Retrieval, Compilation, and Organization of Vertebrate and Vascular Plant Voucher Specimens Originating from National Parks

Andrew Gilbert, U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland 20708 (current address: U.S. Geological Survey, Patuxent Wildlife Research Center, Gardiner, Maine); andrew_gilbert@usgs.gov

Allan O'Connell, U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland 20708; allan_o'connell@usgs.gov

Introduction

Natural history collections are the fundamental source for understanding and interpreting biodiversity, but their value is unappreciated and poorly supported (Cotterill 1995). An important but often overlooked component of this information is the efficient retrieval and compilation of records available. Museums and herbaria have not done a good job of marketing their resources and services (Alberch 1993), and, as a result, the initial appeal and application of their data are limited. In fact, specimens in museums and herbaria are an enigma to many outside the museum environment because the information is often difficult to access. Furthermore, analyses using only partial datasets may provide results different from those obtained from a full complement of records. For example, in the use of butterfly lists to make biodiversity comparisons in Oregon, less than half of the dataset was used because only that portion was computerized (Fagan and Kareiva 1997).

Computerization of records in natural history collections is still in a long way from completion, and even when accessing collection data that are computerized, navigating the computer interfaces can be awkward without adequate guidance. Management of collections also varies greatly from site to site, often making retrieval of information a complicated process. Retrieval and access of specimen data will be necessary to objectively evaluate current inventory and monitoring efforts of our biological resources in the near future. Thus, our objectives in this study were to locate, compile, and organize specimen records originating within and around 14 national parks throughout the northeastern United States. (Table 1). We used a variety of strategies and techniques to search natural history collections for four different taxa in three vertebrate groups (mammals, birds, and reptiles and amphibians) and vascular plants. We developed procedures for assembling collection records into one of four locality categories in a manner that established a database of historical diversity for the National Park Service (NPS) at increasing scales, from within park boundaries to outside park boundaries at county and state scales.

Methods

We obtained information about vertebrate (except fish) and vascular plant natural history collections by first searching two web-accessible databases of natural history collections: the Index Herbariorum (IH; www.nybg.org/bsci/ih/ih.html) and the Directory of Research Systematics Collections (DRSC; www.nbii.gov/datainfo/syscollect/drsc/). We also sent out requests for information about collections to several e-mail listservs (TWS-L, NHCOLL-L, ORNITH-L) and obtained a list of museum contacts from John Karish (NPS, Philadelphia Support Office) from a similar project. Additional collection information was found by searching web sites of regional biology departments.

We mailed requests for data to 274 collection managers curating 299 natural history collections and 8 state natural heritage programs. We specifically requested data for specimens originating within the 14 northeastern national parks. Information about natural history collections was recorded in a Microsoft Access 2000 database. Collection

Table 1. National parks searched for vertebrate and vascular plant voucher specimens.

National Park (Code)	State(s)	Size (Ac)	Year Est.
Acadia National Park (ACAD)	ME	46,784	1916
Marsh-Billings-Rockefeller National Historical Park (MABI)	VT	555	1992
Minute Man National Historical Park (MIMA)	MA	967	1959
Morristown National Historical Park (MORR)	NJ	1,685	1933
Roosevelt-Vanderbilt National Historic Site (ROVA)[1]	NY	683	1940
Saint-Gaudens National Historic Site (SAGA)	NH	150	1964
Saugus Iron Works National Historic Site (SAIR)	MA	9	1968
Saratoga National Historical Park (SARA)	NY	3,406	1938
Weir Farm National Historic Site (WEFA)	CT	60	1990
Assateague Island National Seashore (ASIS)	MD	39,732	1965
Cape Cod National Seashore (CACO)	MA	43,604	1961
Fire Island National Seashore (FIIS)	NY	19,580	1981
Gateway National Recreation Area (GATE)	NY, NJ	26,610	1972
Sagamore Hill National Historic Site (SAHI)	NY	83	1963

[1] ROVA was consolidated from Eleanor Roosevelt National Historic Site (ELRO, est. 1977, 181 ac), Home of Franklin D.Roosevelt National Historic Site (HOFR, est. 1945, 290 ac) and, Vanderbilt Mansion National Historic Site (VAMA, est. 1940, 212 ac).

information was separated by taxa (e.g., Cornell University Museum of Vertebrates ornithology collection) where taxa-specific data were available. Information such as size of collection, percentage computerized, contact person and address, web address, and notes about the collections were recorded. We determined that much of the information provided in the two natural history collection databases were out of date; therefore, we checked contact information for all institutions through web sites or by contacting institutions directly and updated information as necessary.

