. We use the standard Normalized Difference Vegetation Index (NDVI) to build the first branch in our decision tree separating areas of high biomass from areas of low biomass as indicated on the left in Figure 1. Using treatments such as the NDVI, Principal Component Analysis (PCA), and Tasseled Cap, we are able to find “cut points” or variables at which we can build branches for our classification tree illustrated in Figure 1.

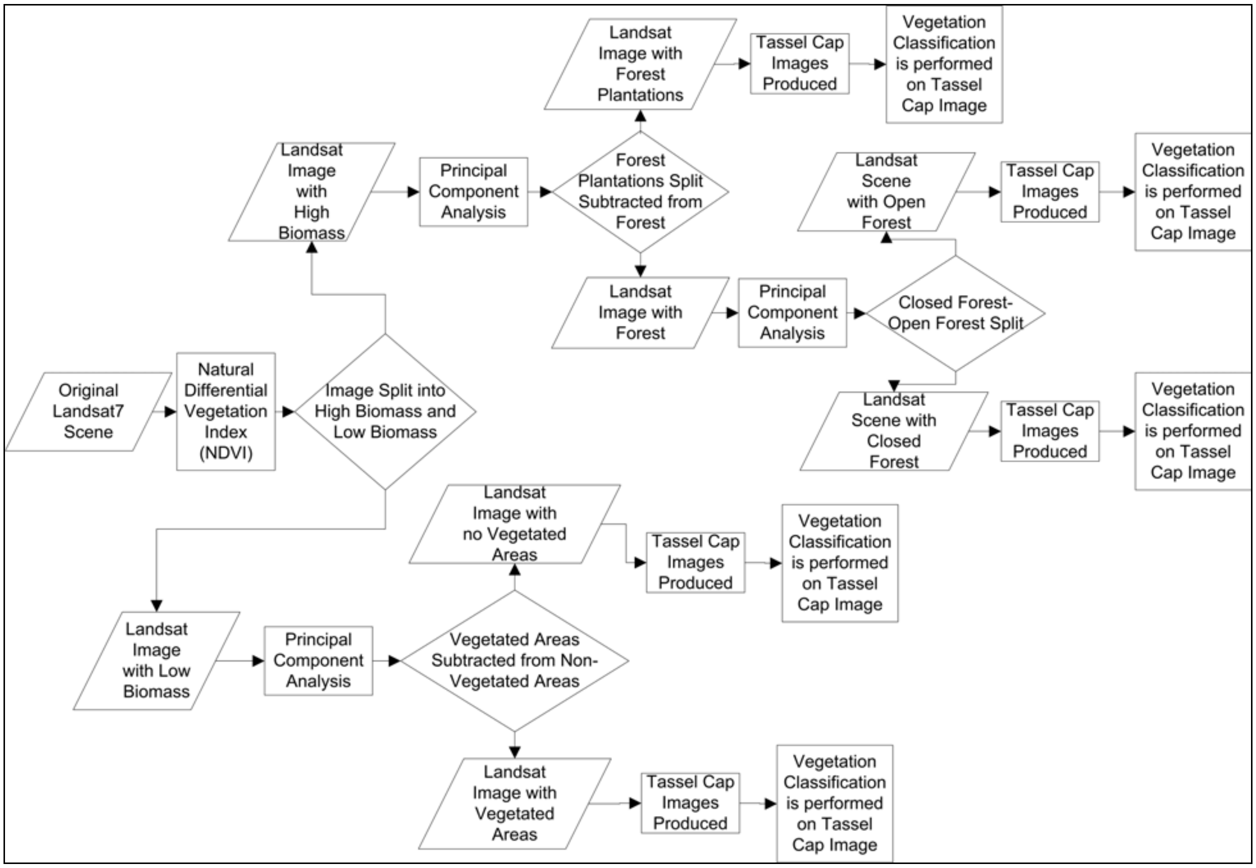


Figure 1. Image Classification Decision Tree. Landsat images used are shown in rhomboid boxes, treatments applied are described in rectangular boxes, and splits into branches are in diamond-shaped boxes.

The tasseled cap treatment is particularly valuable in the Hawaiian High Islands ecosystem because of its utility in revealing brightness, greenness, and wetness. These variables are strong identifiers for Hawaiian vegetation communities, since their distributions are closely tied to moisture availability, exposure, and nutrient availability (Pratt and Gon 1998, Wagner et al. 1999).

Establishing the spectral decision tree within the Knowledge Engineer in one area enables the analyst to test and refine the decision trees in similar areas. When applying a decision tree to a new scene, it has often been found that only minor adjustments are needed to apply it in different places—or the same place under different conditions. However, when scenes are used from different seasons, we find significant adjustments are required in the classification, particularly in areas of grasslands and invasive shrubs, where changes in greenness due to "green-up" phenology are substantial. Being aware of seasonal variation and its effects on particular vegetation types enabled the HI-GAP team to adapt the classification to take advantage of seasonality differences in the tropics.

Results

The spectral decision tree classification has been implemented on the Big Island of Hawaii, Maui, Lanai, and Molokai with positive results. The Big Island of Hawaii was the first island that was classified using this methodology, because it provided a wide variety of vegetation types with which to develop the decision tree methodology. The results for the southern forested regions of the Big Island were assessed using field point data gathered during helicopter surveys. The preliminary results indicate a 90% accuracy level for this region. Spectral properties for specific vegetation types were taken from the Big Island and applied to Maui and Molokai. The spectral values needed small adjustments to achieve the desired results. Results from the first draft have been reviewed and approved by various partners. Based on the work from the first draft on Maui and Molokai it is clear that spectral values gathered from the Big Island can be applied to other islands as initial hypotheses and then adjusted according to ancillary data and expert knowledge.

The minimum mapping unit for the HI-GAP project is 90 m^2. The methodology described above has provided accurate results at this scale where it has been tested. The same methodology was recently tested on a small area of vegetation on the Big Island of Hawaii. The results from this test indicate the methodology can be applied at higher resolutions than 90 m^2. The results from this study produced a detailed vegetation map with 30-meter pixel resolution. The significance of these results indicates the methodology being developed can be applied to small areas and is capable of producing detailed vegetation classifications of these areas.

Ecological Significance

Establishing a model of initially large, then progressively refined vegetation classes is an approach that matches the current hierarchical vegetation classification system developed for the Hawaiian Islands. Broad elevation, moisture, and physiognomic categories are initially established, within which canopy dominants are used to identify alliances and associations (Pratt and Gon 1998). This allows for specific analysis to be confined to geographic subregions on the basis of elevation as well as moisture, stratifying and thereby separating complex signature sets that might otherwise be indistinguishable on signature alone. Several such broad subdivisions can be readily defined and are relevant to ecological considerations such as species ranges and ecophysiology along gradients. Typical examples are wet windward vs. dry leeward, coastal/lowland vs. montane, and closed vs. open/sparse canopy variants of a given dominant canopy species.

A broader geological age gradient is also apparent across the archipelago from the youngest island, Hawaii (less than half a million years) to the oldest of the main islands, Kauai, at the opposite end of the chain (over 5 million years). There are also well-documented floristic differences among the islands due to their isolation from one another, which further complicates vegetation classification. This isolation factor results in the need for adjustments in decision criteria for spectral signatures appropriate to one island when applied to another. At the same time, key dominant species in the canopy are consistent from one end of the archipelago to the other, so an initial hypothesis of applicability of decision criteria from an island to a neighboring island may be valid but requires testing and adjustment for each island.

Finally, although there are over 150 described vegetation associations in the Hawaiian Islands, their distribution in space is strongly correlated with elevation, moisture, and soil, so that in a given class of elevation, moisture, and formation there is a manageably small set of associations that are typically present, especially in landscapes dominated primarily by native vegetation types. The situation is greatly complicated in alien-dominated or mixed native-alien vegetation, where an unstable disturbance-driven mosaic greatly increases the number of possible vegetation associations that can be present.

Conclusion

The Knowledge Engineer approach to developing a classification tree enabled the HI-GAP team to readily test refinements and alternative classification decisions and to analyze the effects of alternative approaches on results. It employs the objectivity and repeatability of CART but augments statistical outcome with the strength of local knowledge and the ability to refine and adapt the decision tree as information becomes available. Results to date indicate the methodology described here provides a new approach to mapping tropical land cover. Where available knowledge is limited, the upper levels of the classification hierarchy can be initially characterized, and as ground-truthing or surveys provide more extensive knowledge on the vegetation composition in the study region, the decision criteria can be further refined without discarding previous work, greatly enhancing the utility of foundation efforts.

Literature Cited

De'ath, G., and K.E. Fabricius. 2000. Classification and Regression Trees: A powerful yet simple technique for ecological data analysis. *Ecology* 81:3178-3192.

Hansen, M., R. Dubayah, and R. Defries. 1996. Classification Trees: An alternative to traditional land cover classifiers. *International Journal of Remote Sensing* 17:1075-1081.

Hegge, K. 2001. Multi-Resolution Landcover Consortium Handbook. EROS Data Center, U.S. Geological Survey, Sioux Falls, South Dakota.

Khorram, S., K. Siderelis, H. Cheshire, and Z. Nagy. 1992. Mapping and GIS development of land use and land cover categories for the Albemarle-Pamlico drainage basin. U.S. Environmental Protection Agency Project # 91-08. March 1992. Raleigh, North Carolina.

Lawrence, R.L., and A. Wright. 2001. Rule-based classification systems using Classification and Regression Tree (CART) Analysis. *Photogrammetric Engineering and Remote Sensing* 67:1137-1142.

Pratt, L.W., and S.M. Gon III. 1998. Terrestrial ecosystems. In S.P. Juvik and J.O. Juvik, editors. Atlas of Hawaii, 3rd ed. University of Hawaii Press, Honolulu.

Roberts, D., M. Gardner, R. Church, S. Ustin, G. Scheer, and R.O. Green. 1998. Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of the Environment* 65:267-279.

Serrano, L., S. Ustin, D. Roberts, J. Gamon, and J. Penuelas. 2000. Deriving water content of chaparral vegetation from AVIRIS data. *Remote Sensing of the Environment* 74:570-581.

Wagner, W.L., D.R. Herbst, and S.H. Sohmer. 1999. Manual of the flowering plants of Hawaii. B.P. Bishop Museum, Special Publication 97. University of Hawaii Press and Bishop Museum Press, Honolulu.

Yang, L., C. Huang, C. Homer, B. Wylie, and M. Coan. In review. An approach for mapping large-area impervious surfaces: Synergistic use of Landsat 7 ETM+ and high spatial resolution imagery. *Canadian Journal of Remote Sensing.*

Acknowledgements

Colin Homer, Kent Hegge, Dar Roberts, Pat Halpin, and Shannon McElvaney.

ANIMAL MODELING

Accuracy Assessment for Range Distributions of Terrestrial Vertebrates Modeled From Species Occurrences and Landscape Variables

GEOFFREY M. HENEBRY, BRIAN C. PUTZ, MILDA R. VAITKUS, AND JAMES W. MERCHANT
Center for Advanced Land Management Information Technologies (CALMIT), School of Natural Resources, University of Nebraska-Lincoln

Introduction

To produce its wildlife-habitat relationship models, the Nebraska Gap Analysis Program (NE-GAP) used recursive partitioning applied to species occurrence data (in the form of museum voucher specimens or curated surveys) and a geodatabase of landscape variables on a hexagonal grid with a nominal spatial resolution of 40 km^2 (Henebry et al. 2001, Holland et al. 2002). Here we describe our approach to accuracy assessment for modeled range distributions.

Methods

To generate the habitat models we used QUEST (Quick, Unbiased, & Efficient Statistical Trees; Loh and Shih 1997), a recursive partitioning algorithm similar to CART (Classification & Regression Trees; Breiman et al. 1984). QUEST has several advantages for habitat modeling: it is much faster than CART, variable selection is unbiased, it handles categorical predictor variables with many categories, and uses automated cross-validation (De'ath and Fabricius 2000, Shih 2002). The motivation for using this strategy is twofold. Not only are the resulting trees of decision points and values that form the models understandable, debatable, and tunable, the nonparametric modeling can handle the multimodality likely to be found in species occurrence data.

The suite of environmental variables (land cover, climate, soils, terrain) included in the modeling process are described in Henebry et al. (2001). Modeling was performed across a hexagonal grid produced by the EPA EMAP program with a cell resolution of about 40 km^2 within Nebraska. Each variable was rescaled from its raster resolution (900 m^2 for land cover, soils, and terrain data and 2.25 km^2 for climate variables) to the coarser hexagonal coverage. All environmental variables contained within the hexagons that intersected BBS routes or CBC circles were associated with the species occurrence data at those sampling locations. Continuous variables were rescaled by area-weighted averaging. Categorical variables were represented as a compositional vector.

Species occurrence data were gathered from route-level composites of the USGS Breeding Bird Survey (BBS; www.pwrc.usgs.gov/bbs) and circle composites of the National Audubon Society's Christmas Bird Count (CBC; www.audubon.org/bird/cbc/) for the period 1970-2000. Given the intensive repeated observations, if a species was not reported along a sampling unit during the study period, it was considered absent. However, it is important to distinguish this inference of absence that is accepted only after many years of observation from an observed absence in a particular year. The use of these absences is different in kind: the former can be used in model construction but the latter is not reliable for accuracy assessment.

Occurrence data and associated environmental variables for each species were submitted to QUEST. Resulting statistical trees were trimmed or pruned interactively by querying the hexagonal coverage of environmental variables to evaluate the sensitivity of the tree splits and assess model generality. The final tree served as the wildlife-habitat relationship model. Using the threshold values of the environmental variables selected in the final model, the geodatabase was queried to produce each species' predicted habitat distribution.

For those species lacking sufficient occurrence data (including all mammals, many birds, and a few reptiles and amphibians), the literature was consulted to identify specific environmental variables that could be used for habitat surrogates. The identified variables were then queried to the geodatabase. The predicted range was assessed visually against the reported range and, if there was a large discrepancy, different variables or variable thresholds were tested. Fitness for both model types was evaluated in two ways: the proportion of the occurrences explained and the visual appearance of the predicted range distribution. Parameter nudging was employed to assess, albeit informally, the sensitivity of specified values to range extent.

Accuracy assessment of the range distributions relied on independent species occurrence data. These independent data had various sources. Literature-based mammal models (n=78) were evaluated against georeferenced voucher specimens collected since 1970 (1,805 unique observations) in the Nebraska State Museum (NSM) and evaluated again at the county level (794 unique observations). The reptile and amphibian models (n=62) were evaluated against voucher specimens collected since 1970 (357 unique observations) in museums other than NSM. The BBS models (n=192) were evaluated against the BBS route level summaries for 2001 and 2002 (1,953 unique observations) and separately evaluated against voucher specimens collected since 1970 from NSM and other museums (733 unique observations).

We focused on rates of omission error because the occurrence data are strictly presence-only. Accuracy assessment was performed at two scales of model representation: the modeling resolution of 40 km^2 hexagons and the reporting resolution of 640 km^2 hexagons. Occurrence data were represented—depending on data source—as county, route, or hexagon. All museum voucher data were scaled to the county level, as it was the only consistent spatial information for many specimens, especially for data from museums other than NSM. The NSM mammal occurrence data that were georeferenced were evaluated as hexagons. The 2001 and 2002 BBS data were evaluated as routes, i.e., the composite of hexagons intersected by the survey route.

We used two different *de minimis* thresholds in the accuracy assessments. We required a "presence" in at least one spatial unit associated with the underlying data (e.g., BBS route, county) for all occurrence data except the georeferenced mammal voucher specimens from NSM. To make those point data comparable to the other data, we first "promoted" the point occurrences to the model hexagon level and required at least five "presence" modeling hexagons to qualify for accuracy assessment. To avoid inflating accuracies, assessments of omission error excluded species with statewide distributions.

Results

The median (mean) omission error rate for the BBS models was 0% (7%), with 90%, 87%, and 80% of the models having omission error rates less than 15%, 10%, and 5% respectively (Figure 1). The median (mean) omission error rate for the reptile and amphibian models was 0% (4%), with 93% of the models having omission error rates less than 5%. The median (mean) omission error rate for the mammal models was 14% (20%), with 55%, 39%, and 21% of the models having omission error rates less than 15%, 10%, and 5%, respectively. Rescaling from the modeling grid to the reporting hexagons (640 km^2) significantly decreased the omission error rates, as expected (Table 1).

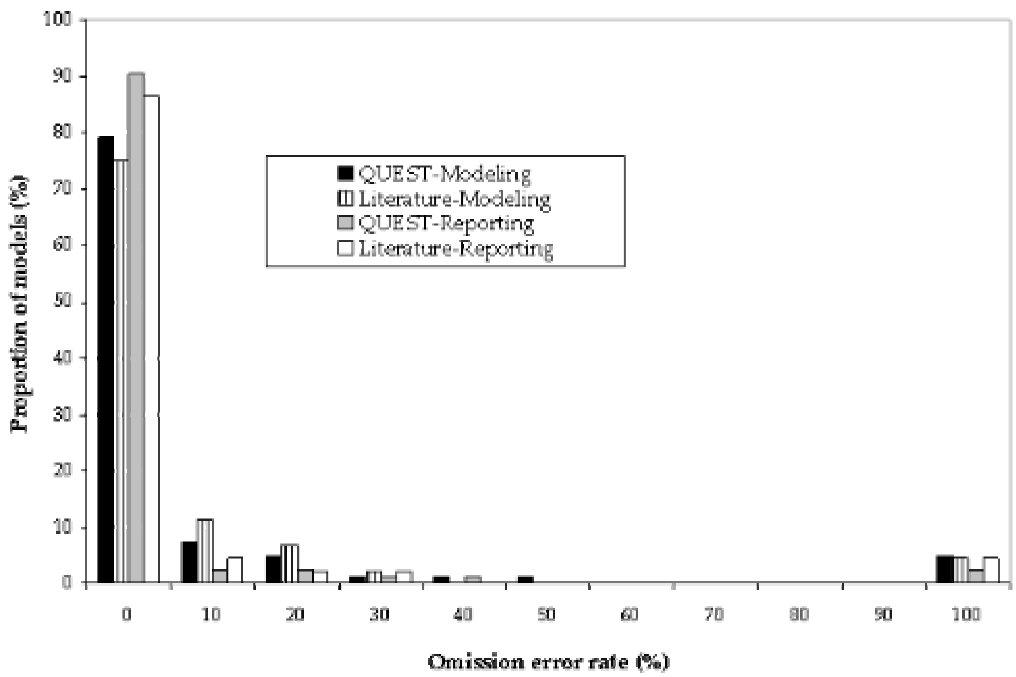


Figure 1. Distribution of omission error rates in QUEST (n=82) and literature (n=44) models assessed using Breeding Bird Survey route level summaries from 2001 and 2002 in modeling (40 km^2) and reporting (640 km^2) EMAP hexagons.

Table 1. Summary of omission error rates for models across taxa, modeling methodology, and spatial scale of accuracy assessment. Percentages in table refer to the proportion of models with omission error rates of less than either 10% or 20%.

Taxon Method Scale	Modeling Hexagons Omission Error Rate <10% | <20%	Reporting Hexagons Omission Error Rate <10% | <20%	No. of Excluded Statewide Species
Birds			
QUEST			
BBS	87% 91%	93% 95%	
County	60% 63%	64% 65%	35
Literature			
BBS	86% 93%	91% 93%	
County	74% 74%	76% 78%	
Reptiles & Amphibians			
QUEST			
County	95% 95%	95% 95%	
Literature			3
County	88% 88%	88% 88%	
Mammals			
Literature			
NSM hexagon	40% 70%	82% 89%	27
County	89% 91%	91% 93%	

Discussion

While the results of the accuracy assessment are encouraging for most species, there remain several challenges to performing accuracy assessment of habitat models using occurrence data. First, few voucher specimens are geospatial data, and the county-level resolution of the older vouchers is so much coarser than the modeling resolution that the results must be interpreted with caution. Second, BBS survey data are temporally sparse for many species, and route level summaries across 30 years may collapse significant trends or fluctuations in population dynamics (cf. Vaitkus et al. 2003, this issue). Third, the discrepancy between the finer spatial resolution of the modeling hexagons and the coarser resolution of the reporting hexagons translates into reduced omission error rates and inflated model accuracy at the reporting resolution. Fourth, a full accuracy assessment requires that commission error be considered, but the lack of true absence data means than neither commission nor correct absence frequencies can be calculated. Fifth, the geographic sampling bias in common among voucher specimens limits the reliability of even omission error rate estimates. We interpret the higher omission error rates for the mammal models as a strong artifact of the clustering of specimens from four locales used for class collecting trips over the years.

