Preface

Version: 2.0.1
Date: February 2007
Peer-reviewed: No

J. Michael Scott
Director, Gap Analysis Project
University of Idaho, Moscow, ID

The purpose of this manual is to provide guidelines to those conducting a Gap Analysis project. The topics covered include program management, building partnerships, data acquisition, constructing data sets, vegetation and vertebrate standards, metadata, validation, analysis, and implementation. The manual is intended to be dynamic in that parts of it will be changed as improvements to the Gap Analysis process are made. Each section has a "version" noted in the footer. Each Gap Analysis state project will be sent new chapter versions as they become available. When users are communicating about handbook material, they are cautioned to refer to the version number of the material they are using.

Gap Analysis is a "bottom-up" approach with national guidelines that allow for creativity and collaboration at the state and local level, where most land management decisions are made. National standards are, however, of enormous importance if we are to provide ecologically meaningful information useful at the bioregional and national levels. Standards are imperative if our information is to be equivalent across the entire range of a species' or vegetation type's occurrence.

Although Gap Analysis is coordinated by the National Gap Operations Office in Moscow Idaho, it is composed of (to-date) more than 200 collaborating organizations including private businesses, special interest groups, universities, as well as state, local, and federal governments. One of the objectives of the Gap Analysis Project is to generate digital thematic maps of existing land cover habitat types and distributions for each vertebrate species at a scale of 1:100,000. By necessity, the effort is multidisciplinary, involving ecologists, geographers, zoologists, botanists, statisticians, sociologists, and others. All map products must be ground-truthed to provide users with an index of reliability.

The intent of Gap Analysis is to provide focus and direction for proactive, rather than reactive, land management activities at the community and landscape levels. For example, GAP provides an ecological context for a hierarchical approach to land management and more detailed ecological studies in the future. We believe that Gap Analysis is one step in a comprehensive land management planning effort that transcends political boundaries.

Gap Analysis products have many different uses. One of the most important uses is to provide an overview of the distribution and management status of selected components of biodiversity (i.e., land cover types and terrestrial vertebrate distributions). The Gap Analysis Project seeks to identify vegetation types and species that are not adequately represented in the current network of special management areas. These are the "gaps" in the present-day overall mix of conservation lands and conservation activities. This information is intended to be used by decision makers for proactive land management planning which will hopefully lead to fewer species becoming endangered, and thus reduce the number of future conflicts regarding natural resource issues.

It needs to be stressed, though, that Gap Analysis is intended to complement, not replace, the species-by-species approach to preserving biodiversity which is so critical to the survival of species now nearing extinction. **The main goal of Gap Analysis is to prevent additional species from being listed as threatened or endangered.**

Because a species or a habitat type is common or abundant today does not guarantee that it will not become endangered with extinction in the future. Gap Analysis is essentially an expanded "coarse filter" approach to biodiversity protection (Noss 1983, Scott et al. 1987, 1993). The vegetation types serve as a coarse filter for land management planning. Landscapes with great vegetation type diversity often are those with high edaphic variety or geographic relief. Where there is substantial geographic relief, a nearly complete spectrum of vegetation types known from a biological region may occur within a relatively small area. Such areas provide habitat for many species, including those that depend on multiple habitat types to meet life history needs (Diamond 1986, Noss 1987).

The intuitively appealing idea of conserving most biodiversity by maintaining examples of all natural community types has never been fully tested empirically. Furthermore, the spatial scale at which organisms use the environment differs tremendously among species and depends on body size, food habits, mobility, and other factors. Hence, no coarse filter can be a complete surrogate of biodiversity protection status and needs. However, species that fall through the mesh of the coarse filter, such as narrow endemics and wide-ranging mammals, may be captured by the safety net of the fine filter (Noss 1990). Again, community-level (coarse filter) protection is a complement to, not a substitute for, protection of individual species which are imperiled.

The analysis part of Gap Analysis is, at this time, often performed at the state level without incorporation of data from adjacent states, because most state projects are still developing basic data. There is, however, an immediate need and strong demand to provide local managers with the information they need to manage natural resources across political boundaries. By merging the state data sets into ecologically defensible units, biologically meaningful conservation strategies may be created by examining the occurrence of a species of vegetation types across an entire ecoregion. Finally, it can't be over-emphasized that there is no single measure of biodiversity that can be used to create a biodiversity plan. Thus, planning must be hierarchical (genes to landscapes) and iterative in nature. Gap Analysis is developing spatial information on the species and natural community elements within that hierarchy.

We solicit any comments you might have for improving the guidelines. Please send these to:

Gap Analysis Project
Idaho Cooperative Fish and Wildlife Research Unit
Department of Fish and Wildlife
University of Idaho
Moscow, ID 83844-1141

Phone: 208-885-6336
FAX: 208-885-9080

Literature Cited

Diamond, J. 1986. The design of a nature reserve system for Indonesian New Guinea. Pp 485-503 In M.

E. Soulé (ed.), Conservation Biology: The Science of Scarcity and Diversity. Sunderland, Mass.: Sinauer Associates.

Noss, R. F. 1983. A regional landscape approach to maintain diversity. BioScience 33(11):700-706.

Noss, R. F. 1987. Protecting natural areas in fragmented landscapes. Natural Areas Journal 7:2-13.

Noss, R. F. 1990. Indicators for monitoring biodiversity : a hierarchical approach. Conserve. Biol. 4:355-364.

Scott , J. M., J. J. Jacobi, and J. E. Estes. 1987. Species richness: a geographic approach to protecting future biological diversity. BioScience 37:782-788.

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. Wright. 1993. Gap Analysis: a geographic approach to protection of biological diversity. Wildlife Monograph 123, 1-41.

This GAP Handbook chapter is provided here for general reference but see the Standards Chapter for current GAP data standards.

Cooperative Structure and Funding for State GAP Analysis Projects

Version: 2.01
Date: 16 February 2007
Peer-reviewed: No

Michael D. Jennings
Idaho Cooperative Fish and Wildlife Research Unit
Moscow, ID

The application of comprehensive geographic information to the conservation of all species and natural communities, rather than an ad hoc approach for single species, represents a philosophical change in natural resource management of historical proportions. Gap Analysis is a method for developing the information needed to actualize this change. The organizational means by which Gap Analysis projects are implemented, it turns out, is equally significant as the underlying philosophical change. To get the job done, the manner in which Gap Analysis projects are implemented is fundamentally different from a single-agency approach.

Of necessity, Gap Analysis is a collaborative, "bottom-up" effort focused on pragmatic near-term conservation gains. Although coordinated by the National Biological Service, the National Gap Analysis Project consists of over 200 cooperating federal, state, and private organizations nationwide as of mid-1994.

Presently, no single organization has the intellectual, technical, or financial capability to carry out an effort of the magnitude and the degree of resolution required for the conservation of native biodiversity of the United States at successive geographic scales and categories of biotic organization. Nor is there reason to expect that a single-organization approach would result in broad acceptance, thus coordinated use, of a cohesive conservation strategy by the many different institutions that, in one way or another, affect the management of biological diversity.

In response to an imperative for conservation and because of limited funding, today as never before, local, state, federal, and private organizations are working together to develop and share information about our natural resources. The sections that follow describe how typical state-level Gap Analysis projects are usually structured, funded, and supported cooperatively.

Structure

Most, though not all, Gap Analysis projects operate under the umbrella of a state Cooperative Fish and Wildlife Research Unit (CFWRU). These research units are joint ventures among the United States Geological Survey (USGS), state fish and wildlife agencies, and state universities and are located at a state university. Our experience shows that the mix of talent from universities, state agencies, and federal agencies, as accessed through the CFWRU structure, is frequently what provides state Gap Analysis projects with their critical mass (Figure 1). However, since not every state has a CFWRU, some Gap Analysis projects are organized through different lead agencies. For example, the Gap Analysis project that covers Maryland, Delaware, and New Jersey is being implemented cooperatively by the Maryland Department of Natural Resources, U.S. Fish and Wildlife Service, and Johns Hopkins University. In all cases, state Gap Analysis projects result from state- or regional-level demand.

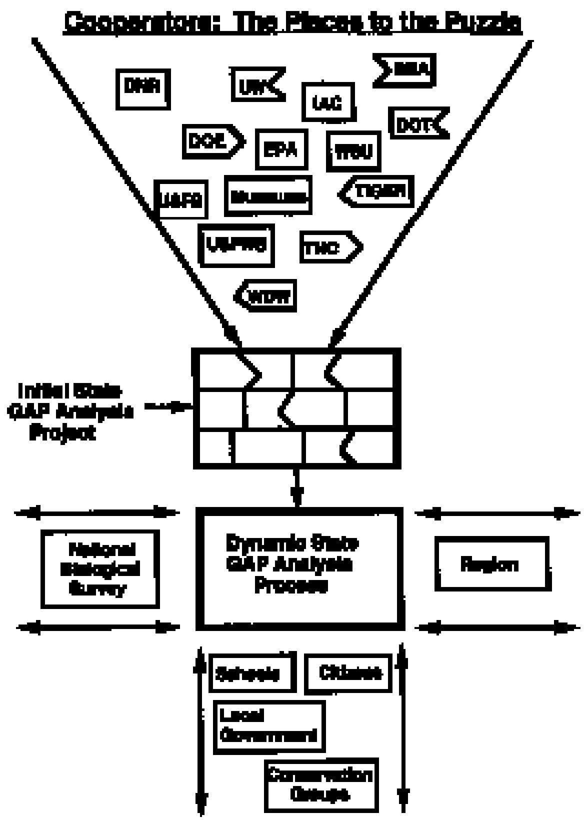


Figure 1. The mix of talent and information from universities, state agencies, federal agencies, museums, etc., provides state Gap Analysis projects with their critical mass. This then feeds two-way information processes with schools, local governments, conservation groups, and others at the regional level. At the national level, this is joined with information from adjacent regions to feed broader processes and perspectives. Graphic courtesy of Karen Dvornich, Washington Gap Analysis Project, Washington Cooperative Fish and Wildlife Research Unit, University of Washington.

Determination of a lead agency and cooperators depends upon how those institutions at the state and regional level organize themselves to answer the following questions: (1) How are the state's biological resources geographically distributed? (2) How are areas managed for the long-term persistence of those resources geographically distributed? (3) How well are all species and natural communities represented within areas managed for their long-term persistence? (4) What is the multi-state regional biogeographic context for the state's biodiversity and biodiversity management areas? The Gap Analysis Project seeks to be user-driven and to meet the demand for, and maximize the utility of, biogeographic information.

State Gap Analysis projects are carried out as research projects. They usually consist of a set of cooperating organizations (Table 1) which form an oversight committee. The cooperators may also form subcommittees for technical advice, outreach, and inter-agency cooperation. Projects are usually administered by a Principal Investigator (PI) or Co-PIs. Day-to-day project implementation is carried out by a Project Leader. Staff usually consist of a remote sensing and geographic information systems expert, a botanist, a zoologist, and a coordinator (Figure 2).

Table 1. Organization types that typically constitute state Gap Analysis Project cooperators.

State fish and wildlife agencies	Herbaria
State departments of natural resources	Bureau of Land Management
State departments of conservation	National Forests
State departments of planning	Universities
State departments of forestry	Environmental Protection Agency
State associations of city and county land-use planning	U.S. Fish and Wildlife Service
State Natural Heritage Programs	National Biological Service
Private businesses	U.S. Geological Survey
Nature Conservancy state offices	U.S. Soil Conservation Service
Museums	Native American Tribes

Additional staff are frequently added as the project increases in intensity. The NBS Division of Cooperative Research's mission includes a commitment to graduate education, and every effort is made to provide opportunities for graduate student involvement in Gap Analysis projects. Although not required, so far in all projects cooperating organizations contribute funds, in-kind support, and materials.

For additional information about the collaboration among Gap Analysis and other land cover inventory programs in the Multi-Resolution Land Characteristics Consortium see the Cooperative Efforts and New Tools in the Imagery Section of this handbook.

Funding

Gap Analysis state projects receive core funding from the United States Geological Survey, through through the National Gap Operations Office. Funding amounts are subject to year-to-year variation, depending upon budget decisions at each of these steps.
Core funding is delivered to the state projects through: (a) a Cooperative Research Unit Research Work Order (RWO); (b) direct transfer to an entity within the NBS, such as a research center; and (c) through an interagency Cooperative Agreement. See Literature Section of this handbook for examples of each of these. A RWO is used if the project is funded through a state Cooperative Fish and Wildlife Research Unit, a Cooperative Agreement is used to transfer funds to a non-NBS entity. In both cases a detailed work plan is required. This must include a statement of work, deliverable products, project management, staffing, schedule, and budget. See the Literature Section of this handbook for examples of work plans.

The funding needs for each state vary. However, based on our experience thus far, some funding considerations follow. Because of the research nature of these projects, an adaptive management approach is suggested. The environment and the condition of resource information in each state differ substantially. There may be unforeseen changes in funding needs as projects are implemented, therefore Gap Analysis cooperators and principal investigators must be creative and flexible in acquiring and assembling a funding package for their project.

which can change year-to-year. Project leaders must depend upon the networked support of cooperators in order to manage the instabilities in both funding amounts and the timing of fund delivery.

The federal fiscal year begins October 1 and ends September 30. However, it is not possible to predict when in fiscal year annual funds will become available. In fiscal year 1993 funds were made available in March, in fiscal year 1994 some limited funds were made available in April, other funds were made available in June. The time of delivery of core funds to Gap Analysis projects is not predictable, and state project cooperators and principal investigators should develop strategies such as structuring their cash flow to run from the middle of the federal fiscal year, and creating periods of source funding overlap. Developing an understanding of these dynamics among the cooperators is important and a commitment among the cooperators for short-term "back-up" funding, if needed, is desirable. Patience is advised.

This GAP Handbook chapter is currently out of date and will be revised during FY 2000. It is provided here for general reference but see the Standards Chapter for current GAP data standards. You should also discuss standards and methodology with the national office prior to using any information in this chapter in conducting a GAP project.

Satellite Imagery, Pattern Delineation and Cooperative Mapping of Land Cover

Version: 2.0.0
Date: 16 February 2000
Peer-reviewed: No

Michael D. Jennings
Coordinator, Gap Analysis Project
National Biological Service
Idaho Cooperative Fish and Wildlife Research Unit
Moscow, ID

In their review of the properties of various types of remotely-sensed data for their utility to Gap Analysis, Scott and others (1993) reported that Landsat Thematic Mapper (TM) offers the best overall source of digital data for Gap Analysis purposes because it has: (a) higher signal-to-noise ratio; (b) higher precision of radiometric data; (c) higher cartographic accuracy; and (d) higher spectral dimensionality. Since then, TM has become a standard for construction of digital base maps upon which the different layers of spatial data are mapped or projected (e.g., land cover, land management categories, predicted distributions of vertebrate species). The use of TM for construction of standard base maps has caused some to believe that TM is the only source of information used for the Gap Analysis land cover data layer. Thematic Mapper imagery is, however, only one of many sources of actual land cover information used to render the land cover maps. Other information sources include direct field observations, air photos and air videos, agency records, current maps, and museum and herbaria records.

Since privatization of the Landsat data in 1984 (Land Remote Sensing Commercialization Act, PL 98-365, also see Lillesand 1983, Arnoff 1985), their high cost and copyrighted protection has severely limited access to and therefore use of these data. This limitation was overcome for Gap Analysis in September 1993, when an agreement was reached between a consortium of six federal programs that were mapping land cover and the Earth Observation Satellite Company (EOSAT) which owns the images. The agreement resulted in the first assembly of a complete TM data set for the United States (U.S.). The six federal programs that acquired these data formed the Multi-Resolution Land

Characteristics Consortium (MRLC). The consortium partners are continuing to work together to achieve a comprehensive land cover map of the U.S., which is discussed in more detail below.

The TM data are archived, preprocessed, spectrally clustered, and distributed by the Earth Resources Observation System Data Center (EDC, an MRLC member). All imagery will be provided to Gap Analysis state projects by EDC with authorization from the Gap Analysis coordinator. Each image will be preprocessed to correct systems errors (e.g., vacant pixels or scan lines, or banding) as well as geographically registered and corrected for terrain distortions. Additionally, each image will be spectrally clustered at EDC using the algorithm developed by Kelly and White (1993, included in the Literature Section of this handbook). The clustered data set will be included with the seven-band image provided to the Gap Analysis state project. This approach will greatly enhance the start up of new projects as well as the compatibility of derived data among state projects and the MRLC partner programs.

The preprocessed seven-band data sets themselves are copyrighted by EOSAT and may only be used by MRLC programs and their cooperators for the purposes of each program. When distributing these data to cooperating organizations, all Gap Analysis state project leaders must obtain a signed acknowledgment from the cooperator stating that they understand the copyrighted nature of the material, that they will use the data only for Gap Analysis purposes, and that they will not reproduce or redistribute the data without the expressed permission of the Gap Analysis project leader.

The clustered data sets, however, are public domain material and may be freely distributed, though project leaders are encouraged to recover the cost of duplication and handling. These clustered data are intended for use with the display software SPECTRUM, described below in the section on cooperative efforts and new tools.

Procedures of land cover pattern delineation

The mapping of land cover is done by delineating areas of relative homogeneity, then labeling these areas using categories defined by the classification system and adding the attributes of the individual areas. Presently there are two different approaches being applied to TM imagery for initial pattern delineation. One of the procedures renders land cover patterns (dominant vegetation types, water, or non-vegetated areas) by visual interpretation of false color digital TM scenes. This procedure uses "on screen" delineation of different dominant vegetation types based on criteria such as color, texture, context and cross-reference to other information sources such as air photos and other maps as interpreted by an analyst (Davis and others 1991, Scott and others 1993). The other procedure delineates vegetation patterns by clustering spectral data using unsupervised or supervised methods (Colwell 1983). In many cases, those rendering maps of dominant vegetation cover use both methods in an iterative way. For more details on these methods see Data Layer Section "Actual Vegetation Layer" by Stoms.

Although concern has been expressed by some that a single mapping procedure is somehow a requisite for adequacy, the use of these two procedures, their hybrids, and the further development of other methods continues at present because of the following: (a) There is a single physical reality yet more than one method for depicting biologically meaningful and comparable abstractions. (b) Vegetation characteristics differ substantially among biogeographic regions, requiring different approaches, especially for interpretation of remotely sensed data. (c) The expertise for vegetation typing and mapping is itself also regional in nature, resulting in different approaches by the state project scientists. (d) Various sources of information are used to render the maps, introducing variability into the product anyway. (e) The current mapping work is a first generation, there is a need to try new methods because an effort of this magnitude, extent, and degree of resolution has not been undertaken before. There is no single proven method. (f) Of necessity, Gap Analysis is a collaborative, "bottom-up" effort focused on pragmatic near-term conservation gains (Gap Analysis actually consists of over 200 collaborating organizations). At present, there is not the institutional support to implement a single "top-down" method. Nor, given present degree and rate of habitat losses (LaRoe and others 1994) and related conflicts over resource uses, is there the time to research and develop a single method, achieve consensus on such a method, then implement a large "top-down" program. Although the level of effort that would be required is beyond current organizational capabilities and financial means, we remain optimistic about a natural convergence of methods, based on the current experience and its outcomes. One strong prospect is the common use of the spectrally clustered data set from EDC with the SPECTRUM display and image manipulation software discussed below.

Rather than limit information to a continent-level least common denominator, there are five ways to deal with differences that could arise in the derived data among state projects which use the two different procedures. First, a standardized method for assessing the accuracy of all land cover maps (described below), regardless of the different sources of information or procedures used to produce them, has been developed and is being applied. Second, a full accounting of the variability among state-level products is provided to users through standardized metadata records (also further described below). Third, a detailed review of data quality is undertaken when edge-matching data from adjacent states; this ensures spatial and thematic consistency. Fourth, since the present effort is a first generation, these data-sets will be improved over time, dampening the amplitude of inter-state variation (see the section on new products and cooperative efforts). Fifth, results from an application of the two procedures to the same area have been compared in an overlap zone along the borders of Montana and Wyoming (Ma and Redmond, unpublished) as well as along the border of Arizona and California. These results strongly suggest a close fit between the two methods. Additional comparisons will be made as the data become available.

Cooperative efforts and new tools for land cover mapping

The Gap Analysis project has joined with five other federally sponsored programs to form the Multi-Resolution Land Characteristics Consortium (MRLC) (Table 1). The overall goal of this consortium is the development of a database that integrates environmental information from multiple sources, distinguishing earth surface characteristics at multiple levels of representation. A key objective of the MRLC is the establishment of an interpreted comprehensive land cover data set based on the TM imagery acquired cooperatively, by integrating the ongoing activities of the partners. With funding and other support from the EMAP-Landscape Characterization group and the involvement of NAWQA and other cooperators, Gap Analysis projects will include the labeling of polygons for major agricultural cropping systems (e.g., "irrigated potatoes/alfalfa/onions") and major urban area types (e.g., "high-density residential").

Traditionally, environmental information has not been gathered in ways intended to maximize its utility for other purposes. One consequence is that data covering large areas is not easily or frequently used among different federal agencies, resulting in limited vision, or duplication of effort and cost. A key advantage to the cooperative approach taken by Gap Analysis is in maximizing the utility of what each of the individual programs is producing by ensuring a geographical and categorical fit. By sharing the intellectual, technical, and financial elements of this effort, these programs can meet a broad demand for computerized land cover information, which most states do not now have. The MRLC effort will not only provide this, but also consistent data on these same conditions across contiguous states, thus the context for what occurs within the state.

A significant new tool for large-area mapping by the MRLC is development of better methods for spectrally clustering and displaying the TM data. These methods were developed by Kelly and White (1993; see Literature Section in this handbook) and applied to MRLC data recently by Benjamin and Gatos (1994 personal communication). The method relies on using a fast-clustering algorithm, producing a relatively large number of spectral clusters (significantly larger than the anticipated number of land cover categories). The MRLC is now clustering all of the acquired TM scenes with this algorithm, and these clustered data sets will be available for the cost of reproduction.

Table 1. The Multi-Resolution Landscape Characteristics Consortium member programs, their areas of interest, and their parent agencies.

Multi-Resolution Landscape Characteristics (MRLC) Programs, Areas of Interest, and Parent Agencies		
Program	Area of Interest	Parent Agency
Environmental Monitoring and Assessment Program (EMAP)	Ecosystem monitoring	Environmental Protection Agency (EPA)
Gap Analysis (GAP)	Biological diversity	National Biological Service
Earth Resources Observation System Data Center (EROS)	Earth resources information management	U.S. Geological Survey
National Water Quality Assessment (NAWQA)	Water quality	U.S. Geological Survey
Coastal Change Analysis Program (C-CAP	Coastal change	National Oceanographic and Atmospheric Administration
North American Landscape Characterization (NALC)	Land cover change detection	Environmental Protection Agency and U.S. Geological Survey

Using the clustered data, efficient display and manipulation of spectral clusters is possible using a software package called SPECTRUM. This package is not copyrighted and can be used with any computer workstation running UNIX and X-windows. The clusters can be analyzed interactively, making best use of both methods of pattern delineation discussed above. To obtain this software over the Internet through an Anonymous File Transfer Protocol see the Data Layer Section, "Additional Data Layers" in this handbook by Cogan, which summarizes nationwide data sets and tools.

This combined effort is presently (June 1994) still in early stages of development, however, we (Gap Analysis and its MRLC partners) anticipate continued consolidation in the production of a comprehensive land cover database. This effort will result in more efficient use of scarce funds, better products of broader utility, and greater stability in the funding of Gap Analysis projects. An MRLC strategic plan is presently under development.

Literature Cited

Arnoff, S. 1985. Political implications of full cost recovery for land remote sensing systems. Photogrammetric Engineering and Remote Sensing, 51(1):41-45.

Colwell, R. N., ed. 1983. The manual of remote sensing. American Society of Photogrammetry. Falls Church, Va. 2143p.

Davis, F. W., J. E. Estes, B. C. Csuti, J. M. Scott, and D. Stoms. 1991. Geographic information systems analysis of biodiversity of California: Final report - Year 1. Department of Geography, University of California, Santa Barbara.

Kelly P. M., and J. M. White. 1993. Preprocessing remotely-sensed data for efficient analysis and classification. In: Applications of artificial intelligence: Knowledge-based systems in aerospace and industry. SPEI, 1993, Vol. 1963: 24-30.

LaRoe, E. T., G. S. Farris, C. E. Puckett, and P. D. Doran, eds. 1994. Our Living Resources 1994: A Report to the Nation on the Distribution, Abundance, and Health of U.S. Plants, Animals and Ecosystems. National Biological Service, Washington, D.C.

Lillesand, T. M. 1983. Issues surrounding the commercialization of civil land remote sensing from space. Photogrammetric Engineering and Remote Sensing, 49(4):495-504.

Ma, X. and R. Redmond. Unpublished. A comparison of pattern delineation techniques for Gap Analysis. Montana Cooperative Fish and Wildlife Research Unit, University of Montana, Missoula.

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, No. 123, The Journal of Wildlife Management 57(1) supplement.

This GAP Handbook chapter is provided here for general reference but see the Standards Chapter for current GAP data standards.

Actual Vegetation Layer

Version: 2.0.0
Date: 16 February 2007
Peer-reviewed: No

David M. Stoms
University of California
Santa Barbara, CA

Vegetation acts as an integrator of many of the physical and biological properties of an area. Gap Analysis maps the extent and distribution of existing vegetation in order to: a) determine species and natural community representation within areas being managed for biodiversity conservation; b) provide spatial data that can be used to model wildlife habitat distributions; and c) provide a single temporal data set for monitoring trends.

No single mapping strategy will work best in all locations. Thus, this handbook is not intended to provide a "cookbook" approach, or require rigid compliance with a certain methodology, but rather describes the required standards for products and summarizes several approaches that have been followed successfully in different states. Gap Analysis Project Manager's are encouraged to be creative and to adapt these strategies into a system that works best in their situation. For more, see Gap Analysis Standards section.

More information on Land Cover Mapping Protocols can be found here.

Current Standards for Gap Analysis Vegetation Maps

Mapping standards for Gap Analysis have been established at various meetings and published in various documents. This section of the handbook attempts to collate and organize them into a single guide. The basis of all the standards is that state vegetation maps will ultimately be aggregated for conducting regional and national gap analyses, and therefore, maps of individual states must be consistent.

As a minimum, vegetation maps will meet or exceed the following:

- nominal map scale of 1:100,000
- minimum mapping unit (MMU) of 100 ha, or 1 km2 for uplands, and 40 ha for special features such as riparian areas (Jennings, 1993)

- the hierarchical classification system will reach at least to the "alliance" level defined by dominant species and be consistent with the UNESCO/TNC system (UNESCO, 1973) down to the "formation" level (for details on classification, see Jennings, 1993 in revision)
- the final product delivered to the National Gap Analysis Data Center in Logan, Utah, will be an ARC/INFO GIS coverage in polygonal (vector) format, complete with map projection parameters and a data dictionary (see Metadata Analysis section in this handbook)
- Landsat Thematic Mapper (TM) imagery will be used as a primary source of data for compiling the vegetation layer. The base year for all TM scenes must be recent (i.e., less than 3 years old at the start of a state project)
- See also: NGAP Standards

It is recognized that there will often be compelling reasons to satisfy additional standards mandated by project collaborators to address issues in specific states (e.g., old growth forests, or local classification systems). So long as the national Gap Analysis standards above are satisfied, and provided that resources are available to meet these local demands, state Principal Investigator's are encouraged to compile vegetation layers with the widest utility to potential users. It will be the responsibility of state Principal Investigator's to collaborate with neighboring states to ensure consistency in polygon boundaries and class labeling across state borders.

TM Data

The Landsat satellites orbit the earth on a 16-day repeat cycle, that is, each location of the surface is imaged approximately twice a month. The TM sensor has a spatial resolution of 28.5 m in six reflective channels plus a thermal infrared channel with 120 m resolution. TM data are typically marketed as individual scenes (approximately 185 km on a side) by the EOSAT Corporation, the sole vendor and copyright holder of the imagery. Gap Analysis, in collaboration with the multi-resolution land characteristics consortium (MRLC) has purchased TM data for the entire country (see Imagery Section, Satellite Imagery and Data Acquisition by Jennings).

Mapping Strategies

Two fundamentally different methods are available for mapping vegetation--computer-assisted classification and visual interpretation. Both are being used successfully for Gap Analysis projects. The preferred method in any state or region will depend on the skills of the cartographic team, the extent and quality of existing maps of vegetation and ancillary data, and the spectral and ecological properties of the vegetation. Making sure a state map is compatible with its neighbors using a different strategy will require care and continuous communication.

Computer-assisted classification Utah Example

Image classification generally is accomplished by cluster analysis (often referred to as unsupervised classification) or by discriminate analysis or pattern recognition (referred to as supervised classification) (Moik, 1982; Richards, 1986). Unsupervised classification involves clustering individual pixels into spectral classes by measured reflectance values in the original channels or in transformations of those channels. The spectral classes (the number of which is selected by the analyst) are then assigned to vegetation cover classes based on information such as field observations, aerial photographs, and existing maps. Generally, many clusters are aggregated into a single vegetation class. In addition, ancillary data is used both to assign clusters to classes and to split clusters containing more than one vegetation type (Brown et al., 1993). Strahler (1981) and Franklin et al. (1986) describe an unsupervised approach to mapping forest vegetation that has been highly successful, in which digital elevation data are used to discriminate between conifer types. It is often useful to mask out urban and agricultural lands (using ancillary sources) before classification, because much of the spectral variation in TM data is typically associated with these cover types. Without masking them, too many spectral clusters will be assigned to them and too few to the natural vegetation types of interest for Gap Analysis. The Utah Gap Analysis project, among others, is using unsupervised classification for some regions of the state.

In supervised classification, the analyst specifies predetermined vegetation cover classes and provides a training data-set for determining the discriminant function based on spectral properties of those classes (Mulder, 1988). The analyst selects representative sites, called training areas, for each vegetation class, where the class is known either from direct field observation of from reliable detailed maps or aerial photos. The Arizona Gap Analysis team used an airborne video camera in conjunction with field sampling to develop a large number of training sites, inexpensively. The geographic location of training sites is digitized and their characteristic spectral "signature" is determined. This process usually takes several iterations to get the maximum separation between the classes of interest. Several different methods of supervised classification have been successful for mapping urban and agricultural features. These methods have not been as successful in mapping natural vegetation because the spectral heterogeneity of classes make specification of an adequate set of training areas difficult.

Converting 0.1 ha pixels in a raster classified image into polygons with a 100 ha minimum mapping unit, as required by national Gap Analysis standards, is not a trivial task. In Utah, the raster was first aggregated into 1 ha pixels and then "polygonized." Following polygonization, coverages were generalized to the 100 ha MMU level by a customized program written in a combination of ARC Macro Language and C to subsume the smaller polygons into the larger polygons iteratively with increasing minimum size thresholds. The elimination of the targeted polygons into larger adjacent polygons was done by weighting for all possible combinations of adjacent cover types so that aggregation made ecological sense. The cartographer has to decide whether to simply label the derived polygon with the most abundant vegetation class or to preserve the composition of all the major types within it.

Visual Interpretation of Satellite Imagery: California Example

Satellite imagery can also be interpreted using the same principles developed for aerial photo interpretation. Humans have some advantages over computers. We detect spatial patterns, while computers typically operate one pixel at a time. Obvious boundaries can be drafted around relatively homogeneous landscapes on the basis of image tone, texture, shape, pattern, site, and association (Estes et al., 1983). As with computer-assisted classification, it may be useful to segment the state into ecological regions prior to interpretation. Two approaches to creating digital coverages of these boundaries have been used in Gap Analysis, depending on the facilities available. The Oregon prototype vegetation map was drafted on mylar on photographic prints of Landsat scenes and then digitized (Kagan and Caicco, 1992). In California (Davis et al., 1994), the map units were digitized directly on-screen over a CRT display of a false color composite of red, near-IR, and mid-IR TM bands using ARC/INFO’s IMAGE INTEGRATOR module. The latter approach has the advantages that mapping can be continuous rather than by sections the size of the photographic prints, and especially that other data layers can be displayed simultaneously to guide interpretation. Also, the image contrast can be enhanced by the analyst to bring out features of interest, whereas you are limited to a single representation with photographic prints. Not all significant ecotones are visible in Landsat TM data. Ancillary information, especially aerial photos and larger-scale vegetation maps, can be used to capture additional compositional changes in vegetation that are not visually obvious in TM imagery. **Final delineation of a landscape unit is an iterative process based on evidence from the satellite imagery, aerial photos, existing vegetation maps and field reconnaissance**. A critical step in the creation of any vegetation map is accuracy assessment. Without that step you have no idea of the reliability of your map,

Dominant species cannot always be identified directly in TM imagery. For labeling community types, as required by Gap Analysis standards, ancillary information is required. The ideal situation is to have recent large-scale vegetation maps that are known to be accurate to draw on, where the interpretation has already been done. In the absence of such sources, the cartographer must rely on their own interpretation of aerial photos, combined with field visits and ecological relationships. For instance, knowledge of species-topographic relationships can sometimes be used with digital elevation data to distinguish spectrally similar species. The California Gap Analysis team found the Vegetation Type Maps of the Forest Service, based on extensive field surveys between 1928 and 1940, to be a valuable source of species information where type conversions had not occurred in the interim (Davis et al., 1994).

Using these various sources, a large amount of information can be collected for each map unit (Davis et al., 1994). Based on the concept of map unit as a landscape mosaic, the California Gap Analysis team recorded a primary species complex, which was the most widespread vegetation type or land cover type in the polygon, a secondary type, and the fraction of the landscape covered by each type. They also recorded the most widespread wetland complex, which was usually riparian vegetation. Each species complex was defined by up to three dominant species. Thus the distribution and management status of individual plant species (where they occur as canopy dominants) can also be derived from the database. Where possible, they also recorded the occurrence of minor species of special conservation concern. Information on pre-burn dominants was recorded to account for fire dynamics (e.g., an area of recently burned chamise chaparral that is presently dominated by herbs would be recorded as sparse chamise canopy codominated by annual herbs).

Sources of Existing Maps of Vegetation

Several potential sources of spatial data should be investigated during the planning phase of a vegetation mapping project. The U. S. Department of Agriculture Forest Service may have timber strata maps produced through aerial photointerpretation or more recently by image classification at the 1:24,000 scale. The emphasis is generally on commercial timber species, with less classification detail for other types. The availability of Bureau of Land Management maps varies greatly from place to place. The most useful maps from the U. S. Fish and Wildlife Service are in the National Wetland Inventory series. Soil Conservation Service soil maps may be useful in areas where vegetation data are lacking. U. S. Geological Survey's Land-Use and Land-Cover maps may be useful for urban and agricultural areas, although they are somewhat dated now. Maps from local and state agencies for specific sites (e.g., a state park or a project area) are also helpful, but they can vary widely in quality. These "maplets" can be used either for compiling the Gap Analysis vegetation map or for its validation. Often the maps described here are at a finer spatial resolution than needed for Gap Analysis. Intelligent generalization to the 100 ha MMU is not always straightforward in existing GIS software (Stoms, 1992).

Map Accuracy

Validation and documentation of accuracy assessment is a complex, but critical, topic. See the Validation section of this handbook for Gap Analysis standards and methods.

Limitations of Gap Analysis Vegetation Maps

Scott et al. (1993) list the following caveats about vegetation maps compiled for Gap Analysis:

1. Vegetation maps do not show habitats smaller than the MMU. Thus, microhabitat elements, such as meadows and wetlands in a forest matrix, are not resolved at this level of effort. Such habitat inclusions must be captured in a subsequent, higher-resolution assessment of potential high-priority biodiversity areas or assigned as polygon attributes without explicit spatial locations.
2. Vegetation maps do not necessarily portray stand age, except for the early successional stages (herb and shrub stages) of forests following clearcutting or stand-replacing fires. Gap Analysis can identify large areas of relatively unfragmented natural forest, but is not designed to indicate how much of that forest is old-growth. However, in some states the cooperators are making special efforts to map stand age characteristics such as old growth.
3. Boundaries between vegetation types along real environmental gradients are seldom as sharp as implied by vegetation maps. Ecotones and subtle gradients must be identified by higher-resolution, landscape scale analysis.

Collaborators and potential users of Gap Analysis databases need to be aware of these limitations.

References

Brown, J. F., T. R. Loveland, J. W. Merchant, B. C. Reed, and D. O. Ohlen. 1993. Using multisource data in global land-cover characterization: concepts, requirements, and methods. Photogrammetric Engineering and Remote Sensing 59: 977-987.

Davis, F. W., P. A. Stine, A. D. Hollander, D. M. Stoms, and M. I. Borchert. 1994. Distribution and management status of terrestrial plant communities in southwestern California, submitted to Madrono.

Estes, J. E., E. J. Hajic, and L. R. Tinney. 1983. Fundamentals of image analysis: Analysis of visible and thermal infrared data, In Manual of Remote Sensing, R. N. Colwell, editor. American Society of Photogrammetry 1:987-1124.

Franklin, J., T. L. Logan, C. E. Woodcock, and A. H. Strahler,. 1986. Coniferous forest classification and inventory using Landsat and digital terrain data. IEEE Transactions on Geoscience and Remote Sensing, GE-24: 139-146.

Jennings, M. D. 1993. Natural Terrestrial Cover Classification: Assumptions and Definitions. Gap Analysis Technical Bulletin 2, U. S. Fish and Wildlife Service, Moscow, ID.

Kagan, J., and S. Caicco. 1992. Manual of Oregon Actual Vegetation. Idaho Cooperative Fish and Wildlife Research Unit, Moscow, ID, 190 p.

Moik, J. L. 1982. Digital Processing of Remotely Sensed Images. National Aeronautics and Space Administration, Houston, NASA SP431, 330 pp.

Mulder, N. J. 1988. Digital image processing, computer-aided classification and mapping, In Vegetation Mapping, A. W. Küchler and I. S. Zonneveld, editors. Kluwer Academic Publishers, Dordrecht, pp. 269-316.

Richards, J. A. 1986. Remote Sensing Digital Image Analysis: An Introduction. Springer-Verlag, New York, 281 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.

Snyder, J. P. 1987. Map Projections--A Working Manual. U. S. Geological Survey, Professional Paper 1395, Washington, 383 pp.

Stoms, D. M. 1992. Effects of habitat map generalization in biodiversity assessment, Photogrammetric Engineering and Remote Sensing 58: 1587-1591.

Strahler, A. H. 1981. Stratification of natural vegetation for forest and rangeland inventory using Landsat imagery and collateral data. International Journal of Remote Sensing, 2: 15-41.

UNESCO. 1973. International Classification and Mapping of Vegetation. United Nations Educational, Scientific and Cultural Organization, Paris, 35 pp.

Assessing Land Cover Map Accuracy

Version: 2.0.1
Date: 16 February 2007
Peer-reviewed: Yes

Patrick Crist[1] and Robert Deitner[2]
[1]Idaho Cooperative Fish and Wildlife Research Unit
University of Idaho, Moscow, ID
[2]New Mexico Cooperative Fish and Wildlife Research Unit
New Mexico State University, Las Cruces, NM

as adapted from Handbook version 1 by

Stoms, D., et al.
University of California
Santa Barbara, CA

- Appendix 1 - Summary of related land cover mapping programs
- Appendix 2 - List of participants in the Land Cover Assessment workshop, February 3-4, 1994, Santa Barbara, CA
- Appendix 3 - List of participants in the Land Cover Assessment workshop, April 9-10, 1996, Denver, CO
- Appendix 4 - Spatial Accuracy
- Appendix 5 - Field Survey Example
- Appendix 6 - Analysis of Land Cover Map Accuracy in Northern Arizona (pdf version)

Introduction

Version 1 of this chapter by Stoms et al. (1994) was based on a workshop conducted in 1994 and made recommendations to guide Gap Analysis Program (GAP) projects in their accuracy assessment of land cover maps. At that time, no project had completed an assessment, and little practical guidance was offered in the literature. By spring 1996, four GAP projects had completed or initiated assessment, and projects nearing completion were seeking additional recommendations for assessment methods. On April 9-10, 1996, another workshop was held in Denver, CO. The number of participants (see Appendix 3 for list) and the goals of the workshop were scaled back from the 1994 meeting. The primary goal of the meeting was to share experience and advice from states involved in assessment projects with states planning assessments. The participants of the 1996 workshop found the general approach and recommendations of the 1994 chapter to be sound but requiring revision based on the actual experience of GAP projects. For that reason, much of the version 1 text was retained and its authors deserve significant credit for this chapter. In the case of errors, responsibility rests with the current authors. We would also like to acknowledge the contribution of our reviewers, particularly Steve Stehman who made numerous suggestions that have been incorporated.

The need for accuracy assessment of GAP data products was recognized by an external peer review of the National Gap Analysis Program conducted in 1994. The panel specifically recommended that "a protocol should be developed that will provide accuracy estimates of the spatial and classification data to other users, so as to enable them to more fully understand its utility for their application" (Peer Review Panel, 1994, page 4). While the panel was sympathetic to the challenges inherent in large-scale validation, such as cost and lack of widely accepted methods, they felt that a national strategy for validation was essential for Gap Analysis to attain credibility. The 1994 Santa Barbara workshop, which addressed that challenge, concluded that there are several reasons that it is not practical to present a national protocol at this time:

- The choice of procedure depends on funding level.
- Opportunities for collaboration with other organizations and for application of existing data sets vary from state to state.
- No single procedure is optimal across all regions and land cover types.
- Any procedure or set of procedures should be field-tested in a set of representative states before being imposed as a national standard for the Gap Analysis Program.

For the most part, the above reasons still hold true, with some progress having been made since then. GAP is committed to funding valid assessments. The advent of aerial videography and use of volunteer cooperator support has brought valid assessments within funding limitations. Several different procedures have been field-tested to the point that confident recommendations can be made for assessment studies. The following discussion and methodology is pertinent to the assessment of the final, deliverable land cover map; assessment or ground-truthing conducted for the mapping process is addressed in the land cover mapping chapter.

How This Chapter Is Organized

This chapter contains some rather lengthy discussion on theory and principles being applied to accuracy assessments. The discussion is necessary to provide sufficient background to the reader because so few published examples of large assessment projects exist. Recent publications have added to the body of work and are recommended reading (Fenstermaker 1994, Mowrer et al. 1996, and Stehman and Czaplewski 1997). The outline of this report is designed to lead the reader from the uses of the Gap Analysis land cover maps and accuracy assessment to a brief literature review of measures of accuracy and constraints on assessment methods. Later sections describe the sampling and measurement strategies developed by the workshop members. We conclude with reporting requirements from the assessment and where it is reported, a discussion of the proposed requirements for data management, and a list of unresolved issues. Assessing accuracy of medium-scale thematic maps is still a research area. We briefly identify several questions needing further research and encourage readers to send comments to Patrick Crist (gap@uidaho.edu, 208/885-3555). We include a set of basic references that should provide readers with access to some of the fundamentals of accuracy assessment. Appendix 1 contains short descriptions of how other agencies' land cover mapping programs relate to Gap Analysis and its accuracy assessment program.

We thank the participants in the workshops for their helpful discussion of the issues and other cooperators for their input and review of this document.

Objectives

1. To obtain a set of reference data suitable for map creation, editing, and assessment.
2. To measure thematic accuracy as a percentage of the land cover map classified correctly overall and by mapped type with a standard error no greater than 8%.
3. To report the results so that users have an understandable measure of map accuracy and enough information to allow assessment of map suitability for individual uses.

Background

First a note on terminology: We tried to keep jargon to a minimum and expect that most terms used in this discussion will be understood by those conducting map creation and assessment. Because GAP projects are creating maps that use either polygons (vector-based) or pixels (raster-based) or both, we use the familiar term "polygons" to describe either a polygon or a group of contiguous pixels of the same primary land cover type attribute. The definition of many other terms can also be found in the Glossary in this handbook.

Purposes of Gap Analysis Land Cover Maps and Accuracy Assessment

The primary use of the GAP land cover maps is to answer the fundamental question in gap analysis: what is the current distribution and management status of the nation's major natural land cover types and wildlife habitats? The assessment gives a measure of overall reliability of the land cover map for gap analysis and also identifies which general classes or which regions of the map do not meet the accuracy objectives for the Gap Analysis Program. Thus the assessment will identify where additional effort will be required when the map is updated. Results of the accuracy assessment will be reported, with the analyses stating that the map is the best available information for the project area.

GAP land cover maps also have numerous secondary uses such as local and regional conservation planning and will likely be used in coordinated national interagency land cover assessments. The maps should be useful for biogeographic analyses and as input to mesoscale models of climate and ecosystem processes. They may well be used to assess the accuracy of smaller-scale land cover maps. These secondary uses also require the per-class accuracy of the maps. Some users would also benefit from having some measure of accuracy by polygon or geographic area, such as where the map is most reliable. We expect that in general, secondary users of the land cover map will need more details about the accuracy assessment to make appropriate use of the map than will primary users (the GAP project researchers) who are more familiar with the data. It is also important to be able to assess the implications of the reported level of accuracy on the kinds of products needed by database users, including estimates of area.

It is impossible for the original cartographer to anticipate all future applications of a land cover map, so the assessment should provide enough information for the users to evaluate fitness for their unique purposes. This can be described as the degree to which the data quality characteristics collectively suit an intended application. The information reported should include details on the coverage's spatial, thematic, and temporal characteristics and accuracy. In fact, these are not independent aspects in area-class maps, because the object boundaries are derived from the same source information as the thematic attributes (Goodchild 1989). That is, if the boundary was drawn differently, the attributes would also be different. Types of accuracy statements include:

- Spatial accuracy: The landscape displays continuous variation in species composition at all scales. Segmenting the earth's surface into map units with relatively homogeneous composition of canopy dominants, at a particular observational scale, is a useful but treacherous activity. The boundaries imply a discontinuity that seldom is as abrupt as a map would suggest. Unlike a road or a survey line, land cover boundaries do not have a "true" location of which the accuracy could be tested against a higher-quality data source. They are simply a subjective interpretation of where between-polygon variation is greater than within-polygon variation. Given the landscape-level view of land cover mapped by Gap Analysis, the contribution of locational errors of boundaries, even if it could be measured, is assumed to be small relative to thematic error. Of greater concern is potential misregistration of the Landsat Thematic Mapper (TM) imagery. This source of error will cause distinctive land use and land cover features such as lakes and agricultural lands to be offset from their location relative to other map sources.

- Thematic accuracy: Thematic (or identification) error occurs when the attribute of a map unit is misclassified with respect to a reference data source assumed true and is generally the type of error users most often mean when they ask about accuracy. There are many potential sources of thematic error. The two most obvious sources include simple misidentification of the class by the interpreter and data entry errors. A more generalized vegetation/land cover map masks fine-scale heterogeneity in the landscape. Thus, a given point within a polygon may actually be incorrectly classified even though the polygon as a whole is correct. Apparent misclassification may also be the result of differences in definitions of classes in the map and reference data, such as a different canopy cover threshold between grassland and woodland. Existing methods for assessing accuracy usually ignore the magnitude of a particular thematic error. For instance, misclassifying a polygon as Ponderosa pine instead of Jeffrey pine is much less dramatic than misclassifying it as agriculture. Clearly, these sources of error are closely related, making it difficult to categorize them cleanly in an error budget. Newer techniques of analysis, such as "fuzzy set," attempt to deal with degrees of error which may provide more useful information to a data user and will be covered in more detail later.
- Topological accuracy: Also called logical consistency, this refers to the fidelity of relationships encoded in the data structure of digital spatial data. Included within this category are tests for valid attribute values, duplicate lines, overshoots and undershoots of lines, and the presence of sliver polygons. This class of errors is associated more with the process of data entry than with interpretation.
- Temporal accuracy: Maps are compiled to portray a relatively static view of land cover at a given time. One of the assumptions of GAP land cover mapping is that map units represent a snapshot in time, but are generally stable over a 10-year update cycle. The finer scale at which sampling must occur will detect more dynamic changes. It is not reasonable to report as errors those areas where land cover changes have occurred after the source data (in this case TM imagery) was gathered. In practice, it may be inevitable that field sampling may not occur until long after the map source materials were acquired. It may not always be obvious whether the discrepancy between the map and the field observation is due to temporal change or simply to misclassification. Samples in which change has taken place may provide some insight into the rate of change, and may indicate an appropriate interval for map updates, but are not helpful in accuracy assessment. Provided that the reference data is collected within a few years of the source data (TM imagery), it is likely that any changes will be obviously caused by human or natural disturbance and those samples can be disregarded.

The remainder of this chapter will deal only with thematic accuracy; spatial accuracy of the source data (TM imagery) is handled by the (Multi-Resolution Land Characteristics Consortium (MRLC) processing at Earth Resources Observation Systems (EROS) Data Center (EDC) and should be assessed during the map development stage. Methods and issues of spatial accuracy from Stoms et al. (1994) are provided in Appendix 4.

Scientific and Programmatic Criteria for the Assessment

The criteria for the accuracy assessment reflect the need to balance the requirements for rigor and defensibility with practical limitations of cost and time. In general, the assessment methods must be:

- Scientifically sound: The process used must be understandable, repeatable (theoretically), and be of sufficient rigor to provide useful results.
- Economically feasible: The cost of the assessment program must be reasonable in light of the cost of the original mapping program, and data collection and archiving should support remapping activities.
- Nationally applicable: States should strive to use similar assessment methods and reporting, regardless of land cover mapping procedure used, to support regional, multistate analyses.

Uses of Gap Analysis Accuracy Assessment Data

Assessment data are valuable for purposes beyond their immediate application to estimating accuracy of a land cover map. GAP requires that the reference data be part of the deliverable product, thereby making it available to other agencies and organizations for use in their own land cover characterization and map accuracy assessments (see Appendix 1 for a sample of these programs). The data set will also serve as an important training data source for later remapping. Archiving of the data is described under Step 4 below.

Even though there will be endpoints in the mapping process, where products will be made available to others, the gap analysis process should be considered dynamic. We envision that maps will be refined and updated on a regular schedule. The assessment data will be used to refine GAP maps iteratively by identifying where the land cover map is inaccurate and where more effort is required to bring the maps up to accuracy standards. In addition, the field sampling may identify new classes that were not identified at all during the initial mapping process.

Methods

An accuracy assessment is conducted by selecting a sample of locations and determining the true land cover classification at these locations. These data are frequently called the reference data set (Stehman and Czaplewski 1997) . This section focuses on the entire process with emphasis on the proper statistical design, analysis, and reporting of the results. Descriptions of field (or video interpretation) methods are necessarily general, owing to the large number of variables among states. Application of aerial videography will be further described in a forthcoming manual in 1998 that will be appended to this chapter. Properly executing an accuracy assessment involves knowing the nature of the created map, identifying the field (or photointerpretation) methods for obtaining the reference data, designing a sound method for selecting reference data, actually collecting the data, conducting analyses, and reporting the results. Each of these steps or phases will be explained in separate sections.

Summary of steps and standards

1. Establish a final land cover map, classification, and description of mapped types that will be assessed.
2. Identify the method(s) for obtaining reference data.
3. Design a sampling protocol that meets the statistical precision indicated in Objective 2 above.
4. Collect the reference data, test their reliability, and archive in a database.
5. Compare the reference data to the map, conduct analyses, and report the results.

Step 1: Establish a final land cover map, classify, and describe mapped types that will be assessed. Most current GAP projects use machine classification (see land cover mapping chapter), and numerous versions of the land cover map can exist with different levels of spatial and thematic detail. It is critical to clearly identify one final version of the land cover map as the official version to be assessed, reported, and distributed. This guarantees that users have an assessment report that is directly applicable to the land cover map that they receive. Knowledge of the characteristics of the map to be assessed is important during the design phase of the study to determine the sample frame (number, size, and classification of polygons). In addition, it is important to have a finalized classification scheme with identified mapped types to insure that the methodology used to collect the reference data matches the classification system of the map.

Step 2: Identify the method(s) for obtaining reference data. There are a variety of sources or methods of obtaining reference data available that differ in cost, access, timing, analysis, and utility for future remapping. To date, GAP projects have typically used ground-truthing, aerial videography, or a combination of both as the source of reference data. Aerial photo interpretation was recommended in the previous version of this chapter and is also suitable but is too costly for this single purpose without substantial cooperator support. Some states are obtaining full coverage of digital orthophoto quads which may also be useful. It is difficult to recommend one technique over the other due to differing circumstances of each project. However, nearly all projects are currently obtaining aerial videography. If videography is the source of the data, there will still be issues of how to select reference samples and apply them to the assessment of the map.

The decision of which technique to use should consider the implications listed above and also be made in consultation with the plant ecologist responsible for developing the land cover classification scheme. This is critical in order to guarantee consistency between the reference data and the needs of the assessment project and future remapping. In other words, the method of collecting the reference data must lend itself to identifying the land cover at the same level of detail as the map and preferably one level more resolved for higher-resolution remapping, e.g., if the alliance was mapped, then the community association should be identified in the reference data if possible. While guaranteeing consistency with the land cover map's classification scheme is obvious, consistency with its resolution is more subtle. Land cover maps produced by GAP have an inherent minimum mapping unit (MMU), and methods used to obtain the reference data must be able to distinguish between the polygon's dominant land cover type and unmapped inclusions of other types within it. One approach is to make the observed unit larger than anticipated inclusions, but this may not always be feasible. Therefore the project plant ecologist should identify a method to detect inclusions *as* inclusions rather than as the primary cover attribute. As GAP land cover maps increase in both spatial resolution (typically 30 m^2 to 2 ha) and thematic resolution (National Vegetation Classification Scheme [NVCS] alliance), this becomes less of a concern. We do, however, strongly recommend that the consulting plant ecologist develop the sampling method suited to each mapped type to assure it can be correctly identified. While this may sound like a complex task to develop and implement, in reality we expect that only a few different protocols need to be created that deal primarily with plot size to be examined. The characteristics of the land cover types that may affect these protocols would be: polygon sizes (small, medium, or large), polygon shapes (linear or nonlinear), and the ecology of the types, such as heterogeneity (degree of patchiness and size of inclusion patches).

Previous recommendations have called for using a sample unit equivalent to the MMU; unfortunately, the MMU is unrelated to polygon shape which can cause difficulty in applying a standard plot over any location on the map. In other words, a square plot equivalent to the MMU will not fit within an irregularly shaped polygon of the same size. The next section argues for the use of a mapped polygon for the sample unit. Any measurement must be based on the entire *mapped* polygon or some subset of that polygon with the understanding that this subset is used as a surrogate for the entire mapped polygon. Stehman and Czaplowski (1997) discuss the relative merits of different sampling units.

Ground-truthing involves physically visiting the site in question to determine its true land cover type. The actual measurement can range from simple reconnaissance to more elaborate vegetative sampling schemes. New Mexico (Thompson et al. 1996) conducted a field inspection of the mapped polygon to identify its true type, while other states (UT, AR, MT) have used a regular grid placed in the interior of polygons. The amount of information gathered per site is directly related to the number and accessibility of sites, time and number of observers/ interpreters available, skill of the observers to consistently record all information, and cost per unit of information (see Tables 1 & 2). Having the field observer identify the dominant type in the mapped polygon most directly identifies that polygon's true class, but issues involving the polygon boundaries become important. A regular grid used for measurement can provide a more objective measure of land cover type, but placing the plot within any polygon can become problematic. Some projects have purposely avoided placing grid sample units near edges; this subject is dealt with further in Step 3. In addition to classifying the sample unit to the one "best" type, it is desirable to indicate a degree of confidence in that decision or list other secondary types that occur within or near the parcel of land in question (Gopal and Woodcock 1994).

Table 1. Comparison of approximate cost per point, per method of assessment. Figures provided by AR-GAP, NM-GAP, and MA-GAP.

Method used	**Project**	**Approximate cost per point**
Ground observation/uniform grids	AR-GAP	$14.00 - $17.00
Ground observation/polygon	NM-GAP	$15.00
Aerial video interpretation	MA-GAP	$1.81 - $6.40 (see Table 2)

Table 2. Cost per point of videography. From MA-GAP.

# of points	**Total cost per point**	**Assessment portion of cost**
1200	$6.40	$3.70
2000	$4.25	$2.62
3000	$3.16	$2.08
4000	$2.62	$1.81

Nearly all GAP projects begun after 1993 have acquired aerial videography; therefore, GAP recommends its use for land cover assessment coupled with field-truthing or other methods needed to distinguish all mapped types. As long as registration and projection problems between the videography and land cover map are solved, this method can provide a more consistent measurement, since interpretation is done in the laboratory with one or a few well trained interpreters rather than in the field by many, frequently volunteer, observers. At one level, videography has the potential for obtaining relatively large amounts of data but is limited by several factors, including: resolution or clarity of the images, spatial accuracy impacted by tilt of the plane, selection of specific locations restricted to flight paths, and the simple fact that it is a form of remote sensing used to assess a map based on remotely sensed data. A site-wide accuracy statement hinges on the representativeness of the flight paths used. We emphasize that some ground-truthing is required to verify the video-interpretation. Typically, representative video frames are field-interpreted and then used as reference for all similar-looking frames. There will likely also be mapped types that cannot be readily distinguished by video alone, therefore other methods must be used to collect reference data for those types. Much of the discussion about ground-truthing will be applicable to the verification of video-interpretation.

Whatever method is used, the resulting measurement needs to be done blindly by the observer. In other words, the collection of reference data should be done without the observer's knowledge of the mapped type. This usually involves publishing a map for the observer with the polygons drawn without their associated land cover type label. The printed map typically has base information such as roads, streams, and locational grids such as Township/Range/Section lines. It is also important to provide ownership information for ground surveys, so that permission can be obtained to enter private lands. It is GAP and BRD policy not to enter private lands without owner permission, and GAP projects have found it beneficial to use observers from agencies such as state fish and game and forestry that regularly visit such lands and have a rapport with the owners.

Lastly, an individual measurement must result in a decision as to whether or not the reference datum agrees with the land cover map's classification of that polygon. Accuracy is the statistical reduction of many samples into a statement of percent agreement. Traditionally, accuracy assessments used a simple agreement between one observed dominant type and mapped type. More general types of agreement and their statistical reduction are termed fuzzy set analyses and are described under the analysis section (Step

5).

