CALIFORNIA NATURAL DIVERSITY DATA BASE

Blair Csuti
California Natural Diversity Data Base
c/o Planning Branch, Department of Fish and Game
1416 Ninth Street
Sacramento, CA 95814

INTRODUCTION

Management of wildlife and wildlife habitat usually implies directed action toward particular goals. Even in those cases where no action is involved, a decision making process has elected this technique to achieve management objectives. In the ideal situation, each management decision will be based on an analysis of the most complete possible data set. Due to the uncoordinated way in which information is frequently gathered and distributed, managers must often make decisions on fragmentary or obsolete data when, in fact, better information exists elsewhere in the resource data management system. The California Natural Diversity Data Base (CNDDB) is one of several components of a resource information retrieval system being developed in California which seeks to improve access of decision-makers to state-of-the-art information prior to making decisions.

It is important that the level of detail at which information is available be compatable with the scope of the management problem being addressed. For example, it is accurate to say that the grizzly bear occurs in western North America; however, this fact is of little use in development of a management plan for a unit of the High Sierra Ranger District. Alternatively, knowing that a pine marten was trapped 3.2 mi W Hobart Mills is of little use when attempting to determine the percentage of forest to be managed for old growth in the Sierra-Cascade mountain system. Within the California resource information system framework, CNDDB functions at the greatest level of resolution, focusing on locality-specific data for a variety of species and communities of management interest in the state.

The California Natural Diversity Data Base was developed through a cooperative effort of The Nature Conservancy and the California Department of Fish and Game. Established in Sacramento in November of 1979, the Data Base functions as a unit of the Department's Planning Branch. The objectives of the Data Base are: 1) assisting in the systematic identification and selection of natural areas deserving protection; 2) serving as a centralized source of information allowing land-use planners and managers to assess the potential of projects to impact species and communities of concern at the planning stage; and 3) identifying gaps in our knowledge of biota requiring research. The methodology for the Data Base has been developed over the last decade by The Nature Conservancy for its Heritage Programs in 27 states. The procedure has three basic activities which are fine tuned through feed-back from accumulating data: classification, inventory, and analysis.

CLASSIFICATION

The initial task of a Data Base dealing with species and communities of concern is to identify those "elements" of natural diversity to be included in the inventory. An "element" is defined as a natural feature (species, plant community, etc.) of particular interest because it is exemplary, unique, threatened or endangered on a statewide or national basis. It is neither practical nor necessarily desireable to manage fine-scale information on every plant and animal species in California. For many very common species (e.g., deer mouse, western fence lizard, scrub jay) tracking detailed information about every specimen record or sighting would quickly surpass the available staff time and the capacity of any practical computer system. Such species are usually abundant in suitable habitat throughout their range and for such elements a range map, habitat description, and status summary provides ample guidance for management. Such elements are best dealt with

through information systems operating at a lower level of resolution. From all the plant and animal species and natural communities in California, the Data Base has identified a subset of "special" elements for inventory purposes. The term "special" is used as opposed to "rare," "endangered," "sensitive," etc., because each of the other words have certain legal connotations. These lists of elements are intended to be inclusive of other species and community lists of elements of concern, and they are continually modified as evidence on element status accumulates.

ANIMALS. The list of Special Animals is a composite of several other lists of species of concern with input from concerned agency and academic biologists. It is intended that the CNDDB list of Special Animals include non-marine animal species addressed by agencies in California. Component lists include the U. S. Fish and Wildlife Service list of federally endangered and threatened species, as well as all proposed and candidate species considered by USFWS. Also included are both U. S. Forest Service and BLM "Sensitive" species, as well as birds and mammals "of special concern" as identified in draft form by the California Department of Fish and Game. Additional nominees for listing were suggested by agency biologists. A draft list was circulated to major museums and research centers for further comment. As a result, over 340 animal taxa are included in the CNDDB inventory. The degree of endangerment varies greatly among them; thus, one of the functions of the Data Base is to rank them based on their current status and available ecological information.

