

Documenting the Biodiversity of the Madrean Archipelago: An Analysis of a Virtual Flora and Fauna

Nicholas S. Deyo and Thomas R. Van Devender
Sky Island Alliance, Tucson Arizona

Alex Smith
School of Natural Resources and the Environment, University of Arizona, Tucson Arizona

Edward Gilbert
SEINet Flora Database, Tucson, Arizona

Abstract—The Madrean Archipelago Biodiversity Assessment (MABA) of Sky Island Alliance is an ambitious project to document the distributions of all species of animals and plants in the Madrean Archipelago, focusing particularly on northeastern Sonora and northwestern Chihuahua, Mexico. The information is made available through MABA's online database ([madrealn.org](http://www.madrealn.org)). The sources of these records are museum collections, herbaria, scientific literature, other databases, observations, and field notes. Many observations come from MABA expeditions to isolated mountain ranges in northeastern Sonora involving taxonomic specialists, professors, students, and others from universities in Arizona and Sonora; and personnel from Mexican agencies. This paper provides an analysis of the current status of species records contained within the MABA flora and fauna databases and the significance of these records for future collection, research, and conservation planning.

Introduction

Project Background

Since 2009, the Sky Island Alliance's Madrean Archipelago Biodiversity Assessment (MABA) has been documenting the biodiversity of the Sky Island region with the aim of supporting conservation and scientific inquiry. The MABA project is a direct result of the Madrean Conferences held in 1994 and 2004. These conferences convened scientists and natural resource managers from the United States and Mexico to discuss the biodiversity and natural resources management of the Sky Island region (DeBano and others 1995; Gottfried and others 2005). One conclusion from these conferences was the lack of biodiversity data needed to support the management and conservation of the Madrean Archipelago, particularly in Mexico. Documenting the biodiversity of the Sky Island region is challenging, as the area contains over 50 mountain ranges/mountain complexes, hundreds of rare and localized species, and numerous habitats dispersed between the Rocky Mountains to the north, the Sierra Madre Occidental to the south, and the Sonoran and Chihuahuan deserts to the west and east. Sky Island Alliance worked with programmer Ed Gilbert, the

SEINet flora database, Symbiota, and the Veolia Foundation to create the MABA online database (<http://www.madrealn.org>); since its creation, the database has become a large repository of information for many groups of species in the Sky Island region and includes both vouchered specimens and observations from herbaria, museum collections, scientific literature, and MABA's own field expeditions. This paper provides an analysis of the current status of species records contained within the MABA flora and fauna databases and the significance of these records for future collection, research, and conservation planning.

Virtual Flora and Fauna

The MABA online database is considered a "virtual flora and fauna," a type of digital resource that is revolutionizing the way biodiversity information is documented and shared. Floras and faunas refer to both the species of plants and animals that inhabit a region as well as the publications that describe them (Heidorn 2004). Historically, floras and faunas have taken the form of print media and contain information such as scientific classifications, species descriptions, illustrations, photos, distribution maps, and identification keys. There are limitations to printed media when it comes to documenting biodiversity, not least of which are the sheer number of organisms in some taxonomic groups; given that there are an estimated 5,300 species of flowering plants in Madrean Pine-Oak Woodlands (Conservation International 2011), a print flora for the entire region would be impractical. Additionally, the cost and labor associated with publishing regional floras and faunas also limit production and distribution of these works; in 2003 the estimated cost of producing a flora with 2,104 species was \$1,579,946 (Heidorn 2004).

In: Gottfried, Gerald J.; Ffolliott, Peter F.; Gebow, Brooke S.; Eskew, Lane G.; Collins, Loa C., comps. 2013. Merging science and management in a rapidly changing world: Biodiversity and management of the Madrean Archipelago III; 2012 May 1-5; Tucson, AZ. Proceedings. RMRS-P-67. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Web-based virtual floras overcome many of these problems: they are less expensive, support unlimited images and species records, allow for complex data queries, and can be corrected and updated continually (Gilbert, unpublished presentation). Web-based flora and fauna have taken databases, once the private domain of individual researchers, and have extended them to be a medium for collaboration between researchers and citizen scientists (Schnase and others 1997). Finally, web-based floras and faunas increase the speed at which biodiversity data is collected and made available, a critical factor considering how quickly species and habitats are being lost to human activities (Heidorn 2004).

