The Condor 92:564–570

AN INTRODUCTION TO THE USE OF GEOGRAPHIC INFORMATION SYSTEMS FOR ORNITHOLOGICAL RESEARCH[1]

DENICE M. SHAW AND SAMUEL F. ATKINSON
Institute of Applied Sciences, Department of Biological Sciences, University of North Texas, Denton, TX 76203

Abstract. Geographic Information Systems (GIS's) facilitate analyses of spatial and temporal data. This paper provides an introduction to the terminology, components, advantages, and current limitations of computerized GIS's. Case studies addressing the identification and characterization of potential habitat for the Golden-cheeked Warbler (*Dendroica chrysoparia*) and Black-capped Vireo (*Vireo atricapillus*) are provided for illustration of the potential utility of GIS's for ornithological research.

Key words: Geographic Information Systems; Golden-cheeked Warbler; Dendroica chrysoparia; *Black-capped Vireo;* Vireo atricapillus; *remote sensing.*

INTRODUCTION

Geographic Information Systems (GIS's) are computerized tools designed to assemble and analyze spatially oriented data. GIS's can also store and analyze descriptive attributes associated with spatial data. These systems can analyze data over larger areas and with more variables than would be possible with traditional (non-computerized) manual methods.

Monitoring and organizing attributes associated with spatial location and time are the nature of most GIS's (Jensen 1986). As an example of the type of analyses which can be addressed, consider a project to monitor nest placement over a specific region. Each nest has a spatial location (defined by X,Y coordinates) and an accompanying set of attributes, for example, species, clutch size, height in tree, all as a function of time (year). Each of the attribute data sets is termed a data layer or theme. When these layers are aligned, or spatially registered, within a GIS they can be simultaneously compared and related (Fig. 1).

This paper presents an introduction to GIS's and describes their potential as tools for ornithological research. Examples are provided from studies which have used GIS's for the description and characterization of potential habitat for the Golden-cheeked Warbler (*Dendroica chrysoparia*) and Black-capped Vireo (*Vireo atricapillus*). Both of these studies, which required evaluation of large volumes of data and multiple variables, were well suited for GIS analyses.

Spatial data within a GIS are stored as one of three basic geometric categories: points (e.g., nest locations), lines (e.g., roads), and polygons (e.g., ponds). Depending on the GIS, data are stored in one of two basic structures: raster or vector. Raster data structure can be perceived as an array of cells. Each cell (designated by an X,Y address) contains a value representing an attribute of the specific location. Examples of commercially available raster-based GIS's are ERDAS (Earth Resources Data Analysis Systems), GRASS, IDRISI, and I²S (Table 1). For vector-based GIS's, data are stored as a sequence of precise X,Y coordinates and vectors (which connect the coordinates), with an accompanying alpha-numerical variable to denote an attribute. ARC/INFO, INTERGRAPH, and MOSS are examples of commercially available vector GIS's (Table 1).

The primary differences between raster and vector formats lie in their accuracy, data volume, and data display. With respect to accuracy, raster-based GIS's are restricted to a preset limit within which locations can be specified. This limit is the cell size. In vector-based systems, data elements can be specifically located without the spatial constraint of a cellular array.

Locational resolution for raster-based systems can be improved by reducing the pixel (cell) size; however, there is a subsequent increase in data volume. Vector-based GIS's, because of the data structure, require less storage space. For example, within a raster GIS, to encode an agricultural field measuring 1,000 m × 100 m with a grid cell size of 25 m × 25 m, 160 pieces of data must be stored. In a vector system, only the coordinates of the four corners are needed to represent the same area.

With respect to data display, raster systems

[1] Received 6 July 1989. Final acceptance 26 February 1990.

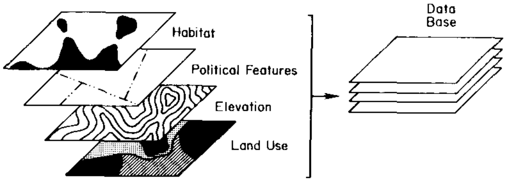

FIGURE 1. Conceptualization of Geographic Information System (GIS) data layers.

are limited in resolution to the pixel size. Vector systems are limited by the resolution of the input data. Figure 2 illustrates how a random polygon is represented in both raster and vector formats.

Despite the apparent advantages of vector-based GIS's Dangermond (1983) proposed that grid cell data structure is more efficient both in data storage and operation of analytic tasks. Additionally, digital remote sensing data are in raster format and therefore easily incorporated into raster GIS's.

Any GIS, to be most effective, should be able to accept, store, retrieve, and display both raster and vector data and have the ability to move between them. A review of the current status and predictions for future integration of these formats is found in Wallace and Clark (1988).