Step 3: Design a sampling protocol that meets the statistical precision indicated in Objective 2. Objective 2 sets a standard of precision for overall accuracy and per land cover class accuracy. It is assumed that if Objective 2 is met for all land cover class accuracies, the precision required for overall accuracy will be met. Consequently, this design phase of the assessment study only considers accuracy by land cover type. There are three tasks within this step. First, the sampling unit and the list of all possible sampling units--the sample frame--are defined. Secondly, the design, including stratification, is decided upon. Lastly, the size and composition of the sample is calculated such that it meets the criteria established in Objective 2.

GAP defines the sampling units to be all areas within the project area geographically contiguous and of homogeneous primary attribute, that is, vector polygons or contiguous raster clusters of the same primary land cover type code. GAP land cover maps are based on algorithmic clustering of TM pixels with the resultant categories being spectrally similar. Therefore, we do not believe that these pixels are independent of each other. In other words, although we do not claim that the polygon boundaries are precise, we do believe they are based on real patterns on the ground and, therefore, the polygon is the feature that has been defined and the feature that should be assessed. For these reasons, we define the sampling unit as any mapped polygon. The sample frame is then the list of all polygons that comprise the final land cover map.

In the workshops and review of this manuscript, sample unit identification has been perhaps the biggest area of debate. It is therefore worth reviewing some of the issues and the GAP positions:

- Differences in polygon size: Polygon sizes for a single mapped type may vary by three or more orders of magnitude, and there is a concern that large polygons should be treated differently than small polygons. The 1996 workshop recommended that further research be conducted on this issue in future assessment projects, and that it may be wise to distribute the samples among different polygon sizes to see if different treatments are warranted. Our assumption at this time, however, is that each polygon describes a homogeneous land cover type (the description of heterogeneity should identify likely inclusions and ecotones) and, therefore, with proper measurement, a single sample site is adequate for both small and large polygons. As presented earlier, the project plant ecologist should be consulted in developing an assessment protocol that can correctly identify the true land cover.
- Errors involving the "boundary" or shape of the polygon: As stated previously, we are not concerned with determining shape accuracy, but it does have an effect on sampling. A 100 ha polygon that is very long and narrow, e.g., riparian types, will require a different sampling approach than more consolidated forms.
- Edge effects: Many assessment projects have avoided placing samples near polygon edges for two reasons: 1) due to locational imprecision, an observation may erroneously be made in an adjacent polygon and 2) as described earlier, boundaries are not precise, and an observer is more likely to encounter ecotonal conditions near the edge, which will not accurately represent the polygon. Others have noted that edge avoidance can remove 50% or more of the map from the sample frame (J. Finn, pers. comm.) and that there are analytical methods for dealing with the ecotonal conditions (S. Stehman, pers. comm.), fuzzy sets among them. As polygons become smaller, edge avoidance becomes very difficult and removes even more of the map from the sampling pool. It is our position, therefore, that edges should not be avoided, but that sufficient buffering between edges and sample plot centroids be used to ensure that the sample is indeed within the polygon, given locational uncertainty between known spatial error of the map and the spatial error of locating the plot. For example, consider the situation where a 100 m x 100 m sample plot is to be observed. The map spatial error is known to be up to 30 m, and nondifferentially corrected GPS used to locate the plot is in error up to 100 m. Then the plot centroid could be buffered from the edge up to 180 m (50 m radius plot centroid + 100 m GPS error + 30 m map error).
- Should secondary attributes be measured? There are several types of secondary attributes that

projects have included in their map, such as disturbance (e.g., burned conifer), canopy closure, age class, and structural class. Projects have also attributed secondary dominant cover or inclusions as a separate attribute. The latter is not a preferred way of describing the mapped type--the classification scheme should include all mappable ecological variations of the plant community, usually described as an alliance; this will also include many of the attributes mentioned above (though not disturbance). The testing of secondary attributes is not likely to be practical through stratification using them as variables, but they may be testable using subpopulation estimates (see later discussion). The 1996 workshop recommended that these attributes only be tested if they have a direct bearing on the vertebrate modeling, but we advise only that the project address them according to their and their cooperators' needs and desires and the ability to do so within project limitations.

The most direct way to measure per class accuracy is to stratify by that classification. Therefore, we recommend that the design of the assessment be stratified by, and only by, land cover types present in the map. Inherent in the argument whether sample units should be derived from the land cover map or some other mosaic of the study area is a difference in the sample frames. In order to stratify by map category, one needs to use the land cover type coverage as the sample frame, hence the sampling units. Again, this is only a recommendation, and scenarios can be envisioned where it is appropriate to stratify by other attributes of the landscape. The GAP standard only requires that the resultant per land cover type accuracy be measured with a certain precision (see Objective 2 above).

Further stratification, while it may be directed at interesting questions, is usually not advisable. Additional stratification results in progressively larger costs and sample size requirements in addition to more difficult analyses. With a limited budget, there is a real danger to "overstratifying," and spreading the limited samples too thinly among the strata. It is sometimes suggested that stratification should be used to guarantee a representative sample, but this is not necessary. Probability sampling of the complete sample frame guarantees a representative sample in probability.

One potentially useful exception is dividing the sample frame by the relative cost or effort required to measure the sampling unit, as was done for UT-GAP and AR-GAP (Edwards, in press; Dzur et al. 1996). In these two cases, distance from the closest road was used as an indicator of cost of observation. The relative effort between types of polygons is adjusted to minimize cost for a given overall precision. A broad overview of this topic can be found in Cochran (1977), and an application to assessment studies can be found in Edwards (in press).

Sample size depends on the design and computation of per class accuracy. With the recommended design, this involves computing per strata accuracy as the proportion of agreements within each strata (land cover class). As with all measures of proportion, this cannot be done without guessing at the unknown true accuracy. An assumed accuracy of 50% is the most conservative strategy, i.e., the widest interval for a given sample size (Cochran 1977). NM-GAP (Thompson et al.1996) used a presumed accuracy of 80% (the GAP mapping objective) for its computations. It is not necessary to use the same presumed accuracy for each land cover type, offering the option for the mapping staff to indicate differences in confidence among types that can aid the statistician in computing sample size per type. Per strata sample size is computed by using the expression for standard error of a proportion:

$$se(\hat{p}) = s = \sqrt{(1-f)\frac{p(1-p)}{n}}$$

(Cochran 1977).

Substituting the desired precision for the standard error and the assumed accuracy (p), then solving for sample size, produces the desired results. The finite population correction factor (1-f, the proportion of the sample frame not sampled) also depends on sample size and this must be accounted for in the algebra. The resulting equation becomes:

$$n = \frac{Np(1-p)}{Ns^2 + p(1-p)}$$

where N is the total number of polygons.

As the number of map polygons (N) becomes large, the finite population correction factor is close to 1, so the previous equation becomes:

$$n = \frac{p(1-p)}{s^2}$$

Cochran (1977) discusses when this approximation is valid.

Using this last sample size formula with a presumed accuracy of 50% (p=0.5) results in the most conservative estimate of sample size, and no more than 40 samples per land cover type are needed to meet the GAP criterion of a standard error of 8%. Standard error is reduced as the total number of mapped polygons per type becomes small, and as observed accuracy differs from 50%.

Step 4: Collect the reference data, test their reliability, and archive the data collected. Once design considerations are resolved, the selection of sites is an easy task. If the frame consists of polygons in a vector map, the selection is from the "polygon attribute table." One can produce maps and statistics based on these selected polygons. The GIS methods of placing sample site centroids or sample grids within the polygon are not routine, but should be manageable for most GIS technicians.

Issues of measurement, though, are difficult, particularly with ground-truthing. The least contentious issues are that the reference set should be done blindly, consistent with the land cover typology used in the map, and by a trained observer. A blind test is when the observer does not know the primary attribute given by the land cover map. Secondly, the observer must classify the polygon according to the scheme used, to avoid the need for subjective cross-walking of observations to the map classification. This could be difficult for some observers who are accustomed to a different type of classification. Production of keys and even sample photographs to aid in identification are especially useful. Related to this is the assumption that the people conducting the ground-truthing or interpretation are capable of producing consistent, repeatable results. Both training of the observers and tests to establish observer/interpreter reliability are important components of the assessment project. Video interpreter training will be covered in the forthcoming manual. Field observers can be trained by using typical techniques for biological inventories.

Typical precision standards for scientific studies result in very ambitious assessments. For example, NM-GAP initially tried to obtain a standard error of 4 percent which roughly translates into 100 polygons per stratum. With the 42 land cover types and reductions due to finite sampling, this amounted to over 3,000 polygons to visit. Time, budget, and personnel constraints prevented the measurement of every selected polygon. In fact, only about half were sampled, resulting in an unknown bias due to a high nonresponse rate. The only way to minimize such biases is to relax the precision or use simpler (less costly) measurement techniques.

For both videography and ground-truthing, identification of the dominant alliance within the sample unit and the location of the sample unit (GPS coordinates) are the only data required by GAP. When possible, it is useful to identify the community association to achieve a more precise identification of the land cover type and to facilitate future higher resolution mapping. Other data fields can include:

- spatial precision of the coordinate
- subdominant or secondary land cover class if the sample unit is heterogeneous or hard to define
- land cover type of adjacent polygons
- structural measures such as percent canopy cover

- evidence of recent disturbances such as fire, harvesting, grazing, etc.
- photographs of the sample units visited
- actual percentages of each land cover class observed within polygon or sampling unit or "maplet" of sample unit

Often there are other field assessment projects being carried out by cooperators, and they may be an attractive source of ground-truthing. These opportunities should certainly be investigated, but there are factors that may limit or eliminate the use of the data for this purpose. First, the data may not be a probability sample and may have been selected for specific reasons. Second, the project may be using a different classification scheme that does not readily cross-walk to the GAP project scheme. Third, the data collection protocol may not be compatible, as many cooperators sample smaller plots than are appropriate for GAP purposes. Preferably, cooperator initiatives would be identified early, and attempts can be made to get them to visit GAP-selected sites within their assessment area and use GAP protocols to collect the data.

Because the reference data set is valuable, the data need to be archived as a GAP data set or actual coverage. Each datum should contain a pointer that identifies its location on the GAP land cover map. In addition to the actual data, metadata should include a description of the method used by the analyst to determine agreement between the map and reference data and a measure of observer reliability in order to replicate the published analysis. Specific archiving requirements are described below under Data Management.

Step 5: Compare the reference data to the map, conduct analysis, and report the results. Traditional analysis of map accuracy begins with tabulation of the error (or confusion) matrix (Congalton 1991). Agreements, or lack thereof, are tabulated in a matrix whose rows represent mapped categories and columns represent observed class. The diagonal elements are observations in agreement with the map, and off-diagonal elements represent disagreement between two land cover types. This matrix is not symmetric. In the recommended stratified design, the rows sum to the known sample sizes, and column sums are random numbers. With no stratification or stratification by another mosaic, both column and row sums are random variables.

The error matrix is a contingency table which represents the probabilities of every possible correct or incorrect classification. If the probabilities across any column sum to 1, then the probabilities on the diagonal are defined as producer accuracies and the off-diagonal entries are the associated errors of ommission. If the probabilities across any row sum to 1, then the probabilities are defined as user accuracy and its associated errors of commission (Congalton 1991). Producer accuracy is the probability that the map classification is correct, given that we know the true land cover classification, and user accuracy is the probability that a parcel of land is correctly classified, given that we know the mapped type. We recommend the tabulation of user accuracy and their associated errors, since this will usually be the type of error a user of the map will be interested in. An example error matrix based on New Mexico's accuracy assessment (Thompson et al. 1996) is presented in Table 3. Reporting of the matrix gives the reader a complete picture of the assessment and allows the reader to derive any measure based on traditional analyses.

Table 3. Example error matrix derived from New Mexico GAP project (Thompson et al. 1996). Rows represent mapped cover types and columns represent observed cover type. Sample size is indicated in the last column. Mapped types have been aggregated for simplicity.

	1000	2000	3000	4000	5000	6000	9000	n
1000 Tundra	44.9	26.53	0	2.04	6.12	0	20.41	49
2000 Forest	1.12	73.18	5.03	9.5	7.82	0	3.35	179
3000 Woodland	3.17	16.93	31.75	13.76	21.69	2.12	10.58	189
4000 Shrubland	0	4.86	9.93	31.35	43.49	0.22	10.15	453
5000 Grassland	0	4.48	11.3	32.2	42.64	0.43	8.96	469
6000 Riparian	0	13.68	7.37	21.05	7.37	17.89	32.63	95

9000 Other	0	0	1.82	14.29	15.2	5.17	65.53	329
Total								1763

In addition to the error matrix, a list of per land cover type accuracies along with an indication of their precision is also in order. Table 4 presents an example based on the error matrix in Table 3. With the recommended design, per class accuracy is simply the number of correct observations within the stratum (mapped classes) divided by the total number of polygons sampled. The standard error of this estimator is found by using the first equation in Step 3, as the column sums are not random variables. If stratification is based on something other than the polygons comprising the land cover map, a different formulation of the estimate is needed.

Table 4. Accuracy by cover type for the example presented in Table 3.

Cover type	Number polygons	Number sampled	Accuracy	Standard error
1000 Tundra	124	49	44.90%	5.53%
2000 Forest	3922	179	73.18%	3.23%
3000 Woodland	5456	189	31.75%	3.33%
4000 Shrubland	6077	453	31.35%	2.10%
5000 Grassland	12693	469	42.64%	2.24%
6000 Riparian	292	95	17.89%	3.23%
9000 Other	1836	329	63.53%	2.40%

Precision is often represented by reporting a confidence interval, but confidence intervals can become problematic when sampling for proportions. Exact confidence intervals based on the hypergeometric distribution are not symmetric. The normal approximation to the hypergeometric distribution is oftentimes valid, particularly when the true confidence intervals become symmetric. When N is large, other possibilities include the binomial distribution or the normal approximation to the binomial distribution. Cochran (1977) discusses the appropriateness of each method. Since the particular confidence interval used depends on each assessment's unique circumstances, the standard refers only to the estimator's standard error which is the same, regardless of which confidence interval formulation is appropriate.

Usually some indication of overall accuracy is given with the per class accuracies. A variety of measures have been suggested for describing overall accuracy, and no consensus has been reached as to which is more appropriate (Stehman, in press). The most useful measure involves a weighted mean of the per class accuracies, i.e.:

$$\sum_{i=1}^{m} \hat{p}_i w_i$$

where $_i$ is the observed users' accuracy of the i'th land cover type and w_i is the weight given to the i'th type (i.e. the number of polygons of type i divided by the total number of polygons). This is the same as Cochran's (1977) estimator for sampling proportions with stratified sampling. Its standard error is:

$$\sqrt{\sum_{i=1}^{m} s_i^2 w_i^2}$$

(Cochran 1977).

If the weights are judiciously chosen, this measure has an unambiguous probabilistic definition. Card (1982) suggested using the fraction area of each mapped type as the proper weights. This provides a measure that estimates the probability that a randomly chosen point or small parcel of land is correctly identified. Another possible set of weights is based on the total number of polygons for each class, resulting in an estimate of the probability that a randomly chosen polygon is classified correctly.

Stehman (in press) describes other measures and the associated merits and drawbacks of each measure.

Describing overall accuracy is a risky undertaking. Ultimately, accuracy depends on the map’s intended use. Any one measure will only be applicable to a certain purpose. It is not possible to devise one measure that will suit the multitude of anticipated and unanticipated uses of the land cover map. An unfortunate corollary to this is that users cannot expect a terse exposition of map accuracy, but rather one must investigate those aspects of the assessment relevant to their particular objectives.

The statistical properties of traditional assessment measures are easy to generate, but a true picture of the nature of the errors can be hard for users to decipher from an error matrix alone. The implications of observed accuracy depend on the nature of the errors. Polygons can be incorrectly classified for a multitude of reasons. They can also be reported inaccurate due to observer uncertainty of the true classification of the tract of land or uncertainty due to the location of the tract of land. In addition, certain errors are more costly or offensive than others. Wrongly attributing a land cover classification with a related alliance can be less costly than applying an attribute from a distantly related class. While accuracy by itself is important to measure, an explanation of the errors and their consequences is also important.

GAP recommends the use of "fuzzy set" analysis to help tease out these issues. The meaning of fuzzy set analysis and justification for its use are well described in Gopal and Woodcock (1994):

> *The use of fuzzy sets in map accuracy assessment expands the amount of information that can be provided regarding the nature, frequency, magnitude and source of errors in a thematic map. The need for using fuzzy sets arises from the observation that all map locations do not fit unambiguously in a single map category. Fuzzy sets allow for varying levels of a set membership for multiple map categories. A linguistic measurement scale allows the kinds of comments commonly made during map evaluations to be used to quantify map accuracy.*

For the purposes of accuracy assessment, fuzzy set analyses are accuracy measurements based on agreements more general than agreement in traditional analyses. This can be based on either observers' descriptions or *a priori* descriptions based on costs of such an error. The linguistic interpretation of accuracy from traditional percent agreement to more general fuzzy descriptions can help describe the nature of the errors and help readers interpret the validity of the map for their own purposes.

In traditional analysis, a polygon is observed and classed as one (hopefully the same) type. For fuzzy set analysis, each possible misclassification is assigned a level that represents the degree of acceptability of that misclassification. This scale can be derived from at least three different sources. If the reference data sets indicate types that are inclusions, secondary dominants, or adjacent, these can be accommodated in fuzzy analyses. Also, locational problems can be incorporated. Lastly, the degree of acceptability can be expressed from *a priori* considerations of the cost of each error. Usually this level is based on a linguistic measurement scale. Gopal and Woodcock (1994) suggested a five-level linguistic scale described below. Lacking other tested alternatives at this time, GAP recommends this system but encourages further research and experimentation with other systems.

The ranks are:

1. Absolutely Wrong: This answer is absolutely unacceptable. Very wrong.
2. Understandable but Wrong: Not a good answer. There is something about the site that makes the answer understandable, but there is clearly a better answer. This answer would pose a problem for users of the map. Not right
3. Reasonable or Acceptable Answer: Maybe not the best possible answer but it is acceptable; this answer does not pose a problem to the user if it is seen on the map. Right.
4. Good Answer: Would be happy to find this answer given on the map. Very right.
5. Absolutely Right: no doubt about the match. Perfect.

The use of the linguistic scale is also influenced by perceived use; that is, someone using the map to specifically analyze the amount and patch arrangement of an alliance because of its specific species content may not be satisfied with a rank of "3" if other alliances ranked "3" do not contain the species of interest. This user would probably want only a level "5" assessment, while another user only needing to distinguish coniferous from broadleaf vegetation may be satisfied with level "3". Fuzzy sets allow these users to be self-selecting. Therefore it is important to provide the above descriptions and to note that the rankings were made according to the intended use of the map for gap analysis. Because the use of fuzzy sets is intended to provide more useful information to the data user (and not make the map appear more accurate than it is), the way the fuzzy sets are constructed is critically important. Little use of fuzzy sets in GAP has occurred as of this writing (see Redmond et al. 1996), so these recommendations must be considered experimental and not standardized. Steps to constructing fuzzy sets:

a. construct a land cover manual (see Land Cover chapter) that describes the mapped types including dominant and codominant vegetation species, ecotones to other alliances, and inclusions of other alliances. These descriptions, developed by a plant ecologist, will provide the information needed to construct the fuzzy sets.
b. for each alliance, rank all other alliances for "correctness" on a scale of 1-5, using the information developed in a) above. Generally, rank should increase for those alliances that are frequently intermixed, share species content, or have similar habitat function. Types that are merely adjacent or are easily confused in image classification would rank as "understandable but wrong" to indicate that the error can be accounted for but is biologically incorrect. For instance, an observation showing the mapped polygon to really be of a type described as an ecotone may be ranked a "2" or "3" depending on intermixing, whereas an alliance that shares a codominant species may be a "4." It is preferred that the plant ecologist who developed the descriptions also develop the fuzzy sets, and that the sets are developed before the reference data are collected or analyzed.

The statistics of fuzzy accuracies are computed in the same way as traditional accuracy with accuracy being the proportion of observed agreements (using the fuzzy agreement criteria) per number sampled in that land cover class. A fuzzy analog to the error matrix, though, is usually not displayed due to its difficulty in interpretation.

Since fuzzy agreements are more general than traditional agreements, they can make a map appear more "acceptable" than through traditional analyses. It is important to emphasize that even though perceived accuracies improve by using fuzzy sets, the accuracy of the map is not altered. Fuzzy sets may, however, better explain the nature of the errors found in the observed reference set and relate them to other types as a measure of biological similarity.

Reporting the Results

Because the assessment protocol follows scientifically accepted standards, the reporting of the assessment study should roughly follow the framework of a scientific article and should provide enough information for persons to duplicate the assessment and derive their own accuracy measures based on their purposes. The methods part of the exposition should include the method of attributing observations, sample design, and sample and strata sizes (number of polygons observed and total number of polygons in the land cover map). The results, at a minimum, should provide the users' accuracy measure (with associated standard error or confidence interval) for each land cover type along with the users' error matrix at each level. Few assessments have been conducted and reported, therefore we encourage projects to be generous in the use of narrative and tables to communicate the results.

Management of Accuracy Assessment Data

The accuracy assessment field data will be valuable for other uses as described in the Data Archive and

Distribution chapter. The data are to be compiled as a GIS coverage containing both the locations of samples and their attributes. This coverage will be transferred by the state principal investigator to the National Gap Analysis office upon project completion. Delivery of video products will be described in the forthcoming manual. The National Center will manage the state coverages and combine them as needed for ecoregion-level accuracy assessments. The Center will also make the coverages available to other users. Therefore, it is essential that standard metadata accompany the coverage. If the data was gathered under a privacy agreement for nondistribution, the data should be treated the same as endangered species point locations from Natural Heritage Programs which can require a range of conditions including destruction of the data after its primary application, generalization, or password protection for limited use.

Limitations

Accuracy assessment is never an easy task. It requires obtaining data of higher quality with adequate coverage of space and classes to test a map. If such data were available at the beginning of a mapping project, they probably would have been used in compiling the map and therefore be unacceptable as test data. Therefore, new test data are required, but our ability to obtain an ideal data set is constrained by practical limits of technology, logistics, and cost. These three general constraints are presented here to establish the context for planning assessments. We realize that, to some degree, the constraints are interrelated. For instance, some technological limitations can be overcome with more funding. Others are beyond practical solution, such as when a private landowner denies access for a field assessment.

Technological constraints: This category of constraints includes all forms of potential measurement error relating to observation of "ground truth." The first of these sources of uncertainty in the observation data set is determining the true location of the sample plot in the field. Identifying one’s location on a topographic map or aerial photo can sometimes be difficult in rugged terrain far from known locations. GPS technology has improved this situation, providing precision at the 100-m level for quick readings. Much greater precision is available at the expense of longer recording time or the cost of using differential GPS. Even GPS has limited utility in some environments, particularly densely forested or canyon landscapes, where few satellites can be detected by the receiver. Fortunately, a small amount of uncertainty in locating the sampling unit in the field is not seriously detrimental to statistical inference in accuracy assessment. It is just a part of measurement error in the test data.

One approach to testing the accuracy of mesoscale land cover maps is to visit large sample units. This design avoids the problem of mistakenly observing a neighboring polygon or detecting heterogeneity well below the intended resolution of the map. The technical problem with this approach is that the observer cannot see the entire unit simultaneously, and it is difficult to systematically visit the entire unit. That is, the observations are only made locally rather than integrating over larger areas. Thus, the field data may inadequately record and integrate all of the attributes of a large sample unit.

Switching from field observation to low-elevation remotely sensed data such as aerial photos or videography can reduce the problem of not seeing the whole sampling unit, but floristic detail often cannot be observed in such reference sources. Consequently, they may not be an acceptable substitute for field observation, leaving us with a trade-off between synoptic perspective and measurement detail.

Logistical constraints: Some locations are simply too far from existing roads or trails to be visited without incurring extravagant expense, such as chartering a helicopter. Access may be denied by the landowner or manager as well. Safety considerations may also preclude access to locations where either the route is too dangerous (e.g., crossing rivers, scaling cliffs) or the sample site is hazardous (e.g., weapons-testing range). If sampling measurements cannot be made from some remote sensing platform, then these sites may need to be dropped from the sampling scheme and either replaced with more accessible ones or reported as nonresponses in the assessment. This is just a fact of life that should be acknowledged when reporting the assessment.

Financial constraints: This constraint is perhaps the least binding (theoretically), because more funding can be supplied if the political will deems it worthwhile. For our purposes here, however, we must accept that the marginal cost of reaching distant sample sites or of narrowing the confidence intervals have practical limits. We can present the options and their associated costs, but we also present what we believe is a reasonable balance between the constraints and the need for rigor and scientific soundness. Partnerships with cooperators that will benefit both parties from a rigorously tested land cover map and acquisition of well-documented reference data can help overcome GAP funding limitations.

Continuing Research

Map accuracy assessment is by no means a standardized procedure. Many research questions relating to field methods over large geographic areas and to statistical techniques and measures are still being explored. The workshop members highlighted a few of these areas that could help move toward a Gap Analysis protocol. Readers can probably think of additional topics. Overriding all these issues is the need to test the proposed strategy in a representative set of states to resolve issues and confirm that the proposal satisfies the criteria stated in the thematic accuracy section.

- How to quantify the contribution of uncertainty from observational error to the apparent map error: A perennial problem with testing the accuracy of land cover maps derived from remote data is that the error of the method used to check the remote data is assumed to be negligible. The incorporation of uncertainty in the observed data set poses several problems and is an open question. If a parcel of land is observed incorrectly, the apparent accuracy is really inaccurate or the apparent inaccuracy may, in actuality, be correct. Inaccurate observations place an upper limit on observed accuracy. An assessment can be no more accurate than either the map or the observer performance.

  In addition, the problem of gradations between cover types introduces apparent error in both the remote and ground data. For example, the transition from shrub to young forest is gradual. The label applied to transitional cover types will depend largely on the interpreter. This source of apparent error can be reduced by clearly defining cover types. To quantify the magnitude of observer variability, a random subset of sampled sites could be visited independently by different observers, or even by the same observer at different times. Since all ground data collection is expensive, multiple site visits might be confined to easily accessible sites, or the technique used only over a limited area to estimate the effect of observer variability.
- The potential of videography: The use of videography eliminates two major problems associated with ground data collection: field inspection of private lands and accessibility limitations because of topography and/or distance from roads. Difficulties with videography were described earlier. We suggest that videography be compared with ground-sampling of polygons in a trial state in which videography transects have already been flown.
- Developing new statistics for using maplets in accuracy assessment: Occasionally, a finer-scale land cover map over a small area will be available for comparison with the statewide land cover map produced by Gap Analysis. The California Gap Analysis Project had access to an independent detailed map of a small area in southern California that favorably compared to the coarser-resolution Gap Analysis polygons (Stoms 1996). However, there is no standard statistical analysis for the technique beyond a simple comparison of the finer with the coarser-scale map.
- Develop new accuracy assessment methods that evaluate magnitude of errors: Existing methods treat comparisons between map and reference data as either correct or erroneous. In vegetation, however, the classes often are similar or even overlap floristically. A disagreement between two such similar classes should be weighted differently than disagreement of greater magnitude. Some progress has been made with the use of fuzzy set theory (Gopal and Woodcock 1994) and a weighted kappa statistic (Agresti 1990) to portray the degree of disagreement in accuracy assessments, but more research will be needed before standard methods can be prescribed.

Literature Cited

Agresti, A. 1990. Categorical Data Analysis. John Wiley and Sons, Inc. New York. 367 pp.

Card, D.H. 1982. Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing* 48:431-439.

Cochran, W.G. 1977. Sampling Techniques. 3rd ed. John Wiley and Sons, Inc. New York.

Congalton, R.G. 1991. A review of assessing the accuracy of classification of remotely sensed data. *Remote Sensing of the Environment* 37:35-46

Dzur, R.S., M.E. Garner, K.G. Smith, W.F. Limp, D.G. Catanzaro, and R.L. Thompson. 1996. Cooperative accuracy assessment strategies for sampling a natural land cover map of Arkansas. In: Spatial Accuracy Assessment in Natural Resources and Environmental Sciences: Second International Symposium, edited by H. Todd Mowrer, Raymond L. Czaplewski, and R.H. Hamre. May 21-23, 1996 Fort Collins, CO, General Technical Report RM-GTR-277.

Edwards, T.C., G.G. Moisen, and D.R. Cutler. In press. Assessing map accuracy in a remotely-sensed, ecoregion-scale cover-map. *Remote Sensing of Environment*

Fenstermaker, L.K., editor. 1994. Remote Sensing Thematic Accuracy Assessment: A Compendium. American Society for Photogrammetry and Remote Sensing.

Goodchild, M.F. 1989. Modeling error in objects and fields. Pp. 107-144 in Accuracy of Spatial Databases. M. Goodchild and S. Gopal, eds. Taylor & Francis, New York

Gopal, S., and C. Woodcock. 1994. Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing* 60:181-188.

Mowrer, H.T., R.L. Czaplewski, and R.H. Hamre. 1996. Spatial Accuracy Assessment in Natural Resources and Environmental Sciences: Second International Symposium. USDA Forest Service General Technical Report RM-GTR-277.

Peer Review Panel. 1994. Peer Review Panel Report of the National Gap Analysis Program for the National Biological Survey. 25 pp.

Redmond, R.L., Z. Ma, T.P. Tady, J.C. Winne, J. Schumacher, J. Troutwine, S.W. Holloway, C.P. McGuire, R.L. Righter, L.G. McWethy, D.W. Schirokauer, P.C. Thornton, C. Tobalske, W. Gustafson, J. Handley, J. Hinsha, A. Jones, S. Stegman, A. Vander Meer, M.M. Hart, K.P. McLaughlin, and W.A. Williams. 1996. Mapping Existing Vegetation and Land Cover Across Western Montana and Northern Idaho: Final Report. Wildlife Spatial Analysis Lab, Montana Cooperative Wildlife Research Unit, University of Montana, Missoula.

Stehman, S.V. In press. Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*.

Stehman, S.V., and R. L. Czaplewski. 1997. Basic structures of a statistically rigorous thematic accuracy assessment. In: Proceedings of the Annual Meetings of the American Society of Photogrammetry and Remote Sensing, Seattle, WA, April 6-10, published by the American Society of Photogrammetry and Remote Sensing, Bethesda, MD, Volume 3, pp. 543-553.

Stoms, D. M. 1996. Validating large-area land cover databases with maplets. *Geocarto International* 112:87-95.

Stoms, D., F. Davis, C. Cogan, and K. Cassidy. 1994. Assessing land cover map accuracy for Gap Analysis. In J.M. Scott and M.D. Jennings, editors. A Handbook for Gap Analysis. Idaho Cooperative Fish and Wildlife Research Unit, University of Idaho. Moscow, ID.

Thompson, B.C., P.J. Crist, J.S. Prior-Magee, R.A. Deitner, D. Garber, and M.A. Hughes. 1996. Gap analysis of biological diversity conservation in New Mexico using geographic information systems. Final Gap Analysis report. U.S. Dept. Interior, New Mexico Coop. Fish and Wild. Res. Unit, Las Cruces, NM.

Selected Bibliography (in addition to above Literature Cited)

Chrisman, N.R. 1991. The error component in spatial data. In: Geographical Information Systems - Volume 1: Principles, edited by D.J. Maguire, M.F. Goodchild, and D.W. Rhind. Longman Scientific & Technical, pp. 165-174.

Cogan, C.B., and T.C. Edwards, Jr. 1994. Metadata standards for Gap analysis. Gap Analysis Technical Bulletin 3. Idaho Cooperative Fish and Wildlife Research Unit, University of Idaho. Moscow, ID.

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20:37-46.

Federal Geographic Data Committee (FGDC). 1994. Content Standards for Digital Geospatial Metadata. 8 June 1994. 66 pp.

Fenstermaker, L.K. 1991. A proposed approach for national to global scale error assessments. In: Proceedings of GIS/LIS'91, Atlanta, Georgia, October 28-November 1, 1991 pp. 293-300.

Foody, G.M. 1992. On the compensation for chance agreement in image classification accuracy assessment. *Photogrammetric Engineering and Remote Sensing* 58(10):1459-1460.

Goodchild, M.F., F.W. Davis, M. Painho, and D.M. Stoms. 1991. The Use of Land Cover Maps and Geographic Information Systems for Assessing Conifer Lands in California, NCGIA Technical Report 91-23, NCGIA, Santa Barbara, California. 75 pp.

Goodchild, M.F., and S. Gopal, editors. 1989. The Accuracy of Spatial Databases. Taylor & Francis, London.

Janssen, L.L.F., and F.J.M. Van der Wel. 1994. Accuracy assessment of satellite derived land-cover data: a review. *Photogrammetric Engineering and Remote Sensing* 60(4):419-426.

Jennings, M.D. 1993. Natural Terrestrial Cover Classification: Assumptions and Definitions. Gap Analysis Technical Bulletin 2, Idaho Cooperative Fish and Wildlife Research Unit, University of Idaho. Moscow, ID. 28 pp.

Jensen, J.R. 1986. Introductory Digital Image Processing. Prentice-Hall, New Jersey.

Merchant, J., L. Yang and W. Yang. 1993. Validation of continental-scale land cover data bases developed from AVHRR data. In: Proceedings of Pecora 12, Sioux Falls, SD, August 24-26, 1993. Bethesda, MD: American Society for Photogrammetry and Remote Sensing, pp. 63-72.

Moisen, G.G., T.C. Edwards, Jr., and D.R. Cutler. 1994. Spatial sampling to assess classification accuracy of remotely sensed data. In: Proceedings of the Symposium on Environmental Information Management and Analysis: Ecosystem to Global Scales, Albuquerque, NM, May 20-22, 1993. National Science Foundation, Washington, D. C.

National Center for Geographic Information and Analysis. 1994. Final Report of the Accuracy Assessment Task Force (California Assembly Bill 1580). National Center for Geographic Information and Analysis, Santa Barbara, CA.

Skidmore, A.K., and B.J. Turner. 1992. Map accuracy assessment using line intersect sampling. *Photogrammetric Engineering and Remote Sensing* 58(10):1453-1457.

Stehman, S. 1996. Estimating the kappa coefficient and its variance under stratified random sampling. *Photogrammetric Engineering and Remote Sensing* 62(4):401-407.

UNESCO. 1973. International Classification and Mapping of Vegetation. United Nations Educational, Scientific and Cultural Organization, Paris. 35 pp.

Appendix 1 - Summary of related land cover mapping programs

Coastwatch Change Analysis Project

The goal of the Coastwatch Change Analysis Project (C-CAP) is to develop a comprehensive, nationally standardized information system for monitoring land cover change and habitat change in the coastal regions of the U.S. C-CAP is part of the Estuarine Habitat Program (EHP) of NOAA's Coastal Ocean Program. Satellite imagery (primarily TM), aerial photography, and field data are interpreted, classified, analyzed, and integrated with other digital data in a geographic information system. The program will delineate coastal emergent and submerged wetland habitats and adjacent uplands and monitor changes in these habitats on a 1-5 year cycle. This type of information and frequency of detection are required to improve scientific understanding of the linkages of emergent and submerged wetland habitats with adjacent uplands and with the distribution, abundance, and health of living marine resources. The monitoring cycle will vary regionally according to the rate and magnitude of change.

North American Landscape Characterization--Pathfinder Project

The North American Landscape Characterization--Pathfinder Project (NALC) generates data products documenting extent of land cover changes on the North American continent. Satellite data used in the NALC project include Landsat MSS data from the years 1991, 1986, and 1973, plus or minus one year. These years represent the best data available and most complete coverage of North America. MSS data cost is relatively low. Criteria for selection of images from these three periods are: similar seasonal periods, 30% or less cloud cover, and high quality images. Triplicate images provide a baseline data set that can be used to: evaluate change between three dates over the last 20 years, detect future geographic shifts in land cover and vegetative communities, determine areas of anthropogenic change in land cover, document changes in biodiversity, contribute to the future mapping of North American biotic community zones, and measure change in carbon abundance over the past 20 years. In addition to the coregistered, multitemporal images, the program is producing spectral cluster images, change detection images, and eventually land cover class maps.

Conterminous U.S. Land Cover Database, USGS

The United States Geological Survey's Earth Resources Observation Systems (EROS) Data Center and the University of Nebraska have developed a United States prototype for a proposed global land cover characteristics database derived from 1-km Advanced Very High Resolution Radiometer (AVHRR) satellite data (Loveland et al. 1993). The database portrays regions composed of similar land cover mosaics, as defined by multitemporal AVHRR NDVI data collected in 1990. Data refinement was made on the basis of several national data sets including elevation, climate, and ecological regions (Loveland et al. 1991, Brown et al. 1993).

The U.S. Forest Service Intermountain Fire Sciences Laboratory in Missoula, Montana, with assistance from the Bureau of Land Management and the U.S. Geological Survey, has directed a volunteer effort to evaluate the EROS land cover characteristics database. During 1993 and 1994, volunteers from several state and federal agencies recorded land cover observations at over 3000 1 km^2 sampling locations in the conterminous United States. Sampling locations were selected using a stratified-random procedure. Details are available in a study plan described by Burgan and colleagues (1993).

Ground information collected as part of the Gap Analysis Program's proposed accuracy assessment plan would serve as valuable supplemental information for assessing the conterminous U.S. land cover characteristics database. In addition, as a new USGS North America land cover product is produced in the next 18 months, the Gap Analysis data may serve as a primary source of ground-based data.

Brown, J.F., T.R. Loveland, J.W. Merchant, B.C. Reed, and D.O. Ohlen. 1993. Using multisource data in global land cover characterization: Concepts, requirements, and methods. *Photogrammetric Engineering and Remote Sensing* 59:977-987.

Burgan, R.E., C. Hardy, D. Ohlen, T.R. Loveland, J.F. Brown, and B.C. Reed. 1993. Land Characterization Map Evaluation. U.S. Forest Service Study Plan No. 440152. 22 pp.

Loveland, T.R., J.W. Merchant, D.O. Ohlen, J.F. Brown, B.C. Reed, and L.T. Steyaert. 1993. Conterminous U.S. Land Cover Characteristics Data Set: 1990 Prototype. EROS Data Center, U.S. Geological Survey CD-ROM [includes Readme file.]

Loveland, T.R., J.W. Merchant, D.O. Ohlen, and J.F. Brown. 1991. Development of a land cover characteristics database for the conterminous U.S. *Photogrammetric Engineering and Remote Sensing* 57:1453-1463.

National Water Quality Assessment Program

The USGS National Water Quality Assessment Program (NAWQA) is designed to describe the status and trends in the quality of the nation’s water resources. The program is organized into 60 hydrological systems called study units, which encompass about 45 percent of the land area of the conterminous United States. The study unit data are aggregated up to provide information about regional and national water quality concerns.

Environmental Monitoring and Assessment Program

The Environmental Monitoring and Assessment Program (EMAP) is an interdisciplinary, interagency program being designed and initiated through the U.S. Environmental Protection Agency’s (EPA) Office of Research and Development. The program’s objectives require that EMAP be an interagency program in which EPA is but one of the participants. EMAP’s mission is to address regional, long-term environmental problems occurring at regional and national scales. This mission will complement the local-scale, shorter-term monitoring programs within state and local agencies.

Multi-Resolution Landscape Characteristics Consortium

EMAP’s land cover mapping effort has joined with the Gap Analysis Program and the other partner agencies of the Multi-Resolution Landscape Characteristics (MRLC) group to generate land cover data for the conterminous U.S. The MRLC is working to develop the classification scheme, accuracy assessment methods, and overall infrastructure to support this effort. EMAP will contribute personnel and resources for the collection of field data to be used by the MRLC for accuracy assessment

Appendix 2 - List of participants in the Land Cover Assessment workshop, February 3-4, 1994, Santa Barbara, CA

Thaddeus Bara
ManTech Envir. Technology, Inc.

Kelly Cassidy
Wash. Coop. Fish and Wildlife Unit

Christopher Cogan
Idaho Coop. Fish and Wildlife Unit

Charles Convis
ESRI Conservation Program

Frank Davis
Department of Geography
U.C. Santa Barbara

Jerry Dobson
Oak Ridge National Lab

Thomas Edwards
UT CFWRU

Allan Falconer
Department of Geography
Utah State Univ.

Michael Goodchild
Department of Geography
U.C. Santa Barbara

Violet Gray
Department of Geography
U.C. Santa Barbara

Dennis Grossman
The Nature Conservancy

Mike Jennings
Idaho Coop. Fish and Wildlife Unit

Gretchen Moisen
USDA Forest Service
FIA-INT Research Station

Tom Muir
National Biological Survey

Tony Olsen
USEPA

Brad Reed
USGS, EROS Data Center

Bill Reiners
Department of Botany

Univ. of Wyoming
Denice Shaw

EMAP Center
USEPA

Mirjam Stadelmann
ESRI Applications Division

Steve Stehman
SUNY-ESF

David Stoms
Department of Geography
U.C. Santa Barbara

Gail Thelin
USGS-WRD

Dorsey Worthy
USEPA/EMSL-LV/AMS

Mike Jennings
Idaho Coop. Fish and Wildlife Unit

Appendix 3 - List of participants in the Land Cover Assessment workshop, April 9-10, 1996, Denver, CO

Patrick Crist
National GAP office
University of Idaho

Bob Deitner
New Mexico GAP
New Mexico State Univ.

Rob Dzur
Arkansas GAP
University of Arkansas

Jack Finn
New England GAP
University of Massachussetts

Bill Halvorson
Arizona GAP
AZ Cooperative Park Studies Unit

Mike Kunzmann
Arizona GAP
University of Arizona

Chris Lauver
Kansas GAP
Kansas Biological Survey

Don Schrupp
Colorado GAP
CO Div of Wildlife

Scot Smith
Florida GAP
University of Florida

Mike Story
NPS Technology Transfer Center

Kathryn Thomas
Arizona GAP
CO Plateau Research Center

Bruce Thompson
New Mexico GAP
NM Coop Fish and Wildlife Research Unit

Appendix 4 - Spatial Accuracy

It is critical that the spatial accuracy of the map is known when comparing it with reference data, so that errors attributable to spatial misalignment are not confused with thematic errors. For instance, if a 1-ha polygon is to be assessed, but the map and/or the reference data cannot be located more precisely than 100 m, there is some probability that the wrong polygon will be assessed due to misalignment. This task is not reported with the thematic assessment tasks as it is a quality control task that should be undertaken earlier in the map creation process.

Points

Circular Map Accuracy Standard: This method assumes a circular model with equal standard errors in the coordinated directions and no correlation, and defines a circle within which the true point is expected to lie nine times out of ten.

Lines

Perkal epsilon band: This measure uses a band of width epsilon surrounding the line such that the band contains any randomly chosen point on the line with a known probability.

Polygons

Measures of point, line, and area serve to assess polygon error.

Gap Analysis locational error assessment can be considered in two basic ways. Land cover polygon boundaries can be assessed, or the map and ground can be compared at point locations. We refer to these as Transect Measure and Benchmark Measure and describe scenarios for each below.

1) Transect Measure

Using a large-scale representation of truth, e.g. maps, video, GPS, etc., apply a randomly oriented transect across a map and measure locations of polygon/transect intersections. At each transect intersection, differences between the boundary locations on the large-scale map and the map being tested can be measured. Resulting values can then be used to determine Root Mean Square (RMS) error in ground unit meters. The length of each transect should be determined by the number of intersections needed for sampling (Skidmore and Turner 1992).

Issues: This method has several drawbacks, most notably the fact that even in the best cases, we are using another map as truth, not actual ground data. It is also important to resolve what to do in the case of missing or additional polygons. Accurate control points are necessary to ensure that the transect is correctly aligned between the map and the other source.

2) Benchmark Measure

Land cover maps produced for Gap Analysis are an abstraction of continuous spectral data gathered from satellite and other sensors. Some of the boundaries can be considered discrete at our mapping resolution, whereas other areas are better described as transition zones. Rather than attempting to precisely document an imprecise set of features, an alternative method of documenting spatial map accuracy describes the accuracy of identifiable points. These "benchmarks" should not be coincident with the ground control points used in the original map production. As described above, the Circular Map Accuracy Standard could be used to document the spatial quality of the Gap Analysis land cover maps.

Issues: This method does not yield direct information about vegetation boundaries, rather it tells us more about geometric registration of the intermediate products used to create a map of vegetation. Also, many of the best locations for benchmarks will have already been selected for use as ground control points, though it should still be possible to locate distinct features visible in the Thematic Mapper imagery. Typical locations which could be useful are dams, corners of logged forests, and other boundaries between anthropogenic and natural areas. Based on a reasonable distance resolution from 1:100,000 scale source materials, map locations within 50 meters of true ground locations could be considered correct (Goodchild et al. 1991).

Related metadata elements (Cogan and Edwards 1994):

Horizontal Positional Accuracy
Horizontal Positional Accuracy Explanation
Vertical Positional Accuracy
Vertical Positional Accuracy Explanation
Distance Resolution
Altitude Resolution
Source Distance Resolution

Appendix 5 - Field Survey Example

1997 GAP Field Sheet Protocol

Data Collector:

Write in your name(s) **and** the name of the agency that you work for.

County Name:

Please record the name of the county where the survey site is located.

T-R-S:

Record the township, range, and section designations for the survey site here. Also please indicate which quarter section or quarter quarter section if possible. **This information is especially important if you do not have access to a GPS unit, so please be as precise as**

possible.

UTM Zone:

If you have access to a GPS unit, please record the UTM zone in which the survey site is located.

GPS UTM Coordinates:

If you are using a GPS unit, please record the UTM coordinates in the space provided.

GPS File Name:

If your GPS rover unit is equipped with an output port, record the file name on the field sheet in the space provided. At the end of the day, please download the data to a 3 ‰" floppy, and return the disk and field sheets to us.

(1) Vegetation Alliance:

Use the vegetation alliance descriptions and key provided with this sheet to assess which is the correct alliance for the area immediately surrounding the GPS point (the survey site).

These cover classifications were developed to resemble the FGDC standard vegetation alliances. Most types of vegetation should be generally represented here, however if you want to further clarify, you can provide more information in the "Notes" section at the bottom or on the back of the sheet.

List the dominant species of the survey site in the spaces provided. Estimate the average height of the vegetation (excluding wooded or forested areas). Also, check the management practices and the physical characteristics listed here that apply to the site.

If you are unsure of your assessment, please provide additional possible alliance types in the spaces provided.

(2) Canopy Cover:

If you are surveying a forest or woodland please provide an estimate of the percent canopy cover. A spherical densiometer should be used to estimate percent canopy cover if one is available. If not, a "best guess" approximation will be sufficient.

(3) Understory:

If the estimate of canopy cover is less than 80%, then please identify the general form of the vegetation comprising the understory. Check all that apply.

(4) Vegetation Cover:

If the survey site is not wooded or forested, please provide a general estimate of the percent vegetation cover.

(5) Soil Color & Texture:

If the vegetation at a site is sparse (ie. ground cover < 75%) and there is bare soil visible, please provide a general description of the soil color using the color chart provided. Also, please give a general description of the soil texture.

(6) Sketch Map, Photos & Notes:

The sketch map area provides a space for a representative drawing of the major features of the survey site.

On the sketch map, please indicate the approximate size and shape of the area that you surveyed. Clearly label the position of the GPS reading. Indicate all visible plant community boundaries surrounding the site. Label any man-made structures (roads, buildings, wells, fences, windmills, etc.), water features, and/or landmarks in the immediate area.

If you have been provided with a disposable camera, please take two or three landscape shots to characterize the setting in which the site occurs. Also, please take a near-vertical shot of the site. Include the location and look direction of the photos taken on site and label them with the corresponding roll and photo numbers. The roll number has been written on the camera body.

The "Notes" section should be used to provide any additional information that you feel will be helpful in the interpretation of the information you have given. If you need more room, please continue on the back of the sheet.

A sample field survey sheet has been provided for you to use as a guide when conducting surveys of your own. However, if you need further clarification, feel free to call Marlen Eve at 402-472-9984 (e-mail: meve@tan.unl.edu) or Mike Bullerman at 402-472-2565 (e-mail: mbuller@tan.unl.edu).

Nebraska GAP Analysis 1997 Field Survey

Nebraska GAP Analysis	**1997**	**Field Survey**

Date: ____________________

UTM Zone: _________

Data Collector: ____

GPS **UTM** Coordinates: _________________________________E

_________________________________N

County Name: ____

Township:______ Range:______ Section:______ Quarter Section:______

GPS File Name: _______________________________________

---

(1) Vegetation Alliance: ____

Dominant Species:____ ____

____ ____

Vegetation Height: 0 - 0.5m____ 0.5 - 1m____ 1 - 2m____ >2m____

Grazed _____ Hayed _____ Irrigated _____ No Visible Management_____

Landscape Position: floodplain____ terrace____ hillslope____ hillcrest____

Aspect: N____ NE____ E____ SE____ S____ SW____ W____ NW____

Slope: 0o -10o____ 10o - 20o____ 20o - 45o____ >45o____

Are you highly confident of your alliance assessment?

Yes No

If No, list other possibilities:

1) _____________________________________

2) _____________________________________

(2a) If forest or woodland, estimate **% Canopy Cover**:

25-40%____ 40-60%____ 60-80%____ >80%____

(2b) If canopy cover is less than 80%, identify **Understory**:

____ herbaceous vegetation ____ small trees (seedlings)
____ saplings ____ shrubs ____ bare ground
other: ____

(3a) If NOT forest or Woodland, estimate **% Vegetation Cover**:

0-25%____ 25-50%____ 50-75%____ 75-100%____

(3b) If vegetation cover is less than 75%, please note:

Soil Color: __________________ WET DRY (circle one)
Soil Texture:

____ sandy ____ silty ____ clay ____ loamy
other:________________________________

---

(4) The Sketch Map should include the following :

- point of GPS reading
- location and look direction of photos labeled with roll & photo number
- clearly delineated crop types and whether they are dry land or irrigated
- clearly delineated plant community boundaries
- man made structures (e.g. wells, fencesbuildings, etc.)

NOTES: ____

____

____

____

____

____

____

DRAFT Heritage Program Key to the Nebraska GAP Alliance Classes - June, 1997

1. Sites with well developed tree canopies (forests & woodlands)

- 2. > 60% tree canopy coverage with interlocking crowns (forests)
    - 3. dominated by ponderosa pine, escarpments in the Panhandle and along the central Niobrara River....**Ponderosa Pine Forest**
      3. dominated by deciduous trees
        - 4. Upland sites dominated by oaks and hickories....**Oak Hickory Forest**
          4. Floodplain sites dominated by cottonwoods and willows.... **Cottonwood-Willow Forest**

- 2. <60% tree canopy coverage, crowns not interlocking (woodlands)
  - 5. Canopy dominated by pines or junipers (coniferous woodlands)
    - 6. Upland sites with canopy of ponderosa pine and subcanopy of junipers....**Pine-Juniper Woodland**
      6. Upland sites dominated solely by junipers....**Juniper Woodland**

    5. Canopy dominated by deciduous trees (deciduous woodlands)
    - 7. floodplain sites dominated by cottonwoods and willows....**Cottonwood-Willow Woodland**
      7. Upland, ravine or terrace sites not dominated by cottonwoods and willows
      - 8. sites dominated by primarily bur oak and other oaks....**Oak-Bluestem Woodland**
        8. Sites dominated by green ash and elms....**Green Ash-Elm Woodland**

1. Sites without tree canopies, dominated by shrubs or herbaceous vegetation (shrublands and grasslands)

- 9. well-developed shrub layer present (shrublands)
  - 10. Sites located on stream floodplains
    - 11. Dominated by willows....**Sandbar willow shrubland**
      11. Dominated by dogwoods....**Dogwood shrubland**

    10. Sites located on uplands or Sandhill valleys
    - 12. Soils peat or muck, located in Sandhill valleys, willows the dominant shrub... **Sandhills fen**
      12. Soil not peat or muck, upland sites

      13. Dominant shrub is sandsage, sites of Sandhill uplands .... **Sandsage shrubland**
      13. Dominant shrub not sandsage
      - 14. Sites associated with deciduous woodlands, dominant shrub is smooth sumac...**Smooth sumac shrubland**
        14. Sites not associated with deciduous woodlands, dominant shrub is western snowberry (buckbrush)....**Western Snowberry Shrubland**

  9. well-developed shrub layer absent (grasslands)
  - 15. Sites located in the Sandhills
    - 16. Upland dune sites, dominated by sand bluestem, little bluestem and needle-and-thread....**Sandhills upland tallgrass prairie**
      16. Interdunal valley sites
      - 17. Dominant grasses are switchgrass and big bluestem, sites not subirrigated and drier....**Sandhill dry valley prairie**

17. Dominant grasses are prairie cordgrass, bluejoint and northern reedgrass, sedges and rushes common, sites subirrigated and wetter....**Sandhill lowland tallgrass prairie**

15. Sites not located in the Sandhill (except for Bulrush-Cattail wetland)

- 18. Lowland or flooded sites
  - 19.Bulrush and cattail dominated marsh.....**Bulrush-Cattail wetland**

    19. Bulrush and cattails do not dominate
    - 20. Soils saline or alkaline, lowland sites dominated by saltgrass, foxtail barley, alkali sakatoon and other salt-tolerant grasses.....**Saltgrass Foxtail Saline Prairie**

      20. Soils not saline or alkaline
      - 21. Sites associated with temporarily flooded upland depressions, primarily in sw Nebraska, buffalo grass and western wheatgrass are common...**Playa**

        21. Sites associated lowlands and stream valleys, big bluestem and prairie cordgrass are common....**lowland tallgrass prairie**

18. Upland sites

- 22. Site dominated by big bluestem, Indiangrass and little bluestem, eastern fourth of Nebraska.....**upland tallgrass prairie**

  22. Sites not dominated by the above listed grasses
  - 23. Sites dominated by the shortgrasses buffalo grass and blue grama, located in the Panhandle....**Blue grama shortgrass prairie**

    23. Sites not dominated by shortgrass, but by midgrasses.
    - 24. Sites located on central Nebraska loess soils, dominant grasses are little bluestem and bluegrama....**Little bluestem-grama prairie**

      24. Sites not located on central Nebraska loess soils.
      - 25. Western wheatgrass is the dominant grass....**Western Wheatgrass mixedgrass prairie**

        25. Western wheatgrass not the dominant grass, needle-and-thread and bluegrama are the dominant grass.... **Needleandthread-grama mixedgrass prairie**

Nebraska Gap Analysis Project
Vegetation Classification Scheme - DRAFT - 6/11/97

General Cover

Cover Characteristic
Physiographic Position
TNC/FGDC Alliance
Associations

NE Gap Class Name and Bailey□s Ecoregions

FOREST Generally >60% tree canopy with crowns interlocking.

Coniferous Forest
Upland
Pinus Ponderosa Forest Alliance
Ponderosa pine/chokecherry forest.

Ponderosa Pine Forest
331F, 331H, 332C

Deciduous Forest
Upland
Quercus Alba-Quercus Rubra-Carya Forest Alliance
Oak/hickory forest.
Quercus Macrocarpa Forest Alliance
Bur oak/american basswood/eastern hophornbeam forest.
Bur oak/american basswood/paper birch canyon forest.

Oak-Hickory Forest
251C, 332B, 332D

Lowland
Fraxinus Pennsylvanica-(Ulmus americana)-Celtis Temporarily Flooded Forest Alliance
Green ash/elm/hackberry forest.

Green Ash-Elm Forest
251B, 251C, 332B, 332D, 332E

Populus Deltoides Temporarily Flooded Forest Alliance
Cottonwood/willow forest.

Cottonwood-Willow Forest
251B, 251C

WOODLAND Open stands of trees comprise 25 - 60% cover

Coniferous Woodland

Upland

Juniperous Spp. (J. Scopulorum, J. Virginiana) Woodland Alliance

Rocky mountain juniper/ricegrass woodland.
Eastern redcedar/ricegrass woodland.

Juniper Woodland
251C, 331C, 331F, 331H, 332C, 332D

Pinus Ponderosa Woodland Alliance

Ponderosa pine/rocky mountain juniper woodland.
Ponderosa pine/little bluestem woodland.

Pine-Juniper Woodland
331H, 332C, 332D

Deciduous Woodland

Upland

Fraxinus Pennsylvanica - (Ulmus Americana) Woodland Alliance

Green ash/american elm/eastern hophornbeam woodland.

Green Ash-Elm Woodland
331F, 331H, 332C, 332E

Quercus Macrocarpa Woodland Alliance

Bur oak/big bluestem - switchgrass woodland.
Bur oak/big bluestem - porcupinegrass woodland.

Oak-Bluestem Woodland
251C, 331C, 332C, 332D, 332E

Lowland

Populus Deltoides Temporarily Flooded Woodland Alliance

Cottonwood - peach leaved willow/sandbar willow floodplane woodland.
Cottonwood/willow/prairie cordgrass woodland.

Cottonwood-Willow Woodland
All

SHRUBLAND Generally shrubs > 0.5 m. tall forming > 25% cover with <25% tree cover.

Evergreen Shrubland

Upland

Artemesia Filifolia Shrubland Alliance

Sand sagebrush/sand bluestem shrubland.

Sandsage Shrubland
331C, 331H, 331F

Deciduous Shrubland

Upland

Rhus Glabra Shrubland Alliance

Smooth sumac shrubland.

Smooth Sumac Shrubland
251C, 332C, 332D, 332E

Lowland

Salix Exigua Temporarily Flooded Shrubland Alliance

Sandbar willow shrubland.
Sandbar willow/mesic graminoid shrubland.

Sandbar Willow Shrubland
All

Cornus Drummondii Temporarily Flooded Shrubland Alliance

Dogwood Shrubland

Dogwood Shrubland

251B, 251C, 332C, 332D, 332E

Symphoricarpos Occidentalis Temporarily Flooded Shrubland Alliance
Western snowberry shrubland.