Conclusions

1. The results suggest that the habitat modeling approach using recursive partitioning has advantages over literature gestalt in producing range distributions with lower rates of omission error.

2. Developing habitat models using statistical trees generated from species occurrence data and environmental variables can lend a greater degree of objectivity to the modeling process, but there is still considerable subjectivity in the pruning stage that is necessary for model generality.

3. There remain significant methodological challenges to accuracy assessment of the predictions of wildlife-habitat relationship models.

Literature Cited

Henebry, G.M., B.C. Putz, and J.W. Merchant. 2001. Modeling reptile and amphibian range distributions from species occurrences and landscape variables. *Gap Analysis Bulletin* 10:22-24.

Holland, A.K., G.M. Henebry, B.C. Putz, M.R. Vaitkus, and J.W. Merchant. 2002. Modeling avian habitat from species occurrence data and environmental variables: Assessing the effects of land cover and landscape pattern. *Gap Analysis Bulletin* 11:25-27.

Loh, W.-Y., and Y.-S. Shih. 1997. Split selection methods for classification trees. *Statistica Sinica* 7:815-840.

Shih, Y.-S. 2002. QUEST User Manual. Department of Mathematics, National Chung Cheng University, Taiwan. April 17, 2002.

Vaitkus, M.R., G.M. Henebry, B.C. Putz, and J.W. Merchant. 2003. Evaluating the use of statistical decision trees for modeling avian habitats and regional range distributions in the Great Plains. *Gap Analysis Bulletin* 12:36-40.

Evaluating the Use of Statistical Decision Trees for Modeling Avian Habitats and Regional Range Distributions in the Great Plains

MILDA R. VAITKUS, GEOFFREY M. HENEBRY, BRIAN C. PUTZ, AND JAMES W. MERCHANT

Center for Advanced Land Management Information Technologies (CALMIT), School of Natural Resources, University of Nebraska, Lincoln

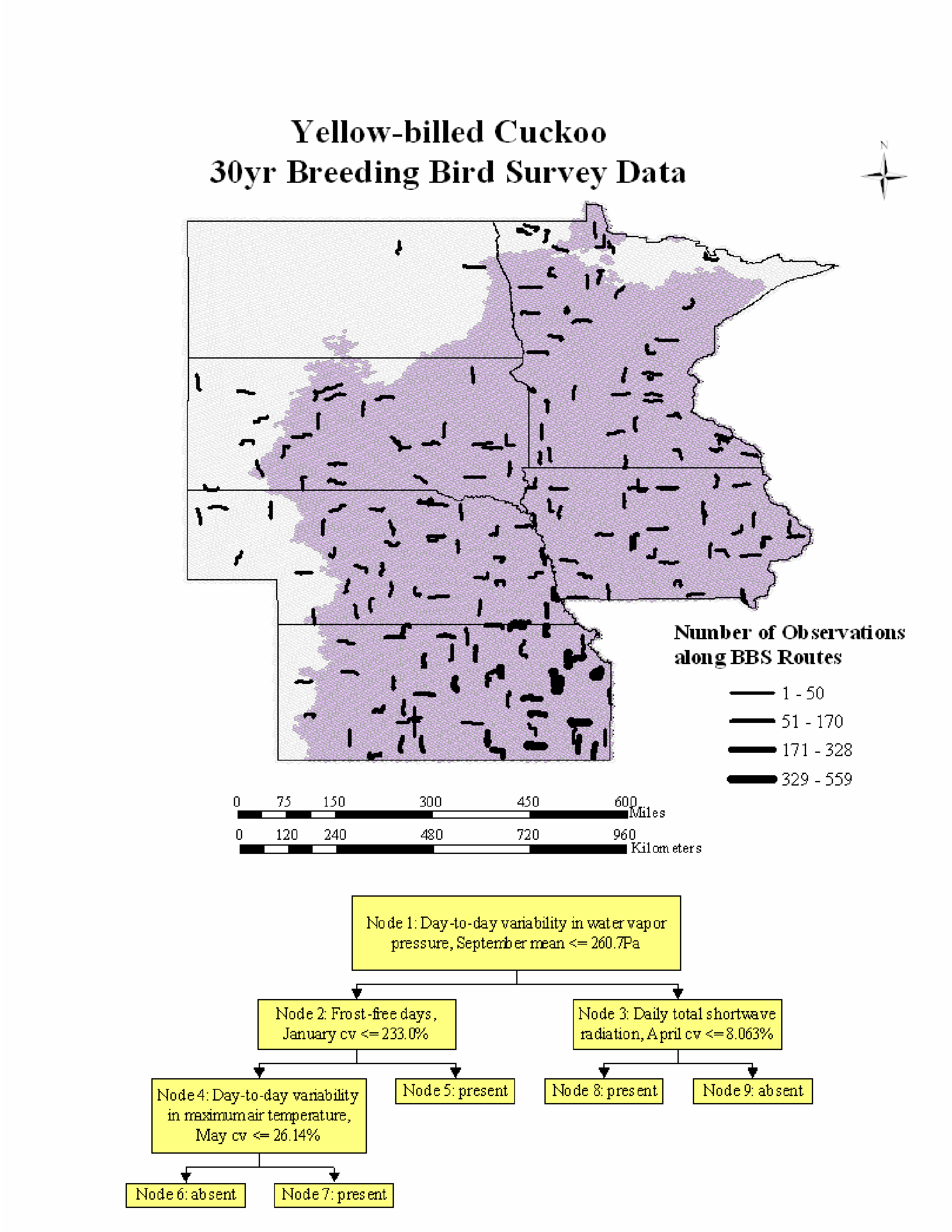
Introduction

Attempts to regionalize species models by mosaicking range distributions produced by neighboring state Gap Analysis projects have been problematic. Variations in habitat modeling result in significant differences in predicted species distributions within and across state lines. Additionally, there is a decided knowledge gap between the spatial and temporal scales used by biogeographers and wildlife managers. Using national geospatial data to map surrogates of habitat, the Nebraska Gap Analysis Project (NE-GAP) examined whether the use of statistical decision trees might help solve these problems.

Methods

We generated regional distributions of 20 selected breeding birds in the six-state GAP Great Plains region (IA, KS, MN, ND, NE, SD) using three recursive partitioning algorithms: QUEST (*Quick, Unbiased, & Efficient Statistical Trees*; Loh and Shih 1997; Shih 2002); CART (*Classification And Regression Trees*; Breiman et al. 1984; De'ath and Fabricius 2000) as an implementation within QUEST; and CRUISE (*Classification Rule with Unbiased Interaction Selection & Estimation*, Kim and Loh 2000, 2001). Breeding Bird Survey (BBS) route level summaries (Sauer et al. 2003) over two time periods (last 10 and 30 years) were used for the occurrence data (presence/absence and abundance), while environmental variables were developed from National Land Cover Data (Vogelmann et al. 1998), Daymet daily climatic means and variances (Thornton and Running 1999), State Soil Geographic (STATSGO) soil texture, and National Elevation Data.

Models were developed on a hexagonal grid produced by the EPA's Environmental Monitoring and Assessment Program (EMAP) with a cell resolution across the Great Plains of approximately 40 km^2. This coverage was intersected with each variable data set to create hexagonal coverages containing averaged values, area-weighted average values, or compositional vectors for each hexagon. These coverage variables were then intersected with the BBS occurrence data. Multiple statistical decision trees were generated for each target species to evaluate the relative strengths and weaknesses of the different algorithms. These statistical trees were then pruned to provide model generality and inverted across the study area to obtain predicted habitat distributions (Figure 1).



Figures 1a and b. Yellow-billed Cuckoo 10 yr (1a) and 30 yr (1b) CART statistical decision trees and the associated predicted range distributions, shown in gray.

Algorithms were compared on the basis of speed of tree identification, interpretability of the cross-validated tree, and plausibility of the range distribution predicted from the tree. Model performance was evaluated by (1) calculating the proportion of species occurrences explained at the first model branch; (2) examining visually how well each model corresponded to published species distributions; (3) assessing correspondence of the model to the spatial distribution of the BBS data; and (4) the computational time required to generate a tree.

Results

CART's exhaustive search of state space took much longer to generate a tree than CRUISE or QUEST (Table 1, Figure 2). All algorithms failed to generate model trees for species with large numbers of observations and/or relatively even distributions across the region (e.g., American Crow, n > 40,000); thus, only 12 of the original 20 species produced sufficient numbers of trees for comparative analysis. CRUISE used fewer observations (number of routes = 340) in its analysis than CART or CRUISE (number of observations/route).

Table 1. Comparison of processing times (CPU-minutes) of the different algorithms. QUEST and CART use # Observations, CRUISE uses #Routes =340.

			Processing Time (CPU-minutes)					
	#Obs		QUEST		CART		CRUISE #Routes = 340	
Common Name	10yr	30yr	10yr	30yr	10yr	30yr	10yr	30yr
Baltimore Oriole	10,343	30,075	12.2	18.3	933.9	2643.5	0.13	0.12
Black Tern	5,713	9,797	9.3	20.5	428.3	627.6	0.12	0.14
Brown Thrasher	53,139	125,960	58.8	76.7	5343.1	no tree	0.16	0.12
Gray Catbird	5,085	11,755	4.4	9.5	1058.0	888.7	0.15	0.13
Great-crested Flycatcher	4,597	11,078	10.8	26.8	83.5	934.9	0.13	0.14
Lark Sparrow	5,025	10,069	7.6	18.3	737.0	782.6	0.13	0.11
Northern Cardinal	9,689	27,114	17.4	62.7	339.0	778.0	0.10	0.09
Northern Harrier	1,593	3,369	1.1	2.1	64.3	232.4	0.13	0.13
Red-bellied Woodpecker	2,498	5,162	3.4	8.0	174.2	254.2	0.12	0.11
Tree Swallow	6,628	6,628	11.5	23.5	598.6	954.4	0.14	0.13
Upland Sandpiper	12,537	12,537	15.5	29.3	189.5	4492.7	0.14	0.14
Yellow-billed Cuckoo	3,268	3,268	5.3	12.1	199.6	859.7	0.13	0.11
Average			8.6	21.4	382.6	1130.8	0.1	0.1

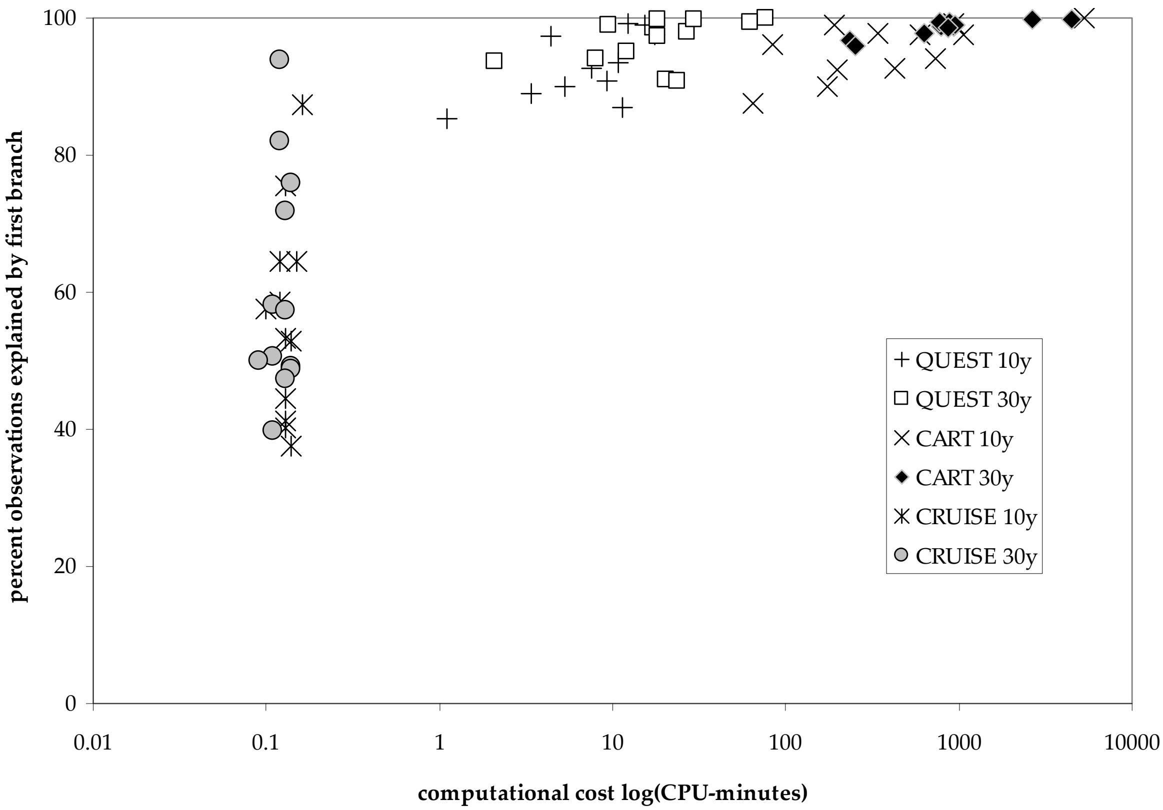


Figure 2. Variation explained at first model branch as a function of computational cost for generation of entire tree.

Trees built from 30 yr data explained a higher percentage of observations at the first model branch than those built from 10 yr data: 30 yr data models averaging 97%, 98%, and 67% versus 10 yr data models averaging 94%, 95%, and 60% for QUEST, CART, and CRUISE, respectively. CRUISE model explanation (avg. 64%) was significantly lower at the first model branch than either QUEST (avg. 96%) or CART (avg. 97%), although the computational costs of the CRUISE models were significantly less (Figure 2). Fewer observations and different configuration of data (routes) between the 10 yr and 30 yr data sets led to differences in inverted geographic ranges (Figure 1). Determining whether this result is due to population trends or less data will require further analysis.

Of the 72 models reported in Table 2, in the first branch of the statistical tree three attempts resulted in no tree, eight models used land cover, eight models used soils or terrain, 15 used water vapor pressure or precipitation, 16 used insolation, and 22 used temperature or frost-free days. Thus, 77% (n=53) of the models relied on climatological variables and of these, 28% (n=15) used climatic variability to model species distribution at the first step of statistical partitioning. Sixty-two percent of the climatological models emphasized the transitional seasons of spring and fall (n=33) over summer (n=12) or winter (n=8). CART and CRUISE selected insolation variables more frequently than QUEST (Table 2), which in turn selected variables more readily interpretable in terms of ecophysiological constraints on bird populations, such as the interannual variability in frost-free days.

Table 2. Synopsis of variable types selected in the first branch of the generated statistical trees (Sp = spring, Su = summer, Fa = fall, Wi = winter).

Common Name	QUEST 10yr	QUEST 30yr	CART 10yr	CART 30yr	CRUISE 10yr	CRUISE 30yr
Baltimore Oriole	% Evergreen Forest	% Evergreen Forest	Mean Su Insolation	% Evergreen Forest	Terrain	Terrain
Black Tern	Mean Wi Vapor Pressure	CV[1] Wi FFD [2]	Mean Fa Insolation	CV Sp min Air Temp	% Emergent Herbaceous Wetlands	CV Sp avg Air Temp
Brown Thrasher	N/T[3]	% Evergreen Forest	% Evergreen Forest	N/T[3]	% Evergreen Forest	Land Covers
Gray Catbird	Terrain	Terrain	Terrain	Terrain	Mean Su Insolation	Mean Su Insolation
Great-crested Flycatcher	Mean Su Insolation	Terrain	Mean Su Insolation	Mean Su Insolation	Mean Fa max Air Temp	Mean Su Insolation
Lark Sparrow	Mean Fa max Air Temp	Soils	CV Wi avg Air Temp	Mean Su Insolation	CV Wi min Air Temp	CV Wi min Air Temp
Northern Cardinal	Mean Sp FFD	CV Su FFD	Mean Sp Vapor Pressure	Mean Sp Vapor Pressure	Mean Sp Vapor Pressure	Mean Sp Vapor Pressure
Northern Harrier	Mean Sp Vapor Pressure	Mean Sp Precipitation freq	Mean Sp Insolation	Mean Sp Precipitation freq	CV Fa max Air Temp	CV Sp Insolation
Red-bellied Woodpecker	Mean Wi Air Temp	CV Fa FFD	Mean Sp Vapor Pressure	Mean Sp Vapor Pressure	Mean Sp Vapor Pressure	Mean Sp Vapor Pressure
Tree Swallow	Mean Wi FFD	Mean Wi FFD	Mean Sp Insolation	Mean Sp Insolation	Mean Wi max Air Temp	N/T[3]
Upland Sandpiper	Mean Su FFD	Mean Su FFD	Mean Su Insolation	Mean Sp min Air Temp	CV Fa Insolation	CV Fa Insolation
Yellow-billed Cuckoo	CV Fa FFD	CV Fa FFD	CV Fa min Air Temp	Mean Fa Vapor Pressure	Mean Fa Vapor Pressure	Mean Fa Vapor Pressure

[1] Frost-free days

[2] Coefficient of variation

[3] No tree was generated by the model

Conclusions

1. Unbiased variable selection in QUEST and CRUISE appeared to facilitate the rapid identification of parsimonious, robust models and plausible range distributions.

2. QUEST trees were generally preferable to CRUISE trees because the latter algorithm relied only upon presence/absence at the route level, while the former considered data on route-level abundance.

3. Developing habitat models using statistical trees generated from species occurrence data and environmental variables can lend a greater degree of objectivity to the modeling process, but there is still considerable subjectivity in the pruning stage that is needed for model generality (Henebry et al. 2001, Holland et al. 2002).

Acknowledgements

This work was supported in part through the GAP Research Project *Evaluating the Use of Statistical Decision Trees for Modeling Avian Habitat and Regional Range Distribution from Occurrence Data and Environmental Variables*.

Literature Cited

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. 1984. Classification and regression trees. Wadsworth and Brooks/Cole, Monterey, California. 358 pp.

De'ath, G., and K.E. Fabricius. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology* 81:3178-3192.

Henebry, G.M., B.C. Putz, and J.W. Merchant. 2001. Modeling reptile and amphibian range distributions from species occurrences and landscape variables. *Gap Analysis Bulletin* 10:22-24.

Holland, A.K., G.M. Henebry, B.C. Putz, M.R. Vaitkus, and J.W. Merchant. 2002. Modeling avian habitat from species occurrence data and environmental variables: Assessing the effects of land cover and landscape pattern. *Gap Analysis Bulletin* 11:25-27.

Kim, H., and W.-Y. Loh. 2000. CRUISE User Manual. University of Wisconsin-Madison, Department of Statistics, Technical Report 989, November 10, 2000.

Kim, H., and W.-Y. Loh. 2001. Classification trees with unbiased multiway splits. *Journal of the American Statistical Association* 96:589-604

Loh, W.-Y., and Y.-S. Shih. 1997. Split selection methods for classification trees. *Statistica Sinica* 7:815-840.

Sauer, J.R., J.E. Hines, and J. Fallon. 2003. The North American Breeding Bird Survey, Results and analysis 1966 - 2002. Version 2003.1. USGS Patuxent Wildlife Research Center, Laurel, Maryland.

Shih, Y.-S. 2002. QUEST User Manual. Department of Mathematics, National Chung Cheng University, Taiwan, April 17, 2002.

Thornton, P.E., and S.W. Running. 1999. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agriculture and Forest Meteorology* 93:211–228.

Vogelmann, J., T. Sohl, and S. Howard. 1998. Regional characterization of land cover using multiple sources of data. *Photogrammetric Engineering and Remote Sensing* 64:45–57.