Both raster- and vector-based GIS's are comprised of four basic features: data entry, data management, data manipulation and analysis, and the generation and display of maps and charts (as well as other visual representations). In this paper, each feature is discussed and the difficulties and limitations encountered with each are presented.

TABLE 1. A selected list of raster and vector Geographic Information Systems.

Raster systems

- ERDAS—Earth Resources Data Analysis Systems
  ERDAS Inc.
  2801 Buford Highway, Suite 300
  Atlanta, GA 30329
- GRASS—Geographical Resources Analysis Support System
  Space Remote Sensing Center
  Building 1103, Suite 118
  NASA Natl. Space Lab, MS 39529
- IDRISI
  J. Ronald Eastman
  Graduate School of Geography
  Clark University
  Worcester, MA 01610
- I²S—International Imaging Systems
  Leuvensesteenweg 613
  B-1930 Zaventem-Zuid
  Belgium

Vector systems

- ARC/INFO
  Environmental Systems Research Institute (ESRI)
  380 New York Street
  Redlands, CA 92373
- INTERGRAPH*
  Intergraph Corporation
  One Madison Industrial Park
  Huntsville, AL 35807
- MOSS
  U.S. Department of Interior
  Bureau of Land Management
  Division of Advanced Data Technology
  Building 50, Denver Service Center
  Denver Federal Center
  Denver, CO 80225

* Also has raster capabilities.

DATA ENTRY

Data entry is one of the most important and labor intensive components of a GIS (Blakeman 1987, Chrisman 1987). It is at this stage that decisions of scale, resolution, and data type must be selected. A distinct advantage of GIS's is the potential for data assimilation from a variety of sources. Potential data can be any geographic information, provided they are referenced and stored using spatial coordinates and a standard geographic projection system (e.g., Latitude/Longitude, Universal Transverse Mercator). Most GIS's are capable of data conversion between commonly used map projections.

There are two major approaches for spatial data input to a GIS: manual (digitization) and through direct input of digital data. Manual digitization requires the use of specialized hardware and software. Depending on the level of detail required, digitizing may be a costly and time-consuming process. Direct acquisition of digital data saves manual input time, but can be costly to acquire as well as time consuming to convert into a compatible format.

Digital data may be acquired from a variety of sources. Table 2 provides examples of data currently available through several federal agencies (Streich and Austin 1986). The cost of these data varies as does availability (with respect to time and geographic area). These data may require preprocessing for system compatibility or to generate data in a particular format. The increasing production and use of digital data on a

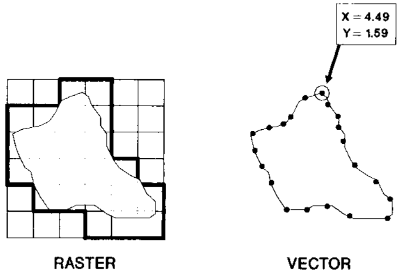

FIGURE 2. Illustration of raster and vector format data display in a Geographic Information System (GIS).

national level by the United States Geologic Survey (USGS), Bureau of the Census, and Environmental Protection Agency (EPA) will encourage simplified data conversion and eventually the establishment of data accuracy standards. It became apparent that improvements in this area are needed, particularly in the establishment of data standards to provide estimations of accuracy, precision, and scale of data (Estes 1986).

Digital data may also be directly entered with the use of electronic scanning equipment. The conversion of analog data sources (e.g., paper maps) to digital format continues to be one of the most significant obstacles to the effective use of GIS (Skiles 1988). Research and development in this technology is underway and major improvements are soon expected. A review of current scanning techniques and tools can be found in Skiles (1988).

DATA MANAGEMENT

Once spatial data have been entered they must be stored in a format that can be readily retrieved. Each data variable (i.e., GIS layer) is archived in digital format as a spatially referenced plane in a data base. When registered to one another, these planes form a data bank composed of n layers (n = the number of GIS layers) which can be queried.

A common problem for both raster- and vector-based systems at this stage is data storage. A potential solution is to use computer compatible tapes (CCT) or optical laser discs for temporary or permanent data storage. Clark and MacGaffey (1988) suggest that optical disk drives, with improved storage capacity, will replace magnetic media and soon become cost effective for GIS users. Another solution, for personal computer (PC) based systems, is an interface with a mainframe computer system or workstation fileserver which would allow access to the large volume remote disks for data storage.

DATA ANALYSIS AND MANIPULATION

Data manipulation in a GIS is the process by which data bases are queried. An extensive review of data manipulations available on most GIS systems can be found in Dangermond (1983). This paper will discuss only two of the most fundamental aspects, spatial and statistical analyses.