Western Snowberry Shrubland
331C, 331F, 331H, 332C

Betula Pumila - (Salix Spp.) Saturated Shrubland Alliance
Meadow willow/sedge - marsh fern shrubland.

Sand Hills Fen
332C

GRASSLANDS Graminoids, forbs and ferns dominate, with trees and shrubs generally less than 25% cover.

Tallgrass Prairie
Upland
Andropogon Gerardii - (Sorghastrum Nutans) Herbaceous Alliance
Big bluestem - indiangrass tallgrass prairie.

Upland Tallgrass Prairie
251B, 331C, 331F, 331H, 332D, 332E

Andropogon Hallii Herbaceous Alliance
Sand bluestem - prairie sandreed tallgrass prairie.

Sand Hills Upland Tallgrass Prairie
331C, 332C

Lowland
Andropogon Gerardii-(Calamagrostis Canadensis, Panicum Virgatum) Herbaceous Alliance
Big bluestem - bluejoint reedgrass - sunflower tallgrass prairie.
Spartina Pectinata Temporarily Flooded Herbaceous Alliance
Prairie cordgrass - reedgrass - sedge wet prairie.

Lowland Tallgrass Prairie
251C, 332B, 332D, 332E

Andropogon Gerardii-(Calamagrostis Canadensis, Panicum Virgatum) Herbaceous Alliance
Big bluestem - switchgrass tallgrass prairie.
Spartina Pectinata Temporarily Flooded Herbaceous Alliance
Bluejoint - rush - sedge wet prairie.

Sand Hills Lowland Tallgrass Prairie
332C

Andropogon Gerardii - Schizachyrium Scoparium Herbaceous Alliance
Big bluestem - little bluestem dry valley prairie.

Sand Hills Dry Valley Prairie
331C, 332C

Typha Spp. - (Scirpus Spp. - Juncus Spp.) Seasonally Flooded Herbaceous Alliance
Bulrush - cattail - burreed shallow marsh.
Typha (Angustifolia, Latifolia) - Scirpus Spp. Semipermanently Flooded Herbaceous Alliance
Bulrush - cattail marsh.
Cattail marsh. Scirpus Acutus - (Scirpus Validus) Semipermanently Flooded Herbaceous Alliance
Bulrush alkaline marsh.
Bulrush marsh Scirpus Pungens Semipermanently Flooded Herbaceous Alliance
Bulrush - sea blite alkaline marsh.
Typha Spp. - Carex Spp. Saturated Herbaceous Alliance
Cattail - scouring rush - sedge seep.

Bulrush-Cattail Wetland
All

Mixedgrass Prairie
Upland
Pascopyrum Smithii Herbaceous Alliance
Western wheatgrass mixedgrass prairie
Western wheatgrass needle-and-thread mixedgrass prairie

Western Wheatgrass Mixedgrass Prairie
331F, 331H

Schizachyrium Scoparium - Bouteloua Curtipendula Herbaceous Alliance

Little bluestem - side-oats grama - hairy grama - spanish bayonet mixedgrass prairie.

Little bluestem - side-oats grama mixedgrass prairie.

Little Bluestem-grama Mixedgrass Prairie
251C, 331C, 332B, 332D, 332E

Stipa Comata - Bouteloua Gracilis Herbaceous Alliance

Needle-and-thread - blue grama - threadleaf sedge mixedgrass prairie.

Needleandthread-Grama Mixedgrass Prairie
331C, 331F, 331H, 332C, 332D

Lowland

Pascopyrum Smithii Intermittently Flooded Herbaceous Alliance

Western wheatgrass - buffalograss.

Polygonum Spp. - Echinochloa Spp. Temporarily Flooded Herbaceous Alliance

Smartweed playa lake.

Playa
331C, 332E

Shortgrass Prairie

Upland

Bouteloua Gracilis Herbaceous Alliance

Blue grama buffalograss shortgrass prairie.

Blue Grama Shortgrass Prairie
331C, 331F, 331H

Lowland

Distichlis Spicata - (Hordeum Jubatum) Temporarily Flooded Herbaceous Alliance

Inland saltgrass - foxtail barley - plains bluegrass - alkali sacaton saline playa.

Inland saltgrass - foxtail barley - plains bluegrass - marsh elder saline prairie.

Inland saltgrass - bulrush - saltwort saline marsh.

Saltgrass-Foxtail Saline Meadows
251C, 331H, 332C

HYDROMORPHIC ROOTED VEGETATION

Potamogeton Spp. - Ceratophyllum Spp. - Elodea Spp. Mixed Permanently Flooded Herbaceous Vegetation

Pondweed - hornwort submerged wetland.

Pondweed-Hornwort Submerged Wetland
All

SPARSE VEGETATION Vegetation typically < 10%

Open Bluff/Cliff

Inland sandstone dry bluff - cliff.
Inland alkaline zeric bluff - cliff.

Rock Outcrop/Butte

Ponderosa pine limestone cliff.
Siltstone - clay butte rock outcrop.

Large Eroding Cliffs

Eroding great plains badlands.

Barren Cliffs and Outcrops
331F, 331H

Sand Flat/Beach

Riverine sand flats - bars.

Barren Sand Flats
All

AGRICULTURAL LANDS

- Cropland
  - Grain Crops
    - Irrigated Cropland
    - Dry Cropland
  - Forage Crops
    - Irrigated Forage
    - Dry Forage (including CRP, pasture and planted brome)

Agriculture - Grains and Forage

Woodland/Forest
Coniferous Woodland or Forest
Grove
Shelterbelt

Agriculture - Coniferous Woodland

Deciduous Woodland or Forest
Orchard
Vineyard
Grove
Shelterbelt

Agriculture - Deciduous Woodland

Mixed Woodland or Forest
Grove
Shelterbelt

Agriculture - Mixed Woodland

<u>URBAN LANDS</u>
Commercial

Urban - Commercial

Residential

Urban - Residential

Recreational

Urban - Recreational

<u>OPEN WATER</u>
Lacustrine
Freshwater Lakes

Alkaline Lakes

Riverine

Streams

Rivers

Canals

Open Water

Analysis of Land Cover Map Accuracy in Northern Arizona

Final Report to the USGS National Gap Analysis Program

US Department of the Interior

Submitted by:

USGS Biological Resources Division

Forest and Rangeland Ecosystem Science Center

Colorado Plateau Field Station

Northern Arizona University

P.O. Box 5614

Flagstaff, Arizona 86011-5614

Dr. Kathryn A. Thomas

Vegetation Ecologist

520-556-7466 ext. 235

Kathryn_A_Thomas@usgs.gov

Sarah R. Jacobs

Research Specialist

520-556-7466 ext. 240

Sarah.Jacobs@nau.edu

May 18, 2000

List of Figures

List of Tables

Executive summary

The Gap Analysis Program (GAP) provides regional evaluation of the conservation status of indigenous natural land cover types and facilitates the application of this information toward land management activities. Gap analysis products have now been brought to completion in eight states, with the rest of the states in active development. The development of regional land cover maps at an unprecedented scale and resolution for such large areas has made a compelling need for statistically based accuracy assessment of these maps. Examination of map accuracy is an important analysis of any mapping project. Map accuracy issues have come under increasing scrutiny by National GAP (Stoms et al. 1994; Crist 1996). Map accuracy is also of importance in informing the process of map updates.

The Arizona Gap Analysis Program (AZ GAP) was begun in 1991 under the direction of Dr. Lee Graham, University of Arizona. Dr. Graham left the program in 1994 after developing the preliminary land cover map and vertebrate distribution models. However, these products were neither accuracy assessed, nor was a gap analysis conservation evaluation conducted. Two USGS Biological Resources Division field stations, the Forest and Rangeland Ecosystem Science Center, Colorado Plateau Field Station (CPFS) in Flagstaff, Arizona and the Western Ecological Research Center, Sonoran Desert Field Station (SDFS) in Tucson, Arizona, were assigned the task of completing an accuracy assessment for the land cover and vertebrate distribution maps. The SDFS was also assigned responsibility for completion of the first AZ GAP final report.

The CPFS assessed the accuracy of the land cover map in northern Arizona and the vertebrate distributions for the entire state. The findings for these accuracy assessments have been submitted to the SDFS for inclusion in the AZ GAP final report. The final report is due in draft form to National GAP; the date to be determined by the SDFS.

The CPFS was authorized (Research Work Order 1445-0009-94-1069 SA#4) by National GAP in March 1998 to conduct additional map accuracy studies for the AZ GAP land cover map for northern Arizona. Work was executed by the project leader, Dr. Kathryn Thomas, and a graduate student at Northern Arizona University, Sarah Jacobs. The study, entitled "Spatial Analysis of Vegetation Accuracy in Northern Arizona", had three objectives:

1. Develop a digital map of accuracy estimates in northern Arizona,

2. Correlate the accuracy estimates with land surface factors that may be related to the pattern of accuracy,
3. Develop recommendations for activities for a gap analysis update in Arizona.

Two deliverables were identified: 1) a report to National GAP and 2) a presentation at the Annual National GAP Meeting describing the findings of the study at the time. This document constitutes the first deliverable in four sections:

1. Results of binary and fuzzy set accuracy assessments for northern Arizona,
2. Development of a spatial view of accuracy using fuzzy set ratings from reference data,
3. Assessment of environmental factors correlated with accuracy assessment, and
4. Recommendations for future work.

A thesis written by Sarah Jacobs, *A Fuzzy Spatial View of Accuracy of a Northern Arizona Land Cover Map* (1999) is also included as a project deliverable.

The second deliverable was achieved in the summer of 1998 at the 8th Annual National GAP Meeting held in Santa Barbara, California. Sarah Jacobs presented a co-authored poster (Sarah R. Jacobs and Kathryn A. Thomas) entitled "Mapping Land Cover Accuracy: A Spatial Assessment of the Northern Arizona GAP Land Cover Map." In addition to these deliverables, a journal paper is being developed that compares the binary and fuzzy set accuracy assessment methods of a land cover map.

Two attributes of a map require accuracy assessment: spatial (cartographic and positional) accuracy and thematic (classification) accuracy. In this report, we examine the thematic accuracy of the northern half of the AZ GAP land cover map using both a binary assessment method, and a method that allows multi-category accuracy ratings based on fuzzy set theory, the fuzzy set assessment method. The fuzzy set assessment proved more valuable than the binary assessment because it provides more information about the type of agreement between reference and map data. The fuzzy set ratings are used to create a spatial view of (thematic) accuracy. The spatial view of accuracy improves upon the fuzzy set assessment by displaying accuracy as it varies across the landscape. The spatial view of accuracy for a national forest illustrates how such information can be used. Factors potentially related to the spatial distribution of the fuzzy set map of accuracy are explored. Unfortunately, none of the factors examined are found to contribute to thematic accuracy of the land cover map. Finally, the spatial view of accuracy is used to make recommendations for improvements on GAP land cover mapping. AZ GAP is participating in a regional GAP among the southwest states (SW ReGAP), which was initiated in the summer of 1999. Insight from the previous project can potentially guide implementation of second generation gap analysis products.

Section 1: BINARY AND FUZZY SET ACCURACY ASSESSMENTS

Introduction

Thematic accuracy assessment of meso-scale (1:100,000 to 1:500,000) land cover maps has traditionally been limited to a binary, right/wrong analysis (Congalton and Green 1999; Congalton 1996). The binary assessment provides information about agreement between land cover types as mapped (classified data) and land cover types as determined by a data source independent of the map (reference data). The reference data are considered to be the true value of land cover.

An alternative method of thematic accuracy assessment uses fuzzy set theory (Zadeh 1965). Fuzzy set analysis was first proposed as an alternative to the binary accuracy assessment approach in Gopal and Woodcock (1994). A fuzzy set analysis provides more data about the *degree* of agreement between the reference and mapped land cover types. Instead of a right/wrong analysis, map labels are considered partially right or partially wrong, generally on a five-category scale. This is more useful for assessing vegetation types that may grade into one another, yet must be classified into discrete types by a human observer (Gopal and Woodcock 1994). The fuzzy set analysis provides a number of measures with which to judge the accuracy of the land cover map.

A simple binary accuracy assessment may be inferior to a fuzzy set accuracy assessment because of the continuous nature of vegetation distribution across the landscape. Distinct boundaries between land cover types seldom exist in nature. Instead, there are often gradations from one vegetation type into another. Confusion results when a reference point can be labeled as more than one land cover type, for example in ecotonal areas or transition zones. Fuzzy set theory aids in the assessment of maps produced from remotely sensed data by analyzing and quantifying vague, indistinct, or overlapping class memberships (Gopal and Woodcock 1994).

The fuzzy set analysis provides insight into the types of errors that are being made. For example, classification of Ponderosa Pine Woodland to Juniper Woodland may be a more acceptable error than classification of the same Ponderosa Pine Forest to a Desert Shrubland. In the first instance, the misclassification may not be important if the map user wishes to know where all coniferous woodlands exist in an area.

In this section of the report, accuracy assessments of the preliminary AZ GAP land cover map are determined by both the binary and fuzzy set methods. Overall accuracy is compared using both methods.

Background

Accuracy Assessment Metrics

The binary assessment is summarized in an error matrix (Congalton and Green 1999). The error matrix is also called a confusion or contingency table. In the matrix, the land cover type predicted by the classified data (map) is assigned to rows and the observed land cover type (reference data) is displayed in columns. The values in each cell represent the count of sample points matching the combination of classified and reference data (Congalton 1996). Errors of inclusion (commission errors, where a type is predicted where it should not be) and errors of exclusion (omission errors, where the map failed to predict a type) for each land cover type, and overall map accuracy can be calculated using the error matrix (Congalton 1996). User's and producer's accuracy are also calculated from the error matrix. User's accuracy corresponds to the area on the map that actually represents that land cover type on the ground. Producer's accuracy represents the percentage of sampling points that were correctly classified for each land cover type.

Accuracy measures for each land cover type produced from fuzzy set data include the Max, Right, and Increase statistics. The Max statistic calculates the same information as user's accuracy because it includes only absolutely right answers. The Right statistic is more inclusive than the Max statistic. For this study, the Right statistic calculates the accuracy of the land cover map to the life form level or better. The Increase statistic reflects the improvement in accuracy associated with using the Right statistic instead of the Max statistic (Gopal and Woodcock 1994). Other statistics that can be calculated in a fuzzy set analysis are Membership, Difference, Ambiguity, and Confusion, which are not used in this analysis.

Land Cover Map

The accuracy assessment was conducted on the *preliminary* AZ GAP land cover map (1994) developed by Dr. Lee Graham, formerly of the University of Arizona in Tucson. Unfortunately, Graham left AZ GAP abruptly without documenting the methodology of creating the land cover map. The following description was put together from verbal explanations provided by other members of the original AZ GAP team.

The land cover map (Figure 1-1) was produced primarily from Landsat Thematic Mapper satellite imagery from the early 1990s. An unsupervised classification of the imagery produced clustered pixels based on similarity in spectral properties. Aerial video and ground measurements facilitated labeling and refinement of the clusters into 105 land cover types. Natural vegetation types were identified using a modification of the Brown, Lowe, and Pase (BLP) classification system (1979); however, the Graham modification of the BLP system was poorly described and documented.

The land cover map consists of polygons labeled with cover types contained in a geographic information system (GIS, Arc/INFO format) with a 40-hectare minimum mapping unit (smaller in some riparian areas). This resolution is best suited for interpretation at the 1:100,000 scale.

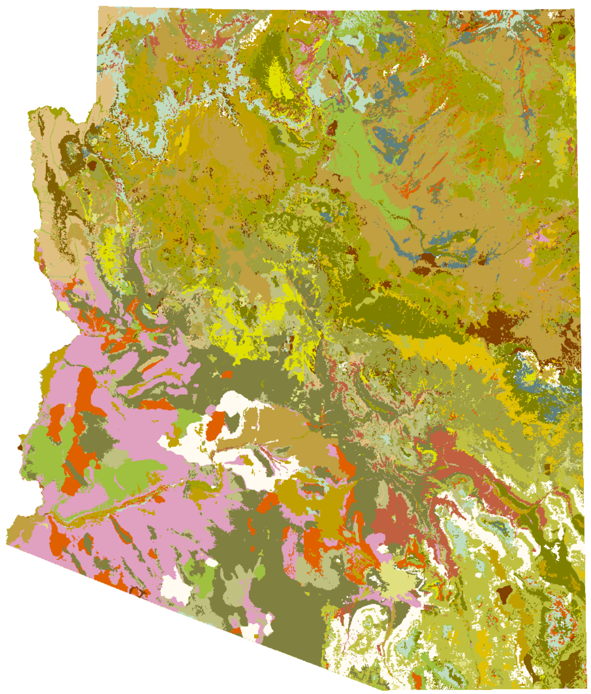


Figure 2-1. Preliminary AZ GAP land cover map. The preliminary map contains 58,170 polygons describing 105 vegetation types. A key is not provided for this illustration.

Reference Data

A stratified random sampling design, with stratification occurring on land cover type, was used to determine the set of polygons to be sampled in the accuracy assessment. A binomial approximation calculation determined the number of target sampling sites for each land cover type. The approximation was calculated with a target accuracy of 80% for the map and a standard error of 0.05. To avoid boundary effects, polygon centroids (or closest approximation) were targeted for sampling (Arundel, pers. comm. 1997). The target sampling sites and their locations were divided between northern and southern Arizona with 1,525 points assigned in northern Arizona. This report addresses the accuracy assessment for the northern portion of the map only.

Methods

Reference Data

A total of 930 sampling sites representing 59 different land cover types in northern Arizona were visited during the summer of 1997. Field technicians located the sampling site in the field by navigating with a Global Positioning System (GPS) to the pre-determined sampling location. They identified dominant, co-dominant, and associate plant species for a one hectare area. In some cases, the target location could not be accessed. In these circumstances either an alternate site was substituted (if known before entering the field), or observations were made at an adjacent area. The field data at each site was assigned to one of the 105 land cover classes by the project plant ecologist using the incomplete definitions provided by Graham (Thomas pers. comm. 1998). The resulting dataset, the reference data, consists of 930 reference points with a field assigned land cover type and associated location (gathered by GPS).

Classified Data

The classified data were obtained from the preliminary AZ GAP land cover map. The UTM locations for each reference point were used to determine the mapped class at the same location.

Binary Accuracy Assessment

The labeling at each site was compared between the reference data and the map. Matches between the two were coded as either 1 for correct or 0 for incorrect. These statistics were incorporated into an error matrix from which user's and producer's accuracy for each land cover type were calculated, as well as overall accuracy of the land cover map. User's accuracy was calculated as:

$$\frac{n_{ii}}{n_{i+}}$$

where

i = land cover type,

n_{ii} = number of matches between map and reference data,

n_{i+} = total number of samples of *i* in the map

Producer's accuracy was calculated as:

$$\frac{n_{ii}}{n_{+i}}$$

where

n_{+i} = total number of sample of *i* in the reference data

Overall accuracy for the map was calculated as:

$$\frac{\sum_{i=1}^{k} n_{ii}}{n}$$

where

k = number of land cover types,

n = total number of reference points

The standard error was estimated for each metric using the formula:

$$\sqrt{(1-f)\frac{p(1-p)}{n-1}}$$

where

p = accuracy (0 to 1.0),

$1-f$ = final population correction factor; proportion of the population not sampled,

n = number of sites sampled

(Cochran 1977).

Fuzzy set accuracy assessment

The plant ecologist developed an accuracy rank matrix that rated each of the 105 land cover classes against each other and assigned an accuracy rank (Table 1-1) to each combination. The accuracy rank matrix was used to assign an accuracy rank to each sample point by comparing its reference data assignment to its classified data assignment within the matrix. Using these ratings for each sample point, the following functions were defined:

Max (M) = number of sites with an absolutely right answer (accuracy rank of 5),

Right (R) =	number of sites with a reasonable, good, or absolutely right answer (accuracy ranks of 3, 4, and 5),
Increase (R —M) =	difference between the Right and Max functions

Membership, Difference, Ambiguity, and Confusion statistics were not calculated because the fuzzy set ratings were predefined rather than assessed at each sampling site. Because of the predefined fuzzy set matrix, a fair amount of information is lost, but a large amount of time in the field is saved.

Table 1-1. Fuzzy set accuracy ranks.

Rank	Answer	Description
1	Wrong	The reference and mapped types do not correspond, and there is no ecological reason for the non-correspondence.
2	Understandable but Wrong	The reference and mapped types do not correspond, but the reason for non-correspondence can be ecologically understood.
3	Reasonable or Acceptable	The reference and mapped types are of the same life form (i.e., formation types).
4	Good	The reference and mapped types are characterized by the same species at the dominant species level.
5	Absolutely Right	The reference and mapped types are exactly the same.

Results

Reference data

The target number of sampling sites was not achieved (930 out of 1,525, 61%) for many of the land cover types due to logistical and financial constraints on the project. In addition, some cover types occur mainly in the southern Arizona data collection area. To fully assess the accuracy of each land cover type for the entire map, the northern Arizona reference data should be combined with the southern Arizona reference data.

Binary Accuracy Assessment

User’s and producer’s accuracy for each cover type, and overall accuracy were low (Table 1-2). The highest producer’s accuracy was for anthropogenic defined cover types (Industrial, 60% and Mixed Agriculture/Urban/Industrial, 80%). Producer’s accuracy for natural cover types ranged between 0 and 50%, indicating that up to 50% of the Encinal Mixed Oak/Mixed Chaparral/Semidesert Grassland-Mixed Scrub and Mohave Blackbush-Yucca Scrub sampling points were correctly classified. Likewise, the highest user’s accuracy was also for anthropogenic defined cover types (Urban, 91% and Industrial, 86%); while natural cover types ranged between 0 and 48%, indicating that up to 48% of the Engelmann Spruce-Mixed Conifer was actually that type on the ground. The standard error was greater than 5% for almost all sampled vegetation types. Overall, map accuracy was 14.8%.

Table 1-2. User's and producer's accuracy using the binary data.

Code	Land Cover Type	# Sites	Producer's Accuracy (%)	Standard Error	User's Accuracy (%)	Standard Error
3	Engelmann Spruce-Mixed Conifer	29	41.2	7.0	48.3	7.2
4	Rocky Mountain Lichen-Moss	1	0.0	0.0	0.0	0.0
5	Rocky Mountain Bristlecone-Limber Pine	2	0.0	0.0	0.0	0.0
6	Pinyon-Juniper-Shrub/Ponderosa Pine-Gambel Oak-Juniper	21	0.0	0.0	0.0	0.0
7	Pinyon-Juniper/Sagebrush/Mixed-Grass-Scrub	34	18.2	6.5	11.8	5.5
8	Pinyon-Juniper-Shrub Live Oak-Mixed Scrub	13	8.0	7.3	15.4	9.6
9	Pinyon-Juniper (Mixed)/Chaparral-Scrub	33	8.3	4.6	3.0	2.9
10	Pinyon-Juniper-Mixed Shrub	18	0.0	0.0	0.0	0.0
11	Pinyon-Juniper-Mixed Grass-Scrub	34	5.3	3.7	2.9	2.9
12	Pinyon-Juniper (Mixed)	41	6.7	3.9	2.4	2.1
13	Douglas Fir-Mixed Conifer	35	38.5	7.2	28.6	6.7
14	Arizona Cypress	8	25.0	12.8	12.5	9.9
15	Ponderosa Pine	45	12.5	4.8	13.3	4.8
16	Ponderosa Pine-Mixed Conifer	23	11.5	5.4	13.0	5.6
17	Ponderosa Pine-Gambell Oak-Juniper/Pinyon-Juniper Complex	36	11.8	5.1	16.7	5.9
18	Ponderosa Pine-Pinyon-Juniper	39	16.7	5.8	5.1	3.3
20	Ponderosa Pine-Mixed Oak-Juniper	3	10.0	18.2	33.3	28.6
21	Encinal Mixed Oak	1	0.0	0.0	0.0	0.0
22	Encinal Mixed Oak-Pinyon-Juniper	5	16.7	18.1	40.0	23.7
23	Encinal Mixed Oak-Mexican Pine-Juniper	2	0.0	0.0	0.0	0.0
24	Encinal Mixed Oak-Mexican Mixed Pine	1	0.0	0.0	0.0	0.0
25	Encinal Mixed Oak-Mesquite	1	0.0	0.0	0.0	0.0
26	Encinal Mixed Oak/Mixed Chaparral/Semidesert Grassland-Mixed Scrub	10	50.0	15.0	10.0	9.0
27	Great Basin Juniper	2	0.0	0.0	0.0	0.0
28	Interior Chaparral Shrub Live Oak-Pointleaf Manzanita	14	20.0	10.7	35.7	12.9
29	Interior Chaparral Mixed Evergreen Schlerophyll	18	33.3	11.0	27.8	10.5
30	Interior Chaparral (Mixed)/Sonoran-Paloverde-Mixed Cacti	1	0.0	0.0	0.0	0.0
31	Interior Chaparral (Mixed)/Mixed Grass-Mixed Scrub Complex	10	0.0	0.0	0.0	0.0
32	Rocky Mountain/Great Basin Dry Meadow	18	20.0	6.4	27.8	7.2
33	Madrean Dry Meadow	22	0.0	0.0	0.0	0.0
34	Great Basin (or Plains) Mixed Grass	20	9.5	6.1	10.0	6.1

Table 1-2 continued. User's and producer's accuracy using the binary data.

Code	Land Cover Type	# Sites	Producer's Accuracy (%)	Standard Error	User's Accuracy (%)	Standard Error
35	Great Basin (or Plains) Mixed Grass-Mixed Scrub	40	8.5	4.2	15.0	5.5
36	Great Basin (or Plains) Mixed Grass-Sagebrush	4	11.1	16.4	25.0	22.7
37	Great Basin (or Plains) Mixed Grass-Saltbush	24	35.7	8.1	20.8	6.9
38	Great Basin (or Plains) Mixed Grass-Mormon Tea	20	11.1	6.8	5.0	4.8
42	Semidesert Mixed Grass-Mixed Scrub	2	0.0	0.0	0.0	0.0
43	Great Basin Sagebrush Scrub	12	0.0	0.0	0.0	0.0
44	Great Basin Big Sagebrush-Juniper-Pinyon	30	20.0	6.5	13.3	5.5
45	Great Basin Sagebrush-Mixed Grass-Mixed Scrub	27	20.0	7.0	22.2	7.3
46	Great Basin Shadscale-Mixed Grass-Mixed Scrub	24	0.0	0.0	0.0	0.0
47	Great Basin Greasewood Scrub	11	37.5	14.8	27.3	13.6
48	Great Basin Saltbush Scrub	7	6.7	9.5	14.3	12.9
49	Great Basin Blackbrush-Mixed Scrub	36	16.0	5.8	11.1	5.0
50	Great Basin Mormon Tea-Mixed Scrub	18	19.4	9.1	33.3	10.9
51	Great Basin Winterfat-Mixed Scrub	11	0.0	0.0	0.0	0.0
52	Great Basin Mixed Scrub	26	9.1	5.6	11.5	6.3
53	Great Basin Mormon Tea/Pinyon-Juniper	16	0.0	0.0	0.0	0.0
55	Mohave Creosotebush-Bursage Mixed Scrub	7	28.6	17.9	28.6	17.9
58	Mohave Blackbush-Yucca Scrub	13	50.0	10.4	23.1	8.7
59	Mohave Saltbush Yucca Scrub	5	0.0	0.0	0.0	0.0
61	Mohave Creosotebush-Brittlebush Mohave Globemallow Scrub	5	0.0	0.0	0.0	0.0
63	Mohave Joshua Tree	1	0.0	0.0	0.0	0.0
64	Mohave Mixed Scrub	9	9.1	9.8	11.1	10.7
75	Sonoran Paloverde-Mixed Cacti-Mixed Scrub	1	0.0	0.0	0.0	0.0
82	Agriculture	1	0.0	0.0	0.0	0.0
83	Urban	11	41.7	14.9	90.9	8.6
84	Industrial	7	60.0	18.9	85.7	13.4
85	Mixed Agriculture/Urban/Industrial	20	80.0	7.7	20.0	7.7
87	Water	2	0.0	0.0	0.0	0.0
	Overall Accuracy = 14.8%					

Fuzzy Set Accuracy Assessment

The Max statistic for the fuzzy set reference data yields the same information as user's accuracy for the binary accuracy assessment (Table 1-3). However, the Right function provides a different view. Accuracy improves across the table for all land cover types because of the Right statistic being more inclusive than the Max statistic. For example, in land cover class 18, Ponderosa Pine-Pinyon-Juniper, the Max statistic indicates this type has very low accuracy, 5%. The Right statistic shows that when assessed at the life-form level this type is correct 74% of the time. The range for Right statistics is large, between 0 and 100%. However, the land cover types are more often correct to the life form (mean 52.7%, standard deviation 33.4%) compared to the Max statistic (mean 13.8%, standard deviation 18.8%). The mean increase in accuracy when viewed at the life form level is 38.8% (standard deviation 31.5%).

Discussion

A binary accuracy assessment may be conservatively biased against a classification system that is poorly defined and numerous in classes (Verbyla 1995). The lack of descriptions in the Graham classification system made labeling the cover type of each reference point difficult. In addition, division of the land cover types of Arizona into 105 classes made distinguishing between types problematic. Therefore, the binary assessment likely assigned a wrong answer to reference locations that may have been partially right.

A traditional binary accuracy assessment using an error matrix provides limited information about the accuracy of a thematic map. In fact, an overall accuracy of 14.8% for the land cover map is dismal and discourages use of the map for any application. However, this is not unexpected given the preliminary nature of the map, high number of land cover types, small sample size compared to number of cover types, and lack of documentation of the Graham vegetation types.

A fuzzy set assessment provides more information about the agreement between the reference data and the map. The Max statistics are disturbing, but are less so when the Right statistics are considered. The Right function indicates that many land cover types are more accurately classified to the life form level. Yet, even for this statistic, accuracy was not the target 80% in most instances. This added information allows the user and producer to judge the value of the land cover map for different applications. For example, for certain cover types, the map will perform adequately to the life form level, and may be used in applications where this determination is all that is required.

Table 1-3. Max-Right classification results using the fuzzy set.

Code	Land Cover Type	# Sites	Max (M) Best #	Max (M) Best %	Right (R) Correct #	Right (R) Correct %	Increase (R - M) #	Increase (R - M) %
3	Engelmann Spruce-Mixed Conifer	29	14	48.3	25	86.2	11	37.9
4	Rocky Mountain Lichen-Moss	1	0	0.0	0	0.0	0	0.0
5	Rocky Mountain Bristlecone-Limber Pine	2	0	0.0	2	100.0	2	100.0
6	Pinyon-Juniper-Shrub/Ponderosa Pine-Gambel Oak-Juniper	21	0	0.0	10	47.6	10	47.6
7	Pinyon-Juniper/Sagebrush/Mixed-Grass-Scrub	34	4	11.8	19	55.9	15	44.1
8	Pinyon-Juniper-Shrub Live Oak-Mixed Scrub	13	2	15.4	11	84.6	9	69.2
9	Pinyon-Juniper (Mixed)/Chaparral-Scrub	33	1	3.0	12	36.4	11	33.3
10	Pinyon-Juniper-Mixed Shrub	18	0	0.0	7	38.9	7	38.9
11	Pinyon-Juniper-Mixed Grass-Scrub	34	1	2.9	18	52.9	17	50.0
12	Pinyon-Juniper (Mixed)	41	1	2.4	21	51.2	20	48.8
13	Douglas Fir-Mixed Conifer	35	10	28.6	28	80.0	18	51.4
14	Arizona Cypress	8	1	12.5	1	12.5	0	0.0
15	Ponderosa Pine	45	6	13.3	28	62.2	22	48.9
16	Ponderosa Pine-Mixed Conifer	23	3	13.0	16	69.6	13	56.5
17	Ponderosa Pine-Gambell Oak-Juniper/Pinyon-Juniper Complex	36	6	16.7	20	55.6	14	38.9
18	Ponderosa Pine-Pinyon-Juniper	39	2	5.1	29	74.4	27	69.2
20	Ponderosa Pine-Mixed Oak-Juniper	3	1	33.3	2	66.7	1	33.3
21	Encinal Mixed Oak	1	0	0.0	0	0.0	0	0.0
22	Encinal Mixed Oak-Pinyon-Juniper	5	2	40.0	3	60.0	1	20.0
23	Encinal Mixed Oak-Mexican Pine-Juniper	2	0	0.0	0	0.0	0	0.0
24	Encinal Mixed Oak-Mexican Mixed Pine	1	0	0.0	0	0.0	0	0.0
25	Encinal Mixed Oak-Mesquite	1	0	0.0	1	100.0	1	100.0
26	Encinal Mixed Oak/Mixed Chaparral/Semidesert Grassland-Mixed Scrub	10	1	10.0	2	20.0	1	10.0
27	Great Basin Juniper	2	0	0.0	0	0.0	0	0.0
28	Interior Chaparral Shrub Live Oak-Pointleaf Manzanita	14	5	35.7	6	42.9	1	7.1
29	Interior Chaparral Mixed Evergreen Schlerophyll	18	5	27.8	7	38.9	2	11.1
30	Interior Chaparral (Mixed)/Sonoran-Paloverde-Mixed Cacti	1	0	0.0	1	100.0	1	100.0
31	Interior Chaparral (Mixed)/Mixed Grass-Mixed Scrub Complex	10	0	0.0	0	0.0	0	0.0
32	Rocky Mountain/Great Basin Dry Meadow	18	5	27.8	5	27.8	0	0.0
33	Madrean Dry Meadow	22	0	0.0	6	27.3	6	27.3
34	Great Basin (or Plains) Mixed Grass	20	2	10.0	3	15.0	1	5.0
35	Great Basin (or Plains) Mixed Grass-Mixed Scrub	40	6	15.0	15	37.5	9	22.5

			Max (M) Best			Right (R) Correct			Increase (R - M)	
Code	Land Cover Type	# Sites	#	%		#	%		#	%
36	Great Basin (or Plains) Mixed Grass-Sagebrush	4	1	25.0		2	50.0		1	25.0
37	Great Basin (or Plains) Mixed Grass-Saltbush	24	5	20.8		10	41.7		5	20.8
38	Great Basin (or Plains) Mixed Grass-Mormon Tea	20	1	5.0		9	45.0		8	40.0
42	Semidesert Mixed Grass-Mixed Scrub	2	0	0.0		0	0.0		0	0.0
43	Great Basin Sagebrush Scrub	12	0	0.0		7	58.3		7	58.3
44	Great Basin Big Sagebrush-Juniper-Pinyon	30	4	13.3		13	43.3		9	30.0
45	Great Basin Sagebrush-Mixed Grass-Mixed Scrub	27	6	22.2		17	63.0		11	40.7
46	Great Basin Shadscale-Mixed Grass-Mixed Scrub	24	0	0.0		13	54.2		13	54.2
47	Great Basin Greasewood Scrub	11	3	27.3		10	90.9		7	63.6
48	Great Basin Saltbush Scrub	7	1	14.3		4	57.1		3	42.9
49	Great Basin Blackbrush-Mixed Scrub	36	4	11.1		24	66.7		20	55.6
50	Great Basin Mormon Tea-Mixed Scrub	18	6	33.3		9	50.0		3	16.7
51	Great Basin Winterfat-Mixed Scrub	11	0	0.0		5	45.5		5	45.5
52	Great Basin Mixed Scrub	26	3	11.5		18	69.2		15	57.7
53	Great Basin Mormon Tea/Pinyon-Juniper	16	0	0.0		10	62.5		10	62.5
55	Mohave Creosotebush-Bursage Mixed Scrub	7	2	28.6		6	85.7		4	57.1
58	Mohave Blackbush-Yucca Scrub	13	3	23.1		11	84.6		8	61.5
59	Mohave Saltbush Yucca Scrub	5	0	0.0		5	100.0		5	100.0
61	Mohave Creosotebush-Brittlebush Mohave Globemallow Scrub	5	0	0.0		5	100.0		5	100.0
63	Mohave Joshua Tree	1	0	0.0		1	100.0		1	100.0
64	Mohave Mixed Scrub	9	1	11.1		9	100.0		8	88.9
75	Sonoran Paloverde-Mixed Cacti-Mixed Scrub	1	0	0.0		0	0.0		0	0.0
82	Agriculture	1	0	0.0		0	0.0		0	0.0
83	Urban	11	10	90.9		11	100.0		1	9.1
84	Industrial	7	6	85.7		7	100.0		1	14.3
85	Mixed Agriculture/Urban/Industrial	20	4	20.0		19	95.0		15	75.0
87	Water	2	0	0.0		0	0.0		0	0.0
	Sums and overall accuracies	930	138	14.8		523	56.2		385	41.4

Section 2: A SPATIAL VIEW OF ACCURACY USING FUZZY SET RANKS

Introduction

Reference data inherently contain spatial information that is usually ignored in both binary and fuzzy set accuracy assessments. For both assessments, spatial information associated with the reference data are not utilized in the summary statistics, and results are given in tabular rather than spatial format. "The most fundamental drawback of the confusion matrix is its inability to provide information on the spatial structure of the uncertainty in a classified scene," (Canters 1997).

Literature on the spatial variability of thematic map accuracy is limited. Congalton (1988) proposed the simplest method of displaying accuracy by producing a difference image of ones and zeros to represent agreement or disagreement between the classified and reference images. Fisher proposed a dynamic portrayal of a variety of accuracy measures (1994). Steele et al. (1998) developed a map of accuracy illustrating the magnitude and distribution of classification errors for a land cover map. They used kriging to interpolate misclassification estimates (produced from a bootstrapping method) at each reference point. The interpolated estimates were then used to construct a contour map showing accuracy estimates over the map extent. This latter work provided a starting point for this project.

Thematic map accuracy, however, may vary spatially across a landscape in a manner partially or totally unrelated to land cover type. In other words, a land cover type may be misclassified more often when it occurs in certain contexts, such as on steep slopes, than others. Or, land cover types that are located near ground control data used in the map development may be more correct than remote areas for which only imagery was used to develop the map.

Using spatial interpolation, a map of thematic or classification accuracy can be produced from the reference data. Spatial interpolation, also called estimation or prediction, refers to making inferences at unobserved locations using data from known sources. There are several types of interpolation procedures, including inverse distance weighting, least-squares estimators, and Gauss-Markov estimators. This last group includes kriging, which is generally accepted as a robust and accurate estimator (Chou 1997, Rouhani 1986).

In this section, the methodology to develop a map of thematic accuracy by kriging is described. The technique is applied to the fuzzy set reference data for northern Arizona to produce a spatial view of accuracy. The UTM coordinates are used as the *x* and *y* variables, and the fuzzy set rank is used as the *z* variable. The application of such a map for land managers is illustrated in the discussion.

Background

Kriging estimates a surface from a scattered set of points. Synonymous with optimal prediction in space, kriging refers to making inferences at unobserved locations from data collected at random locations. Kriging exploits the spatial dependence among the variable to be predicted.

The fuzzy set reference data must be explored and modeled prior to kriging. The spatial autocorrelation of the fuzzy set reference data must be calculated (semivariance) and plotted to understand the extent to which neighboring accuracy ranks influence each other. Positive spatial autocorrelation indicates that similar values are clustered together in space, and is usually evident in geostatistical data, or natural processes (Odland 1988). Then, a model fit to the semivariance provides direction in kriging accuracy ranks across northern Arizona.

Semivariance describes how accuracy ranks are related with distance. Semivariance, g , is defined as half the average squared difference in *z* values between points separated by a distance *h* (Matheron 1963). Specifically, the semivariance is calculated as:

$$\gamma(h) = \frac{1}{2|N(h)|} \sum_{N(h)} (z_i - z_j)^2$$

where

g*(h)* = sample semivariance,

h = distance measure with magnitude only,

N(h) = set of all pairwise Euclidean distances $d(x_i, x_j) = h$,

|*N(h)*| = number of distinct pairs in *N(h)*,

z_i *and* z_j = accuracy ranks at spatial locations *i* and *j*

(Kaluzny et al. 1996). In this study, accuracy rank is a qualitative estimate, and does not have any units. Therefore, there are no units for g*(h)*.

A robust estimator of the semivariance reduces the effect of outliers in the data without removing specific data points (Cressie and Hawkins 1980). The estimation is based on the fourth power of the square root of the absolute differences in *z* values.

$$\hat{\gamma}(h) = \frac{\left[\frac{1}{2|N(h)|} \sum_{N(h)} \sqrt{|z_i - z_j|} \right]^4}{0.457 + \frac{0.494}{|N(h)|}}$$

where

$\hat{\gamma}(h)$ = robust estimator of the semivariance,

(Kaluzny et al. 1996).

A semivariogram is constructed from the semivariance values. The semivariogram relates each data point to all other data points with respect to the lag distance between them (Clark 1979). Semivariance are shown on the y-axis, and lag distance on the x-axis (Figure 2-1). The plot shows the way in which spatial dependence varies with spatial separation. Typically, data of natural processes are positively correlated at small distance intervals, becoming uncorrelated at larger distances.

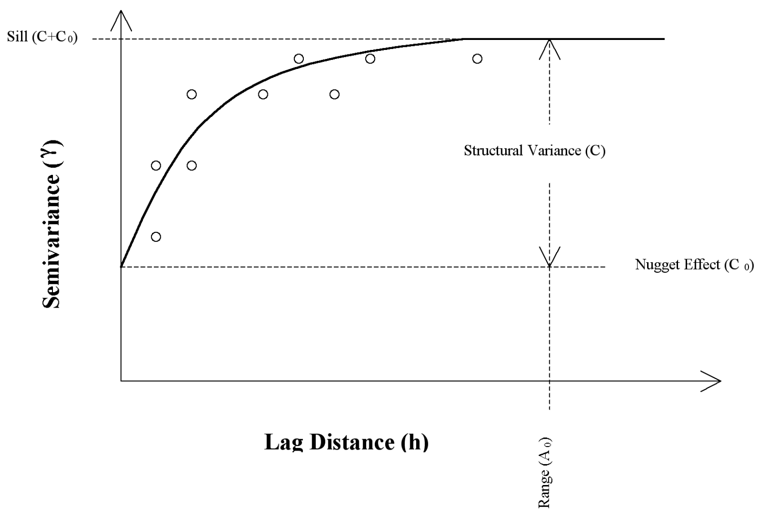


Figure 2-1. Semivariogram, including semivariance (circles) and generic model (heavy line).

A model is fit to the semivariance values to direct the kriging process. Three parameters of the model are usually recognized and interpreted for geostatistical data: the nugget effect (C_0), the sill ($C+C_0$), and the range (A_0) (Figure 2-1). In theory, the semivariance is zero at a lag distance of zero (a value measured at exactly the same location should be the same value). However, several factors may cause sample values separated by extremely small distances to be quite dissimilar. The nugget effect describes the discontinuity at the origin. The sill is the point of maximum variance, and is the sum of the structural variance (semivariance attributed purely to the process) and the nugget effect. It is the plateau that the semivariogram reaches at the range. The range, or range of influence, defines the smallest distance at which the variable becomes uncorrelated (Royle 1980).

Possible semivariogram models include exponential, spherical, gaussian, linear and power (Kaluzny et al. 1996). The linear and power models increase without bounds, and are not typical of geostatistical data because they imply that the variation between data points continually increases with distance. However, most geostatistical data hit a sill of maximum variance at a certain range. The fit of the model to the semivariogram values is determined by the residual sum of squares (RSS). The values of the three main parameters are changed iteratively to reduce the RSS value and fit the model.

Kriging is performed using the fuzzy set reference data, semivariogram model, associated parameters, and spatial locations for prediction. The result is a lattice of points with predicted *z* values.

Methods

Using S-PLUS and S+SPATIALSTATS, the spatial autocorrelation of the fuzzy set reference data were modeled by conducting a semivariance analysis on the accuracy rank (Table 1-1) at each reference point. The semivariogram was computed using the robust estimator for semivariance. Accuracy rank was assigned to the *z* variable and modeled as a function of the easting and northing UTM coordinates gathered by GPS at each reference site.

The semivariance values were plotted on a semivariogram, and a graphical model was fit to the data. A visual examination of the semivariogram determined the initial model (spherical, exponential, or gaussian) and values of the three main parameters (nugget effect, sill, and range). The model and parameters were then changed iteratively to minimize the RSS value.

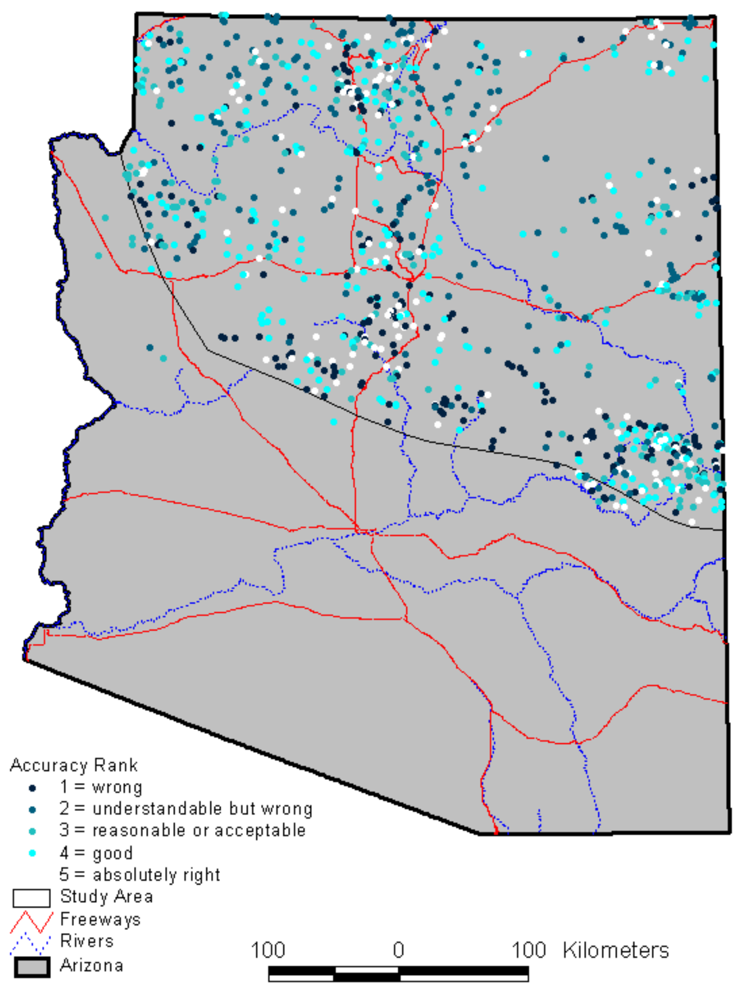
Ordinary kriging was performed on the fuzzy set reference data, chosen model, and parameters to produce a regularly spaced lattice of points representing accuracy ranks. Kriging predicted continuous (rather than ordinal) accuracy ranks ranging from one to five. The lattice of points for input into the kriging process covered northern Arizona, and were located at the center of 1 km square cells. This resulting tabular file of UTM locations and predicted accuracy ranks were converted to an Arc/INFO grid, with predicted accuracy rank as the value. The result is the fuzzy spatial view of accuracy, a map of predicted accuracy ranks for northern Arizona. The continuous accuracy rank estimates were reclassified into an ordinal variable for ease of interpretation and display. A frequency histogram was produced from the predicted accuracy ranks.

Results

The fuzzy set reference data (Figure 2-2) are positively spatially autocorrelated at shorter distance separations (Figure 2-3). This is evidenced by lower semivariance values at shorter lag distances. As typical of geostatistical data, the semivariance values reach a sill at range where they become uncorrelated.

The spherical model provides the best fit to the semivariance data (Figure 2-3). Using this model, the range (A_0) is 22.6 km. This means that the accuracy ranks are correlated at a broad scale. The sill ($C+C_0$), or semivariance peak, is 1.4081. The nugget effect (C_0) is 0.6638, or about 47% of the sill.

As a continuous variable, the predicted accuracy ranks produced from kriging do not reach the extremes of 1 (wrong) and 5 (absolutely right). Instead, they range from a minimum of 1.039 to a maximum of 4.934 (Table 2-1). The mean and median are very close to 3.



Accuracy Rank
1 = wrong
2 = understandable but wrong
3 = reasonable or acceptable
4 = good
5 = absolutely right
Study Area
Freeways
Rivers
Arizona
100
0
100 Kilometers

Figure 2-2. Fuzzy set reference data.

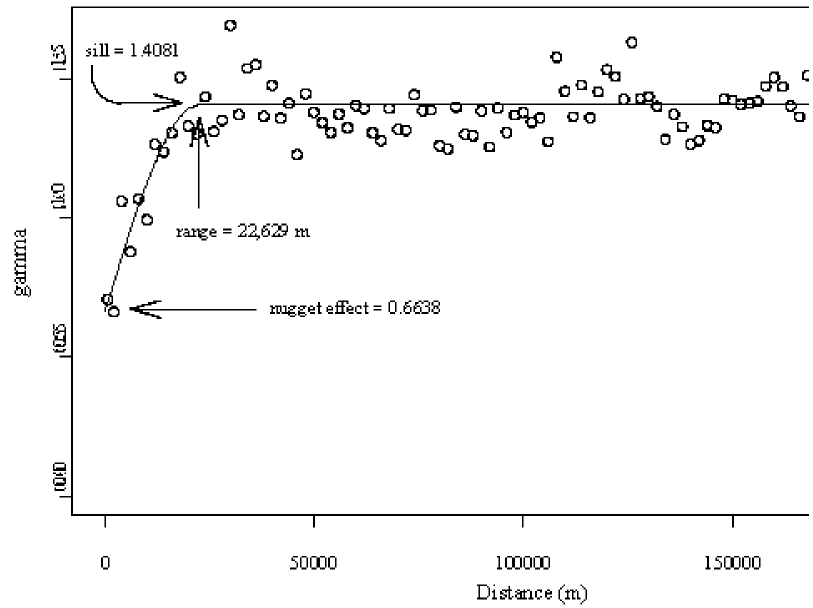


Figure 2-3. Semivariogram and spherical model of the fuzzy set reference data (930 points).

Table 2-1. Summary statistics of predicted accuracy ranks.

Accuracy Rank Standard Error	
Minimum	1.039 0.1785
Maximum	4.934 1.1880
Mean	2.901 1.1070
Median	2.900 1.1220

Figure 2-3. Semivariogram and spherical model of the fuzzy set reference data (930 points).

The fuzzy spatial view of accuracy displays the predicted accuracy ranks reclassified as an ordinal variable (Figure 2-4). High accuracy is lighter in color than low accuracy. The frequency histogram of accuracy ranks (Figure 2-5) shows that about 85% of the fuzzy spatial view of accuracy has a rank of 3, 4, or 5. In ecological terms, the land cover map is accurate to the life form level or better for a majority of the study area.

Discussion

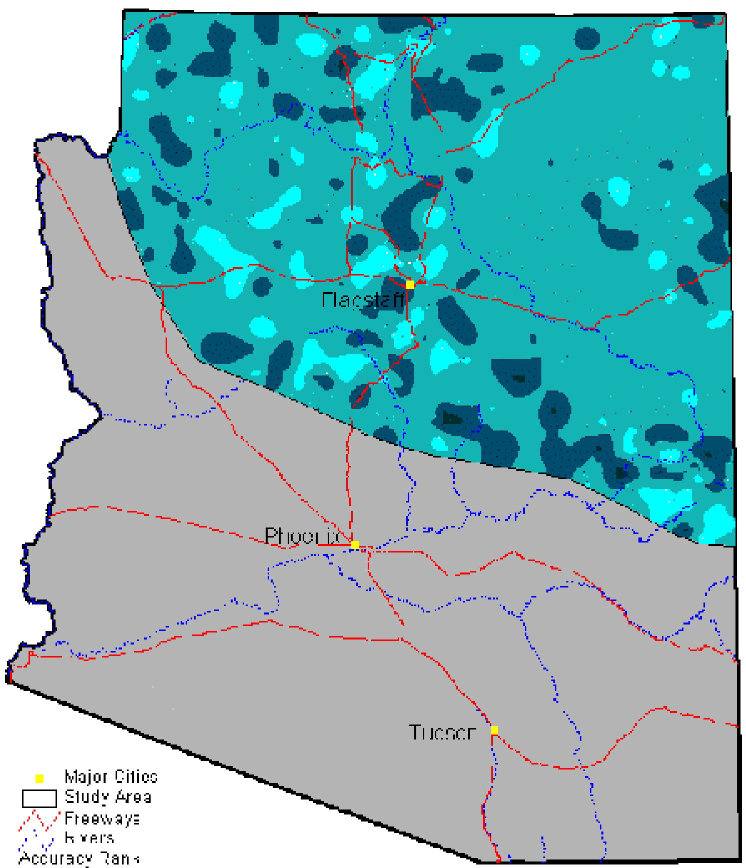
The fuzzy spatial view of accuracy created from fuzzy set theory and spatial interpolation provides spatial data about the accuracy of the preliminary AZ GAP land cover map. Not only is accuracy displayed as it varies across the northern Arizona landscape, but the degree of accuracy is conveyed by fuzzy set ranks. Overall, the fuzzy spatial view of accuracy indicates that the land cover map is accurate to the life form level, with locations of higher and lower accuracy. The histogram of accuracy ranks for northern Arizona shows that the interpolated accuracy is 85% at the life form level for all land cover types. However, where classification required identification of the dominant species, accuracy remained low (8%).

An Example: The Coconino National Forest

Foresters of the Coconino National Forest (NF) will likely be interested in the portion of the AZ GAP land cover map that includes the Coconino NF and a buffer area. Using a binary accuracy assessment, they will consult Table 1-2 to glean information about the accuracy of land cover types that occur in the forest. As an example, they may review user's accuracy for types 15 through 20, the land cover types that dominate the Coconino NF around Flagstaff.

The user's accuracy for these land cover types range from 5.1% for Ponderosa Pine-Pinyon-Juniper to 33.3% for Ponderosa Pine-Mixed Oak-Juniper. Because these percentages are so low, the foresters should put little faith in the land cover map and not use it.

Foresters will get more information on the type of accuracy of the land cover map if they use the fuzzy spatial view of accuracy. They look up the fuzzy set accuracy by land cover type (Table 1-3), and conclude that while the map and reference data seldom agree perfectly, they often agree



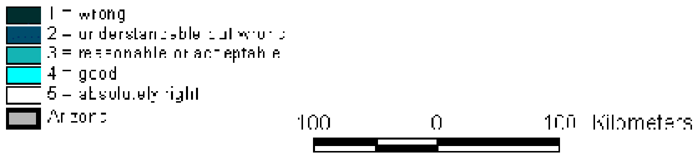


Figure 2-4. Fuzzy spatial view of accuracy.

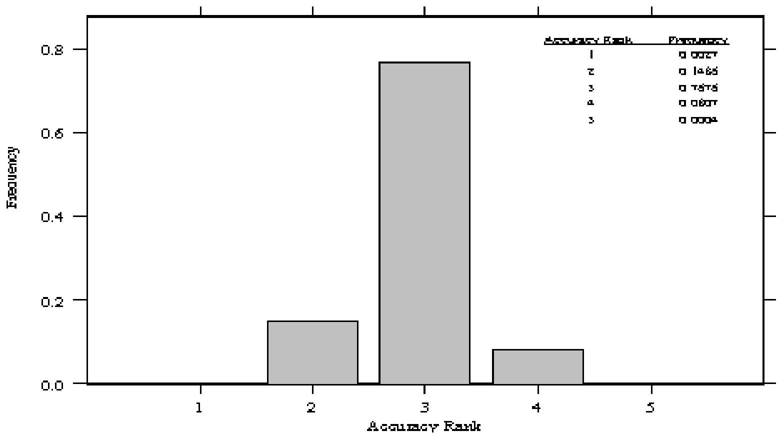


Figure 2-5. Frequency histogram of accuracy ranks in the fuzzy spatial view of accuracy.

to the life form level or better. Right statistics for codes 15 through 20 range from 55.6% for Ponderosa Pine-Gambell Oak-Juniper/Pinyon-Juniper Complex to 74.4% for Ponderosa Pine-Pinyon-Juniper. The foresters at the Coconino NF may use this additional data to decide whether the land cover map is accurate enough to use in their applications.

Not only will the foresters be able to look up fuzzy set accuracy by land cover type, but they also can view accuracy for the Coconino NF (Figure 2-6). They are able to see patches of high accuracy (accuracy ranks of 5 and 4) within the forest boundaries. From the frequency histogram of accuracy ranks for the Coconino NF (Figure 2-7), the foresters conclude 75% of the forest is classified correctly to life form or better (accuracy rank 3, 4, and 5) but only 15% is classified so that the reference and mapped types are characterized by the same dominant species (accuracy ranks 4 and 5). The foresters can decide if this accuracy is good enough to justify using the land cover map in their applications.

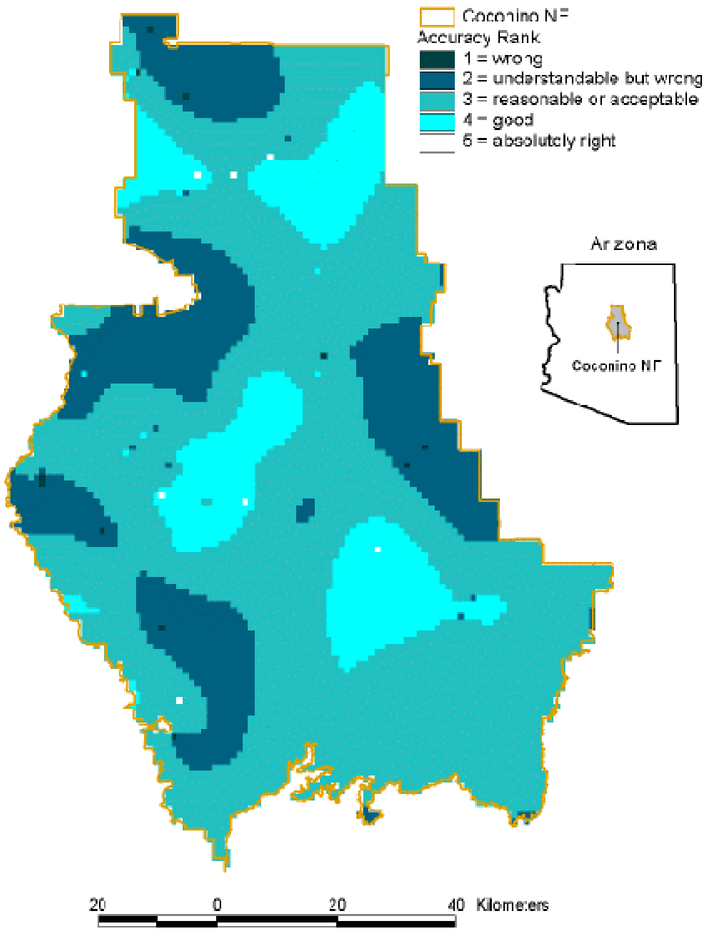


Figure 2-6. Fuzzy spatial view of accuracy for the Coconino NF.