The last two decades have seen a worldwide effort to digitize herbaria and museum collections, resulting in flora and fauna databases containing on the order of 10^4 and 10^6 records (Soberón and others 2002). Access to such large datasets has spurred the burgeoning field of biodiversity informatics (Soberón and Peterson 2004), and has led to scientific applications such as modeling species ranges and evolutionary analyses (Petersen and others 1999; Zhong 1999); locating areas for biodiversity exploration (Jones and Gladkov 1999; Soberón and others 2004); prioritizing for protected areas (Godown and Peterson 2000; Kelley and others 2002), and modeling routes for the spread of invasive species (Higgins and others 1999; Peterson 2004).

No matter how sophisticated the design of a virtual flora or fauna, it is only as good as the data it contains. A database with a comprehensive set of species records allows for comparison of species lists across sites; observation of species richness over time, including before and after habitat modifications; and analyses such as those mentioned in the previous paragraph (Soberón and others 2004). Adding records to the MABA database has been ongoing since the program began in 2009. This research reflects the status of the MABA database through analyzing species records in reference to Sky Islands/complexes, reporting the sources of records, and identifying which time periods have been the most productive for collection efforts. The results of this paper are useful for targeting areas for future collection and data acquisition efforts and better understanding how the MABA database can be used for research and conservation planning.

Methods

Defining Sky Island Ranges and Complexes

In order to perform count and density analyses on species data for each Sky Island/complex, a GIS polygon layer was created defining the physiographic boundaries of each range or complex containing oak woodland. Generally, the boundaries of ranges or complexes were digitized using physiographically defined mountainous features containing contiguous or closely clustered areas of oak woodland that were isolated from the main Sierra Madre and Rocky Mountains. Previously created maps of the region have classified Sky Islands exclusively on the presence of vegetation (Marshall 1957; Brown and Lowe 1980). Range/complex borders were defined by drawing along a 7 percent slope gradient threshold (see fig. 1). In areas of ambiguous slope, borders were drawn to include 50 percent or greater digital elevation model (DEM) cells with a 7 percent slope gradient or higher. These boundaries emphasize the physiographic continuity of complexes or ranges, including areas outside of the oak woodland boundaries that are still mountainous. Physiographic ranges and complexes were subdivided when they showed significant gaps in oak woodland presence. These divisions were drawn along low points of elevation in those gaps. Additionally, two ranges that did not show oak woodlands in the Geographic Information System (GIS) vegetation database used in this analysis, the Mule Mountains

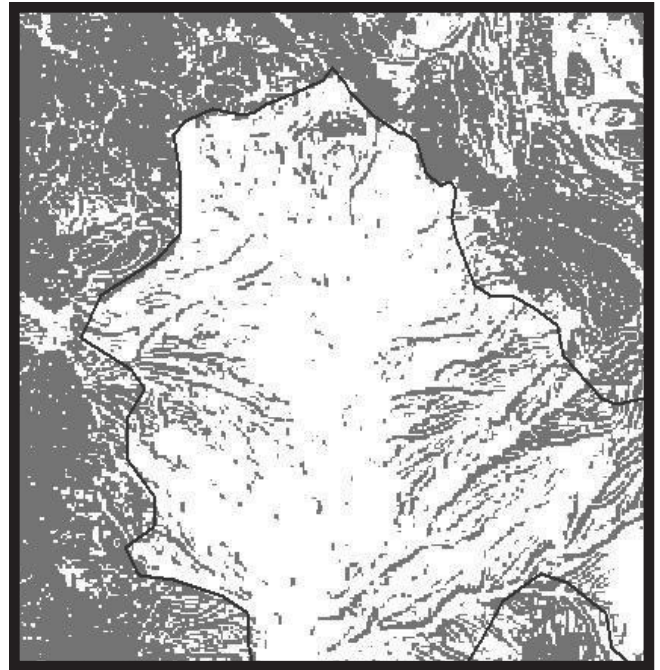


Figure 1—This image shows areas above 7 percent grade (white) that were used to help digitize the boundary of Sky Island ranges and complexes.