PLANTS. The California Native Plant Society (CNPS) has been developing files on rare plant species in California for more than a decade. The rare plants being addressed by CNDDB include all those taxa listed in the CNPS 1980 inventory of rare plants and supplements. Those taxa considered by the CDFG's Endangered Plant Program and those identified by the USFWS in the Federal Register are included in the list. Through a cooperative agreement with CNPS, the society's rare plant files are housed at the Data Base under the supervision of a CNPS botanist. Since over 1700 taxa are involved in the inventory, initial attention has been to those plants appearing on state and federal lists of endangered, threatened, or rare species, followed by those additional species of highest biological concern in the state. Changes to the list are reflected by CNPS supplemental lists of taxa.

NATURAL COMMUNITIES. In addition to species of plants and animals, there are certain definable assemblages of flora and fauna that form vanishing elements of California's natural diversity. Identification of these begins from the same perspective as that of species, except that a defensible classification must be adopted so that endangered elements can be located within the milieu of California's natural communities. For classification purposes, CNDDB has adopted a modified form of a natural community classification developed by Dr. N. Cheatham and Dr. R. Haller of the University of California. From this list, certain sets of vanishing natural communities (e.g., vernal pools, coastal dune vegetation) have been selected for preliminary data processing attention. Ultimately, occurrences for all of California's vanishing natural communities will be tracked by the Data Base.

OTHER ELEMENTS. Other components of the state's diversity, such as geologic features and archaeological sites may ultimately be addressed by the CNDDB. Plans call for the addition of an aquatic biologist to the staff in the near future to address the problems of an aquatic classification system for California and identification of unique or unusual aquatic features of natural diversity.

INVENTORY

Once a classification is established, Data Base biologists gather information about the biology, status, and distribution of each element, beginning with the set of highest priority elements. Information is accumulated on each element in "Element Files" and this is condensed into a series of biological status reports (Element Abstracts). Information on each element is searched for distributional data, each locality being known as an "Element Occurrence." An Element Occurrence (EO) is the documented location of that element in California. The biological parameters used to define an EO vary from element to element and are documented in the Element Abstract. For example, each location of a salamander may constitute an element occurrence, but it may be more reasonable to combine several sightings of more mobile species into a single occurrence. For some species (such as birds of

prey) only sites fulfilling specific ecological requirements such as nesting may be considered EO's.

SOURCES OF INFORMATION. Data for each element comes from a variety of sources, although most information comes, directly or indirectly, from agency or academic sources. A notable exception is rare plant field survey data, which comes primarily from CNPS volunteer botanists and from U. S. Forest Service botanists who directly supply the Data Base with their observations.

For Special Animals, much of the knowledge of distribution resides in major systematic collections. California is fortunate to have many major and minor collections within its boundaries, and many of these have been consulted for information on localities of Special Animals or for access to their catalogs so that we may obtain these data at a later date. Cooperating systematic collections include: Museum of Vertebrate Zoology, University of California, Berkeley; Dickey Collections, UCLA; California Academy of Sciences; Los Angeles County Museum; and the San Bernardino County Museum of Natural History. Records from the Cornell Laboratory of Ornithology have been acquired for avian species on our list.

Publications and species files of the CDFG Nongame Wildlife Investigations Program and Inland Fisheries Branch have provided the bulk of recent information on high priority species. Information gathered by other agencies, such as the Bureau of Land Management (especially the Desert Plan Staff), U. S. Forest Service, Army Corps of Engineers, and Department of Defense have been incorporated into our Element Files as well. Other sources of information include field survey reports and reports published in the literature.

Much of the information on the status and distribution of rare plant species in California has been compiled by the California Native Plant Society. Photographic records of the specimen labels from the major herbaria of California have been accumulated by CNPS. Recent collection information is often available in CNPS files or on field survey forms coordinated through the CNPS botanist at the Data Base. Agency and academic contracts are maintained to insure receipt of reports from recent field surveys and research.

Data on distribution of Natural Communities is also available from a variety of sources, but tends to emphasize information from recent field surveys aimed at particular communities (e.g., riparian areas, vernal pools, coastal dunes, desert oases). An additional source of information for part of the state is vegetation maps prepared by the U. S. Forest Service in the 1930's (Wieslander Maps). Coinciding with 15' series topographic maps, the locations of uncommon communities can be identified in many chaparral and woodland areas of coastal and southern California using this source. As highest priority natural communities are identified, locality information for each of them is being integrated into the Data Base in a systematic fashion; however, the difficulties of classification and nomenclature of communities requires the gradual building of this data base. Unlike information on rare plant species, no pre-existing compilation of these data was available to serve as a nucleus for CNDDB efforts.