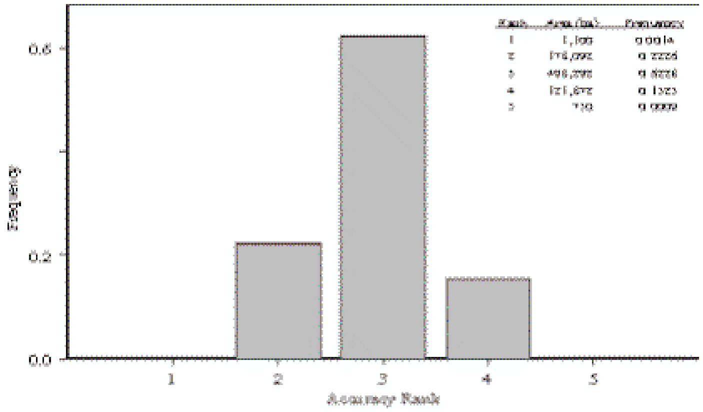


Figure 2-7. Frequency histogram of accuracy ranks in the Coconino NF.

Contribution of the Fuzzy Spatial View to Evaluating Map Accuracy

A spatial assessment improves upon both the fuzzy set and binary assessments by displaying accuracy as it varies across the landscape. Therefore, a combined approach of using fuzzy set accuracy assessment in a spatial analysis brings together the benefits of both procedures. The fuzzy spatial view of accuracy displays the degree of agreement between reference and classified data as it varies spatially over the landscape. It presents a different view of accuracy than that of a fuzzy set assessment alone.

Based on the fuzzy spatial view of accuracy, the land cover map has reached the target accuracy (80%) for mapping at the life form level of classification (accuracy rank of 3 or higher). This estimate is different from the average Right statistic (56% from Table 1-3) because the average Right statistic was not weighted for number of samples. Each cover type was given the same weight in the average, regardless of how many samples were included. However, the fuzzy spatial view of accuracy shows that at the dominant species level (accuracy rank of 5), the land cover map is even less accurate (8%) than determined by overall binary accuracy (14.8% from Table 1-2).

Section 3: Correlation of Accuracy with Land Surface Factors

Introduction

One potential source of error in a land cover map derived from remote sensing may be difficulty in classifying the spectral values in the imagery. Factors that can cause confusion during unsupervised classification may be correlated with the accuracy of the resulting map. Such factors include the slope of the terrain, the underlying soil reflectance, and possibly, the land stewardship as it may influence the complexity of vegetation in an area.

Quantitative comparisons of the fuzzy set reference data to spatially variable processes can provide insight into the spatial variability of accuracy. A linear regression tests for the relationship between a response (such as accuracy rank) and one or more predictor variables. If the relationship is deemed statistically significant, a prediction equation can be interpreted, and the spatial variable can be used to identify potential problem areas in a re-mapping effort.

In this section, a map of the fuzzy spatial view of accuracy is compared visually with slope, land status, and soil order to inspect for similarity in patterns. Then, the fuzzy set reference data is used as the response variable in a linear regression with slope, land status, and soil order, respectively, to quantitatively examine the relationship between the fuzzy set accuracy rank and the selected terrain variables.

Background

Most classical statistical analyses assume that observations are taken under identical conditions and independently from one observation to another. However, the spatial dependency of the reference data violate these assumptions and may cause bias in classical statistical tests.

Spatial correlation between variables reduces the effective sample size (degrees of freedom) for classical statistical tests (Clifford et al. 1989). Since data spatially close to each other tend to have similar values, fixing one of the data values influences the values around it. In a simple linear regression, spatial correlation between the variables results in the standard errors of the parameter estimates (slope and intercept) being underestimated. The standard error is used in a classical test for statistical significance.

$$t = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)}$$

where

t = test statistic for comparison to a t_{n-2} distribution,

$\hat{\beta}_1$ = slope parameter estimate of the regression line,

$se(\hat{\beta}_1)$ = standard error of the slope parameter estimate

An underestimated standard error of the slope parameter causes a larger test statistic, and results in an increased probability of Type I errors (rejecting the null hypothesis when it is true). Therefore, if a statistically significant relationship occurs where the null hypothesis is rejected (either true or a Type I error), corrections for the spatial autocorrelation must be made.

The strength and statistical significance of the regression defines the relationship between the response variable (such as accuracy rank) and predictor variable (such as slope, soils, and land stewardship). The statistical significance of a linear relationship is determined by the *t* statistic and corresponding *p* value. The *p* value represents the probability of obtaining the current result or more extreme given the null hypothesis (no relationship) is true. The null hypothesis is usually rejected when the *p* value is below 0.05.

If the relationship is statistically significant, the strength of the relationship is determined by the coefficient of determination (R^2). R^2 is the proportion of variability in the response variable (such as accuracy rank) explained by a linear relationship with the predictor. For example, an R^2 value of 0.50 means that 50% of the variability in accuracy rank is explained by a linear relationship with slope. This may be considered a strong relationship between the two variables, depending on the analysis.

Slope Influence on Accuracy

From the perspective of a satellite, high slope locations have a small surface area compared to low slope locations. The reflectance data of high slope locations are condensed into fewer pixels than flatter terrain, which can complicate land cover classification, resulting in poor accuracy for high slope areas. This leads to the hypothesis that there is an inverse relationship between slope and accuracy of the land cover map. That is, as slope increases, accuracy of the land cover map decreases.

Slope may be a predictor of accuracy. In order to examine the effect of slope on accuracy of the land cover map, a simple linear regression can be performed.

Land Management Status Influence on Accuracy

Different management objectives for the land affect the complexity of the landscape. Areas preserved for biodiversity (status 1 in Table 3-1) will likely have greater landscape complexity than areas managed for human use (status 4). For example, areas managed for timber growth and harvest (status 3) will be fairly homogeneous in plant species because the managers will discourage growth of non-productive species. This would be easier to map than the land cover of a more complex, natural landscape. This leads to the hypothesis that the land cover map is less accurate in areas preserved for biodiversity.

Land status may be a predictor of accuracy. In order to examine the effect of status on accuracy of the land cover map, a simple linear regression can be performed in a similar fashion as the comparison of accuracy with slope.

Table 3-1. Description of land status categories (Crist 2000).

Land Status	Description
1	An area having permanent protection from conversion of natural land cover and a mandated management plan in operation to maintain a natural state within which disturbance events (of natural type, frequency, intensity, and legacy) are allowed to proceed without interference or are mimicked through management.
2	An area having permanent protection from conversion of natural land cover and a mandated management plan in operation to maintain a primarily natural state, but which may receive uses or management practices that degrade the quality of existing natural communities, including suppression of natural disturbance.
3	An area having permanent protection from conversion of natural land cover for the majority of the area, but subject to extractive uses of either a broad, low-intensity type (e.g., logging) or localized intense type (e.g., mining). It also confers protection to federally listed endangered and threatened species throughout the area.
4	There are no known public or private institutional mandates or legally recognized easements or deed restrictions held by the managing entity to prevent conversion of natural habitat types to anthropogenic habitat types. The area generally allows conversion to unnatural land cover throughout.

Soils Influence on Accuracy

Soil order may be a predictor of spatial accuracy. While the spectral characteristics of soils and vegetation are different, they can have similar properties in certain bands of Landsat Thematic Mapper imagery. In land cover classification, confusion can result from soils dominating or heavily influencing the spectral properties of a particular location (Murphy and Wadge 1994). This is especially evident in areas where soils cover greater than 40% of the pixel, a common occurrence in the arid southwest. This leads to the hypothesis that there is a relationship between soil order and accuracy of the land cover map. It is proposed that areas with more arid soils (aridisols) are less accurate than other areas.

Soil order may be a predictor of accuracy. In order to examine the effect of soil order on accuracy of the land cover map, a linear regression can be performed.

Methods

Maps of Slope, Land Status, and Soil Order

A grid map of slope angles was derived from 1:24,000 digital elevation models for the study area using Arc/INFO. Slope was transformed by taking the natural logarithm twice to get a more normal distribution. However, because of the preponderance of values at zero, one was added before each logarithmic transformation to get a more normally distributed variable. While not necessary, it is desirable to have a normally distributed variable for comparison with accuracy rank. This prevents extreme values from having a disproportionate effect on results in classical statistical tests. Frequency histograms of slope values and transformed values were derived using S-PLUS.

A digital map of land management status for all of Arizona was obtained from the Sonoran Desert Field Station and was clipped to the study area. This map contained four categories of land stewardship (Table 3-1) as defined by the National GAP Handbook (Crist 2000).

The digital map of soil orders (Table 3-2) was obtained from Natural Resources Conservation Service (NRCS) State Soil Geographic (STATSGO) Database (USDA 1994).

Table 3-2. Soil order and major characteristics (adapted from Brady 1990).

Order	Major characteristics
Alfisols	Silicate clay accumulation and/or may be high in sodium, columnar or prismatic structure; high to medium base saturation
Aridisols	Dry soil, light colored, low organic content; sometimes silicate clay accumulation and/or may be high in sodium, columnar or prismatic structure
Entisols	Little profile development; light colored, low organic content, may be hard and massive when dry
Inceptisols	Embryonic soils with few diagnostic features; low organic content, strong structure; changed or altered by physical movement or by chemical reactions
Mollisols	Thick, dark colored, high base saturation, strong structure some with silicate clay accumulation and/or high in sodium, columnar or prismatic structure
Vertisols	High in swelling clays; deep cracks when dry

Regression Analysis

In Arc/INFO, the fuzzy set reference data (930 points) were intersected with the slope, land status, and soils maps. From the resulting point coverage, a tabular file of location, accuracy rank, slope, land status, and soil order for each point was produced.

Two simple linear regressions were calculated for accuracy rank (response variable) using 1) transformed slope and 2) land status as the predictor variables. Parameter estimates and standard errors, as well as the R^2, t statistic, and p value were determined for each regression in S-PLUS.

Unlike transformed slope and land status which are ordinal variables, soil order is a nominal variable, requiring a slightly different statistical test for association with accuracy. To represent the effects of a categorical predictor variable that takes on m possible levels, m-1 indicator variables must be used (Weisberg 1985). An indicator, or dummy, variable can take a 0 or 1 value. Therefore, since there are 6 soil orders, 5 indicator variables (D_i) were used (Table 3-3). Using these indicator variables as the predictor variables, accuracy rank was used as the response variable in a linear regression. A test of an overall relationship between the indicator variables and accuracy rank was given by the F statistic and corresponding p value. In addition, specific variables were tested using the t statistic.

	Indicator Variable				
Soil Order	**D_1**	**D_2**	**D_3**	**D_4**	**D_5**
Alfisols	0	0	0	0	0
Aridisols	1	0	0	0	0
Entisols	0	1	0	0	0
Inceptisols	0	0	1	0	0
Mollisols	0	0	0	1	0
Vertisols	0	0	0	0	1

Results

Visual comparison of the slope map with the fuzzy spatial view of accuracy (Figure 3-1), the land status map with the fuzzy spatial view of accuracy (Figure 3-2), and the soil order map with the fuzzy spatial view of accuracy (Figure 3-3) do not reveal any apparent correlation patterns.

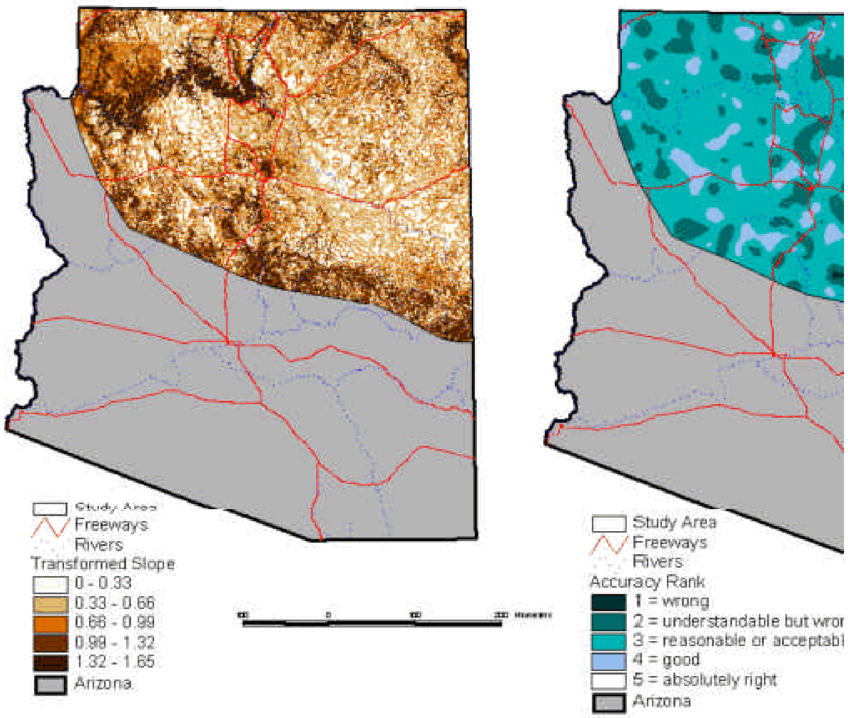


Figure 3-1. Transformed slope (left) and the fuzzy spatial view of accuracy (right).

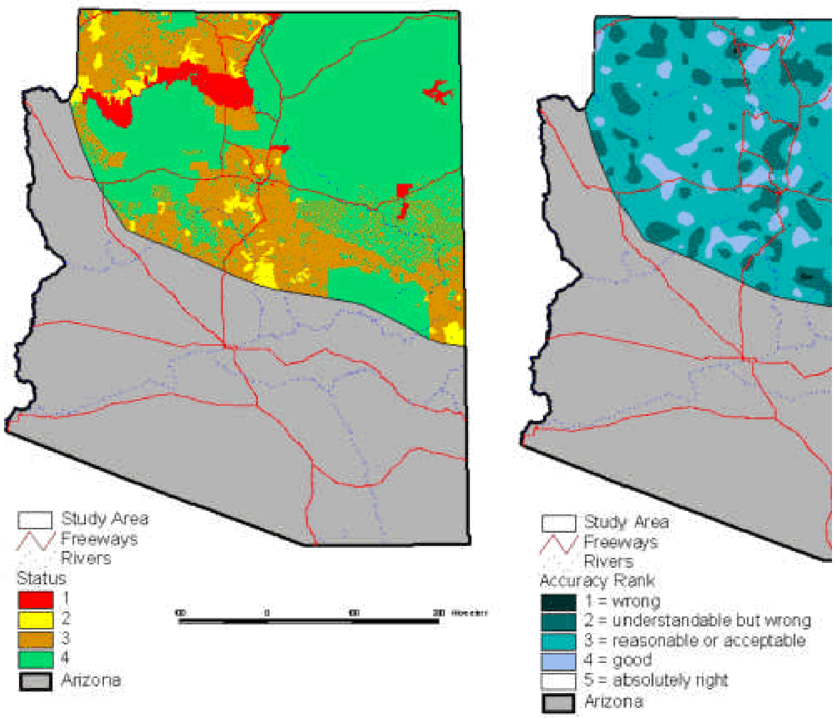


Figure 3-2. Land status (left) and the fuzzy spatial view of accuracy (right).

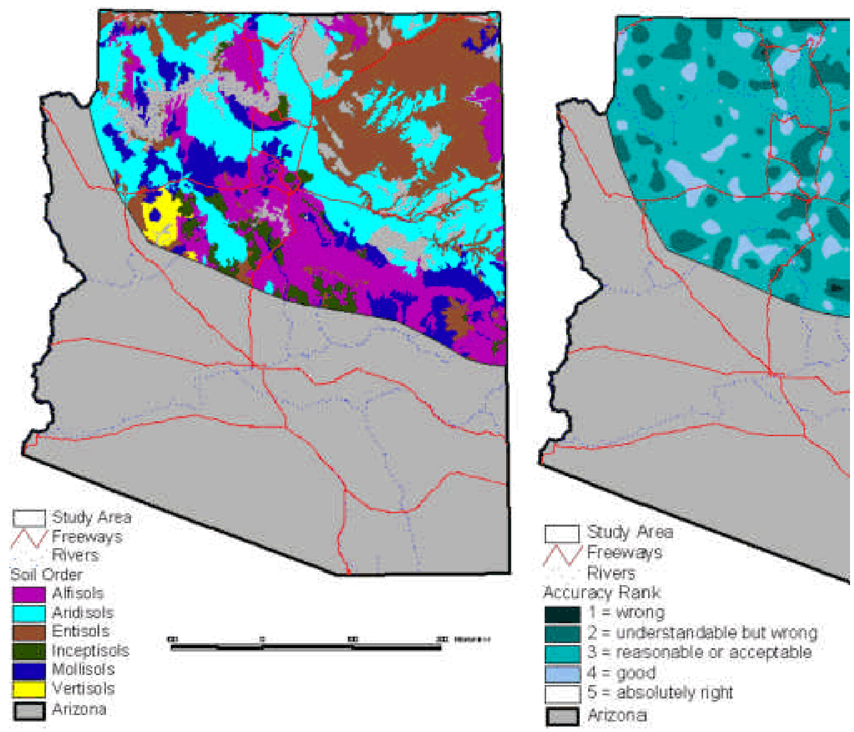


Figure 3-3. Soil order (left) and the fuzzy spatial view of accuracy (right).

A frequency histogram of slope for the study area shows a right skewed distribution with a preponderance of values at zero (Figure 3-4). This implies that much of the terrain in the study area is relatively flat. Obvious exceptions are the Grand Canyon and the San Francisco Peaks. The frequency histogram of transformed slope angles shows a fairly normal distribution with many values at zero (Figure 3-4).

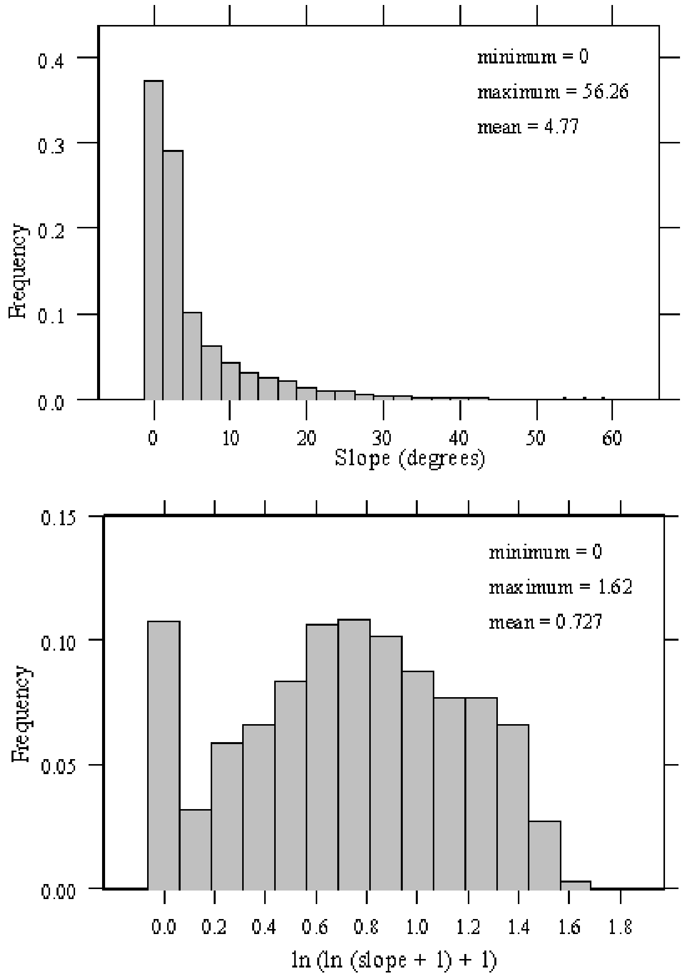


Figure 3-4. Frequency histograms of slope (top) and transformed slope (bottom).

The regression analysis of transformed slope on accuracy rank shows no significant relationship (Table 3-4). The *t* statistic for the slope parameter is not significant. Therefore, the null hypothesis, that there is no relationship between slope and accuracy rank, can not be rejected.

Table 3-4. Simple linear regression of slope on accuracy.

Multiple R^2 = 0.00005532				N = 930
Prediction Equation: $Accuracy\ Rank = 2.8871 + .0220(\ln(\ln(slope + 1) + 1)$				
Variable	**Parameter**	**Standard Error**	**t value**	**p value**
Intercept	2.8871	0.0828	34.8541	0.0000

Transformed Slope	0.0220	0.0988	0.2233	0.8234

The regression analysis of land status on accuracy rank does not show a linear relationship between the two variables (Table 3-5). The *t* statistic for the status parameter is not significant. Therefore, the null hypothesis, that there is no relationship between land status and accuracy rank, can not be rejected.

Table 3-5. Simple linear regression of land status on accuracy.

Multiple $R^2 = 0.0004764$ N = 930

Prediction Equation:

$Accuracy\ Rank = 2.7645 + 0.0413(land\ status)$

Variable	Parameter	Standard Error	t value	p value
Intercept	2.7645	0.2155	12.8311	0.0000
Land Status	0.0413	0.0631	0.6549	0.5127

The regression analysis of soil order on accuracy rank shows that overall, the relationship is not significant (Table 3-6) because the F statistic does not have a p value less than 0.05. However, individual t tests of each indicator variable show that D2 and D3 are statistically significant. Both of these soils, entisols and inceptisols, are younger soils. Their presence is correlated with map accuracy; however, the overall R2 of the regression is low so that their contribution to the variability in map accuracy is low.

Table 3-6. Linear regression of soil order on accuracy.

Multiple $R^2 = 0.01274$ N = 930

Prediction Equation:

$Accuracy\ Rank = 2.9000 + 0.0729(D_1) + 0.1022(D_2) - 0.0973(D_3) + 0.0193(D_4) + 0.0200(D_5)$

Variable	Parameter	Standard Error	t value	p value
Intercept	2.9000	0.1418	20.4508	0.0000
D_1	0.0729	0.0608	1.1999	0.2305
D_2	0.1022	0.0430	2.3771	0.0177
D_3	-0.0973	0.0504	-1.9314	0.0538
D_4	0.0193	0.0314	0.6143	0.5392
D_5	0.0200	0.1335	0.1500	0.8808
			$F_{5,924} = 2.086$	0.0651

Two issues about the statistical significance of D2 and D3 are brought into question. First, the spatial autocorrelation of accuracy rank may be causing a Type I error (rejecting the null hypothesis when it is true). Corrections should be made to the test, which are not possible at this level of analysis. Second, if the variable is truly correlated with accuracy rank, there is very little value to the information. For example, where entisols are present, the predictive equation becomes:

Accuracy Rank = 2,9000 = 0.1022 = 3.0022

Since the average value of the fuzzy spatial view of accuracy is 2.9, accuracy actually increases where entisols are present, but only very slightly (recall that an accuracy rank of 5 is absolutely right). A similar situation exists for inceptisols. Therefore, the statistical significance of these two variables does not mean that they are practically significant. That is, entisols and inceptisols provide very little, if any, information about accuracy of the land cover map.

Discussion

Accuracy of the land cover map was not correlated to slope and land status and only weakly correlated with soil type. The underlying cause of error in the vegetation map may be complex. While the analysis in this section does not identify land surface variables strongly correlated with error in the vegetation map, other spatial factors can not be eliminated. One source of error may be inherent in the manner ground truthing data was collected and analyzed for classification of the original imagery. Ground truthing for this map consisted of interpreted video images and field based observations. Distance from the video flight lines or from field plots may be correlated with the accuracy of the land cover map. However, that correlation was not examined in this analysis.

Section 4: Recommendations

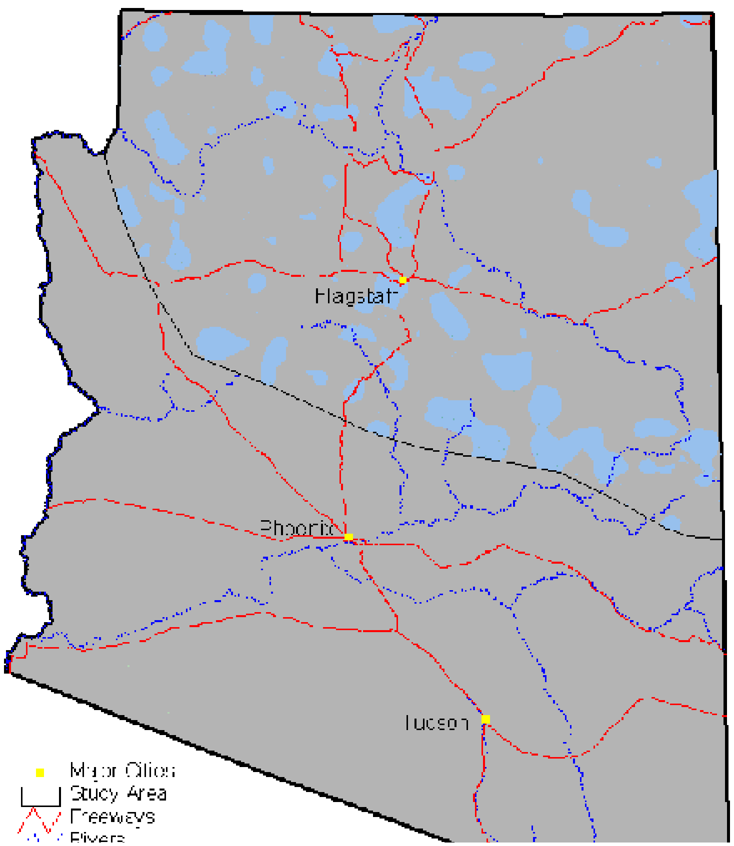
Fuzzy set theory suits the continuous nature of land cover especially well. When the land cover classification scheme describes numerous types that are difficult to separate from similar types, fuzzy set assessment seems most appropriate. By allowing for multiple set memberships, more information is gleaned from an assessment providing a more detailed analysis of accuracy.

However, fuzzy set theory is not useful and practical in all situations. Where few, broad categories exist, confusion over which class to assign a location becomes less of an issue. Stehman asserts that while a fuzzy set accuracy assessment provides a lot more data, it is unclear whether there is more usable information in the data (Stehman 1998). However, he acknowledges that some modifications to the traditional binary assessment need to be made.

The fuzzy spatial view of accuracy provides a visual map of accuracy of the preliminary AZ GAP land cover map. This map delineates locations of poor and good accuracy on the land cover map. This is important for users of the land cover map to decide if the land cover map is accurate enough in their planning area to justify using it. We recommend that producers of a land cover map consult the fuzzy spatial view of accuracy to determine where and how to focus map updating.

We highly recommend to spatially interpolate fuzzy set reference data to create a map of accuracy rankings. The fuzzy spatial view of accuracy provides a wealth of information to users and producers of the land cover map, which allows for more informed use and production of the map.

The fuzzy spatial view of accuracy delineates locations of poor accuracy (accuracy ranks of 1 and 2) (Figure 4-1). Ground observations should be intensified in these areas during SW ReGAP. Also, because the AZ GAP map proved to be 85% accurate at the life form level, preliminary versions of the SW ReGAP map in Arizona could be compared to the original AZ GAP map at the life form level for a first order validation.



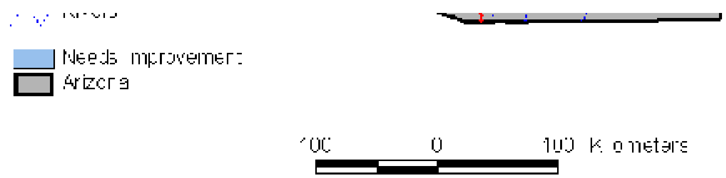


Figure 4-1. Areas recommended for intensified land cover mapping.

Citations

Arc/INFO Ver. 7.2.1. 1998. Environmental Research Systems Institute, Inc., Redlands, CA.

Arundel, T. 1997. Geographer.USGS-BRD Colorado Plateau Field Station.

Brady, N. C. 1990. The Nature and Properties of SOILS. 10th Edition. New York: MacMillan.

Brown, D. E., C. H. Lowe, and C. P. Pase. 1979. A digitized classification system for the biotic communities of North America, with community (series) and association examples for the Southwest. Journal of the Arizona-Nevada Academy of Science 14, Supplemental issue 1: 1-16.

Canters, F.. 1997. Evaluating uncertainty of area estimates derived from fuzzy land-cover classification. Photogrammetric Engineering & Remote Sensing 63, no. 4: 403-14.

Chou, Y. 1997. Exploring Spatial Analysis in Geographic Information Systems. Santa Fe, NM: OnWord Press.

Clark, I. 1979. Practical Geostatistics. London: Elsevier Applied Science.

Clifford, P., S. Richardson, and D. Hémon. 1989. Assessing the significance of the correlation between two spatial processes. Biometrics 45: 123-34.

Cochran, W. G. 1977. Sampling Techniques. 3rd Edition. New York: Wiley & Sons.

Congalton, R. G. 1988. Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. Photogrammetric Engineering & Remote Sensing 54, no. 5: 587-92.

???. 1996. Accuracy assessment: A critical component of land cover mapping. in Gap Analysis: A Landscape Approach to Biodiversity Planning. Editors J. M. Scott, T. H. Tear, and F. W. Davis, 119-31. Bethesda, MD: American Society for Photogrammetry and Remote Sensing.

Congalton, R. G., and K. Green. 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Mapping Science Series, ed. John G. Lyon. Boca Raton, FL: Lewis Publishers.

Cressie, N. 1993. Statistics for Spatial Data. Revised Edition. New York: Wiley & Sons.

Cressie, N., and D. M. Hawkins. 1980. Robust estimation of the variogram: I. Mathematical Geology 12: 115-25.

Crist, P. 1996. Final summary of the land cover accuracy assessment workshop, Denver, CO.

Crist, P. 2000. Mapping and categorizing land stewardship. Version 2.1.0 (February 23, 2000). A handbook for conducting Gap Analysis. Internet WWW page, at URL: <http://www2.gap.uidaho.edu/handbook/Stewardship/default.htm>

Fisher, P. F. 1994. Visualization of the reliability in classified remotely sensed images. Photogrammetric Engineering & Remote Sensing 60, no. 7: 905-10.

Gopal, S., and C. Woodcock. 1994. Theory and methods for accuracy assessment of thematic maps using fuzzy sets. Photogrammetric Engineering & Remote Sensing 60, no. 2: 181-88.

Graham, L. 1994. Preliminary Arizona Gap Analysis Program Land Cover Map. Tucson, AZ: University of Arizona.

Jacobs, S. 1999. A Fuzzy Spatial View of Accuracy of a Northern Arizona Land Cover Map. Master's Thesis. Flagstaff, AZ: Northern Arizona University.

Kaluzny, Stephen P., Silvia C. Vega, Tamre P. Cardoso, and Alice A. Shelly. 1996. S+SPATIALSTATS User's Manual. Seattle: MathSoft.

Matheron, G. 1963. Principles of geostatistics. Economic Geology 58: 1246-66.

Murphy, R. J., and G. Wadge. 1994. The effects of vegetation on the ability to map soils using imaging spectrometer data. International Journal of Remote Sensing 15, no. 1: 63-86.

Odland, J. 1988. Spatial Autocorrelation. Newbury Park: SAGE.

Rouhani, S. 1986. Comparative study of ground-water mapping techniques. Ground Water 24, no. 2: 207-16.

Royle, A. G. 1980. Why Geostatistics? in Geostatistics. A. G. Royle, I. Clark, P. I. Brooker, H. Parker, A. Journel, J. M. Endu, R. Sandefur, D. C. Grant, and P. Mousset-Jones. New York, NY: McGraw Hill.

Steele, B. M., J. C. Winne, and R. L. Redmond. 1998. Estimation and mapping of local misclassification probabilities for thematic land cover maps. Remote Sensing of Environment 66, no. 2: 192-202.

Stehman, S. 1998. Thoughts on "fuzzy" versus "hard" land-cover classes in accuracy assessment.

Stoms, D., F. Davis, C. Cogan, and K. Cassidy. 1994. Assessing vegetation/land cover map accuracy for gap analysis, Santa Barbara, CA.

Thomas, K. A., Vegetation Ecologist. 1998.
Verbyla, D. L., and T. O. Hammond. 1995. Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids. International Journal of Remote Sensing 16(3): 581-587.
USDA, Natural Resources Conservation Service. 1994. State Soil Geographic (STATSGO) data base for Arizona. Fort Worth, TX: USDA NRCS. Internet WWW page, at URL: URL:http://www.ftw.nrcs.usda.gov/stat_data.html
Weisberg, S.. 1985. Applied Linear Regression. 2nd Edition. New York: Wiley & Sons.
Zadeh, L. A. 1965. Fuzzy sets. Information and Control 8: 338-53.

Methods for Developing Terrestrial Vertebrate Distribution Maps for Gap Analysis

Version: 2.0.0
Date: 16 February 2000
Peer-reviewed: No

Blair Csuti & Patrick Crist
Idaho Cooperative Fish and Wildlife Research Unit
University of Idaho, Moscow, ID

Introduction

The gap analysis process compares the distribution of several elements of biological diversity with areas managed primarily for native species and natural ecosystems (Scott et al. 1993). To accomplish this task, it is necessary that the Gap Analysis Program (GAP) develop detailed, high-confidence maps of the distribution of individual animal species for comparison with maps of land stewardship and management status. These comparisons are used to assess the habitat area and relative percent of distribution of a species in the different categories of stewardship and biodiversity management status. This document describes methods and standards for producing high-confidence vertebrate distribution maps. It was developed from the previous version of this chapter, the GAP projects, and in consultation with a committee of GAP researchers and cooperators. Because this is a synthesis of the work from several GAP projects, the origin of the method is referenced as "(see XX-GAP)," with the XX indicating the two character code for the state, e.g., ID = Idaho. For additional information from those states, please refer to their reports, journal articles, and contacts found on the GAP Web site.

Objectives

1. Provide maps of known confidence that predict the distributions of all terrestrial vertebrate species in the project area in order to support analysis of conservation status.
2. Develop a database of locational records, geographic range, wildlife habitat associations, and predicted distribution of each vertebrate species for the long-term utility for GAP and its cooperators.

Background

Most information on species distribution has typically been collected at the scale of individual field sites and is often extrapolated to create small-scale range maps for state, regional, or national reference works and field guides. Typically lacking for most species is a meso-scale (approximately 1:100,000) expression of a spatially detailed distribution map versus a range map which only depicts coarse, regional representation. Prior to the availability of detailed GIS coverages (including GAP land cover maps), there simply was no way to depict species distribution at the meso-scale. There are three types of distribution expressions: actual distribution which is based on exhaustive, long-term surveys that are very rare; known distribution which is based on current knowledge of where the species has been found and is usually incomplete, and predicted distribution which combines known distribution and knowledge of habitat associations of the species to extrapolate to areas where the species is expected to occur. The basic assumption of GAP's predicted species distribution maps is that a species has a high probability of occurring in appropriate habitat types that are within its predicted range. GAP links species' ranges to satellite-derived land cover maps and other physical data, which are intermediate in scale between representation in field sites and field guides. It is simply impractical to map the distribution of species through intensive field surveys for entire states, regions, or nations (Scott et al. 1986). GAP therefore makes use of existing information on range limits and refines it to develop spatial statements of the probability of a species being present in map polygons that represent appropriate habitat as understood from current knowledge of the species and the ability to map its habitat. (Csuti and Scott 1991, Scott et al. 1991, 1992, 1993, Butterfield et al. 1994).

No matter what their scale, *all* range and distribution maps are predictions about the presence of a species in a particular geographic area. The accuracy of those predictions generally improves as the size of the area and length of sampling period are expanded because temporal scale and heterogeneity of large areas makes it more likely that a species will be found to occur there. GAP maps of predicted distributions are currently intended for use and validation at the landscape scale–an area kilometers across (Forman and Godron 1986), but new efforts are able to attribute species to "patches" as small as 2 hectares. For some species such resolution may be desirable to allow more precise estimation of habitat area, while for other species such small patches may be biologically meaningless. For the majority of species, the ability to map at 2 ha resolution probably exceeds our knowledge of their biology. In 1993, we began research to determine the defensibility, utility, and ability to validate such high-resolution predictions (Edwards et al. 1993, Scott et al. 1993, Krohn 1996, see also Continuing Research in this chapter).

Another challenge to increasing the precision of species' predicted distributions is the association of some species to microhabitat features (standing dead trees, cliff faces, wetlands) that are irregularly distributed. This is a different issue than increasing the thematic and spatial detail of land cover maps, because these features are irregularly distributed and not possible or appropriate to map as part of the current level of effort. Species requiring these features must therefore be predicted for those land cover types or other regions (such as certain slopes or soil types) known to contain those features, albeit with less precision than those species with less specialized requirements. "At relatively large scales [= areas]... one can assume that many of the special habitat elements are present... however, for site specific projects... reasonable accuracy of predictions is present only when special habitat elements are considered" (Mayer and Laudenslayer 1988). This raises the issue of inconsistency between methods and levels of confidence in the maps, but all maps of species distribution are predictions with probabilities of accuracy that change depending on the scale of the inquiry (how accurate is the prediction for this polygon versus this county?) and year to year (how accurate this year versus next, or over ten years of field assessment?). Therefore it is critical that users of GAP data heed the statements of data limitations and evaluate the suitability of the data for their intended purpose. Treatment of rare species is described under the Methods section below.

Despite the constraints mentioned above, predicting species distributions by relating them to environmental features that can be mapped from remotely sensed data is a rapid and efficient approach to estimating the distribution and management status of elements of biodiversity (when compared to field survey approaches).

Methods

<u>Summary of steps and standards</u>: A more detailed narrative follows which includes acronym definitions.

Step 1. Establish a list of species to be mapped. Use The Nature Conservancy (TNC) list of names and codes.

Step 2. Obtain locational records for each species and attribute their source in a database. Attribute records for date and collector, if applicable.

Step 3. Delineate the range extent for each species using the best available information, and subdivide the range extent into areas or units of known recent occurrence and those of extrapolated or predicted occurrence using the location records in Step 2. Conduct an expert review of the range extent maps. If the Environmental Monitoring & Assessment Program's (EMAP) hexagons are not used for the range extent, attribute these with the final GAP-predicted occurrence for each species.

Step 4. Identify or develop a database of habitat association information for each species. Provide references for all information.

Step 5. Develop a Wildlife/Habitat Relationship Model (WHRM) for each species, based on available GIS coverages, and document in a database. Use the National Hierarchical Framework of Ecological Units (ECOMAP) sections or subsections to stratify regional differences in habitat associations. Conduct an expert review of the WHRMs.

Step 6. Integrate range limits and habitat associations into a draft predicted species distribution map attributed by known and predicted occurrence from the range extent map. Conduct expert review of the draft maps.

Step 7. Conduct final mapping. Use the general GAP standards for data format and documentation.

Step 8. Prepare the distributions for comparison to ownership and management. Use the Analysis section standards.

These steps are diagrammed in Fig. 1. The following narrative provides details for each step in the process. Treatment of threatened and endangered (T&E) species and microhabitat specialists are described at the end of each step, if applicable.

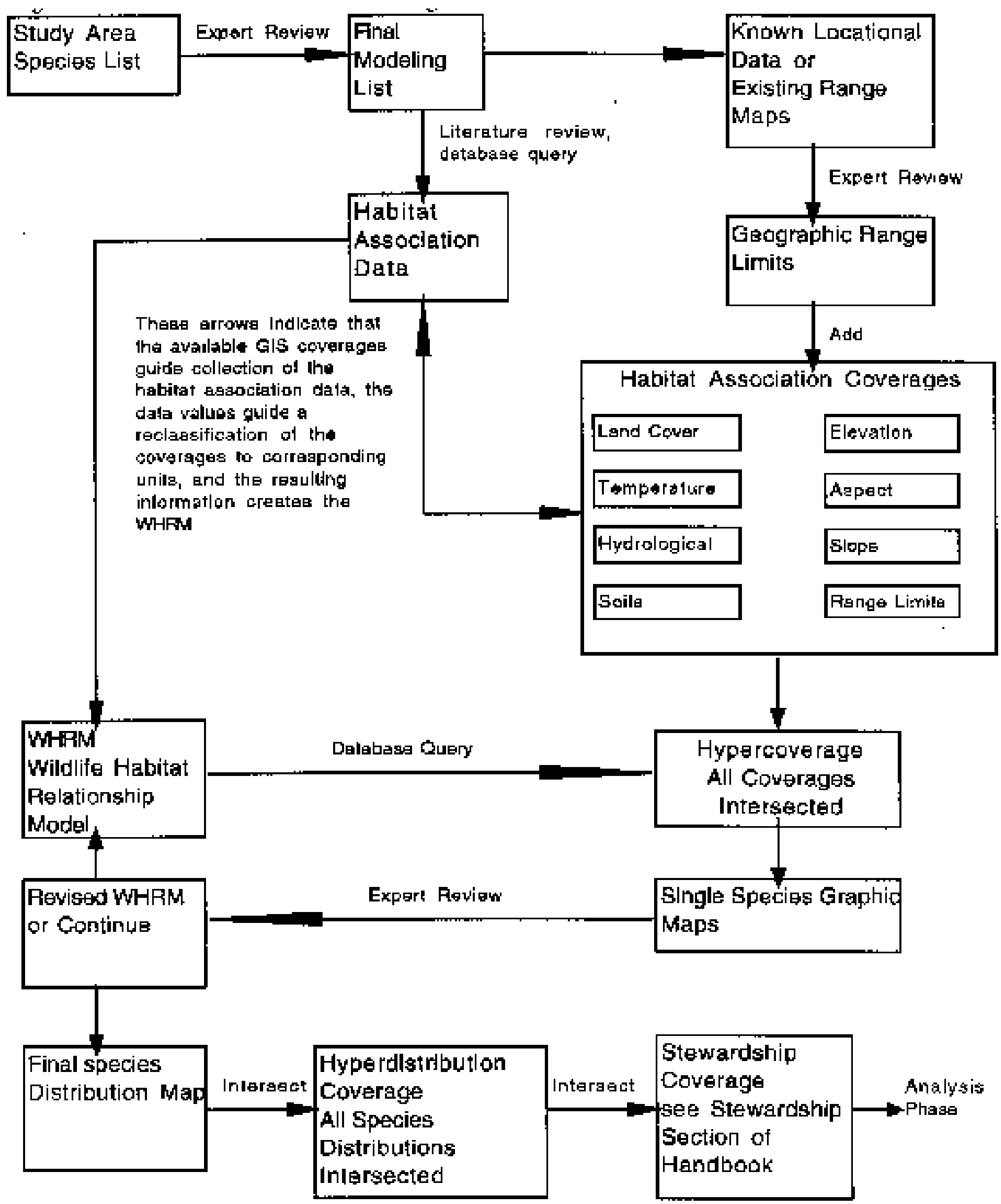


Figure 1. Flow Chart of Vertebrate Modeling Process.

Step 1: Establish a list of species to be mapped. Use the TNC "Selected Nature Conservancy Scientific Data Sets–Vertebrate Animals" (available from the World Wide Web URL http://www.consci.tnc.org/src/zoodata.htm) and review among GAP project cooperators to establish a list of species to be modeled. The goal is to develop a list that will be reasonably certain to include at a minimum those species known to breed in the project area and that are regularly occurring non-accidentals. Species that are extinct or extirpated from the study region should not be included. It is important to distinguish native breeding residents as an attribute in the database. Individual scientists are likely to have different criteria for determining "breeding" status. A species must have bred in the project area consistently over a period of time. Evidence of breeding five out of the past ten years is a suggested criterion.

Where the presence of exotic species impacts native species, we require their inclusion as a secondary gap analysis factor beyond management status (see Management Status section in the Handbook). Their presence may be attributed to management status units rather than modeling their distribution, though a distribution map may be of greater utility. Some state agencies maintain good maps of exotic species locations, especially game species. Consult those agencies for spatial data.

After establishment of the GAP list, use the TNC database for official names, codes, and status attributes. The English and Latin names in this database follow standard reference works by herpetologists (Collins et al. 1990), ornithologists (American Ornithologists Union 1983 and Supplements), and mammalogists (Jones et al. 1992). A state may accommodate in-state variations of names and codes by adding additional fields in the database that will allow use of either system to identify the same species. Though current status attributes are best obtained from the TNC database at the time of data use, including them in the species distribution coverage as of the GAP modeling will serve as an important record to justify inclusion in the list, modeling techniques used, and to identify change over time. The TNC database also indicates species that are adapted to human-modified landscapes and are indicated by "urban/edificarian" (associated with structures), "suburban/orchard", and "cropland/hedgerow". These distinctions may prove useful in the analysis stage (see Analysis chapter).

Step 2: Identify locational records for each species. The goal is to develop a "picture" of where the species is known to occur to guide the modeling process in the important step of establishing range limits and to serve as a visual method of comparison between those records and GAP-predicted distributions to support editing, evaluation, and future field sampling activities. This step involves compiling the available information on a species' distributional limits into a common spatial database but does not expect projects to be exhaustive in obtaining such records without substantial cooperator support. GAP does not rely solely on locational records but combines them with subjective expert opinion to develop a picture of range extent. This step describes acquisition of locational records; use of other information follows in Step 3.

The basic sources of information include: 1) existing range maps, which will usually be printed at very small scales; 2) agency records; and 3) other reliable records such as those from conservation institutions. These sources may not be completely independent. For example, the ranges published in most books and field guides are based to some extent on specimen locality records, and therefore they cannot be assumed to be independently validating each other. Additional and/or more recent records can often be obtained from agencies and museums (see WA-GAP, see Appendix 1). Since not all museums have computerized their collection records, localities obtained directly from museums will be a subset of the information collectively housed in the nation's museums. While it is possible to visit museums whose collections are not computerized, this is usually impractical. Other records may include localities for specific species published in the scientific literature, records based on specimens or observations by local experts, and U. S. Fish and Wildlife Service Breeding Bird Survey observations. In some cases these records have already been summarized into publications or databases. Examples include breeding bird atlases and, for globally or state imperiled and vulnerable species, Conservation Data Centers of the Natural Heritage Programs and the Multi-state Databases (see CO-, OR-, NM-GAP). For states lacking otherwise available data in digital format, it is worthwhile to pursue a cooperative project to make this conversion (see TX-GAP). If analog records are converted to digital spatial records, attribute to the smallest defensible unit. Examples include State Plane Sections, grid blocks of coordinate systems, or–for riparian species–the Environmental Protection Agency’s (EPA) River Reaches. Some states have been successful in obtaining recent species locations from industry-produced studies, e.g., environmental impact statements (EISs) (see WA-GAP). Regardless of source, it is important to store and document the information to retain its utility beyond the GAP application.

For species listed as rare, threatened, or endangered (RTE) by state or federal agencies, obtain point locations from the appropriate database, such as National Heritage Program (NHP), U.S. Fish and

Wildlife Service (USFWS), or state game and fish department. An agreement with the owner of the data establishing rules of data distribution and use (typically a spatial generalization) is usually required (see Appendix 2) (see NM-, WA-GAP).

Step 3: Delineate the range extent for each species. Use the best available information, but at a minimum you must provide EMAP hexagons (see below) attributed by GAP-predicted occurrence of all mapped species. The goal is to establish the limits of a species' range (through documented sources) where it can be expected to be found in the appropriate habitat within those limits. Projects are encouraged to use the best available information and method of delineation, even if it diverges from the standard recommended approach below. If current high-confidence maps already exist, those should be used. The following method describes creation of new range extent maps.

Create a spatial database and map of the geographic range based on the information developed in Step 2 that will allow flexibility for updating (as opposed to monolithic polygons that must be redrawn for each update). The GAP standard unit is the EPA's EMAP hexagon, a regular, equal-area 635 km^2 hexagonal grid developed for the program (White et al. 1992). The hexagon cell is an intermediate size between the very large western counties and the smaller eastern townships or parishes and provides the advantage over county-of-occurrence maps of being an equal area unit that can be matched across project area boundaries. It also has the advantage of permanence over units derived from natural characteristics such as watersheds or soils units which may be subject to reinterpretations and are not easily aggregated or tessellated into hierarchical units. Additionally, several NHPs are currently populating the hexagon grid–thus reducing or eliminating that task for GAP–or are conducting the task in cooperation with GAP (see ID-, MT-, OK-, OR-, PA- WA-, WY-GAP). As the minimum mapping unit (MMU) of GAP land cover maps decreases, the range unit must also decrease to support that resolution (see Upper Midwest-GAP). The hexagon coverage for the U.S. may be accessed from the GAP home page and can be tessellated into finer units by a constant 1/7 factor as desired. Other units or hand-delineated range polygons may be more advantageous for species distribution modeling, but GAP will require attributing the grid with the GAP projects' predicted occurrence as a minimum deliverable.

After establishing the range unit, it must be attributed with both species occurrence and whether occurrence is based on a recent collection or observation, or is extrapolated from expert opinion. If using the hexagon process, follow the methods established by WY-GAP found in Appendix 1. Conduct expert review of the draft range extent maps. GAP uses the subjective review process because of the lack of current or complete occurrence data. It is helpful to reviewers to provide overlays of county boundaries, known occurrence locations, and major landscape features such as rivers and mountain ranges. The results of expert reviews should be documented; for instance, if a range unit is added because of expert opinion, the species, unit, and reviewer should be noted in a metadata file for that species. For migratory species, the range units must also be attributed as "breeding" or "nonbreeding" occurrence. If the unit is sufficiently large that both attributes occur, you may include both or default to "breeding" occurrence.

Step 4: Acquire or develop a database of wildlife habitat associations. The goal is to identify, through documented sources, those physiographic characteristics with which a species is associated, that are mapped or can be derived from existing maps. While known distribution can only be derived from locality records, the next step in predicting complete species distributions requires building a database that identifies appropriate habitat for each species. Some states will be fortunate to have an existing digital database of species habitat associations that should require only minimal translation for the modeling process. Most states, however, will have to construct one from scratch, though they may benefit from use of an adjacent state's database. This chapter provides general instructions for development of a habitat associations database, but the technical aspects of database development require expertise in that area. GAP also encourages states to cooperate on development of the database when appropriate. The database, whether existing or GAP-created, uses peer-reviewed literature about the ecology and natural history of individual species or groups of species as the primary source of information about the habitats in which a species can be expected to occur. These literature sources represent the cumulative, published experience of specialists in the field. An almost universal shortcoming, however, is that habitat descriptions are rarely provided in detail comparable to land cover map descriptions, or describe microhabitat features but do not identify land cover type. For example, "mixed coniferous forests" must be equated with many different vegetation types dominated by coniferous species. Some of these types may occur in climatic zones inappropriate for the species in question. For this reason, association information besides land cover type is necessary to reduce errors of commission in the final list of species predicted to occur at a given site. Also, many species, especially herpetiles, are not strongly associated with vegetation and therefore require other mappable associations to create a predicted distribution map.

As noted, some state and regional institutions have developed wildlife/habitat relationship (WHR) tables and databases. These documents are usually peer-reviewed and predict the presence of a species in a list of simplified habitat types. As is the case with literature accounts, the generalized habitats described in most WHR tables and databases must be matched to the larger number of GAP land cover types. Some WHR databases have been found to have errors which are frequently the result of extrapolation from other regions of the species range (see NM-, OR-GAP). Errors can be identified and corrected most efficiently by extracting the information required for modeling and submitting it to experts for review (see NM-, WY-GAP). A more thorough update can be undertaken with buy-in from the agency managing the WHR database. For RTE species and microhabitat specialists which are to have a modeled distribution, identify the habitat characteristics required that are attributed to other mapped features; for example, species associated with rock outcrops can be mapped to soil polygons attributed as containing rock outcrops. Though GAP currently does not predict habitat viability, attributes of home range or minimum patch size should be recorded when available to support future analysis.

Species with ranges that span physiographic regions frequently exhibit differences in habitat associations between regions. Failure to incorporate these differences in the modeling process can result in substantial errors of omission or commission. We recommend that the state be stratified by ECOMAP sections or subsections (see Keys et. al. 1995) as appropriate for the species, and the habitat associations indicate the differences according to those units (see WA-GAP). This will allow a state to make more accurate distributions based on regional differences and to greatly facilitate distribution edge-matching (or remapping) by ecoregion as adjacent states are completed.

It may be more efficient to collect all available GIS coverages (see Step 5) before conducting this step to narrow the habitat association data that must be collected and to provide literature reviewers or expert consultants with a list of database codes from the coverages to eliminate the need for later translation. Keep in mind, however, that during the timeframe of carrying out this methodology, new GIS coverages may become available. Therefore, it is worthwhile to record attributes of all foreseeably mappable associations, as it is far easier to record these during the literature search rather than going back to the sources later.

The tasks for creating the database of habitat associations are as follows:

a. Identify the GIS coverages available for modeling. Several but not all of the possible coverages are shown in Figure 1 (see Appendix 1). National Wetland Inventory (NWI) data are also helpful, but their availability in digital form is highly variable. Confirm metadata and quality of the coverages and extract a legend from each, so that the habitat associations can be matched to the coverages' thematic descriptions.
b. Collect the habitat association information. If using an existing database, query for the fields applicable to the available coverages and export the file for WHRM development in Step 5. If creating a new database, develop an electronic and paper form for literature reviewers and experts to fill out associations according to the GIS coverages' legend units (see sample format, Appendix 1) (NM-, WY-GAP). In other words, the reviewer can translate between the literature and the coverages directly, which should reduce interpretation errors by others later. It is critical to also provide the source citation for every entry, so that they can be reassessed if problems occur in the modeling phase. The translation will include cross-walking the habitat associations from the

literature to the GAP land cover types. For coverages with user-definable units, such as the digital elevation model (DEM), it is easier to reconfigure the coverage to match the associations. For instance, DEMs can be stratified into any desired contour interval in meters or feet and slope in degrees or percent (see NM-, WY-GAP, see Appendix 1). After normalizing all species' elevation associations to a single unit of measure, e.g., use a 100-meter contour interval to associate all species to elevation, enter the associations into the database.

The value of this database for future GAP updates and the regular use of researchers and managers can be substantial if well planned and documented. Efforts should be made to obtain cooperative assistance and funding from those entities that would benefit from access to such a system. Institutions that traditionally manage this kind of data should be offered the opportunity to become the steward for the database by making regular updates and offering free and open access to it. Step 5: Develop a Wildlife/Habitat Relationship Model (WHRM) for each species, based on available GIS coverages (Scott et al. 1993 and Krohn 1996). The goal is to convert the habitat associations identified in Step 4 to a simply-stated model that can be input into a computer query to integrate the associations with the GIS coverages for the project area. Steps 4 and 5 are better viewed as interacting tasks, but for simplicity they are described separately. The distinction is that Step 4 collects the raw data, while Step 5 processes it into a format readily used (and repeated) in the modeling process. Good planning and database expertise can facilitate combining of these steps to increase efficiency.

To create a map of the species distribution, it is necessary to interface the GIS coverages with a species' habitat association in the form of a query or model. The WHRM can be considered a query made to a spatial database. To establish the WHRM:

a. Extract the associations for each species developed in Step 4 and determine if they are inclusive or exclusive. For example, association with land cover types is inclusive ("include all polygons of x,y, and z types"); whereas association with a distance from perennial streams is exclusive ("exclude all polygons outside of this buffer distance"). Failure to make this distinction will lead to gross errors of commission.
b. Write the WHRM in the form of a database query that will select and exclude polygons according to the associations (see NM-, WY-GAP) and ecoregions (see WA-GAP). For an example query see Appendix 1. Conducting an expert review of the WHRM is recommended to identify any misinterpretations, erroneous extrapolations from other regions, or the need to separate regional differences as described in Step 4 and adjust the model accordingly. Errors in the model may be more obvious after a draft map is produced (Step 6), but generation of draft maps can easily require 1000 hours of CPU time. Therefore, if expert review of the models is not conducted, the sample tests in Step 6 are critical to avoiding unnecessary iterations of the entire set of species maps.

Step 6: Integrate the range limits and habitat associations into a draft predicted species distribution map. The goal is to achieve a polygon map based on the appropriate habitat association, as limited (but not truncated) by the range limits. In other words, there must be consistency in the type of boundaries of the species' distribution; if they are derived from maps of physiographic elements, they should not also contain random boundaries from counties or coordinate grids. This can pose some challenges to the GIS process and substantially complicate the query development described in Step 5. Further discussion on these challenges and possible solutions are found below. There are many different approaches to accomplishing this step through GIS and database query techniques; the following tasks are provided as an example and were incorporated from a number of successful approaches used by GAP projects. For a substantially different but successful approach to distribution modeling, see WA-GAP.

a. Prepare the GIS coverages for query by the WHRM. For example, generate hydrologically defined buffers (see Appendix 1) based on known associations, and create an elevation map from the DEM based on the finest minimum contour interval needed (see NM-, WY-GAP).
b. Intersect all coverages to be used to predict the distribution, including the ecoregion boundaries if

applicable. This intersected coverage is called a hypercoverage (Garber, pers. comm.) and will be very large and complex. It may be necessary to tile the hypercoverage if its complexity exceeds the GIS software limitations (see NM-GAP). As of this writing, some projects were experimenting with raster-formatted data and are confident that the process will run more efficiently in that mode (see NM-, MT-, TX-GAP). Raster format has the advantages of retaining the same number and location of spatial records regardless of the number of GIS coverages intersected, simpler generation of graphic output, and filtering operations for polygon size and connectivity. Additionally, the equal area unit may prove valuable for other analyses (see Analysis section, version 2).

c. Export the database file of the hypercoverage to a database program capable of supporting a matrix with polygon or raster ID number on the y-axis, attributes on the x-axis, and attribute values in the body (see NM-GAP).

d. Run the WHRM queries. Tests of small groups of species are encouraged for each modeling step to ensure that the source coverages were coded properly, the hypercoverage is functioning properly, and the queries were written properly. It is recommended that one or more species from each taxonomic group is selected for testing and that species are included that will test every coverage incorporated in the process. The query will establish the presence/absence of each species in each polygon or raster. This process will attribute the hypercoverage with presence/absence per species per spatial record (of the polygon or raster) in the database. The newly attributed database can then be linked back to the hypercoverage spatial data to generate a map. Polygons should also take on the attributes from the range units (see step 4) so that species maps can be drawn and queried according to levels of confidence (see WA-GAP, see Appendix 1). For example, some researchers may only want to generate species' distributions based on known occurrence, and some may want a more liberal prediction.

e. Generate graphic files of individual animal distribution for review and/or provide prints to experts, or set up a password-protected Internet site for reviewers to access (see Appendix 1). The password protection is recommended to discourage undesired access and distribution of draft maps.