FINAL REPORT SUMMARIES

Iowa Gap Analysis Project

KEVIN L. KANE
Iowa Cooperative Fish and Wildlife Research Unit, Iowa State University, Ames

Introduction

The Iowa Gap Analysis Project (IA-GAP) began in 1997 to identify areas in the state where vertebrate species richness lacked adequate protection under existing land ownership and management regimes.

To accomplish this goal, the IA-GAP team prepared an assortment of data sets that led to three main pieces of information:

- Iowa vegetation types
- Iowa vertebrate/habitat relationship models for 288 species
- Iowa land stewardship (ownership and management)

When the project began, there were few statewide data sets available that provided the type of data needed for this project. Consequently, much effort was devoted to building the previously mentioned key data layers at a sufficiently fine scale and resolution for subsequent analysis. At the completion of the project, these data became freely available, with the intent that they will be used by those responsible for managing the state's valuable natural resources and by the public, so that everyone can be better informed. With this in mind, we emphasize that these data are dynamic and, in some places, already out-of date. Nonetheless, the data and analyses that constitute IA-GAP represent an important first step toward understanding the status of vertebrates and land cover in Iowa and planning for the conservation of their biodiversity.

Data Development

Land Cover

The land cover of Iowa was mapped by a two-phase, digital classification procedure, which was applied independently to 12 Landsat 5 Thematic Mapper (TM) images covering the state. All TM images were from mid-April to early October between 1990 and 1994. In the first phase, classification of the satellite TM imagery was done by the Iowa Department of Natural Resources, Geological Survey Bureau between 1997 and 1998. In this phase the land cover was separated into six cover types: cropland, grassland, trees, artificial, barren, and water, using unsupervised classification. The resolution of the satellite imagery and resulting classified image was a 30 m pixel, and all subsequent classifications were done at that same resolution. Phase two of the image analysis further differentiated certain land cover classes generated in phase one into one of 29 vegetation alliance aggregations. These alliance aggregations are part of an Iowa vegetation alliance list developed for IA-GAP by experts within the state.

Ground-reference data were used in an unsupervised classification to label each mapping unit according to its land cover type. A total of 29 different land cover types were mapped across the state. Digital National Wetland Inventory (NWI) data was used to provide the wetland data for Iowa GAP. An NWI aggregating model lumped the many classes of wetlands into five general groups: temporary, seasonal, semipermanent, permanent, and water.

The single most extensive cover type was cropland, which comprised almost 60% of the state. As a group, grasslands covered about 28% of the state. Six and a half percent of the state was forested, the upland deciduous forest type making up 5.7% of that total.

Several factors influenced our decision to use existing data sets for accuracy assessment and forego the implementation of a statewide data collection effort. Two data sets were available for use from the same time period as the land cover map: a partial statewide coverage of land cover from the Iowa county offices of the Natural Resource Conservation Service (NRCS) and the joint USGS and EPA generated National Land Cover Dataset (NLCD). Overall accuracies were 77% and 75%, respectively. Both data sets were aggregated to the Andersen Level 1, seven classes, and then compared to Iowa GAP land cover aggregated to the same seven classes.

Predicted Vertebrate Distributions

Distributions of 288 terrestrial vertebrate species were predicted, including 21 amphibians, 44 reptiles, 170 birds, and 53 mammals. The modeling process involved five steps. First, hexagon-based range limits for each species were determined, based on the location of species records or breeding bird survey blocks. This step included input from experts in the field. Next, associations between each species and its habitat, features such as land cover, soil types, and distance to water, were researched and summarized in a Wildlife-Habitat Relationship Model (WHRM) database. After preparing the necessary GIS layers to represent these habitat features, a raster-based modeling approach was used to determine the predicted distribution; the distribution grid was clipped to the extent of the range map. The range maps and WHRM for each species were reviewed by various experts within the state. After review, any necessary changes were made to the range limits and model rules. No accuracy assessment was done for vertebrate species. It is hoped that this report will form the basis for future accuracy assessment studies.

Geographic patterns of species richness generally suggest higher diversity in the northeastern and southeastern portions of the state, with the lowest diversity found in regions where farming is predominately intense row cropping in north central and northwest Iowa. Greater species diversity occurred in the most heavily forested counties in northeastern Iowa and along the major streams and rivers associated with the Mississippi River system. Grasslands in south central Iowa and in the Loess Hills Region also showed greater diversity across taxa.

Considering the issue of scale, we feel confident that our models performed reasonably well for Iowa land cover types. With this coarse-scale model approach, errors of commission will be more common than errors of omission. In other words, overestimation of a species distribution is more likely. Failure to predict a species' presence in an area where it actually occurs may cause inadvertent harm if land-use decisions are made without that species in mind. If, however, a species is predicted to occur where it has never been recorded, it is more likely that the species will be targeted in future surveys and also considered in subsequent land-use decisions.

Land Stewardship and Management

The term "stewardship" is used in place of "ownership" because legal ownership, especially in the case of public lands, does not necessarily identify the entity responsible for management of the land resource. At the same time, it is necessary to distinguish between stewardship and management status, because a single land steward may manage portions of its lands differently.

The digital land stewardship layer was created by incorporating various administrative boundaries into a base layer of land ownership obtained from various sources. State lands were obtained from the Iowa Department of Natural Resources as an ARC/INFO coverage. County lands were done by conducting an extensive mail survey through the Iowa Association of County Conservation Boards (IACCB). Individual counties submitted data on paper maps or as ArcView shapefiles if they possessed GIS capabilities. Each map feature in the stewardship layer was assigned a management status code and other required National GAP attributes. Status codes were determined by consulting management plans if they existed, talking with agency personnel, or looking at legislation that pertained to a particular land designation such as the State Preserves System.

Lands were assigned to one of four management classes based on the relative degree to which land stewards were responsible for maintaining biodiversity values. Status 1 lands reflected the highest, most permanent level of restrictive management; such lands included National Monuments, lands designated as a State Preserve, Nature Conservancy Preserves, and some National Wildlife Refuges where multiple uses were not permitted. Management could be changed more easily on Status 2 lands, such as wildlife management areas and National Wildlife Refuges where multiple uses were permitted, but it was still more restrictive than the remaining multiple-use public lands or private lands, which were assigned to Status 3. Status 4 included lands with no irrevocable easement or mandate to preserve biodiversity values or where the status otherwise could not be determined.

Private land makes up approximately 98% of land in Iowa. Public lands administered by federal, state, and county agencies consist of less than 2% of the state. Other than a few exceptions, most of Iowa's public land consists of relatively small, disjunct areas within a vast amount of private land. Exceptions are areas along the Mississippi and Missouri Rivers, reservoirs along the Des Moines, Cedar, and Iowa rivers, and a scattering of larger complexes managed by many agencies and private individuals. Status 1 and 2 lands occupy less than 0.5 %. Status 3 and 4 lands, which actually actively contribute to the state conservation system, occupy less than 2% of the state; half of this is managed by the Iowa Department of Natural Resources.

Analyses

Once the requisite statewide data were assembled, the actual gap analysis involved intersecting the GIS layers of land cover and predicted vertebrate distributions with land stewardship. These results form the basis of GAP's mission to provide landowners and managers with the information necessary to conduct informed policy development, planning, and management for the long-term maintenance of biodiversity. A practical solution to the problem of defining adequate representation for vegetation or vertebrate species is to report both percentages and absolute area of each element in management areas and allow the user to determine which types are adequately represented in areas under active management.

Land Cover

Being an agricultural state, most land in Iowa is privately owned, and it was expected that the gap analysis results would reflect this situation. Cropland is 99% privately owned as is 98% of the grassland types, 90% of the forest types, and 86% of the herbaceous wetlands. Open water had the lowest private ownership at 53%. All 29 land cover types have less than 10% of their managed areas in Status 1 and 2. Actually, all 29 land cover types have less than 0.5% of their managed areas in Status 1 and 2. The area-weighted average percentage for all status 1 and 2 land in Iowa is 0.05%. Herbaceous wetlands as a group fared the best with 0.22% of their total area in status 1 or 2 land. Forest types follow with 0.17% of their total area in status 1 or 2 land.

Predicted Vertebrate Distributions

Greater than 90% (95.75%) of the predicted habitat for all species modeled in Iowa were on private lands followed by state lands (2.00%), and then federal lands (1.03%). The total amount of land falling into the status 1 and 2 categories was very small (< 0.5 % or 6,678 ha) and reflected in the amount of predicted habitat within these categories. For almost all species (98.26%, 283), the amount of predicted habitat within status 1 and 2 areas was less than 1.0%. The remaining five species (1.74%) modeled were found in the category of 1-<10% of predicted distribution in status 1 and 2 lands.

Conclusions

Intensive agriculture, urban development, drainage, soil erosion, deforestation, channelization of streams and rivers, and an extensive grid of transportation corridors have reshaped Iowa's landscapes since the beginning of European settlement more than a century ago. The tallgrass prairies that covered the state's highly productive soils have been reduced by more than 99%, and about 95% of the once abundant prairie potholes have been drained. Over half of the original forest has been lost, and the remainder has been severely fragmented and disturbed. Most of the natural areas that remain have experienced some kind of disturbance by grazing, fire suppression, or drainage.

Only a tiny proportion (2%) of the land area of Iowa is in public ownership, and only a few tracts are larger than a few thousand acres. Scattered remnants of prairies, forests, and wetlands have been preserved in state and county parks, preserves, wildlife management areas, state forests, and a few privately owned areas. Most public lands are managed for multiple uses, and few areas are managed for biodiversity conservation.

Much of Iowa's biodiversity occurs along stream corridors where the land is less suitable for agriculture. Bluffs and bottomlands along the Mississippi River on the eastern border of the state, and the Loess Hills and Missouri River on the western border represent some of the best of the remaining natural habitats. These major rivers together with smaller rivers and stream corridors are important for species movement for both terrestrial species and migratory birds. The Des Moines river corridor, the Loess Hills, grassland areas in the northwest and south central sections of the state, the Iowa Great Lakes, and the northeast paleozoic plateau are also important centers of biodiversity and have potential for restoration and management.

Because of Iowa's fertile soils and favorable climate, it is likely that the land will remain in agriculture and private ownership in the foreseeable future. Gap analysis can assist natural resource planners with identifying existing centers of biodiversity so that conservation efforts can be directed where they will do the most good. Large tracts of land for biodiversity management are seldom available; therefore, ways must be found to protect biodiversity on private lands such as through long-term conservation easements and other voluntary initiatives.

Kentucky Gap Analysis Project

KEITH WETHINGTON
Kentucky Department of Fish and Wildlife Resources, Frankfort

Land Cover Mapping

A 48-class land cover map of Kentucky was developed as one of the primary inputs for vertebrate distribution mapping. The classification was developed based on the National Vegetation Classification System (NVCS): *International Classification of Ecological Communities: Terrestrial Vegetation of the Southeastern United States* (Weakley et al. 1998) but used alliance aggregations as final map units. The map units were derived after consideration of the availability of statewide ancillary data sets for modeling, success achieved by other states with similar vegetation types, and expert review by state professionals.

The map was created using various combinations of input data, including
1. a vegetation classification based on satellite images,
2. a statewide Digital Elevation Model (DEM),
3. the National Wetlands Inventory (NWI) of Kentucky,
4. various other coverages including ecoregions, hydrology, soils, and urban boundaries, and
5. a computer model to integrate the data.

The vegetation classification used 29 Landsat Thematic Mapper (TM) images and a hybrid of supervised and unsupervised classification routines and preclassification image stratification to produce the primary modeling input. A DEM of Kentucky was created by mosaicking 721 USGS DEM quadrangles comprising Kentucky and a 10 km buffer zone surrounding the state. Derivatives of the DEM including slope, aspect, and landform were used extensively during the vegetation modeling procedures. Other data were used where appropriate and available. The final computer models were developed independently for each ecoregion in the state with the specific input dependent on data availability and the differing physiographic and geomorphic properties of the region.

Final map accuracy on a pixel-by-pixel, per-class basis was 51% with user's and producer's accuracy ranging from 8-100%. A "fuzzy" accuracy assessment was also conducted. The accuracy using this method was 75% with user's and producer's accuracy ranging between 25-100%.

Predicted Vertebrate Distributions

The distributions of 365 native terrestrial vertebrate species, including 52 amphibians, 52 reptiles, 63 mammals, 152 breeding birds, and 111 wintering birds (65 bird species occurred in both the breeding and wintering groups) were predicted. Several steps were required to complete the modeling process. Experts from around the state reviewed the products from each step before the next step was taken. First, we determined range limits for each species based on current information about species' presence or absence within counties and quads, or the Environmental Protection Agency's (EPA) hexagon grid system. Second, the association of each species with habitat features such as land cover, water, edge, and elevation was researched and compiled in a Wildlife-Habitat Relationships (WHR) database. Third, the necessary GIS layers to represent these habitat features were prepared. Fourth, a raster-based modeling approach was used to combine the species' ranges and WHR databases into maps of predicted distributions for each species at a resolution of 30 m grid cells. The final step was to determine the accuracy of the predicted species distributions.

Accuracy assessment was conducted at three spatial scales. In the first level, records of species occurrences were obtained within specific areas of the state. The data were based on species checklists from 51 validation areas, including 47 Breeding Bird Survey routes, one national park, one national forest, one national recreation area, and one state park. Predicted species' presence and absence were then compared with those indicated on the species checklists. In the second level, an accuracy assessment of physiographic provinces was conducted. A predicted species occurrence list was created for each province and compared with the occurrences documented in the species checklists. Finally, an accuracy assessment at the state level was conducted by comparing the predicted species occurrences with observational data compiled from databases around the state. The latter two assessment processes only evaluated species' presence and not species' absence because of our inability to distinguish between observational errors and true absence of species. Thus, these two assessments were not complete.

Geographic patterns of richness of terrestrial vertebrate species indicated that biodiversity (i.e., predicted species richness) was generally higher in the western portion of the state. Highest mean biodiversity was associated with land cover of wet forested habitats (i.e., riparian, bottomland, and floodplain forests) and drier deciduous/coniferous forest habitats. Lowest biodiversity was predicted in the Cumberland Plateau and the northernmost region of the state. Not surprisingly, lowest biodiversity among the habitat types was predicted for areas with little or intermittent vegetative cover (i.e., nonforested mine lands, agricultural lands, and high-density urban areas) across the state.

Comparisons between predicted and observed species presence/absence at 51 validation areas indicated high agreement rates overall. Low rates of omission errors (failure to predict a species at a location in which it has been recorded) occurred, averaging 4% for all taxa combined. Commission error rates (prediction of a species in a location in which it has not been recorded) were also relatively low (< 10%) for all groups except breeding birds. High rates of commission errors for breeding birds (32%) showed that the models were more likely to overpredict bird distributions than to underpredict them. High commission errors are preferable to high omission errors in the context of management decisions. Failure to predict a species in an area in which it actually occurs (omission errors) can lead to management decisions that inadvertently harm the state's biodiversity. In contrast, if a species is predicted to occur where it has never been recorded (commission errors), then that species can be targeted for future surveys and can be considered in land use decisions.

Land Stewardship

The Gap Analysis Program (GAP) uses a scale of 1 through 4 to denote the relative degree of management for biodiversity maintenance for each tract of land, with "1" being the highest, most permanent and comprehensive level of maintenance and "4" being the lowest, or unknown, status. Status codes were assigned to land parcels by Kentucky Department of Fish and Wildlife Resources staff based on conversations with managing entities and field staff regarding management goals and practices. A flow chart adapted from the Gap Analysis Handbook was used to make final status determinations. The gap analysis of Kentucky consisted of intersecting stewardship lands with vegetation and vertebrate distribution. Vertebrate species and natural vegetation types were considered as underrepresented (i.e., "gaps") if < 1% of their statewide distribution fell within Status 1 and 2 lands.

Less than 2% of Kentucky was classified as Status 1 and 2 lands. None of the 35 natural vegetation map units were considered well (> 50%) protected. Two (5.7%) of the 35 units had < 1% of their area in Status 1 and 2 lands. One of these, dry-oak forest in the Upper East Gulf Coastal Plain, is relatively common in other physiographic regions. The other underrepresented vegetation unit (Cumberland highlands forest) is a good candidate for future protection efforts. As a whole, vegetation units exhibited moderate to low levels of protection; 21 of 35 (60% of vegetation units) had < 10% of their area in Status 1 and 2 lands. Other candidates for further protection include floodplain, riparian, and bottomland forest types, because they were consistently associated with higher vertebrate diversity.

Of the 428 native terrestrial vertebrate species assessed, 27 (6.6%) had >10% of their predicted distributions within lands assigned management Status 1 or 2, while 57 (13.3%) species had < 1%. Only 1 (0.2%) species had > 50% of predicted distribution in Status 1 and 2 lands. The distribution of this species (red-breasted nuthatch) was restricted to a single 7.5' USGS quadrangle. Similarly to vegetation units, vertebrate species on the whole showed low to moderate levels of protection.

The GAP analysis of Kentucky confirms that few vertebrate species or vegetation types have high levels of long-term protection. More land under public ownership is warranted, but budgets require land purchases to be strategically located. Data products from this report provide the basic tools needed to examine and compare the merits of potential land purchases. More involvement from private landholders is also needed to improve protection where land for purchase is not available. Finally, continued research that builds upon this present work is needed to monitor long-term trends in vertebrate species and habitat protection.

Maryland, Delaware, New Jersey Gap Analysis Project

D. ANN RASBERRY
University of Maryland Eastern Shore

The Maryland, Delaware, and New Jersey Gap Analysis Project (MDN-GAP) was initiated as a cooperative effort between the U.S. Geological Survey (USGS) Biological Resources Division, the Maryland Department of Natural Resources, and the U.S. Fish and Wildlife Service (USFWS). Late in 1999, the New Jersey Department of Environmental Protection (NJDEP) was added to facilitate work in that state. Various state, federal, and private natural resources entities in the three states were embraced as cooperators throughout the project.

The objectives were to (1) develop databases which describe current land cover, management of lands managed for conservation of biological diversity, and land stewards; (2) identify elements of biological diversity that are underrepresented or not represented in the existing network of lands managed for long-term maintenance of biological diversity; and (3) provide the databases in a format accessible to land managers and stewards, natural resources groups, and others interested in long-term maintenance of biological diversity.

As is true throughout the eastern United States, beginning over 200 years ago with the settlement by western Europeans, the landscape has changed dramatically from heavily forested regions to the current configuration. The middle Atlantic is highly fragmented and continues to be a rapidly urbanizing corridor. This complex matrix of elements made landscape characterization a challenging undertaking.

The 5,051,578 hectares (12,482,449 acres) in the study area were mapped into 62 classes (30 wetland and beach classes, 23 upland forest classes, 7 urban/disturbed/agricultural classes, and 2 water classes). The classification scheme follows the physiognomic hierarchical structure of the National Vegetation Classification System (NVCS). Out of necessity because of the resolution of the satellite imagery and ancillary data, most of the classes ended as aggregates of alliances based on ecological similarities. Landsat TM imagery from 1991-1993 was used as the basis for the classification along with aerial videography flown in 1996-1997.

Stewardship data sets were generated to allow a management status class to be assigned to all public and private conservation properties in the project area. The characteristics used to determine status were

- permanence of protection from conversion of natural land cover to unnatural (anthropogenic habitats, human-induced barren, exotic-dominated, arrested succession),
- relative amount of the tract managed for natural cover,
- inclusiveness of the management, i.e., a single feature or species versus all biota, and
- type of management and degree that it is mandated through legal and institutional arrangements.

This data layer combined with the land cover layer allowed the gap analysis to be performed. As the purpose of the analysis was to identify gaps in current protection of biological diversity, MDN-GAP took a conservative approach in assignment of status categories. Every parcel was considered on the merits of the actual use and management activities for that parcel without an initial bias of ownership. A result of the extensive multiple use policies and practices in these areas is that properties may have received a status characterization different from that which the land manager would have assigned.

