EMPLOYMENT OF GEOGRAPHIC INFORMATION SYSTEMS FOR DETERMINING THE ACCURACY OF MUSEUM VOUCHER SPECIMEN DATA

KELLY E. ALLEN, ROBERT D. BRADLEY, R. RICHARD MONK, OLEKSIY V. KNIAZHNITSKIY, NICK C. PARKER, DAVID J. SCHMIDLY, AND ROBERT J. BAKER

Voucher specimens in museums are the foundation of our understanding of systematics, biodiversity, evolutionary genetics, biogeography, ecology, etc. Today, many economic, political, and scientifically important decisions are made based on the synthesis of knowledge of biological organisms and associated biocomplexity. Therefore, the identification of voucher specimens and, subsequently, the conclusions drawn from knowledge based on these specimens is critical to many aspects of society.

Classically, identification of voucher specimens was verified and maintained through the efforts of specialists for a given taxonomic group or geographic region. Changes in nomenclature, taxonomic status, and geographic distribution or regional status were determined by the specialist and easily updated in the museum collection. The constant addition of specimens over time has produced museum collections of remarkable sizes. As these collections increased in size, curation of collections, maintenance of taxonomic order, and the ability to choose or develop museum computer databases have proven to be challenging and labor intensive (Sarasan, 1981; Sarasan and Neuner, 1983; Monk, 1996; Monk, 1997). The classical ‘hands-on’ approach has been taxed by the size of collections and the ever-decreasing number of qualified specialist in a given field or discipline. For example, a recent survey of mammal collections in the Western Hemisphere indicates 17% or approximately 58 collections consist of 10,000 or more specimens (Hafner et al., 1997). Additionally, at Texas Tech University a Recent mammal collection was initiated in the early 1960’s, contained 5000 specimens in 1967, and today includes over 80,000 voucher specimens making the management of this collection a complex task.

The development of computers and computational technology since the 1960’s has provided a means for better management of large sources of data. Increased memory, data storage, software, manipulation, retrieval, analysis capabilities, and archival quality printing capabilities have substantially improved since the initial development of the computer (DeMers, 2000). Collection of digital environmental data through aerial photographs, satellite imagery, global positioning systems (GPS), and geographic information systems (GIS) has provided a means to assess issues of biocomplexity that previously were difficult or impossible (Baker et al., 1998; Parker et al., 1998). In addition, recent technological advances not only provide a means to store and manipulate large data sets, but permit new insights into potential ecological and taxonomic variation for voucher specimens and thereby provide a means to identify possible problems associated with some voucher specimens.
At Texas Tech University, comparison of documented distributions of Texas mammal species (Davis and Schmidly, 1994) with maps representing distributions created from habitat associations and known collection localities (Allen, 2000) resulted in the identification of several voucher specimens that were either possibly misidentified, represented significant range extensions (Figs. 1a, 1b, 1c), or required nomenclature changes. Several possibilities could result in the observed mismatches of the known collection localities. A full summary of these possibilities is given by Allen (2000), but can be narrowed to (1) the expansion of a species' range extent, (2) a limitation of the digital data resolution or scale, (3) temporal changes in habitat, (4) a nomenclature change, and (5) the misidentification of specimens. In this paper, the misidentification of voucher specimens refers to the assignment of an incorrect taxonomic name for a voucher specimen within the museum archive. This may be the product of misidentification of the voucher specimen or an error due to data handling (Monk, 1997). In this study, we examined voucher specimens for which there appeared to be a mismatch with either the documented range, predicted habitat, or actual specimen locality to explore which possibilities accounted for the apparent inconsistencies observed in the data sets.

METHODS

Geographic coordinates were assigned to all mammalian voucher specimens collected in Texas and housed within the Natural Science Research Laboratory (NSRL), Museum of Texas Tech University. Coordinates were assigned using a computer algorithm that determines the Universal Transverse Mercator (UTM) coordinate based upon the traditional descriptor location (Knyazhnitskiy et al., 2000). Assignment of UTM coordinates to voucher specimen localities provided a geographic location for that voucher specimen that could be easily translated by a GIS.