To reduce search time and increase the number of responses from institutions, we broadened search criteria to county-wide locality requests. This approach also had the benefit of including locations that were misspelled or used historic names. We sent institutions a list of parks and localities by state and county. We requested that the following data fields be provided: park name, taxonomic name, common name, catalogue number, accession number, condition of specimen, collector's name, date of collection, locality information, latitude–longitude, and comments. We e-mailed follow-up requests for data to 177 collection managers who did not respond within six weeks of the initial request for data. We logged responses into the collection database as they were received. We established two databases: one in Microsoft Access for collections, and the other in Microsoft Excel for specimen records we located.

Results

We received a 70% response rate form the curators we queried and tallied information from 78 collections. We assembled 31,110 specimen records (30,833 categorized 1–4 by locality; Table 2) of which 4,745 (15%) are from within park boundaries (category 1) and an additional 4,552 (15%) may be from within park boundaries (category 2), but for which we do not have enough information to determine their exact location. We gathered the most specimen records for plants, followed by birds, mammals, and amphibians and reptiles. Within the four taxa, specimens comprised 260 families, 909 genera, and 2,055 species/species hybrids. Plant specimens represented the highest diversity of taxa with the greatest number of categories from species/species hybrids to families and genera. More than one-third of all records were from Acadia National Park, the largest and oldest park in this study. Acadia also had the most category 1 and 2 specimens (4,615) followed by Cape Cod National Seashore (2,180). We were unable to corroborate taxonomic identification due to time constraints. Most transfers of specimens to other institutions were catalogued as accessions, but in some cases disposal of specimens was not recorded.

The software EstimateS 6 (Colwell 2001) generated estimates of species richness for plant diversity at Acadia (Figure 1) using several different estimators (and functions).

Discussion

The staff of most natural history collections were unable to search records themselves because of the lack of time and resources to fulfill such requests, which understandably places the responsibility of searching upon the organization requesting the data. To conduct efficient manual searches, we offer several recommendations. Preparation is the key. Knowledge of the historic names for the localities for which you are searching will be helpful in identifying relevant specimens. In addition, lists of potential species for a region can help narrow the search field, although care must be taken not to exclude rare, extinct, and vagrant species.

Efficiency in searching is also important. We suggest searching specimen tags if the collection is divided by locality. In most large collections, specimens were divided regionally into separate folders (for plants) or trays (for vertebrates). Although size alone can make the largest collections overwhelming, they often were the easiest to search because they possessed enough specimens to be divided into smaller discrete geographic regions. Smaller collections tended to be divided into local specimens, the rest of North America, and foreign specimens, thus requiring searching most, if not all, specimens. Searching specimen tags can be tedious, but has the advantage of having updated taxonomy and the assurance that specimens are still in the collection. Tags are often very difficult to read, particularly for vertebrate specimens with small tags and old writing. Additionally, handling specimens

Table 2. The number of specimen records received in each proximity category for all parks.

Park code	Number of specimen records[1]				Total (%)[2]
	Category 1	Category 2	Category 3	Category 4	
ACAD	3,392	1,223	7,739	149	12,503 (40.6)
MABI	1	199	273	20	493 (1.6)
MIMA	72	408	1,797	78	2,355 (7.6)
MORR	0	119	905	46	1070 (3.5)
ROVA	237	4	251	485	977 (3.2)
SAGA	0	10	102	19	131 (0.4)
SAIR	0	17	722	0	739 (2.4)
SARA	180	6	115	423	724 (2.3)
WEFA	12	15	983	8	1,018 (3.3)
ASIS	471	1	197	3	672 (2.2)
CACO	186	1,994	1,806	6	3,992 (12.9)
FIIS	109	276	4,026	0	4,411 (14.3)
GATE	30	277	1,107	75	1,489 (4.8)
SAHI	55	3	201	0	259 (0.8)
Total (%)	4,745 (15.4)	4,552 (14.8)	20,224 (65.6)	1,312 (4.3)	30,833

[1] Category 1 = within park boundaries, 2 = may be within park boundaries, 3 = in county, 4 = in state.
[2] Totals are reduced by 277 specimens (0.89%), because we were unable to identify current locality based on a historic place name, there were discrepancies in the locality data, or they could not be assigned to any one park.