and the Cerro Cobachi, were included based on the field experience of Dr. Tom Van Devender. A small number of ranges and complexes that show oak woodland connections to the Sierra Madre were included in this analysis because of historical precedent. These “peninsulas” of the Sierra Madre are noted in table 1.

Names of complexes were determined by combining the northernmost and southernmost range names. Range names were acquired from U.S. Geological Survey (USGS, 2002) and Mexican National Institute of Statistics and Geography (INEGI) topographic maps. In some cases, a more commonly known name was chosen over the naming convention above. A single unhyphenated name was used for isolated ranges and smaller subdivisions of Sky Island complexes. Figure 2 provides a map of the Sky Island ranges/complexes used in this analysis.

Data Acquisition and Manipulation

Species point locations and associated data were obtained by SQL queries to the separate MABA flora and fauna databases. Queries were made using a web-based MySQL interface to facilitate downloading large datasets; however, the same data can be accessed using the MABA online database (<http://www.madrean.org>). Latitude and longitude coordinates bounded queries to a region containing Arizona, New Mexico, Sonora, and Chihuahua. Records lacking geographic coordinates were eliminated. Each acquired dataset was then imported into a GIS and projected into the Albers Conformal Conic projection using the NAD 83 datum. A spatial join was then created to link species data with Sky Islands/complexes boundaries.

Queries were made for the six separate classes of animals: Arthropoda, Aves, Mammalia, Osteichthyes, Amphibia, and Reptilia. Amphibia and Reptilia classes were then combined. For plant records, the entire database was downloaded and classified by family name into three separate datasets for angiosperms, gymnosperms, and ferns.

Table 1—Species records by Sky Island or Sky Island complex.