Spatial analyses include computations such as overlays, proximity searches, measurements, and the use of filters and reclassifications. Overlay analyses, including Boolean queries and/or algebraic manipulations, are used to integrate two or more data layers to create a third. Overlay analyses are useful for identifying spatial or temporal coincidence. For example, a temporal comparison of nest sites and brood size over time can be accomplished with overlay analyses.

Proximity or neighborhood analyses are performed within adjacent cells of one GIS layer (whereas spatial coincidence analyses are performed to establish relationships between different GIS layers). Proximity analyses can be used to identify the selective juxtaposition of cells with respect to one another (Johnson et al. 1988). An example of this would be the extension of a buffer around individual nests to identify the tree species within 500 m. For both proximity and overlay analyses, raster-based systems are limited with regards to the accuracy for calculations of area, perimeter, and volume; larger grid cells result in more error. This is not a problem for vector-based systems.

Statistical analyses in a GIS include the generation of histograms, frequencies, regressions, and correlations as well as the ability to interface with standard statistical packages such as Statistical Analysis Systems (SAS) or Statistical Package for Social Sciences (SPSS). Currently, extensive manipulations are required to get data into a format compatible with a statistical package such as SAS. Improvements to facilitate data transfer between the GIS and these statistical packages are still needed.

DATA DISPLAY AND PRODUCT GENERATION

Clear and effective graphic output is important for communicating the results of GIS analyses

TABLE 2. Digital data sources available from the U.S. Federal Government.

Agency	Data type
Bureau of the Census Customer Services Data Users Services Division Washington, DC 20233	Agriculture, livestock, demographic, and economic
EROS User Services EROS Data Center Sioux Falls, SD 57198	Digital Landsat Imagery
National Climatic Data Center (NCDC) Federal Building Asheville, NC 28801	Meteorological, soil temperature, and storm tracking
National Geophysical Data Center NOAA-E/CGI 315 Broadway Boulder, CO 80303	Seismic, gravity, geologic, magnetic, and geothermal
National Oceanographic Data Center (NODC) User Services Branch NOAA/NESDIS E/OC21 Washington, DC 20235	Current, oceanographic station, marine mammal, marine birds, primary productivity
USDA Soil Conservation Service Cartography and GIS Division 14th and Independence, W.W. P.O. Box 28990 Washington, DC 20013	Soil mapping
United States Geological Survey (USGS) National Cartographic Information Center 536 National Center Reston, VA 22092	Elevation, political boundaries, and transportation

to the user. For this reason GIS's generally include the capability for displaying maps, charts, graphs, and tabular information. The hardware on which this output is generated includes line printers, color monitors, laser printers and plotters, ink jet printers, thermal wax printers, electrostatic plotters, video recorders, and photographic and microfilm devices. The display media on which they are produced include mylar, photographs, and paper.

EXAMPLES

To illustrate the use of GIS's for ornithological research, habitat studies conducted with the Golden-cheeked Warbler and Black-capped Vireo are described. Two raster-based (ERDAS and IDRISI) and one vector-based GIS's (ARC/INFO) were used for the analyses.

A variety of data sources was used for these studies. One source, common to both, was the acquisition of digital Landsat satellite imagery (Table 1). For the Golden-cheeked Warbler research, these data were employed for the identification of a specific vegetative community. For the Black-capped Vireo, satellite imagery was used for the identification of a generalized category of shrub vegetation commonly utilized as nesting habitat by the Black-capped Vireo.

Two types of Landsat data were acquired, Multispectral Scanner (MSS) data (80 m spatial resolution) and Thematic Mapper data (30 m spatial resolution). Both of these data types are multispectral, denoting that radiant energy from the earth's surface was recorded for multiple portions of the electromagnetic spectrum, termed bands (four portions for MSS and seven for TM).

For the Golden-cheeked Warbler study MSS data were used. Coarser resolution data were deliberately chosen to minimize computer storage requirements and computer processing time required for analyses. Portions of four Landsat MSS scenes were required to cover the entire study area. Storage for each Landsat scene required approximately 37 megabytes of space.

TM data were used for the Black-capped Vireo research. These data were preferred because of the improvement in both spatial and spectral resolution which was necessary to identify shrub vegetation. Because of the small study area, despite the improved resolution and seven bands of TM, data storage was not a problem. All the digital satellite imagery data were analyzed with the image analysis capabilities of ERDAS software.