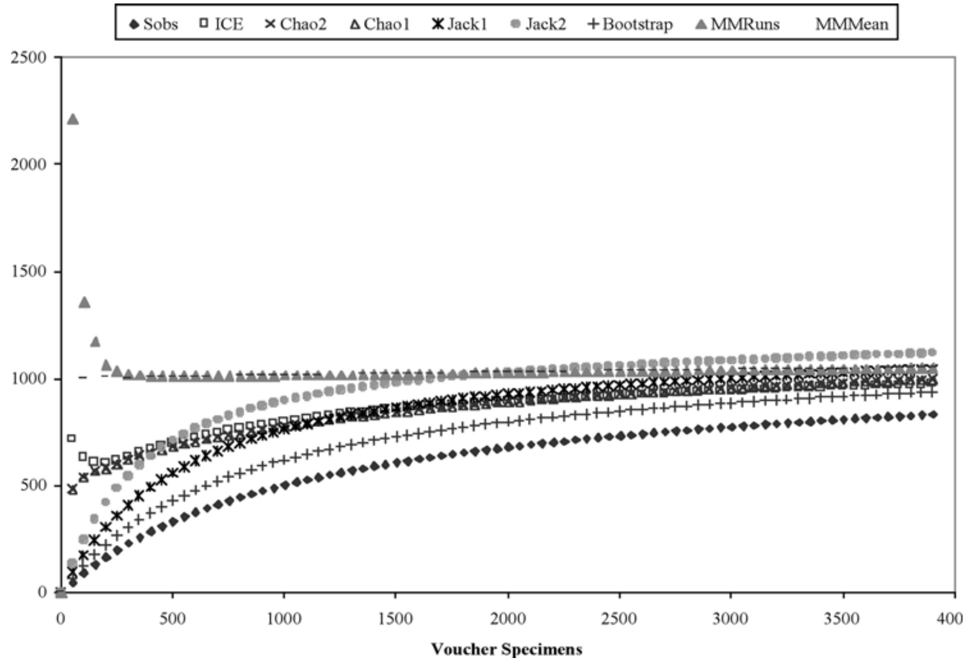

Figure 1. Plant species diversity for Acadia National Park plotting the number of voucher specimens identified (category 1 & 2) using the program EstimateS 6. Results are based on actual observations (Sobs) and eight numerical estimators. A detailed description of individual estimators is in Colwell 2001.

degrades them and may be irritating to the searcher because of harsh chemicals (i.e., arsenic) that may have been used for their preservation. Searching by catalogue is much faster, but data are less reliable and taxonomic updates are not usually made to catalogue entries. If time permits, we recommend searching catalogues, then checking and referencing those records against specimens in the collection. Ultimately, every collection is managed differently, which will affect the search strategy. Flexibility in search strategy is important for determining the best method to search for specimens at a particular site.

Natural history specimens originating from NPS lands acquired prior to 1984, and stored in a non-NPS repository, typically are not catalogued in the NPS national catalogue and are not tracked (i.e., on loan from NPS to non-NPS institutions). A 1984 regulation requires that specimens collected in parks and permanently retained in collections (even in non-NPS facilities) be catalogued into the NPS national catalogue. Although most NPS natural history catalogue records are now recorded in the NPS Automated National Catalog System (ANCS+), catalogue records created prior to 1987, when the automated system began, continue to be input into ANCS+, a project the NPS expects to complete within the next two years. Most collections we searched were in non-NPS repositories in an attempt to locate records unknown to the NPS. In a few cases, we also searched institutions considered NPS repositories as part of our overall effort to compile all available specimen records. For example, the College of the Atlantic in Maine maintains the herbarium collection from Acadia National Park and serves as an official NPS repository and, as such, has records recorded both locally and in ANCS+. The herbarium is on loan from the park.

Natural history specimens originating from NPS lands and acquired prior to 1984 typically are not catalogued in the NPS national catalogue and are not tracked (e.g., those on loan from NPS to non-NPS institutions). A 1984 regulation requires that specimens collected in parks and permanently retained in collections (even in non-NPS facilities) be catalogued into the NPS national catalogue. Although most NPS natural history catalogue records are now recorded in the ANCS+, data acquired prior to 1984 continue to be input into ANCS+, a project that NPS expects to complete within the next two years. Most collections we searched were in non-NPS repositories in an attempt to locate records unknown to NPS. In a few cases, we also searched institutions considered to be NPS repositories as part of our overall effort to compile all available specimen records. For example, the College of the Atlantic in Maine maintains the herbarium collection from Acadia National Park and serves as an official NPS repository and, as such, has records recorded both locally and in ANCS+.

