



Elevational gradient analyses and the use of historical museum specimens: a cautionary tale

Rebecca J. Rowe*

Committee on Evolutionary Biology, University of Chicago, Chicago, IL, USA and Division of Mammals, The Field Museum, Chicago, IL, USA

ABSTRACT

Aim The value of biodiversity informatics rests upon the capacity to assess data quality. Yet as these methods have developed, investigating the quality of the underlying specimen data has largely been neglected. Using an exceptionally large, densely sampled specimen data set for non-flying small mammals of Utah, I evaluate measures of uncertainty associated with georeferenced localities and illustrate the implications of uncritical incorporation of data in the analysis of patterns of species richness and species range overlap along elevational gradients.

Location Utah, USA, with emphasis on the Uinta Mountains.

Methods Employing georeferenced specimen data from the Mammal Networked Information System (MaNIS), I converted estimates of areal uncertainty into elevational uncertainty using a geographic information system (GIS). Examining patterns in both areal and elevational uncertainty measures, I develop criteria for including localities in analyses along elevational gradients. Using the Uinta Mountains as a test case, I then examine patterns in species richness and species range overlap along an elevational gradient, with and without accounting for data quality.

Results Using a GIS, I provide a framework for *post-hoc* 3-dimensional georeferencing and demonstrate collector-recorded elevations as a valuable technique for detecting potential errors in georeferencing. The criteria established for evaluating data quality when analysing patterns of species richness and species range overlap in the Uinta Mountains test case reduced the number of localities by 44% and the number of associated specimens by 22%. Decreasing the sample size in this manner resulted in the subsequent removal of one species from the analysis. With and without accounting for data quality, the pattern of species richness along the elevational gradient was hump-shaped with a peak in richness at about mid-elevation, between 2300 and 2600 m. In contrast, the frequencies of different pair-wise patterns of elevational range overlap among species differed significantly when data quality was and was not accounted for.

Main conclusions These results indicate that failing to assess spatial error in data quality did not alter the shape of the observed pattern in species richness along the elevational gradient nor the pattern of species' first and last elevational occurrences. However, it did yield misleading estimates of species richness and community composition within a given elevational interval, as well as patterns of elevational range overlap among species. Patterns of range overlap among species are often used to infer processes underlying species distributions, suggesting that failure to account for data quality may alter interpretations of process as well as perceived patterns of distribution. These results illustrate that evaluating the quality of the underlying specimen data is a necessary component of analyses incorporating biodiversity informatics.

*Correspondence: Rebecca J. Rowe, Division of Mammals, The Field Museum, 1400 South Lake Shore Drive, Chicago, IL 60605, USA.
E-mail: browe@uchicago.edu

Keywords

Biodiversity, biogeography, bioinformatics, data quality, digital elevation model, elevational gradient, GIS, MaNIS, small mammals, Utah.

INTRODUCTION

Identifying patterns of biodiversity is fundamental to increasing our understanding of ecological and evolutionary processes and informing conservation and management decisions. Specimen identification and locality information are the basic units of biodiversity; regional or larger-scale data on biodiversity derive principally from museum or herbarium specimens and associated data bases. Indeed, natural history collections provide the bulk of information about what we know concerning many aspects of biological diversity. As a result, recent emphasis has been placed on updating and improving access to natural history collection data, and increasing collaborations across museums, research institutions, and universities (Bisby, 2000; Kaiser, 2000; Krishtalka & Humphrey, 2000; Sugden & Pennisi, 2000; Murphy *et al.*, 2004; Stein & Wicczorek, 2004).

This union of biodiversity data and bioinformatics has improved the accessibility of both historical and contemporary museum specimen data, which has the potential to expedite research across subdisciplines of ecology and evolutionary biology (Kress *et al.*, 2001; Canhos *et al.*, 2004; Causey *et al.*, 2004; Edwards, 2004; Wheeler *et al.*, 2004). Improved access to accurate specimen locality data promotes research across spatial scales and at greater spatial extents (e.g. Gaston & Blackburn, 2000), and may also increase sampling resolution by providing additional specimens for analysis (e.g. Bieler & Mikkelsen, 2004; Godoy *et al.*, 2004). Moreover, integrating historical records with contemporary data can provide evidence of population change (e.g. Shaffer *et al.*, 1998) and distributional shifts (e.g. Parmesan *et al.*, 1999) in response to changing climate conditions or anthropogenic causes, and may facilitate better predictive modelling of species distributions or invasions (e.g. Davies *et al.*, 1999; Peterson, 2001; Raxworthy *et al.*, 2003).

With increased development of, and greater access to biodiversity data bases, it is necessary to step back and evaluate data quality and to determine the sufficiency of the data for addressing biodiversity research questions. This is particularly important with increasing application of meta-analyses to conservation issues and the growing urgency for decision-making in biodiversity conservation. Challenges inherent in using data from natural history collections have long been identified, the most common of which are taxonomic inaccuracies, and incomplete data, in particular artefactual absences. While there has been much debate on how to address taxonomic concerns (e.g. Ruedas *et al.*, 2000), and how to assess data completeness and properly analyse presence/

absence or abundance data gleaned from museums and data bases (Grayson & Livingston, 1993; Kodric-Brown & Brown, 1993; Ponder *et al.*, 2001; Anderson, 2003; Skelly *et al.*, 2003; Solow & Roberts, 2003), quantifying the potential magnitude of errors or uncertainty in spatial metadata has been neglected. Only recently have spatial errors in specimen data been included in the discussion of data quality and data sufficiency (Knyazhniyskiy *et al.*, 2000; Bachman *et al.*, 2004; Graham *et al.*, 2004; Murphy *et al.*, 2004; Soberon & Peterson, 2004; Wicczorek *et al.*, 2004).

Spatial error accompanies the documentation of every collecting event, but its significance is only apparent in the context of a particular analysis. Specifically, the scale of a given analysis determines the needed precision of localities within a sampling domain. Historically, most collecting events or localities were recorded as textual descriptors and were not accompanied by collector-assigned coordinates. Even when coordinates were recorded, the source from which the localities were derived and associated levels of precision (map scale, datum, GPS) are often not preserved in the collection data base (Wicczorek *et al.*, 2004). Here, I use spatial error to refer to the imprecision of the original recorded locality, whether a textual description or set of coordinates. Because the majority of museum and herbarium specimens have inexact locality descriptors, incorporating specimen data in historical or spatial analyses is often facilitated by retrospective georeferencing. Georeferencing assigns coordinates and estimates of uncertainty to locality descriptors *post-hoc*. For georeferenced localities, spatial error is expanded to include the uncertainty estimate associated with the assigned coordinates and errors in the georeferencing process.

Data base initiatives incorporating a standardized, collaborative georeferencing component, such as the Mammal Networked Information System (MaNIS, 2001), increase the functionality of historical and contemporary specimens and their associated collection information. These networked data permit the incorporation of historical information and/or expansion of the spatial extent, sampling density, or resolution of a study. In addition, georeferencing permits the integration of specimen data in a geographic information system (GIS) for mapping or analysis.

In this paper, I examine patterns of uncertainty associated with georeferenced coordinates from museum collections with and without collector-documented coordinates. Specifically, I treat two aspects of data quality: precision of areal location recorded at specific collecting events, and precision of estimates in elevational error derived from a digital elevation model. I demonstrate significant consequences that may result from

uncritical incorporation of georeferenced localities when estimating patterns of species richness and species range overlap along elevational gradients. Elevational gradient analyses are frequently employed as a way of evaluating factors structuring species distributions (e.g. Fleishman *et al.*, 2000; Heaney, 2001; Rickart, 2001; Vetaas & Grytnes, 2002; Bhattarai *et al.*, 2004; McCain, 2004), rendering a robust analysis of the consequences of incorporating data of unknown quality highly pertinent.

As an example, I use MaNIS georeferenced data on non-flying small mammals from Utah. This system is an empirical 'best-case scenario' as it is one of the most well-sampled regions for small mammals world-wide, providing a high number of spatially dispersed localities for assessment (e.g. Durrant, 1952; Rickart, 2001). In evaluating uncertainty, I emphasize the need for 3-dimensional error estimates and provide a framework for *post-hoc* 3-dimensional georeferencing that can also serve as a tool for detecting georeferencing errors.

METHODS

Measures of uncertainty

MaNIS is a consortium of 17 North American research institutions and museums developing a network of distributed data bases of mammal specimen locality data (Stein & Wieczorek, 2004). One objective of MaNIS is collaboratively to georeference localities at which museum-based mammal specimens were collected world-wide. In this study I use the georeferenced coordinates and associated uncertainty estimates assigned by MaNIS to localities in Utah.

To standardize georeferencing across regions and specimens, and to promote consistent and reproducible results, MaNIS researchers employed the point-radius method for assigning latitude-longitude coordinates and uncertainty (Wieczorek *et al.*, 2004). This method considers seven sources of uncertainty in locality descriptions: locality extent, GPS accuracy, unknown datum, map scale, and imprecision in distance, coordinate and direction measurements. Ultimately, coordinates are assigned and the uncertainty estimate for each locality is summarized as a single value, the maximum error distance. This distance describes the radius of a circle extending from the assigned coordinates, within which the locality lies (Wieczorek *et al.*, 2004). In accounting for each source of uncertainty, no recording conventions were assumed, resulting in conservative estimates. Most notably, conservative estimates are obtained when establishing the extent of a named place from which distance and direction are recorded in the locality descriptor. As the area defined by a place name often increases over time, extents will be overestimated for many historical localities because they will be based on modern maps (Wieczorek *et al.*, 2004). In addition, extent measurements typically record the greatest extent for a given place.