Some additional notes on WHRM and map development: The query must be able to extrapolate the polygons of the association coverages beyond the boundaries of the range extent. Otherwise, the incompatible boundaries of the range extent (such as county, degree blocks, hexagons, etc.) become the outside boundary with a fine, patchy interior boundary. Extrapolation outside the range extent to achieve consistency in boundary types must be carefully controlled to avoid gross errors of commission. This has been especially true for early GAP projects that used land cover maps with very large polygons, and simple extrapolation of an entire polygon intersecting the range extent was undesirable. This will generally not be a problem as the land cover MMU decreases, thematic detail increases, and additional filters are incorporated. As of this writing however, the above method (include all selected polygons that fall inside or intersect with the range extent) is the only one tested. If this method proves unsatisfactory (produces unacceptable extrapolation), the following ideas may be tested and applied to the species (all species, groups, or individuals):

- Set a random buffer distance beyond the range extent for the purpose of including whole habitat patches that intersect with the range extent. The query should select only those polygons that fall completely within that buffer.
- Use a physiographic boundary such as watersheds as the outside limit with the same function as the random buffer above.
- After these automated methods are used, those few species that still end up with unacceptably artificial truncations can be reviewed for alternative solutions, including hand digitizing a limit that can be biologically defended. This should only occur for "generalist" species found in land cover types with typically large polygons such as grasslands or shrub-steppe.

Another issue is the potential incompatibility between the range unit size and the hypercoverage MMU (or simply consider the land cover MMU). As the land cover MMU decreases, it is likely that simple extrapolation of the hypercoverage polygons that intersect with the range unit boundary will produce a jagged straight line. This undesirable effect has been noticed for projects intersecting 100-ha MMU land cover maps with county boundaries (see NM-GAP). The best solution has been to adopt a smaller range unit such as the EMAP hexagon, but as land cover MMU decreases to the 2-ha size, even the hexagon may prove too coarse. One option is to tessellate the hexagon into smaller units. For second generation GAP projects (those having previously completed animal distribution maps), the use of the first generation distribution maps as the range extent for the next generation should eliminate incompatibility between range unit size and land cover MMU. This would only be acceptable if the first generation maps were assessed (see Vertebrate Distribution Assessment chapter), or the distributions underwent expert review for the purpose of depicting range extent as described in Step 3.

This methodology will likely produce very small patches that would be desirable to filter out. Identification of such patches would preferably be accomplished through viability analysis, but patch-size viability is rarely addressed in the species literature and, without specific data, would be a subjective decision. Therefore, we do not require any steps to address this issue. In other words, let the maps reflect the process used to create them. The existence of these patches should be addressed in your data limitations description and resolution of the issue, for now, should occur in the data application. We do, however, encourage further research and experimentation to establish simple, biologically defensible solutions. One recommendation is to use species' body size and/or home range to establish a minimum polygon size. These polygons can then be "eliminated" in vector maps or filtered out in raster maps.

There are many considerations when dealing with rare species, and it is difficult to establish a single set of rules to determine what method to use to map those species. GAP requires a spatial map for all terrestrial vertebrates as described in Step 1. This map may be either a predicted distribution as established through the standard methodology presented here, or it may be a generalized map of the known point locations as obtained from the state NHP, Conservation Data Center (CDC), or other agency responsible for such data. The degree of generalization required to allow public distribution of the map will be left to negotiation between the GAP project and the data owner, but we recommend a random unit such as the USGS 7.5 minute quad versus a circular unit with the point as a centroid. This is to prevent the easy estimation of the actual point coordinates by users who should more appropriately obtain that specific information from the data owner. When generalized maps are created, the actual point data should still be used in the Analysis phase, as they will be more precise in the estimation of management status, understanding that area measurements will not be possible or appropriate in either case. The decision as to which type of map should be generated will best be made through discussion between the GAP project and the data owner. If the data owner has high confidence that the points represent the total distribution (or large majority) of the species, the generalized points should be used as they will be more precise and defensible than a predicted distribution. If there is not a high level of confidence, or if very little is known about the species' distribution, a map that predicts its distribution should be developed. In either case, it is important to document all decisions and any caveats to the data user for each rare species. For the generalized point-derived map, the basic caveat is that it is a generalized map of the known locations for the purpose of protecting those locations, and that the data owner should be contacted regarding distribution of the original data. For predicted distributions of rare species, specific caveats would be useful, indicating the likely discrepancy between the prediction and true occurrence, e.g. current collections and studies are insufficient to establish any confidence in the known occurrence, or the species' life history does not result in repeated regular occurrence in the same location, etc.

<u>Step 7</u>: Expert review of draft maps and final mapping. The goal is to identify more specific problems in the model than improper range extent or basic habitat association that may require addition or deletion of a coverage from a species' WHRM, or ecoregion changes to the model. Some reviewers may not be able to provide a useful review of these spatially detailed maps, as it will likely be the first time they have seen their species of interest depicted this way. For this reason, this step is not required for all species,

but it has been useful for GAP projects and is strongly encouraged. It is helpful to reviewers to provide overlays of range maps from the literature and for more difficult species, overlays of particularly important GIS coverages with which they are associated.

Correct the WHRMs as needed after expert review and rerun the queries. Dissolve the final hypercoverage to eliminate all unnecessary boundaries. Retain species presence/absence and range confidence (known or predicted occurrence) attributes to create a hyperdistribution coverage. This step merely reduces the spatial complexity of the map by removing boundaries separating adjacent and identically attributed polygons. It is unnecessary for raster data, though for this format, some filtering operations may be desirable to eliminate "salt-and-pepper" cells as with the land cover map (see Land Cover Mapping chapter).

Step 8: Prepare the distributions for comparison to ownership and management (see Analysis section). Intersect the hyperdistribution coverage with the stewardship coverage (see Land Stewardship chapter) to produce a database with both attributes. Conduct the analysis as described in that section of the Handbook.

Step 9: Prepare an individual coverage for each species. Use the TNC element code as the coverage name. The goal is to allow easier access to distribution maps of individual species for the generation of GAP products and especially for the users of GAP data. This is accomplished by selecting the polygons from the hyperdistribution coverage attributed for each species and creating a new coverage from those selected units. Creation of an individual coverage can also facilitate additional investigations of patch size, proximity, viability, etc.

Accuracy Assessment: This task is described in an individual chapter in progress. It is expected to be available by January 1997.

Edge-matching: GAP encourages but does not require individual projects to conduct animal distribution edge-matching with neighboring projects (refer to your project contract for specific requirements). Testing of methods is being conducted in the northwest region which will hopefully lead to standard recommendations. For the purpose of stimulating development of methods for edge-matching, the following untested suggestions are provided:

a. Join the individual species predictions from each project and print the resulting map at a suitable small scale (e.g., 11 x 17) and identify areas of misalignment. Species with very minor misalignment may be able to be edge-matched through raster-smoothing operations. Some apparent misalignment may be from actual habitat conditions caused by differences in land use policies. These are often obvious in the satellite imagery and land cover map which should be reviewed prior to pursuing additional corrective steps. Otherwise, proceed with remodeling as follows:
b. Establish a new range extent for the total project area by hexagon; at a minimum by intersecting the hexagon grid with the individual GAP-predicted species distributions. Compare the hexagon ranges and identify apparent mismatches between adjoining states. Submit the hexagon ranges and predicted distributions for expert review to create a hexagon range map across the state boundary using all practically available information of occurrence and species checklists to justify.
c. Compare WHRMs used in each state and identify any differences, for instance in elevation limits or vegetation type associations. Check these associations against the cited literature or wildlife database to identify errors or regional differences not already incorporated and correct. Develop new WHRMs for the entire project area or per ecoregion.
d. Edge-match all coverages to be used in remodeling.
e. Remodel per steps identified previously in this chapter.

Limitations

GAP projects will report specific limitations of their data; the following is a description of the limitations of the method described here.

The methods followed by GAP make use of the best available information to create meso-scale maps of predicted species distribution, but rarely involve the collection of new specimens or field observations. Because locational records of collection and observation of species are grossly incomplete, GAP relies strongly on the use of expert opinion in the development of range limits, habitat association, and review of predicted distributions. The resulting maps then are testable hypotheses that we encourage field biologists to assess whenever conducting field studies.

These procedures work best for species with habitat preferences that can be described in terms of land cover and other digitally mapped features or characteristics. It provides less accurate predictions of the presence of species that are highly variable in spatial and/or temporal occurrence (Krohn 1996), and it will work for habitat specialists only if their specific habitat requirements are available as mapped features or are well associated with other mapped characteristics such as land cover types.

An additional caution is that species with very restricted distributions (occurring in one or a few locations of a size below the GAP MMU) *cannot* be predicted to occur in seemingly appropriate habitat within their distributional limits. Because of their rarity, these species are often the subject of special attention from state and federal resource agencies and are ranked G1-G2 by the Natural Heritage Central Databases (see ranks and their description at http://www.heritage.tnc.org). The specific locations where they are known to occur are usually tracked by NHPs, CDCs, and the USFWS. GAP makes use of the data from NHPs and CDCs to report the presence of populations of such species within a geographic unit. For security purposes, the exact locations of these populations should be distributed only by the data owner.

Continuing Research

Ongoing research efforts include:

- Accuracy assessment at varying levels of thematic and spatial resolution (J.M. Scott, pers. comm.) and analysis of life history and behavioral traits (nocturnal, diurnal, home range, size, abundance, conspicuousness, etc.) for species that have been found to be accurately predicted by GAP maps and those that have not, in an attempt to identify mappable habitat features that might be used to improve reliability of the GAP maps (Krohn 1996).
- Future research efforts will focus on linking GAP maps of presence/absence with maps of abundance and reproductive success in an attempt to identify source/sink habitats. Additionally, the maps of single species presence/absence will be used as testable hypotheses, and field studies are planned to test that assumption.
- Additional studies will use GAP maps as a sampling universe from which unbiased samples will be taken to estimate species abundance and reproductive success over the entire range of the species. Additionally, the importance of minimum mapping unit (2-200 ha) in predicting occurrence of a species will be tested in field studies.
- TNC and the NHP science programs plan to relate vertebrate species to vegetation alliances in their databases in cooperation with GAP (L. Master, pers. comm.). This is a first step to building regional and national databases of species habitat associations, collection and observation records, range extent, and predicted distributions.

Literature Cited

American Ornithologist's Union. 1983. Check-list of North American birds, 6th ed. Allen Press, Lawrence, Kansas. 877 pp. (as modified by any subsequent supplements and corrections, the latest published July '95 in the *Auk* 112:819-830).

Butterfield, B. R., B. Csuti, and J. M. Scott. 1994. Modeling vertebrate distributions for Gap Analysis. Pages 53-68 in Mapping the diversity of nature (R. I. Miller, ed.). Chapman and Hall, London.

Collins, J. T., et al. 1990. Standard common and current scientific names for North American amphibians and reptiles, 3rd ed. Society for the Study of Amphibians and Reptiles. Herp. Circ. No. 19. 41 pp.

Csuti, B., and J. M. Scott. 1991. Mapping wildlife diversity for Gap Analysis. *Western Wildlands* 17(3):13-18.

Edwards, T.C., Jr., J.M. Scott, C.G. Homer, and R.D. Ramsey. 1993. Gap Analysis: A geographic approach for assessing national biological diversity. *Natural Resources and Environmental Issues* 2:65-71.

Forman, R. T. T., and M. Godron. 1986. Landscape ecology. John Wiley & Sons, New York.

Gaines, S. D. and W. R. Rice. 1990. Analysis of biological data when there are ordered expectations. *American Naturalist* 135:310-317.

Hall, E. R. 1981. The mammals of North America, 2nd Edition, Vols. 1 & 2. John Wiley & Sons, New York.

Hoffmeister, D. F. 1986. Mammals of Arizona. The University of Arizona Press, Tucson, and the Arizona Department of Game and Fish, Phoenix.

Jones, J. K., Jr., et al. 1992. Revised checklist of North American mammals north of Mexico, 1991. Occ. Pap. Mus. Texas Tech Univ., Lubbock, TX. 23 pp.

Keys, Jr., J., Carpenter, C., Hooks, S., Koenig, F., McNab, W. H., Russell, W., and Smith, M.L. 1995. Ecological units of the eastern United States– First approximation. USDA Forest Service, Atlanta, GA.

Krohn, W.B. 1996. Predicted vertebrate distributions from Gap Analysis: Considerations in the designs of statewide accuracy assessments. Pages 153-168 in J.M. Scott, T.H. Tear, and F.W. Davis (editors). Gap Analysis– A landscape approach to biodiversity planning. American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

Maser, C., J. W. Thomas, and R. G. Anderson. 1984. Wildlife habitats in managed rangelands–the Great Basin of southeastern Oregon: the relationship of terrestrial vertebrates to plant communities, Part 1, Text. U.S.D.A. Forest Service, General Technical Report PNW-172.

Mayer, K. E., and W. F. Laudenslayer, Jr. 1988. A guide to wildlife habitats of California. California Department of Forestry and Fire Protection, Sacramento.

Mittermeier, R. A., and A. B. Forsyth. 1994. Setting priorities for biodiversity conservation. Pp. 125 in Principles of conservation biology (G. K. Meffee and C. R. Carroll, eds.). Sinauer Associates, Inc., Sunderland, Massachusetts.

Morrison, M. L., B. G. Marcot, and R. M. Mannan. 1992. Wildlife-habitat relationships: concepts and applications. The University of Wisconsin Press, Madison.

Noss, R. F. and B. Csuti. 1994. Habitat fragmentation. Pp. 237-264 in Principles of conservation biology (G. K. Meffee and C. R. Carroll, eds.). Sinauer Associates, Inc., Sunderland, Massachusetts.

O'Neil, T. A., R. J. Steidl, W. D. Edge, and B. Csuti. 1995. Using wildlife to classify vegetation communities for conserving biodiversity. *Conservation Biology* 9:1482-1491.

Scott, J. M., B. Csuti, and F. Davis. 1991. Gap analysis: an application of geographic information systems for wildlife species. Pp. 167-180 in Challenges in the conservation of biological resoures: A practitioner's guide (D. J. Decker, M. E. Krasny, G. R. Goff, C. R. Smith, and D. W. Gross, eds.). Westview Press, Boulder, Colorado.

Scott, J. M., B. Csuti, D. Stoms, and F. Davis. 1992. Remote sensing for nongame wildlife habitat management. Trans. 56th North American Widlife and Natural Resources Conference (1991):134-140.

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* No. 123:1-41.

Scott, J. M., S. Mountainspring, F. L. Ramsey, and C. B. Kepler. 1986. Forest bird communities of the Hawaiian islands: Their dynamics, ecology, and conservation. Studies in Avian Biology, No. 9. Cooper Ornithological Society, Lawrence, Kansas, 431 pp.

Thomas, J. W. 1979. Wildlife habitats in managed forests: The Blue Mountains of Oregon and Washington. USDA Forest Service, Agriculture Handbook No. 553, Washington, D.C.

Udvardy, M. D. F. 1969. Dynamic zoogeography: with special reference to land animals. Van Nostrand Reinhold Company, New York.

White, D., A. J. Kimerling, and W. S. Overton. 1992. Cartographic and geometric components of a global sampling design for environmental monitoring. *Cartography and Geographic Information Systems* 19:5-22.

Appendix 1 - Building Hexagon Range Maps

From WY-GAP Final Report (Merrill et al. 1996)

3.2.2 Occurrence of Species within Hexagons

Counties and latilongs are common units used to document the general location of species. Wyoming consists of 23 counties (average size = 10,950 km^2) and 28 latilong blocks (average size = 9,004 km^2). Using either of these geographic units to make species predictions would have overestimated distributions of species in cases where a species' range extended only partly into a county or latilong. To reduce this problem, we mapped the distributional limits of species using smaller, hexagon units (635 km^2) which are part of a global hexagonal grid system developed by the EPA (White et al. 1992). Advantages to using the hexagon grid include its equal area sampling structure, its independence from political and administrative boundaries (resulting in more consistent mapping of animal distributions), and its hierarchical structure which can facilitate increasing or decreasing grid densities in future analyses (White et al. 1992).

Species were recorded within each of the 436 hexagons for Wyoming using 1 of 7 definitions (Table 3.1). We adopted the first 3 definitions of species occurrence from the Biodiversity Research Consortium (Master et al. 1995), which is a complementary effort coordinated by EPA's Habitat/Biodiversity Program whose objective is to identify areas of the country where risks to biodiversity are greatest. The remaining 4 definitions (Table 3.1) were developed to enhance the species-hexagon database and are shown as part of the vertebrate species maps (Merrill et al. 1996b). We used only the data classified in the first 4 categories to conduct our gap analysis. Statement of probabilities in these descriptors were used as guidelines to subjectively qualify the occurrence of a species within a hexagon consistent with the descriptions in Table 3.1. At this time, they do not represent a quantified analysis of the probability of occurrence. Future refinements to the database may allow a quantified probability statement of species occurrence.

Three primary sources of information were used to document the occurrence (or expected occurrence) of a species within a hexagon: (1) species locality records, (2) published range maps, and (3) the opinions of experts. Species locality records (i.e., recorded occurrences of observed, trapped, or killed individuals) were obtained from 16 existing wildlife databases collected from state and federal agencies, conservation groups, museums, and outdoor science schools in Wyoming (Table 3.2). Fifteen of the species databases were non-spatial, tabular databases which included Public Land Survey System (PLSS) descriptions or coordinates for the location of observed species. PLSS locational descriptions were converted to latitude-longitude coordinates for import into Arc/Info using a fortran program called TR-LL (Morgan and McNellis 1965). Hexagons encompassing locality records with a date $\geq$ 1950 were coded as Confirmed, while those populated with locality records < 1950 were coded as Historical. Historical hexagons that were immediately adjacent to other hexagons coded as Confirmed, Probable, or Possible, were initially included within a species' current distribution. In cases where the historical hexagon was geographically isolated from a species' contiguous range, the hexagon was initially excluded from the species' current distribution, but was not removed from the species-hexagon range maps. Later, when expert reviewers examined the maps (see below), they were given the chance to modify historical records as necessary.

Table 3.1. Categories used to qualify species occurrence within hexagons used to predict vertebrate species distributions. The first three definitions were adopted from the Biodiversity Research Consortium (Master et al. 1995). The remaining four definitions were developed for use in Wyoming Gap Analysis.

CONFIRMED (C)	The species is confidently assumed (> 95% certain) or known to occur in the hexagon. Information sources confirming occurrence within a hexagon included species locality records and expert opinion.	
PREDICTED (PR)	The species is predicted to occur in the hexagon based on the "fact-pattern" (i.e., presence of suitable habitat or conditions and historical record and/or presence in adjacent hexagons[s]); at least 80% certain that the species occurs in the hexagon. Information sources used to document a species within a hexagon included expert opinion only.	
POSSIBLE (PO)	The species possibly or potentially occurs in the hexagon; its estimated likelihood of occurrence in the hexagon is thought to be between 80% and 10% (or less for extremely rare species where suitable habitat or conditions may be present). Information sources used to document a species as Possible within a hexagon included expert opinion and published range maps.	
HISTORICAL (H)	The species is confidently assumed (> 95% certain) or known to have occurred in the hexagon prior to 1950. The historical presence within the hexagon was included as part of the species' current distribution. Information sources used to document a species as historical (included) within a hexagon included species locality records and expert opinion.	(Included)
HISTORICAL (Hx)	The species is confidently assumed (> 95% certain) or known to have occurred in the hexagon prior to 1950. The historical presence within the hexagon was not included as part of the species' current distribution. Information sources used to document a species as historical (excluded) within a hexagon included species locality records and expert opinion.	(Excluded)
QUESTIONABLE (?)	The occurrence of the species within a hexagon was still in question after having been reviewed by experts. Hexagons coded as questionable were not included as part of the species' current distribution. Information sources used to document a species as questionable within a hexagon included expert opinion only.	(Excluded)
EXCLUDED (X)	The documented occurrence of a species was excluded by expert review after once having been coded as confirmed, predicted, or possible. Information sources used to document a species as excluded within a hexagon included expert opinion only.	

Range maps published by Clark and Stromberg (1987) and Baxter and Stone (1985) also were used to document the occurrence of species within hexagons for mammal and herptile species. Wyoming-specific range maps for birds did not exist. For mammals and herptiles, the geographic range of each species was manually transferred from paper maps to the computerized hexagon grid using a mouse to select the hexagons which overlapped with range map polygons. Hexagons populated in this manner were coded as Possible.

Table 3.2. Databases used to document species occurrence within hexagons.

Database	Source of Database	No. of Records	Date of Acquisition
Wildlife Observation System*	Wyoming Game & Fish Department	666,567	5/92
Element occurrence Database*	Wyoming Natural Diversity Database	2,880	7/94
Vertebrate Museum Database	Museum Databases	4,389	6/93
Wildlife Observation Database	Grand Teton National Park	6,668	3/92
Devils Tower Fauna Database	Devils Tower National Monument	199	4/92
Green River Sage Lek Database	BLM -Green River Resource Area	128	9/92
Green River Raptor Database	BLM -Green River Resource Area	1,577	9/92
Lander Raptor Database	BLM -Lander Resource Area	162	3/92
Kemmerer Raptor Database	BLM -Kemmerer Resource Area	125	2/92
Cody Raptor Database	BLM -Cody Resource Area	1,060	7/92
Cody Nongame Bird Database	BLM -Cody Resource Area	225	7/92

Grizzly Bear Database	NPS -Interagency Study Team	9,338	3/92
M.A.P.S Database	Teton Science School	332	10/92
Amphibian Survey Database	Teton Science School	35	10/92
Wind River Wildlife Database	U.S. Fish & Wildlife Service	2,775	3/93
Great Divide RA Raptor Database	BLM - Great Divide Resource Area	3,266	3/93

* Includes additional records from 1994 or 1995 for specific areas and/or taxonomic groups

Species-hexagon range maps developed from locality records and published range maps were reviewed by over 60 acknowledged experts consisting of federal and state biologists, university professors, and Audubon Society members (Appendix 3.1). Reviewers were asked to check, and if necessary, correct the hexagon occurrences that were based on questionable locality records or range maps. Reviewers were also given the opportunity to add animal occurrences within hexagons using the definitions in Table 3.1. The 1994 review of the species-hexagon range maps represented the first of two distinct map reviews.

Maps of species richness within hexagons were derived by totaling the number of species documented/expected to occur within hexagons and do not reflect species distributions modeled using habitat associations. For this analysis, we used only species occurrences which qualified as one of the first four definitions in Table 3.1. The five categories of species richness identified in the maps were determined using an equal-interval classification.

In developing the database for species distributions for Wyoming, we did not differentiate between breeding and winter ranges for bird species. Seasonal information for birds existed only by latilong blocks and interpolation of breeding ranges to the hexagon level within these larger units would have represented an unreasonable refinement of scale. The refinement to seasonal ranges also would have complicated the review process beyond reasonable time demands of the reviewers since most bird reviewers reviewed all 291 bird distribution maps. Further, the conservation of bird species must consider the maintenance of habitat throughout the year (Csuti 1996). Future refinements to the bird distribution maps should separate breeding and wintering ranges and incorporate new information on seasonal habitat use by individual bird species.

Appendix 2 - Sample Data Use Agreement

The pertinent language in the agreement to allow a Gap Analysis Project to use and distribute locational data for RTE species should contain:

1. The project will be able to retain a copy of the actual coordinate data for the duration of the project for the purpose of modeling and assessing models of the species.
2. The data will be stored in a secured location that is inaccessible except to project research staff.
3. The project may not transfer or distribute the data to anyone not identified in this agreement.
4. The project may distribute the data set after it has been altered such that the species locations may not be identified at a finer unit than (e.g., Township, or 7.5 minute USGS quad), and shall not occur as a center point of the unit such that the exact location could be determined through geometric calculation.
5. Upon project conclusion this agreement may be extended at the option of the data owner, otherwise the data set must be returned and any copies destroyed.

Appendix 5 - Species Associations Data Forms and Queries

This appendix contains 4 example algorithms to illustrate an amphibian, reptile, bird, and mammal as well as the spectrum of simplicity or complexity of the predictive algorithms.

Also, the 3 reference lists (herpetofauna, birds, mammals) included in this appendix contain number-coded entries that match with codes included in the electronic files of complete algorithms for all species. The number codes for information source references in the appendix lists are not necessarily sequential because of removal of some references during final editing of algorithms to remove unnecessary information or redundancies. Original numbers were retained for all references to maintain consistency with codes in 584 algorithms. Reference codes are always in parentheses in the algorithms. Electronic files of all algorithms are not included in this report; they are accessible as described in Appendix F and the metadata section in Chapter 7.

Field names on the algorithms are:

BISON_CODE = Code number from Biota Information System of New Mexico.
SPEC_NAME = Scientific name in standard listing from The Nature Conservancy.
MOUNTAIN = Codes for associations with Mountain Range coverage.
HYDRO = Codes for associations with Watershed coverage.
VEG_TYPE = Codes for associations with Land Cover classes coverage.
TEMP = Not used.
PRECIP = Not used.
ELEV_M = Elevation range limits.
SLOPE _ASP = Slope limitation (Aspect originally considered but not used).
WATER = Text describing association with perennial or intermittent water and pertinent buffer distances.
SOIL = Codes for associations with soil classes derived from STATSGO data.
EVOLUTION = Text describing changes in the algorithm during review.
TEMP_MIN = Minimum winter temperature coding (ultimately not used).
TEMP_MAX = Maximum summer temperature coding (ultimately not used).
ELEV_M_MIN = Minimum elevation (m).
ELEV_M_MAX = Maximum elevation (m).
SOIL FILE = Codes to translate soil association data to polygon selection process.
WATER EXP. = dBase expression used to select water hypercoverage polygons during distribution prediction.
POLY _SEL = dBase expression that selected qualifying cells in the land cover hypercoverage files to predict distribution.

Sample Species Algorithm #1

```
BISON_CODE 020080
SPEC_NAME Spea multiplicata
```

MOUNTAIN
HYDRO   11020010,11040001,11040002,11080001,11080002,11080003,11080004,
11080005,11080006,11080007,11080008,11090101,11090102,11090103,
11090104,11100101,11120101,11120102,12050001,12050002,12050005,
12080001,12080003,12080004,12080006,13010002,13010005,13020101
13020102,13020201,13020202,13020203,13020204,13020205,13020206,
13020207,13020208,13020209,13020210,13020211,13030101,13030102,
13030103,13030201,13030202,13050001,13050002,13050003,13050004,
13060001,13060002,13060003,13060004,13060005,13060006,13060007
13060008,13060009,13060010,13060011,13070001,13070002,13070007,
14080101,14080103,14080104,14080105,14080106,14080107,14080201,
14080204,15020001,15020002,15020003,15020004,15020006,15040001,
15040002,15040003,15040004,15040006,15080302,15080303, VEG_TYPE
3121,3122,5121(48),5220,5221,5222,2121,2122,9420(31),4131(3),
9110(1,48),9120(1,48),3211,2211,5211(56),5212(56),4211(56), 4212(56)
TEMP
PRECIP
ELEV_M 900-2600m (48)
SLOPE_ASP <10% (54)
WATER
SOIL
EVOLUTION
TEMP_MIN -99
TEMP_MAX -99
ELEV_M_MIN 900
ELEV_M_MAX 2600
SOIL FILE
WATER EXP.

POLY_SEL (hydro_code$trim(spalg->hydro) .or.
hydro_code$trim(spalg->hydro2) .or.
hydro_code$trim(spalg->hydro3)) .and.
vegcode$spalg->veg_type .and. contour>
800.and.contour<2600

Sample Species Algorithm #2

BISON_CODE 030235
SPEC_NAME Leptotyphlops humilis
MOUNTAIN
HYDRO   13020203,13020204,13020208,13020209,13020210,13020211,13030101,
13030102,13030103,13030201,13030202,13050001,13060011,13070001,
13070002,15040001,15040002,15040003,15040004,15040006,15080302, 15080303
VEG_TYPE 3211,3222,4110,4121,4211,4212,4220,4221,4222,5220,5221,5222,
6120(51)
TEMP
PRECIP
ELEV_M 900-1500m(48)
SLOPE_ASP
WATER
SOIL
EVOLUTION Eliminated hydro units 13050004, 13050003,
13070007, 13060007, 13060009, 13060010
TEMP_MIN -99
TEMP_MAX -99
ELEV_M_MIN 900
ELEV_M_MAX 1500
SOIL FILE
WATER EXP.

POLY_SEL (hydro_code$trim(spalg->hydro)) .and.
vegcode$spalg->veg_type .and. contour>
800.and.contour<1500

Sample Species Algorithm #3

BISON_CODE 040050
SPEC_NAME Agelaius phoeniceus
MOUNTAIN
HYDRO
VEG_TYPE 6120,9110,9120,6211,9410(2,107),6110(107),9310(107),9320(107) TEMP
PRECIP
ELEV_M
SLOPE_ASP
WATER   permanent/intermittent water w/50m buffe
SOIL
EVOLUTION Counties 001, 003, 005, 007, 011, 013, 015, 017,
021, 023, 025, 028, 029, 031, 033, 035, 037, 039, 041, 043, 045, 047, 049,
053, 055, 059, 061 from BISON-M (Reference # 1). Put in all counties per
expert review (106).
Eliminated vegetation types 4220, 4221, 4222, 5110, 5121, 5122, 9320 (107, 108).
Eliminated vegetation types 5212 (117).
TEMP_MIN -99
TEMP_MAX -99
ELEV_M_MIN -99
ELEV_M_MAX -99
SOIL FILE
WATER EXP. (p50=1 .or. i50=1)

```
  POLY_SEL vegcode$spalg->veg_type
```

Sample Species Algorithm #4

```
BISON_CODE 050115
  SPEC_NAME Castor canadensis
  MOUNTAIN
  HYDRO   11040001,11080002(1,12),11080003(1,36),11080004,11080006,11080007,
  13020101(1,36),13020102(1,36),13020201,13020202(1,36),13020203, 13020204
  13020211,13030101,13030102(1,36),13060001(1,36),
  13060003(1,36),13060007,13060008(1,36),13060011,13070002 14080101,
  14080104(1,36),14080105,14080107,14080201,15020004,15040001(1,36),
  15040002,15040004(1,36)
  VEG_TYPE 6110(1,90,107,108),6120,6211,9410(1,29,107,108) TEMP
  PRECIP
  ELEV_M
  SLOPE_ASP <15%(19)
  WATER   Permanent water w/100m buffer
  SOIL
  EVOLUTION Eliminated hydro units 11020010, 11040002,
  11080001, 11080005, 11080008, 11090101, 11090102, 11090103, 11090104,
  11100101, 11120101, 12050001, 12080001, 12080003, 13010002, 13010005,
  13020205, 13020206, 13020207, 13020208, 13020209, 13020210, 13030103,
  13030201, 13030202, 13050001, 13050002, 13050003, 13050004, 13060002,
  13060004, 13060005, 13060006, 13060009, 13060010, 13070007, 14080103,
  14080106, 14080204, 15020001, 15020002, 15020003, 15020006, 15040003 (All
  95, 96).
  TEMP_MIN -99
  TEMP_MAX -99
  ELEV_M_MIN -99
  ELEV_M_MAX -99
  SOIL FILE
  WATER EXP. p100=1

  POLY_SEL (hydro_code$trim(spalg->hydro) .or.
  hydro_code$trim(spalg->hydro2) .or.
  hydro_code$trim(spalg->hydro3)) .and.
  vegcode$spalg->veg_type
```

WA-GAP Wildlife Habitat Relation Model by Ecoregion

Eco-Region	Veg. Zone	Land Cover	AMGR	AMMA	AMTI	ASTR	BUBO	BUWO
			Species Codes					
6	40	960	4	4	0	0	4	0
6	40	962	4	0	0	0	4	0
6	40	972	4	0	0	4	4	0
6	40	982	4	0	0	4	4	0
6	82	411	2	5	0	0	5	0
6	82	524	2	5	0	0	5	0
6	82	534	2	5	0	0	5	0
6	82	610	0	7	0	0	7	0
6	82	612	0	7	0	0	7	0
6	82	620	0	0	0	0	7	0
6	82	621	0	0	0	0	7	0
6	82	624	0	0	0	0	7	0
6	82	625	0	0	0	0	7	0

Table 5. A subset of the amphibian habitat association modeling matrix (Northwestern Salamander, Long-toed Salamander, Tiger Salamander, Tailed Frog, Western Toad and Woodhouse's Toad). Ecoregion 6 is Outer Olympic Peninsula, and Vegetation Zones 40 and 82 are Silver Fir and Mountain Hemlock, respectively. Land Cover codes are a series of 3 numbers (Table 3). These codes are the labels of land cover polygons occurring in ecoregion 6 and in Vegetation Zones 40 and 82: 960, 962, 972, 982 are conifer forests (9), early (6), mid (7) and late (8) seral, with open and closed (0) and closed (2) canopy. Lakes (411) are open water. Wetlands (5) include freshwater marshes (2) and riparian areas (3), both surrounded by conifer trees (4). Non-forested classes (6) were disturbed (1) or non-disturbed (2), sparsely vegetated (1), grasslands (2), shrublands (4), tree savannas (5) or mixed types (0).

Assignment of habitat quality codes was obviously subjective. Models, however, based on presence alone imply that habitats of different quality are equally important in management. Lack of confidence in assigned codes was more likely to be based on lack of knowledge rather than the subjective nature of code assignment. We went through the process of eliminating habitats that we felt confident would not be used, then began a review process of the remaining habitats. Amphibians were primarily assigned suitability codes indicating that their occurrence was contingent upon the availability of suitable microhabitats, if there was a possibility that suitable habitats were present below our minimum mapping unit.

Methods for Assessing Accuracy of Animal Distribution Maps

Blair Csuti and Patrick Crist [1]

Adapted from version 1 by

Kelly Cassidy, Edward O. Garton, William B. Krohn, L. Scott Mills, J. Michael Scott, Ken Williams, and Blair Csuti

1 Gap Analysis Program, University of Idaho, Moscow, ID

Introduction

Assessing the accuracy of GAP-predicted vertebrate distribution maps is subject to many of the same problems as assessing land cover maps, but involves additional serious challenges. It is, however, necessary to provide some measure of confidence in the results of the gap analysis (comparison of element s distribution with stewardship and management) for each species, and to allow users to judge the suitability of the distribution maps for their own uses. We, therefore, feel it is important to provide users with a statement about the accuracy of GAP-predicted vertebrate distributions in relation to both spatial and temporal variables. Here we provide some guidelines for assessing the reliability of GAP vertebrate distribution maps. We acknowledge that distribution maps are never finished products, but are continually updated as new information is gathered. However, we feel that assessing the accuracy of their current iteration provides useful information about their reliability to potential users and the suitability of the modeling process for particular species. We realize that not all projects have the resources to carry out complete accuracy assessments, and we encourage collaboration between projects to test accuracy on a regional rather than state basis.

We especially encourage wildlife biologists to treat the predicted distributions as testable hypotheses and engage in the process of validation and iterative modeling. Our goal is to produce maps that predict distribution of terrestrial vertebrates and, from that, total species richness and species content of selected areas with an accuracy of 80% or higher. Failure to achieve this accuracy indicates the need to refine the data sets and models used for predicting distribution (see the Vertebrate Modeling chapter). For some species, no modeling method may be suitable for the project area and available digital data.

Objectives

1. To provide a sense of the predicted distribution accuracy for individual species and groups of

species based on comparison to valid checklists for geographic areas and/or independent field observations.

2. To provide a statistically valid measure of accuracy for selected individual species and groups of species for geographic areas based on field investigation when adequate resources are available.
3. To accumulate information on field-verified distributions of species based on research by other agencies and organizations and, as these independent data accumulate, use them to assess the accuracy of GAP species distribution maps.
4. Identify weaknesses in the available data and modeling methods to benefit future iterations.

Background

While GAP vertebrate distribution maps can be useful planning tools, they represent a prediction at one point in time, and it is useful to know the accuracy of those predictions. (GAP does not advocate the development of policies or regulations, or the enforcement of such, based on these predictions without supporting site surveys). The methodology for assessing the accuracy of the predicted vertebrate distribution maps is not unlike that for land cover maps—samples of the map are compared to data for the same geographic location, which is assumed to represent "truth," and the results are assumed to be valid for all like elements. In the case of land cover (see Land Cover Assessment chapter), stratified random samples are drawn from the map, and then truth data are collected from site visits and/or aerial photos or videographs of that location. Ultimately, this would be the preferred procedure for assessing the vertebrate maps as well, but there are a number of difficulties with this approach:

- Species' habitat occupation varies diurnally, seasonally, and annually, thus requiring repeated visits over many years to detect all species occurring there (Cooperrider et al. 1986).
- Species' behaviors and habitat use such as burrowing, evading potential predators, or using only tree canopies makes detection difficult. The mere presence of the assessor(s) could cause many species to hide or leave the sample site, thus producing erroneous results.
- Species assessment in the field suffers from the same difficulty of obtaining access to private and military lands which can be largely overcome with aerial imagery for land cover, but not for animal species.

Field investigations to inventory species for single study sites have indicated the need for repeated visits during different times of day and seasons for a period of several years to achieve a complete inventory (Ralph and Scott 1981). When considering the number of species modeled in a typical GAP project (300 - 700) and the size of the region to assess (millions of hectares), it is infeasible to recommend a thorough, field-based approach for the entire set of species and area of the project. Therefore, rather than randomly selecting sample locations and then gathering truth data, GAP identifies locations that have compiled a high-confidence checklist of species occurrence over a period of time and uses these as the truth data. There are a number of limitations to this approach which are described in the methods section.

GAP vertebrate distribution maps incorporate a number of spatial and temporal assumptions that must be considered in accuracy assessment:

1. Species are assumed to occur within a polygon representing potential habitat but are not predicted to occur at any particular point within that polygon.
2. Species are assumed to be present within a polygon, but no assumptions are made about the abundance of the species in the polygon.
3. Species are assumed to be present in a polygon at least once in the last 10 years but need not be present every year in the last decade.
4. Species are assumed to be present during some portion of their life history, not necessarily during the entire year.

Prior to any assessment initiative, GAP distributions will have been subject to expert review (see

Vertebrate Modeling chapter). While we consider such review the most important aspect for establishing confidence in the maps, we realize that it is subjective and unquantifiable. We therefore provide these recommendations to establish quantifiable, if incomplete, measures of accuracy.

Methods

The following describes methods for acquiring validation data and conducting the comparison to the distributions. The methods are presented in order of increasing cost and difficulty.

Requirements for reporting results are described under that section.

Method 1: Comparison with species checklists

This is the required method for validating GAP distribution maps through comparison of the lists of species GAP predicts to occur in an area with a comprehensive, recent, and reliable list of species known to occur in the same area that is based on field observations and collections. This method cannot be considered a true accuracy assessment, but is rather a measure of agreement between the predicted distributions and high-confidence data. This method of validation depends on the availability of existing checklists from a reasonable number of areas (5 or more) for comparison. There may be only a couple or no reliable checklists for any individual state. In this case, assessment may have to be deferred until regionalized distribution maps are created (see Vertebrate Modeling) and reliable checklists can be accumulated from a larger region.

Reliable checklists of species from sites such as national wildlife refuges, national parks, and other management areas should not be used to build distribution maps. These checklists represent the minimum data required to assess the accuracy of distribution maps, and therefore must be withheld for this purpose. Lists of observed species come from previously developed checklists that are considered accurate and complete (e.g., wildlife refuge lists based on field observations, faunal surveys of discrete study areas, or selected breeding bird survey results).

Checklists vary in reliability by site and taxonomic group. Only reliable checklists, as confirmed by consultation with the area biologist, should be used (GAP projects have found management area staff to be very forthright about checklist quality). Checklists constructed from actual records (specimens, observations) of species known to occupy the area during the last decade are considered reliable (excluding accidentals). Unsatisfactory checklists are those constructed from field guides or based on predictions of species expected to be present. The source, date, and method of constructing the checklist should be recorded in the project metadata. There is, however, some evidence that high-quality checklists will indicate false commission errors when further surveys are conducted (Thompson et al. 1996), indicating that checklists will also have an unknown degree of omission. Another suggestion for assessing the reliability of surveys is to obtain a description of field methods used. This can help determine the appropriateness of the comparison for particular species or taxa (Jeff Waldon, pers. comm.).

Scale is an important consideration, as checklist comparison for a very small area (less than 1,000 hectares) will probably report a high degree of commission error simply because GAP distributions are not created for accuracy at fine resolutions. In most cases, checklists will deal with areas that contain several habitat types, even though species are not usually associated by habitat type in such lists. Comparisons must therefore be made with predictions for the entire area of the checklist, not for particular habitat polygons within that area.

Comparison of predicted species lists with checklists yields several results:

1. Raw data: number and identity of species predicted and of species observed.
2. Number and identity of species predicted to occur but not observed (errors of commission).
3. Number and identity of species observed but not predicted (errors of omission).
4. Size of area included in checklist (the differences between predicted and observed species lists may

be dependent on the size of the area covered by the checklist).

Several assessments using this approach to map validation have been carried out by GAP investigators (Scott et al. 1993, Edwards et al. 1996, Thompson et al. 1996, Merrill et al. 1996). In a prototype application of this method, Scott et al. (1993:30) compared predicted species lists for three managed areas in the state of Idaho. The results indicated that for most terrestrial vertebrates GAP modeling methods worked reasonably well. For a total of 419 species, the overall omission error was 10.7% (45 species), and the commission error was 21.0% (88 species). Some commission errors may be due to incomplete area lists and, in such cases, it may be worthwhile to review the checklist with the area management staff to reevaluate confidence for those species.

Edwards et al. (1996) compared GAP-predicted distributions and observed species lists from eight national parks in Utah and found differences in omission and commission errors between terrestrial vertebrate classes. Overall accuracy by taxonomic group ranged from 60% to 86% for 15 amphibians, 81% to 95% for 353 birds, 78% to 85% for 131 mammals, and 70% to 83% for 67 reptiles. Mean percent accuracy ranged from a high of 85% for birds to a low of 72% for amphibians. Some directed field research is called for to clarify the factors contributing to each type of error, but initial evidence suggests that some commission errors reflect species that are present but difficult to detect (e.g., shrews, bats).

Thompson et al. (1996) compared GAP-predicted animal presence lists with records of occurrence for birds of San Juan County, New Mexico, and for amphibians, reptiles, and birds of the White Sands Missile Range, New Mexico. No suitable lists were available in New Mexico to allow an accuracy assessment of mammal predictions. The overall accuracy of 236 GAP predictions of birds for San Juan County was 88.6%, once accidental and nonnative species were excluded. The accuracy of GAP-predicted animal lists for the White Sands Missile Range varied by class: 53.8% for amphibians, 67.2% for reptiles, and 77.1% for birds. However, the biological staff believed many more species of amphibians and reptiles were likely but not yet documented on the Missile Range. If all reptile species documented or likely on the missile range were considered, comparison with GAP predictions would be 53 matches, 1 omission error, and 5 commission errors, for an accuracy of 89.9%.

Thompson et al. (1996) feel that their results appear to corroborate the general adequacy of using habitat relation models to estimate species distributions, as demonstrated and discussed for more locations by Edwards et al. (1996). They also consider it important that there is no assurance that the lists they used to assess animal accuracy are independent from their distribution modeling process, because people consulted during algorithm development were aware of at least some of the species detections enumerated on these lists. The most independent data were bird records extracted from research conducted on White Sands Missile Range in 1996 after the algorithms were completed, which documented 44 species which otherwise would have been errors of commission.

Merrill et al. (1996) carried out a similar assessment of vertebrate distribution map accuracy in Wyoming, using lists of all four classes of vertebrates from three study areas and lists of birds and/or mammals from an additional five areas. As in New Mexico, Wyoming GAP researchers also followed analysis protocols described by Edwards et al. (1996). The overall accuracy for all species across all sites was 79.5%. Errors of omission averaged 12.2% (range 0 - 36.6%) for all taxonomic groups and were often high for birds, indicating that their models tended to underpredict the presence of bird species. Errors of commission averaged 8.3% (range 0 - 34.8%) for all taxonomic groups and were highest for mammals, indicating that their methods tended to overpredict the presence of mammal species. They did not find strong evidence that error rates decreased with increasing size of the assessment area as suggested by Edwards et al. (1996).

Reporting results

Tables 1 and 2 below are the minimum information that should be provided to communicate the results of the checklist comparison. Because it is unlikely that you will obtain sufficient samples (checklists) to derive a meaningful result by individual species, Table 1 aggregates the results by taxonomic group and can be combined with Table 3 if it is used. Table 2 provides a record of the comparison by species and checklist. Table 3 (optional) provides results for individual species assessments when sufficient samples to do so are available. A map indicating the locations from where the comparisons were made is also required. The WY-GAP final report (Merrill et al. 1996) is a good example of reporting the methods and results of the assessment.

Table 1. Summary of comparison of predicted animal distributions to checklists by taxonomic group. The numbers after the name of the taxonomic group are examples for the number of species assessed/the number mapped.

	Matches		**Errors of Commission**		**Errors of Ommission**	
Taxonomic group:	n	%	n	%	n	%
Birds (145/189)						
Mammals (44/83)						
Amphibians (5/15)						
Reptiles (16/22)						
Overall (210/309)						

Table 2. Results of checklist comparison for individual species. Under each comparison location, the result is indicated as M (match), C (commission error), or O (omission error). No comparison was made if the species was not identified to occur on either the checklist or the GAP map. Note: these should be grouped taxonomically, in the same order as your master species list. This table should be provided as an appendix.

Species Code	**Species Name scientific/common**	**Lava Mountain N.P. 50,400 ha**	**Duckville NWR 880 ha**	**Ft Brigade Army Base 512 ha**
Birds				
42525	Picoides tridactylus Three-toed woodpecker	M	C	

Table 3. Accuracy of select species. nM (number of matches), %M (percent matches), nC (number of commissions), %C (percent commissions), nO (number of omissions), %O (percent omissions).

Species Code	**Species Name scientific/common**	**nM**	**%M**	**nC**	**%C**	**nO**	**%O**
Birds							
42525	Picoides tridactylus Three-toed woodpecker	15	75	3	15	2	10

Limitations

- Checklists do not represent a probability sample of the state, therefore, inferences about accuracy of a species distribution outside of the checklist area, about the taxonomic group, or the entire collection of predicted distributions will be open to a variety of interpretations. The results may support general inferences about taxonomic groups and even the collection of distribution maps, but as with all GAP results, specific field verification is required prior to any decision-making for individual situations.

- Depending on the method used to construct the maps, there may be circularity between the map creation and validation if experts that developed the range maps were aware of the checklist contents. There is no practical way to prevent this, and it should merely be recognized as a limitation.
- While mismatches of predicted and actual occurrences indicate inaccuracies in GAP predictions, matches may not be indicative of overall map accuracy. One reason for this is the effect of habitat management on accuracy. Because the checklists will be obtained primarily from areas managed for biodiversity values (GAP status 1 & 2 areas), they may show a higher degree of agreement with the GAP distributions than areas managed otherwise. This is because GAP typically does not include habitat quality in the prediction model. In other words, an area that may be degraded but is mapped with the same habitat characteristics as the comparison area may show less agreement because the species are absent due to the effects of unknown management practices.

Method 2: Comparison with species occurrence records

Records for use in this assessment can come from two sources, those that were withheld from use in the development of the distribution maps, and those that were discovered or developed after completion of the maps. Obviously, because the existence of the second set cannot be assumed, a sample of the first set should be withheld if possible. For many species there may be so few records that all must be used to develop range maps (see Vertebrate Modeling chapter). Records used for assessment must be independent of those used to build species distribution maps and allow statistical comparison of predicted versus observed distributions. Washington GAP was successful in obtaining a number of recent observation points from the forest industry. Members of the Public Utilities Managers Association have indicated that utilities would be a good source of such data because of the frequent inventories they conduct for environmental impact statement (EIS) requirements.

After identification of occurrence records, the comparison can be handled efficiently in a GIS intersection to identify those records that are congruent with the distribution polygons. For species with highly fragmented or patchy distributions, a proximity analysis may be necessary to establish congruency such that an occurrence falling among an area of patchy distribution, but not directly in a distribution polygon, will yield an accurate match. One suggestion may be to establish a buffer around the observed point representing the size of the species home range. Establishment of rules to govern a spatially "fuzzy" assessment will be dependent on scale and MMU and requires further research. Reporting of results is the same as for comparison with checklists.

Method 3: Field surveys

Gathering new information on occurrence and species richness with field surveys has several advantages over other methods, but such surveys are very costly and difficult to conduct properly. For this reason, we do not describe the data collection procedure but instead identify some critical issues and suggested approaches and provide an example (forthcoming) in Appendix 1. Properly designed field surveys provide a primary source for completely independent data on current species occurrence (Scott et al. 1986) and richness. Field surveys of areas selected by probability sample designs, such as simple random, stratified random, or cluster sampling, allow us to make inferences about the entire state or region (sample universe) (Scott et al. 1986). However, obtaining high-quality species lists for large sample units with diverse communities (habitats, etc.) with limited resources of qualified personnel, funds, and time is difficult.

A comprehensive assessment of map accuracy incorporates species-to-species variations over the landscape based on spatially distributed field data. Because of limitations of time and money, without careful design field surveys can quickly become infeasible. Several approaches are possible, depending on assessment objectives, the nature of the reference data to be used, and the underlying biological and statistical assumptions.

Obtaining complete species lists for sample units in a short period of time with reasonable effort is probably impossible in most habitats. The temporal variability of species occurrence and the desire to include species present during the most recent decade suggests that species lists gathered during a single year's sampling, when compared to predicted lists, will lead to apparent errors of commission that would be found correct with longer-term sampling.

All of these constraints combine with the reality of very limited funding to suggest that few state GAP projects will be able to consider field surveys as a practical way to validate their species distribution maps. This is not meant to imply that field survey data cannot be accumulated from other sources over time and used by GAP for accuracy assessment. Many of our cooperating agencies (U.S. Forest Service, Bureau of Land Management, National Park Service, U.S. Fish and Wildlife Service, state wildlife agencies, and academic institutions) regularly engage in field investigations, inventory, and monitoring. A careful inspection of the methods and results of these studies can lead to numerous independent assessments of GAP map accuracy. Every time a species is detected in a biotic survey or during standard ecological research, a comparison with the GAP map tests the hypothesis that the species was or was not predicted to be present. The accumulation of dozens or hundreds of such results can provide a powerful way of assessing map accuracy.

Continuing Research

Every GAP project provides the opportunity to further test and refine the approach described in this chapter. A number of limitations were described above; additional research needs are summarized here:

- Testing if the results of comparisons made with geographically limited checklists can be generalized to distributions over the entire project area. In particular, testing the hypothesis that commission errors are lower in areas with measurably higher-quality habitat conditions.
- Testing the bias in the process caused by knowledge of the checklists by those involved in creating and reviewing the predicted distribution maps.
- Testing the ability to generalize the results of assessment to those species untested but related by modeling process, taxonomically, by body weight, guilds, etc.
- Creating a method for ongoing data capture and subsequent assessment and model improvement in cooperation with state, federal, and institutional partners that collect animal data.
- Conducting directed field research to clarify the factors contributing to each type of error.
- Finding evidence that error rates decrease with increasing size of the assessment area, as suggested by Edwards et al. (1996).
- Establishing rules to govern a spatially "fuzzy" assessment (dependent on scale and MMU).

Reviewers

Much thanks to the following people that provided review of the draft manuscript:

Robert Deitner, Tom Kohley, Ann Rasberry, Mike Scott, and Jeff Waldon.

Literature Cited

Cooperrider, A. Y., R. J. Boyd, and H. R. Stuart, eds. 1986. Inventory and monitoring of wildlife habitat. U.S. Dept. of the Interior, Bureau of Land Management Service Center, Denver, Colorado.

Edwards, T. C., Jr., E. T. Deshler, D. Foster, and G. G. Moisen. 1996. Adequacy of wildlife habitat relation models for estimating spatial distributions of terrestrial vertebrates. Conservation Biology 10:263-270.

Merrill, E. H., T. W. Kohley, M. E. Herdendorf, W. A Reiners, K. L. Driese, R. W. Marrs, and S. H. Anderson. 1996. Wyoming Gap Analysis Project Final Report. Department of Zoology and Physiology, Department of Botany, and Wyoming Cooperative Fish and Wildlife Research Unit, USGS Biological Resources Division, University of Wyoming, Laramie.

Ralph, C. J., and J. M. Scott. 1981. Estimating numbers of terrestrial birds. Studies in Avian Biology No. 6, Cooper Ornithological Society, Lawrence, Kansas.

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 No. 123:1-41.

Scott, J. M., S. Mountainspring, F. L. Ramsey, and C. B. Kepler. 1986. Forest bird communities of the Hawaiian Islands: their dynamics, ecology, and conservation. Studies in Avian Biology No. 9:1-431. Cooper Ornithological Society, Lawrence, Kansas.

Thompson, B. C., P. J. Crist, J. S. Prior-Magee, R. A. Deitner, D. L. Garber, and M. A. Hughes. 1996. Gap analysis of biological diversity conservation in New Mexico using geographic information systems. New Mexico Cooperative Fish and Wildlife Research Unit, USGS Biological Resources Division, New Mexico State University, Las Cruces.

Mapping and Categorizing Land Stewardship

Version: 2.1.1
Date: February 2007007
Peer-reviewed: No

Patrick J. Crist
Idaho Cooperative Fish and Wildlife Research Unit
University of Idaho, Moscow, ID

adapted from Version 1 by

Thomas C. Edwards, Jr., Collin G. Homer, and Scott D. Bassett
Utah Cooperative Fish and Wildlife Research Unit
Utah State University, Logan, UT

and including Working Group results by

Bruce C. Thompson
New Mexico Cooperative Fish and Wildlife Research Unit
New Mexico State University, Las Cruces, NM

& special thanks to **Elisabeth Brackney** for research

Introduction

Two primary goals of GAP are to provide an assessment of the management status for certain elements of biodiversity (vegetation communities and animal species) throughout their U.S. range, and to provide land stewards with information on the representation of these elements on their land so they can make informed decisions about their management practices regarding biodiversity. To accomplish this, the mapped distribution of vegetation communities (see Land Cover and terrestrial vertebrate species (see Vertebrate Distribution Mapping chapter) are compared to a map of land stewardship (see Analysis chapter). In GAP, the land stewardship map combines attributes of ownership, management, and a measure of intent to maintain biodiversity. As explained in the Analysis chapter (expected spring 1997), these comparisons do not consider viability, but are a start to assessing the likelihood of future threat to a biotic element from habitat conversionóthe most obvious cause of biodiversity decline (Noss et al. 1995).

We use the term "stewardship," because legal ownership of a land area does not necessarily equate to the entity charged with managing the resource. Though we record the management and ownership entities of public lands and privately owned conservation lands (as best can be determined), we also acknowledge that these attributes are complex and change rapidly (see Continuing Research). At the same time, it is necessary to distinguish between stewardship and biodiversity management status in that a single land steward, such as a national forest, may subdivide its land into units that may be managed for different purposes that affect biodiversity. There are three primary pieces of information involved in developing the Stewardship coverage: 1) Geographic boundaries of public land ownership (and voluntarily provided private conservation lands, e.g., Nature Conservancy Preserves), 2) the manager/owner attributes of each mapped land unit, and 3) the biodiversity management status category of each mapped unit. While stewardship is the most straightforward and quantifiable product of GAP, the mix of owners, lessees, and managers involved with any one land unit and the process of categorizing its management can be confusing and daunting.

The stewardship coverage is no less important than the land cover or vertebrate species distributions, and, like them, a complete statewide map has probably never been produced before. The importance of this map in the GAP process is not only for the spatial documentation of the existing network of conservation lands, it is also the base map from which futureóand hopefully rationalódesigns for the conservation network will come. This chapter will present the standards, data sources, and recommended methodology to complete the task with a minimum of problems. The key to successful completion is to begin this task at the start of the project to allow ample time to obtain and review stewardship documents and involve all potential land management interests.

Terminology

There are many terms used to describe units of land such as "parcel," "tract,"jurisdiction," "subdivision," etc. All of these terms mean different things to different people depending on context, therefore we use "land unit" to describe areas of homogeneous attributes of stewardship and GAP status. We infer no association to a particular type or legal description of an area of land. The terms "map" and "coverage" are used interchangeably to describe a digital spatial map and accompanying database. The term "elements" refers to biotic elements which, for the purposes of GAP, are terrestrial vertebrate species and vegetation communities. "GAP status" is the classification scheme or category that describes the relative degree of management for biodiversity maintenance.

Objectives

1. To develop or obtain a complete digital map (coverage) of land ownership boundaries that describe: a) public land ownership categories and their internal biodiversity management area (land unit) boundaries, b) voluntarily provided private biodiversity management areas, and c) the entities responsible for management of all identified land units.
2. To attribute each mapped land unit with categories of management status 1 through 4 (and subcategories, if desired), current at the time of analysis for the purpose of describing the management status of elements of biodiversity and identifying potential gaps (see Analysis chapter).