It was not surprising, given the preponderance of fragmented and urbanized landscapes in the eastern U.S., to see that 87% of the area in Maryland, Delaware, and New Jersey was held in private ownership. Federal and state lands made up 10% of the area, while regional and local lands were less than 2%. Private conservation organizations owned less than 1% of the land area.

There are a number of public and privately held lands with some degree of biological diversity management activities not included in this data set. These include state university, correctional facilities, and hospital lands, and local land trusts and easements. These data were just being developed during the timeframe encompassing this project, and subsequent updates should make an effort to include these properties in the assessment.

The results of the gap analysis of the land cover are generally described here for MDN-GAP. These are arranged by major land cover class for simplicity; specific details are found within the report.

<u>Forest classes:</u> For the evergreen forest category, there were 10 mapped classes included. Status 1 or 2 lands were approximately 5,299 ha (13,094 ac), which comprised 2.8% of those classes' total area. There were 16 deciduous forest mapped classes with an area of 70,270 ha (173,637 ac) listed as Status 1 or 2 lands, representing 5.6% of those classes' total area. The mixed forest category contained six mapped classes and covered an area of 20,235 ha (50,001 ac) in Status 1 or 2 lands. This was approximately 4.2% of those classes' total area. For forest classes as a whole, there were 32 mapped classes with 95,804 ha (236,732 ac) in Status 1 or 2 classification, which represents 5% of those classes' total area.

<u>Shrubs:</u> The shrub category contained eight mapped classes covering an area of 1,946 ha (4,809 ac) in Status 1 or 2 categories. This represents 6.8% of those classes' total area.

<u>Herbaceous:</u> The herbaceous category contained 11 mapped classes covering an area of 14,544 ha (35,938 ac) in Status 1 or 2 categories. This represents 7.9% of those classes' total area.

<u>Sparsely vegetated:</u> There was only one mapped class included here, the Sparsely Vegetated Beach Alliances (SVBA). Other classes that may have been included were determined to be more transitional than permanent and therefore counted as anthropogenic and not included in the analyses directly. All other classes may be found in the data sets which accompany this report. The SVBA mapped class found in Status 1 or 2 lands had an area of 451 ha (1,114 ac), which represented 22.4% of that class' total area.

The vertebrate predictive distribution modeling aspect of MDN-GAP is under development by a separate GAP project. Expected completion of these models is 2004. As a result, the gap analysis conducted and reported on in this report is limited to the land cover mapping. MDN-GAP provides a baseline upon which more detailed studies and efforts needed for the long-term conservation of biodiversity in the area may be placed. It is hoped that the data sets provided with this report benefit a wide variety of uses.

South Dakota Gap Analysis Project

JONATHAN A. JENKS
South Dakota State University, Brookings

Land Cover

Seventy vegetation alliances were identified in South Dakota using the National Vegetation Classification System. SD-GAP used a regional approach to mapping land cover, separating the state by biogeographic regions: eastern South Dakota (e.g., prairie potholes and Northern Great Plains), western South Dakota (e.g., Northern Great Plains, Badlands, Sand Hills, minus the Black Hills), and the Black Hills. Landsat 5 Thematic Mapper satellite imagery with a resolution of approximately 30 m was obtained through the Multi-Resolution Land Characteristics Consortium. Vegetation was mapped at a scale of 1:100,000 with a minimum mapping unit (MMU) of 2 ha for vegetation. Special features, such as temporary wetlands, also were preserved with a MMU of 2 ha. Satellite images covering eastern South Dakota were interpreted using training data from Farm Service Agency aerial photo section maps (2.56 km^2), which identified agricultural and perennial vegetation categories. A GIS database containing four wetland classes was overlaid onto the classified image. Remaining land cover classes and individual vegetation alliances were mapped by reclustering perennial vegetation and subsequently interpreting the clusters into vegetation types, as well as by onscreen digitizing. Satellite images of the Black Hills were interpreted using training data obtained from a USDA Forest Service GIS that included land cover categories for the Black Hills National Forest. Vegetation was mapped using a GIS summary that contained the percentage of land cover types comprising each spectral cluster, and a regression tree analysis incorporating ancillary databases. Western South Dakota, however, lacked training data over the majority of the land area, and therefore this region was interpreted using known data from only two locations, northwestern South Dakota and Wind Cave National Park (TNC Nature Mapping Program).

Once all regions were classified, they were mosaicked, and population information from CIESEN (University of Missouri, Columbia) was overlaid to represent towns, cities, and industrial areas. The SD-GAP land cover classification identified 35 categories, including 9 grassland, 3 shrubland, 1 dwarf-shrubland, 2 woodland, 5 forest, 6 water and wetland, 3 barren or badland, and 6 disturbance categories. Combined grassland categories dominated the South Dakota landscape, accounting for 56.1% of the land area. Agriculture comprised 31.2% of the entire land area of South Dakota. Creeping juniper woodland, bur oak forest, cottonwood woodland, and shale barren slope sparse vegetation were among the smallest categories. Six water categories comprised 4.5% of the land area; eastern water categories were primarily prairie potholes. Badlands comprised less than 1.3% of the land area, while the forest categories of the

Black Hills made up only 2.7%. A total of 931 locations were assessed for accuracy in eastern South Dakota. Of those, 797 were correctly classified for an overall accuracy of 85.6%.

Terrestrial Vertebrate Distributions

Environmental Protection Agency (EPA) Environmental Monitoring and Assessment Program (EMAP) hexagons were used to create distribution maps for terrestrial vertebrate species throughout the state of South Dakota. Each hexagon in the state was coded with a 0, 1, 2, or 3 using ARC/INFO or ArcView software. Species richness was determined by overlaying all distribution maps using EMAP hexagons. Species richness was determined for all taxa of species and with an overall richness for terrestrial and aquatic species. Habitat models were created by researching habitat use described in papers published with habitat information on each species. Over 1,200 papers were reviewed for both terrestrial vertebrates and aquatic species. Each species' habitat type was evaluated for the most appropriate GIS method. Land cover types were selected from the South Dakota land cover map. Each map was clipped along the species' range boundary, if needed. Land cover was limited by appropriate elevation boundaries, wetland or woodland buffers, soils, or precipitation boundaries for each species by clipping the land cover with the boundary, if necessary. Range maps and models were created for 78 mammals, 215 birds, 30 reptiles, and 15 amphibians occurring in South Dakota. Overall species richness ranged from 177 to 249 of the 338 species present in South Dakota.

Aquatic Systems

After initial processing was completed, habitat variables were added to the reach files. Variables used were temperature, stream size, flow, geology, groundwater potential, relative gradient, size discrepancy, floodplain reach, and dam or stream reach. These are essential variables believed to affect the community structure of the stream system. Each variable was combined to form a valley segment code. This unique valley segment code identified the properties of each individual stream. Fish distribution predictions were based on this code. When a known location was plotted on a given reach, fish species found in that reach were assumed to be in suitable habitat. If all 10 habitat variables matched another stream reach, that stream also was assumed to have suitable habitat. Additional habitat information was located with pertinent literature related to each species. When specific habitat information was found, the attributes were also considered to be suitable habitat and used to predict additional stream types where the species may be found. The literature was incorporated into the species reach models by substituting values in the known valley segment types that were different from data found in the literature. Concatenating habitat variables for each river reach resulted in greater than 6,200 unique valley segment codes. For the analysis, temperature, stream size, flow, gradient, and groundwater potential were used to create a reduced number of concatenated valley segment types. This resulted in 127 unique types in South Dakota. Range maps were created for 116 fish species using South Dakota's 11-digit hydrologic units. After predicted river reaches were identified for 97 viable species, range maps were updated to identify watersheds with predicted reaches present. Species richness ranged from 0 to 87. Highest richness was in watersheds found along the main channel of the large rivers present in the state. Few fish species were predicted to occur in headwater streams, possibly indicating a data gap. Due to the river continuum concept, which states that as stream order increases the number of fish species increases, large rivers are not necessarily in need of protection through public lands. Headwater streams likely need higher levels of protection through public lands to maintain biological diversity.

Land Stewardship

Digital coverages of public land boundaries were obtained from the respective agencies. When coverages were not available, land ownership boundaries were obtained in paper format and digitized to create coverages. All coverages were error-checked, attributed, and combined to create a statewide stewardship map for gap analysis. The stewardship layer identified 20 land ownership categories. Over 70% of the land area of South Dakota was privately owned and managed. Federal and state entities owned approximately 9.5% and 2.1% of the land area in South Dakota, respectively. Status codes also were assigned and evaluated. Greater than 85% of the land area in South Dakota was classified as Status 4 land, meaning that conversion to unnatural land cover types is likely. Just over 1% of the land was considered "highly conserved" (Status 1 & 2 combined).

Gap Analysis

Gap analysis was conducted by intersecting land cover types with conservation status and ownership codes and exporting resulting tables. Square kilometers and percentages of each land cover type were reported within each status and ownership code. Each land cover type or species protection status was reported as 0<1%, 1<10%, 10<20%, 20<50%, or >50% of the species total land area found in Status Code 1 or 2. Three South Dakota land cover types (burned pine, vegetated badlands, and unvegetated badlands) had greater than 10% but less than 20% of land area in Status 1 and 2 management. No valley segment code was protected in at least 10% of its range. No mammal species was protected in at least 10% of its range. Five bird species (bufflehead, great egret, rednecked grebe, white-throated swift, and wood duck) were protected in at least 10% but less than 20% of their ranges. One amphibian species (mudpuppy) was protected in 15% of its range. Twelve fish species were protected in greater than 10% but less than 20% of their ranges, including American eel, lake herring, logperch, longnose gar, mottled sculpin, rainbow smelt, silverband shiner, silver chub, silver lamprey, spottail shiner, suckermouth minnow, and yellow perch. One fish species (blue catfish) was protected in 100% of its range. Protection from conversion currently is lacking in most areas in the state (<1% of South Dakota's land is classified in Status 1 or 2 conservation lands), especially in eastern South Dakota where most land (49%) has been converted to agriculture. None of 127 revised valley segment codes were protected in at least 10% of their ranges. This includes 17 rare valley segment types that received no Status 1 or 2 protection on any portion of the streams. Of these 17, 11 (65%) were present entirely on Status 4 (private) land. Of 55 Black Hills cold-water stream types, 44 (80%) had no Status 1 or 2 protection. Of 72 warm-water stream types, 27 (38%) had no Status 1 or 2 protection. Based on this analysis, South Dakota's biological stream conditions were not considered protected. Overall, 20 of 435 (5%)

vertebrate species in South Dakota were protected in 10% or more of their ranges. Of the terrestrial vertebrates, only 6 of 338 (2%) had 10% or more of their range in Status 1 and 2 lands, while 14 of 97 (14%) of aquatic vertebrates were protected at this level.

Texas Gap Analysis Project

NICK C. PARKER
Texas Cooperative Fish and Wildlife Research Unit, Texas Tech University, Lubbock

The National Gap Analysis Program is a nationwide effort to assess and document the spatial distribution of biodiversity elements and evaluate its conservation status across the U. S. The Texas Gap Analysis Project (TX-GAP) has been conducted at the Texas Cooperative Fish and Wildlife Research Unit (TX-CFWRU) at Texas Tech University (TTU) in Lubbock, Texas. The objectives of the project were (a) to develop a map of current land cover of Texas from recent Landsat TM images that was at least 80% accurate in predicting the true vegetation types; (b) to estimate the potential distribution of Texas wildlife vertebrate species with at least 80% accuracy; (c) to depict and map land stewardship categorized by level of conservation management; and (d) to combine the above data layers in a GIS and perform analyses of species richness patterns relative to known levels of land conservation and management.

Due to the size of Texas and variability in environments, its vegetation is both diverse and complex. In addition, more than 90% of the land is privately owned, and access to the land for field verification procedures is limited. The land cover map was generated through digital classification of satellite imagery supported by field surveys and ancillary information. Accuracy assessment involved a statistical comparison of subset samples from the classified scene to ground observations and using airborne videography. Eighty-three land cover/land use types were mapped, four of which are not natural vegetation (water, bare soil, cropland, and urban). Vertebrate distribution predictions were modeled from known location data and information on species-habitat associations. The distribution maps produced were verified through expert review, by comparing predictions with specific areas where detailed inventories exist, and by conducting field surveys. Land stewardship and management were inventoried and mapped. This information was used as reference for the analysis of vegetative communities and wildlife species conservation status.

For the final analysis, all these data were combined in the GIS to evaluate how vegetation communities, sites with maximum number of species overlap (richness), or even single species distributions were represented in existing managed areas. This allowed the identification and delineation of potential “gaps” in conservation and their potential risks.

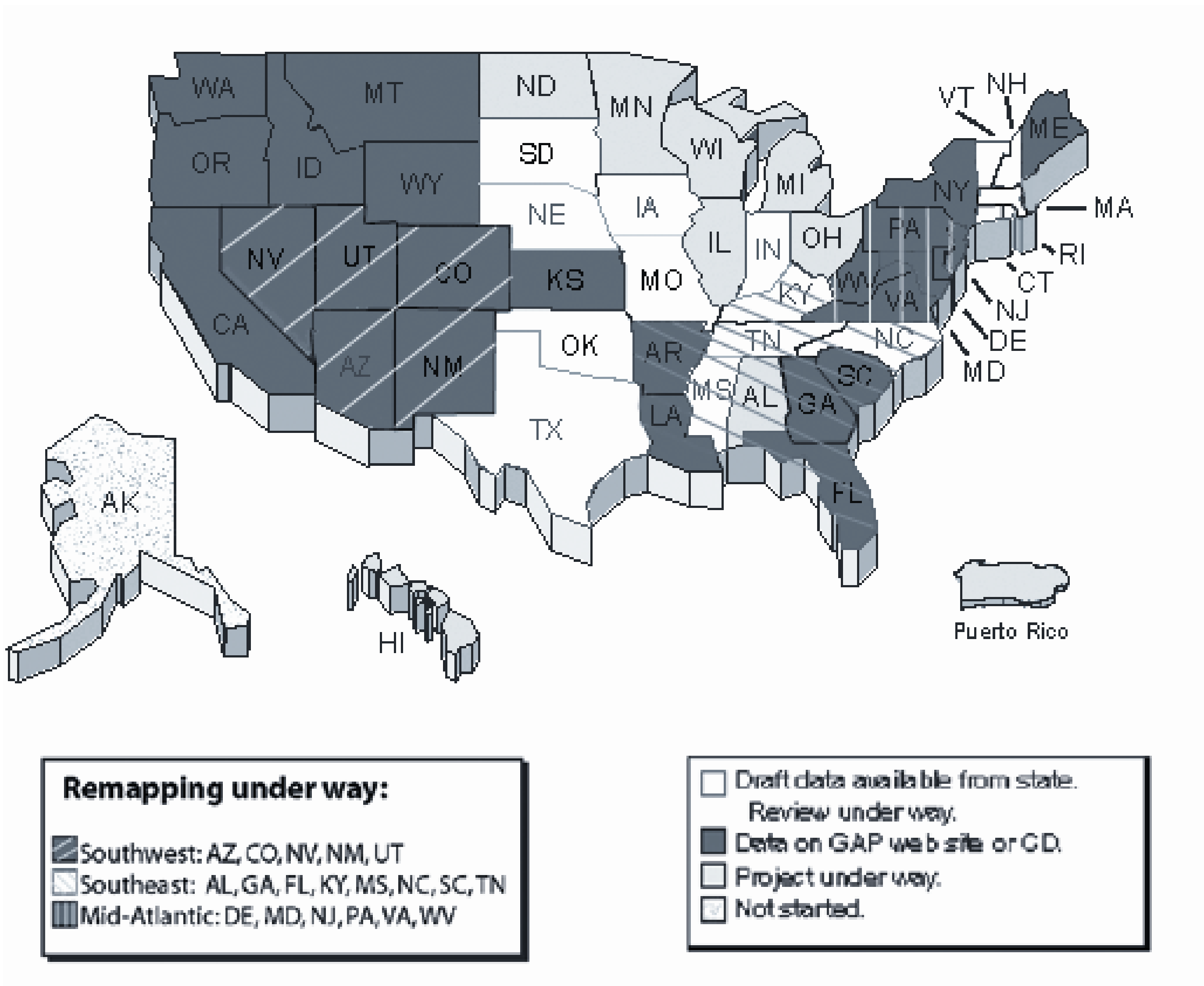
The gap analysis for land cover types showed that the most protected land cover type is the consolidated rock sparse vegetation which overlaps 80% with conservation lands. If we isolate that type, we find that it occurs as rocky cliffs almost exclusively in Big Bend National Park and the Guadalupe Mountains. Although it would appear to be very well protected, it has a limited distribution and only accounts for a very small portion of the total land cover of Texas (0.02%). Other examples of types with restricted areas and very low amounts of protection (i.e., <1%) are juniper woodlands, playa lakes, wetlands, Harvard oak shrublands, and water oak forests. Riparian areas, hardwood bottomlands, and wetlands have experienced some of the most serious reductions in occurrence since human settlement. The grassland prairie types, while they may represent a much larger percentage of the total area (approximately 20% combined) have little overlap with conservation lands (generally <10%). In some cases, they have been overgrazed to the point of creating a type conversion to cactus shrubland types. Despite the widespread occurrence of the prairie grasslands in Texas, their low representation on conservation lands is a concern.

The gap analysis of the vertebrate species shows that most species fall into the 1-10% level of protection. Ninety percent of the 628 species are less than 10% protected. Even more sobering, we do not know what the current distribution of these animals represents in habitat quality or what portion this is of their historic distributions. The majority of the animals in the category that is greater than 50% are species whose predicted distributions are confined to the Big Bend Park area. Three of the species have endangered or threatened listings: the reticulated gecko, the greater long-nosed bat, and the spotted bat. Their level of protection does not necessarily indicate that the populations are viable. However, it does suggest how very important a natural area such as Big Bend is to the continued conservation of these species of interest. On the other hand, some endangered species have very little overlap on protected lands, such as the Comal blind salamander (0.1%) and the Texas kangaroo rat (basically 0%). Texas has a very small amount of land that is set aside for protection of biodiversity. Consequently, 85% or more of each taxon's species are less than 10% protected. As a group, amphibians are least represented.

Because Texas is so dominated by private lands, we clearly need to consider more solutions that offer economic and social incentives to landowners and encourage them to participate in creative management activities that would benefit our natural resources. Education and outreach programs should be developed and promoted throughout the state, and partnerships between agencies, corporations, and the private sector should be encouraged.

STATE PROJECT REPORTS

(Status as of December 2003)



All completed products and reports will be available through the GAP Web site at http://www.gap.uidaho.edu/Projects/Data.asp. Drafts and other products may be obtained from the state project PI as noted.

Alabama

Project under way

Anticipated completion date: December 2006

Contacts: James B. Grand, PI
Leader, Alabama Cooperative Fish and Wildlife Research Unit
Auburn University, Auburn
bgrand@acesag.auburn.edu, (334) 844-4796

Amy L. Silvano, Project Coordinator
Alabama Cooperative Fish and Wildlife Research Unit
Auburn University, Auburn
silvaal@auburn.edu, (334) 844-9295

Land cover: Land cover mapping responsibilities were expanded in June 2003 to include the entire East Gulf Coastal Plain (Figure 1), integrating our mapping efforts with those of the Southeast Gap Analysis Project (SE-GAP). Regionalized mapping methods were tested, and a graduate research assistant was hired in January 2003 to develop methods for spectrally differentiating longleaf pine from other pine species. Nearly all aerial video from Alabama has been mosaicked, and approximately 6,000 training points were compiled from the video and additional field work. The preliminary training points have been imported into a decision tree model to create a series of rules to classify our satellite imagery. The initial decision rule sets were drafted in July 2003 and will be modified with training points collected throughout 2004 to further develop the land cover map. In addition, efforts to build ancillary data layers for the East Gulf Coastal Plain began in November 2003 and will continue in 2004 with the completion of a riparian/wetland layer and landform models.