Georeferencing by MaNIS was based upon textual locality descriptors and was 2-dimensional. Collector-derived elevations were not incorporated, even if they were available,

because the source from which elevation was recorded and the precision with which it was recorded are rarely known. In instances where this information was documented, it was often preserved in collectors' field notes rather than in museum data bases. As a result, ascertaining the error in recorded elevations, when possible, requires access to data sources beyond the primary data base. For these reasons, researchers associated with MaNIS chose not to incorporate elevation in the initial georeferencing process. Since georeferencing by MaNIS was 2-dimensional, I use maximum error distance to refer to areal maximum error distance, accounting for uncertainty along the horizontal plane, which is separate from maximum error in elevation. Using the areal maximum error distances provided by MaNIS, I computed maximum error in elevation using a geographic information system (GIS).

I conducted *post-hoc* 3-dimensional georeferencing as follows. I developed a seamless digital elevation model (DEM) of Utah at maximum 30 m resolution in ArcInfo Workstation (ESRI, Redlands, CA, USA) from 10 and 30 m grids purchased through GeoCommunity. Formatted as decimal degrees, MaNIS localities were plotted on the DEM of Utah, and elevation at each locality extracted from the DEM using 3D Analyst (ESRI). Estimates of maximum error distance output by the georeferencing calculator maintained the same units as in the original textual locality descriptor (Wieczorek *et al.*, 2004). Not all maximum error distances were recorded in the same units, potentially introducing error into the measurements. The majority were recorded in miles. I converted all estimates of maximum error distance to the metric scale. Error associated with this conversion is 2 m, far less than the resolution of the DEM, and therefore contributes insignificantly to the overall error. Associated error in elevation was obtained by running a batch process through an Arc Macro Language (AML) script written for ArcInfo Workstation and later converted to the Python programming language for ArcGIS v. 9.0 (ESRI). Taking one locality at a time, a buffer was created in 3-dimensional space based on the maximum error distance. The minimum and maximum elevation in each 3-dimensional buffer were computed to the nearest integer. Points with a maximum error distance less than the grid cell size of the DEM, 30 m, received values equivalent to the elevation for the input coordinates. Estimated errors in elevation are as conservative as the maximum error distances on which they were based. Maximum error distances are reported in kilometres and error in elevation reported in metres.

Localities

Localities incorporated in this study were limited to those associated with the non-flying small mammals (shrews, lagomorphs, and rodents) of the Rocky Mountain and Colorado Plateau regions, Utah (Fig. 1a,c). All localities with a specific textual or coordinate locality descriptor were considered, with the exception of textual descriptors recorded only to the county level. However, localities with specific descriptors that were too vague or could not be located resulting in the

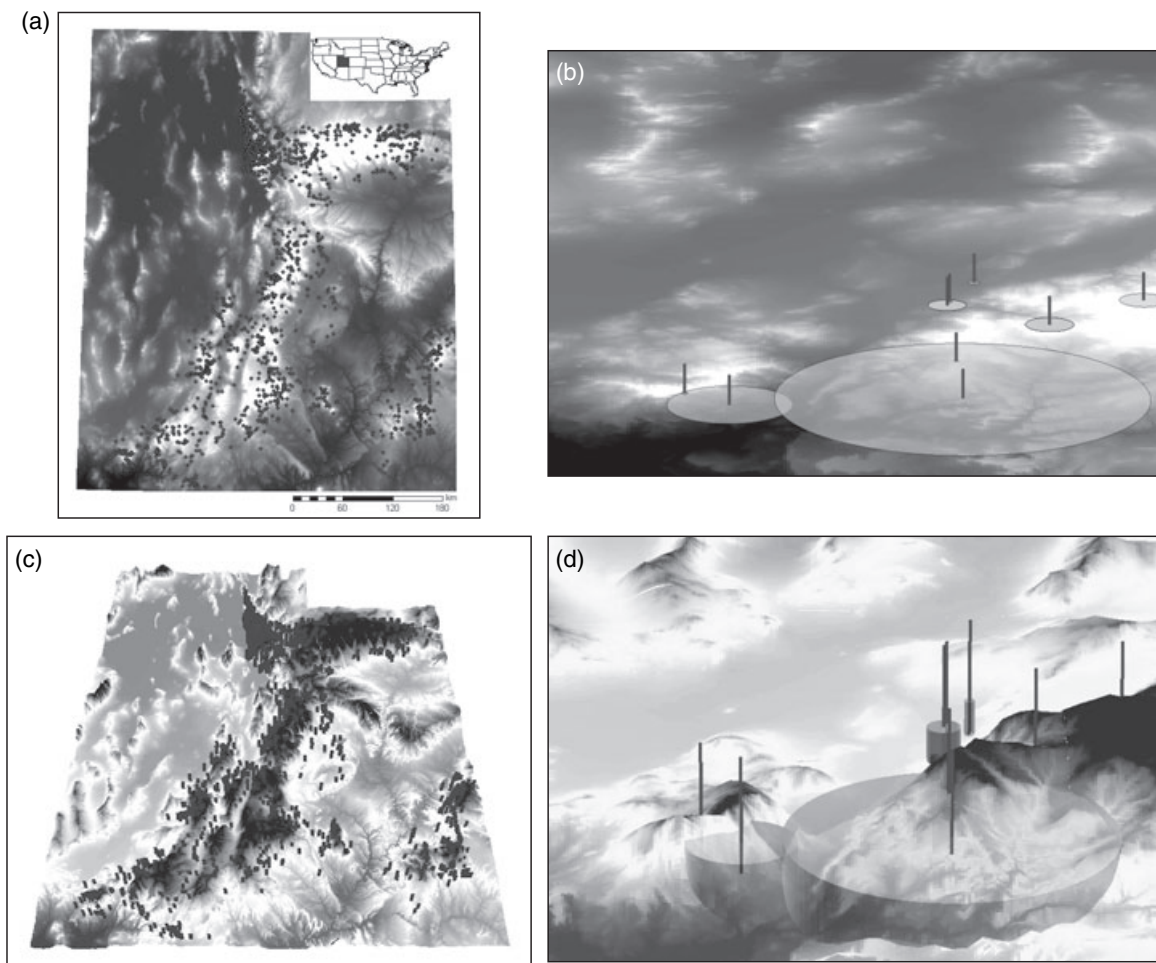


Figure 1 A 2-dimensional (a) and 3-dimensional (c) representation of Utah depicting localities georeferenced by MaNIS that fall within the study area. The maximum error distances for a subset of localities are also illustrated in both 2 (b) and 3-dimensions (d). Lines indicate localities. Circles represent maximum error distances in area for localities in b; cylinders represent maximum error in both area and elevation for the same set of localities in d.

georeferencing of that point to the county level were included in these analyses as they represent real variation in locality description quality. Overall, 3144 Utah localities georeferenced by MaNIS fit the criteria. Each locality represents a distinct locality descriptor taken from an affiliated museum data base. In order to identify all localities to be georeferenced without applying any *a priori* notions of similarity among localities, MaNIS isolated each locality descriptor varying in any manner from other descriptors, including spelling and punctuation. The voucher specimens associated with these localities are housed at 12 museums in North America. Appendix 1 is a list of the number of these voucher specimens housed at each of the contributing museums.

Different filters were applied to the 3144 localities to produce the appropriate unique combinations of localities for each subsequent analysis (Table 1). Of the 3144 localities included in this study, 1903 (61%) unique combinations of latitude–longitude coordinates and uncertainty are represented (e.g. 37.05755° N, 112.60516° W, 4.841 km & 37.05755° N, 112.60516° W, 5.420 km). Locality duplication commonly

results from syntax errors when locality descriptors are entered by hand into a given museum data base. Therefore, isolating unique series of latitude, longitude, and maximum error distance removes site duplication as a result of syntax errors while maintaining sites georeferenced to the same coordinates based on locality descriptors of differing precision. In addition, this approach removes site duplication in instances where multiple MaNIS identification numbers were attributed to identical locality descriptors because specimens collected at those localities were housed in more than one museum. Sixty-four instances (0.05%) of this type of site duplication were present in this data set.