Maps illustrating predicted habitat for 127 native, terrestrial mammals were acquired from the Texas GAP Analysis Program. Predictive habitat maps were generated with the use of a database of known habitat associations for each mammalian species (Figs. 1b and 1c). A GIS was used to query predictive habitat from digital data sets consisting of vegetation, soils, hydrography, and documented range extents (Fig. 1a).

RESULTS AND DISCUSSION

Approximately 30,000 specimens were plotted in association with their respective species’ documented range extent and predicted habitat. Of these, 1073 (3.6%) voucher specimen records did not correspond with their respective documented range or predicted habitat. As a result, 221 voucher specimens were examined by mammalogists and 290 specimen records were examined using a GIS. In addition, 527 records of Perognathus flavus were recognized as mismatches with the documented range extent. However, these could not be verified at this time due to the complicated nature of the Perognathus complex. An additional 35 specimens were unaccounted for because either they were missing, had been deaccessioned, or were not pulled because it was obvious the mismatch resulted from a database error. After verification, 332
PINON MOUSE
*Peromyscus truei*
AMAFF03130

![Map of Texas showing documented range extent, predicted habitat, and NSRL collection localities.](image)

Documented Range Extent ¹

Predicted Habitat ²

Distribution Model ²

- NSRL Collection Locality


² Shaded area is the predicted distribution based on modeling habitat associations following Texas GAP Analysis Project.

Figure 1. Predicted habitat models for *Peromyscus truei* model. Models were created using a Geographic Information System and reflect (a) the documented range extent as published in Davis and Schmidly (1994), (b) predicted habitat occurring throughout the state and (c) predicted habitat restricted by the documented range extents.
Figure 2. Illustration of mismatches observed between voucher specimen localities and their representative documented range extents for (a) *Peromyscus truei*, (b) *Geomys bursarius*, and (c) *Perognathus flavus*. Shaded regions represent the documented range extents digitized from Davis and Schmidly (1994).
(31%) of the voucher specimen records initially identified as possible mismatches and comprising 20 species, 10 genera, and 5 families were found to be incorrectly identified to species or required nomenclature changes. Of these 332 specimens records, 313 (94%) required changes in nomenclature whereas 20 (6%) were the result of misassignment upon cataloguing. Examples of voucher specimens misidentified as a result of changes in nomenclature include Dipodomys ordii, Peromyscus attwateri, Peromyscus difficilis, Peromyscus boylii, and Peromyscus nasutus, and the Geomys complex. Examples of voucher specimens misidentified upon initial cataloguing include Tayassu tajacu, Myotis volans, Lasiurus ega, and Reithrodontomys fulvescens.

Two special cases of extensive mismatch of specific assignment of voucher specimens in the museum database and the ecological/geographic range involved the Geomys bursarius complex (Fig. 1.b) and the Perognathus flavus complex (Fig. 1.c). Both of these involved a substantial number of voucher specimens, Geomys complex, 290 records, and P. flavus, 527 records. The problem associated with Geomys resulted from specimens previously recognized as a single morphologically highly variable species (Hall, 1981) being subdivided into six different species (G. attwateri, G. breviceps, G. bursarius, G. knoxjonesi, G. personatus, and G. texensis) (Baker et al., 1989; Block and Zimmerman, 1991; Bradley, et al., 1991; Honeycutt and Schmidly, 1979; Tucker and Schmidly, 1981). Taxa within the genus Geomys exhibit allopatric distribution. Due to a lack of concensus regarding taxonomic arrangements, the museum database had not been updated. In addition to the specimens of G. bursarius being misidentified due to nomenclature changes, several specimens were found to be misidentified based upon the geographic location of the collection locality. Revisions of the museum database were based on the published ranges recognized in the above-cited papers and in The Mammals of Texas (Davis and Schmidly, 1994).

The problem with the Perognathus flavus complex is more difficult because this species complex involves two sibling (cryptic) species (P. flavus and P. merriami) that are often difficult to distinguish based on traditional skin and skull characteristics (Lee and Engstrom, 1991). Based upon karyology and biochemical methods, P. flavus and P. merriami have been documented to be reproducitively isolated but karyological and biochemical data are not available for all museum specimens. The proper assignment of voucher specimens must await more detailed studies of the museum specimens and the development of morphological methods of species identification. At present, it seems best to recognize that in the P. flavus/merriami complex “specific” assignments in this complex are tenuous unless biochemical or chromosome data are available and that ecological and other conclusions based on these specimens will need to be viewed in light of this situation.