DATA PROCESSING. A standardized system of data flow has been established for all element classes. This system integrates information from a wide variety of sources into a single set of manual and computerized files, using the Element Occurrence as the basic unit of information. CNDDB is a primary user of the Department of Water Resources Interactive Graphics System (IGS). This electronic data processing system allows integrated textual and graphic retrieval of information stored in the Data Base. Using a PDP 11/70 central processing unit and associated peripherals, this hardware-software system enables cooperating agencies to querry up to 63 levels of graphically stored data on a single display. Available at a variety of scales, this system offers considerable potential for analysis of complex planning problems as its total user-oriented data base is expanded. Rare element localational data from CNDDB is available as one set of these graphically stored data.

Each new source of information is assigned a "Document Code" as it is examined for locality-specific information about the existence and description of rare elements of diversity in California. This code facilitates tracking of information retrieval systems. A standard computer transcription form (Element Occurrence Record=EOR) is prepared for each location of a special element. All additional reports about the element at that place are included

on the same EOR. It is important to distinguish between locations of rare elements and reports about rare elements. The Data Base tracks the places elements are seen, and combines multiple sightings or sightings over a period of time onto the same occurrence record. While the EOR format is identical for all elements, the biological parameters significant for the various elements vary. For that reason, each occurrence record contains space for general comments about habitat features, element status, land use, threats, etc. Additional site detail can be stored in non-computerized form on the reverse of the EOR.

Following transcription, each EOR is mapped onto 7½' USGS topographic sheets (15' series maps are used where 7½' is not available). Two important events occur at this step: 1) each occurrence of a rare element is assigned a unique identification number for future retrieval; and 2) the precision of the locational information for each occurrence is quantified. Three levels of precision are used to describe the level of uncertainty present in geographical information. Point occurrences are those for which information is accurate to 1/8th of a section (approximately a 1/5 mile radius). This category includes all occurrences from sources providing mapped boundaries for occurrences. Another example would be a point record referenced to a landmark. For example, Cottonwood Creek, 4280 ft, can refer to only one place, whereas "mouth of Cottonwood Creek" would refer to an area of undetermined size and location around the lower terminus of the same creek. A non-specific occurrence is accurate from 1/8 to 3 sections, or a 1 mile radius, and a general occurrence is accurate to an area of a 7½' topographic sheet or a 5 mile radius. Interpretation of retrieved information must recognize that existing knowledge of rare species and community locations is often imprecise, and that computerized retrieval cannot add precision to such data.

In effect, each Element Occurrence is a statement of the probability of finding a species or community within a given radius of some point. When recent field survey information documents precise locations, those data will be more useful than turn of the century museum records from a broad area. The Data Base can only retrieve the best available knowledge of a specified area. Since that knowledge is often vague, refinement in the form of updated field survey information is incorporated as it becomes available.

Entry of textual information on localities takes place following mapping. Information is keyed directly from the EOR into a remote terminal with a small memory. These data are then transmitted directly to the mainframe computer, and are immediately user-accessible. Next, the locations of each occurrence from topographic sheets are entered into the graphics system, a process which supplies various geo-based parameters to the EOR (e.g., latitude and longitude, county, biogeographic provence, watershed). Once entered into the graphics system ("digitized"), each occurrence becomes part of the systems' memory for that portion of California and is accessible for graphics retrieval and manipulation. Finally, each graphics record is attached to its corresponding textual record, providing documentation for localities in the graphics system and facilitating user applications.

USE AND ANALYSIS

The California Natural Diversity Data Base is in its third year of development. Currently over 10,000 occurrences of rare plants, animals, and natural communities have been transcribed, including nearly all available occurrences on state and federally listed plant and animal taxa. More than 6,000 of these have been entered into our manual map files, and 4000 entries have been made into the textual component of our computer system. At this writing (March, 1982) rapid progress is being made toward the goal of having all 6,000 occurrences for listed taxa available through both textual and graphics computer retrieval. Several public and private agencies have become regular subscribers of the Data Base, while individual searches of our files have been conducted for several hundred other users.