The 1903 unique coordinate–uncertainty pairs were used for examining the distribution of maximum error distances across localities, and the relationship between maximum error distance and associated error in elevation for a given point. Year of collection was associated with specimens at 3116 of the 3144 localities. Of those 3116 localities, 2862 (92%) were sampled during only 1 year and 254 (8%) across multiple years. Of these 3116 localities, 1889 unique coordinate–

Table 1 For each analysis, the 3144 localities georeferenced by MaNIS were filtered to provide the appropriate set of unique factors. This table outlines the variables used as filters for each analysis and the total number of resulting unique combinations

Analysis	Variables				Number of unique combinations
	Coordinates	Uncertainty	Year	Collector's elevation	
1. Distribution of maximum error distance	X	X			1903
2. Maximum error distance vs. error in elevation	X	X			1903
3. Temporal distribution of maximum error distance	X	X	X		2620*
4. DEM elevation vs. collector elevation	X			X	1984

*Localities (coordinate–uncertainty pairs) sampled across multiple years were included in each relevant time interval.

uncertainty pairs are represented. Localities sampled during multiple years were included in all relevant time intervals, resulting in the analysis of 2620 unique coordinate, uncertainty, and year combinations. For 2604 (83%) of the 3144 localities, collector-documented records of elevation were available, although the source from which elevation was taken was rarely recorded. Of these 2604 records, 1984 (76%) represent unique coordinate–elevation pairs. This set of records was used when comparing DEM and collector-derived measures of elevation.

Eighty-nine recently surveyed (non-MaNIS) localities were included in conjunction with 42 localities georeferenced by MaNIS to validate DEM-derived elevations and to determine the extent at which DEM-derived elevation is an appropriate proxy for collector elevation. All 131 localities were georeferenced in the field by the same set of collectors and elevation recorded from either a handheld GPS receiver or 1:24,000 USGS map. In each case, the source from which elevation was recorded is known. Maximum error estimates were computed following the MaNIS georeferencing guidelines (Wieczorek, 2001).

Elevational gradient analyses

To determine whether the uncritical incorporation of georeferenced data affects observed patterns of species richness and species range overlap along elevational gradients, I chose the Uinta Mountains as a test case and used the georeferenced localities associated with collecting events that occurred between 1960 and 1969. Only one time interval was chosen to avoid possible complications associated with species distributions shifting over time in response to climate change or changes in land-use practices that may also affect the structure of biodiversity patterns. Using the data quality criteria outlined below, a subset of these localities was extracted. Observed patterns of species richness and species range overlap were then compared across the complete Uinta Mountains data set and this subset.

All localities whose coordinates fell within the Uinta Mountains and which were surveyed from 1960 to 1969 defined the complete data set. In determining which localities fell within the Uinta Mountains, I used major stream channels

as the lowland boundaries. Stream data for the state of Utah were obtained from the Utah Automated Geographic Reference Center, at scale 1:24,000. Evanston and Firehole quadrangles for Wyoming at 1:100,000 were incorporated and taken from the Wyoming Geographic Information Science Center. The major streams were traced using ArcGIS v 8.3 and converted to polygons.

The test subset of that complete data set was based on the following criteria. (1) Using a GIS, each locality was buffered by 75% of its maximum error distance. Points were included only if, when buffered, they were completely contained within the Uinta Mountains. (2) In addition, one of the following three criteria must have been met: (i) the maximum error distance was less than or equal to 0.805 km, (ii) elevational uncertainty was less than or equal to 100 m, or (iii) the absolute difference between the DEM and collector-derived elevation was less than or equal to 100 m, and the collectors' value fell within the possible range of elevations from the DEM. The maximum error distance of 0.805 km was used because it reflects the greatest uncertainty estimate attributed by MaNIS to a point locality, such as a spring or campground. Error in elevation was limited to 100 m because preliminary analyses along elevational gradients in Utah, which range from 1500 to 3000 m in extent, indicate that this is the required precision of localities within a sampling domain. At bin sizes greater than 100 m, the decrease in the number of bins analysed renders the detection of patterns in range overlap along the gradients problematic at best. In addition, bins of 100 m elevation are becoming common procedure when museum or herbarium specimens are used in analyses of species richness along elevational gradients (e.g. Vetaas, 2000; Fleishman *et al.*, 2001; Nor, 2001; Bachman *et al.*, 2004).

Gradient analysis was limited to the 35 non-flying small mammals that were surveyed along the Uinta Mountains from 1960 to 1969. This complete data set consisted of 1275 specimen records from 158 unique coordinate–uncertainty pairs, of which 996 specimens from 89 unique coordinate–uncertainty pairs were included in the subset. Appendix 2 provides the full scientific names of the species included in this study as well as a list of the number of specimens and unique coordinate–

uncertainty pairs associated with each species in both the complete data set and subset; the taxonomy follows Wilson & Reeder (1993). For both the complete data set and subset, species' upper and lower elevational range margins were summarized from the specimen data for each incorporated locality and used to define species' elevational ranges. Therefore, the distributions of species along the elevational gradient were assumed to be continuous between the highest and lowest documented localities for each species. The observed species richness pattern for each data set was determined by counting the number of species present in each 100-m interval of elevation along the gradient. To compare the shape and mode of species richness patterns derived from the subset and complete data set, I fitted a quadratic equation to each distribution and examined similarity among equations and modes.

Intraspecific variation in elevational range between the complete data set and subset was evaluated to determine whether shifts in species' perceived distributions were differentially biased based on whether they had large or small ranges under the complete data set. For each species, I used the number of 100-m intervals of elevation in which it was present as a proxy for range size. This proxy was necessary to account for species known from only one locality for which no elevational range could be obtained. A Mann–Whitney *U*-test was conducted comparing species' range sizes under the complete data set across the following categorization of range shifts between data sets: (1) no change, both elevational range margins were maintained, or (2) at least one range margin shifted. When present in both data sets, species known from only one locality were classified as unchanged. Because the complete data set and subset differed in species richness, the *U*-test was evaluated twice, including and excluding the one species present in the complete set but absent from the subset. When included, this species was counted in the shifted range category.

To evaluate whether changes in species' elevational ranges across data sets impact species range overlap patterns, pair-wise frequencies of the following three categories of range overlap were computed and compared across data sets: no range overlap, range overlap without shared margins, and range overlap with at least one range margin in common. Statistical analyses were conducted using XLSTAT version 7.5.2 (Addinsoft SARL, Paris, France) and SPSS version 13.0 (SPSS Inc., Chicago, IL, USA).

RESULTS

Patterns of areal uncertainty

Localities are distributed non-uniformly both throughout the study area (Fig. 1) and over time. Sites were surveyed from 1887 to 2001, with the majority of sites (59%) worked during the 1950s and 1960s (Fig. 2). The relatively small number of sites surveyed between 1990 and 2001 may be an underestimate as a result of museums cataloguing specimens collected during this time interval after the locality 'snap-shot' for the MaNIS collaborative georeferencing project that was conducted in mid-2001.

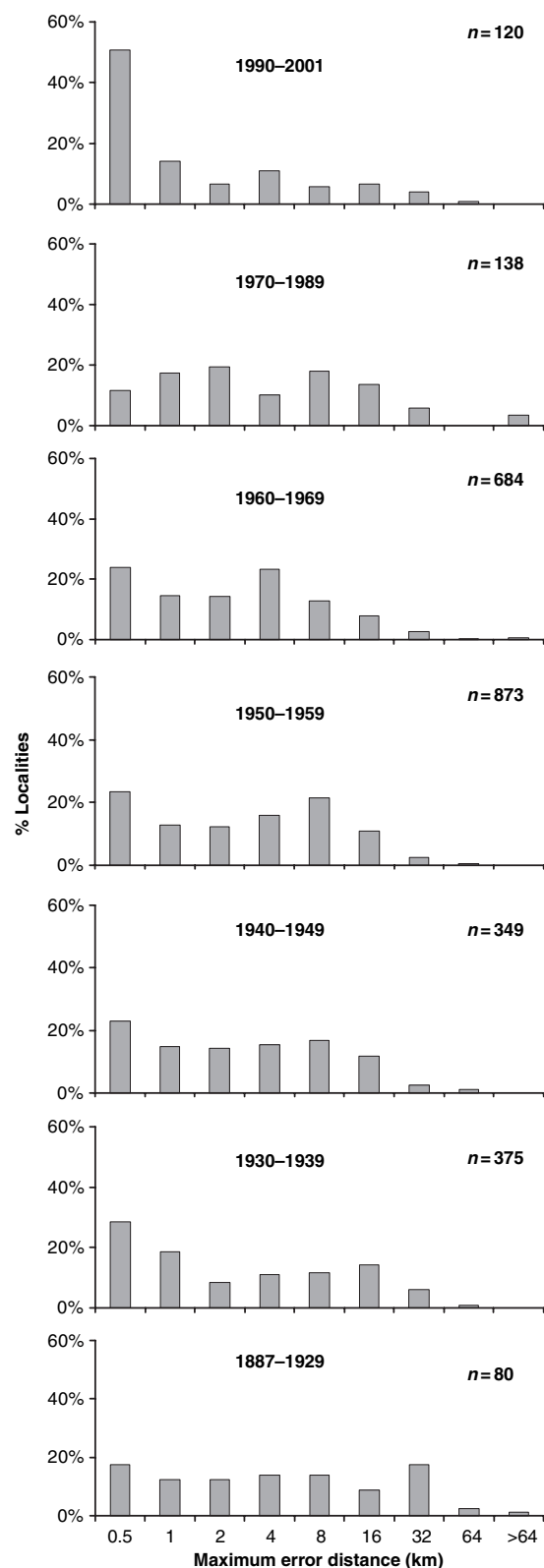


Figure 2 Distribution of the localities georeferenced by MaNIS across bins of maximum error distance for different time intervals. Localities sampled across multiple years were included in each relevant time interval. Total number of localities (*n*) per time interval is listed on each plot.