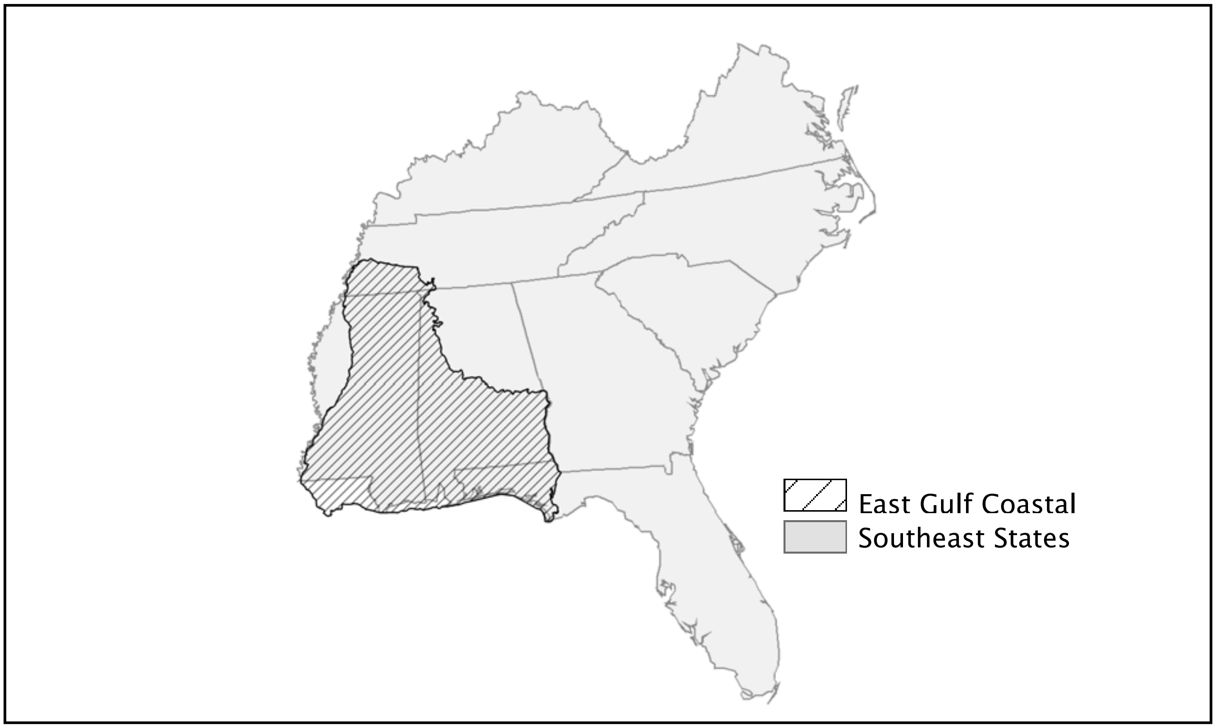


Figure 1. Southeast GAP mapping region.

Animal modeling: Animal modeling has been ongoing since mid-2002. In June 2003, AL-GAP partnered with NC-GAP and GA-GAP to produce vertebrate models for a nine-state area in the southeast (Figure 1) as part of the SE-GAP effort. In total, 594 species will be modeled for SE-GAP, and Alabama will be responsible for constructing models for 246 of those species. To facilitate construction of species range extents, we have developed a Visual Basic application within ESRI's ArcObjects interface to automate range delineations. Statewide hexagon range extents have been completed for all 372 species indigenous to Alabama, and regionwide extents are currently being drafted. Development of regional range extents will continue throughout 2004 as will literature reviews for all habitat relationship data. In addition, we anticipate drafting the predicted habitat distribution models by the end of 2004.

Land stewardship mapping: Stewardship mapping is under way. Digital boundary files and ownership data have been compiled from various public and private agencies through cooperative arrangements. Building of this layer will continue through the duration of the project and will be finalized in the last year (early 2006) to provide the most up-to-date data for our gap analysis.

Reporting and data distribution: Report writing will be ongoing through the duration of the project. Project updates and current information can be found on our Web site at http://www.auburn.edu/gap.

Other accomplishments and innovations: In June 2003, SE-GAP was initiated. Within the mapping area spanned by SE-GAP (Figure 1), Alabama is one of the few states that has yet to complete a state-level project. This provided AL-GAP the unique opportunity to work within the scope of SE-GAP and integrate our state mapping and modeling efforts with those of the regional project. As a result, our land cover mapping area and vertebrate modeling range expanded, extending our project completion date to December 2006. Consequently, AL-GAP will not only contribute to SE-GAP but also ensure the development of state spatial data sets that are regionally consistent, which virtually avoids edge-matching issues and allows for seamless boundaries to be created with our bordering states.

Alaska

Not started

Arizona

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD. Remapping under way (see Southwest Regional GAP).

Arkansas

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

California

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Colorado

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD. Remapping under way (see Southwest Regional GAP).

Connecticut

(see Massachusetts, Connecticut, and Rhode Island)

Delaware

(see Maryland, Delaware, and New Jersey)

Florida

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Georgia

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Hawaii

Project under way
Anticipated completion date: June 2005

Contact: Megan Laut
Hawaii Natural Heritage Program, University of Hawaii, Honolulu
mlaut@hawaii.edu, (808) 587-8591

Land cover: The Hawaii Gap Analysis Project (HI-GAP) has spent the past year developing a decision tree land cover classification approach that is repeatable and accurate. The decision tree classification is implemented in ERDAS Imagine software, using the Knowledge Engineer platform. Different Landsat7 enhancements are used during the process to divide vegetation types into specific categories, thereby increasing overall accuracy.

The land cover classification for the Big Island of Hawaii has been completed. Accuracy assessment of the first draft is currently taking place. While accuracy assessment is being completed for the Big Island of Hawaii, land cover drafts for Maui, Molokai, Oahu, and Kauai are being developed and will be completed in the next several months.

Animal modeling: Data sources have been identified, and data have been collected for bird species. Standardization and normalization of data are currently under way. Data sources for selected representative invertebrates have been contacted, and data collection is under way. Species distribution modeling has been initiated for native and nonnative freshwater aquatic species of vertebrates and selected macroinvertebrates.

Land stewardship mapping: A first draft of the stewardship map has been completed both for the terrestrial and marine environment. GIS data will be made available on the ARC IMS Web site of the Pacific Basin Information Node (PBIN) of the National Biological Information Infrastructure (NBII).

Analysis: Analysis is currently scheduled for FY05. We plan to employ a simulated annealing approach to optimizing biological viability goals while minimizing exposure to degradation of ecological integrity and socioeconomic factors. The gap analysis for our project is anticipated to employ the use of SITES for design of a comprehensive ecosystem conservation approach.

Reporting and data distribution: Data are available for mapping of survey information on aquatic species as well as for stewardship mapping. Contact the Hawaii Natural Heritage Program for details.

Other accomplishments and innovations: HI-GAP has developed partnerships with the National Park Service of Hawaii in an effort to develop vegetation maps for the state parks lands in Hawaii. Planning meetings have taken place to develop this partnership. The University of California at Santa Barbara has partnered with HI-GAP to implement land cover mapping using AVIRIS data for land cover mapping.

HI-GAP has developed an aquatic species distribution modeling approach unique to the Hawaiian Islands' aquatic biota. The modeling approach uses GIS-derived variables in combination with SAS statistical software to cluster watersheds into unique classes. Aquatic species distribution modeling at the stream segment level will be completed for the clustered watersheds.

Idaho

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Illinois

Project under way. Draft data available from state (http://www.inhs.uiuc.edu/cwe/gap/).
Anticipated completion date: March 2004

Contact: Tari Tweddale, Coordinator
Illinois Natural History Survey, Champaign
tweicher@uiuc.edu, (217) 265-0583

Land cover: Complete.

Animal modeling: Complete.

Land stewardship mapping: Complete.

Analysis: Nearly complete.

Reporting and data distribution: Digital coverages will be submitted in early 2004. The IL-GAP team is now in the process of compiling the final report, which will be submitted for peer review in mid-2004.

Indiana

Draft data available from state contact. Review under way.
Anticipated completion date: June 2004

Contact: Forest Clark
U.S. Fish and Wildlife Service, Bloomington
forest_clark@fws.gov, (812) 334-4261 x206

Land cover: Complete.

Animal modeling: Complete.

Land stewardship mapping: Complete.

Analysis: Nearly complete.

Reporting and data distribution: Final report in progress.

Other accomplishments and innovations: The Indiana Biodiversity Initiative's Regional Biodiversity Assessment (RBA) project, which draws heavily on Indiana Gap Analysis data, has moved into the implementation stage with the support of many organizations and individuals and major funding from the Efroymson Fund of the Central Indiana Community Foundation. Four of the seven modified Natural Regions of Indiana have complete RBAs, with the remaining three near completion. Pilot release of the data from the RBAs is planned to occur in southwest Indiana in the winter/spring of 2004. Implementation in other natural regions will follow in 2004 and 2005. The goal is an initial blueprint to guide protection and restoration efforts toward conservation of biodiversity in Indiana. Our goal is to increase recognition of conservation opportunities and in particular to increase opportunities for coordination among agencies and organizations interested in conservation.

Iowa

Draft data available from state (http://www.ag.iastate.edu/centers/cfwru/iowagap/). Review under way.
Anticipated completion date: February 2004

Contact: Kevin Kane

Director, GIS Support and Research Facility
Iowa State University, Ames
kkane@iastate.edu, (515) 294-0526

Kansas

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Contact: Jack Cully
USGS, Manhattan
bcully@ksu.edu, (785) 532-6534

Kentucky

Draft data available from state contact. Review under way.

Contacts: Keith Wethington, PI
Kentucky Department of Fish & Wildlife Resources, Frankfort
keith.wethington@mail.state.ky.us, (502) 564-7109

Tom Kind, Co-PI
Murray State University, Murray
tom.kind@murraystate.edu, (270) 762-3110

Louisiana

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Maine

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Maryland, Delaware, and New Jersey

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Massachusetts, Connecticut, and Rhode Island

Draft data available from state contact. Review under way.

Contact: Curtice Griffin

University of Massachusetts, Amherst
cgriffin@forwild.umass.edu, (413) 545-2640

Michigan

Project under way
Anticipated completion date: June 2004

Contact: Mike Donovan
Michigan Department of Natural Resources
Wildlife Division, Lansing
donovanm@state.mi.us, (517) 335-3445

Land cover: Land cover mapping followed the Upper Midwest GAP protocol (ftp://ftp.umesc.usgs.gov/pub/misc/umgap/98-g001.pdf). Mapping of the existing natural and seminatural land cover of Michigan, in cooperation with the DNR's Integrated Forest Monitoring Assessment and Prescription (IFMAP) project, was completed in 2003. The existing land cover classification for the state (from original MRLC imagery) has been cross-walked to the NVCS.

Animal modeling: Wildlife Division research faculty at Michigan State University (MSU), in cooperation with the Michigan Natural Features Inventory (MNFI) and other Wildlife Division staff, completed species modeling in the fall of 2003.

Land stewardship mapping: The stewardship data layer was completed in the fall of 2003.

Analysis: The gap analysis has begun and will be completed by June 2004.

Reporting and data distribution: Land cover data and stewardship data are available from the USGS Upper Midwest Environmental Sciences Center. Contact Kirk Lohman at (608) 783-7550 x58 or klohman@usgs.gov.

Minnesota

Project under way
Anticipated completion date: September 2004

Contact: Gary Drotts
Minnesota Department of Natural Resources, Brainerd
gary.drotts@dnr.state.mn.us, (218) 828-2314

Land cover: Land cover mapping followed the Upper Midwest GAP protocol (ftp://ftp.umesc.usgs.gov/pub/misc/umgap/98-g001.pdf). The state Department of Natural Resources (DNR) completed classification of the entire state and, with the assistance of NatureServe, cross-walked the classification to the NVCS.

Animal modeling: Hexagon species range maps have been developed for Minnesota and delivered to the USGS Upper Midwest Environmental Sciences Center (UMESC). The animal modeling coordinator for the Minnesota DNR is Jodie Provost (Jodie.provost@dnr.state.mn.us). Meetings were held to conduct expert review of predicted distribution maps for mammals, open landscape, forest, and water birds. Vertebrate distribution mapping will be completed in 2004.

Land stewardship mapping: Stewardship mapping is completed, and a draft version is available from UMESC.

Analysis: Gap analysis will be completed in 2004.

Reporting and data distribution: Draft land cover data and stewardship coverages are available from UMESC. Contact Kirk Lohman at (608) 783-7550 x58 or klohman@usgs.gov.

Mississippi

Draft data available from state contact. Review under way.

Contacts: Francisco J. Vilella, PI
USGS Biological Resources Division
Cooperative Fish and Wildlife Research Unit
Mississippi State University, Mississippi State
fvilella@cfr.msstate.edu, (662) 325-0784

Richard B. Minnis, Coordinator
Department of Wildlife and Fisheries
Mississippi State University, Mississippi State
rminnis@cfr.msstate.edu, (662) 325-3158

Missouri

Draft data available from state contact. Review under way.

Contact: Timothy L. Haithcoat
Geographic Resources Center

University of Missouri-Columbia
HaithcoatT@missouri.edu, (573) 882-2324

Montana

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Nebraska

Draft data available from state contact (http://www.calmit.unl.edu/gap/).
Anticipated completion date: May 2004

Contacts: Geoffrey M. Henebry, Coordinator
CALMIT, University of Nebraska-Lincoln
ghenebry@calmit.unl.edu, (402) 472-6158

James W. Merchant, PI
CALMIT, University of Nebraska-Lincoln
jmerchant1@unl.edu, (402) 472-7531

Land cover: The land cover map has been completed.

Animal modeling: Animal models have been completed.

Land stewardship mapping: Land stewardship mapping has been completed.

Analysis: Gap analyses have been completed.

Reporting and data distribution: Draft report, species atlases, GIS coverages, and metadata under review by state experts before delivery.

Other accomplishments:

1. Henebry, G.M., B.C. Putz, M.R. Vaitkus, and J.W. Merchant. 2003. Accuracy assessment for range distributions of terrestrial vertebrates modeled from species occurrences and landscape variables. *Gap Analysis Bulletin* 12: this issue.
2. Vaitkus, M.R., G.M. Henebry, B.C. Putz, and J.W. Merchant. 2003. Evaluating the use of statistical decision trees for modeling avian habitats and regional range distributions in the Great Plains. *Gap Analysis Bulletin* 12: this issue.
3. Henebry, G.M. 2003. Avian habitat and range distribution modeling in the Nebraska Gap Analysis Project. USFWS Platte-Kansas Rivers Ecosystem Prioritization and Planning workshop, Hays, Kansas, November 5-6. (talk)

4. Henebry, G.M., B.C. Putz, W. Chen, and J.W. Merchant. 2003. Interannual variation in the land surface phenology of the USFWS Platte-Kansas ecosystem. USFWS Platte-Kansas Rivers Ecosystem Prioritization and Planning workshop, Hays, Kansas, November 5-6. (poster)
5. Henebry, G.M., B.C. Putz, M.R. Vaitkus, and J.W. Merchant. 2003. Accuracy assessment for range distributions of terrestrial vertebrates modeled from species occurrences and landscape variables. National Gap Analysis Program Annual Meeting. Ft. Collins, Colorado, October 6-9. (talk)
6. Henebry, G.M., B.C. Putz, W. Chen, and J.W. Merchant. 2003. Exploiting land surface phenology for regional land cover modeling using AVHRR and MODIS image time series. National Gap Analysis Program Annual Meeting. Ft. Collins, Colorado, October 6-9. (poster)
7. Vaitkus, M.R., G.M. Henebry, B.C. Putz, , and J.W. Merchant. 2003. Regional models of avian range distributions for the Great Plains: A comparison of methods and results. National Gap Analysis Program Annual Meeting. Ft. Collins, Colorado, October 6-9. (poster)
8. Vaitkus, M.R., G.M. Henebry, B.C. Putz, and J.W. Merchant. 2003. Evaluating the use of statistical decision trees for modeling avian habitats and regional range distributions in the Great Plains. ESA annual meeting, Savannah, Georgia, August 3-8. (poster)
9. Henebry, G.M., W.W. Hargrove, F.M. Hoffman, B.C. Putz, and J.W. Merchant. 2003. Delineating and resolving ecoregions statistically: Sorting out contexts for wildlife habitat. Nebraska GIS Symposium, Lincoln, Nebraska, May 13-15. (poster)
10. Holland, A.K., G.M. Henebry, B.C. Putz, M.R. Vaitkus, and J.W. Merchant. 2003. Modeling avian habitat from species occurrence data and environmental variables: Assessing the effects of land cover and landscape pattern. Nebraska GIS Symposium, Lincoln, Nebraska, May 13-15. (poster)
11. Putz, B.C., G.M. Henebry, M.R. Vaitkus, A.K. Holland, and J.W. Merchant. 2003. Modeling range distributions of terrestrial vertebrates from species occurrences and landscape variables: Data integration and GIS implementation issues. Nebraska GIS Symposium, Lincoln, Nebraska, May 13-15. (poster)
12. Vaitkus, M.R., G.M. Henebry, B.C. Putz, and J.W. Merchant. 2003. Assembling an environmental geodatabase for regional modeling of avian habitats and range distributions. Nebraska GIS Symposium, Lincoln, Nebraska, May 13-15. (poster)
13. Holland, A.K., G.M. Henebry, B.C. Putz, M.R. Vaitkus, and J.W. Merchant. 2003. Modeling avian habitat from species occurrence data and environmental variables: Assessing the effects of land cover and landscape pattern. US-IALE annual meeting, Banff, Canada, April. 2-6. (poster)
14. Providing GAP land cover, species models, range distributions, and stewardship data to Nebraska decision-makers, federal agencies (Park Service, USDA/APHIS), and researchers outside the state.

Nevada

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD. Remapping under way (see Southwest Regional GAP).

New Hampshire

(see Vermont and New Hampshire)

New Jersey

(see Maryland, Delaware, and New Jersey)

New Mexico

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD. Remapping under way (see Southwest Regional GAP).

New York

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

North Carolina

Draft data available from state contact. Review under way.
Anticipated completion date: June 2004

Contact: Alexa McKerrow
North Carolina State University, Raleigh
mckerrow@unity.ncsu.edu, (919) 513-2853

Land cover: The land cover map and the assessment are complete and under review.

Animal modeling: Models for the 416 vertebrate species that breed in the state have been completed and are under review.

Land stewardship mapping: The stewardship layer is complete. GAP stewardship assignments are being incorporated into the Lands Managed for Open Space, being updated and maintained by the Center for Geographic Information and Analysis.

Analysis: The analysis of land cover is complete and under review. The species-specific GAP status results have been completed and are under internal review.

Reporting and data distribution: The land cover and stewardship chapters are complete and in review. The vertebrate modeling and analysis chapters are in preparation.

Other accomplishments and innovations: To meet their obligations under the State Wildlife Grants program, states must develop a Comprehensive Wildlife Conservation Strategy (CWCS) and submit it to the U.S. Fish and Wildlife Service by October 2005. In North Carolina, the Wildlife Resources Agency is incorporating GAP data into the State Wildlife Conservation Plan.

North Dakota

Project under way
Anticipated completion date: June 2004

Contact: Larry Strong
USGS Northern Prairie Wildlife Research Center, Jamestown
larry_strong@usgs.gov, (701)253-5524

Land cover: The land cover map was completed. The map was produced from analysis of 48 May, July, and September Thematic Mapper images in a per-pixel, supervised classification procedure using a sequential series of classification tree analyses. National Wetland Inventory data were inserted into the land cover classification. The legend for the land cover map includes 12 prairie, 3 shrubland, and 10 woodland plant communities. The final land cover map is fine-grained (0.09 ha pixels) with an extent of 183,103 sq km. An accuracy assessment of the land cover map is approximately 50% completed. Early results reveal the stratified random, single-stage cluster sample design is providing useful information about the spatial distribution of classification accuracy.

Animal modeling: Individual species models, co-occurrence (hypergrids), and species richness grids for birds, mammals, amphibians, and reptiles have been completed. Reference databases, spreadsheets, and metadata for vertebrate models are completed. Accuracy assessment for amphibians, reptiles, and mammals is complete and near completion for birds.