Another source of data for both studies was Digital Elevation Models (DEM), obtained from the USGS. DEM data consist of an array of elevations referenced on the latitude/longitude coordinate system (USGS 1987). With computer processing, elevation data can be used to extract slope and aspect data for a given region.

For both studies, manual digitizing was necessary to obtain some data not available in digital format. Major water bodies, towns, and cities as well as political features such as county boundaries were digitized from USGS 1:250,000 and 1:24,000 scale maps. These features were used primarily for spatial reference but may be useful for future analyses.

GOLDEN-CHEEKED WARBLER

The research for characterization of habitat for the Golden-cheeked Warbler was a component of the U.S. Fish and Wildlife Service (USFWS) species status survey (Wahl et al. 1989). The study area included the entire breeding range of the species, approximately 73,000 km^2 in central Texas. In the preliminary stages of this research, digital Landsat satellite data were classified to identify the characteristic Ashe juniper (*Juniperis ashei*)—oak woodland utilized as nesting habitat by the species.

Classification of Landsat data was conducted with ERDAS image processing software. Once the Ashe juniper-oak woodland was identified, a GIS was employed to estimate total available habitat, generate a baseline inventory, and describe some of the spatial parameters of habitat patches in selected portions of the study area (Shaw 1989, Wahl et al. 1989).

A variety of GIS spatial analytical techniques were employed for this research. Overlay algorithms were used to subset the data base with respect to individual counties. From this, measurement calculations were used to estimate hectares of potential habitat per county.

For a series of test regions across the study area, a set of indices was employed to describe spatial characteristics of the patches of potential habitat. These analyses could only be conducted with the data processing and storage capabilities of a GIS; thousands of habitat patches were evaluated within each county. One index was the evaluation of the size distribution of patches for each of the test regions. Another index, a modification of a gravity model (Hartshorn 1980) was used to relate patch size to the distance between adjacent patches for each region. The third index, termed a configuration index (Shaw 1989) described the deviation of the perimeter of a patch from the circumference of a circular patch of equivalent area. This index was used to measure the proportional amount of edge. The information from these three indices provided an assessment of the spatial distribution and characteristics of habitat patches for each of the test areas. When these indices were compared between the areas, distinct differences emerged. For example, patches in the southeastern portion of the study area were the largest and most contiguous but had the highest proportional edge to volume ratios. Moving north, the patches were smaller, more fragmented, and had lower edge to volume ratios. These differences are most likely attributable to geologic features which in turn influence land use and development patterns (Shaw 1989).

BLACK-CAPPED VIREO

The Black-capped Vireo research was a segment of the U.S. Army Corps of Engineers' assessment of potential vireo nesting habitat on Camp Bullis, a 113-km^2 military installation in central Texas. One objective of the investigation was the utilization of a GIS to identify and characterize potential nesting habitat for the Black-capped Vireo. For this GIS's were used to estimate total available habitat, generate a baseline inventory, and integrate the spatial locations of nest sites with variables such as underlying geology, slope, and aspect.

Habitat for known locations of Black-capped Vireo, was identified through satellite interpretation and field census during the 1989 breeding season. Once territories were identified, a GIS was used to investigate relationships between topographic and geologic features and the locations of Black-capped Vireo territories on Camp Bullis (Shaw et al. 1989). To do this, territory locations were merged with a data layer of underlying geologic formation, aspect, and slope.

This investigation identified apparent relationships between locations of Black-capped Vireos and slope, aspect, and underlying geologic formation. Once these relationships were identified, additional areas were located on Camp Bullis which met these criteria. Final analyses included acreage estimations of potential habitat that were identified through the integration of the classified Landsat TM imagery with slope, aspect, and underlying geologic formation data.

For both the Black-capped Vireo and the Golden-cheeked Warbler studies, GIS's were used to generate data tables as well as maps and three-dimensional images indicating the distribution of potential habitat. Maps, including those identifying potential habitat, were produced on mylar and paper maps at several scales. The maps were produced for use in the field to verify the results of the Landsat image classification as well as for display and demonstration. Limitations encountered at this phase in both projects were with the accuracy with which maps could be generated. For example, despite calculations which indicated data were correctly registered between the computer maps and the paper, maps produced by an inkjet printer could not be completely aligned with all features on the map. This type of problem may be caused by both software and hardware limitations.

DISCUSSION

For the research described here, the use of GIS's has offered improvements over manual approaches in the areas of data management, data organization, and the opportunity for description and modeling of spatial relationships. Computer technology to support GIS's operation is evolving to ever higher levels of capability and capacity. The technology has improved from a point at which previously only large institutions could support GIS technology to one in which researchers and resource managers can have GIS capability in a desktop personal computer (Estes 1986). Future research directions in GIS's involve integration with artificial intelligence and expert systems, improvement of spatial statistics, hardware developments for data processing and storage, and simplification of operating languages used for GIS's.