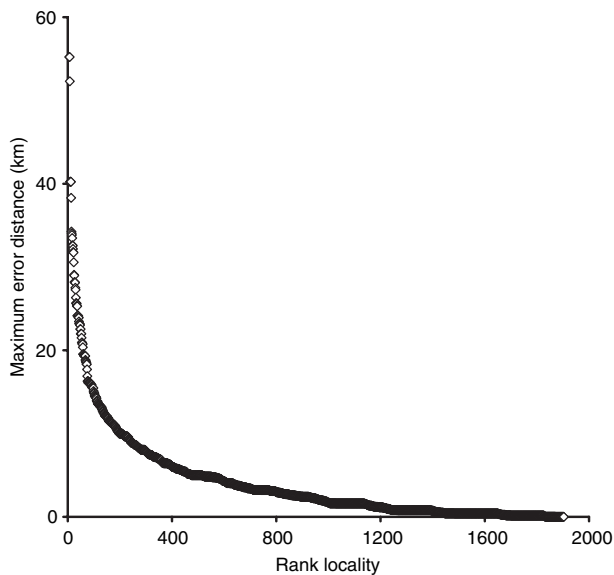


Figure 3 Distribution of maximum error distance in area across localities. The graph was limited to a maximum error distance of 60 km. As a result, four points do not appear on the graph: 72.43, 76.44, 80.48, and 112.7 km. Mean = 4.348 km, mode = 0.4023 km, standard deviation = 7.152 km.

There is considerable variation in maximum error distance. However, there is no spatial bias in data quality, as the magnitude of maximum error distance is not correlated with the elevation, latitude, or longitude of localities ($r^2 = 0.000$, 0.022, and 0.001, respectively). Of the 1903 unique coordinate–uncertainty pairs, more than 95% have maximum error distances less than 16 km, about 50% of all localities have an error less than 2 km, and 637 (33%) were attributed an error equal to or less than 0.805 km (Fig. 3). The greatest maximum error distance is 112.7 km. Very high maximum error distances result from vague locality descriptors, such as named places without distance and direction measurements (e.g. ‘Emery County, Huntington Canyon’), very imprecise distance and direction measurements taken from named places (e.g. ‘Daggett County, 30 mi N Vernal’), or instances where places could not be identified and the localities were georeferenced to the county level. For example, the locality ‘Emery County, Huntington Canyon’ has an error of 24.1 km. The coordinates for this locality were set at the midpoint of the c. 48-km long canyon. In addition, the unique coordinate–uncertainty pairs with maximum error distances between 72.4 and 112.7 km reflect instances where the named place in the locality descriptor could not be located on a map or gazetteer and the locality was therefore georeferenced to the centre of the designated county.

Data quality varies within and among time intervals, showing improvement in only the most recent interval. During each time interval, localities are distributed across at least 8 of the 9 bins of maximum error distance, with absence of localities limited to the two largest error bins, > 32–64 or > 64 km (Fig. 2). In addition, the smallest proportion of localities consistently occupies these two largest error bins,

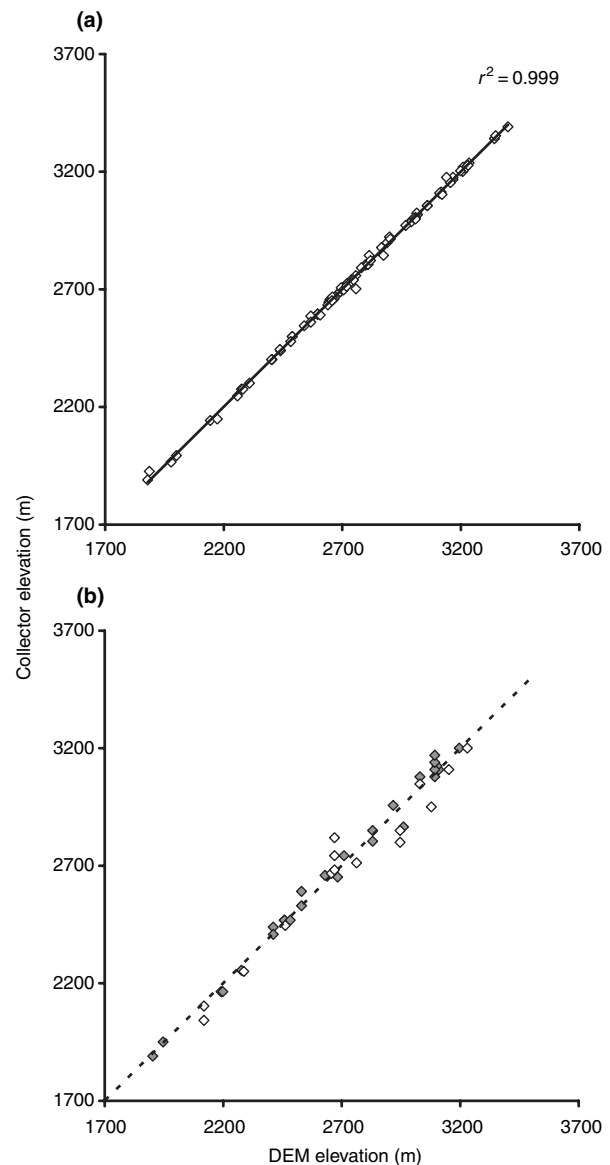


Figure 4 (a) Comparison of collector and DEM-based elevations for 86 localities in this study area surveyed from 2000 to 2004. All localities were georeferenced with a GPS in either decimal degrees (0.00001 precision) or degrees decimal minutes (0.01 precision). Overall precision is less than or equal to 30 m, which is equivalent to the maximum DEM grid cell size. (b) Comparison of collector and DEM based elevations for 42 localities in this study area surveyed from 1990 to 2001. Localities were georeferenced from a 1:24,000 USGS map, datum unknown, and recorded either as degrees minutes seconds (1.0 precision), \blacklozenge or degrees minutes (1.0 precision), \diamond . Dashed line represents equivalence. Maximum error distance < 118 m for degrees minutes seconds and < 2.53 km for degrees minutes.

> 32–64 km and > 64 km, representing at most 5% of the localities individually or 6% collectively. In five of the seven time intervals, the relative frequency of localities is greatest in the smallest error bin, > 0–0.5 km. In two of those five time intervals, the frequency of localities within the smallest error bin is greater than 25%: 29% in 1930–1939, and 51% in 1990–

2001 (Fig. 2). Restricting the analysis of this data set to a particular time period(s), other than the most recent, will not improve overall data quality as estimated by maximum error distance.

Collector and DEM elevation

The DEM employed in these analyses is an appropriate proxy for collector-derived elevations at high resolution. However, it fails consistently to represent collector-derived elevations within the required 100 m of precision when the collector-derived coordinates were recorded at low resolution. The relationship between collector and DEM-derived elevations for the 86 recent localities with less than 30 m estimated maximum error distance is near-perfect with $r^2 = 0.999$ (Fig. 4a). For the 42 localities with maximum error distances of either 118 m or 2.53 km, deviation from a 1:1 relationship between collector and DEM-derived elevations varied from 0 to 149 m (Fig. 4b). Three localities had residual values greater than 100 m, each of which was recorded in degree minutes (1.0 precision, 2.53 km error).

The comparison between collector and DEM-derived elevations for the 1984 unique coordinate-collector elevation pairs in this data set shows substantial deviation from a perfect relationship (Fig. 5a). Absolute values of the residuals range from 0 to 1630 m, as shown in Fig. 5b, of which 821 (41%) are greater than 100 m. In Figs 4b and 5a correlation strength is not provided because the question being addressed is the magnitude of the residual values.

Areal and elevational uncertainty

As estimates of maximum error distance increase, the lower and upper limits of possible error in elevation increase. While the correlation between these two measures of uncertainty is strong, $r^2 = 0.639$, there is a wide range of potential error in elevation for a given maximum error distance (Fig. 6). Variance in elevational error estimates increases with increased maximum error distance. However, as standard error of the mean increases with maximum error distance estimates, the number of localities decreases (Table 2). This comparison between maximum error distance and uncertainty in elevation was limited to the 95% of localities with less than 16 km in estimated maximum error distance.

Across the 1903 localities, elevational error estimates range from 0 to 2753 m, with 51% and 20% of localities containing greater than 400 and 1000 m of estimated error in elevation, respectively. Similar to maximum error distance, the magnitude of error in elevation is not correlated with the elevation, latitude, or longitude of localities ($r^2 = 0.005$, 0.004 and 0.002, respectively).