The methods employed in this study assisted the curators by increasing accuracy of the museum database. The methods are limited, however, in that they only identify ecological or geographical outliers. Specimens from the assigned geographic and ecological ranges also may be misidentified. These misidentifications cannot be recognized using the methodology described in this paper because the voucher specimens appear to occur within the appropriate geographic or ecological range. Nonetheless, the methods employed here can be added to those available to museum curators to strengthen the accuracy of museum databases.

In addition, the authors believe this method provides a rapid, cost-effective means for the non-specialist to identify potential problems. A non-specialist here refers to someone that is not capable of verifying the vouchers species for determining the accuracy of a museum’s database. Obviously this method requires museums to have already converted descriptor locations of each voucher specimen to some format that is acceptable for a GIS. The conversion of a descriptor locality to one that can be interpreted by a GIS can indeed be very costly and time consuming to a museum. In this situation, the non-specialist would have to assume the museum’s specimens are correctly identified and reflect current nomenclature status. However, if the conversion to digital data already has been made and the location of voucher specimens is known by the museum, the comparison of documented range extents with voucher locations is easily done using an algorithm that automates the process of mapping the documented range extent with the geographic location of voucher specimens. Visual identification of mismatches can then be completed in a matter of hours.
The authors would like to stress that the application of GIS is no substitute for a knowledgeable specialist, a workable key, and an accurate range map for the purpose of identifying specimens. The strength of any set of data is a function of the collector and the means of processing the data. Any identified mismatches using the methods outlined in this paper should be further addressed by a taxonomic specialist.

ACKNOWLEDGMENTS

The authors thank two anonymous reviewers for their review of this manuscript. The U.S. Geological Survey-Biological Resources Division through the National GAP Analysis Program and Texas Tech University through the Natural Sciences Research Laboratory supported this research.

LITERATURE CITED


Addresses of authors:

**KELLEY E. ALLEN**

Texas Cooperative Fish & Wildlife Research Unit
Texas Tech University
Lubbock, TX 79413-2120

e-mail: rbradley@ttacs.ttu.edu

**ROBERT J. BRADLEY**

Department of Biological Sciences and Museum of Texas Tech University
Lubbock, TX 79409-3131

e-mail: rbradley@ttacs.ttu.edu

**OLEKSY V. KNYAZHNITSKIY**

Department of Biological Sciences and Museum of Texas Tech University
Lubbock, TX 79409-3191

e-mail: alexey_khn@hotmail.com

**R. RICHARD MONK**

Museum of Texas Tech University
Box 43191
Lubbock, TX 79409-3191

e-mail: rich.monk@ttu.edu

**Current Address:**

Central Oregon Community College
2600 NW College Way
Bend, OR 97701

e-mail: kallen@cocc.edu

**NICK C. PARKER**

Texas Cooperative Fish & Wildlife Research Unit
U.S. Geological Survey-Biological Resources Division, Texas Tech University
Lubbock, TX 79413-2120

e-mail: nparker@ttu.edu

**DAVID J. SCHMIDLY**

President, Texas Tech University
Department of Biological Sciences and Museum of Texas Tech University
Lubbock, TX 79409-3131

e-mail: schmidly@ttu.edu

**ROBERT J. BAKER**

Department of Biological Sciences and Museum of Texas Tech University
Lubbock, TX 79409-3131

e-mail: rjbaker@ttu.edu
It was through the efforts of Horn Professor J Knox Jones, as director of Academic Publications, that Texas Tech University initiated several publications series including the Occasional Papers of the Museum. This and future editions in the series are a memorial to his dedication to excellence in academic publications. Professor Jones enjoyed editing scientific publications and served the scientific community as an editor for the Journal of Mammalogy, Evolution, The Texas Journal of Science, Occasional Papers of the Museum, and Special Publications of the Museum. It is with special fondness that we remember Dr. J Knox Jones.

Institutional subscriptions are available through the Museum of Texas Tech University, attn: NSRL Publications Secretary, Box 43191, Lubbock, TX 79409-3191. Individuals may also purchase separate numbers of the Occasional Papers directly from the Museum of Texas Tech University.

ISSN 0149-175X

Museum of Texas Tech University, Lubbock, TX 79409-3191