However, despite the obvious advantages of GIS's, limitations still exist and therefore caution must be exercised with their use. The major limitations include the lack of available digital data, lack of standards for digital data quality, and the lack of spatial statistics and models designed to address ecological phenomena. Proper use of a GIS requires careful attention to planning and organization. If input data contain errors, these errors will be passed on in cumulative and compounded fashion to the final results. This is especially important in a GIS, since data of differing accuracy and scales are combined into a final product.

GIS's offer a powerful tool for use in ornithology or any discipline for which interests are coupled with spatial data. GIS's can improve research efforts through better organization and documentation of project materials and methodology. GIS's also facilitate spatial and temporal analysis of larger areas and greater numbers than would be feasible through traditional methods. Current trends indicate that GIS capabilities for data input and management will continue to provide improvement to help in ornithological research.

ACKNOWLEDGMENTS

We would like to thank P. Marr, M. Reakes, and D. Whalley for their helpful comments on this manuscript and T. Nelson for help with the graphics. The studies addressed in this manuscript were funded by the Texas Park and Wildlife Department and the United States Army Corps of Engineers Fort Worth District.

LITERATURE CITED

BLAKEMAN, D. A. 1987. Some thoughts about GIS data-entry, p. 226–233. *In* Proceedings GIS/LIS '87—San Francisco. American Society for Photogrammetry and Remote Sensing, Falls Church, VA.

CHRISMAN, N. R. 1987. Efficient digitizing through the combination of appropriate hardware and software for error detection and editing. Int. J. Geogr. Inf. Syst. 1:265–277.

CLARK, S. R., AND N. MACGAFFEY. 1988. Optical disk technology: potential in geographic information systems, p. 541–549. *In* Proceedings GIS/LIS '88—San Antonio. American Society for Photogrammetry and Remote Sensing, Falls Church, VA.

DANGERMOND, J. 1983. A classification of software components commonly used in geographic information systems, p. 70–91. *In* D. Peuquet and J.

O'Callaghan [eds.], Design and implementation of computer-based geographic information systems. International Geographical Union Commission on Geographic Data Sensing and Processing, Amherst, NY.

ESTES, J. E. 1986. A perspective on the use of geographic information systems for environmental protection, p. 3–23. *In* Geographic Information Systems for environmental protection. Environmental Protection Agency, Washington, DC.

HARTSHORN, T. A. 1980. Interpreting the city: an urban geography. John Wiley and Sons, New York.

JOHNSON, L. B., C. A. JOHNSTON, AND J. PASTOR. 1988. Raster and vector data in ecological research applications, p. 386–394. *In* Proceedings GIS/LIS '88—San Antonio. American Society for Photogrammetry and Remote Sensing, Falls Church, VA.

JENSEN, J. R. 1986. Digital image processing, a remote sensing perspective. Prentice-Hall, Englewood Cliffs, NJ.

SHAW, D. M. 1989. Applications of GIS and remote sensing for the characterization of habitat for threatened and endangered species. Ph.D.diss., Univ. of North Texas, Denton.

SHAW, D. M., J. C. COFFEY, AND S. F. ATKINSON. 1989. Status report: the Black-capped Vireo on Camp Bullis, Fort Sam Houston, Texas. Report prepared for the U.S. Army Corps of Engineers, Fort Worth District.

SKILES, J. M. 1988. Tools and techniques used in scan conversion of geographic and facilities data, p. 58–67. *In* Proceedings GIS/LIS '88—San Antonio. American Society for Photogrammetry and Remote Sensing, Falls Church, VA.

STREICH, T., AND I. AUSTIN. 1986. Selection of geographic information systems and related services for environmental protection applications, p. 170–201. *In* Geographic information systems for environmental protection workshop proceedings. Environmental Research Center, Las Vegas, NV.

UNITED STATES GEOLOGICAL SURVEY. 1987. Digital elevation models—data users guide. Department of the Interior, U.S. Geological Survey. Data Users Guide 5.

WAHL, C. R., D. D. DIAMOND, AND D. M. SHAW. 1990. The Golden-cheeked Warbler: a status review. Report submitted to U.S.F.W.S., Fort Worth, TX.

WALLACE, T., AND S. R. CLARK. 1988. Raster and vector data integration: past techniques, current capabilities, and future trends, p. 418–426. *In* Proceedings GIS/LIS '88 San Antonio. American Society for Photogrammetry and Remote Sensing, Falls Church, VA.