Land stewardship mapping: The stewardship vector was completed. The data set identifies 16 stewards and 33 stewardship categories. Estimates of the area for status 1 and status 2 lands are 38,200 ha and 259,900 ha, respectively.

Analysis: The gap analysis is near completion.

Reporting and data distribution: The final report and CDs of products for the National Gap Analysis Program are in progress, and their completion will be the major activity in the winter of 2003/04. The report and data will also be made available to North Dakota GIS Technical Committee for distribution on the North Dakota GIS Hub.

Ohio

Project under way
Anticipated completion date: September 2006

Contact: S. Alex Covert, Coordinator
U.S. Geological Survey, Columbus
sacovert@usgs.gov, (614) 430-7752

Land cover: The goal of acquiring 60,000 digital aerial photographs was completed in 2003. These images have been georeferenced for about two-thirds of Ohio, with the remaining images to be ready in March 2004. Fieldwork to verify the aerial photographs was performed at about 600 locations throughout Ohio. An Anderson Level II classification as well as an unsupervised classification was completed for Ohio. The final vegetation land cover map will be classified using the "Terrestrial Ecological Systems of the United States" produced by NatureServe. Ohio has worked with Shannon Menard of NatureServe to cross-walk classifications from the National Vegetation System to the Ecological System. A draft supervised classification has been completed for about half of the state, mainly western and northern Ohio, with the rest of the state to be completed in June 2004. All classified areas will be merged, and an accuracy assessment will be performed in the next year.

Animal modeling: The hexagon range maps for Ohio reptiles were completed in 2003. Range maps for all 308 breeding terrestrial vertebrate species have therefore been constructed and reviewed. A hexagon range map was produced that shows preliminary species richness for all terrestrial vertebrate species (Figure 1). Literature review for habitat affinity data has been completed as well. Efforts to model species distributions using scripts created by West Virginia GAP will begin in 2004. Expert review will follow.

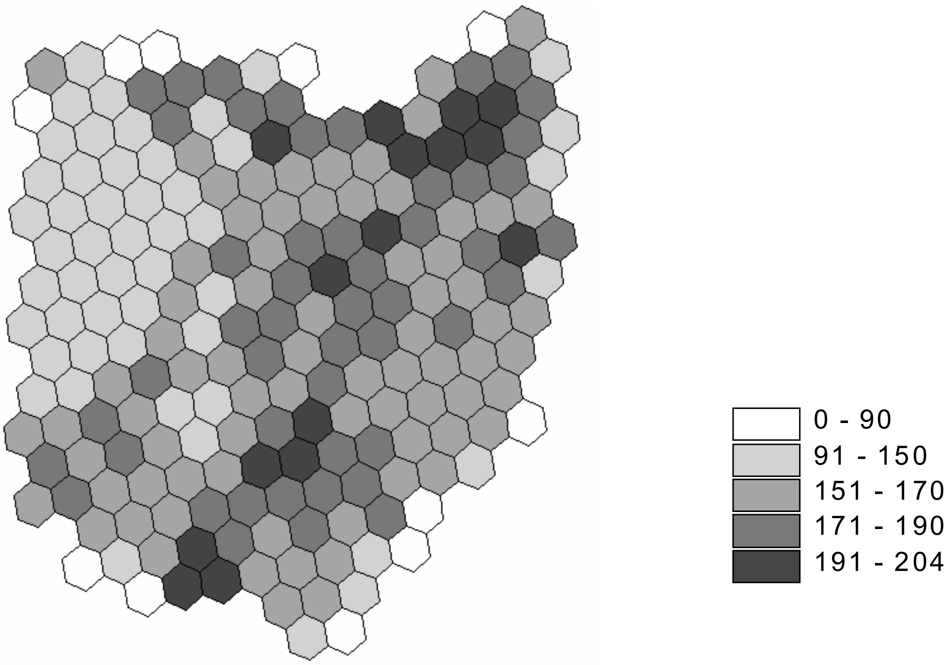


Figure 1. Preliminary species richness for all terrestrial vertebrate species in Ohio.

Land stewardship mapping: Digital maps of all Ohio conservation lands were obtained and compiled into one map. Each land parcel was attributed with a GAP land-status code. The map was reviewed and finalized in 2003.

Reporting and data distributions: Hexagon range maps for reptiles were released on the Ohio GAP Web site (http://oh.water.usgs.gov/ohgap/ohgap.html) in 2003. Two stakeholders meeting were held in June and December 2003.

Oklahoma

Draft data available from state. Review under way.

Contact: William L. Fisher
Oklahoma Cooperative Fish and Wildlife Research Unit, Stillwater
wfisher@okstate.edu, (405) 744-6342

Oregon

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Pennsylvania

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Puerto Rico

Project under way
Anticipated completion date: December 2005

Contacts: William Gould, PI
USDA Forest Service, International Institute of Tropical Forestry, San Juan
wgould@fs.fed.us, (787) 766-5335 x209

Brick M. Fevold, Project Coordinator
USDA Forest Service, International Institute of Tropical Forestry, San Juan
bfevold@fs.fed.us, (787) 766-5335 x204

Land cover: The International Institute of Tropical Forestry (IITF) developed a semiautomated process to create a Landsat-7 ETM+ image mosaic based on 2001-2003 satellite imagery that is 97.5% cloud- and cloud-shadow free (Martinuzzi et al. 2003b). The semiautomated process is capable of performing routine regional updates (e.g., where clouds and cloud-shadows remain) as new imagery becomes available. The resulting composite is being classified to provide us with the most current land cover classification and habitat map of Puerto Rico for our vertebrate distribution modeling and mapping (Gould et al. 2003). Initial classification includes mapping the extent of four classes of urban cover in Puerto Rico (Martinuzzi et al. 2003c). Urban cover comprises nearly 15% of the land surface in Puerto Rico, and the urban forest and low- and high-intensity urban land cover classes are important in both our habitat modeling and in understanding the dynamics of land cover change and threats to habitat sustainability and biodiversity. We have compiled vegetation descriptions from the plant community level and organized them into a hierarchical structure along gradients of climate, substrate, and topographic position (Carrero et al. 2003). We have developed two new maps that will be useful in both our land cover mapping and modeling efforts. These include an updated map of the physiography of Puerto Rico (Gould et al. in prep.) and an analysis and map of landforms (slope position) of Puerto Rico (Martinuzzi et al. 2003a).

Animal modeling: Our original list of 437 vertebrate species has been through expert review and now consists of 426 species known to occur across Puerto Rico or its off-shore islands. The collection of species occurrence information has been, and continues to be, an arduous process. A large proportion of Puerto Rico's vertebrate fauna is composed of species dependent upon aquatic and/or coastal-marine habitat. With this in mind, we are developing our relational database model with the understanding that the aquatic and marine species are important components of the landscape and have good potential for gap analysis after the completion of the terrestrial Gap Analysis project. We have identified a subset of 168 species to include in the terrestrial component of the gap analysis. This list contains those species considered endemic, resident, breeding migratory, or are species with special conservation importance that have become established through human introductions (e.g., Asian mongoose) or range expansion (e.g., Hispañolan parrot). PR-GAP adopted a modification of the USFS Forest Inventory and Analysis (FIA) hexagon grid of the Caribbean as the minimum mapping unit for creating species' geographic range maps. We feel the smaller hexagon size (24 km^2) is a valid scale for representing species distribution while considering the challenge of representing Puerto Rico's diverse and heterogeneous landscape. We are collaborating with the Puerto Rico Ornithological Society to assist in the development of field survey methods for a Breeding Bird Atlas for Puerto Rico and to incorporate PR-GAP data, maps, and analyses into the Atlas. Expert review of species geographic range maps are currently in progress.

Land stewardship mapping: We are currently establishing an interagency collaborative effort to update an existing, but incomplete land stewardship layer of Puerto Rico. We will be identifying land management areas, contacting land managers to determine management policies, classifying land parcels into the management strategies used in the GAP program, and developing a land management geospatial database in order to facilitate the final GAP analyses.

Analysis: Gap analyses will begin as we complete our vertebrate models and database in 2004.

Literature cited:

Carrero, G., W. Gould, B. Fevold, G. González, and S. Martinuzzi. 2003. Hierarchical vegetation classification for the Puerto Rico Gap Analysis Project: Integrating climate, substrate, topography, and species composition in a land cover map legend. Poster presented at the National GAP Annual Meeting, October 6-9, 2003. Fort Collins, Colorado.

Gould, W., S. Martinuzzi, and O. Ramos. 2003. Image analysis and land cover mapping for Puerto Rico. Poster presented at the National GAP Annual Meeting, October 6-9, 2003, Fort Collins, Colorado.

Gould, W., S. Martinuzzi, B. Edwards, and O. Ramos. 2004. Physiography, geology, and the distribution of landforms in Puerto Rico: Shaping land use and vegetation. In preparation.

Martinuzzi, S., W. Gould, and O. Ramos. 2003a. Integrating remote sensing and GIS for land cover mapping and analysis in the Karst area. Presented at the Second Symposium of Karst Research, September 27, 2003, Interamerican University in Bayamon, Puerto Rico.

Martinuzzi, S., W. Gould, and O. Ramos. 2003b. Cloud and cloud shadow removal in the creation of a cloud-free composite Landsat ETM scene in tropical landscapes. Poster presented at the National GAP Annual Meeting, October 6-9, 2003, Fort Collins, Colorado.

Martinuzzi, S., W. Gould, and O. Ramos. 2003c. Urban cover estimates from image analysis and land cover mapping of Puerto Rico. Presented at the 2nd Congreso de Ecourbanismo, Centro de Bellas Artes, November 18-19, 2003, Caguas, Puerto Rico.

Rhode Island

(see Massachusetts, Connecticut, & Rhode Island)

South Carolina

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

South Dakota

Draft data available from state contact. Review under way.

Contact: Jonathan A. Jenks
South Dakota State University, Brookings
jonathan_jenks@sdstate.edu, (605) 688-4783

Southeast Regional GAP

Update under way for the thirteen-state region.
Anticipated completion date: June 2006

Contacts: Alexa J. McKerrow and Steven G. Williams
Biodiversity and Spatial Information Center
North Carolina State University, Raleigh
mckerrow@unity.ncsu.edu, (919) 513-2853
steve_williams@ncsu.edu, (919) 513-2853

Elizabeth R. Kramer
Natural Resource and Spatial Analysis Laboratory
University of Georgia, Athens
lkramer@arches.uga.edu, (706) 542-3577

Amy L. Silvano
Alabama Cooperative Fish and Wildlife Research Unit
Auburn University, Auburn, Alabama
silvaal@auburn.edu, (334) 844-9295

The Southeast Gap Analysis Project is working hard to develop the baseline data sets and tools for use by the conservation community within the region. Currently we have two focus areas: (1) production of regionally consistent and current data sets and (2) use of existing state GAP project data to create regional products for use by partner agencies. The Southeast Gap Analysis Project started actively mapping and modeling in July 2003. The regional effort involves thirteen states (Figure 1) throughout the southeastern U.S. The goal is building on the state GAP experiences to develop consistent land cover and vertebrate models. The regional work is being coordinated through the Biodiversity and Spatial Information Center (BaSIC) at North Carolina State University with partner researchers at the Natural Resource and Spatial Analysis Laboratory (NaRSAL) at the University of Georgia and the Alabama GAP Project at the University of Auburn.

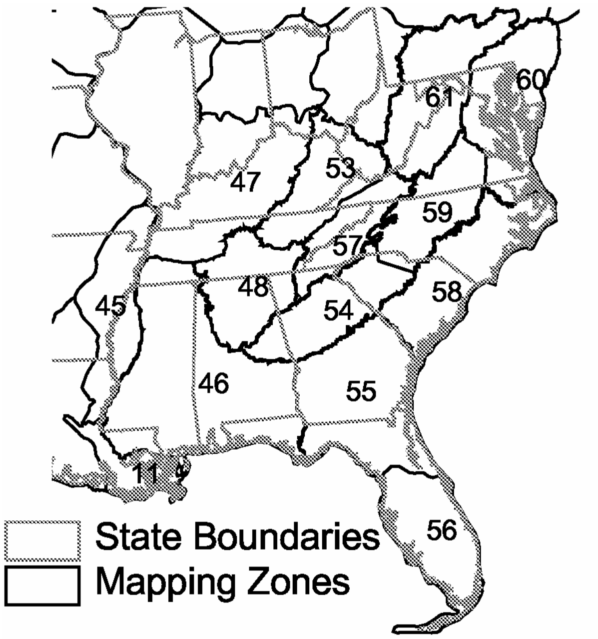


Figure 1. Mapping zones and states Southeast GAP is working in.

Land cover: Products being developed include land cover maps and impervious surface and canopy closure estimations. The land cover maps will be created at two levels of thematic detail: one compatible with the National Land Cover Dataset (NLCD 2001) and the second a detailed vegetation map based on the Ecological Systems described by NatureServe (2003). We are actively working with EROS Data Center (EDC) methodologies to create the general land cover products based on the NLCD 2001 protocols. Currently we are working on the general land cover in four, and on impervious surface and canopy estimations in seven of the ten Southeast mapping zones (Table 1).

Table 1. Lead responsibilities and timeline for completion of general and detailed land cover mapping for the Southeast GAP Project.

Southeast Mapping Zone	Lead Project		Year	
	NLCD	GAP	NLCD	GAP
46. Gulf Coastal Plain	AL	AL	2004	2006
48. Interior Low Plateaus	NC	GA/NC	2005	2006
57. Blue Ridge, Ridge & Valley	GA	GA	2005	2006
59. Southern Piedmont	GA	GA	2004	2005
54. Northern Piedmont	GA	GA	2004	2005
55. Southern Coastal Plain	NC	NC	2005	2005
58. Northern Coastal Plain	NC	NC	2004	2005
53. Eastern Highlands & Plains	Other	NC	N/A	2006
47. Western Highlands	Other	GA	N/A	2006
56. Southern Florida	Other	NC	N/A	2006

In preparation for the detailed land cover mapping, we have gathered over 30,000 digital photographs over the region. The camera setup consists of a Kodak 645 Pro Digital Back coupled with a Hasselblad H1 camera body and 80 mm lens. Other sensors include a Watson Inertial Measurement Unit and a Trimble GeoExplorer 3 GPS unit. In addition, digital video is also being collected as a backup data source in case of camera troubles. Software for rectification of the photographs is near completion at the University of Georgia. With a swath of 600 meters and a resolution higher than 0.2 meters, these photographs will be used to build an extensive sample set for use in the detailed land cover mapping. The approach is similar to that used by many of the GAP projects previously, with a subset of photos being visited on the ground to verify the cover types being seen. Those photographs will then be used as a reference library for labeling the remaining photos throughout each mapping zone.

NatureServe ecologists are assisting throughout the process to guarantee that the Ecological Systems, as they have been described, are being appropriately labeled in the field as well as in the computer labs. Alabama GAP, having started prior to the development of the regional effort, has hit the ground running with field visits based on previously flown videography. They are starting with field visits using the new photographs in March 2004. The field visits for creation of the reference library for the remaining zone are scheduled for late spring 2004.

Regional data sets being compiled to support the land cover mapping efforts include Ecological Land Units, 1997 Census of Agriculture (USDA 1997), the National Wetland Inventory (NWI) data, golf courses, and mines. With over 1,000 NWI quadrangles not available in digital format for most of Alabama and Mississippi, Alabama GAP initiated a tremendous effort to start scanning those quadrangles. Given the scope of the effort, the utility of those data to a broad user base, and the ultimate goal of having those data in a vector format, we are working with collaborators to identify ways to get that work done.

Animal modeling: The design of the Southeast Gap Analysis Vertebrate Database has primarily been the responsibility of the BaSIC personnel. While design of the database has been centralized, the responsibility for development of the ranges and the habitat suitability models has been split up between the three laboratories based on the specific expertise and interests of each of the vertebrate biologists. A total of 608 terrestrial vertebrate species are being modeled in the Southeast.

The ranges for each of these species will be hand-delineated based on the existing hexagon-based data from the individual state efforts, as well as a review of the literature for each of the species. In order to facilitate a common approach in range delineations, a common set of spatial data layers (ecological region boundaries, watersheds, hydrology, outerbanks, tidal/non-tidal boundary) has been compiled for the three labs to use when line work from an existing layer describes the range limits. An internal review for these new ranges will be conducted in the summer of 2004.

The habitat database has been designed in Access, and the literature reviews from each of the state GAP efforts have been compiled. A unified set of habitat relationships for the region will be created with relationships to both the detailed and general land cover map units being developed. As the land cover data for each of the mapping zones becomes available, the habitat models for that zone will be created and reviewed internally. By 2006 the detailed land cover for the entire region will be available and used as the basis for the final habitat maps.

Other accomplishments and innovations: A pilot study has been initiated between the Southeast GAP Project and the U.S. Fish and Wildlife Joint Ventures Program. This project is a direct response to discussions at the October 2003 National GAP meeting, in which an opportunity for collaboration between ongoing GAP projects and USFWS Joint Venture bird conservation planning efforts were identified. In this project SE-GAP will augment the responsiveness of ongoing mapping and modeling work to specific data and analysis needs of the USFWS. Two goals of the project include (1) developing a map of the historic distribution of longleaf pine for the Atlantic and East Gulf Coastal Plain and (2) working with USFWS personnel to refine habitat models for priority bird species by identifying and creating habitat content and context variables specific to key southeastern habitats.

Literature cited:

NatureServe. 2003. A working classification of terrestrial ecological systems in the coterminous United States. International Terrestrial Ecological Systems Classification. NatureServe, Arlington, Virginia. 61 pp. + appendices.

U.S. Department of Agriculture. 1997. Census of Agriculture Geographic Area Series. National Agriculture Statistics Service.

Southwest Regional GAP (SWReGAP)

Update under way for the five-state region encompassing Arizona, Colorado, Nevada, New Mexico, and Utah. State coordination for the project is facilitated through the SWReGAP Web site (http://leopold.nmsu.edu/fwscoop/swregap/default.htm).

Anticipated completion date: May 2005

Contacts:

Julie Prior-Magee, SWReGAP Coordinator
USGS/BRD, Las Cruces, NM
jpmagee@nmsu.edu, (505) 646-1084

Arizona: Kathryn A. Thomas, PI
USGS/BRD Southwest Biological Science Center
Colorado Plateau Research Station, Flagstaff
Kathryn_A_Thomas@usgs.gov, (928) 556-7466 x235

Colorado: Donald L. Schrupp, PI
Colorado Division of Wildlife
Habitat Resources Section, Denver
hqwris@lamar.colostate.edu, (303) 291-7277

Nevada: David F. Bradford, Co-PI
U.S. EPA, Office of Research and Development, Las Vegas
bradford.david@epa.gov, (702) 798-2681

William G. Kepner, Co-PI
U.S. EPA, Office of Research and Development, Las Vegas
kepner.william@epa.gov, (702) 798-2193

New Mexico: Ken Boykin, Co-PI

NM Coop. Fish & Wildlife Research Unit, Las Cruces
kboykin@nmsu.edu, (505) 646-6303

Utah: R. Douglas Ramsey, PI
RS/GIS Laboratory, Utah State University, Logan
dougr@cnr.usu.edu, (435) 797-3783

Land cover: The RS/GIS Lab at Utah State University is the regional land cover mapping lab for the five-state southwest region. Coordination with the other four states is facilitated through a Web page that allows access to spatial data, procedural documents, and an Internet Map Server (http://www.gis.usu.edu/docs/projects/swgap).

Land cover mapping methods – Landsat 7+ imagery for three dates (spring, summer, fall) spanning the years 1999-2001 is being used, along with ancillary DEM-derived data to map land cover for the five-state region. The land cover mapping protocol follows the approaches employed by EROS Data Center (EDC) for the National Land Cover Database (NLCD). We are using the CART Imagine module developed for EDC by EarthSat Corp. along with the classification tree software See5 (Rulequest). Geographic stratification of the region is accomplished through *mapping zones* and *functional units*. Both mapping zones and functional units represent ecoregional divisions of the landscape. Mapping zones are smaller mapping areas with similar ecological and spectral characteristics and nest within functional units. Functional units are broader units used to aid in tracking and reporting mapping progress for the five participating states (Figure 1).