Background

Comparing Elements to Managing Entities

The loss of biodiversity in the U.S. can frequently be characterized as a gradual and incremental fragmentation of habitats, where no single management decision causes significant harm, but the cumulative impact of many stewards' decisions contributes to endangered species crises. One cause of this incremental decline is that, prior to GAP, the distribution of elements among different land stewards and management conditions was essentially unknown. Therefore, the impact of an individual management decision on the total distribution of an element was also unknown. The purpose of comparing element distributions with stewardship is to provide a method by which land stewards can assess their relative amount of responsibility for the management of an element in relation with other stewards who share that responsibility. Lacking management, acquisition, or regulatory authority, GAP seeks to help stewards become the best biodiversity managers that they want to be through informed decision making. This information can reveal opportunities for cooperative management of elements, which directly supports the primary mission of GAP to provide objective, scientific information to decision makers and managers. In some situations, a steward that has previously borne the major responsibility for managing an element may, through such analyses, identify a more equitable distribution of that responsibility. We emphasize, however, that GAP only identifies private land as a single homogeneous category and does not differentiate individual private land units or owners, unless the information was provided voluntarily to recognize a long-term commitment to biodiversity maintenance through binding conservation easements, covenants, or institutional dedication.

Comparing Elements to Management Status

After comparing elements' distributions with stewardship attributes, it is then necessary to compare them with categories of management status. The purpose of this comparison is to indicate the potential need for change in management status for individual elements or areas important to biodiversity maintenance. Such changes can be accomplished in many ways that do not affect the stewardship (ownership and managing entity) status (Steiner 1991). While it will eventually be desirable to identify specific management practices for each land unit and their benefit or harm to each element, GAP currently uses a scale of 1 through 4 to denote the relative degree of management commitment to maintenance of biodiversity for each land unit. A status of "1" denotes the highest, most permanent level of commitment, and "4" represents the lowest level of commitment or unknown status. We recognize a variety of limitations in our approach, however, we maintain certain principles in assigning the status level. The first principle is that prescribed management, not land ownership, is the primary determinant in assigning status. The second principle is that while data are imperfect, and all land is subject to changes in ownership and management, we can use the intent of a land steward as evidenced by legal and institutional factors to assign status. In other words, if a land steward institutes a program backed by legal and institutional arrangements that are intended for permanent biodiversity maintenance, we use that as the guide for assigning status. The concept of permanence implies that there is no identified termination date for the current management status as there are for example with Conservation Reserve Program lands. Areas that are established through management plans that are renewable, such as riparian protection zones in national forests are, for our purposes, considered permanent because the legal mandates that established the area through the planning process will likely maintain their status indefinitely through plan renewals.

The GAP Status Classification Scheme

An ad hoc work group was assembled in FY 96 to address the need for a more explicit and repeatable method for categorizing management status for GAP. The group specifically reviewed a dichotomous key developed for this purpose by Crist and Thompson (Crist et al. 1996) and which is provided in Table 2 and Figure 1. The key was accepted by the group as a recommended tool, with the recognition that improvements be incorporated as more experience is gained, particularly in the application to states dominated by private land. Other recommendations of the group are incorporated in this chapter.

The following are some criteria and assumptions used by GAP to determine biodiversity management status for individual land units:

- Permanence of protection from conversion of natural land cover to unnatural (human-induced barren, cultivated exotic-dominated, or arrested early succession). The assumption is that retention of natural land cover is of prime importance to maintaining biodiversity (Noss et al. 1995). We define "natural land cover" simplistically as areas not maintained in an unnatural state (defined above) by human activities, as visible from remote sensing or revealed through management documents. That is, past alterations may be evident, but ongoing alteration is not evident or intended. Management to support biodiversity is permitted (see Type of management below). We do not consider the viability of the land unit to support elements due to size or neighboring impacts, as these are highly element-dependent and are best left to specific analyses of individual elements. We also do not consider pre-Columbian human effects in our definition because of lack of data and concensus on that subject.

- Relative amount of the land unit managed for natural cover. The assumption is that the majority of a land unit must be maintained in a natural state for the reasons stated in the above criterion. We arbitrarily set 5% as the maximum amount of a land unit that can be managed in an unnatural state (as defined in the above criterion) for it to be considered "natural" for a Status 1 rank. All other status ranks allow human disturbance to varying degrees. We do not currently attempt to measure actual conversion of land cover in a land unit and rely instead on the management documents. Anthropogenic land cover <5% within a land unit otherwise managed for biodiversity can be considered an "inclusion" that has an effect but does not dictate the management status.
- Inclusiveness of the management, i.e., single feature or species versus all biota. The assumption is that a land unit managed to retain all of its elements will maintain biodiversity better than a land unit managed only for a single species (oftentimes at the expense of other species). If management is for a "keystone" species for which the majority of the land unit must be maintained in a natural state, we consider that to be inclusive of all elements.
- Type of management (e.g., suppresses or allows natural disturbance) and degree that it is mandated through legal and institutional arrangements. The assumptions are that management which allows or mimics natural disturbance regimes such as fire will maintain biodiversity better than land units that suppress disturbance (Christensen et al. 1996). Our use of the term "mimic" when referring to disturbance is intended for those practices that produce similar overall ecosystem effects as the natural disturbance. This is different than substituting fire with an anthropogenic disturbance such as logging that does not leave the same "legacies" that greatly influence habitat quality and occupation by species (Franklin 1993). We do not, however, address remote anthropogenic disturbance effects such as climate change or exotic species invasion because both effects are at scale extremes incompatible with our immediate analyses. Such analyses are components of a more comprehensive analysis at the reserve design phase (see Analysis chapter).

Using the above criteria, the four biodiversity management status categories can generally be defined as follows (after Scott et al. 1993, Edwards et al. 1994, Crist et al. 1996):

> **Status 1:** An area having permanent protection from conversion of natural land cover and a mandated management plan in operation to maintain a natural state within which disturbance events (of natural type, frequency, intensity, and legacy) are allowed to proceed without interference or are mimicked through management.
>
> **Status 2:** An area having permanent protection from conversion of natural land cover and a mandated management plan in operation to maintain a primarily natural state, but which may receive uses or management practices that degrade the quality of existing natural communities, including suppression of natural disturbance.
>
> **Status 3:** An area having permanent protection from conversion of natural land cover for the majority of the area, but subject to extractive uses of either a broad, low-intensity type (e.g., logging) or localized intense type (e.g., mining). It also confers protection to federally listed endangered and threatened species throughout the area.
>
> **Status 4:** There are no known public or private institutional mandates or legally recognized easements or deed restrictions held by the managing entity to prevent conversion of natural habitat types to anthropogenic habitat types. The area generally allows conversion to unnatural land cover throughout.

Several work group members identified the desirability to add subcategories within each general status category, e.g., 3a, 3b, 3c, etc., to distinguish finer differences in management status than can be resolved with the standard categories described above. This choice and method for accomplishing it are left to the individual projects, but some minimal guidelines are provided later in this text.

Relation To Other Categorization Schemes

The International Union for the Conservation of Nature (IUCN) established 8 categories of protected areas (IUCN 1994). The descriptions are intentionally vague to allow flexibility for global application, therefore the crosswalk to the four GAP categories is ambiguous. The name of those categories and the approximate GAP equivalents are:

ICUN Category	Description	GAP status/comments
I	Scientific reserve/strict nature reserve	1
II	National Park	1 or 2 based on level of development and use
III	Natural monument/natural landmark	1,2, or 3 depending on biologic nature and use
IV	Managed nature reserve/wildlife sanctuary	1,2, or 3 depending on use
V	Protected landscapes	3 or 4
VI	Resource reserve	1,2, or 3 depending on use
VII	Natural biotic area/anthropological reserve	2 or 3 depending on use
VIII	Multiple-use management area/managed resource area	3 or 4 depending on level of protection

Methods

Summary of steps and standards

1. Obtain or develop a map of land ownership boundaries with a minimum mapping unit (MMU) of 40 acres/16 ha (quarter-quarter Township section) or smaller.
2. Attribute land units with general ownership categories (see Table 1 below).
3. Obtain internal management boundaries as applicable and attribute for management authority and type.
4. Categorize all land units for management status using GAP status levels 1-4 and subcategories, if desired.

Detailed descriptions for each step follow:

Step 1: Obtain or develop a map of land ownership boundaries. This is generally a two-stage process. First, the more general ownership categories, such as the major state and federal land agencies, are identified. Second, the internal management units and voluntarily provided private conservation lands are added. The second stage tasks are described in more detail in Step 3.

The Minimum Mapping Unit

The 40 acre (16 ha) MMU has some western bias, as it is based on the State Plane Grid or Public Land Survey System (PLSS) which dominates the land ownership patterns in the western U.S. This unit is equal to a "quarter-quarter section" which is 1/16 of a square mile "section" component of a 6 x 6 mile "Township." The Bureau of Land Management (BLM) maps of land ownership generally operate at this same MMU, but also include parcels as small as 10 acres (4 ha). As with GAP land cover maps, the MMU should be considered a minimum standard, in other words, the coverage must be able to resolve at least a 40-acre land unit, but smaller areas may be included. If you obtain a more detailed ownership coverage from another party, we recommend it be maintained at full resolution to support in-state applications.

Obtaining Existing Ownership Maps

To begin the task of creating a map of ownership boundaries, you should first determine if there is an entity responsible for mapping land ownership/management in the state. These entities range from state or federal natural resource agencies, state economic affairs agencies, to state geographic reference centers. If you are fortunate, the entity will have a digital state ownership map already on file. You may only need to establish an agreement to acquire a digital copy of the map. This can require payment of fees or an exchange of in-kind services. In the western states, the BLM has produced detailed and reasonably accurate maps of public land ownership, usually in conjunction with other agencies. The BLM work can provide a substantial headstart, but it may require updating and verification. The more difficult work will be the identification and incorporation of agencies' internal management boundaries and private conservation areas, which is described in more detail below.

If you obtain the land management data layer from another party, be sure it includes detailed information on the map's lineage. In particular, what sources were used, who did the work, when was the work done, and have there been subsequent modifications to the map since it was first produced? This information will be necessary to produce the metadata required for the stewardship coverage. The older the map, the more likely it is that significant changes in stewardship have occurred and these will have to be incorporated into the map. Some projects have encountered a proprietary problem with ownership maps developed by the state, as they have a cost-recovery system and require a fee for each copy released. This can severely hamper the distribution of the data to GAP data users and for later ecoregional analyses. There are two ways to accommodate proprietary conditions: first, you can make an agreement that allows transfer of the ownership layer to at least one additional source (National GAP) without requiring the purchase of a second copy. This will allow the National Office to carry out ecoregional analyses, and the office can guarantee, if necessary, that it will not transfer the original map to any other parties. Second, the map can be altered or generalized in some way acceptable to the proprietor to allow its general distribution through GAP. The latter is preferred.

Creating A Map From Scratch

If no statewide digital map exists, you will be faced with the task of creating the map from scratch. The degree of difficulty and time expenditure will be highly variable, depending on the amount of local government and private conservation lands in the state, and the degree to which digital maps of ownership exist. We estimate roughly 1-1.5 person years of combined efforts are required for all tasks. There are two primary pieces of information you will need to create the base map: 1) an accurate base map of grids and physical features which typically define ownership boundaries in the state such as the PLSS, state-plane grid with Township-Range-Section designations, political boundaries for counties and municipalities, and roads and hydrologic features, and 2) ownership maps from all public lands agencies (see Step 3 for combining map request efforts). Though most agencies are rapidly adopting GIS to map their lands, the degree of disarray that many organizations' land ownership files are in cannot be overstated. Responses to this request may range from "we don't have any maps like that," to receiving a pile of hand-sketched boundaries or vague descriptions. Where ownership is strongly tied to the PLSS or state-plane grids, it is a fairly easy task to attribute grid polygons with ownership categories.

In other cases, each land unit will have to be scrutinized for the best method to develop a digital boundary through incorporation of agencies' digital files, digitizing of paper maps, or coordinate geometry (COGO) input of legal descriptions. Simply intersecting these digital files with the base map boundaries usually causes slivers that must be manually corrected. If, for instance, a land unit boundary is tied to a water body, it will be better to use the GAP hydrologic data set to derive the boundary to eliminate the generation of slivers in future intersects. All submissions should be reviewed for the amount of work versus benefit to the gap analysis process. If you do not receive an easily digitized map, you should not feel obligated to include the land unit. Many organizations, however, will view digital conversion of their boundaries as a highly beneficial service which can be used to negotiate in-kind services or funding to your project in return. In larger states, it is convenient to partition the state into 1:100,000 tiles when digitizing. These are reasonable storage and working units and were selected principally because the BLM ownership maps used by western states are available in this format. Once each tile is completed, they should be transformed into the appropriate UTM zone coordinates. The transformed coverages can then be appended to each other and edge-matched. Small discrepancies in polygons between map edges (usually caused by differences in map dates) can be simply closed off.

Mapping Internal Management Boundaries and Private Conservation Lands

Most public lands, such as national forests, have internal management boundaries that are either legislatively created, such as wilderness areas, or established through a management plan. Though both of these management types could be changed through either congressional vote (the former) or administrative decision (the latter), GAP considers both to have "long-term intent," and therefore they are critical to identifying potential "gaps." Likewise, there are numerous types of private lands that are managed in ways beneficial to biodiversity that need to be recognized, ranging from muncipal watersheds where activities are tightly controlled, to institutional and corporate lands and private land easements managed by conservation institutions. Because both public land internal management boundaries and private conservation land boundaries are difficult to obtain and incorporate digitally, GAP projects may need to be selective in what to include. A simple rule is that if the land unit is likely to be categorized in a GAP management status higher than the status of the land in the same general ownership, it should be included. For example, most multiple-use public land is categorized as GAP status 3. If the internal management unit is not managed to rank status 2 or better, there is no reason to distinguish it from the surrounding land. The management status categorization process is explained in detail in Step 3, and this specific issue will be explained in greater detail there.

Step 2: Attribute land units with general ownership and management authority categories (see Step 3 for attributing specific management areas and the Standards chapter for database file attribute standards). Because legal ownership is the attribute typically applied to existing maps or most easily obtained, we begin with that category. The BLM maps for the West are attributed with a variety of public ownership categories (though usually not all), many of which may not be pertinent to gap analysis and can be aggregated. For projects developing their own map, 1:100,000 printed maps of the attributed ownership and management should be sent to the appropriate agencies to confirm or correct the boundaries and management responsibilities (this can be combined with a check of management status categorization, see Step 3). It is important to note that the intent of GAP is to identify general boundaries accurate to USGS 1:100,000 spatial accuracy standards, and not to produce a cadastral-quality product for fixing boundaries on the ground. Although NGOs such as The Nature Conservancy are private land owners, they frequently act as a partner with federal conservation efforts and are thus recognized as a separate category.

The following table provides the first comprehensive (however incomplete) GAP coding scheme for land units in the U.S. A four digit scheme was developed to allow sufficient flexibility to add land types as needed by state projects. For categories of lands not listed but within an identified group, e.g. BLM, use the XX90-99 range to code those lands, e.g. for a BLM type not listed, code it as 1190, the next one 1191, etc. This will help us identify these as miscellaneous BLM management categories. The uniform application of this scheme by all GAP projects will greatly reduce the need for cross-walking database attributes when states are edge-matched. Not every type needs to be included in a project coverage, even if it occurs in the state, e.g., Conservation Reserve Program lands. It is up to the state to determine how important these lands are to establishing GAP status for elements in the state, though any areas meeting the MMU and ranking GAP status 1 or 2 should be included. For clarification on project-specific situations, contact the national GAP office. The types of management units were provided by the respective agencies, and the national GAP office will continue to refine the table as more information is acquired.

Table 1. Management coding system for land stewardship. This table is not all-inclusive but allows projects to insert additional codes for further breakdown of categories. Please submit suggestions for category additions to the author. Status codes are explained above.

Class			Land Management Descriptor	GAP Status
1000			Federal Lands[1]	1,2, or 3
	1100		Bureau of Land Management (BLM)	1,2, or 3
		1101	Area of Critical Environmental Concern (ACEC)	1 or 2
		1102	Globally Important Bird Area	1,2, or 3
		1103	National Conservation Area	1,2, or 3
		1104	National Monument	1,2, or 3
		1105	National Natural or Historic Landmark	1,2, or 3
		1106	National Outstanding Natural Area	1,2, or 3
		1107	National Recreation Area	2 or 3
		1108	National Scenic-Research Area	1,2, or 3
		1109	Research Natural Area (RNA)	1
		1110	Significant Cave and Cave System	1,2, or 3
		1111	Wild, Scenic, and Recreation River	1,2, or 3
		1112	Wilderness Area	1 or 2
		1113	Wilderness Study Area (WSA)	2 or 3
		1114	World Heritage and Biosphere Site	1,2, or 3
	1200		Bureau of Reclamation (BOR)[1]	1,2, or 3
		1201	National Recreation Area	2 or 3
		1202	Wildlife/Recreation Management Area	1,2, or 3
	1300		Fish and Wildlife Service (FWS)	1,2, or 3
		1301	National Wildlife Refuge (NWR)	1,2, or 3
		1302	Waterfowl Production Area	1,2, or 3
		1303	Wilderness	1,2, or 3
		1304	Conservation easement	1,2, or 3
	1400		Forest Service (USFS)	1,2, or 3
		1410	Archaelogical Area	1,2, or 3
		1402	Botanical Reserve	1 or 2
		1403	Geological Area	1,2, or 3
		1404	Municipal Watershed	2
		1405	National Game Refuge	1,2, or 3
		1406	National Monument	1,2, or 3
		1407	National Primitive Area	1,2, or 3
		1408	National Recreation Area (NRA)	2 or 3
		1409	National Scenic-Research Area	1,2, or 3
		1410	Research Natural Area (RNA)	1
		1411	Wild and Scenic River	1 or 2
		1412	Wilderness Area	1 or 2
		1413	Wilderness Study Area	2 or 3
	1500		Department of Defense (DoD)[1] and Department of Energy (DOE)	1,2, or 3
		1501	Ecological Reserve	1,2, or 3
		1502	National Wildlife Refuge Overlay	1,2, or 3
		1503	Special Resources Area/Research Natural Area	1,2, or 3
		1550	Army Corps of Engineers[3]	?
		1560	Department of Energy (DOE)	?
	1600		National Park Service (NPS)[1,2]	1,2, or 3
		1601	International Historic Site	1,2, or 3
		1602	National Battlefield	1,2, or 3
		1603	National Battlefield Park	1,2, or 3
		1604	National Battlefield Site	1,2, or 3
		1605	National Historical Park	1,2, or 3

		1606	National Historic Site	1,2, or 3
		1607	National Lakeshore	1,2, or 3
		1608	National Memorial	1,2, or 3
		1609	National Military Park	1,2, or 3
		1610	National Monument	1 or 2
		1611	National Park	1 or 2
		1612	National Preserve	1,2, or 3
		1613	National Recreation Area	2 or 3
		1614	National Reserve	1,2, or 3
		1615	National River and Wild and Scenic Riverway	1,2, or 3
		1616	National Seashore	1,2, or 3
		1617	Wilderness Area	1 or 2
	1700		Natural Resources Conservation Service	any
		1701	Conservation Easement[3]	?
		1702	Conservation Reserve Program Land (optional)[3]	?
		1703	Wetland Reserve Program Land[3]	?
		1704	Wildlife Habitat Incentive Program Land[3]	?
	1800		Bureau of Indian Affairs (BIA)	any
		1801	Wildlife Reserve[3]	?
	1900		National Oceanic and Atmospheric Administration (NOAA)[3]	?
		1901	National Estuarine Research Reserve[3]	?
2000			Native American Lands	any
3000			State Lands	any
	3100		State Parks & Recreation Areas	any
	3200		State School Lands	any
	3300		State Wildlife Reserves	1,2, or 3
4000			Regional Government Lands	any
5000			Local Government Lands	any
	5100		City Parks	any
	5200		County Parks	any
6000			Non-Governmental Organization Lands (NGO)	1,2, or 3
	6100		Audubon Society Preserves	1 or 2
	6200		Local Land Trust Preserve/Easement[1]	1,2, or 3
	6300		The Nature Conservancy (TNC)	1,2, or 3
6301			TNC Easement	1,2, or 3
		6302	TNC Preserve	1,2, or 3
7000			Private Land	any
	7100		Private Conservation Easement/Conservation Deed Restriction	1,2, or 3
	7200		Private Institution-Managed for Biodiversity	1,2, or 3
	7300		Private Unrestricted for Development/No Known Restrictions	4
8000			Water[4]	na
0000			Unknown	4

1. Status rank given is for land units with substantial natural land cover. Land units with substantial human alterations should be downgraded, and primarily developed land units of all categories should be the lowest status allowed for that category.
2. National Parks are status 1 unless they are dominated by visitor facilities or other developments.
3. There was insufficient information to provide a status range as of this writing.
4. Management of water sources is addressed in the aquatic component of GAP.

For attributing water bodies, simply attribute all occurrences with code "8000." The mix of ownership, water rights, and managing entities for water bodies is extremely complex and will be addressed by the aquatic component of GAP.

Frequently, the entity charged with managing the land is different from the owner. This information is useful, because different management entities operate under different rules that may influence management status categorization. In the case of private lands subject to deed restriction or easement, the entity holding the easement should be attributed as the management authority (data field attributes are described below). For very complex arrangements, or where the actual owner is not readily identified, you may attribute "Owner" the same as "Manager." Use Table 1 to assign attribute codes to land units.

You may need as many as four data fields (plus others described elsewhere) to attribute a particular land unit: "Manager" is the primary entity charged with managing the land unit plus the type of land unit it is, with the first two digits indicating the managing entity and second two the land unit type, e.g., a wilderness area within a USFS land unit will be coded 1412; "Owner" is the legal owner and will usually be indicated by 2 digits followed by two zeros such as 1400 in the above example to identify the U.S. Forest Service; "Division" is the primary unit comprised of the internal management units of the managing entity; and "Unit" is the name or code the managing entity uses to identify the land unit or internal management unit. Using the above example, a USFS land unit would be attributed as "*1412*" for "Manager" and "*1400*" for "Owner", "*Big Mountain N.F.*" for "Division, and "3A" for "Unit." The last two fields are only required for land units categorized as GAP status 1 or 2, but if they are simple to attribute for other land units, they will increase the utility of the coverage for other applications. Previous GAP projects counsel new projects to think through carefully the need to extract particular boundaries and to ensure that the attributes will allow the flexibility to do so.

Step 3: Categorize all polygons for management status using status levels 1-4 and subcategories, if desired. The method of categorizing lands is straightforward, but it is time-consuming to obtain and review all relevant information. The key is to make the effort to identify all possible stewards of conservation lands and be clear what is required for their inclusion. The tasks are:

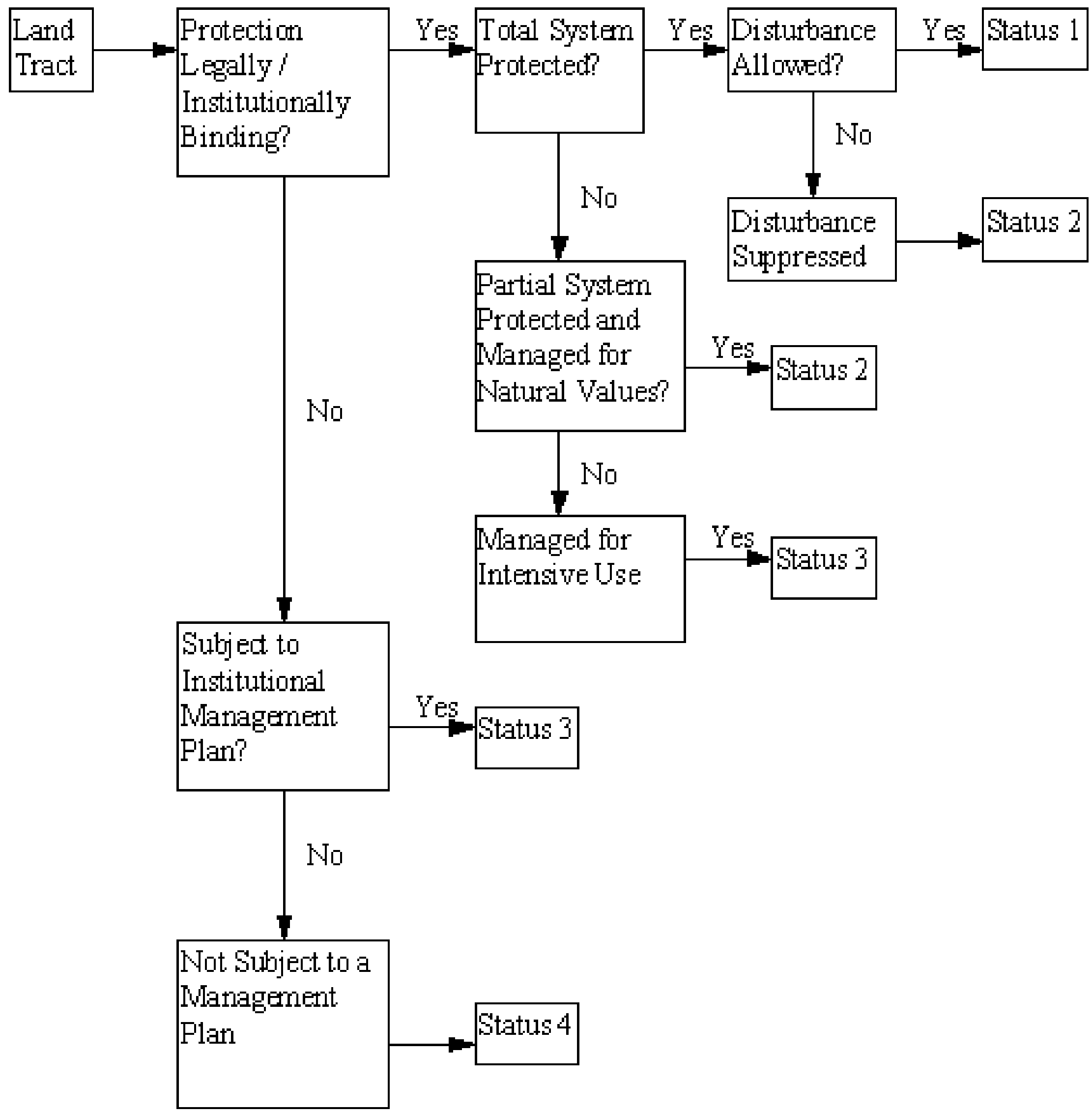
a. Contact all organizations that may have lands that can be categorized as status 3 or better and request boundaries and, if available, written management plans for each conservation land unit. This step should be combined with any requests for general land boundaries per Step 1. Note that differentiation of general ownership into specific management land units may have substantial effects on GAP status. For instance, a large land unit with mining allowed in a single location would be ranked Status 3, whereas, distinguishing an internal management boundary for the mining site may render the greater land unit Status 2. It is not necessary to map internal management boundaries that will not have a Status higher than the general category, e.g., BLM lands generally are status 3, therefore only map those internal management areas that are Status 2 or higher. Occaisionally, areas that are essentially managed the same have very complex checkerboard ownership patterns that would require great effort to map. If the outcome for the gap analyses is likely to be the same, these areas can be treated as a single management unit as long as the appropriate documentation of the mixed ownership is maintained. An example of these areas is the land surrounding many reservoirs which may be distributed among five or more stewards, but may be managed jointly or following similar rules.
b. Use the "GAP status" code from Table 1 to attribute the land unit if it is explicitly stated (e.g., "2") or to bound the possible choices (e.g., ">3" meaning Status 1 or 2). Final determination of the status can then be made with the dichotomous key (task d). GAP uses explicit status rankings for stewardship categories when management is uniform throughout the U.S. for that category. Otherwise, a range is provided as a guide so that projects may use specific management plans to determine the final status rank.
c. Review plans or interview the manager for the criteria necessary to establish the management status.
d. Use a dichotomous key (see example below) to consistently rank all management areas.

e. An optional step is to use subcategories to distinguish finer differences in management status than can be resolved with the standard 1 through 4 categories. Assignments may be added under a field "Status_sub" and should use lower case characters in a descending hierachy of management for biodiversity such that a "3a" is of higher management status than a "3b".

Table 2. A dichotomous key for categorization of biodiversity management status of land units.In using the terms "permanent" and "legally enforceable" we recognize that all conditions are subject to change, even in wilderness and national parks, but the intent is for the condition to be of very long term.

A-1:	If the management intent can be determined through agency or institutional documentation GO TO A-2, if not, GO TO A-5
A-2:	If the land unit is subject to statutory or legally enforceable protection from conversion to anthropogenic use of all or selected biological features by state or federal legislation, regulation, private deed restriction, or conservation easement intended for permanent status, GO TO B-1; if not, GO TO A-3
A-3:	If ecological protection is not legally enforceable, temporary, or lacking but managed by a plan intended for permanent status, GO TO A-4; if not, GO TO A-5
A-4:	Management to benefit biological diversity is provided by a written plan in place or in process under an institutional policy requiring such management **- Status 3**
A-5:	Not subject to an adopted management plan or regulation that promotes biological diversity, or management intent is unknown - Status 4
B-1:	If the total system in the land unit is conserved for natural ecological function with no more than 5% of the land unit in anthropogenic use, GO TO B-4; if conservation provisions apply only to selected features or species, GO TO B-2
B-2:	If management emphasizes natural processes including allowing or mimicking natural ecological disturbance events, but also allows low anthropogenic disturbance, renewable resource use, or high levels of human visitation on more than 5% of the land unit - **Status 2**; if not, GO TO B-3
B-3:	Management allows intensive, anthropogenic disturbance such as resource extraction, military exercises, or developed or motorized recreation on more than 5% of the land unit, but includes ecological management for select features **- Status 3**
B-4:	If management strives for natural processes including allowing or mimicking natural ecological disturbance events - Status 1 ; if not, GO TO B-5
B-5:	Managed for natural processes, but some or all disturbance events are suppressed or modified - Status 2

Figure 1: The dichotomous key in graphic flow-chart format.



The above key should not be used in the same strict terms as a scientific key in that you will rarely know enough about a land unit to categorize it absolutely. It is intended as an aid to make a subjective process less so. For that reason, the user is allowed to go back to a higher level if the result is counter-intuitive to what the user knows about the land unit. Distinguishing between status 2 and 3 is especially difficult for some situations and critical to the identification of "gaps"; therefore we prefer a thoughtful deliberation over rigid application of the key. This, of course, increases the level of subjectivity in the process, but the key is simply an aid to establish a status that can be rationalized given current documentation. We encourage projects to categorize all land units by two persons independently of each other and then to resolve any differences through discussion between themselves and in consultation with the land unit managers if the criteria are still unclear.

Metadata for this process should include (see Standards chapter):

1. Name or code of the management area as designated by the area manager.
2. Date of establishment of the management area in its current status.
3. Effective date and expiration of the current management plan, if any.

4. Citation for the management plan or name and position of person consulted.
5. Comments.

Some areas under temporary management programs can be included for more localized, short-term

assessments, but for national GAP assessments, only those land units with programs intended for long-term biodiversity management and backed by legal or institutional arrangements are considered. Distinguishing between these is a judgment call, but a general guideline for use is that if a land unit is subject to a temporary management plan but is owned by the managing institution which has a biodiversity management mission, the current management intent may be considered long-term. An example would be USDA Forest Service Botanical Reserves, which are not legislatively created, but nonetheless were established through a biodiversity management program. Another criteria to distinguish areas that should be mapped and attributed for biodiversity management status is the likelihood that a change in the current management intent would be identified through an update to the map in the 5-10 year update cycle of GAP. In the case of the Forest Service, such management changes are easily identified.

Stewardship is dynamic. In the absence of a state agency solely devoted to updating and maintaining ownership records, there is no way to keep abreast of the many changes that occur in stewardship. Consequently, you must select a cutoff date for incorporation of changes. An exception to this rule might be, for example, if a new wilderness area was designated, or The Nature Conservancy was to purchase a large tract of land. Such transfers are of extreme importance in estimating the management status of biotic elements, and it should not present an onerous task to incorporate them immediately prior to conducting the analyses.

Reporting the Results

It is useful to produce a simple cross-tabulation of managing entities with GAP status to report the representation of stewardship among the status categories, and statewide. An example summary table is provided below. You must also document information about the status 1 and 2 management areas and the source of information used to make the categorization.

Table 1. This table indicates the area in square kilometers (km sq) (multiply by 100 for hectares and by 247 for acres) for the following measurements: 1) area of the land stewardship category in each management status and total for the state, 2) the steward's percent of total area in each management category and of the state area, 3) total area of each stewardship category in the state and its percent of state area, and 4) mean elevation of the lands in each status category. Note projects- the categories provided are only an example. Include only categories that represent a fairly large area of the state or a large portion of any one status category. Lump all those not shown categorically into a column that says "other," to indicate that the list is incomplete.

Steward	Status 1	Status 2	Status 3	Status 4	Total
USFS	1685 0.4%	2497 0.6%	28861 6.7%	0 0.0%	33043 7.7%
Other					
Total	3561 0.8%	3122 0.7%	99521 23.2%	323828 75.3%	430032 100.0%
Mean Elevation	2667	2440	1122	988	

Table X. The following table identifies all management areas identified by XX-GAP categorized as a management status 1 or 2. It also identifies the entity charged with management implementation, and the source of information used to categorize the area. Analyses of representation (see page_) were based on these identified places, therefore places not identified here (or changes in management plans that would influence the assigned status) will affect the analyses. Please use a written plan citation, or if it is a pers. comm., identify the person, and provide a date for either. If the list is very lengthy, you may place it in an appendix.

Name of Status 1&2 Areas/size (sq km)	Established	Managing Entity	Source Used to Categorize
use official map name/size in sq km	date est.		

Limitations

The described process does not account for land units smaller than the MMU and is valid only at the point in time that the data were acquired. Because only conservation areas with an intent of permanence are typically included, it is unlikely that status for a land unit would be downgraded, but conservation areas established after the coverage is produced are, of course, excluded from the analysis. We acknowledge that many public and private lands have high value for biodiversity and are managed responsibly, but the intent of GAP is to assess long-term potential for habitat conservation, and therefore management that is not backed by legal enforcement or institutional mandate is not included. We also do not consider actual management practices such as pesticide spraying, which may have considerable influence on biodiversity. At the scale of gap analyses, it is currently impractical to map such attributes, which also change very rapidly. Further limitations of the application of these data in analysis are described in the Analysis section of this handbook.

Maintaining the Stewardship Coverage

When conducted according to these standards and suggested methods, a good deal of effort will have been applied to the development of this coverage. Because this coverage is probably the most dynamic of all GAP products, it is especially worthwhile to establish an arrangement for its regular update by a state entity. The Natural Heritage Programs (NHP) are generally charged with tracking managed areas, and several have expressed an interest in maintaining the coverage. Facilities and resources are highly variable among the NHPs, however, and other state institutions may have to be considered. Keeping up on the changes to the coverage would probably amount to only a few hours a month and would allow continued and improving utility of the coverage for a variety of uses, while letting it lapse until a major GAP update would require nearly starting from scratch. Such regular maintenance would also provide an opportunity to gradually incorporate smaller and shorter-term conservation arrangements that could better address immediate risk to biotic elements.

Continuing Research

- Because GAP provides an assessment of individual elements, it would be of greater utility and precision to evaluate the effects of management practices on individual species and plant communities. The National Gap Analysis Program at the University of Idaho is examining the feasibility of attributing actual management practices for comparison to vertebrate species' known sensitivities to such practices.
- An additional attribute that would be of enormous potential utility for implementation would be "Management Interest" to identify those entities other than the "Owner" or "Manager" that exert routine oversight of the land unit. These entities are not those already built into the managers' institutional system of plan development, but those groups that seek to influence the manager both for conservation and resource exploitation.
- Additional assessment of how land stewards and decision makers regard the status rank assignments is warranted to understand how data users may accept or reject gap predictions.

Literature Cited

Christensen, N.L., A.M. Bartuska, J.H. Brown, S. Carpenter, C. D'Antonio, R. Francis, J.F. Franklin, J.A. MacMahon, R.F. Noss, D.J. Parsons, C.H. Peterson, M.G. Turner, R.G. Woodmansee. 1996. The report of the Ecological Society of America Committee on the Scientific Basis for Ecosystem Management. Ecological Applications 6:665-691.

Crist, P.J., B. Thompson, and J. Prior-Magee. 1996. Land management status categorization for Gap Analysis: A potential enhancement. Gap Analysis Bulletin #5, National Biological Service, Moscow, ID.

Edwards, T.C, C. Homer, and S. Bassett. 1994. Land management categorization: A users' guide. A Handbook for Gap Analysis, Version 1, Gap Analysis Program.

Franklin, J.F. 1993. Lessons from old-growth. Journal of Forestry 91:10-13.

IUCN. 1994. Guidelines for protected area management categories. International Union for the Conservation of Nature. Gland, Switzerland. 261 p.

Noss, R.F., E.T. LaRoe, and J.M. Scott. 1995. Endangered ecosystems of the United States: A preliminary assessment of loss and degradation. Biological Report 28. National Biological Service. U.S. Department of the Interior.

Scott, J. M., F. Davis, B. Csuti, R. Noss, B. Butterfield, S. Caicco, C. Groves, T. C. Edwards, Jr., J. Ulliman, H. Anderson, F. D'Erchia, and R. G. Wright. 1993. Gap Analysis: a geographic approach to protection of biological diversity. Wildlife Monographs No. 123.

Steiner, F. R. 1991. The living landscape. McGraw-Hill, New York

KANSAS GAP SURVEY

Status__________________

Name of Individual Completing Survey ____________________________________Date______________

Name of Land Unit __________________________________ Total Acreage of Land/Water Unit___________

Owner (Agency) _____________________ Total Area(land) ________ Total Area (water)_______________

Manager (Agency) _______________ Total Area(land) ________ Total Area (water)_______________

What is the primary use of the land? ____

Is the area accessible to the general public by road? **Yes No** Other Access (trails)? _____________
Number of miles of PAVED roads ______, GRAVEL roads ___________, PRIMITIVE roads __________

Number of Parking areas ______ Total acreage of Parking areas ______ Distribution_______________

Is there Permanent protection from conversion of natural land cover? **Yes No**

Is there Statutory or legally enforceable protection from conversion to anthropogenic (related to or influenced by the impact of humans on nature) use of <u>all or selected</u> biological features by state or federal legislation, regulation, private deed restriction or conservation easement intended for permanent status?
Yes No

If YES, please 'X' the applicable criteria:

Type of protection	**Protection of ALL features**	**Protection of SELECTED features**
State or Federal LEGISLATION		
State or Federal REGULATION		
Private deed restriction		
Conservation easement intended for permanent status		

Are there Natural Elements Present? **Yes No**
Is there a Refuge present? **Yes No** Total acreage of Refuge_________________
Activities Allowed on Refuge ___________________________________
Length of time refuge is open ___________________________________
Is there a Natural area present? **Yes No** Total acreage of Natural Area_____________
Activities Allowed on Natural Area______________________________
Length of time Natural Area is open______________________________

What Types of Habitats are present in the area (e.g., deciduous forest, tallgrass prairie)? ___________________

Has there been Exotic invasion in the area? **Yes No**
If yes, what species?____

Are there Wetlands present? **Yes** **No** Total acreage of Wetland ________________________

Activities Allowed in wetlands __________________________________

Length of time wetlands are open __________________________________

Is the Primary focus for conserving biodiversity? **Yes** **No**

Is there a Single species focus? **Yes** **No** If yes, what species?___________

Does the species require natural land cover? **Yes** **No**

Is there a multiple-species focus? **Yes** **No** If yes, what species? ___________________

Do the species require natural land cover? **Yes** **No**

Is there habitat protection for threatened and endangered species? **Yes** **No**

If **yes**, are they Federally listed or State listed species? **(circle one)** Federal State Both

Is there a Mandated management plan in operation to maintain a natural state? **Yes** **No**

Are Natural disturbance events allowed to proceed without interference? **Yes** **No**

Are Natural disturbance events mimicked through management practices? **Yes** **No**

Are there Prescribed Burns? **Yes** **No**

Yes / No	**% of area burned annually**	**Length of time between consecutive burns**

Is Grazing permitted? **Yes** **No**

Types of Grazing	**Yes / No**	**Acres**	**Length of Grazing (days, months)**	**Intensity (low, intermediate, high)**	**Number of animal units**
Bison					
Cattle					
Other livestock					

Are Natural Resources Removed from the area? **Yes** **No**

If **YES,** Please complete the following chart:

Activity	**Yes/No**	**Acres**	**Intensity (week, month or seasonal)**
Mowing frequently during growing season			
Mowing and Baling			
Logging			
Mining			
Leasee Agriculture			
CRP land			
Grain Production			
Forage Production			
Conversion of natural habitat to cool season grass pasture			
Irrigation			

Food plots			
Shelterbelts, hedgerows, woody plantings			

Are recreational opportunities available? **Yes** **No**
If **Yes**, Please complete the following table:

Activity	Yes/ No	Usage per year (high, intermediate, low)	Other
Hunting			Type of animal:
Shooting Range			
Fishing			
Tournaments			Frequency:
Trapping			
Camping/Picnicking			
Primitive (no facilities)			
Low intensity (fire rings, picnic tables)			
High intensity (pads, electrical hookups)			
Boating			
Restrictions			
Number of Ramps			
Hiking Trails			Miles of Trails:
Biking Trails			Miles of Trails:
Nature Trails			Miles of Trails:
Horseback Riding Trails			Miles of Trails:
Other public use facilities			Types:
			Number of facilities:

Thanks for completing this survey! If you have other comments, please use the space below or attach additional sheets.

National Gap Analysis Project Standards

Version: 2.0.1
Date: 16 Februar 2007
Peer-reviewed: No

Patrick Crist and Michael Jennings
Idaho Cooperative Fish and Wildlife Research Unit
University of Idaho, Moscow, ID

- Appendix 1 - Data Delivery Standards for GAP (PDF)
- Appendix 2 - Project Close-out Protocol (PDF)

Introduction

This document provides a summary of standards for conducting a Gap Analysis Project. As a collaborative "bottom-up" effort, the National Gap Analysis Program recognizes the need for a flexible approach at the state level. However, in order to provide ecologically meaningful information at the bioregional and national levels, certain minimum project standards are essential. The standards described below apply to all products delivered to the national program (see Appendix 2 for project close-out and data delivery process). Detailed descriptions of standards for individual components are provided in the *Handbook for Gap Analysis* (Scott and Jennings 1994), but note that this section references information in specific chapters that may be in progress and therefore not yet available. For clarification, contact the National GAP office.

Projects shall indicate in their proposals, contracts, and modifications the standards in effect at the time the agreement is developed (version number and date). They will not be subject to later revisions of the standards but are strongly encouraged to meet the standards in effect at the time of project completion when doing so will not result in delay of completion or increase in project cost. In particular, this applies to file and database field names and attribute codes, which are easily revised and critical for regionalization of the data. A note on terminology:

Because this is a summary chapter, we do not define terms or provide citations or sources of information that are already provided in the Handbook chapters for those specific items. See also the Glossary in the Handbook for definitions of many terms and acronyms. With this chapter we introduce the use of the term "resolution" which replaces Minimum Mapping Unit (MMU). GAP previously required land cover to be aggregated to a MMU of 100 ha, which suggested that, nothing smaller than 100 ha should be portrayed. We no longer make that requirement and instead state that "the resolution of this coverage is at least X ha" to indicate the smallest mapped unit that a project should be able to portray, but it may portray smaller features as well. This is further explained in this chapter and will be treated in detail in the forthcoming revision of the Land Cover chapter (see Handbook table of contents for chapter status).

Basic GAP Components and Deliverables:

The basic GIS coverages and associated databases required to complete Gap Analysis are identified below. The list of deliverables is substantially expanded to: a) support data regionalization and analyses, b) provide a permanent archive for the entire set of data needed for replication, and c) provide an adequate starting point should a different facility be required to complete the project, and d) provide a more user-friendly and complete CD-ROM/DVD distribution product. All items are required deliverables unless otherwise described:

1. Land Cover Data:

a. A GIS coverage of actual land cover by NVCS (alliance level or aggregation of alliances).

b. The digital reference data (ground points, aerial video, etc.) used to label and assess the map. Generally these are a deliverable if they can be delivered in a compact digital form (CD or DVD). Consult with National GAP for your particular situation.

c. A digital manual describing the mapped land cover types.

d. Any imagery obtained with project funds and not included in the MRLC set.

2. Animal Modeling Data:

a. A database file of predicted species occurrence in EMAP hexagons.

b. A database of specimen collection or sighting localities, if available.

c. A database of species habitat associations with references.

d. GIS coverages of the individual predicted distributions of each terrestrial vertebrates species.

e. A "hyperdistribution" of all species in a single coverage.

3. Land Stewardship Data:

a. A GIS coverage of land stewardship (ownership and management).

b. A database or lookup table of all status 1 & 2 management areas with reference data.

4. Analysis Data:

a. A database of the analysis statistics for land cover and animal species.

5. Ancillary GIS coverages and digital tabular data used to create GAP coverages, including the USGS 30 x 60 minute quadrangle coverage.

6. One digital copy of the final project report according to the standard template.

7. Graphic image files of all coverages, including all individual animal species distribution maps, tables, charts, and outcome analyses maps produced for the written report.

8. A graphic file of the logo for your organization.

Additional General Requirements:

The following are not meant to replace specific project contract language but to inform potential GAP investigators of performance expectations.

1. Projects shall deliver a CD-ROM/DVD of all interim or completed data with their annual and final reports. Annual deliveries prior to the final delivery do not require ancillary data or imagery, only data actually produced by the project.

2. Projects shall deliver all coverages according to the standards in the Data Delivery Standards, Appendix 1.

3. Projects shall follow all applicable rules of the USGS Biological Resources Division (BRD) as stated in their RWO or Cooperative Agreement when conducting Gap Analysis.

4. Projects shall make substantial efforts to involve all organizations that can likely contribute to the project in financial, in-kind, or advisory roles. At a minimum, workshops should be held for all potential cooperators with invitations going to the directors as well as key program and GIS leaders within the institutions. A newsletter is also highly recommended. They must also participate in regional GAP coordination efforts to achieve thematic and cartographic match between states.

5. Projects shall make substantial effort to identify an institution in-state to serve and maintain the GAP data. Data distribution shall include metadata and the final report, including the BRD data disclaimer.

6. Projects shall maintain a fire-safe backup of all data current to within one week of the last update.

7. In addition to the final report, projects shall produce progress reports according to their project agreement. The Principal Investigator will also notify the GAP office whenever there is a significant change in the status of: principal and secondary investigators; staff and affiliated investigators; cooperators; project funding; equipment and imagery; mapping; and analysis and data archive.

8. Projects shall follow the project close-out protocol (see Appendix 2) to achieve sign-off on project completion.

Data Production Standards:

General

a. Follow the overall Gap Analysis process described in *A Handbook of Gap Analysis* using the version current on the GAP home page (http://gapanalysis.nbii.gov) or in consultation with the national office if a chapter is not current or available.

b. Projects may exceed these standards (e.g., use finer resolution, more detailed classification levels) provided that:

> 1) the refinements do not increase the original project budget or are funded by others; and
>
> 2) do not cause extension of the project time-line without prior consultation with National GAP.

c. For digital spatial data (GIS coverages), geographic position accuracy will meet or exceed the standards of the USGS for 1:100,000 scale products follow standards identified in the

Handbook chapters. Because production of GAP data products requires the incorporation or "overlay" of numerous data layers, it is critical that they align such that the total spatial variation between all data does not exceed that allowed by the 1:100,000 standard.

d. When digitizing, it is suggested that each 1:100,000 tile should have a maximum root mean square error tolerance of <.005 digitizing inches (–13). Imbedded, missing, or extra polygons in vector coverages must be corrected and cleaned.

1a. Land Cover Coverage

1. Use the National Vegetation Classification System (NVCS) to identify alliances or aggregations of alliances to be mapped in the project area. Note that descriptions of alliances are currently available for the East and will be available for the West approximately end of calendar year 1998. See the Land Cover chapter for additional information on application of the NVCS, or contact the national office.

2. Use Landsat TM imagery at 30 meter pixel resolution as the data source for map development. Do not resample the imagery to a different pixel size.

3. Use the methodology and standards described in the Land Cover Mapping chapter of the Handbook for both rendering a digital base map of land cover pattern and deriving thematic data with a by-type accuracy objective of 80%.

4. Map nonvegetated cover to the minimum categories described in the Land Cover Mapping chapter of the Handbook.

5. Resolution is at least 2 ha* although a tiled, vectorized version of the coverage must be provided that may use a coarser resolution to achieve vectorization and allow efficient operation of a tile on the current version of PC ArcView. The "Merge" program available from National GAP is recommended for aggregation as needed. * Note that spatial versus thematic resolution will be discussed in detail in the updated Land Cover Mapping chapter. Spatial resolution will generally be a result of the thematic level mapped and the nature of the plant community itself. The goal should be to achieve a reasonable portrayal of the mapped units geographic extent and distribution patterns, which may result in variable spatial resolutions by type, ranging from a few pixels to several hundred hectares.

6. Projects shall map land cover beyond their project boundary for a distance of 10 kilometers or to the edge of their TM imagery, whichever is less, provided the entire project area is mapped. Projects may map further than this requirement if they wish and are encouraged to discuss an appropriate amount of overlap with adjoining states.

7. Conduct an accuracy assessment of the land cover map to provide a measure of accuracy by mapped type. Follow the Land Cover Assessment chapter guidelines in the Handbook.

1b. Land Cover Reference Data

1. Obtain data that represents "ground truth" for use in map creation and assessment (see 1a above). This can be near-remote-sensing-based data such as aerial photographs or videography (digital is recommended), accuracy-assessed large-scale vegetation maps, or field-gathered data. Digital data gathered by the

GAP project are considered a deliverable and should be documented and archived as described in this chapter and the Information Archiving and Distribution chapter.

1c. Land Cover Manual

1. Produce a Land Cover Manual that describes the mapped types in the context of the hierarchical classification scheme. Refer to the Land Cover section of the Handbook for minimum Manual components.

1d. TM Imagery

1. Generally, projects will be supplied with sufficient imagery to create the land cover map. However, many projects obtain additional imagery. If this imagery is obtained with project funds, it must be delivered to the National GAP office. For more details see the Land Cover Mapping chapter or contact the National GAP office.

2a. Database File of Predicted Species Occurrence in EMAP Hexagons

1. Use the TNC standard list of names and codes available from the TNC home page (see the link on the GAP home page). Note that around 1/98 we anticipate switching to the Integrated Taxonomic Information System (ITIS) system (http://www.itis.gov/) and will update this standard at that time. This caveat applies also to all other animal naming and coding standards in this chapter.

2. Include, at a minimum, all terrestrial vertebrates that breed or use habitat for a substantial part of their life history in the study area, including migration stopovers.

3. Use the Environmental Protection Agency (EPA) Ecological Mapping and Assessment Program (EMAP) hexagon grid (~635 km2) as the geographic unit (available from the GAP home page).

4. If using the hexagons as the source for delimiting range extent in the predicted vertebrate distributions, follow the attributing protocols found as an appendix in the animal modeling chapter of the Handbook.

5. If not using the hexagons in the predicted distribution modeling process, then at a minimum intersect the predicted distribution maps with the hexagon coverage to attribute the grid with GAP-predicted occurrence of each species mapped.

6. Export the resulting INFO file to dBASE such that each hexagon cell is represented by its ID number followed by the occurrence attribute for each modeled animal species.

2b. Database of Specimen Collection or Sighting Localities (if available)

1. Obtain available digital vertebrate point locations to use in vertebrate range mapping and verification (see Vertebrate Modeling and Assessment chapters).

2. Develop a digital database of locations from hard copy tabular or map products when this can be done in partnership with other institutions.

2c. Database of Species Habitat Associations:

1. Use the peer-reviewed literature, existing habitat association databases, and expert opinion to attribute each species with its known habitat associations.

2. Enter all associations into a database such that each association for each species is referenced to the source of that information.

2d. Predicted Distributions of Terrestrial Vertebrate Species:

1. Use the ECOMAP sections or subsections as available from USFS to bound regional differences in habitat associations for each species if they are known.

2. Develop a distribution map for each species with an accuracy objective of 80% based on comparison to high-confidence checklists for sample areas.

3. Resolution will be variable and a consequence of the resolution of the various coverages used in the modeling.

4. Conduct an accuracy assessment of the species distributions according to the guidelines in the Animal Distribution Assessment chapter of the Handbook.

2e. A "hyperdistribution" of all species in a single coverage:

1. Create a raster coverage such that each cell is attributed for presence/absence of each modeled species. This coverage allows simple queries of richness or coincidence of any group of species. Other attributes are allowed but not required at this time (see the vertebrate modeling chapter).

3a. Land Stewardship Coverage:

1. Map land management authority by the major categories of

    a. Publicly owned lands, attributed by the name of the agency that is responsible for managing the land parcel, and secondarily with the actual owner if different from the managing entity.

    b. Privately owned lands categorized in status 3 or higher when identified voluntarily by the land owner.

    c. All other privately owned lands are to be attributed "private."

    d. Subdivide general ownership/management categories with internal management boundaries that will have a different management status ranking, for instance, a Research Natural Area is an internal management boundary within Forest Service lands.

2. The resolution is at least 16 ha (40 acres), but smaller units may be used.

3. The management authority and management status attributes shall occur as one GIS coverage that can be queried to produce separate maps. Use the code system in the Stewardship chapter.

4. Categorize management according to the methods described in the Stewardship chapter of the Handbook using the status levels 1-4 described therein. You may attribute with subcategories as described in that chapter.

3b. Status 1 & 2 Management Areas:

1. Document the name of the area, date of establishment under its current status (an area may have been established as a national monument and then upgraded to national park; use the date of establishment of the latter), managing entity, owner if different than the managing entity, and source of management information for status 1 & 2 areas. Documentation for status 3 tracts is encouraged but not required.

2. Enter the data in a database program.

4a. Analysis Statistics for Land Cover and Animal Species:

1. Intersect the land cover and animal distribution coverages with the stewardship coverage and export the resulting coverage table to a database file.

2. For each element (land cover type and animal species), report its total project area distribution in square kilometers and the area and percent of its total distribution for each major land steward by management status level (see example in the Analysis chapter of the Handbook). Identify elements that have less than 10%, 20%, and 50% of their distribution in status 1 & 2 areas. Identify the measured accuracy level of each element.

5. Ancillary Data Used to Create Gap Coverages:

1. All data must be obtained from a referenced source with known methods used to create it and the location where the original data may be obtained.

2. Text, tabular, and hard copy map data must be referenced as typical for literature citations (see the GAP Handbook citations for style).

3. GIS data must include metadata sufficient to determine its fitness for use in deriving GAP products.

6. The Final Project Report:

1. Develop the final project report using the standard final report template available from the National GAP office (see Report chapter in this Handbook).

2. It is recommended that projects use the template during project production phases and annual reporting so that all pertinent information is captured when it is fresh and production staff are still available.

7. Graphic Image Files of All Coverages:

1. All coverages, including individual animal ranges by hexagon and predicted distribution shall be provided as a graphic image in addition to the GIS coverages. **Maps and figures using less than 256 colors must be in .GIF format. Files using more than 256 colors must be in either .TIF or .JPG format.** This is for CD-ROM production. Format guidelines are described in Appendix 1.

Acknowledgements:

We thank the following reviewers for their contributions to this document.

- Robert Deitner
- Daniel Fitzpatrick
- Liza Fox
- Steve Hartley
- Frederick Limp
- Leonard Pearlstine
- Roly Redmond
- Kathryn Thomas

Literature Cited:

Anderson, J. R., E. E. Hardy, J. T. Roach, and R. E. Witmer. 1976. A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey Professional Paper 964, U.S. Government Printing Office, Washington, D.C. 28 pp.

Cowardin, L. M., V. Carter, F. C. Golet, and E. T. LaRoe. 1979. Classification of wetlands and deepwater habitats of the United States. U.S. Fish and Wildlife Service Office of Biological Services No. 79/31. U.S. Government Printing Office, Washington, D.C., stock number GPO 024-010-00524-6. 103 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 Monograph 123. 41 pp

Scott, J. M., and M.D. Jennings. 1994. A Handbook for Gap Analysis. National Gap Analysis Program, University of Idaho, Moscow, ID 83843, http://www.gap.uidaho.edu/gap/.

National Gap Analysis Project Standards

Appendix 1 - Data Delivery Standards for GAP

General:

Media: All digital information is to be delivered on CD-ROM. A standard directory structure is forthcoming and will be provided as an appendix to this chapter.

File naming: In file naming described below, "XX" denotes the 2-digit mail code for the state project, e.g., ID means Idaho. For projects not conforming to single state boundaries, consult with National GAP on an appropriate name.

GIS coverage format: Note, this item is highly dynamic, please consult with your national coordinator prior to delivery. All GIS coverages are to be delivered in ARC/INFO GRID and vector export (.E00) format when indicated. When SDTS (see http://mcmcweb.er.usgs.gov/sdts/) becomes operational, we will adopt that federal standard. If tiling is required, as noted for each coverage, use the USGS 30 x 60 minute quadrangle boundaries. Though some very small tiles may result from a quad boundary falling just inside a project area boundary, please use the exact quad boundaries.

GIS coverage projection: Deliver coverages in the projection they were produced *and* that is commonly used in the state and compatible with the projection of ancillary data commonly available in the state; this is usually UTM. If coverages need to be transformed, all information on the transformation process, such as the original coordinate system and parameters of projection (e.g., units, parallels, meridians, projection origin, false easting or northing), must be documented. Coverages will be re-projected by National GAP to match regional parallels and meridians, and all original projection information is needed to perform this transformation accurately.