Given the volume of information, we did not error-check data, assuming correct identification of specimens with accurate supporting information. Given that taxonomic revisions occur frequently, verification of identity may be necessary. Furthermore, data such as locality or date can lack specific information or be missing, particularly for older specimens. Locality names can change over time and historic names need to be checked to ensure compatibility between the past and current locations. Despite these limitations, these data are useful as a set of tools for exploring changes in biodiversity, especially when records date back over a century or more.

Estimation of species richness has become an important topic in community ecology and monitoring (Cam et al. 2002) and is an important component of evaluating biodiversity (Colwell and Coddington 1994). Species accumulation curves (Soberon and Llorente 1993) have been used to estimate species diversity, but the use of phenomenological models to plot species accumulation data has been criticized because there is no mechanistic basis to correct for sampling effort (Fagan and Kareiva 1997; Cam et al 2002). The EstimateS 6 program is vulnerable to the criticisms posed above. However, plotting the relationship of the number of voucher specimen records against the number of species identified in these records can be a useful exploratory tool to view the "thoroughness" of sampling conducted in an area and compare sampling across regions (Fagan and Kareiva 1997) or, in this case, park units.

For parks such as Acadia with intensive sampling over several decades, voucher specimens records may provide species richness estimates that are nearly asymptotic for true species diversity. Recent statistical procedures, such as the information-theoretic approach (Burnham and Andersen 1998), can provide further objectivity in selecting a particular estimator (and function) to determine the accuracy of species accumulation data, assuming a reasonable a priori model set (Cam et al. 2002). Additionally, techniques for estimating species richness that are preferred over the function-fitting approach employed by EstimateS 6 (Cam et al. 2002) can also be used. The lognormal distribution of species abundances (Fagan and Kareiva 1997), models of detection probability (Cam et al. 2002), and others based on capture–recapture theory are preferred by some authors (see Nichols and Conroy 1996; Boulinier et al. 1998). These models can estimate the size of species assemblages—an important consideration in the design of biological inventories and monitoring programs. We recommend further exploration of how to use these techniques with voucher specimen data so that species inventory results can be objectively evaluated in the context of temporal change in species diversity.

References

Alberch, P., 1993. Museums, collections, and biodiversity inventories. *Trends in Ecology and Evolution* 8, 372–375.

Boulinier, T., J.D. Nichols, J.R. Sauer, J.E. Hines, and K.H. Pollock. 1998. Estimating species richness: the importance of heterogeneity in species detectability. *Ecology* 79, 1018–1028.

Burnham, K.P., and D.R. Andersen. 1998.

Model Selection and Inference: A Practical Information-Theoretic Approach. Berlin: Springer-Verlag.

Cam, E., J.D. Nichols, J.R. Sauer, and J.E. Hines. 2002. On the estimation of species richness based on the accumulation of previously unrecorded species. *Ecography* 25, 102–108.

Colwell, R.K. 2001. EstimateS: Statistical estimation of species richness and shared species from samples. Version 6.0b1. User's guide and application published on-line at http://viceroy.eeb.uconn.edu/estimates.

Colwell, R.K., and J.A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society London (Series B)* 345, 101–118.

Cotterill, F.P.D. 1995. Systematics, biological knowledge, and environmental conservation. *Biodiversity and Conservation* 4, 183–205.

Fagan, W.F., and P.M. Kareiva. 1997. Using compiled species lists to make biodiversity comparisons among regions: a test case using Oregon butterflies. *Biological Conservation* 80, 249–259.

Nichols, J.D., and M.J. Conroy. 1996. Estimation of species richness. In *Measuring and Monitoring Biological Diversity: Standard Methods for Mammals.* D.E. Wilson, F. R. Cole, J.D. Nichols, R. Rudran, and M.S. Foster, eds. Washington, D.C.: Smithsonian Institution Press, 226–234.

Soberon, J.M., and J.B. Llorente. 1993. The use of species accumulation functions for the prediction of species richness. *Conservation Biology* 7, 480–488.