Elevational gradient analyses

In the Uinta Mountains test case, species richness patterns were qualitatively similar across data sets, both showing a

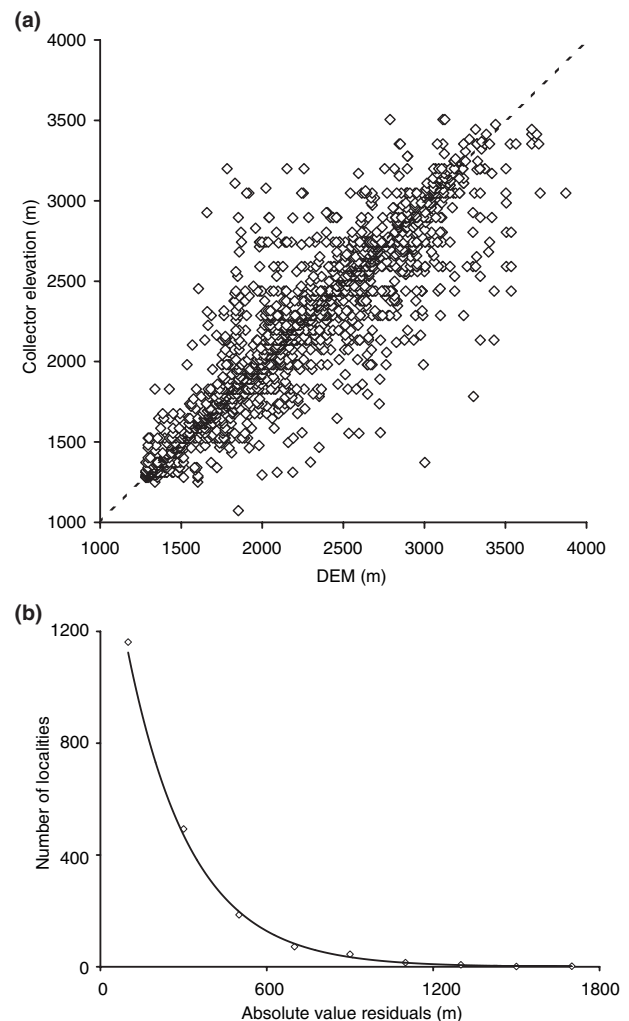


Figure 5 (a) Comparison of collector and DEM-based elevations. Collector elevations were available for 2604 localities of which the 1984 unique coordinate–elevation pairs are represented. Dashed line represents equivalence. (b) Absolute value of residuals plotted against number of localities for the 1984 localities in (Fig. 5a). Residual values range from 1 to 1630 and were grouped in bins of 200 m. More than half of the records, 1163 localities or 59% had residual values less than or equal to 100 m.

hump-shaped pattern with a peak in richness at about mid-elevation between 2300 and 2600 m (Fig. 7a). I fitted a quadratic equation to both distributions, and the resulting equations and modes are similar. For the complete data set, $y = 105.599 - 0.102x - 2.0 \times 10^{-5}x^2$, mode = 2550 with richness = 24.45; for the subset, $y = 124.23 - 0.115x - 2.3 \times 10^{-5}x^2$, mode = 2500 with richness = 19.52. These results indicate that incorporating localities with a high degree of elevational uncertainty does not impact the observed pattern of species richness along an elevational gradient.

Maximum species richness was 26 and 23 in the entire data set and the subset, respectively. Fewer species were recovered in the subset than represented in the entire data set at all elevational intervals (Fig. 7a). The difference in species richness between data sets was most pronounced at the lowest and

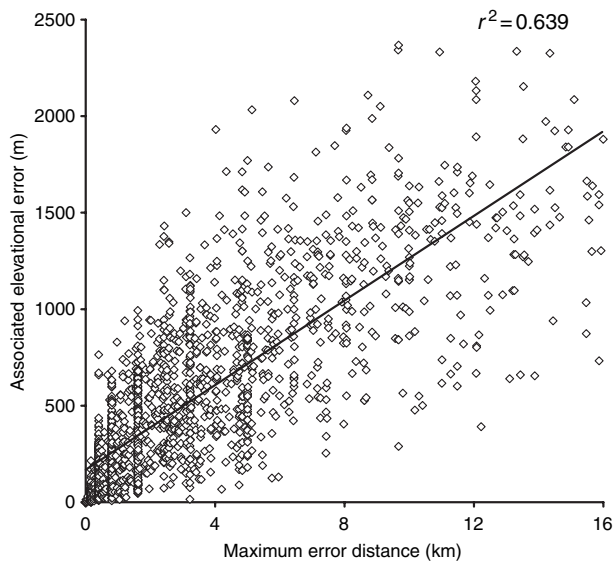


Figure 6 Maximum error distances range up to 112.7 km. 95% of which fall below 16 km of error. This figure portrays the maximum error distance in area and associated error in elevation for these 1817 unique coordinate-maximum error distance pairs.

Table 2 Standard error of the mean for values of error in elevation in bins of 2-km maximum error distance. Limited to the 1817 localities with maximum error distance less than 16 km and displayed in Fig. 6

Max error distance (km)	Number of localities	Standard error of the mean
> 0–2	932	6.76
> 2–4	337	16.15
> 4–6	235	24.77
> 6–8	104	35.61
> 8–10	87	44.62
> 10–12	57	50.76
> 12–14	41	71.95
> 14–16	24	80.39

highest elevations, with no species present in the lowest and two highest intervals in the subset. A greater decrease in richness at the ends of the gradient is expected, as it is the endpoints of species' ranges that were altered across data sets. In addition, it is important to note that because uncertainty estimates are not correlated with elevation, the criteria used in assessing data quality are not biased with respect to elevation.

Although the species pool as a whole was largely maintained across data sets, the majority of species' distributional patterns along the gradient were altered. Overall, 34 of the 35 species in the entire data set were present in the subset (Fig. 7b). The two species known from only one locality in the entire data set remained present in the subset. Of the remaining 32 species, 11 show no change in elevational range between data sets, while 21 species' ranges shifted. Four of those 21 are present at only one locality in the subset: two at their original upper range margin and two at their lower (Fig. 7b). Of the remaining 17

species whose ranges shifted, 5 shifted both range margins and 12 shifted only one range margin. Of those 12 species, the upper range margin shifted in 10 cases (83%) and the lower margin in 2. The tendency for ranges which extended to the highest elevations to shift downward over several hundred metres in the subset likely reflects lower sampling effort at high elevations due to difficulty in reaching sites at those elevations during the 1960s.

Range size, as measured by the number of occupied 100 m intervals in elevation, varied from 1 to 20, out of a maximum possible of 25, for the 35 species in the complete data set. Given this initial variation in extent of the elevational range, a Mann–Whitney *U*-test was conducted to determine whether the truncation of species' ranges from the complete data set to the subset was correlated with elevational range size. The results were significant ($P = 0.029$ and $P = 0.018$, including and excluding the 1 species absent from the subset, respectively) indicating that, on average, elevational range extent was greater for species whose ranges varied between data sets compared to those that did not change. However, the magnitude of the shift in extent is not tightly correlated with the initial elevational range size, $r^2 = 0.234$.

Changes in species elevational distributional patterns can lead to changes in the patterns of species range overlap. In the complete data set there were 94 (16%) pair-wise combinations of non-overlap, 407 (68%) of range overlap (either one range embedded within the other or partial overlap without a shared range margin), and 94 (16%) instances of range overlap including at least one shared range margin. In contrast, the subset showed a significantly different distribution of frequencies across the three categories (Table 3). In the subset the frequency of mismatches or non-overlaps increased (133, 24%), as did the frequency of shared range margins (114, 20%), and the remaining number of range overlaps decreased (314, 56%). The pattern of increased occurrences of shared range margins relative to other patterns of overlap was maintained when rare species, those present in only one interval, were removed from the analysis (Table 3). Although species range overlap patterns varied among data sets, the overall order of species' first and last occurrences along the elevational gradient is maintained across data sets; Kendall's rank correlation $\tau\text{-}b = 0.714$, $P < 0.01$ for first occurrences and 0.867, $P < 0.01$ for last occurrences.

DISCUSSION

Patterns in areal uncertainty

Given (1) the time period over which localities were surveyed, (2) variability in locality recording methods such as named places, offsets from named places, and combinations of the two, (3) different levels of precision with which those named places, distances, and directions were recorded, and (4) potential limitations when entering locality information into collection data bases (e.g. limits on character length which may have resulted in truncated descriptions), it is not surprising to