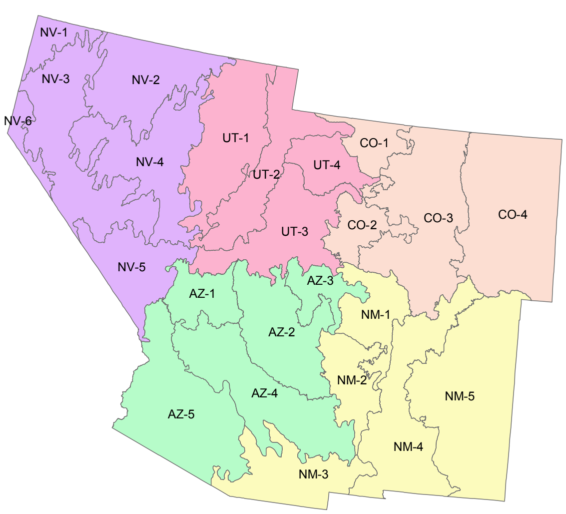


Figure 1. Functional units for SWReGAP.

More detailed information on the mapping process can be found in the SWReGAP Land Cover Handbook, available at
http://www.gis.usu.edu/%7Eregap/download/documents/LCHandbook112603.doc.

Field data collection – Field data collection for SWReGAP land cover mapping was completed in the 2003 field season. Each of the five states had at least two field crews working this season, resulting in approximately 35,000 field samples collected for the region. Additional sample data, collected from various agencies and other projects in the region, augmented the sampling effort of the project. It is estimated that approximately 50,000 samples were collected in total for the five-state region.

Land cover mapping progress by functional unit – As of December 31, 2003, preliminary maps were in progress for the following functional units: CO-1, CO-2, CO-3, NV-3, NV-4, NV-5, NV-6, NM-1, NM-2, NM-3, NM-4, NM-5, UT-3, and UT-4. Preliminary maps were complete for functional units AZ-1, AZ-2, AZ-3, AZ-4, AZ-5, CO-4, NV-1, and NV-2. Final maps were complete for UT-1 and UT-2.

Goals for the coming year – The regional land cover mapping timeline was revised during 2003 and, in accordance with this timeline, each state will complete their state responsibility area by May 2004. Utah State University will mosaic the five areas, and the regional land cover map will be completed by June 2004.

In March 2004, a special SWReGAP session is planned for the International Association for Landscape Ecology's regional conference in Las Vegas, Nevada. Utah State University will coordinate the development and preparation of presentations for the land cover portion of that special session.

From July to December 2004, USU will coordinate the development and compilation of the land cover portion of the SWReGAP final written report. Also during this time period, USU will organize and archive all spatial and tabular databases associated with the land cover effort for the five-state region.

Animal habitat modeling: The New Mexico project is providing regional animal habitat modeling coordination. The primary objectives include (1) defining wildlife-habitat relationships, (2) finalizing the list of taxa to model, including review and finalizing decision rules, (3) allocating taxa modeling responsibilities among the projects, (4) identifying multiple modeling techniques that may be of use for the project, (5) creating a habitat modeling protocol to facilitate data collection and consistency within the region, (6) creating a Web interface for data transfer to the regional lab, (7) creating a database to facilitate association compilation, expert review and modification, and potential end user application, and (8) conducting a regional animal habitat modeling workshop in Fort Collins, Colorado, in October 2003.

Defining wildlife-habitat relationship models – To assure regional consistency, the New Mexico project proposed several definitions for Wildlife-Habitat Relationship (WHR) for the region. In summary, a wildlife-habitat relationship is a statement describing resources and conditions present in areas where a species persists and reproduces or otherwise occurs. Relationships can be modeled to predict habitat composition and, if the relationships are represented in a cartographic plane, they can predict the presence of habitat spatially.

Decision rules and modeling allocation – Taxa inclusion into the modeling process was determined by a series of decision rules. These rules initially identified 839 species to be modeled in the SWReGAP effort. Exclusionary rules removed species if they had only incidental, accidental, or vagrant occurrence. As modeling continues, taxa can be eliminated if they meet one of the exclusionary decision rules. Currently the total number of species to be modeled is 836.

A taxa allocation decision rule was created to distribute initial taxa modeling responsibilities among all projects in a manner that capitalizes on previous modeling experience, is localized to the distribution of taxa experts, and is sensitive to the greatest awareness about local conditions applicable to more restricted taxa. All projects have opportunity for input on modeling approach and results among taxa, regardless of the lead assignment. The current allocation of taxa modeling responsibility is: Arizona – 189 taxa, Colorado – 157, Nevada – 73, New Mexico – 378, and Utah – 39.

Habitat modeling database – The New Mexico project has created an Access database to compile taxa-specific information for modeling. The intent is to create a data set that manages information and is used to construct each taxon's wildlife habitat relationship model. The database addresses several concerns of the regional group regarding expert participation and end user functionality. Included within the database is a user-friendly method to define range limits using the 8-digit hydrologic unit code (HUC). Each HUC is designated using a 3-character coding system based on historic/recent distribution as either known (K), potential (P), or extirpated (X). We developed a coding system based on reproductive use (breeding, nonbreeding, both) and seasonal use (migratory, wintering, summering, wintering and summering). The database also incorporates the core data layers the region had identified to be minimally addressed in each wildlife habitat relationship model. These core data layers are land cover, elevation (minimum and maximum), slope, aspect, soils, hydrology (distance to and association with permanent water), and patch size. Other layers specifically addressed in the database are mountain ranges, temperature (minimum and maximum), and precipitation. The database allows further data layers to be incorporated into the model-building process.

Incorporated in the habitat modeling database is the ability to model species beyond the current overlay process. We are continuing to review modeling techniques that can be applied to gap analysis habitat association information. Within the database we have the option of applying a weighted index overlay procedure in addition to the standard Boolean overlay procedure. Index overlay offers a subjective consideration of the relative value of habitat variables, and fuzzy sets allow for the inclusion of ambiguity at the habitat boundaries. If applicable, two products will be produced: nonbinary representations incorporating uncertainty and the traditional GAP binary representations.

The database was presented to the region at a workshop held prior to the National GAP Meeting in Fort Collins, Colorado, in October 2003. This workshop detailed the process of populating the database. It also provided a forum to identify additional needed database refinements. The database will be modified through the course of the project to ensure the most functionality possible at project completion.

<u>Habitat modeling progress</u> – While the regional lab has been creating the habitat-modeling database, each state has been contributing habitat association information through a Web interface. This interface is based on the hard copy form protocol and allows the habitat modeler to input the data into the database. The initial focus has been on species that may not need land cover to be modeled successfully. As of December 30, 2003, 86% of species data collection has been completed (716 of the 836 species). Approximately 80% of models completed also have their ranges delineated. The initial models will be reviewed internally and then be reviewed by species experts. Completion of land cover mapping is projected for June 2004 and will impact when models will be run and predicted animal habitat distributions will be mapped.

<u>Expert review</u> – The region is finalizing the process that will be used for expert review. Each state is identifying species experts and contacting these experts to gauge their level of interest. These lists will then be provided to the regional laboratory so that coordination of the regional approach can begin. Because of the number of species, we will be using a variety of methods to capture expert knowledge. These methods may include state expert review panels and regional expert review panels.

<u>Accuracy assessment</u> – The region will complete the standard gap analysis habitat modeling measure of agreement as well as a measure of agreement with existing species occurrence records. States are currently identifying qualified species lists for the standard measure of agreement. These lists will then be provided to the regional laboratory. In the next year, the regional laboratory and the Arizona project will identify a procedure to use existing data to measure the degree of concordance between habitat models and species occurrence.

Land stewardship mapping: Land stewardship mapping began regionwide during 2003 with the development of a regional workplan outlining the steps involved in the process. The New Mexico project hired the regional Stewardship Coordinator, Andrea Ernst, who began the process of data collection and consolidation within the five-state region. Base data layers were gathered from sources such as the Bureau of Land Management and the Conservation Biology Institute, which developed the Protected Areas Database. The land stewardship mapping effort is also coordinating with an existing project in Colorado that has previously gathered detailed land stewardship information for the state. The Stewardship Coordinator has also initiated contact with various federal and state agencies to gather more detailed internal management boundary information and associated management plans for the region.

Analysis: Analysis for SWReGAP will take place when the mapping tasks are completed. Land cover analysis will begin in July 2004 and animal habitat modeling analysis in January 2005.

Reporting and data distribution: All products derived from the Southwest Regional Gap Analysis Project are scheduled to be complete by approximately May 2005.

Tennessee

Draft data available from state. Review under way.

Contact: Jeanette Jones
Tennessee Wildlife Resources Agency, Nashville
Jeanette.Jones@state.tn.us, (615) 781-6534

Texas

Draft data available from state contact. Review under way.

Contact: Clint W. Boal
Texas Coop. Fish and Wildlife Research Unit
Texas Tech University, Lubbock
cboal@ttacs.ttu.edu, (806) 742-2851

Utah

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD. Remapping under way (see Southwest Regional GAP).

Vermont and New Hampshire

Draft data available from state contact. Review under way.

Contact: David E. Capen
University of Vermont, Burlington
dcapen@snr.uvm.edu, (802) 656-3007

Virginia

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Washington

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

West Virginia

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

Wisconsin

Project under way

Anticipated completion date: September 2004

Contact: Kirk Lohman
U.S. Geological Survey
Upper Midwest Environmental Sciences Center, La Crosse
klohman@usgs.gov, (608) 783-7550 x58

Land cover: Land cover mapping followed the Upper Midwest GAP protocol (ftp://ftp.umesc.usgs.gov/pub/misc/umgap/98-g001.pdf). Land cover mapping is completed, and a draft version is available from the USGS Upper Midwest Environmental Sciences Center (UMESC). With the assistance of NatureServe, the classification has been cross-walked to the NVCS.

Land stewardship mapping: The Wisconsin DNR has finished compiling data for state, county, and U.S. Forest Service lands. UMESC acquired coverages of DOI lands and compiled the complete stewardship coverage.

Reporting and data distribution: Land cover and stewardship coverages are available from UMESC. Contact Kirk Lohman at (608) 783-7550 x58 or klohman@usgs.gov.

Wyoming

Data on GAP Web site (http://www.gap.uidaho.edu/Projects/Data.asp) or CD.

AQUATIC GAP PROJECT REPORTS

Great Lakes Regional Aquatic GAP

Anticipated completion date: September 2007

Contact: Jana Stewart, Regional Coordinator
USGS, Middleton, Wisconsin
jsstewar@usgs.gov, (608) 821-3855

Michigan: Stephen S. Aichele, Co-PI
USGS, Lansing, Michigan
saichele@usgs.gov, (517) 887-8918

Dora R. Passino-Reader, Co-PI
USGS, Ann Arbor, Michigan
Dora_Reader@usgs.gov, (734) 214-7229

New York: James E. McKenna, PI
USGS, Cortland, New York
Jim_McKenna@usgs.gov, (607) 753-9391 x21

Ohio: S. Alex Covert, PI
USGS, Columbus, Ohio
sacovert@usgs.gov, (614) 430-7752

Wisconsin: Jana S. Stewart, PI
USGS, Middleton, Wisconsin
jsstewar@usgs.gov, (608) 821-3855

The Great Lakes Aquatic GAP began as a regional project in 2001, to be completed in 2007 in the states of Michigan (MI), New York (NY), and Wisconsin (WI). The Ohio (OH) Aquatic GAP pilot project has been in progress since early 2000 (see separate status report for Ohio in this section). The objectives of the regional project are to develop an Aquatic Gap Analysis for riverine systems in all eight states in the Great Lakes Region by 2009. Projects are planned sequentially, with new projects starting up when existing ones are nearing completion. In addition, a Coastal Pilot project is under way to develop a habitat classification framework, aquatic biota database, and initial gap analysis for near-shore coastal systems of the Great Lakes. Two pilot studies are currently under way in western Lake Erie and eastern Lake Ontario.

Development of a regionally consistent database and spatial data layers, with uniformity across state boundaries, has been a major focus of the Great Lakes Aquatic GAP project. A central relational database is being developed to accommodate stream habitat characteristics, aquatic biota, and habitat affinity data in a consistent manner across all states in the Great Lakes region. Processing of stream habitat characteristics has also been coordinated across state boundaries, so assessments can be completed at the regional, state, and local levels.

Central Database development: The Central Database uses Oracle Discoverer and Oracle 9i software with capabilities for data sharing on the Web through user-client and Web-interface clients and is housed at the USGS Great Lakes Science Center in Ann Arbor, MI. Fish data have been acquired, organized, formatted, and reviewed for MI, NY, OH, and WI, with plans to load all data into the Central Database by June 2004. Habitat characteristics of stream segments for MI and WI will be loaded by July 2004 and those of streams in the Great Lakes drainages of NY by September 2004. Fish life history and habitat affinity information is in the process of being acquired and will be loaded into the Habitat Affinity section of the Central Database for validation of predicted fish distributions and analysis of fish community ecology. Plans are under way to develop additional data tables to store habitat characteristics for the Coastal Pilot studies. Future plans also include acquisition and review of invertebrate and freshwater mussel databases for possible incorporation into the Central Database.

A Web-based map application prototype has been developed to produce dynamic species distribution maps for the WI Aquatic project, with future plans to apply the map interface to the Central database. Using a relational database and spatially enabling the data (Oracle Spatial), a user is dynamically able to query the database via a Web browser through a graphical user map interface.

Stream habitat classification and modeling: The stream habitat classification methods developed and used by the Missouri (MoRAP) and OH Aquatic GAP projects were reviewed and modified for use by the Great Lakes Aquatic GAP project to improve habitat characterization, model predictions and analysis, and regional consistency. The modifications were made to better reflect factors contributing to physical habitat at all scales, including the channel segment, riparian buffer, and entire upstream contributing area of each channel segment. These modifications include (1) preservation of interval/ratio data rather than grouping all variables into categories, (2) calculation of habitat characteristics for not only the channel segment but the riparian buffer and watershed for each segment, (3) preservation and attribution of lake and double-line stream features, (4) addition of land cover, (5) and development of temperature and flow models to predict real values of temperature and flow for each stream segment. ....

Numerous processing scripts were developed by the MI Aquatic GAP team to expedite processing of stream habitat characteristics and provide consistency in methods and results. Streams have been attributed with habitat characteristics in MI and WI and will be 80% complete for the Great Lakes drainages of NY by June 2004. As part of habitat characterization, regression models have been developed to predict stream temperature for every stream segment in MI and WI. Fish sample locations will also be linked to stream segments by June 2004. Initial exploration of modeling methods for fish-environment relationships has begun, and a significant array of example results should be available for evaluation by September 2004.

Coastal GAP pilot project: A conceptual framework for identification and classification of coastal habitat types has been developed and applied to the western Lake Erie pilot study area. Databases of fish distributions in western Lake Erie and eastern Lake Ontario have been acquired. A substantial amount of fieldwork, designed to help assess the efficacy of the classification framework and to collect data from unsampled and important habitat types, was completed.

Outreach and meetings: A USGS fact sheet describing the Great Lakes GAP project was published in June 2003 and is available on the Great Lakes GAP Web page (http://www.glsc.usgs.gov/GLGAP.htm). Numerous papers describing the Great Lakes Aquatic GAP project were presented in a special session entitled "Biodiversity Conservation in the Great Lakes Region" at the International Association for Great Lakes Research in Chicago, IL, in June 2003. A Great Lakes Aquatic GAP poster was presented at the Society for Conservation Biology meeting in Duluth, MN, in June 2003, and a number of papers and a poster were presented at the National GAP meeting in Fort Collins, CO, in October 2003. Contributed and symposia papers describing the Central Database, preliminary modeling results, and the overall projects have been submitted for presentation at the 134th Annual Meeting of the American Fisheries Society in Madison, WI, in August 2004.

The Great Lakes Aquatic GAP team has worked closely with an EPA Star Grant group on ecological classification of the rivers of IL, MI, and WI to share expertise and develop regional methods for habitat classification and species modeling. Individual state projects continue to work closely with stakeholder agencies, including the MI Institute of Fisheries Research (MIFR), the New York State Department of Environmental Conservation, and the WI Department of Natural Resources (WDNR).

Hawaii Aquatic GAP

Anticipated completion date: May 2005

Contact: Michael H. Kido
University of Hawaii, Hawaii Stream Research Center, Kapaa
mkido@hawaii.edu, (808) 956-4907

In the past year the HI-GAP project has established a methodology for mapping the distribution of aquatic species. The methodology is unique and accounts for Hawaii's topographic complexity and climate. The model has been implemented on Kauai and Maui, and preliminary results indicate the methodology is producing an accurate indication of alien and native aquatic species distribution. The model is also providing indicators of stresses on individual watersheds to identify watersheds with degraded stream habitat. The project has also been successful in bringing together stream researchers in Hawaii through interest meetings and positive collaboration among partners.

Analysis: The modeling methodology being implemented utilizes cluster analysis modeling to group watersheds per island into similar entities. The clustered watersheds are then analyzed, and aquatic species distribution along the stream continuum is mapped. The watershed clustering is based on variables derived from the GIS, which include average slope, average elevation, aspect values, watershed size, stream length, percentage of perennial streams, percentage of intermittent streams, and land cover types. Each value is summarized per watershed and used as inputs to cluster the watersheds into similar entities. Initial results indicate between 5 and 8 different watersheds exist per island.

The aquatic species distribution along each stream is then mapped, using GIS attributes derived for all streams. The variables used to define species ranges include slope variation, elevation ranges, waterfall locations, and distance from the mouth of the stream. Based on extensive observation data, each aquatic species is assigned physical attributes that define its desired habitat and range. The assigned habitat parameters are then used to define the distribution of each aquatic species along the stream continuum. The results for the Island of Kauai indicate the methodology is accurate.

Future work: In an effort to protect the remaining native aquatic species and identify the location of alien aquatic species, the Hawaii Gap Analysis Project will be providing the derived data on an ArcIMS server to allow public access to this information for academic research and public awareness. HI-GAP is also planning on conducting several analyses to prioritize streams for future conservation. The overall goal of the HI-GAP project is to approach the assessment of conservation from an integrated system of terrestrial, aquatic, and marine ecosystems.

Lower Missouri River Basin Aquatic GAP

a. Iowa

Anticipated completion date: December 2004

Contact: Kevin Kane
Iowa State University, Ames
kkane@iastate.edu, 515/294-0526

Recent achievements

a) Biological assessment:

- completed draft range maps for all fish species
- completed test models for Eastern Broadleaf Forest subregion using AnswerTree software, including determining habitat variables
- completed the professional review of one third of all species
- submitted an article to the 2003 GAP Bulletin
- presented a poster on the development of the biological inventory database at the Iowa Water Monitoring Conference
- submitted an abstract for the Ecological Society of America annual meeting.

b) GIS coverages:
The NHD linework was received from MoRAP in two separate coverages. MoRAP had processed the IA linework we sent them early last year to produce a variety of attributes that would possibly be used to model fish distributions. Roughly 25 attributes were created and assigned variables using ARC aml scripts. The two coverages were modified to be able to be mapjoined, and the resulting coverage was processed to add three additional attributes deemed necessary for modeling. Not all 28 attributes will be used in modeling; a subset of about 9 variables will be used.

Several attributes have been identified as necessary already (flow, temperature, size discrepancy, gradient, and downstream link), and the linework is being checked for valid values in those attributes. Flow is a variable directly from the original NHD, and it has over 900 zero values in the 57 watersheds for Iowa GAP. Those are being corrected using aerial photography and 24K topo maps. It is possible several secondary channels that were not part of the attribute calculations at MoRAP will have to be used in our final modeling due to the need to keep the fish samples on those reaches. The model variable values for those secondary channels will have to be calculated, and the process and source data are being investigated and compiled.

Tasks to complete

We have made excellent progress in many areas thus far, including what has been referred to as one of the premier state historical fish record databases in the country. The spatial data portion of the project has progressed more slowly, however. We just received what we hope is the final river database from MoRAP, our data provider.