Metadata:

Proper documentation of information sources and processes used to assemble GAP data layers is central to the successful application of GAP data. Metadata documents the legacy of the data for new users. The Federal Geographic Data Committee (FGDC 1994, 1995) has published standards for metadata and NBII (http://www.nbii.gov) has updated those standards to include biological profiles. Executive Order 12906 requires that any spatial data sets generated with federal dollars will have FGDC-compliant metadata.

Each spatial data layer submitted must be accompanied by its metadata (*.html file) in the same directory. You must also include an additional directory (called 'meta_master') which will include each metadata file in three forms (*.txt, *.html, and *.sgml). These are readily created in MetaMaker (http://www.nbii.gov/about/factsheet5.html). The redundancy in format is to provide one file for error checking (*.txt) one for presentation on the Internet (*.html) and one for indexing elements for the spatial data clearinghouse (*.sgml). Remember, metadata describe the development of the spatial dataset being documented. If there are companion files to the GIS data, use metadata to reference (reports, spreadsheet, another GIS layer).

USGS personnel conduct metadata training to meet FGDC standards and to include biological data. See the Internet site, http://www.nbii.gov/metadata/training/index.html for more information.

Database files: All database files other than INFO files should be delivered in dBASE format.

Graphic files: All graphics files for maps and figures using less than 256 colors must be in .GIF format. Graphics files using more than 256 colors must be in .TIF or .JPG format.

Text documents: Other than "readme" files as described below, documents should be delivered in either Microsoft Word or WordPerfect format.

Readme files: Every coverage, folder, or directory should contain a "readme" file (readable on both PCs and UNIX workstations) that identifies the other items required to use that product. For example, all readme files should identify the final report and, for GIS coverages, the metadata file and any necessary database or INFO files. This will ensure that the user has been made aware of all items they should acquire to use the product intelligently.

Deliverables directory: The CD should have a readme file titled "directory" that lists all delivered items by file name or logical name if the item is not digital, e.g., "video tapes." The name should be followed by a brief descriptor. The directory should be arranged exactly according to the actual data directory structure on the CD. We anticipate that the National GAP office will produce a required directory structure which will be provided as an appendix when complete.

1. Land Cover Data

1a. Land Cover Coverage:

Product Type: Digital GIS coverages. You must provide a project-wide coverage in 30 m raster format that includes the required overlap area and both a raster and ARC/INFO vector (export version), tiled using the USGS 30 x 60 minute quads (statewide vector version is optional). The vector-tiled version may be aggregated as needed to assure functionality of individual tiles on desktop ArcView 3.

Coverage Name: "XXgapveg" (coverage directory name)

When tiling the coverage into individual files, use the 3-digit USGS GEODATA code followed by "veg.E00". The 3-digit code is a unique 2-character abbreviation for the 1:250,000 tile, followed by a 1, 2, 3, or 4 to denote which 1:100,000 tile it is (NW=1, NE=2, SW=3, SE=4). The file with the codes is available on the GAP home page.

Attributes Field Names: "Covertype_code" is the internal code the project uses to identify the mapped types, and "Covertype_name" is the logical name used if these are different than the NVCS. "NVCS_code" will denote the code according to the National Vegetation Classification System, and "NVCS_name" will denote the official name. For land cover maps produced prior to NVCS adoption, mapped vegetation types that cannot be directly cross-walked to one of the named alliances must provide the NVCS code and name for the closest match such as the formation group. In addition, the alliances that comprise the mapped type must be listed (codes only) under the field "NVCS_equivx" where the "x" is a numeric value for the number of NVCS equivalent types, e.g., "NVCS_equiv1" is the most dominant alliance equivalent, "NVCS_equiv2" is the second, and so forth. As of the writing of this document, it was anticipated that there would also be an accepted GAP code and name for nonvegetated land cover which should appear under "Covertype_code"; if they are not available prior to your completion date, you may provide your own. "xxxxxxxx_code" should be used to provide equivalents for nonvegetated cover under other established land cover classification schemes where the "x"s indicate a logical name for the scheme for example "Anderson_code" can denote the Anderson land cover classification code.

1b. Land Cover Reference Data:

Product Type: Variable. Analog video tape, digitally captured video frames, digital video, digital database file of coordinates and attributes.

File Name: For the database file use "XXrefdat.dbf"

1c. Land Cover Manual:

Product Type: Digital and hard copy text document

File Name: "XXvegman.doc"

Attributes Field Names: Use the fields in the NVCS alliance descriptions (forthcoming). This is a minimum; other fields may also be provided.

2. Animal Modeling Data

2a. Predicted Species Occurrence in EMAP Hexagons:

Product Type: Digital database file

File Name: "XXhex.dbf"

Attributes Field Names: "hex_id" is the ID number for each hexagon cell. Use the full TNC element code for the animal species as the field names. The attributes will follow the coding system by L. Master (in development as of this writing) and will be accessible through the GAP home page.

2b. A database of specimen collection or sighting localities, if available:

Product Type: Digital database file

File Name: "XXspploc.dbf"

Attributes Field Names: "TNC_code" is the full element code from the TNC database, "TNC_name" is the common name from the TNC database, "State_name" is the common name in state usage if different, "TNC_rank", "State_rank", and "Fed_rank" are the TNC, state, and federal status/ranks, respectively, as of the coverage creation date. "GAP_status" is the percent representation of the element s total distribution in GAP management status 1 & 2 lands expressed as "xxxx" (use 4 digits with the first digit indicating the whole number, e.g., 1000 for 100% or 0xxx for the fraction carried to three decimals, e.g., 0352 for 35.2 percent).

2c. Species Habitat Associations:

Product Type: Digital database file

File Name: "XXhabtat.dbf"

Attributes Field Names: There are no standards at this time except for the citation for each association. Fields used by your state Natural Heritage Program or the Multistates Database are recommended.

2d. Animal Distributions:

Product Type: Digital GIS coverages in 30 or 90 m raster, vector format optional.

Coverage Name: "xxxxxxx" (insert for "x"s the 2nd through 9th numeric digits of the element code used by TNC. We do not anticipate the need to tile individual species maps.

Attributes Field Names: use the same identification codes as in 2b.

2e. A "hyperdistribution" of all species in a single coverage:

Product Type: Digital GIS coverage in 30 m raster format.

Coverage Name: "XXanimal"

Attributes Field Names: use the same identification codes as in 2b.

3. Land Stewardship Data

3a. Land Stewardship Coverage:

Product Type: Digital GIS coverages in both vector and 30 m raster.

Coverage Name: "XXstwrd"

We do not anticipate the need to tile this coverage. If needed, use the same convention as for land cover but substitute "stwrd" for "veg" (see 1a above).

Attributes Field Names: The required fields are as follows; their descriptions and instructions for populating them are found in the current version of the Handbook Stewardship chapter: Manager, Owner, Division, Unit, Status, and if desired, Status_sub.

3b. Status 1 & 2 Management Areas Reference Data:

Product Type: Digital database file

File Name: "XXmgmt12.dbf"

Attributes Field Names: "Unit" is the official name of the management area, "Date" is the date the area was established, "Manager" and "Owner" are the same as the Land Stewardship coverage attributes, "Source" is the source of information used to assign the management status.

4. Analysis Data

4a. Analysis Statistics for Land Cover and Animal Species:

Product Type: Digital database file

File Name: "XXstats.dbf"

Attributes Field Names: "Covertype_code" is the numeric code used to identify mapped land cover types and "TNC_code" is for the animal species, "Area" is the total mapped distribution of the element, "xxxxx_area" is the area in square kilometers of each element s distribution by manager category where the "x"s are the "Manager" code from the Stewardship coverage, "Statusx_area" is the area in square kilometers of each element s distribution by management status category where "x" is the 1-4 status code. If you used subcategories, those can be included, e.g., "Status3b_area."

5. Ancillary Data:

Product Type: Digital GIS coverages in 30 m raster and vector if they originated in vector or were converted for project purposes, tabular data in database format.

File Name: For coverages use "XXxxxxxx" where "x"s indicate a logical name such as "road" or "stream." For other tabular data use a logical name followed by ".dbf."

Attributes Field Names: There are no GAP standards.

6. The Final Project Report:

Product Type: Digital text file

File Name: "XXreport.xxx" where "x"s indicate the file extension of the native word processor program. Reports must be in either Word or Wordperfect (PC, not Apple versions).

7. Graphic Image Files of All Coverages:

Product Type: Digital graphic image file in JPEG, TIF, or GIF format.

File Name: Use the same file names as the coverages except use ".jpg," "tif," or "gif," for the file extension to denote that it is a graphic file, not a coverage.

Appendix 2 - Project Close-out Protocol

Version: 2.0.0
Date: 6/20/00

All time frames are stated in months before close-out.

1. At three months you will be notified of your impending completion date, and a list of deliverables will be provided. If there is any change in the completion date or question about deliverables, please call your coordinator at the National GAP office.

2. At two months provide to the National GAP office for review a hardcopy of the draft final report, and all digital data on CD-ROM. The draft report will be reviewed by your coordinator at the national office and selected others; the data will be examined at the University of Idaho GIS lab for consistency with project standards.

3. At one month or sooner you will be provided with a review of the report and list of any data deficiencies.

4. After making all necessary changes, resubmit the final CD with the final project report on it in WORD format, and two hard copies of the report.

5. After final data review, a CD will be sent to the EROS Data Center (EDC) for archive and Internet distribution along with a digital version of the report (see below).

6. The National GAP office will produce both HTML and Adobe Acrobat versions of your final report and graphics and master on CD-ROM. Approximately 200 copies will be pressed (the number will be decided in consultation with the project). You will be provided with the CDs to distribute free within your state. If you have a state GIS distribution center, they may request additional copies to sell under a negotiated arrangement.

7. The CD masters will be sent to a "press-on-demand" center linked through the GAP home page for the public to order CDs and "print-on-demand" hard copy reports.

Gap Analysis

Version: 2.13
Date: 16 February 2007
Peer-reviewed: No

Patrick Crist
Idaho Cooperative Fish and Wildlife Research Unit
Moscow, ID

as adapted from Handbook version 1 by

Blair Csuti
Idaho Cooperative Fish and Wildlife Research Unit
Moscow, ID

- Appendix 1 - Distribution of Wolverine in Wyoming (courtesy WY-GAP)

Introduction

The term and concept of gap analysis was defined in the Preface to this handbook, but briefly, a "gap" is the lack of representation or inadequate representation of a biotic element (plant community or animal species mapped by GAP) in areas managed primarily for natural values. Identification of a gap indicates potential risk of extinction or extirpation unless changes are made by land stewards in the management status of the element. Biodiversity has always been a product of a changing environment, subsequent evolution, and natural extinction events. The need to identify potential gaps stems not so much from the natural pace of evolution and extinction, but rather from recent and accelerating effects of human-caused changes to the natural environment that threaten biodiversity. Noss and Cooperrider (1994) identified four categories of human influences on biodiversity: direct exploitation, indirect exploitation, disturbance, and indirect disturbance such as global warming. Of these, indirect exploitation and disturbance are the primary causes of the alteration of habitat type and structure which are probably the most prevalent, largely irreversible, and controllable of human influences (Bingham and Noon 1997, Noss and Cooperrider 1994). The likely presence or potential for these influences are identified and categorized by GAP into four levels of management status to maintain biodiversity (see Stewardship chapter). Therefore the purpose for conducting gap analysis is to identify gaps that have great potential for mitigation by land stewards and decision-makers. Another substantial contribution of gap analysis is the provision of regional or range-wide analysis of an element's representation that can allow stewards to assess their role and responsibility in the greater context for those elements occurring on their lands. In other words, rather than basing stewardship decisions only on local, observable conditions, stewards can take into account an element's status over a larger geographic region (their state or ecoregion) or throughout an element's mapped range.

Objectives

1. To determine the representation of each mapped element (land cover types and animal species as identified in the land cover and animal modeling chapters) in different categories of land stewardship (management authority) and management status.
2. To interpret the results from Objective 1 in a way that is useful for land stewards in their development of land use and management plans that conserve or restore those biotic elements or avoid causing further impairment.
3. To provide the analysis and data sets in suitable form for additional modeling, analyses, biodiversity assessment, and planning activities by a broad group of users.

Background

There have been a variety of analyses proposed for and by GAP to aid in the identification of both elements at risk and geographic areas that can be managed to reduce those risks. As experience and science progress, the proposed analyses have changed to reflect current knowledge and capabilities. This revised version of the original chapter focuses on the basic Gap Analysis Program requirements, which call only for identification of those biotic elements with potential gaps (see GAP standard final report). It is worthwhile, however, to briefly review some of the previous recommendations and objectives.

The identification of "species-rich areas" as an additional filter was described in Scott et al. (1993) and became widely associated with GAP as focusing on "species-rich hot spots". Since then, however, studies show that the species-rich areas of major taxa do not co-occur (Prendergast, 1993, Lawton et al. 1994). This emphasis has been replaced with the goal of identifying complementary sets of geographic units that could represent all (mapped) elements, rather than "hot spots" that would always leave out some elements (Williams 1996). Complementarity analysis is, however, still in the research mode and will therefore be addressed in the Applications chapter (forthcoming). GAP still requires the production of species richness maps for comparative value (see Final Report chapter), however, there is no required analysis at this time, and the subject, therefore, is not covered in this chapter (however see the NM-GAP final report, Thompson, et al. 1996). The ultimate purpose of the application of GAP information by data users is to identify specific geographic locations that can "plug" the gaps through changes in management practices, planning, policy, or stewardship status (Scott et al. 1993). The identification of such geographic locations, however, is relevant to the process of reserve selection which is a complex issue and is described in the Applications chapter (forthcoming). Many researchers, planners, and managers are beginning to use GAP for conservation planning (Scott and Csuti, 1997) and as the science and techniques evolve, they will be incorporated into future revisions of this chapter when adopted as a GAP standard, or the Applications chapter (forthcoming) when provided as a recommended use of the data.

Methods

Summary of steps and standards

Step 1: Prepare the land cover, animal distribution, and stewardship coverages for final analysis according to the respective chapters' standards.

Step 2: Intersect the land cover and animal distribution (element) coverages with the stewardship coverage so that the element coverages incorporate the stewardship boundaries and attributes.

representation of individual elements in each stewardship and management category.

Step 4: Prepare additional analyses as applicable from the list in the standard final report.

Step 5: Incorporate the results into the standard final report and generate example maps for elements of note or state interest.

The following narrative provides details for each step in the process:

Step 1: Prepare the land cover, animal distribution, and stewardship coverages for final analysis according to those chapters□standards. It is imperative that all coverages are considered to be the final deliverable versions prior to beginning the analyses. If correctable errors are detected during or after analysis, the analysis must be either be run again so that the results apply to those delivered coverages or, if the error was not actually corrected, it should be identified in the metadata and report.

Step 2: Intersect the land cover and animal distribution (element) coverages with the stewardship coverage so that the element coverages incorporate the stewardship boundaries and attributes. This intersection will subdivide the element distributions by the stewardship boundaries and apply the stewardship attributes to each segment. The resulting attribute table will allow the calculation of percent of total element distribution by ownership and management categories and support numerous other analyses.

Step 3: Using statistics from the above intersections, generate a table reporting representation of individual elements in each ownership and management category as described in detail later in this chapter. The representation of elements by stewardship categories should be reported per Table 1 (one for land cover and one for animal species). The table for mapped animals should be organized by taxonomic group. Note that text is added for explanation in producing the table, you may alter the text within the table to that needed for the reader.

Table 1

Suggested caption: This table provides the gap analysis results for each natural/semi-natural land cover type.The first column under each major land steward is the area of distribution in square kilometers (multiply by247 for acres) by GAP management status. The second column is the percent representation (area in columndivided by total state distribution). Statewide totals are included along with measured map accuracy of theelement. This table may be used to identify important land stewards for each element, their relativeresponsibility for management as a function of area and their current intent to manage for the element as afunction of distribution by GAP management status.

Projects: Please include all important stewards plus private land and a "misc." category if needed so totalsums to the total mapped area of the element. Use GAP standard fonts (see report instructions) and min 10pt font size. The use of a calculation template in a spreadsheet or database program will make it simple togenerate percentages and totals.

IA8Nc3 Picea Glauca Forest Alliance: Fuzzy set level 5 Accuracy 73%

	USFS		BLM		NPS		Tribal		State DNR		Misc Public		Private		Total	
Status 1	28.07	15.98	0.00	0.00	6.39	3.64	0.00	0.00	0.00	0.00	0.00	0.00	2.23	1.27	36.69	20.88
Status 2	23.09	13.14	5.37	3.06	0.00	0.00	3.20	1.82	2.83	1.61	0.00	0.00	1.98	1.13	36.47	20.76
Status 3	16.00	9.11	2.20	1.25	0.00	0.00	11.40	6.49	17.10	9.73	6.21	3.53	0.00	0.00	52.91	30.12
Status 4	0.00	0.00	0.00	0.00	0.00	0.00	6.50	3.70	0.00	0.00	0.00	0.00	43.11	24.54	49.61	28.24
Total	67.16	38.23	7.57	4.31	6.39	3.64	21.10	12.01	19.93	11.34	6.21	3.53	47.32	26.94	175.68	100.00

The above table should appear as an appendix because of its length, but the standard final report asks for lists of species with 0-<1%, 1-<10%, 10-<20%, and 20-<50% of their distribution in management status 1 or 2 within the body of the report. A workshop convened in fall 1995 concluded that, without better methods for establishing adequate levels of conservation representation, these arbitrary levels, which have been published in the literature, will allow the data user to judge which species may be at risk depending on the criterion applied. (Noss and Cooperrider 1994, Noss 1991, Odum 1972, Specht, Roe, and Boughlon 1974, Ride 1975, Miller 1994).

This information can also be easily extracted from a query of the intersection table. For example, for element X, sum the area of its distribution in status 1 and 2 and divide by the total area for element X. Select for elements with resulting value <0.1, <0.2, and <0.5 and report per Table 2. Also, in that table, provide the status ranks according to federal, state, and The Nature Conservancy (TNC) systems. Narratives that explain or hypothesize about the results for select elements are important to give meaning to the raw numbers. There is currently no standard for which elements should have narratives or what the narratives should contain, but see the WY-GAP final report (Merrill et al. 1996) for an excellent example. We expect to provide more specific guidelines on this as the collective experience of GAP expands and our cooperators indicate what information is of most use to them.

Table 2. Elements with less than 0-<1, 1-<10%, 10-<20%, and 20-<50% representation in GAP management status 1 and 2 and their status according to federal and state governments and The Nature Conservancy (TNC) as of the date of analysis.

Element code	Element name	Federal Status	State Status	TNC Status	accuracy	< 1%	< 10%	< 20%	< 50%
IA8Nc3	Picea glauca Forest Alliance				73%			13.3%	

Step 4: Prepare additional analyses as applicable from the list in the standard final report. Report the results per Table 1. As of this writing, the optional analyses include:

- Analysis of Important Statewide Species Assemblages. For example, if migratory waterfowl are of special state interest, report the analysis results for the group.
- Analysis of Statewide Endemics.
- Special Features Analysis. For example, riparian vegetation as a group can be analyzed.
- Ecoregional Analyses. GAP encourages results to be reported by Bailey’s 1995 Provinces or Sections, especially if such a unit’s sole U.S. representation lies wholly within the study area. If available, the multiagency ECOMAP (McNab and Avers, 1994) subsection boundaries should be used to develop the ecoregion boundary. Analysis shall be reported the same as described in step 3, but will identify the ecoregion as the area of analysis.

The statistics necessary for these analyses should be readily extractable from the attribute table resulting from Step 3. The Ecoregional Analyses will require the additional intersection of the ecoregion boundaries to segment element distributions and stewardship information by those units.

Step 5: Incorporate the results into the standard final report and generate example maps for elements of note or state interest. The above tables and lists are the required items for the report, but example maps showing the element’s distribution colored (or patterned) by management status and ownership categories are useful graphics (See Appendix 1), Maps of elements can be selected for scientific interest, such as species believed either to be especially at risk, or not at risk and shown otherwise by the gap analyses, or species generally of interest in the state. Refer to the standard final report template for resolution of inconsistencies with this document and for formatting requirements, with the final report template taking precedence.

Limitations

The limitations, described in the Gap Analysis Program Description (see the GAP home page) and the chapters for land cover, animal modeling, and stewardship, should be consulted in addition to the following:

- There is no accuracy assessment of the analyses results. Measures of mapped accuracy for element distributions is provided only as a guide for the general reliability of the results. Further research is needed on how to interpret the analysis results for a particular element relative to the mapped accuracy. In some cases, the conservation status of an element may be either better or worse than reported, depending on the nature of the mapping error (see Animal Modeling, and Land Cover Assessment chapters). It is for this reason we state that no on-the-ground actions should be taken based solely on these analyses, and that ground-based surveys should be conducted prior to taking a specific management action.
- Because historic distribution of an element is usually unknown, the measure of percent of total distribution in high stewardship status may not account for a large loss in its historic range (see Noss et al. 1995). In other words, if an element has already been reduced by 90% from its historic distribution, and gap analysis indicates 50% occurrence in status 1 or 2 areas, the result is that only 5% of its historic distribution is represented in conservation lands.
- GAP currently does not predict element viability. For most species and plant communities, viability measures - such as habitat quality and species abundance - and population trends are unknown and cannot be assessed given current knowledge (Boyce, 1992 and Shaffer, 1994). Therefore, GAP only provides information on representation, with the objective of highlighting at-risk species.

Generally, analysis of animal species must be used with more caution than analysis of land cover types for various reasons, including: a) land cover maps are actual, typically with a statistically valid accuracy assessment, while animal distributions are predicted and difficult to validate; b) land cover types are stationary and change relatively slowly, while animal species are mobile and expand and contract geographically over relatively short time spans; and c) effects of management status on land cover types are generally easier to predict than on animal species.

Continuing Research

To date, GAP has devoted substantial effort to the science of producing the element maps, but has only recently had sufficient data to devote more effort to the analysis phase. Concurrently, the objectives for the analysis, and thus the techniques needed, have changed. Continuing research for developing more resolved and accurate element distribution maps will always be required to increase confidence in the analysis results, but specific research needed for analysis includes:

- Providing interpretations of the reported map accuracy for an element relative to its management status.
- Providing better stewardship information to allow analysis that includes better measures of risk for individual elements. For example, some species will thrive in status 3 areas while others may perish in status 2 areas.
- Incorporating a time factor that can account for current management status relative to risk. For example, elements may occur in status 4 land that is currently well managed and not foreseeably threatened with habitat conversion. This information could help focus priorities on species that are more imminently threatened. Some initial work in this area was conducted by CA-GAP (Davis and Stoms 1996).

- Developing maps of historic range or distribution to better assess the amount of historic loss or gain in an element’s distribution relative to GAP status.

For current research being funded by GAP and its cooperators, see the Research section of the home page.

Literature Cited

Bailey, R.G. 1995. Description of the ecoregions of the United States. 2nd ed. Rev. and expanded (1st ed. 1980). Misc. Publ. No. 1391 (rev.), Washington, DC: USDA Forest Service. 108 p. With separate map at 1:7,500,000.

Bingham, B.B., and B.R. Noon. 1997. Mitigation of Habitat "Take": Application to Habitat Conservation Planning. Conservation Biology, pg 127-139, v. 11 n. 1.

Boyce, M. S. 1992. Population viability analysis. Annual Review of Ecology and Systematics 23:481-506.

Davis, F.W. and D.M. Stoms. 1996. Sierran Vegetation: A Gap Analysis. Sierra Nevada Ecosystem Project: Final Report to Congress, v II. Assessments and Scientific Basis for Management Options. Davis, University of California, Centers for Water and Wildland Resources pg 671-689.

McNab, W.H. and P.E. Avers. 1994. Ecological subregions of the United States: Section Descriptions. Administrative Publication WO-WSA-5. Washington, D.C.: U.S. Dept of Agriculture, Forest Service.

Merrill et al. 1996. The Wyoming Gap Analysis Project Final Report. Wyoming Cooperative Fish and Wildlife Research Unit, University of Wyoming, Laramie, WY.

Miller, K.R. 1994. In J.A. McNeely and K.R. Miller, editors. National parks conservation and development. Smithsonian Institution, Washington, D.C.

Noss, R.F. 1991. Report to the Fund for Animals in Washington, D.C.

Noss, R.F., and A.Y. Cooperrider. 1994. Saving nature’s legacy. Island Press, Washington, D.C.

Noss, R.F., E.T. LaRoe, and J.M. Scott. 1995. Endangered Ecosystems of the United States: A Preliminary Assessment of Loss and Degradation. U.S. Department of the Interior, National Biological Service. Washington, D.C.

Odum, E.D., and H.T. Odum. 1972. Proceedings of the North American Wildlife and Natural Resources Conference 39:178.

Prendergast, J.R., R.M. Quinn, J.H. Lawton, B.C. Eversham, and D.W. Gibbons. 1993. Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature* 365:335-337.

Ride, W.L.D. 1975. In F. Fenner, editor. A national system of ecological reserves in Australia. 64 p.

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.

Scott, J. M., and B. Csuti. 1997. Gap Analysis for biodiversity survey and maintenance. Pages 321-340 in M. L. Reaka-Kudla, D. E. Wilson, and E. O. Wilson, editors. BioDiversity II: Understanding and protecting our biological resources. Joseph Henry Press, Washington, D.C.

Shaffer, M. L. 1994. Population viability analysis: determining nature’s share. Pages 195-196 in Principles of Conservation Biology (Gary K. Meffe and C. Ronald Carroll, editors). Sinauer Associates, Inc., Sunderland, Massachusetts.

Specht, R.L., E.M. Roe, and V.H. Boughlon. 1974. Australian Journal of Botany Supplement Series. Supplement No. 7.

Thompson, B.C., P.J. Crist, J.S. Prior-Magee, R.A. Deitner, D.L. Garber, and M.A. Hughes. 1996. Gap Analysis of Biological Diversity Conservation in New Mexico Using Geographic Information Systems. New Mexico Cooperative Fish and Wildlife Research Unit. New Mexico State University, Las Cruces, NM.

Williams, P., D. Gibbons, C. Margules, A. Rebelo, C. Humphries, and R. Pressey. 1996. A comparison of richness hotspots, rarity hotspots and complementary areas for conserving diversity using British birds. Conservation Biology 10: 155-174.

Appendix 1 - Distribution of Wolverine in Wyoming (courtesy WY-GAP)

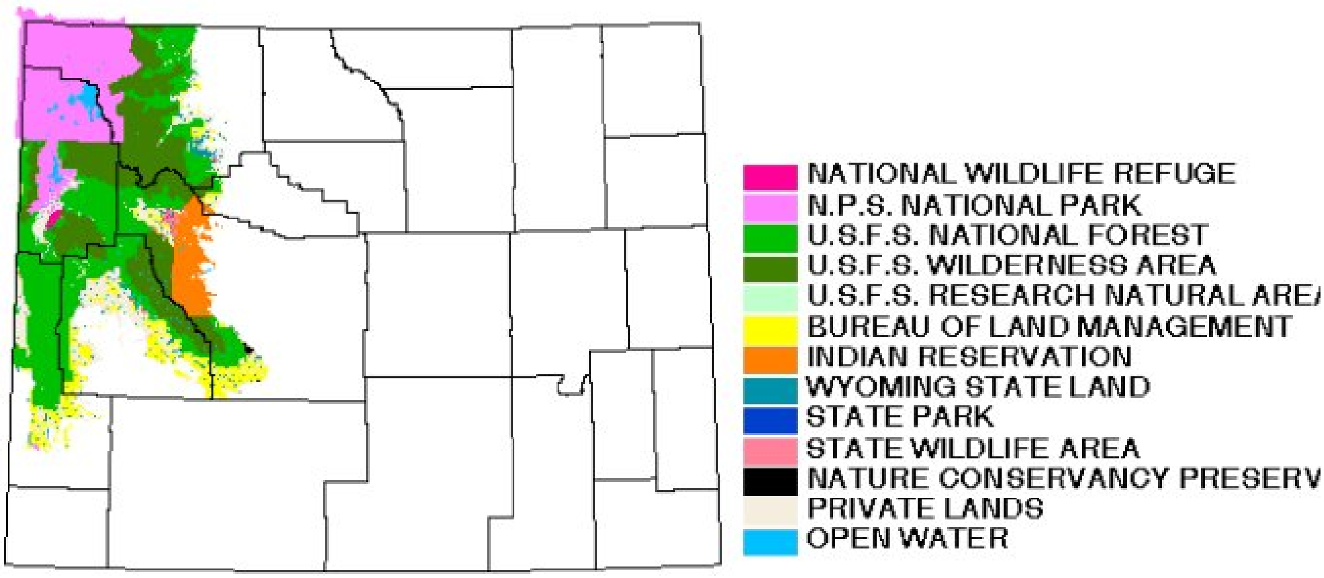


Figure 1: Distribution by Management Authority

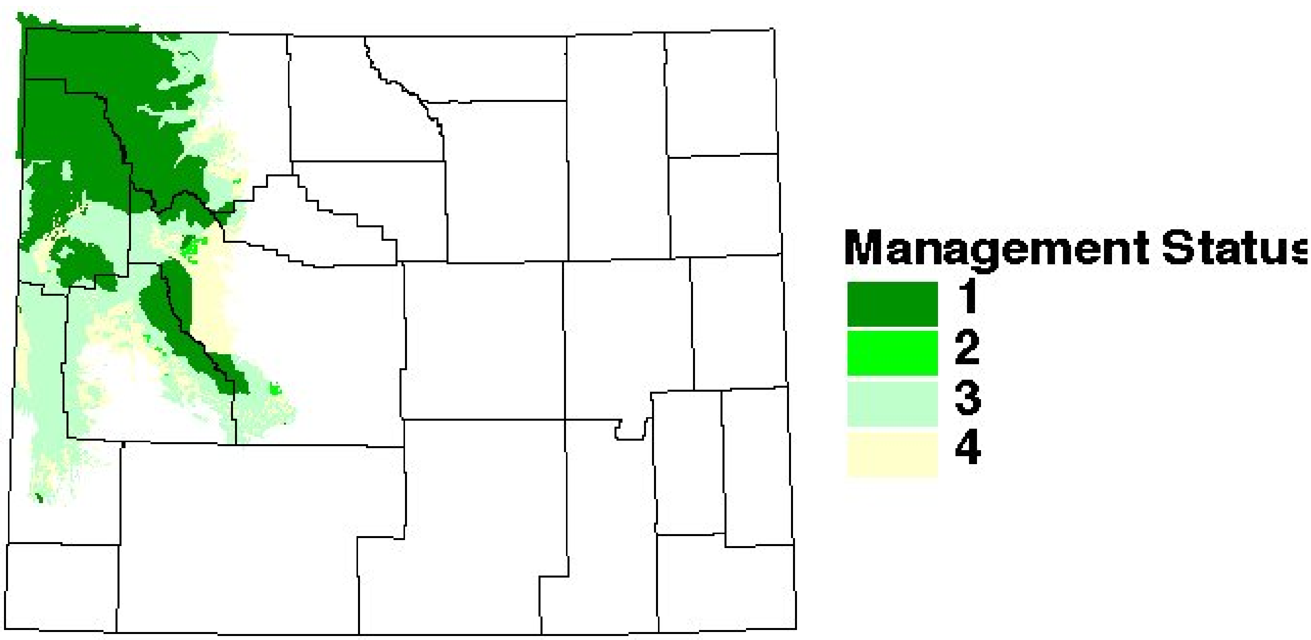


Figure 2: Distribution by Management Status

Project Reporting Requirements

Version: 1.0.1
Date: 13 Febuary 2007
Peer-reviewed: No

There are two types of reporting required from all GAP projects. This document describes those requirements to aid understanding of their purpose and our expectations. We know that reporting is a chore and have endeavored to make the process as painless and productive as possible. Suggestions on how to improve the process are always welcome. Reports are expected on time. The annual Bulletin status statement is not a substitute for your Progress Report. Funding requests or project closeout will not be processed until reporting obligations are met.

Annual Progress Report

The purpose of this report is to provide National GAP and BRD staff with an understanding of the progress since the last report for each component of the project, plus status overall and planned progress for the following year. A secondary purpose is to aid development of the final project report by capturing information as it is produced rather than relying on memory at the end of these multiyear projects. For example:

"Since June 1999, six TM scenes were processed, labeled, ground-truthed and merged into the previously completed three scenes. The project land cover map is now 75% complete with three scenes remaining (see figure 1). These should be completed by November 2000."

You should of course include more detail than that and report any difficulties that resulted in time delay or unexpected costs that may impact the project--the national program as well as you can learn from these experiences. The progress report should follow the same structure as the standard final report template but should not include any boiler plate text. Please include an executive summary and graphic evidence of the mapping progress (e.g., a TIFF of the completed TM scenes mosaicked together and example animal species and stewardship map). We do not require citations, but if you wish to include them in the "building" of your final report, by all means do so.

Annual Bulletin

This report is the National Program's newsletter to all program researchers, cooperators, and agency staff. It describes program status at the state level as well as overall and communicates developments in methods with the GAP community. Projects are provided with a reporting template once per year to concisely describe status.

Appendix 1: the Final Report Template

Reporting the Results of Gap Analysis

Version: 2.1.3
Date: 20 July, 2000
Peer-reviewed: No

Patrick J. Crist and Jill Maxwell
National GAP Office, Moscow Idaho

Introduction

This is the template for the final report of a state Gap Analysis project. It is provided in this Handbook for states to keep abreast of the continuing revisions, however, you must obtain an electronic template from the GAP website. These are available in WORD and can be downloaded from the Report handbook section.

Web master note: The Final Report Template can be found at the bottom of this page.

The reason for having a standard final report format is to provide consistency with the nationwide program so that someone picking up a report from any state will find the information they needed in the same location and same format. We do not, however, want to limit innovation in presenting the results, therefore, consider the contents of this chapter to be a minimum standard. The "boiler plate" text may be changed provided you present the concepts to the same or better degree of detail and intent. You may not, however, change the headings or their order; you may add subheadings as needed.

We strongly recommend that you begin writing your annual reports according to this format so that the final report will gradually "write itself" during the course of the project. Early experience has shown that states have a very difficult time completing the report if the information is not captured while the work is ongoing- staff move on, details are forgotten, etc., and have required about six months to complete the task. Users of the first template found it a big help in organizing and completing the report efficiently, and we hope that by following the above advice and using the template, the task will require no more than 6-8 weeks. Remember, this report will be widely distributed and will stand as THE report on biodiversity in your project area for several years; please make sure the quality level is the highest possible.

Instructions

Note: these instructions are under frequent revision, consult with national GAP prior to beginning your final draft.

Instructions for using the template

1. Your state project is identified as **XX-GAP**, do a universal replace of "**XX**" with your state postal code, for multistate regions or other situations where a postal code does not apply, consult with the national office on an appropriate acronym. For further information on the use of GAP terminology, please refer to the style guide http://www.gap.uidaho.edu/handbook/StyleGuide.
2. 2. For the body text of the report, use the font *Times New Roman* with point size 12. (If possible use the Adobe Type 1 version of the font, not the true type version.) Smaller point sizes and sans serif fonts may be used within tables as necessary. Headings are in this same format with these exceptions:

3. main headings are centered in all caps bold;

   subheadings are in bold with the first letter of every word in caps;

   third level headings are plain text underlined.

   If you should require a fourth level heading (discouraged) then follow the second level format but indent one tab.

   For subheadings and their associated text, select the heading and the first line of text below it and use the *Keep Together* function found in WORD. This is to ensure that your headings will not be left alone on the last line of the page after PostScript file conversion. Chapters (main headings) should start on a new page.

4. All references to other chapters, sections, tables, appendices, etc. within the body text are to be put in parentheses, using the phrase "see _______". For example: "(see Figure 2-2)." Linked text that you wish to put into a different format than the previous format is to be documented in a separate file named "links". This file is to include the page number, paragraph, sentence, and word in which these desired links are to be located-- this word will be hyper-linked within the final report CD.
5. Small tables, charts, and figures may be included within the body of the chapter text, while large tables and appendices should be provided as separate files with location and caption referenced within the body of the text per #3. As a general guide, if the feature can fit on one page it should go within the body, if more than that make an appendix. This holds for multiple associated figures or tables--create an appendix of the collection; this will lead to a smoother read of the chapter.
6. Table cell borders are shown only for clarity in the template, we suggest borders when there are multiple lines of information per cell, or the table is especially complex, otherwise do not use borders.
7. Provide graphics, such as maps, as separate files with location and caption referenced within the body of the text per #3. If too much detail will be lost by placing a map on an 81/2 X 11 page, maps (such as land stewardship and management status) should be place on an 11 X 17 page. Regardless of their size, graphic files must be saved in .GIF, TIF., or JPEG format without compression. GIF if less than 256 colors, otherwise .tif or .jpg.
8. Text files must be in Microsoft WORD saved in a format no older than the version previous to the current release.
9. Tables produced in spreadsheet formats must be imported into WORD. For inline images such as figures, use the "Insert/Object" function and select the "create from file" option in the associated window of this function. **Please note:** table headings should be centered in Times New Roman bold and should appear above the table. Figure headings should be centered in Times New Roman bold and should appear below the figure.
10. If figures are landscaped, the captions should also be landscaped and should be placed below the figure. However, leave the page numbers in the same position as on the non-landscaped pages.
11. Provide each chapter as a separate file. Save each file with a filename containing the prefix "CH_#", where '#' is equal to the corresponding chapter. For example: "CH_1, CH_2, CH_3..."
12. Page numbers should be in the bottom center of the page. Be sure to embed page numbers into each chapter starting with the next page number of the last page of the previous chapter. This can be done choosing "Format" under the "Insert/Page Numbers" function. Appendices should be numbered independently. The required page numbering format for each appendix is "appendix_letter-#". For example: "A-29, C-3, S-21..."
13. For GAP to convert your report into CD and web site formats, it is imperative that the template and instructions be followed. If you have an idea for a deviation, please consult with national GAP

before making any changes. We expect the report to evolve, but random unilateral changes will cause significant cost and time increases for publication. Remember also that most readers will receive the report as an electronic PDF but that the report must also be practical in hard copy format.

14. The report begins with the required template for the cover of the Final Report. Insert your State Name in the title replacing "state name". Custom logos and images are to be contained within the middle window. This is your work space to be creative. Anything outside of the window is to be left in the same format and appearance.

Instructions for submission

1. Please do a thorough in-house review of the report for spelling, grammar, and terminology prior to submitting a draft to the national office and your peer reviewers; a professional technical editor is highly recommended. You are also required to do a peer review by at least two people for every chapter-either two individuals read the entire thing or two people read individual chapters. We recommend your primary cooperators as reviewers since they will be the primary users, but anyone knowledgeable of the state ecology and who did not directly contribute to creating a data set or part of the report would be suitable. Your reviewers can download instructions and a comment form here.
2. According to the GAP Closeout Protocol (see that handbook chapter) submit a hard-copy draft of the report. We will review it and provide comments. We may also have all or portions of it reviewed by others. If you disagree with changes that we feel are necessary for USGS publication, we will insert them with an editor's note within the document in brackets and identify them as such. Call to discuss such items.
3. Provide a final copy in digital form (with graphic files of all coverages) and two unbound hard copies (with sample graphics) as specified in the closeout protocol found in the Standards chapter. If you have funds available, we also strongly encourage you to provide a copy to each of the active GAP projects, but we will provide you with up to 50 free CDs for that purpose. It is also intended that one-off hardcopies will be available for sale nationally (probably through GPO).

Final Report Template

- Microsoft Word (3425 kB)
- PDF (358 kB)

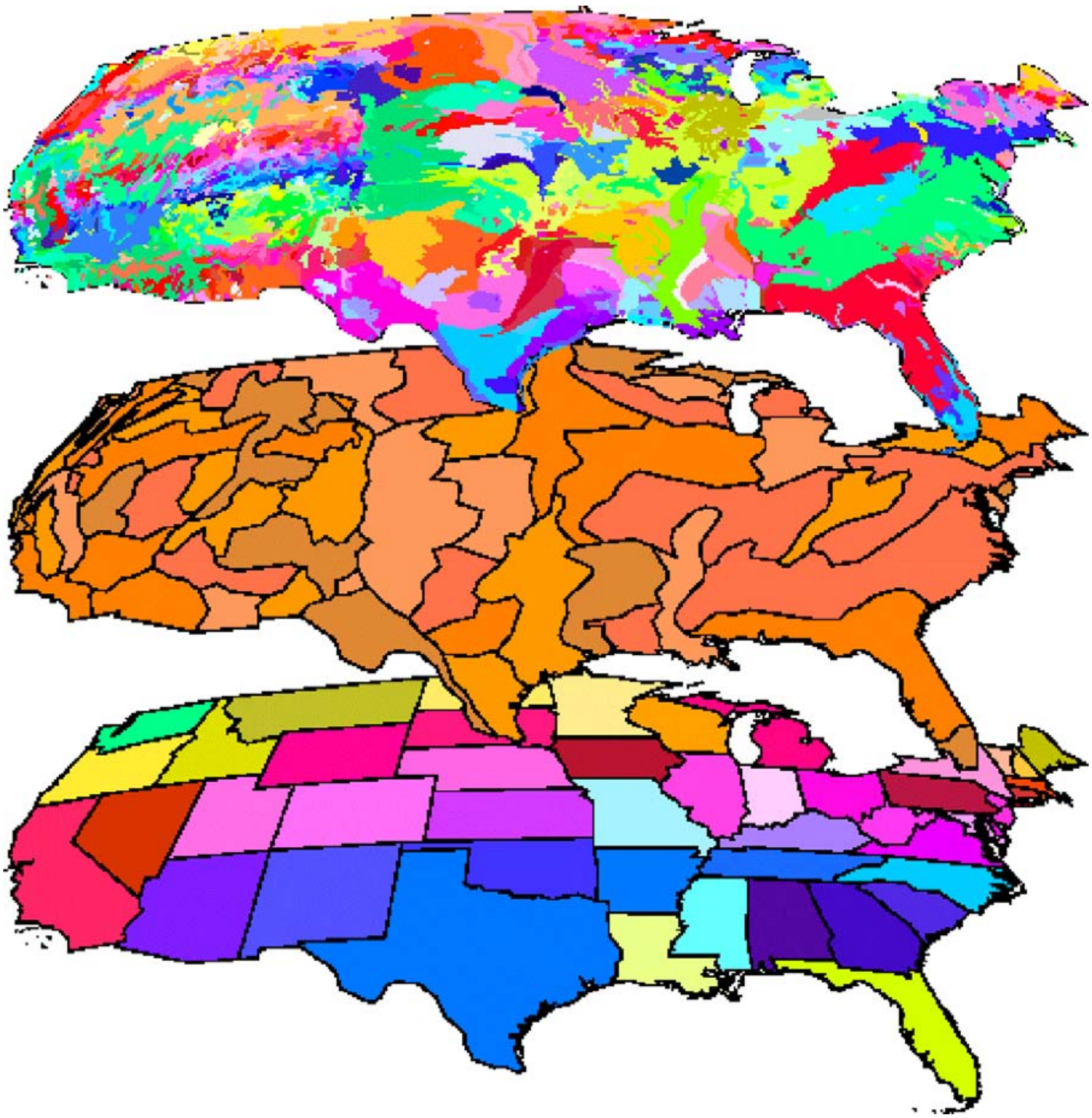
Front Cover

The following page provides a template for the required front cover. Though some elements and margins are not shown per USGS requirements, the template is provided for you to insert the appropriate text and graphics. Your submission will be revised by the national office using a graphics program. If the project logo style is the one you want to use, we can build it for you; if you desire this, delete the graphic and insert "please provide project logo." You may provide an alternative graphic if you wish. You may also wish to include graphics of plant communities and animal species.

> For the remainder of this chapter, all text in italics are instructions, other text should appear in the report.

A GAP ANALYSIS OF XXX*(Insert State Name Here)*

XXXX Final Report *(Insert date here)*



A GEOGRAPHIC APPROACH TO PLANNING FOR BIOLOGICAL DIVERSITY

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

The following information should appear on the title page.

THE *insert state name here*
GAP ANALYSIS PROJECT

FINAL REPORT

insert date of final version here

insert names and affiliations of PI(s), staff, associated researchers here

Contract Administration Through:
insert text

Submitted by:
insert principal investigator

Research Performed Under:
Cooperative Agreement No. *Insert number*
Research Work Order No. *Insert number if applicable*

Insert managing organization and location

U.S. Geological Survey
Biological Resources Division
Gap Analysis Program

Insert national GAP logo and USGS logo
LIST OF PROJECT AFFILIATES
*
*
*

United States Geological Survey Biological Resources Division

TABLE OF CONTENTS

LIST OF FIGURES

You will need to develop a separate list for digital versus hard copy if you are (most likely) providing only example printed maps versus all maps available on the CD version.

LIST OF MAPS

DEDICATION

Insert state text here; see previous reports for suggestions.

EXECUTIVE SUMMARY

Insert state text here. See NM-GAP for example. As a style guide--the summary should be understandable by your congressperson; it should be written to stand on its own and not be copied and pasted sections from your report body. You may treat the report body in more technical terms but it too should be understandable by someone with basic knowledge in the various fields--reserve the most highly technical writing for published journal articles.

INTRODUCTION

Each item is to be followed by specific state text

How This Report is Organized

This report is a summation of a scientific project. While we endeavor to make it understandable for as general an audience as practicable, it reflects the complexity of the project it describes. A glossary of terms is provided to aid the reader in its understanding, and for those seeking a detailed understanding of the subjects, the cited literature should be helpful. The organization of this report follows the general chronology of project development, beginning with the production of the individual data layers and concluding with analysis of the data. It diverges from standard scientific reporting by embedding results and discussion sections within individual chapters. This was done to allow the individual data products to stand on their own as testable hypotheses and provide data users with a concise and complete report for each data and analysis product.

We begin with an overview of the Gap Analysis Program mission, concept, and limitations. We then present a synopsis of how the current biodiversity condition of the project area came to be, followed by land cover mapping, animal species distribution prediction, species richness, and land stewardship mapping and categorization. Data development leads to the Analysis section, which reports on the status of the elements of biodiversity (natural community alliances and animal species), for *insert project area name*. Finally, we describe the management implications of the analysis results and provide information on how to acquire and use the data.

The Gap Analysis Program Mission

The mission of the Gap Analysis Program is to prevent conservation crises by providing conservation assessments of biotic elements (plant communities and native animal species) and to facilitate the application of this information to land management activities. This is accomplished through the following five objectives:

1) map actual land cover as closely as possible to the alliance level (FGDC 1997).
2) map the predicted distribution of those terrestrial vertebrates and selected other taxa that spend any important part of their life history in the project area and for which adequate distributional habitats, associations, and mapped habitat variables are available.
3) document the representation of natural vegetation communities and animal species in areas managed for the long-term maintenance of biodiversity.
4) make all GAP project information available to the public and those charged with land use research, policy, planning, and management.
5) build institutional cooperation in the application of this information to state and regional management activities.

To meet these objectives, it is necessary that GAP be operated at the state or regional level but maintain consistency with national standards. Within the state, participation by a wide variety of cooperators is necessary and desirable to ensure understanding and acceptance of the data and forge relationships that will lead to cooperative conservation planning.

State objectives for GAP: *optional section, insert state text here*

The Gap Analysis Concept

The Gap Analysis Program (GAP) brings together the problem-solving capabilities of federal, state, and private scientists to tackle the difficult issues of land cover mapping, animal habitat characterization, and biodiversity conservation assessment at the state, regional, and national levels. The program seeks to facilitate cooperative development and use of information. Throughout this report we use the terms "GAP" to describe the national program, "GAP Project" to refer to an individual state or regional project, and "gap analysis" to refer to the gap analysis process or methodology.

Much of the following discussion was taken verbatim from Edwards et al. 1995, Scott et al. 1993, and Davis et al. 1995. The gap analysis process provides an overview of the distribution and conservation status of several components of biodiversity. It uses the distribution of actual vegetation and predicted distribution of terrestrial vertebrates and, when available, invertebrate taxa. Digital map overlays in a GIS are used to identify individual species, species-rich areas, and vegetation types that are unrepresented or underrepresented in existing management areas. It functions as a preliminary step to the more detailed studies needed to establish actual boundaries for planning and management of biological resources on the ground. These data and results are then made available to the public so that institutions as well as individual landowners and managers may become more effective stewards through more complete knowledge of the management status of these elements of biodiversity. GAP, by focusing on higher levels of biological organization, is likely to be both cheaper and more likely to succeed than conservation programs focused on single species or populations (Scott et al.1993).

Biodiversity inventories can be visualized as "filters" designed to capture elements of biodiversity at various levels of organization. The filter concept has been applied by The Nature Conservancy, which established Natural Heritage Programs in all 50 states. The Nature Conservancy employs a fine filter of rare species inventory and protection and a coarse filter of community inventory and protection (Jenkins 1985, Noss 1987). It is postulated that 85-90% of species can be protected by the coarse filter without having to inventory or plan reserves for those species individually. A fine filter is then applied to the remaining 15-10% of species to ensure their protection. Gap analysis is a coarse-filter method because it can be used to quickly and cheaply assess the other 85-90% of species. GAP is not designed to identify and aid protection of elements that are rare or of very restricted distribution; rather it is designed to help "keep common species common" by identifying risk far in advance of actual population decline. These concepts are further developed below.

The intuitively appealing idea of conserving most biodiversity by maintaining examples of all natural community types has never been applied, although numerous approaches to the spatial identification of biodiversity have been described (Kirkpatrick 1983, Margules and Nicholls 1988, Pressey and Nicholls 1989, Nicholls and Margules 1993). Furthermore, the spatial scale at which organisms use the environment differs tremendously among species and depends on body size, food habits, mobility, and other factors. Hence, no coarse filter will be a complete assessment of biodiversity protection status and needs. However, species that fall through the pores of the coarse filter, such as narrow endemics and wide-ranging mammals, can be captured by the safety net of the fine filter. Community-level (coarse-filter) protection is a complement to, not a substitute for, protection of individual rare species.

Gap analysis is essentially an expanded coarse-filter approach (Noss 1987) to biodiversity protection. The land cover types mapped in GAP serve directly as a coarse filter, the goal being to assure adequate representation of all native vegetation community types in biodiversity management areas. Landscapes with great vegetation diversity often are those with high edaphic variety or topographic relief. When elevational diversity is very great, a nearly complete spectrum of vegetation types known from a biological region may occur within a relatively small area. Such areas provide habitat for many species, including those that depend on multiple habitat types to meet life history needs (Diamond 1986, Noss 1987). By using landscape-sized samples (Forman and Godron 1986) as an expanded coarse filter, gap analysis searches for and identifies biological regions where unprotected or underrepresented vegetation types and animal species occur.

More detailed analyses were not part of this project, but are areas of research that GAP as a national program is pursuing. For example, a second filter could combine species distribution information to identify a set of areas in which all, or nearly all, mapped species are represented. There is a major difference between identifying the richest areas in a region (many of which are likely to be neighbors and share essentially the same list of species) and identifying areas in which all species are represented. The latter task is most efficiently accomplished by selecting areas whose species lists are most different or complementary. Areas with different environments tend to also have the most different species lists for a variety of taxa. As a result, a set of areas with complementary sets of species for one higher taxon (e.g., mammals) often will also do a good job representing most species of other higher taxa (e.g., trees, butterflies). Species with large home ranges, such as large carnivores, or species with very local distributions may require individual attention. Additional data layers can be used for a more holistic conservation evaluation. These include indicators of stress or risk (e.g., human population growth, road density, rate of habitat fragmentation, distribution of pollutants) and the locations of habitat corridors between wildlands that allow for natural movement of wide-ranging animals and the migration of species in response to climate change.

General Limitations

Limitations must be recognized so that additional studies can be implemented to supplement GAP. The following are general project limitations; specific limitations for the data are described in the respective sections:

1. GAP data are derived from remote sensing and modeling to make general assessments about conservation status. Any decisions based on the data must be supported by ground-truthing and more detailed analyses.

2. GAP is not a substitute for threatened and endangered species listing and recovery efforts. A primary argument in favor of gap analysis is that it is proactive: it seeks to recognize and manage sites of high biodiversity value for the long-term maintenance of populations of native species and communities before they become critically rare. Thus, it should help to reduce the rate at which species require listing as threatened or endangered. Those species that are already greatly imperiled, however, still require individual efforts to assure their recovery.

3. GAP data products and assessments represent a snapshot in time generally representing the date of the satellite imagery. Updates are planned on a 5-10 year cycle, but users of the data must be aware of the static nature of the products.

4. GAP is not a substitute for a thorough national biological inventory. As a response to rapid habitat loss, gap analysis provides a quick assessment of the distribution of vegetation and associated species before they are lost, and provides focus and direction for local, regional, and national efforts to maintain biodiversity. The process of improving knowledge in systematics, taxonomy, and species distributions is lengthy and expensive. That process must be continued and expedited, however, in order to provide the detailed information needed for a comprehensive assessment of our nation's biodiversity. Vegetation and species distribution maps developed for GAP can be used to make such surveys more cost-effective by stratifying sampling areas according to expected variation in biological attributes.

The Study Area

A Brief Description of ***insert here state name*. *Provide a brief description of the geologic, climatic, and human history and processes that have resulted in the current landscape and biodiversity condition. Tourist guides produced by states frequently have a good brief description of the landscape and human history. This should be 1-3 pages. See NM-GAP and PA-GAP reports for example. A map of the state showing topography, major roads, rivers, and municipalities is also suggested--see WY-GAP for example.***

insert state text here

LAND COVER CLASSIFICATION AND MAPPING

Introduction

Mapping natural land cover requires a higher level of effort than the development of data for animal species, agency ownership, or land management, yet it is no more important for gap analysis than any other data layer. Generally, the mapping of land cover is done by adopting or developing a land cover classification system, delineating areas of relative homogeneity (basic cartographic "objects"), then labeling these areas using categories defined by the classification system. More detailed attributes of the individual areas are added as more information becomes available, and a process of validating both spatial pattern and labels is applied for editing and revising the map. This is done in an iterative fashion, with the results from one step causing re-evaluation of results from another step. Finally, an assessment of the overall accuracy of the data is conducted. The final assessment of accuracy will show where improvements should be made in the next update (Stoms,1994).

In its "coarse filter" approach to conservation biology (e.g., Jenkins 1985, Noss 1987), gap analysis relies on maps of dominant natural land cover types as the most fundamental spatial component of the analysis (Scott et al. 1993) for terrestrial environments. For the purposes of GAP, most of the land surface of interest (natural) can be characterized by its dominant vegetation.

Vegetation patterns are an integrated reflection of the physical and chemical factors that shape the environment of a given land area (Whittaker 1965). They also are determinants for overall biological diversity patterns (Franklin 1993, Levin 1981, Noss 1990), and they can be used as a currency for habitat types in conservation evaluations (Specht 1975, Austin 1991). As such, dominant vegetation types need to be recognized over their entire ranges of distribution (Bourgeron et al. 1994) for beta-scale analysis (*sensu* Whittaker 1960, 1977). These patterns cannot be acceptably mapped from any single source of remotely sensed imagery, therefore, ancillary data, previous maps, and field surveys are used. The central concept is that the physiognomic and floristic characteristics of vegetation (and, in the absence of vegetation, other physical structures) across the land surface can be used to define biologically meaningful biogeographic patterns. There may be considerable variation in the floristics of subcanopy vegetation layers (community association) that are not resolved when mapping at the level of dominant canopy vegetation types (alliance), and there is a need to address this part of the diversity of nature. As information accumulates from field studies on patterns of variation in understory layers, it can be attributed to the mapped units of alliances.

Land Cover Classification

Land cover classifications must rely on specified attributes, such as the structural features of plants, their floristic composition, or environmental conditions, to consistently differentiate categories (Küchler and Zonneveld 1988). The criteria for a land cover classification system for GAP are:

- an ability to distinguish areas of different actual dominant vegetation;
- a utility for modeling animal species habitats;
- a suitability for use within and among biogeographic regions;
- an applicability to Landsat Thematic Mapper (TM) imagery for both rendering a base map and from which to extract basic patterns (GAP relies on a wide array of information sources, TM offers a convenient meso-scale base map in addition to being one source of actual land cover information);
- a framework that can interface with classification systems used by other organizations and nations to the greatest extent possible; and
- a capability to fit, both categorically and spatially, with classifications of other themes such as agricultural and built environments.

For GAP, the system that fits best is referred to as the National Vegetation Classification System (NVCS) (FGDC 1997). The origin of this system was referred to as the UNESCO/TNC system (Lins and Kleckner 1996) because it is based on the structural characteristics of vegetation derived by Mueller-Dombois and Ellenberg (1974), adopted by the United Nations Educational, Scientific, and Cultural Organization (UNESCO 1973) and later modified for application to the United States by Driscoll et al. (1983, 1984). The Nature Conservancy and the Natural Heritage Network (Grossman et al. 1994) have been improving upon this system in recent years with partial funding supplied by GAP. The basic assumptions and definitions for this system have been described by Jennings (1993).

Insert here and within the boiler plate variations in your project's approach, particularly the differences in classification scheme, i.e., note if your project developed before the above scheme existed. ***Provide the full descriptive and hierarchical scheme with crosswalks as necessary as an appendix.***

Methods

insert any introductory state text here

Mapping Standards and Data Sources: *Describe here the standards used to create the map applicable at the time of your project, e.g., MMU, spatial and thematic accuracy, data source--e.g.,* ***describe form of MRLC and any other imagery acquired including near-remote sensing systems such as videography. Provide a map showing the outlines of the actual imagery used--see the WY-GAP report for example.***

Land Cover Map Development: ***Insert state text here. Provide a detailed but concise description of steps used in creating the map that would allow a theoretical duplication of the effort. This requires you to be scientifically accurate but understandable by non-remote sensing specialists. Terms used that are not already in the Glossary should be inserted there.***

Special Feature Mapping: *insert state text here (optional-- describe process for mapping of any specialized features that was conducted as a separate effort such as a riparian mapping project).*

Results

Insert state text here. Provide a map in maps section

Table X. The land cover types mapped, their area mapped in the state in square kilometers (multiply by 100 for hectares or 270 for acres), and the percentage of the state's total area represented by the mapped type.