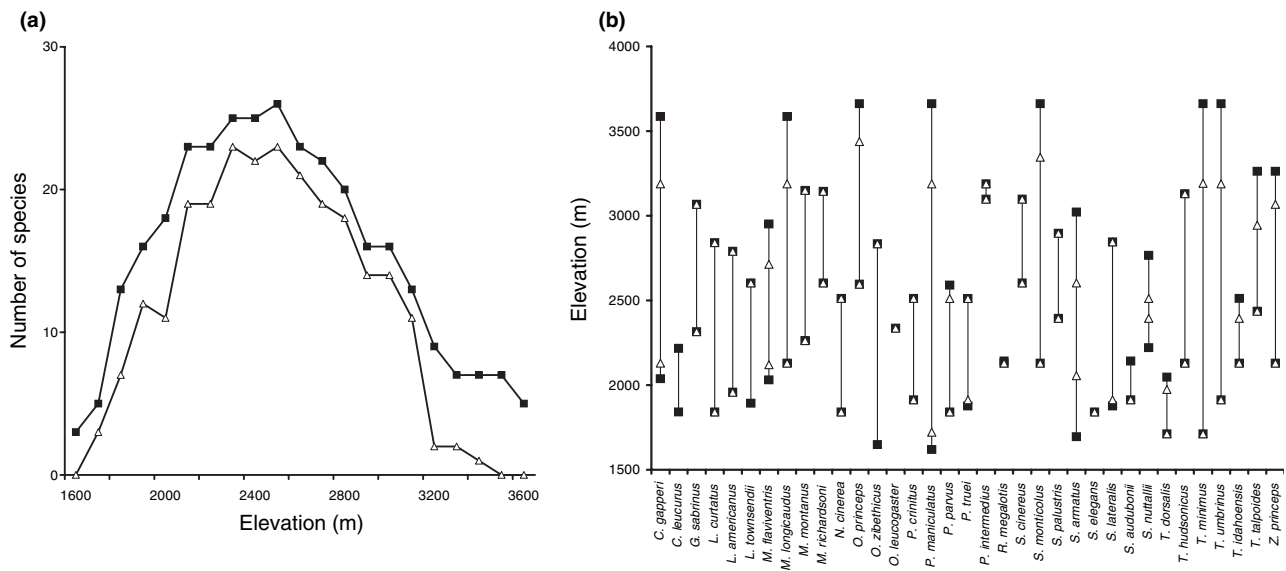


Figure 7 (a) Species richness patterns of non-flying small mammals along the Uinta Mountains 1960 to 1969. Two relationships are depicted, one for all localities georeferenced by MaNIS (■), and the other for a subset of those points representing points of highest quality (△). (b) Elevational range endpoints for 35 species from the two data sets in Fig. 7a. Complete data set (■), subset (△). Full scientific names are provided in Appendix 2, as well as the number of specimens and localities associated with each species across both the complete data set and subset.

Table 3 Frequency of patterns in range-overlap among species pairs under both the complete data set and subset. Significance tested using a contingency table

	Complete	Subset
Including all species [$P < 0.001$]		
No. range overlap	94	133
Range overlap	407	314
Range overlap with shared margin(s)	94	114
Excluding species known from only one interval [$P = 0.015$]		
No range overlap	53	44
Range overlap	364	246
Range overlap with shared margin(s)	79	88

find a wide range of uncertainty measures associated with georeferenced locality data (Fig. 3). This variability underscores the fact that data quality, as estimated by maximum error distance, will be less of a limiting factor in analyses at coarse spatial scales than at fine spatial scales.

What is at first surprising is that locality description quality did not improve over time until very recently (Fig. 2). From 1887 to 1989 there is no evidence of an increase in the frequency of localities with smaller uncertainty measures. The pulse of localities with low maximum error distances in the 1990s is largely attributed to collector-recorded coordinates, whether derived from a map or GPS. When coordinates were provided they were converted to decimal degrees under MaNIS, not georeferenced based on their textual descriptors. As a result, it is likely that this increase in localities with low uncertainty values is

not a result of improved textual descriptors. Perhaps it is unfair to have expected to see improvement over time. After all, the spatial scale at which research questions within this region have been addressed has been variable, and it is reasonable to assume that many collecting events were recorded at a scale corresponding to the question(s) asked. In addition, past perceptions of what constitutes a sufficient locality descriptor may have been quite different from today's as a result of our increased understanding of how map scale and datum affect precision. Even today, with this heightened awareness, collectors are not universally recording those features, and those that have often do so only in their field notes as many collection data bases do not have the resources to include these additional fields and retroactively input the data.

Thinking in 3-dimensional space

The location of a point in space is a 3-dimensional problem and is best illustrated when comparing a 2-dimensional and 3-dimensional perspective of maximum error distance (Fig. 1b,d, respectively) where the extent of elevation within the areal uncertainty is evident. Limiting measures of uncertainty to 2 dimensions is adequate if the research question is at a coarse enough scale that determining horizontal position is the only concern, for example, grouping points within regions. However, finer scale analyses of biodiversity may require horizontal and vertical positional uncertainty measures. This is particularly the case when biodiversity is examined along an elevational gradient, where the placement of a locality along the elevational gradient must be known within a certain degree of precision, typically at most 100 m elevation for analyses

along gradients of 1500 to 3000 m in extent (e.g. Vetaas, 2000; Fleishman *et al.*, 2001; Nor, 2001; Bachman *et al.*, 2004).

Although MaNIS did not georeference in 3 dimensions, this analysis illustrates that elevation values for coordinates and uncertainty in elevation for a given maximum error distance can be obtained *post-hoc* in a GIS. Obtaining estimates for elevation and uncertainty in elevation improve the description of where a locality falls in space and explicitly permits the analysis of uncertainty in 3 dimensions. In addition, despite potential error in collector-recorded elevations, they can be used as an error-detection method for identifying georeferences that require inspection and perhaps reassessment.

Accuracy of a DEM-derived elevation for coordinates depends upon the resolution of the DEM and the precision with which the coordinates were determined initially. DEM resolution affects precision since each grid cell contains only one value of elevation. The DEM used in these analyses is high resolution, with a maximum grid cell size of 30 m. If coordinates are determined within 30 m accuracy then a near-perfect relationship between a DEM and collector-derived elevation is expected, as illustrated in Fig. 4a. These results validate the use of this DEM to derive elevation and elevational uncertainty values for georeferenced localities and interpret those values with reference to the uncertainty associated with the coordinates.

Due to spatial autocorrelation among elevation values for cells in a DEM, disagreement between collector and DEM-derived elevation is expected to increase with the inaccuracy of the coordinates. When coordinate imprecision is increased, deviations from a one-to-one relationship between DEM and collector-recorded elevations are observed (Fig. 4b), with three localities having residual values greater than 100 m in elevation. These results suggest that at low levels of precision (degree minutes, 1.0 precision), DEM-derived elevations represent collector-recorded elevations less accurately, indicating that DEM-derived elevations are not always appropriate proxies for collector-recorded elevations. What is appropriate will depend on the scale of analysis. As mentioned previously, resolution of at least 100 m is often necessary for elevational gradient analyses to ensure appropriate resolution for pattern detection. When a collector has properly recorded elevation using a map, altimeter, or GPS unit and the original coordinate precision is known and is high, then the collector's original elevation can be used in place of the DEM-derived elevation. However, determining the precision of the original measurement can be difficult. Museum data bases may incorrectly reflect that precision by requiring values in all degree, minute, and second fields or by automatically converting coordinates to decimal degrees without a verbatim record of the collectors' coordinates. As a result, greater agreement between collector and DEM-derived elevation may at times be obtained when localities are georeferenced based on their textual descriptors rather than using the collector-assigned coordinates.

When DEM and collector-recorded elevations are compared for this data set, residual values range from 0 to 1630 m in elevation, with 1163 (59%) of the 1984 localities having

residual values less than or equal to 100 m (Fig. 5a,b). Substantial disagreement between these two estimates most likely occurs for two reasons: (1) the collector 'guesstimated', incorrectly recording the elevation based either on vegetation or a general understanding of where they perceived themselves to be on the landscape, or (2) there was considerable error in georeferencing the locality. Less severe deviations probably result from collectors' rounding elevations (e.g. to the nearest hundred feet) independently or in conjunction with expected deviations resulting from coordinate imprecision.

Both collector and DEM-derived methods of defining elevation may produce inaccuracies. As a result, previous authors (Bachman *et al.*, 2004) have used both methods when examining patterns of species richness along an elevational gradient. Despite relatively poor DEM resolution (c. 1 km), they found that both techniques of measuring elevation produced broadly similar patterns of species richness. As discussed earlier, assessing the accuracy of collectors' elevations is problematic due to difficulties in identifying the source (if any) from which a collector recorded elevation and the precision with which elevation was recorded. In contrast, this analysis indicates that assessing the accuracy of DEM-derived elevations can be achieved. Using a high resolution DEM, areal maximum error distances can be accurately converted to elevational uncertainties in a GIS. This approach provides two partially correlated measures of uncertainty from which the user can evaluate data quality and define the appropriate subset of data for biodiversity analyses at finer spatial scales. Since the localities incorporated in this analysis are all from a mountainous region, variability between the maximum error distance associated with a locality and the elevational relief encompassed within that area is expected and is observed (Fig. 6). Therefore, in this system, evaluating data quality based purely on maximum error distances would be inappropriate as sites with relatively large error distances may have narrow elevational uncertainties and vice versa.

Although evaluating the accuracy of collector-recorded elevations can be problematic, comparing collector and DEM-derived elevations can be a valuable error-detection technique for identifying georeferenced coordinates that may be questionable and need to be reassessed. Due to the myriad combinations of errors resulting in variation between DEM and collector elevations, deviations of different magnitude are likely to occur across different levels of coordinate uncertainty. As a result, not all records with high deviations will have been georeferenced incorrectly. For example, a named place that is identifiable on a map and small in extent, such as 'Garfield County, Bown's Reservoir' georeferenced with 0.576 km of uncertainty, has a DEM elevation of 2259 m and a collector-recorded elevation of 6000 ft or approximately 1829 m, resulting in a 430 m difference between the two elevation measures. In this case, where field notes confirm that the correct reservoir was georeferenced, it is fair to assume that the collector guessed the elevation. In this instance the georeference is correct and the elevation associated with the specimen is misleading. Developing error-detection modules is an

important component of biodiversity informatics as it can improve the overall quality of the resource as well as identifying misinformation associated with the specimens themselves, which is of great benefit to participating museums and institutions (Graham *et al.*, 2004; Wieczorek *et al.*, 2004). Close agreement can exist between collector and DEM elevations as a spurious result when a locality was incorrectly georeferenced but at an elevation matching the collector's record. For such reasons, comparing collector and DEM elevations should be used in conjunction with other error-detection modules.