We are leveraging Aquatic GAP funding with the Iowa DNR, which continues to fund the development of the complimentary Iowa Rivers Information System (IRIS) this year.

a) GIS databases and analysis:

- Prepare final database
- Categorize variables based on preliminary biologic analysis
- Create analysis variables
- Run all analyses
- Create and print preliminary analysis maps
- Create and print final analysis maps
- Update stewardship layer
- Stewardship/prediction analysis

b) Biological assessment:

- Complete aquatic inventory data set
- Compile aquatic range maps from current and historical data
- Build habitat models
- Professional review of ranges and habitat
- GIS analysis for predictive distributions (three months)
- Report writing

c) Access and Interface:

- Create IA Aquatic GAP Web site for information and data/analysis/map delivery
- Create IMS interface for users (extension to IRIS IMS)

b. Kansas

Anticipated completion date: May 2005

Contact: Keith Gido
Kansas State University, Manhattan
kgido@ksu.edu, (785) 532-6615

The Kansas Aquatic GAP project is currently in its third year. To date we have acquired data on the distribution and abundance of fish and mussel species from over 4,000 localities in Kansas. These data have come from a variety of sources including the Kansas Department of Wildlife and Parks, Kansas Department of Health and Environment, University of Kansas Museum of Natural History, Kansas Natural Heritage Inventory, Sternberg Museum of Natural History, The Nature Conservancy, and various individuals with scientific collections in the state. With the help of MoRAP, we have modified the existing National Hydrology Database in our state to include a suite of environmental variables for over 100,000 individual valley segments. These data sets are being combined to construct predictive models of species occurrences across the state. In addition, we are working with other colleagues involved in Aquatic GAP projects in the Missouri River Basin to standardize efforts among states.

Our data have been compiled, and species distribution maps have been constructed for peer review and are available on our Web page (www.ksu.edu/aquaticgap). As a pilot study, we have evaluated several modeling approaches and different suites of environmental variables in the Big Blue River basin. Although some species models do not perform well in this region, several models for species of special concern worked quite well and are promising tools for conservation. We are continuing to explore (1) the feasibility of different modeling approaches, (2) the use of different independent variable sets, and (3) what scale of analyses is appropriate for this region.

c. Missouri
Anticipated completion date: December 2005

Contact: Scott Sowa
MoRAP, University of Missouri, Columbia
scott_sowa@usgs.gov, (573) 441-2791

MoRAP is completing the aquatic ecological classification system for the Lower Missouri River Basin, which includes the states of Iowa, Kansas, and Nebraska. This past year we completed a draft version of Aquatic Subregions and Ecological Drainage Units for the lower basin. These draft units will have to be verified and modified through analyses of the biological data sets currently being developed in each of the states. We have completed the Valley Segment Classifications for Iowa and Kansas, and we are currently working on this same coverage for Nebraska. Continued funding from the Gap Analysis Program has allowed us to begin the classification of Aquatic Ecological System Types in all three states. We will also be generating human stressor statistics for each of the AESs and generating local, upstream riparian, and overall watershed ownership statistics (by GAP stewardship category) for each of the stream segments within the lower basin. All of these data sets are scheduled to be completed in December 2005.

Missouri Aquatic GAP

Anticipated completion date: October 2004

Contact: Scott Sowa
MoRAP, University of Missouri, Columbia
scott_sowa@usgs.gov, (573) 441-2791

Aquatic ecological classification: We are currently working on detailed biophysical descriptions for each of the Aquatic Subregions, Ecological Drainage Units, and Aquatic Ecological Systems. A draft version of the specific classification procedures for each level of the hierarchy has been completed.

Species modeling: We developed 517 separate models in order to generate predictive distribution maps for 315 species of fish, mussels, and crayfish. The vast majority of predicted models were developed empirically using Decision Tree Analyses. However, due to data limitations, some models were based on contingency-table analyses combined with habitat-affinity information compiled from existing literature. A hyperdistribution MS Access database has been completed, which spatially links to our statewide valley segment coverage via unique segment identifiers. This relational database also contains taxonomic, ecological, and conservation status information for each species.

We also completed an accuracy assessment of our models using independent data sets for each taxon. Table 1 provides a breakdown of the overall, commission, and omission errors. Omission errors are relatively low, averaging less than 10%, while commission errors are relatively high, averaging over 50%. However, these accuracy statistics are very misleading. There are many problems associated with this accuracy assessment related to spatial and temporal sampling "inadequacies" of the independent data sets and with the inherent difference in what we are trying to predict (i.e., biological potential) versus the fact that most of the stream segments sampled in these independent data sets were degraded to some degree. In fact, some of the sites are highly degraded, and in such instances we would expect very little correspondence between our predicted assemblage and the assemblage that presently occupies the site. Based on a separate evaluation of two fish collection sites, where the data are more temporally and spatially comprehensive, we found our overall accuracy to increase to nearly 70%, mainly due to a significant decrease in commission errors. A proper evaluation of the accuracy of our models will require a separate project that identifies relatively high-quality sites, which are then sampled intensively throughout long stretches of stream during several seasons and over a period of several years.

Table 1. General accuracy assessment statistics for predictive models based on assessment of independent data sets.

Taxa	Overall	Commission	Omission
Crayfish	48	52	11
Fish	51	48	10
Mussel	36	64	6
Average	45	55	9

Habitat-affinity reports were completed for each species. These reports can be viewed at the MoRAP Web site, http://www.cerc.usgs.gov/morap/projects.asp?project_id=1. These are stand-alone reports, which provide images of the species, the predicted distribution map (for species that are not state-listed as either rare, threatened, or endangered), state range description, habitat affinity information extracted from the literature, the predictive model(s), and literature pertaining to that species. A draft version of the methods used to develop the predictive models was also completed.

Stewardship: Identifying conservation gaps for riverine ecosystems is no straightforward task. Simply assessing whether a stream segment is within public ownership provides insufficient information, since each segment is influenced by everything occurring within the surrounding watershed or upstream riparian area. Even a segment flowing through a national park is not really being conserved if most of its watershed is urbanized. For the Missouri Aquatic GAP Project we calculated three sets of ownership statistics (by GAP stewardship category) for every individual stream segment: (1) local ownership, (2) percent of the upstream riparian area, and (3) percent of the watershed in public ownership. Since streams do not respect political boundaries, we had to combine the stewardship coverages from Kansas and Iowa with that of Missouri in order to generate these statistics. All three statistics are important for assessing conservation gaps and provide decisions makers with a suite of information for effective conservation planning.

Human stressors/threats: In addition to assessing multiple forms of public ownership, we must also consider existing human stressors, since even ownership of a substantial portion of a watershed does not ensure effective conservation. To account for this, we generated statistics for nearly 50 individual human stressors (e.g., percent urban, lead mine density, degree of fragmentation) for each Aquatic Ecological System, which are ecologically defined watershed units. We then used correlation analyses to reduce this overall set of metrics into a final set of 11 relatively uncorrelated measures of human disturbance or stress. Relativized rankings (range 1 to 4) were then developed for each of these 11 metrics. A rank of 1 is indicative of relatively low disturbance for that particular metric, while a rank of 4 indicates a relatively high level of disturbance. These rankings were based on information contained within the literature or simply quartiles when no empirical evidence on thresholds was available. For instance, rankings for percent urban were 1: 0-5%, 2: 6-10%, 3: 11-20%, and 4: >20%, based on the collective results of various studies that have examined the effects of urban land cover on the ecological integrity of stream ecosystems. However, existing research for percent agriculture has not identified clear thresholds, suggesting a more or less continual decline in ecological integrity with each added percentage of agriculture in the watershed. For this measure of human stress we simply used quartiles, 1: 0-25%, 2: 26-50%, 3: 51-75%, and 4: >75%. The relativized rankings for each of these 11 metrics were then combined into a three-number Human Stressor Index. The first number reflects the highest ranking across all 11 metrics (range 1 to 4). The last two numbers reflect the sum of the 11 metrics (range 11 to 44). This index allows you to evaluate both individual and cumulative impacts. For instance, a value of 418 indicates relatively low cumulative impacts (i.e., last two digits = 18 out of a possible 44), however, the first number is a 4, which indicates that one particular land use is potentially severe. This index is an admittedly crude measure of human disturbance, however, it is well suited for a coarse-filter assessment since it does act as a red flag. This purpose of the HSI is to further evaluate those locations that appear to be well represented within the existing matrix of public lands, since ownership does not ensure effective conservation within riverine ecosystems. Using this index, we have found a handful of instances where stream segments flowing within GAP 1 or 2 lands are not being adequately conserved due to human disturbances occurring outside the boundaries of the public lands.

Gap analysis: We found a major error in the STATSGO soil coverage for Missouri, which affected the classification of our Aquatic Ecological Systems and resulted in a trickle-down effect on our already completed gap analysis for Missouri. We have since fixed the problem with the STATSGO coverage, reclassified our AES-types, and are now redoing the gap analysis. Although the errors in the STATSGO coverage were major and will change the specific geographic context of our gap analysis, the general conclusions from the original analysis will not change.

It is quite evident from our original analysis that only a tiny fraction of the stream resources in Missouri are being adequately represented in the current matrix of public lands. This statement does not pertain so much to the amount of the stream resource base currently within public lands as it does to the spatial arrangement of public ownership. From a local ownership perspective, 5% (9,365 km) of total stream length in Missouri (173,074 km) are currently flowing through public lands. Based on our definition of what constitutes a gap (in terms of local ownership), this is nearly twice as much as is minimally required to represent the full spectrum of ecosystem, community, and species diversity within the state. The problem lies in the fact that ownership patterns are highly fragmented and do not holistically represent interacting systems, as well as being overly redundant in terms of their representation of ecosystem and community types.

The benefits of our project to future conservation efforts are not so much related to our ability to document conservation gaps. More importantly, we have developed the data necessary to make informed decisions that incorporate fundamental principles of stream ecology and conservation biology into an overall conservation planning strategy. We are currently working with the Missouri Department of Conservation to use these data to develop a statewide comprehensive conservation plan for conserving riverine biodiversity. For each of the 15

Ecological Drainage Units in Missouri we are identifying a set of focus areas that collectively represent the full spectrum of abiotic and biotic diversity. These focus areas then serve as a geographic template for a variety of conservation actions, including new land acquisitions, changes in management designations, or private land conservation initiatives. Although not completed for the entire state, present results suggest that an idealized network of reserves, which focuses on local ownership of critical stream segments, would require approximately 5,000 to 6,000 km of stream to represent the full spectrum of biological diversity in the state. Certainly, this says nothing about the difficulty of securing enough land to protect the watersheds of these critical segments, which would be a mind-boggling number (millions of acres) for a Midwestern landscape. However, it is our contention that efforts must initially focus on protecting critical stream segments in public lands or through intensive private land conservation, since even the most ambitious watershed protection measures can be circumvented by local disturbances (e.g., channelization, point-source pollution, impoundments, etc.) to the stream segments to be protected.

Ohio Aquatic GAP

Anticipated completion date: March 2005

Contact: S. Alex. Covert, Ohio GAP Coordinator
U.S. Geological Survey, Columbus, Ohio
sacovert@usgs.gov, (614) 430-7752

Animal modeling: The crayfish database was completed in 2003. An expert review of distributions of 92 freshwater mussel and clam species and 20 crayfish species was also completed. An effort to model potential species distributions for freshwater mussels, clams, and crayfish was started in 2003 using the same modeling approach as was used for fish in Ohio, i.e., GARP. An analysis that integrates all aquatic biota will be completed in 2004.

Land stewardship mapping: All digital maps of Ohio conservation lands were obtained and compiled into one map. Each land parcel was attributed with a GAP land-status code. The map was reviewed and finalized in 2003.

Analysis: Potential fish distributions for 148 species were displayed at the USGS 14-digit Hydrologic Unit (HUC) for each of three stream-size classes. The number of unique fish species predicted for each 14-digit HUC (1,790 in Ohio) was calculated and compared to watersheds in the larger, 8-digit HUC (44 in Ohio) that the 14-digit HUC occupied. This effort helped to identify watersheds with high numbers of predicted fish species, as well as to ensure a distributed assessment unit throughout the state. The number of predicted fish species, areas predicted for Ohio's endangered species, and areas predicted to have high-quality, cold-water fish assemblages will be used to perform a gap analysis in 2004. Areas that have never been sampled were identified to help guide future sampling efforts. Likewise, areas identified as having high species numbers despite having been sampled abundantly with no captures of particular species were identified for further analysis of limiting conditions.

Reporting and data distribution: A data-CD entitled "Fish Distribution and Valley Segment Type Data from Ohio Aquatic Gap Analysis Project (GAP)," was published in 2003. This CD contains Ohio fish sampling data and derived valley segment data that was used to model potential fish species distributions. Ohio Aquatic GAP has presented methods and progress to the Upper Midwest Environmental Science Center in Onalaska, Wisconsin, at the joint annual International Association for Great Lakes Research (IAGLR) and International Lake Environment Committee (ILEC) meeting, at the National GAP meeting in Fort Collins, Colorado, to the Ohio EPA, and at two stakeholder meetings.

Southeast Aquatic GAP

a. Alabama

Anticipated completion date: August 2004

Contact: Elise Irwin
Alabama CFWRU, Auburn University, Auburn
eirwin@acesag.auburn.edu, (334) 844-9190

We are in the final stages of project completion for the entire Tallapoosa River basin. We have recently completed QA/QC of newly delineated 12-digit HUCs, which is our mapping and modeling unit. All landscape-level GIS layers are complete, including geology, hydrology (1:100,000 NHD), LU/LC (Anderson Level I classification; 1992 NLCD) and a small impoundment layer (from DOQs). Watershed-level characters are also complete for each 12-digit HUC. These include but are not limited to drainage density, road density, and physiographic province. Stream reach characters have also been quantified, including site elevation, stream gradient and aspect, link magnitude, and downstream link magnitude. Fauna data include fish collections from over 300 sites in the basin. Crayfish, snail, and mussel data are sparse but available from over 100 sites. We are estimating detection probabilities for species using the program CAPTURE. Detection probabilities will be used as model weights in K-nearest neighbor (KNN) models to account for incomplete detection of species. We concluded that KNN classification was more appropriate than other techniques to determine significant predictors for species occurrence. These empirical models will be valuable for development of decision support systems.

b. Georgia
Completed. Draft data available from state contact.

Contact: James Peterson
Georgia CFWRU, University of Georgia, Athens
peterson@smokey.forestry.uga.edu, (706) 542-6032

Upper Missouri River Basin Aquatic GAP

Anticipated completion date: October 2004

Contacts: Jonathan A. Jenks and Charles R. Berry, Jr.
South Dakota State University, Brookings
jonathan_jenks@sdstate.edu, (605) 688-4783
charles_berry@sdstate.edu, (605) 688-6121

Status: We have acquired all necessary data sets from state, federal, and international agencies for physical habitat, fish distribution, and stewardship. We have completed attributing valley segments with ten physical habitat affinities, which include temperature, stream size, flow regime, channel gradient, size discrepancy, floodplain interaction, surficial geology, elevation, stream connectivity, and groundwater input. We are now performing a quality check of valley segment habitat attributes and are about 75% finished. Fish distribution data is also undergoing quality control. We have edge-matched and merged land cover data from states and provinces and are in the process of merging stewardship layers. We have completed delineation of watersheds similar to 10-digit hydrological units for portions of North Dakota within our study area and are combining these with 10-digit hydrological units provided by the North Dakota Department of Health, Division of Water Quality. We have completed field sampling of fish within the Frenchman River watershed, which flows from Saskatchewan, Canada, into Montana, and also sampled the Sweet Grass Creek watershed within Montana. Field data is being summarized and will be used to test the accuracy of fish distribution models for these areas. We also have collected invertebrate data from sampled stream reaches in these watersheds and will use this information to model invertebrate distributions and biodiversity.

Future plans: We will continue to quality-check our valley segment and fish distribution databases. We plan to use decision tree analysis (Answer Tree, SPSS) to produce fish-habitat models similar to methods used by the Lower Missouri River Basin Aquatic GAP Project. Fish will be modeled for seven regions based upon ecoregions and major drainages. Accuracies of fish-habitat models will be evaluated using field data and also statistical methods (e.g., boot strapping). Fisheries experts from each region will also review models and fish distribution maps.

Upper Tennessee River Basin Aquatic GAP

Anticipated completion date: May 2004

Contacts: Paul Angermeier
Virginia CFWRU, Virginia Tech, Blacksburg
biota@vt.edu, (540) 231-4501

Jeff Waldon
Conservation Management Institute, Virginia Tech, Blacksburg
fwiexchg@vt.edu, (540) 231-7348

In 2001, researchers from the Department of Fisheries and Wildlife Sciences and the Conservation Management Institute of Virginia Tech began an aquatic gap analysis of the upper Tennessee River basin (UTRB), which is shared by Virginia, Tennessee, North Carolina, and Georgia. In 2003 we assembled available GIS coverages on (*a*) biota (mostly fishes but some data on mussels and crayfishes), (*b*) land and water use (e.g., dams, roads, effluents, urban areas, row crops, pastures), and (*c*) physical landscape features (e.g., physiography, elevation, hydrography). We are using these coverages to develop models that (*a*) predict species occurrence and (*b*) estimate threat to watershed health.

We are building two types of models to predict species distribution. The type built for a particular species depends on data availability. We identified 524 assemblage samples containing a total of 126 fish species, 71 samples containing a total of 11 crayfish species, and 66 samples containing a total of 5 mollusk species. For species occurring in many samples we are building a suite of logistic regression models, then choosing the best model. Useful predictor variables differ considerably among species. Elevation (73%), state (67%), and stream order (61%) are most frequently useful for predicting fishes. Elevation (81%) was the most frequently useful predictor of crayfish occurrence. Stream order (60%), elevation (60%), and sinuosity (60%) were most frequently useful in predicting mussel occurrence. We are also developing less precise (descriptive) models based on species accounts in recent books on the region's fish fauna. These models largely reflect known distributions as described by drainage, physiography, and stream size. For poorly sampled species these descriptive models are the only model type available. For well-sampled species the logistic regression models can be used to calibrate the reliability of the descriptive models.

A main research focus is to develop more powerful protocols to assess threats to aquatic biota. We anticipate that the standard stewardship data layer used in gap analyses, which is based on land ownership, will not provide an informative assessment of protective status. Threats to biota vary in scope of origin (nonpoint vs. point source), frequency of occurrence (accidental spill vs. permitted effluent), and severity (heavy metal contamination vs. nutrient enrichment).

Moreover, most threats to aquatic biota emanate from outside the aquatic environment. Thus, we have developed an integrative protocol that assesses a wide array of threats to stream biota and converts degree and extent of threat into a numerical form that facilitates ranking among watersheds. We have assembled georeferenced data layers on dams, roads, railroads, pipelines, waste disposal sites, permitted water discharges, agriculture, urban areas, and industrial sites. We ranked each human activity based on its potential impact on flow regime, water quality, habitat quality, energy sources, and biological interactions. Then we estimated the frequency of each activity within individual catchments. Finally, an index based on impact and frequency was computed for each catchment. This protocol enabled us to develop maps of severity for individual threats as well as for cumulative threat, and to identify large-scale patterns of protective status across the entire UTRB. We are in the final stages of map development and are preparing to examine how protective status based on stewardship data compares to protective status based on our new protocol.

The Gap Analysis Bulletin is published annually by the USGS Biological Resources Discipline's Gap Analysis Program. The editors for this issue are Elisabeth S. Brackney and Michael D. Jennings. To receive the bulletin, write to: Gap Analysis Bulletin, USGS/BRD/Gap Analysis Program, 530 S. Asbury Street, Suite 1, Moscow, ID 83843, fax: (208) 885-3618, e-mail: brackney@uidaho.edu. A digital version of the Bulletin, containing additional graphics, is available on the Internet at http://www.gap.uidaho.edu/gap/Bulletins/12/.

Suggested citation: Brackney, E.S., and M.D. Jennings, editors. 2003. Gap Analysis Bulletin No. 12. USGS/BRD/Gap Analysis Program, Moscow, Idaho.