Mapped Type	NameTotal Area mapped	% of state area

Accuracy Assessment

Introduction: GAP land cover maps are primarily compiled to answer the fundamental question in gap analysis: what is the current distribution and management status of the nation's major natural land cover types and wildlife habitats? Besides giving a measure of overall reliability of the land cover map for Gap Analysis, the assessment also identifies which general classes or which regions of the map do not meet the accuracy objectives for the Gap Analysis Program. Thus the assessment identifies where additional effort will be required when the map is updated. We report the results of the accuracy assessment, believing that the map is the best map currently available for the project area.

The purpose of accuracy assessment is to allow a potential user to determine the map's "fitness for use" for their application. It is impossible for the original cartographer to anticipate all future applications of a land cover map, so the assessment should provide enough information for the user to evaluate fitness for their unique purpose. This can be described as the degree to which the data quality characteristics collectively suit an intended application. The information reported includes details on the database's spatial, thematic, and temporal characteristics and their accuracy.

Assessment data are valuable for purposes beyond their immediate application to estimating accuracy of a land cover map. The reference data is therefore made available to other agencies and organizations for use in their own land cover characterization and map accuracy assessments (see Data Availability for access information). The data set will also serve as an important training data source for later updates.

Even though we have reached an endpoint in the mapping process where products are made available to others, the gap analysis process should be considered dynamic. We envision that maps will be refined and updated on a regular schedule. The assessment data will be used to refine GAP maps iteratively by identifying where the land cover map is inaccurate and where more effort is required to bring the maps up to accuracy standards. In addition, the field sampling may identify new classes that were not identified at all during the initial mapping process.

<u>Methods:</u>
Insert state text here .Provide a thorough description of the methods used to collect truth data, quality assess it, and analyze the map accuracy. Provide maps and tables to document the truth data collection. See also the reporting requirements in the Handbook chapter.

<u>Results:</u>
Insert state text here. Some tables are provided below, interpretive maps are also suggested.

Limitations and Discussion

insert state text here

PREDICTED ANIMAL SPECIES DISTRIBUTIONS AND SPECIES RICHNESS

Introduction

All species range maps are predictions about the occurrence of those species within a particular area (Csuti 1994). Traditionally, the predicted occurrences of most species begin with samples from collections made at individual point locations. Most species range maps are small-scale (e.g., >1:10,000,000) and derived primarily from point data to construct field guides which are suitable, at best, for approximating distribution at the regional level or counties for example. The purpose of the GAP vertebrate species maps is to provide more precise information about the current predicted distribution of individual native species according to actual habitat characteristics within their general ranges and to allow calculation of predicted area of distributions and associations to specific habitat characteristics.

GAP maps are produced at a nominal scale of 1:100,000 or better and are intended for applications at the landscape or "gamma" scale (heterogeneous areas generally covering 1,000 to 1,000,000 hectares and made up of more than one kind of natural community). Applications of these data to site- or stand-level analyses (site--a microhabitat, generally 10 to 100 square meters; stand--a single habitat type, generally 0.1 to 1,000 ha; Whittaker 1977, see also Stoms and Estes 1993) will likely reveal the limitations of this process to incorporate differences in habitat quality (e.g., understory condition) or necessary microhabitat features such as standing dead trees.

Gap analysis uses the predicted distributions of animal species to evaluate their conservation status relative to existing land management (Scott et al. 1993). However, the maps of species distributions may be used to answer a wide variety of management, planning, and research questions relating to individual species or groups of species. In addition to the maps, great utility may be found in the consolidated specimen collection records and literature that are assembled into databases used to produce the maps. Perhaps most importantly, as a first effort in developing such detailed distributions, they should be viewed as testable hypotheses to be confirmed or refuted in the field. We encourage biologists and naturalists to conduct such tests and report their findings in the appropriate literature and to the Gap Analysis Program such that new data may improve future iterations.

Previous to this effort there were no maps available, digital or otherwise, showing the likely present-day distribution of species by habitat type across their ranges. Because of this, ordinary species (i.e., those not threatened with extinction or not managed as game animals) are generally not given sufficient consideration in land-use decisions in the context of large geographic regions or in relation to their actual habitats. Their decline, because of incremental habitat loss can, and does, result in one threatened or endangered species "surprise" after another. Frequently, the records that do exist for an ordinary species are truncated by state boundaries. Simply creating a consistent spatial framework for storing, retrieving, manipulating, analyzing, and updating the totality of our knowledge about the status of each animal species is one of the most necessary and basic elements for preventing further erosion of biological resources.

insert state text here

Methods

insert state text here, general statements about approach used.

Mapping Standards and Data Sources:
Describe here the standards used to create the map applicable at the time of your project, e.g., MMU, spatial and thematic accuracy, data source--e.g., ***describe types of data used in modeling.***

Mapping Range Extent:
describe the process of mapping general geographic ranges, e.g., EMAP hexagons.

Wildlife Habitat Relationships:
describe the process of developing WHRMs

Distribution Modeling:
Insert state text here. Describe the general process and if there are significant differences among taxa worth noting, use optional section below. Provide an appendix (suggest matrix style) of experts that helped and their affiliation, taxonomic group they worked on, and task, e.g., model development, range extent attributing, map review.

Table X. GIS coverages used in the animal species modeling process. Refer to the metadata accompanying the digital data for more complete descriptions. *please provide digital maps for each, printed copies in the hardcopy version are optional.*

Coverage NameSource of AcquisitionDescription*insert file name, e.g., epariverWhere did you get itprovide brief description such as "EPA River Reach 3--perennial rivers"*

Optional--provide specific methods descriptions including: range extent delineation method, habitat association development, and particular modeling approaches.

Mammals:

insert state text here

Birds:

insert state text here

Reptiles:

insert state text here

Amphibians:

insert state text here

Add here other taxa if mapped

Results

Insert state text here, any general statements about particular success or difficulties in the mapping, e.g., reptiles proved relatively simple to map whereas amphibians were not because of lack of NWI, etc. Below note the number of spp mapped per group and include example maps.

Mammals:

insert state text here

Birds:

insert state text here

Reptiles:

insert state text here

Amphibians:

insert state text here

Add here other taxa if mapped

Species Richness

GAP has often been associated with the mapping of species-rich areas or "hotspots." Richness maps identify where the same numbers of elements co-occur in the same geographic locations. (In the case of our data, where numbers of animal species are mapped for the same grid cells.) These are color coded or shaded in intensity from the highest numbers of co-occurrence (richness), to the lowest. While we continue to perform this useful pattern analysis, it is only one of many that may be conducted using the data. Richest areas may or may not indicate best conservation opportunities. They may occur in already protected areas or may represent mostly already protected species or those not at risk. Still, they are often a useful starting point to examine conservation opportunities in combination with other analyses described in this report's Introduction and in the Analysis section. We also feel they may be useful for other rewarding applications such as identifying places of interest for wildlife observation and study.

Richness is depicted for all mapped taxa (figure X), and by taxonomic groups (figures X-X). *Describe here more detail on map production, color coding, and then results and interpretation. See OR-GAP for example.*

Insert figures of species richness

Accuracy Assessment

Assessing the accuracy of the predicted vertebrate distributions is subject to many of the same problems as assessing land cover maps, as well as a host of more serious challenges related to both the behavioral aspects of species and the logistics of detecting them. These are described further in the Background section of the GAP Handbook on the national GAP home page. It is, however, necessary to provide some measure of confidence in the results of the gap analysis for species collectively, if not individually or by taxonomic group (comparison to stewardship and management status), and to allow users to judge the suitability of the distribution maps for their own uses. We, therefore, feel it is important to provide users with a statement about the accuracy of GAP-predicted vertebrate distributions within the limitations of available resources and practicalities of such an endeavor. We acknowledge that distribution maps are never finished products but are continually updated as new information is gathered. This reflects not only an improvement over the modeling process, but also the opportunity to map true changes in species distributions over time. However, we feel that assessing the accuracy of the current maps provides useful information about their reliability to potential users.

Our goal was to produce maps that predict distribution of terrestrial vertebrates and from that, total species richness and species content with an accuracy of 80% or higher. Failure to achieve this accuracy indicates the need to refine the data sets and models used for predicting distribution. There is a conscious effort in the GAP process, however, to err on the side of commission. In other words, to attribute species as possibly present when they are not. There are two primary reasons for doing so: first, few species have systematic, unbiased known ranges and we believe science is best served by identifying a greater potential for sampling and investigation than a conservative approach that may miss such opportunities; second, in conducting the analysis of conservation representation (see the Analysis section), we believe it most appropriate to identify a species that may need additional conservation attention that is then refuted by further investigation rather than identifying a species as sufficiently protected that is discovered not to be by its subsequent loss.

The methods for validating and assessing the accuracy of the vertebrate distribution maps are presented below along with the results.

<u>Methods:</u>
Insert state text here

<u>Results:</u>
Insert state text here. insert tables per Handbook chapter.

Limitations and Discussion

Insert state text here

LAND STEWARDSHIP

Introduction

To fulfill the analytical mission of GAP, it is necessary to compare the mapped distribution of elements of biodiversity with their representation in different categories of land ownership and management. As will be explained in the Analysis section, these comparisons do not measure viability, but are a start to assessing the likelihood of future threat to a biotic element through habitat conversion--the primary cause of biodiversity decline. We use the term "stewardship" in place of "ownership" in recognition that legal ownership does not necessarily equate to the entity charged with management of the resource, and that the mix of ownership and managing entities is a complex and rapidly changing condition not suitably mapped by GAP. 0 At the same time, it is necessary to distinguish between stewardship and management status in that a single category of land stewardship such as a national forest may contain several degrees of management for biodiversity.

The purpose of comparing biotic distribution with stewardship is to provide a method by which land stewards can assess their relative amount of responsibility for the management of a species or plant community and identify other stewards sharing that responsibility. This information can reveal opportunities for cooperative management of that resource, which directly supports the primary mission of GAP to provide objective, scientific information to decision makers and managers to make informed decisions regarding biodiversity. It also is not unlikely that a steward that has previously borne the major responsibility for managing a species may, through such analyses, identify a more equitable distribution of that responsibility. We emphasize, however, that GAP only identifies private land as a homogeneous category and does not differentiate individual tracts or owners, unless the information was provided voluntarily to recognize a long-term commitment to biodiversity maintenance.

After comparison to stewardship, it is also necessary to compare biotic occurrence to categories of management status. The purpose of this comparison is to identify the need for change in management status for the distribution of individual elements or areas containing high degrees of diversity. Such changes can be accomplished in many ways that do not affect the stewardship status. While it will eventually be desirable to identify specific management practices for each tract, and whether they are beneficial or harmful to each element, GAP currently uses a scale of 1 to 4 to denote relative degree of maintenance of biodiversity for each tract. A status of "1" denotes the highest, most permanent level of maintenance, and "4" represents the lowest level of biodiversity management, or unknown status. This is a highly subjective area, and we recognize a variety of limitations in our approach, although we maintain certain principles in assigning the status level. Our first principle is that land ownership is not the primary determinant in assigning status. The second principle is that while data are imperfect, and all land is subject to changes in ownership and management, we can use the intent of a land steward as evidenced by legal and institutional factors to assign status. In other words, if a land steward institutes a program backed by legal and institutional arrangements that are intended for permanent biodiversity maintenance, we use that as the guide for assigning status.

The characteristics used to determine status are as follows:

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

The four status categories can generally be defined as follows (after Scott et al. 1993, Edwards et al. 1995, Crist et al. 1995):

Status 1: An area having permanent protection from conversion of natural land cover and a mandated management plan in operation to maintain a natural state within which disturbance events (of natural type, frequency, and intensity) are allowed to proceed without interference or are mimicked through management.

Status 2: An area having permanent protection from conversion of natural land cover and a mandated management plan in operation to maintain a primarily natural state, but which may receive use or management practices that degrade the quality of existing natural communities.

Status 3: An area having permanent protection from conversion of natural land cover for the majority of the area, but subject to extractive uses of either a broad, low-intensity type or localized intense type. It also confers protection to federally listed endangered and threatened species throughout the area.

Status 4: Lack of irrevocable easement or mandate to prevent conversion of natural habitat types to anthropogenic habitat types. Allows for intensive use throughout the tract. Also includes those tracts for which the existence of such restrictions or sufficient information to establish a higher status is unknown.

Mapping Standards

Describe here the standards used to create the maps ***and criteria for including species*** *applicable at the time of your project.*

Methods

Note, there is no accuracy assessment section for stewardship--please describe your quality control process within this section.

<u>Stewardship Mapping</u>:
Insert state text here. Indicate sources and methods used to map public land ownership and voluntarily provided private lands (unless obtained from a public domain source).

Be sure to document land types identified but not included and reason, e.g., too small, did not qualify for >4 ranking, steward could not/would not provide boundaries.

Management Status Categorization:

insert state text here *(if you used a flow chart or key of some kind, provide description in appendix).*

Results

Insert state text here. Provide maps for both stewardship and management status.

The following tables present summary statistics of area representation of stewardship and management categories in the state. We begin by comparing representation of various stewardship categories in management status categories. Table 1 provides information on the proportional make-up of management status categories by stewardship and vice-versa, so that land stewards can see to what degree their lands generally contribute to biodiversity maintenance.

Table X. *Insert table from Stewardship chapter. Note-- the categories provided are only an example. If the table will not fit on one page, provide a generalized summary table here that includes only categories that represent a fairly large area of the state or a large portion of any one status category and lump all those not shown categorically into a column that says "other," to indicate that the list is incomplete. Provide the complete table as an appendix..*

Table X. ***Insert status 1 &2 areas documentation table from Stewardship Chapter of the Handbook.***

Limitations and Discussion

This map is a compilation of ownership maps provided by a variety of sources that are individually responsible for their accuracy. It was created solely for the purpose of conducting the analyses described in this report and is not suitable for locating boundaries on the ground or determining precise area measurements of individual tracts.
insert additional state text here

ANALYSIS BASED ON STEWARDSHIP AND MANAGEMENT STATUS

Introduction

This chapter describes the method and results of the gap analysis as used by the Gap Analysis Program. As described in the general introduction to this report, the primary objective of GAP is to provide information on the distribution and status of several elements of biological diversity. Although GAP "seeks to identify habitat types and species not adequately represented in the current network of biodiversity management areas" (GAP Handbook, Preface, Version 1, p. I), it is unrealistic to create a standard definition of "adequate representation" for either land cover types or individual species (Noss et al. 1995). A practical solution to this problem is to report both percentages and absolute area of each element in biodiversity management areas and allow the user to determine which types are adequately represented in natural areas. There are many other factors that should be considered in such determinations such as:

- historic loss or gain in distribution,
- nature of the spatial distribution,
- immediate versus long term risk, and
- degree of local adaptation among populations of the biotic elements that are worthy of individual conservation consideration.

Such analyses are beyond the scope of this project, but we encourage their application coupled with field confirmation of the mapped distributions.

Currently, land cover types and terrestrial vertebrates are the primary focus of GAP's mapping efforts, however, other components of biodiversity, such as aquatic organisms or selected groups of invertebrates may be incorporated into GAP distributional data sets. Where appropriate, GAP data may also be analyzed to identify the location of a set of areas in which most or all land cover types or species are predicted to be represented. The use of "complementarity" analysis, that is, an approach that additively identifies a selection of locations that may represent biodiversity rather than "hot spots of species richness" may prove most effective for guiding biodiversity maintenance efforts. Several quantitative techniques have been developed recently that facilitate this process (see Pressey et al. 1993, Williams et al. 1996, Csuti et al. 1997, for details). These areas become candidates for field validation and may be incorporated into a system of areas managed for the long-term maintenance of biological diversity.

The network of Conservation Data Centers (CDCs) and Natural Heritage Programs (NHPs) established cooperatively by The Nature Conservancy and various state agencies maintain detailed databases on the locations of rare elements of biodiversity. GAP cooperatively uses these data to develop predicted distributions of potentially suitable habitat for these elements, which may be valuable for identifying research needs and preliminary considerations for restoration or reintroduction. Conservation of such elements, however, is best accomplished through the fine-filter approach of the above organizations as described in the introduction. It is not the role of GAP to duplicate or disseminate Heritage Program or CDC Element Occurrence Records. Users interested in more specific information about the location, status, and ecology of populations of such species are directed to their state Heritage Program or CDC.

Methods

The gap analysis is accomplished by first producing: maps of land cover (see Figure X), predicted distributions for selected animal species (see Figure X), and land stewardship and management status (see Figure X). Intersecting the land stewardship and management map with the distribution of the elements results in tables that summarize the area and percent of total mapped distribution of each element in different land stewardship and management categories. ***insert state text here to provide more technical explanation of methods.***

Results

The data are provided in a format that allows users to carry out inquires about the representation of each element in different land stewardship and management categories as appropriate to their own management objectives. This forms the basis of Gap's mission to provide land owners and managers with the information necessary to conduct informed policy development, planning, and management for biodiversity maintenance.

As a coarse indicator of the status of the elements, we provide a breakdown along five levels of representation (0-<1%, 1-<10%, 10-<20%, and 20-<50%; >=50%). The <1% level indicates those elements with essentially none of their distribution in a protected status while levels of 10%, 20%, and 50% have been recommended in the literature as necessary amounts of conservation (Noss and Cooperrider 1994; Noss 1991; Odum and Odum 1972; Specht et al. 1974; Ride 1975; Miller 1994).

<u>Land Cover Analysis</u>

A summary of the analysis according to the thresholds described above is shown in Table X below. The table found in Appendix X provides the area in square kilometers (multiply by 100 for hectares, 247 for acres) of each types' mapped distribution by management status and land steward, and the percent of the types' total distribution in each category. For example, a typical entry may indicate that a particular woodland type has 35.23 km2 (3523 ha) in state lands that are ranked Status 2, which represents 9.3% of that type's total distribution. *Projects--please replace this example with an actual example from your table. The table format is found in the Analysis chapter. If it is small enough, you may include it here rather than as an appendix.*

Table X. see example in analysis chapter of handbook.

As explained above, we provide results according to thresholds of representation advocated in the literature to conserve biodiversity. The values in the table will allow users to set any desirable threshold to identify elements requiring more protection according to their criteria. The following summaries highlight potential gaps and conservation needs; Appendix X provides figures of stewardship representation for all land cover types by land steward and management status.

The narratives should summarize the stewardship condition (representation among managing entities and GAP status) of select types for each group below to explain the table results in a more informative way. In particular, it is useful to compare the gap analysis results to the current notions within the state about what needs more protection. Consider if this analysis had been carried out several years ago, would it have provided forewarning of current problems? WY and OR-GAP provide a good qualitative model for this section.

Land Cover with 0-<1% representation in GAP status 1 and 2:
insert state text here

Land Cover with 1-<10% representation in GAP status 1 and 2:
insert state text here

Land Cover with l0%-<20% representation in GAP status 1 and 2:
insert state text here

Land Cover with 20% -<50% representation in GAP status 1 and 2:
insert state text here

Land Cover with at least 50% representation in GAP status 1 and 2:
insert state text here

<u>Limitations and Discussion for Land Cover Analysis</u>

Assessing the conservation status of natural land cover is limited by several confounding factors:

GAP has typically found the accuracy of the mapped distributions of natural communities at the floristic (e.g., alliance) level to be substantially lower and more variable than that of animal distributions;

any aggregation of biotic units (e.g., above species) is a surrogate for species or lower levels of biotic organization and will underrepresent conservation need (Pressey and Logan 1995); and

for the most part we cannot distinguish the degree of natural condition or value of the mapped units due to management manipulation, exotic invasion, or spatial configuration. Considering an aggregation of species such as we have mapped to be sufficiently represented in existing conservation areas cannot be determined solely by the percentage of the community represented because the aggregation has unmapped variation in species composition that we could not measure. Until individual plant species distributions can be mapped, it is not possible to assure that the full range of vegetation biodiversity is represented, and surrogates must be used.

Insert state text here

<u>Predicted Animal Species Distributions Analysis</u>

A summary table is not provided due to the large number of species analyzed, but some generalizations and examples of species results by the various thresholds are provided below. The complete Animal Species Distributions Analysis Table found in Appendix _ provides the area in square kilometers (multiply by 100 for hectares, 247 for acres) of the species' mapped distribution by management status and land steward, and the percent of the species' total distribution in each category. For example, a typical entry may indicate that a particular species has 35.23 km2 (3523 ha) of potential habitat in state lands that are ranked Status 2, which represents 9.3% of that species' total distribution. *Projects--please replace this example with an actual example from your table. The table format is found in the Analysis handbook chapter. If it is small enough, you may include it here rather than as an appendix.*

This section should summarize the stewardship condition of each taxonomic group with additional narrative on species of interest.

Species with 0-<1% of predicted distribution in status 1 or 2:
Insert state text here--this is a list followed by explanatory narrative of a few species

Species with 1-<10% of predicted distribution in status 1 or 2:
Insert state text here

Species with l0%-<20% of predicted distribution in status 1 or 2:
Insert state text here

Species with 20% -<50% of predicted distribution in status 1 or 2:
Insert state text here

Species with at least 50% representation in GAP status 1 and 2:
insert state text here

Analysis of Important Statewide Species Assemblages: ***insert state text here***

Analysis of State Endemics: ***insert state text here (if applicable)***

Special Features Analysis: ***insert state text here (optional)***

Ecoregional Analysis: ***insert state text here (optional)***

Limitations and Discussion
When applying the results of our analyses, it is critical that the following limitations are considered: 1) the limitations described for each of the component parts (land cover mapping, animal species mapping, stewardship mapping) of the analyses, 2) the spatial and thematic map accuracy of the components, and 3) the suitability of the results for the intended application (see Appropriate and Inappropriate Use below). ***insert additional state text here***

CONCLUSIONS AND MANAGEMENT IMPLICATIONS

Insert state text here. See previous project reports for examples; this is not to be a summary of the report. Generally, this should have statements about the condition of biodiversity in the state, the need for additional data and analyses, and any obvious management implications that don't overstep the information mission (versus policy) that can help foster the additional work and cooperative implementation activities. This is a chance to tell your cooperators "where we go from here."

PRODUCT USE AND AVAILABILITY

How to Obtain the Products

It is the goal of the Gap Analysis Program and the USGS Biological Resources Division (BRD) to make the data and associated information as widely available as possible. Use of the data requires specialized software called geographic information systems (GIS) and substantial computing power. Additional information on how to use the data or obtain GIS services is provided below and on the GAP home page (URL below). While a CD-ROM of the data will be the most convenient way to obtain the data, it may also be downloaded via the Internet from the national GAP home page at:

http://www.gap.uidaho.edu/

The home page will also provide, over the long term, the status of our state's project, future updates, data availability, and contacts. Within a few months of this project's completion, CD-ROMs of the final report and data should be available at a nominal cost--the above home page will provide ordering information. To find information on this state GAP project's status and data, follow the links to "Current Projects" and then to the particular state of interest.

Insert state text here describing availability through your state GIS library, directly from you, etc. including hardcopy volumes and any associated products such as atlases.

Minimum GIS Required for Data Use: ***insert state text here. Please indicate the minimum type of setup to use the data and a lab in the state where someone can get GIS services conducted e.g., special cut-outs, analyses, packaged with other data, etc.***

Disclaimer

Following is the official Biological Resources Division (BRD) disclaimer as of 29 January 1996, followed by additional disclaimers from GAP. Prior to using the data, you should consult the GAP home page (see How to Obtain the Data, above) for the current disclaimer.

Although these data have been processed successfully on a computer system at the BRD, no warranty expressed or implied is made regarding the accuracy or utility of the data on any other system or for general or scientific purposes, nor shall the act of distribution constitute any such warranty. This disclaimer applies both to individual use of the data and aggregate use with other data. It is strongly recommended that these data are directly acquired from a BRD server [see above for approved data providers] and not indirectly through other sources which may have changed the data in some way. It is also strongly recommended that careful attention be paid to the content of the metadata file associated with these data. The Biological Resources Division shall not be held liable for improper or incorrect use of the data described and/or contained herein.

These data were compiled with regard to the following standards. Please be aware of the limitations of the data. These data are meant to be used at a scale of 1:100,000 or smaller (such as 1:250,000 or 1:500,000) for the purpose of assessing the conservation status of animals and vegetation types over large geographic regions. The data may or may not have been assessed for statistical accuracy. Data evaluation and improvement may be ongoing. The Biological Resources Division makes no claim as to the data's suitability for other purposes. This is writable data which may have been altered from the original product if not obtained from a designated data distributor identified above.

Metadata

Proper documentation of information sources and processes used to assemble GAP data layers is central to the successful application of GAP data.

Metadata is a description of the content, quality, lineage, contact, condition, and other characteristics of data. It is a valuable tool that preserves the usefulness of data over time by detailing methods for data collection and data set creation. It greatly minimizes duplication of effort in the collection of expensive digital data and fosters sharing of digital data resources. Metadata supports local data asset management such as local inventory and data catalogs, and external user communities such as Clearinghouses and websites. It provides adequate guidance for end-use application of data such as detailed lineage and context. Metadata makes it possible for data users to search, retrieve, and evaluate data set information by providing standardized descriptions of geospatial and biological data.

The Federal Geographic Data Committee approved the Content Standard for Digital Geospatial Metadata (FGDC-STD-001-1998) in June 1998 and NBII (<http://www.nbii.gov>) developed the Biological Data Profile (approved in 1999) that adds fields for biological information such as taxonomy, analytical tools, and methodology to the FGDC standard core set of elements. <http://www.nbii.gov/datainfo/metadata/standards/> Executive Order 12906 requires that any spatial data sets generated with federal dollars will have FGDC-compliant metadata.

Each spatial data layer submitted must be accompanied by its metadata (*.xml or .sgml file) in the same directory. You must also include an additional directory (called "meta_master') which will include each metadata file in four forms (*.txt, *.xml, *.html, and *.sgml).

There are many tools available for metadata creation. For some examples, see <http://www.nbii.gov/datainfo/metadata/tools/> Please note that some tools are free, and some are not. The redundancy in output format is to provide one file for error checking (*.txt), one for presentation on the Internet (*.html), and two for indexing elements for the spatial data clearinghouse (*xml, *.sgml). Remember, metadata describes the development of the spatial data set being documented. If there are companion files to the GIS data, use metadata to reference (reports, spreadsheet, another GIS layer).

USGS (NBII and FGDC) personnel conduct metadata training to meet FGDC standards and to include biological data. The metadata workshop provides an introduction to the metadata standard with hands-on practice producing documentation for a sample data set using appropriate software: Intergraph's "Spatial Metadata Management System (SMMS)" and USDA Forest Service North Central Research Station's "Metavist" are commonly used. The focus of the workshop is an understanding of the metadata standard, but other topics will include the metadata clearinghouse, metadata development tools, and strategies for metadata production. See <http://www.nbii.gov/datainfo/metadata/training/> for more information and access to the training calendar.

Appropriate and Inappropriate Use of These Data

You may adapt this section freely to local conditions and the resolution and confidence in your data.

All information is created with a specific end use or uses in mind. This is especially true for GIS data, which is expensive to produce and must be directed to meet the immediate program needs. For GAP, minimum standards were set (see A Handbook for Gap Analysis, Scott et al. 1993) to meet program objectives. These standards include: scale or resolution (1:100,000 or 100 hectare minimum mapping unit), accuracy (80% accurate at 95% confidence), and format (ARC/INFO coverage tiled to the 30' x 60' USGS quadrangle). For complete project standards, refer to appendix XX.

Recognizing, however, that GAP would be the first, and for many years likely the only, source of statewide biological GIS maps, the data were created with the expectation that they would be used for other applications. Therefore, we list below both appropriate and inappropriate uses. This list is in no way exhaustive but should serve as a guide to assess whether a proposed use can or cannot be supported by GAP data. For most uses, it is unlikely that GAP will provide the only data needed, and for uses with a regulatory outcome, field surveys should verify the result. In the end, it will be the responsibility of each data user to determine if GAP data can answer the question being asked, and if they are the best tool to answer that question.

<u>Scale:</u> First we must address the issue of appropriate scale to which these data may be applied. The data were produced with an intended application at the ecoregion level, that is, geographic areas from several hundred thousand to millions of hectares in size. The data provide a coarse-filter approach to analysis, meaning that not every occurrence of every plant community or animal species habitat is mapped, only larger, more generalized distributions. The data are also based on the USGS 1:100,000 scale of mapping in both detail and precision. When determining whether to apply GAP data to a particular use, there are two primary questions: do you want to use the data as a map for the particular geographic area, or do you wish to use the data to provide context for a particular area? The distinction can be made with the following example: You could use GAP land cover to determine the approximate amount of oak woodland occurring in a county, or you could map oak woodland with aerial photography to determine the exact amount. You then could use GAP data to determine the approximate percentage of all

oak woodland in the region or state that occurs in the county, and thus gain a sense of how important the county's distribution is to maintaining that plant community.

Appropriate Uses: The above example illustrates two appropriate uses of the data: as a coarse map for a large area such as a county, and to provide context for finer-level maps. Specific case-study examples are provided in appendix XX, but following is a general list of applications:

- Statewide biodiversity planning
- Regional (Councils of Government) planning
- Regional habitat conservation planning
- County comprehensive planning
- Large-area resource management planning
- Coarse-filter evaluation of potential impacts or benefits of major projects or plan initiatives on biodiversity, such as utility or transportation corridors, wilderness proposals, regional open space and recreation proposals, etc.
- Determining relative amounts of management responsibility for specific biological resources among land stewards to facilitate cooperative management and planning.
- Basic research on regional distributions of plants and animals and to help target both specific species and geographic areas for needed research.
- Environmental impact assessment for large projects or military activities.
- Estimation of potential economic impacts from loss of biological resource-based activities.
- Education at all levels and for both students and citizens.

Inappropriate Uses: It is far easier to identify appropriate uses than inappropriate ones, however, there is a "fuzzy line" that is eventually crossed when the differences in resolution of the data, size of geographic area being analyzed, and precision of the answer required for the question are no longer compatible. Examples include:

- Using the data to map small areas (less than thousands of hectares), typically requiring mapping resolution at 1:24,000 scale and using aerial photographs or ground surveys.
- Combining GAP data with other data finer than 1:100,000 scale to produce new hybrid maps or answer queries.
- Generating specific areal measurements from the data finer than the nearest thousand hectares (minimum mapping unit size and accuracy affect this precision).
- Establishing exact boundaries for regulation or acquisition.
- Establishing definite occurrence or non-occurrence of any feature for an exact geographic area (for land cover, the percent accuracy will provide a measure of probability).
- Determining abundance, health, or condition of any feature.
- Establishing a measure of accuracy of any other data by comparison with GAP data.
- Altering the data in any way and redistributing them as a GAP data product.
- Using the data without acquiring and reviewing the metadata and this report.

LITERATURE CITED

Austin, M.P. 1991. Vegetation: Data collection and analysis. Pages 37-41 in C.R. Margules and M.P. Austin, editors, Nature conservation: Cost effective biological surveys and data analysis. Australia CSIRO, East Melbourne.

Bourgeron, P.S., H.C. Humphries, R.L. DeVelice, and M.E. Jensen. 1994. Ecological theory in relation to landscape and ecosystem characterization. Pages 58-72 in M.E. Jensen and P.S. Bourgeron, editors, Ecosystem management: Principles and applications, Volume II. Gen. Tech. Rep. PNW-GTR-318. USDA Forest Service, Pacific Northwest Research Station, Portland, Oregon. 376 pp.

Crist, P., B. Thompson, and J. Prior-Magee. 1995. A dichotomous key of land management categorization, unpublished. New Mexico Cooperative Fish and Wildlife Research Unit, Las Cruces, New Mexico.

Csuti, B., S. Polasky, P.H. Williams, R.L. Pressey, J.D. Camm, M. Kershaw, A.R. Kiester, B. Downs, R. Hamilton, M. Huso, and K. Sahr. 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation* 80:83-97

Csuti, B. 1994. Methods for developing terrestrial vertebrate distribution maps for Gap Analysis (version 1). In J.M. Scott and M.D. Jennings, editors. A handbook for Gap Analysis. Idaho Cooperative Fish and Wildlife Research Unit, University of Idaho, Moscow.

Davis, F.W., P.A. Stine, D.M. Stoms, M.I. Borchert, and A.D. Hollander. 1995. Gap Analysis of the actual vegetation of California - 1. The southwestern region. *Madroño* 42:40-78.

Diamond, J. 1986. The design of a nature reserve system for Indonesian New Guinea. Pages 485-503 in M.E. Soulé, editor. Conservation biology: The science of scarcity and diversity. Sinauer Associates, Sunderland, Massachusetts.

Driscoll, R.S., D.L. Merkel, D.L. Radloff, D.E. Snyder, and J.S. Hagihara. 1984. An ecological land classification framework for the United States. Miscellaneous Publication 1439. USDA Forest Service, Washington, DC.

Driscoll, R.S., D.L. Merkel, J.S. Hagihara, and D.L. Radloff. 1983. A component land classification for the United States: Status report. Technical Note 360. U.S. Department of the Interior, Bureau of Land Management, Denver, Colorado.

Edwards, T.C., Jr., C.H. Homer, S.D. Bassett, A. Falconer, R.D. Ramsey, and D.W. Wight. 1995. Utah Gap Analysis: An environmental information system. Technical Report 95-1, Utah Cooperative Fish and Wildlife Research Unit, Utah State University, Logan, Utah.

Federal Geographic Data Committee (FGDC). 1994. Content standards for digital geospatial metadata. 8 June 1994. Federal Geographic Data Committee, Washington, DC.

Federal Geographic Data Committee. 1995. Content standards for digital geospatial metadata workbook, FGDC, Washington, DC.

Federal Geographic Data Committee, Vegetation Subcommittee. 1997. FGDC Vegetation Classification and Information Standards--June 3, 1996 Draft. FGDC Secretariat, Reston, Virginia. 35 pp.

Forman, R.T.T., and M. Godron. 1986. Landscape ecology. John Wiley and Sons, New York.
Franklin, J.F. 1993. Preserving biodiversity: Species, ecosystems, or landscapes? *Ecological Applications* 3(2):202-205.
Gap Analysis Program. 2000. A handbook for conducting Gap Analysis. USGS Gap Analysis Program, Moscow, Idaho. http://www.gap.uidaho.edu/gap/handbook.
Grossman, D., K.L. Goodin, X. Li, C. Wisnewski, D. Faber-Langendoen, M. Anderson, L. Sneddon, D. Allard, M. Gallyoun, and A. Weakley. 1994. Standardized national vegetation classification system. Report by The Nature Conservancy and Environmental Systems Research Institute for the NBS/NPS Vegetation Mapping Program. National Biological Service, Denver, Colorado.
Jenkins, R.E. 1985. Information methods: Why the Heritage Programs work. *The Nature Conservancy News* 35(6):21-23.
Jennings, M.D. 1993. Natural terrestrial cover classification: Assumptions and definitions. Gap Analysis Technical Bulletin 2. Idaho Cooperative Fish and Wildlife Research Unit, University of Idaho, Moscow.
Kirkpatrick, J.B. 1983. An iterative method for establishing priorities for the selection of nature reserves: An example from Tasmania. *Biological Conservation* 25:127-134.
Küchler, A.W., and I.S. Zonneveld, editors. 1988. Vegetation mapping. Kluwer Academic Publishers, Dordrecht, The Netherlands. 635 pp.
Levin, S.A. 1981. The problem of pattern and scale in ecology. *Ecology* 73:1942-1968.
Lins, K.S., and R.L. Kleckner. 1996. Land cover mapping: An overview and history of the concepts. Pages 57-65 in J.M. Scott, T.H. Tear, and F. Davis, editors, Gap Analysis: A landscape approach to biodiversity planning. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland.
Margules, C.R., A.O. Nicholls, and R.L. Pressey. 1988. Selecting networks of reserves to maximize biological diversity. *Biological Conservation* 43:63-76
Miller, K.R. 1994. In J.A. McNeely and K.R. Miller, editors. National parks conservation and development. Smithsonian Institution, Washington, DC.
Mueller-Dombois, D., and H. Ellenberg. 1974. Aims and methods of vegetation ecology. John Wiley & Sons, New York. 547 pp.
Nicholls, A.O., and Margules. 1993. An upgraded reserve selection algorithm. *Biological Conservation* 64:165-169.
Noss, R.F. 1987. From plant communities to landscapes in conservation inventories: A look at The Nature Conservancy (USA). *Biological Conservation* 41:11-37.
Noss, R.F. 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4:355-364.
Noss, R.F. 1991. Report to the Fund for Animals in Washington, DC.
Noss, R.F., and A.Y. Cooperrider. 1994. Saving nature's legacy. Island Press, Washington, DC.
Noss, R.F., E.T. LaRoe III, and J.M. Scott. 1995. Endangered ecosystems of the United States: A preliminary assessment of loss and degradation. Biological Report 28, National Biological Service, Washington, DC.
Odum, E.D., and H.T. Odum. 1972. Proceedings of the North American Wildlife and Natural Resources Conference 39:178.

Pressey, R.L., and A.O. Nicholls. 1989. Application of a numerical algorithm to the selection of reserves in semi-arid New South Wales. *Biological Conservation* 50:263-278.

Pressey, R.L., C.J. Humphries, C.R. Margules, R.I. Vane-Wright, and P.H. Williams. 1993. Beyond opportunism: Key principles for systematic reserve selection. Trends in Ecology and Evolution 8:124-128.

Pressey, R.L., and V.S. Logan. 1995. Reserve coverage and requirements in relation to partitioning and generalization of land classes: analyses for Western New South Wales. *Conservation Biology* 9(6):1506-1517.

Ride, W.L.D. 1975. In F. Fenner, editor. A national system of ecological reserves in Australia. 64 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 G. Wright. 1993. Gap analysis: A geographic approach to protection of biological diversity. *Wildlife Monographs* 123.

Specht, R.L. 1975. The report and its recommendations. Pages 11-16 in: F. Fenner, editor, A national system of ecological reserves in Australia. Australian Academy of Sciences Report No. 19. Canberra, Australia.

Specht, R.L., E.M. Roe, and V.H. Boughlon. 1974. Australian Journal of Botany Supplement Series. Supplement No. 7.

Stoms, D., and J. Estes. 1993. A remote sensing research agenda for mapping and monitoring biodiversity. *International Journal of Remote Sensing* 14:1839-1860.

Stoms, D.M. 1994. Actual vegetation layer. In J.M. Scott and M.D. Jennings, editors. A handbook for Gap Analysis. Idaho Cooperative Fish and Wildlife Research Unit, University of Idaho, Moscow.

UNESCO. 1973. International classification and mapping of vegetation. Paris.

Williams, P.H., D. Gibbons, C. Margules, A. Rebelo, C. Humphries, and R. Pressey. 1996. A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds. *Conservation Biology* 10:155-174.

Whittaker, R.H. 1960. Vegetation of the Siskiyou mountains, Oregon and California. *Ecological Monographs* 30(3):279-338.

Whittaker, R.H. 1965. Dominance and diversity in land plant communities. *Science* 147: 250-259.

Whittaker, R.H. 1977. Species diversity in land communities. *Evolutionary Biology* 10:1-67.

GLOSSARY

insert current file from GAP Handbook

GLOSSARY OF ACRONYMS

insert current file from GAP Handbook and add any additions from your text

MAPS

APPENDICES

Appendix _ List of Example GAP Applications

Businesses and Non-government Organizations:
The following are some examples of applications of GAP data by the private sector:

- The Wyoming Natural Heritage Program (a private non-government organization) transformed the endangered and sensitive species database into a spatially referenced digital geographic information system using the GAP digital base map and other GAP spatial data.
- Hughes Corp. is experimenting with the Utah and Nevada GAP digital base maps, simulating images to aid the development of new space-based remote sensing devices.
- The Nature Conservancy used the Wyoming GAP data to develop a map of ecoregions of Wyoming.
- Weyerhaeuser Corp. is using the Arkansas GAP data in managing their lands in Arkansas.
- IBM Corp. is funding a project at the University of California-Santa Barbara that, in part, uses GAP data in the development of visualization software.
- NM-GAP vegetation data is being used for an environmental assessment of a proposed spaceport, a state/private venture.

County and City Planning:
Some other examples of the use of GAP by local governments are:

- CA-GAP biological data were combined with the Southern California Association of Governments (SCAG) land ownership data to show which ownerships and jurisdictions were needed for joint conservation planning and management of a particular natural community or species, maximizing efficiency and minimizing the potential for yet another conservation crisis.
- In California, county and city planners of several jurisdictions, wildlife agencies, developers of the 4S Ranch property, and the state Natural Communities Conservation Planning program used the GAP regional data, as well as more detailed information, to conserve 1,640 acres of habitat within a 2,900-acre planned development.
- Day-to-day county planning operations in Piute, Grande, and Washington counties, Utah.
- County planners in Piute County, Utah, used GAP data to optimize the siting of a proposed sawmill for aspen with respect to the distribution of aspen stands.
- Missoula County, Montana, used the GAP land cover map of the area as a base map for its comprehensive long-range plan.
- Snohomish County, Washington, used the GAP land cover map in meeting state requirements for a growth management plan.
- The City of Bainbridge Island, Washington, used GAP data to assist them in development of a watershed planning project.

State Uses:
The following are some examples of uses of GAP data by state agencies.

- The GAP database of species habitats was used by the Tennessee Wildlife Resources Agency (TWRA) to update its book "Species in Need of Management."
- Images of land cover derived from GAP TM data are used by TWRA for locating particular habitat types. Information on the locations of these habitat types is provided by TWRA to the public for a wide variety of public service functions, from education to cooperative resource management.
- Early GAP data developed by TWRA were used to help identify an extremely important area of the state with high biodiversity that was subsequently purchased by the state for conservation.
- Preliminary findings from GAP were used by TWRA to develop three resource management initiatives.
- The Tennessee GAP project, which is being carried out primarily by TWRA, is the foundation of a multi-agency, long-term biodiversity program for Tennessee.
- GAP data have been used by the Tennessee Forestry Stewardship Program to help develop a district program for nine conservation planning districts, outlining Best Management Practices (BMPs) for biological conservation on private lands.
- GAP data are being used extensively by TWRA in the preparation of project proposals to the North American Waterfowl Conservation Program. These proposals require that biodiversity issues are addressed in specific detail. The use of GAP data on occurrence of land cover types and terrestrial vertebrates has made this possible.
- The Wyoming Department of Fish and Game used GAP data to assist them in transforming the Wildlife Observation System database into a spatially referenced geographic information system.
- The Utah Division of Wildlife Resources and the Bear River Water Conservancy District used the Utah GAP land cover map in a resource management assessment for mitigating conflicts between a proposed groundwater withdrawal project and the maintenance of an elk calving area in the Uinta Mountains.
- The Utah Division of Wildlife Resources, the Rocky Mountain Elk Foundation, and Sheik Safari International used the Utah GAP land cover map to identify critical elk habitat. The environmental profile of these areas was then used to identify other similar areas for elk habitat enhancement.
- The Utah Division of Wildlife Resources used the Utah GAP land cover map for a rapid ecological assessment of the Echo Henefer Wildlife Management Area.
- The Washington Department of Fish and Wildlife used GAP data to develop a breeding bird atlas and an atlas of mammals of Washington State.
- The Washington Department of Fish and Wildlife uses GAP data to operate an integrated landscape management program.
- The Washington Department of Fish and Wildlife uses GAP data from Eastern Washington to assist with an innovative program that brings the forest products industry, state agency biologists, non-government organizations, and tribal biologists together in the field to jointly determine the appropriate management practices for any particular site of concern (Timber, Fish & Wildlife Program).
- The Idaho Department of Fish and Game used GAP data to evaluate the impact from expanded military training activities on public lands in Southern Idaho.

- The Idaho Department of Fish and Game uses GAP data for regional planning efforts on a regular basis.

<u>Statewide Planning</u>:
Biodiversity planning programs or projects are now under way in Arizona, California, Colorado, Maine, Missouri, Nevada, Oregon, and Tennessee. It is likely that similar efforts will develop in other states. These activities were the subject of the State Biodiversity Programs meeting discussed on page _ in this report. In some cases, these efforts grew out of the state GAP project, however, in most cases, the GAP data are being used to meet a previously defined need. In all cases, GAP data are central to their development and operations. The goals of each of these programs or projects are presented briefly below.

<u>Federal Agency Applications</u>:
Some examples of applications of GAP data by federal agencies follow:

- GAP data are being supplied to all military installations in the Great Basin ecoregion for integrated management of the natural resources. These installations constitute a very large amount of land area. Much of it is of high value for native species.
- The Ouachita National Forest used the Arkansas GAP data to help them develop an ecosystem management plan.
- The Wyoming GAP data were used by NASA to calibrate a model that predicts vegetation types based on climate and soil variables.
- The potential contributions to biodiversity conservation of four different options proposed for new wilderness designation in Idaho were quantified by the Idaho Cooperative Fish and Wildlife Research Unit in cooperation with the Park Studies Unit.
- The potential contributions to biodiversity conservation of four different options proposed for new national park designation in Idaho were quantified by the Idaho Cooperative Park Studies Unit.
- The U.S. Forest Service in Booneville, Arkansas, used the Arkansas GAP data land cover maps in a 3-dimensional presentation to provide the public with a visual representation of the region and to enhance the public's involvement with the National Forest planning process.
- The U.S. Fish and Wildlife Service regularly uses the GAP data for Southern California for habitat evaluation and management.
- The U.S. Forest Service, Bureau of Land Management, and National Park Service are using the GAP data for a wide variety of natural resource management operations in Utah. For example, the entire Utah GAP database is directly linked with existing National Park Service databases for use by National Parks.
- The Bureau of Land Management uses the Wyoming GAP data for managing the Buffalo Resource Area.
- The U.S. Forest Service used the Utah GAP data to help assist them in evaluating human-induced impacts to forested lands surrounding ski resorts in central Utah.
- The U.S. Fish and Wildlife Service in Delaware used GAP data to help identify potential habitat for the federally endangered Delmarva fox squirrel. These maps

Glossaries of Terms and Acronyms

Version: 2.0.0
Date: 16 February 2000
Peer-reviewed: No

Terms

aerial videography - video images of the land surface taken from an airplane

algorithm - a procedure to solve a problem or model a solution (In GAP typically refers to a GIS procedure used to model animal distributions.)

alliance level - a land unit made up of an "alliance" of natural communities that have the same dominant or co-dominant plant species or, in the absence of vegetation, by the dominant land cover typically described according to the Anderson land cover classification (see "Natural Community Alliance" in Grossman et al. 1995)

alpha diversity - a single within-habitat measure of species diversity regardless of internal pattern, generally over an area of 0.1 to 1,000 hectares (see Whittaker 1960, 1977) -

Anderson Level II - the second hierarchical level in the Anderson land cover classification system (see Anderson et al. 1976)

anthropogenic - caused by man

assemblages - a group of ecologically interrelated plant and animal species

band, spectral - a segment of the electromagnetic spectrum defined by a range of wavelengths (e.g. blue, green, red, near infrared, far infrared) that comprise the Landsat TM imagery

beta diversity - the change in species diversity among different natural communities of a landscape; an index of between-habitat diversity (see Whittaker 1960, 1977)

biodiversity - generally, the variety of life and its interrelated processes

biogeographic - relating to the geographical distribution of plants and animals

biological diversity - see biodiversity

cartographic - pertaining to the art or technique of making maps or charts

classify - to assign objects, features, or areas on an image to spectral classes based upon their appearance as opposed to ‘ classification’ referring to a scheme for describing the hierarchies of vegetation or animal species for an area

coarse filter - the general conservation activities that conserve the common elements of the landscape matrix, as opposed to the "fine filter" conservation activities that are aimed at special cases such as rare elements (see Jenkins 1985)

community - a group of interacting plants and animals

cover type - a non-technical higher-level floristic and structural description of vegetation cover

cross-walking - matching equivalent land cover categories between two or more classification systems

delineate - identifying the boundaries between more or less homogenous areas on remotely sensed images as visible from differences in tone and texture

delta diversity - the change in species diversity between landscapes along major climatic or physiographic gradients (see Whittaker 1977)

digitization - entering spatial data digitally into a Geographic Information System

ecoregion - a large region, usually spanning several million hectares, characterized by having similar biota, climate, and physiography (topography, hydrology, etc).

ecosystem - a biological community (ranging in scale from a single cave to millions of hectares), its physical environment, and the processes through which matter and energy are transferred among the components

edge-matching - the process of connecting polygons at the boundary between two independently created maps, either between TM scenes or between state GAP data sets

element - a plant community or animal species mapped by GAP. May also be referred to as "element of biodiversity".

error of commission - the occurrence of a species (or other map category) is erroneously predicted in an area where it is in fact absent

error of omission - when a model fails to predict the occurrence of a species that is actually present in an area

exact set coverage - a basic optimization problem to determine the best method for identifying general areas that, when selected sequentially, would have the greatest positive cumulative impact on attaining adequate representation of any or all biotic elements of interest

extinction - disappearance of a species throughout its entire range

extirpation - disappearance of a species from part of its range

fine filter - see "coarse filter"

floristic - pertaining to the plant species that make up the vegetation of a given area.

formation level - the level of land cover categorization between Group and Alliance describing the structural attributes of a land unit, for example, "Evergreen Coniferous Woodlands with Rounded Crowns" (see Jennings 1993b)

gamma diversity - the species diversity of a landscape, generally covering 1,000 to 1,000,000 hectares, made up of more than one kind of natural community (see Whittaker 1977)

gap analysis - a comparison of the distribution of elements of biodiversity with that of areas managed for their long-term viability to identify elements with inadequate representation

geographic information systems - computer hardware and software for storing, retrieving, manipulating, and analyzing spatial data

Global Positioning System (GPS) - an instrument that utilizes satellite signals to pinpoint its location on the earth's surface

greedy heuristic - an algorithm for exact set cover analysis (see Kiester et al., in press)

ground truthing - verifying maps by checking the actual occurrence of plant and animal species in the field at representative sample locations

habitat - the physical structure, vegetational composition, and physiognomy of an area, the characteristics of which determine its suitability for particular animal or plant species

hectare - a metric unit of area of 10,000 square meters and equal to 2.47 acres

hex/hexagon - typically refers to the EPA EMAP hexagonal grid of 635 square kilometer units

hyperclustering - a efficient, interactive method for accurately analyzing and classifying remotely-sensed data that reduces data size and computational requirements while retaining the integrity of the original data

lotic - flowing, e.g., water in a stream or river

metadata - information about data, e.g., their source, lineage, content, structure, and availability

minimum mapping unit - the smallest area that is depicted on a map

neotropics - the zoo-geographic region stretching southward from the tropic of Cancer and including southern Mexico, Central and South America, and the West Indies

phenology - the study of periodic biological phenomena, such as flowering, breeding, and migration, especially as related to climate

phenotype - the environmentally and genetically determined observable appearance of an organism, especially as considered with respect to all possible genetically influenced expressions of one specific character

physiognomic - based on physical features

physiographic province - a region having a pattern of relief features or land forms that differ significantly from that of adjacent regions

pixel - the smallest spatial unit in a raster data structure

polygon - an area enclosed by lines in a vector-based Geographic Information System data layer or a region of contiguous homogeneous pixels in a raster system

preprocessing - those operations that prepare data for subsequent analysis, usually by attempts to correct or compensate for systematic, radiometric, and geometric errors

pro-active - acting in anticipation of an event as opposed to reacting after the fact

range - the geographic limit of the species

range unit - a spatial, geographic unit to record and display species geographic range.

reach - a stream or river segment between inflowing tributaries

registration, spatial - matching different images to each other by finding points on the images that can be matched to known points on the ground

remote sensing - deriving information about the earth's surface from images acquired at a distance, usually relying on measurement of electromagnetic radiation reflected or emitted from the feature of interest

resolution - the ability of a remote sensing system to record and display fine detail in a distinguishable manner or: the smallest feature that can be distinguished or resolved on a map or image, such as a TM pixel

scale, map - the ratio of distance on a map to distance in the real word, expressed as a fraction; the smaller the denominator, the larger the scale, e.g. 1:24,000 is larger than 1:100,000

sensitivity analysis - the consideration of a number of factors involved in the mathematical modeling of an ecosystem and its components. These include feedback and control, and the stability and sensitivity of the system as a whole to changes in some part of the system. Predictions can be made from the analysis.

simulated annealing - an algorithm used for set coverage analysis (see Kiester et al., in press)

species richness - the number of species of a particular interest group found in a given area

spectral cluster - a group of adjacent pixels that are uniform with respect to their brightness values

supervised classification - the process of classifying TM pixels of unknown identity by using samples of known identity (i.e., pixels already assigned to informational classes by ground truthing or registration with known land cover) as training data

synoptic - constituting a brief statement or outline of a subject; presenting a summary

tessellation - the division of a map into areas of equal and uniform shape such as the EPA- EMAP hexagon

Thematic Mapper - a sensor on LANDSAT 4 and 5 satellites that records information in seven spectral bands, has a spatial resolution of about 30 m x 30 m, and represents digital values in 256 levels of brightness per band

transect - a transversely cut line along which physical and biological observations are made

trophic structure - the various levels in a food chain, such as producers (plants), primary consumers (herbivores), and secondary consumers (carnivores)

Universal Transverse Mercator - one of several map projections or systems of transformations that enables locations on the spherical earth to be represented systematically on a flat map

Universal Transverse Mercator grid - a geographic reference system used as the basis for worldwide locational coding of information in a GIS or on a map

unsupervised classification - the definition, identification, labeling, and mapping of natural groups, or classes, of spectral values within a scene. These spectral classes are reasonably uniform in brightness in several spectral channels.

vector format - a data structure that uses polygons, arcs (lines), and points as fundamental units for analysis and manipulation in a Geographic Information System

virtual reality - a computer-generated simulation of reality with which users can interact using specialized peripherals such as data gloves and head-mounted computer graphic displays

wildlife habitat relationship model - a method of linking patterns of known habitat use by animal species with maps of existing vegetation, thereby identifying the spatial extent of important habitat features for use in conservation and management.

Acronyms

ACSM American Congress on Surveying and Mapping
ADAMAS Aquatic Database Management System
ADEM Alabama Department of Environmental Management
AML ARC/INFO Macro Language
ASPRS American Society for Photogrammetry & Remote Sensing
AVHRR Advanced Very High Resolution Radiometer (satellite system)
BEST Biomonitoring of Environmental Status and Trends
BLM Bureau of Land Management
CAFF Conservation of Arctic Flora and Fauna
C-CAP Coastwatch Change Analysis Program (NOAA)
CDC Conservation Data Center
CEC Council on Environmental Cooperation
CENR Committee on Environment and Natural Resources
CERES California Environmental Resources Evaluation System
CIESIN Consortium for Internat'l Earth Science Information Network
CODA Conservation Options and Decision Analysis (software)
CRMP Coordinated Resource Management Plan
CRT Cathode ray tube (?)
CRUC Cooperative Research Unit Center
DLG-E Digital line graph - enhanced
DOI Department of the Interior
EDC EROS Data Center
ECOMAP The National Hierarchical Framework of Ecological Units mapping project of the USDA Forest Service
EMAP Environmental Monitoring & Assessment Program
EMAP-LC EMAP-Landscape Characterization (USEPA)
EMSL Environmental Monitoring & Systems Laboratory (USEPA)
EMTC Environmental Management Technical Center (NBS)
EOS Earth Observing System
EOSAT Earth Observation Satellite Company (the commercial operator of the Landsat satellite system)
EOSDIS EOS Data & Information System
ERL Environmental Research Laboratory, Corvallis (USEPA)
EROS Earth Resources Observation Systems (USGS)
ESRI Environmental Systems Research Institute
ETM+ Enhanced Thematic Mapper plus
FGDC Federal Geographic Data Committee
FTP file transfer protocol
FY Fiscal Year
GAO General Accounting Office (Congress)
GAP Gap Analysis Program
GCDIS Global Change Data and Information System
GLIS Global Land Information System (USGS)
GLOBE Global Learning and Observations to Benefit the Environment
GPS Global Positioning System
GRASS Geographic Resources Analysis Support System
GRIS Geographic Resource Information Systems
HRMSI High Resolution Multispectral Stereo Imager
IALE International Association of Landscape Ecology
IDRISI A GIS developed by Clark University

LAPS Land Acquisition Priority System
LC/LU Land Cover/Land Use (USGS)
MIPS Map and Image Processing System
MOU Memorandum of Understanding
MMU Minimum mapping unit
MRLC Multi-Resolution Land Characteristics Consortium
MSS Multi-Spectral Scanner
MTPE Mission to Planet Earth
NAFTA North American Free Trade Agreement
NALC North American Landscape Characterization (USEPA, USGS)
NAWQA National Water Quality Assessment (USGS)
NBII National Biological Information Infrastructure
NBS National Biological Service
NCCP Natural Communities Conservation Planning program (in CA)
NDCDB National Digital Cartographic Data Base
NERC National Ecology Research Center (Ft. Collins, CO)
NMD National Mapping Division
NPS National Park Service
NSDI National Spatial Data Infrastructure
NSTC National Science and Technology Council
NWI National Wetlands Inventory (USFWS)
OMB Office of Management and Budget (Administration)
OSIS Oregon Species Information System
PARC Public Access Resource Center
PI Principal Investigator
SAB Science Advisory Board (USEPA)
SCICOLL Scientific Collections Permit Database
SDTS Spatial Data Transfer Standard
SGID State Geographic Information Database
SNEP Sierra Nevada Ecosystem Project
SOFIA Southern Forest Inventory and Analysis
SPOT Système Pour l'Observation de la Terre
RMSE Root mean square error
TIGER Topologically Integrated Geographic Encoding and Referencing system (used for U.S. census)
TM Thematic Mapper
TNC The Nature Conservancy
UNESCO United Nations Educational, Scientific, and Cultural Organization
URISA Urban and Regional Information Systems Association.
URL Universal Resource Locator
USFS US Forest Service
USFWS US Fish & Wildlife Service
UTM Universal Transverse Mercator
UVM University of Vermont
WHRM Wildlife/habitat relationship model
WISCLAND Wisconsin Initiative for Statewide Cooperation on Landscape Analysis and Data