Implications for elevational gradient analyses

The Uinta Mountains test case illustrates the impact that failure to account for data quality can have on analysis of patterns of biodiversity along elevational gradients. The overall pattern of species richness is similar across data sets with respect to the elevation at which richness peaks and the hump-shaped relationship (Fig. 7a, see results for quadratic equations and modes). Commonality in the observed pattern of species richness across data sets can largely be attributed to the lack of correlation between the extent of uncertainty (both maximum error distance and elevational error) and the elevation, latitude, or longitude of localities. In addition, the order in which species appear and disappear along the elevational gradient is similar across data sets ($P < 0.01$ for both first and last elevational occurrences).

Although accounting for data quality did not affect the observed trend in species richness along the elevational gradient, it did affect the absolute numbers. The number of species present within each elevational interval and at the greatest peak in richness varied across data sets. In general, the uncritical incorporation of localities in analyses of species richness along elevational gradients is biased toward inflating the number of species within a given elevational interval. Adding localities that have large vertical uncertainties is likely to extend the elevational distributions of species. Fewer species per elevational interval in the subset would be expected given the decreased sample size and non-biased distribution of uncertainty among localities along the gradient. These results suggest that identifying an emergent pattern of species richness may not vary when data of uneven quality are incorporated, but that the quantitative assessments of richness will vary.

While the overall pattern of species richness and order of species' first and last elevational occurrences is similar across data sets, closer inspection of the underlying distributional patterns of species indicates important differences among data sets. Both community composition at a given elevation and the distribution of species with respect to one another vary across data sets (Fig. 7b). Changes in the placement of species' range endpoints can result in changes in patterns of species range overlap: one type of species co-occurrence pattern (Table 3). Interpreting species co-occurrence patterns is central to community ecology and biogeography. Empirically, patterns of range overlap across spatial scales have often been used to

infer underlying abiotic and biotic mechanisms structuring such distributional patterns (e.g. Brown, 1971; Diamond, 1975; Whittaker, 1975; Pielou, 1977; Legendre *et al.*, 1997; Dale, 1999; Anderson *et al.*, 2002). As a result, considering the implications that misleading patterns of range overlap can have on interpretations of process, underscores the importance of accounting for data quality when historical museum or herbarium specimens are used in establishing species ranges.

For the Uinta Mountains test case, localities were pooled across the entire mountain range rather than from a single transect controlling for sampling, area, and slope effects. However, the same principles and findings illustrated here would be relevant to patterns of species richness and species range overlap along a single transect when museum or herbarium specimens are used either in comparison to, or to supplement modern collecting events where localities were georeferenced in the field.

When the same criteria outlined in the Uinta Mountains test case are applied to the entire 1903 unique coordinate–uncertainty pairs incorporated in this analysis, 1045 (55%) meet the criteria and are considered of high quality. The fact that more than half of the localities georeferenced by MaNIS in this region can be including in analyses at such a fine spatial scale (without requiring reassessment) highlights the benefits that can be attained through biodiversity informatics.

CONCLUSIONS

Recent technological innovations have promoted the development of integrated biodiversity data bases, providing open access to specimen data. Such biodiversity informatics programs can enhance the value of natural history collections and promote previously unfeasible biodiversity analyses. In particular, georeferencing initiatives augment the information associated with the specimen data, while providing the user with the opportunity to incorporate more readily specimen data in analyses, especially spatial analyses. In addition, georeferencing provides the user with estimates of uncertainty that are essential in evaluating data quality in order to identify the appropriate subset for analysis. The ability to evaluate data quality is the limiting factor in how robust an analysis can be and is therefore a measure of the overall utility of biodiversity informatics. How much of a limiting factor depends on the spatial scale of analyses and subsequent necessary levels of precision.

This paper illustrates that species richness patterns along elevational gradients can be exceedingly similar with and without incorporating moderate-to-substantial error in locality elevation. In contrast, failure to account for uneven data quality can yield misleading estimates of species richness and community composition at sites along an elevational gradient, as well as patterns of elevational range overlap among species. These results indicate that a fundamental component of analyses drawing upon biodiversity data bases should be an evaluation of the quality of the underlying specimen data.

As biodiversity informatics expands, it is necessary to recognize the responsibility of the data user to determine

what subset of the data is appropriate for use, to report errors in taxonomic identity or location to the original institution and data base supervisor, and if possible to provide corrections to those errors. In addition, collectors should take great care to provide detailed locality descriptors, both textual descriptors with distance and direction measurements recorded with high precision, as well as coordinates with their respective datum and source. Finally, it should be borne in mind that additional fieldwork to supplement historical data is often essential.

ACKNOWLEDGEMENTS

I thank everyone involved with MaNIS, especially E. A. Rickart, S. L. Robson, A. Baugh, and W. King for georeferencing Utah's localities, and the curators and collection managers of the museums which provided specimen records. Special thanks to T. Schuble for programming the batch analysis of elevational uncertainty estimates and to L. R. Heaney, M. A. Leibold, S. Lidgard, T. Price, E. A. Rickart and W. T. Stanley for critical reviews of earlier versions of the manuscript. In addition, this manuscript was greatly improved by comments from O. R. Vetaas and an anonymous reviewer. I also thank J. Finarelli, T. E. Roberts, P. J. Wagner and J. R. Wiczorek for insightful comments and discussion. Support for this research was provided by a Science to Achieve Results (STAR) Fellowship from the Environmental Protection Agency (award number U91613701), the Chicago chapter of the ARCS foundation, the American Society of Mammalogists, and the Theodore Roosevelt Memorial Fund of the American Museum of Natural History.

REFERENCES

- Anderson, R.P. (2003) Real vs. artefactual absences in species distributions: tests for *Oryzomys albigularis* (Rodentia: Muridae) in Venezuela. *Journal of Biogeography*, **30**, 591–605.
- Anderson, R.P., Peterson, A.T. & Gomez-Laverde, M. (2002) Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos*, **98**, 3–16.
- Bachman, S., Baker, W.J., Brummitt, N., Dransfield, J. & Moat, J. (2004) Elevational gradients, area and tropical island diversity: an example from the palms of New Guinea. *Ecography*, **27**, 299–310.
- Bhattarai, K.R., Vetaas, O.R. & Grytnes, J.A. (2004) Fern species richness along a central Himalayan elevational gradient, Nepal. *Journal of Biogeography*, **31**, 389–400.
- Bieler, R. & Mikkelsen, P.M. (2004) Marine bivalves of the Florida Keys: a qualitative faunal analysis based on original collections, museum holdings and literature data. *Malacologia*, **46**, 503–544.
- Bisby, F.A. (2000) The quiet revolution: biodiversity informatics and the Internet. *Science*, **289**, 2309–2312.
- Brown, J.H. (1971) Mechanisms of competitive exclusion between two species of chipmunk. *Ecology*, **52**, 305–311.
- Canhos, V., Souza, S., Giovanni, R. & Canhos, D. (2004) Global biodiversity informatics: setting the scene for a 'new world' of ecological modeling. *Biodiversity Informatics*, **1**, 1–13.
- Causey, D., Janzen, D.H., Peterson, A.T., Vieglais, D., Krishtalka, L., Beach, J.H. & Wiley, E.O. (2004) Museum collections and taxonomy. *Science*, **305**, 1106–1107.
- Dale, M.R.T. (1999) *Spatial pattern analysis in plant ecology*. Cambridge University Press, Cambridge.
- Davies, N., Villablanca, F.X. & Roderick, G.K. (1999) Bioinvasions of the medfly *Ceratitis capitata*: source estimation using DNA sequences at multiple intron loci. *Genetics*, **153**, 351–360.
- Diamond, J.M. (1975) Assembly of species communities. *Ecology and evolution of communities* (ed. by M.L. Cody and J.M. Diamond), pp. 342–344. Belknap Press, Cambridge, MA.
- Durrant, S.D. (1952) *Mammals of Utah: taxonomy and distribution*. University of Kansas Museum of Natural History Publications, Lawrence.
- Edwards, J.L. (2004) Research and societal benefits of the Global Biodiversity Information Facility. *Bioscience*, **54**, 485–486.
- Fleishman, E., Fay, J.P. & Murphy, D.D. (2000) Upsides and downsides: contrasting topographic gradients in species richness and associated scenarios for climate change. *Journal of Biogeography*, **27**, 1209–1219.
- Fleishman, E., Austin, G.T. & Murphy, D.D. (2001) Biogeography of Great Basin butterflies: revisiting patterns, paradigms, and climate change scenarios. *Biological Journal of the Linnean Society*, **74**, 501–515.
- Gaston, K.J. & Blackburn, T.M. (2000) *Pattern and process in macroecology*. Blackwell Science, Oxford.
- Godoy, J.A., Negro, J.J., Hiraldo, F. & Donazar, J.A. (2004) Phylogeography, genetic structure and diversity in the endangered bearded vulture (*Gypaetus barbatus*, L.) as revealed by mitochondrial DNA. *Molecular Ecology*, **13**, 371–390.
- Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution*, **19**, 497–503.
- Grayson, D.K. & Livingston, S.D. (1993) Missing mammals on Great-Basin mountains – Holocene extinctions and inadequate knowledge. *Conservation Biology*, **7**, 527–532.
- Heaney, L.R. (2001) Small mammal diversity along elevational gradients in the Philippines: an assessment of patterns and hypotheses. *Global Ecology and Biogeography*, **10**, 15–39.
- Kaiser, J. (2000) Fossil databases move to the Web. *Science*, **289**, 2307.
- Knyazhniksiy, O., Monk, R., Parker, N. & Baker, R. (2000) Assignment of global information system coordinates to classical museum localities for relational database analyses. *Occasional Papers Museum of Texas Tech University*, **199**, 1–15.