A primary use of the Data Base is the objective identification of areas of high natural values. Both The Nature Conservancy and the Significant Natural Areas Program (SNAP) of the Department of Fish and Game are concerned with the most efficient use of limited private and public conservation resources. CNDDB uses methodology developed by The Nature Conservancy for assessing the status of each element of natural diversity. Localities for

highest priority elements are identified and presented in abbreviated format for conservation groups. Documentation supporting the selection of elements and sites is available from CNDDB staff.

From a resource manager's viewpoint, the centralization and accessibility of rare element data for California may be the most valuable function of CNDDB. The Data Base can perform textual and/or graphics searches of specified areas and prepare reports in a variety of formats. Complete details are available on request from the Data Base. Textual locality information can be searched by mapping units (county, T-R-S coordinates, latitude and longitude polygons, etc.) or by a variety of descriptors (species name, family name, etc.). Output varies from a complete printing of all computerized information on each record to a simple list of occurrences by Township-Range-Section. Associated graphics output is available at several scales. While data can be plotted to any scale on special order, standard products are available at the 7½' (1:24,000) or 15' (1:62,500) scale, at the 1:250,000 scale, or at a 8½x11" format for California (1:5,500,000). Graphics products of the Data Base may take advantage of overlaying any of the data sets stored in the DWR IGS system onto a given map. At the moment these include such features as county outlines, biogeographic provences, Bailey Ecoregions, and watersheds. For selected areas, river and stream systems and soil types are also available. The boundaries of some managed areas (land held by a public or private group for some designated use - e.g., CDFG Ecological Reserves) are currently digitized, and many other categories of managed areas will ultimately be digitized.

There are a variety of applications of these products for land-use planning and impact assessment. For the first time existing knowledge of an area can be easily retrieved for all elements (rare plant and animal species and communities) by consulting a single source. The CNDDB functions to transfer data spread throughout volumes of documents into mapped localities which can be searched manually or electronically. For a given area, known past occurrences of rare species can be retrieved and used (in conjunction with knowledge of the habitat of the species) to predict the potential for encountering that element. Timber cuts, grazing leases, energy exploration and development projects, highway or commodity-route sighting, and habitat conversions are types of projects whose planners may benefit from a search of the Data Base.

It cannot be overemphasized that, while the Data Base is an efficient method for assessing existing knowledge, it can never substitute for on-site field verification. Given the rate of habitat change in California, even recent information may already be obsolete. Similarly, older records may still represent extant populations of rare species. Only field biologists can say for sure, and even they are subject to seasonal and climatic limitations and stochastic errors. Often rare species are especially elusive on the day of the scheduled field survey! Further, while the Data Base attempts to transcribe original data as accurately as possible, misidentifications exist even in the best data sources. Documenting the presence of an element on a given site remains the responsibility of the user. A corollary is that the absence of records of rare elements in a search area only means that no such information has been processed by the Data Base, not that nothing of concern is present in a given area.

Returning to the problem of scale of information retrieval systems, there are two ways to assess the probability of any element (species, community) occurring in a given area. The first is the CNDDB approach: plot all known records of occurrence and determine if any fall near an area of interest. Here the probability of occurrence deminishes as one travels farther from known records. The second approach is to prepare a range map from the known limits of distribution, then to determine whether the area in question falls within the range of the element. Range maps still can only assess the probability of encountering a particular element of interest, and their predictive ability is enhanced by matching the habitat of the species with those habitats present within a search area. The success of each method rests on a knowledge of habitat and habitat requirements combined with field investigation. Information retrieval systems such as the CNDDB are helpful in efficiently providing the manager with existing knowledge of an area or element. It is unreasonable to expect all local or regional resource managers to maintain comprehensive files on every possible element of concern in their areas of responsibility. However, the CNDDB can supply

those data, allowing decision makers to then devote their resources to acquisition of key pieces of information not yet available, or to recognize that sufficient data do exist for the decision at hand and to procede accordingly.

With the incorporation of the CNDDB into the Department of Fish and Game as of January 1, 1982, this information retrieval system has the potential for serving a variety of department and inter-agency data needs. While the scope of elements addressed by the CNDDB will remain limited, it offers fine-scale information on species and communities of high interest to California land-use managers. Other resource information management systems in the state deal with more widely distributed organisms and communities at a lower level of resolution, and the integration of these various systems into resource planning for the state offers a challenge for wildlife and habitat resource managers of the coming decade.