- Kodric-Brown, A. & Brown, J.H. (1993) Incomplete data sets in community ecology and biogeography – a cautionary tale. *Ecological Applications*, **3**, 736–742.
- Kress, W.J., Miller, S.E., Krupnick, G.A. & Lovejoy, T.E. (2001) Museum collections and conservation efforts. *Science*, **291**, 828c–829.
- Krishtalka, L. & Humphrey, P.S. (2000) Can natural history museums capture the future? *Bioscience*, **50**, 611–617.
- Legendre, P., Galzin, R. & HarmelinVivien, M.L. (1997) Relating behavior to habitat: solutions to the fourth-corner problem. *Ecology*, **78**, 547–562.
- MaNIS (2001) <http://elible.cs.berkeley.edu/manis/>.
- McCain, C.M. (2004) The mid-domain effect applied to elevational gradients: species richness of small mammals in Costa Rica. *Journal of Biogeography*, **31**, 19–31.
- Murphy, P.C., Guralnick, R.P., Glaubitz, R., Neufeld, D. & Ryan, J.A. (2004) Georeferencing of museum collections: a review of problems and automated tools, and the methodology developed by the Mountain and Plains Spatio-Temporal Database-Informatics Initiative (Mapstedt). *Phyloinformatics*, **3**, 1–29.
- Nor, S.M. (2001) Elevational diversity patterns of small mammals on Mount Kinabalu, Sabah, Malaysia. *Global Ecology and Biogeography*, **10**, 41–62.
- Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J.K., Thomas, C.D., Descimon, H., Huntley, B., Kaila, L., Kullberg, J., Tammaru, T., Tennent, W.J., Thomas, J.A. & Warren, M. (1999) Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, **399**, 579–583.
- Peterson, A.T. (2001) Predicting species' geographic distributions based on ecological niche modeling. *Condor*, **103**, 599–605.
- Pielou, E.C. (1977) The latitudinal spans of seaweed species and their patterns of overlap. *Journal of Biogeography*, **4**, 299–311.
- Ponder, W.F., Carter, G.A., Flemons, P. & Chapman, R.R. (2001) Evaluation of museum collection data for use in biodiversity assessment. *Conservation Biology*, **15**, 648–657.
- Raxworthy, C.J., Martinez-Meyer, E., Horning, N., Nussbaum, R.A., Schneider, G.E., Ortega-Huerta, M.A. & Peterson, A.T. (2003) Predicting distributions of known and unknown reptile species in Madagascar. *Nature*, **426**, 837–841.
- Rickart, E.A. (2001) Elevational diversity gradients, biogeography and the structure of montane mammal communities in the intermountain region of North America. *Global Ecology and Biogeography*, **10**, 77–100.
- Ruedas, L.A., Salazar-Bravo, J., Dragoo, J.W. & Yates, T.L. (2000) The importance of being earnest: what, if anything, constitutes a “specimen examined”? *Molecular Phylogenetics and Evolution*, **17**, 129–132.
- Shaffer, H.B., Fisher, R.N. & Davidson, C. (1998) The role of natural history collections in documenting species declines. *Trends in Ecology & Evolution*, **13**, 27–30.
- Skelly, D.K., Yurewicz, K.L., Werner, E.E. & Relyea, R.A. (2003) Estimating decline and distributional change in amphibians. *Conservation Biology*, **17**, 744–751.
- Soberon, J. & Peterson, A.T. (2004) Biodiversity informatics: managing and applying primary biodiversity data. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **359**, 689–698.
- Solow, A.R. & Roberts, D.L. (2003) A nonparametric test for extinction based on a sighting record. *Ecology*, **84**, 1329–1332.
- Stein, B.R. & Wicczorek, J. (2004) Mammals of the world: MaNIS as an example of data integration in a distributed network environment. *Biodiversity Informatics*, **1**, 14–22.
- Sugden, A. & Pennisi, E. (2000) Diversity digitized. *Science*, **289**, 2305.
- Vetaas, O.R. (2000) Comparing species temperature response curves: population density versus second-hand data. *Journal of Vegetation Science*, **11**, 659–666.
- Vetaas, O.R. & Grytnes, J.A. (2002) Distribution of vascular plant species richness and endemic richness along the Himalayan elevation gradient in Nepal. *Global Ecology and Biogeography*, **11**, 291–301.
- Wheeler, Q.H., Raven, P.H. & Wilson, E.O. (2004) Taxonomy: impediment or expedient? *Science*, **303**, 285.
- Whittaker, R.H. (1975) *Communities and ecosystems*. Macmillan, New York, NY.
- Wicczorek, J.R. (2001) *Georeferencing calculator* (<http://elible.cs.berkeley.edu/manis/GC.html>).
- Wicczorek, J., Guo, Q.G. & Hijmans, R.J. (2004) The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *International Journal of Geographical Information Science*, **18**, 745–767.
- Wilson, D.E. & Reeder, D.M. (eds) (1993) *Mammal species of the world: a taxonomic and geographic reference*. Smithsonian Institution Press, Washington D.C..

BIOSKETCH

Rebecca J. Rowe is a graduate student in the Committee on Evolutionary Biology at the University of Chicago and the Field Museum of Natural History with research interests in the fields of biogeography, macroecology, community ecology, and conservation biology. In particular, her research focuses on patterns of species richness, abundance, and the distribution of small mammals along ecological gradients.

Editor: José Alexandre F. Diniz Filho

Appendix 1 16,692 specimens from the orders Rodentia, Lipotyphla, and Lagomorpha, were collected at the 3144 non-unique localities included in this study. The distribution of those specimens across museums is presented below. Of these 16,692 specimens, approximately 11,470 were retained after data quality assessment

Collection	Number of specimens
Utah Museum of Natural History	12,083
The Field Museum	2433
University of Michigan Museum of Zoology	922
Kansas University Museum of Natural History	391
Museum of Vertebrate Zoology	317
University of Washington Burke Museum	266
Texas Tech University	107
Museum Southwestern Biology	62
Louisiana Museum of Natural History	57
Los Angeles County Museum	33
James R. Slater Museum of Natural History,	19
Univ. of Puget Sound	
University of Alaska Museum	2

Appendix 2 The number of specimens and unique coordinate–uncertainty pairs (labelled localities) associated with each of the 35 species in the Uinta Mountains test case. Numbers are provided separately for both the complete data set and subset

Species	Complete data set		Subset	
	Specimens	Localities	Specimens	Localities
<i>Clethrionomys gapperi</i>	45	14	41	12
<i>Cynomys leucurus</i>	3	3	0	0
<i>Glaucomys sabrinus</i>	6	4	5	3
<i>Lemmys curtatus</i>	13	8	13	8
<i>Lepus americanus</i>	4	4	2	2
<i>Lepus townsendii</i>	7	4	1	1
<i>Marmota flaviventris</i>	13	10	9	6
<i>Microtus longicaudus</i>	87	32	75	26
<i>Microtus montanus</i>	88	30	72	22
<i>Microtus richardsoni</i>	29	10	26	7
<i>Neotoma cinerea</i>	4	3	4	3
<i>Ochotona princeps</i>	23	10	20	8
<i>Ondatra zibethicus</i>	5	4	1	1
<i>Onychomys leucogaster</i>	2	1	2	1
<i>Peromyscus crinitus</i>	15	2	15	2
<i>Peromyscus maniculatus</i>	366	64	277	42
<i>Perognathus parvus</i>	29	7	28	6
<i>Peromyscus truei</i>	17	5	10	3
<i>Phenacomys intermedius</i>	12	3	12	3
<i>Reithrodontomys megalotis</i>	2	2	1	1
<i>Sorex cinereus</i>	3	3	3	3
<i>Sorex monticolus</i>	28	16	25	13
<i>Sorex palustris</i>	15	6	15	6
<i>Spermophilus armatus</i>	21	13	14	6
<i>Spermophilus elegans</i>	7	1	7	1
<i>Spermophilus lateralis</i>	27	16	23	13
<i>Sylvilagus audubonii</i>	3	2	2	1
<i>Sylvilagus nuttallii</i>	7	6	2	2
<i>Tamias dorsalis</i>	9	6	7	4
<i>Tamiasciurus hudsonicus</i>	44	24	40	21
<i>Tamias minimus</i>	143	48	103	34
<i>Tamias umbrinus</i>	94	35	57	24
<i>Thomomys idahoensis</i>	10	4	8	2
<i>Thomomys talpoides</i>	49	19	36	13
<i>Zapus princeps</i>	45	18	40	13