	Sky Island/Complex Name	Area (sq km)	Total Animals	Animal Density ^a	Total Plants	Plant Density ^a	Amphibia- Reptilia	Arthropoda	Aves	Mammalia	Osteichthyes	Angiosperms	Asteraceae	Poaceae	Fabaceae	Gymnosperms	Ferns and Allies
U.S. Sky Island Complexes	ALAMO HUECO	99	0	0.00	80	0.81	0	0	0	0	0	77	12	16	4	0	3
	ANIMAS ^b	543	1	0.00	1221	2.25	1	0	0	0	0	1170	226	183	117	31	20
	BIG HATCHET	153	0	0.00	112	0.73	0	0	0	0	0	108	8	3	12	1	3
	COYOTE - BABOQUIVARI	552	9	0.02	3065	5.56	8	0	0	1	0	2865	463	429	312	17	183
	DOS CABEZAS - CHIRACAHUA	1,630	10	0.01	10871	6.67	4	0	0	6	0	10265	2123	1231	958	314	292
	DRAGOON	180	1	0.01	1274	7.08	0	1	0	0	0	1120	150	204	100	25	129
	GALIURO	1,500	56	0.04	2136	1.42	12	2	37	3	2	2057	275	191	238	29	50
	HUACHUCA - PATAGONIA	1,083	174	0.16	10930	10.09	71	64	12	6	21	10122	1999	1506	967	230	578
	LITTLE DRAGOON	84	0	0.00	143	1.71	0	0	0	0	0	136	20	29	26	0	7
	LITTLE HATCHET	115	0	0.00	169	1.47	0	0	0	0	0	165	24	34	15	1	3
	MULE	335	10	0.03	3835	11.46	10	0	0	0	0	3710	642	193	448	50	75
	PELONCILLO	618	5	0.01	227	0.37	2	1	0	1	1	222	35	22	19	1	4
	PERILLA	192	14	0.07	209	1.09	4	10	0	0	0	202	29	38	18	1	6
	PINALEÑO	829	6	0.01	5053	6.10	0	0	0	3	3	4714	725	658	307	165	174
	RINCON	877	5	0.01	4854	5.53	0	2	0	3	0	4614	656	907	463	28	212
	SANTA CATALINA	770	5	0.01	9989	12.97	0	2	0	3	0	9341	1610	1170	908	175	473
	SANTA RITA	811	11	0.01	6737	8.31	3	2	0	4	2	6430	1152	929	762	86	221
	SANTA TERESA	954	0	0.00	1818	1.91	0	0	0	0	0	1725	233	253	157	31	62
	SIERRITA	148	0	0.00	117	0.79	0	0	0	0	0	109	16	35	6	2	6
	SWISSHELM	133	37	0.28	218	1.64	0	37	0	0	0	211	41	9	29	3	4
	TORTOLITA	132	1	0.01	80	0.61	0	0	0	1	0	76	8	9	7	2	2
	WHETSTONE	210	35	0.17	577	2.75	3	0	32	0	0	546	88	99	47	12	19
	WINCHESTER	288	0	0.00	79	0.27	0	0	0	0	0	78	20	2	6	1	0
Mexican Sky Island Complexes	ACONCHI	2,267	560	0.25	288	0.13	54	66	427	3	10	274	37	18	66	2	12
	ANIBACACHI	102	10	0.10	853	8.40	10	0	0	0	0	816	116	122	86	3	34
	ATASCOSA - CIBUTA	2,443	570	0.23	5218	2.14	87	6	428	6	43	4955	753	681	686	34	229
	BATAMOTE	37	0	0.00	0	0.00	0	0	0	0	0	0	0	0	0	0	0
	CERRO COBACHI	261	5	0.02	14	0.05	0	0	0	5	0	14	3	0	6	0	0
	CERRO EL TIZNADO	370	31	0.08	17	0.05	4	0	27	0	0	17	1	1	8	0	0
	CERRO HUEQUECHI	241	19	0.08	145	0.60	6	0	7	0	6	139	15	16	20	1	5
	CHIVATO	340	68	0.20	61	0.18	0	0	67	0	1	58	2	6	6	3	0
	EL CARMEN - VERDE	2,702	234	0.09	413	0.15	60	0	154	2	18	408	46	24	128	2	3
	EL GATO ^b	1,243	741	0.60	779	0.63	67	174	465	29	6	667	74	26	75	64	48
	EL HUMO	122	115	0.94	172	1.41	1	0	114	0	0	169	21	26	19	0	3
	EL MANZANAL	980	41	0.04	393	0.40	12	5	4	0	20	390	53	54	57	2	1
	EL MAVIRO	412	158	0.38	438	1.06	97	0	39	5	17	434	43	15	136	0	4
	EL PAJARITO - LA CAÑADA	441	139	0.32	98	0.22	5	20	110	1	3	98	19	3	22	0	0
	EL PINITO	944	556	0.59	337	0.36	272	53	219	0	12	316	40	46	44	0	21
	EL TIGRE ^b	3,277	1476	0.45	1768	0.54	428	55	927	28	38	1645	222	145	198	91	32
	ENMEDIO	129	0	0.00	0	0.00	0	0	0	0	0	0	0	0	0	0	0
	LA HUERTA	525	301	0.57	177	0.34	195	2	88	0	16	177	7	6	56	0	0
	LA MADERA - CUCURPE	806	309	0.38	585	0.73	22	43	240	4	0	567	59	49	94	11	7
	LAS CALABAZAS - MARTINEZ	560	72	0.13	302	0.54	7	3	61	1	0	297	30	6	80	0	5
	LAS GUIJAS - LA SEÑORITA	1,414	171	0.12	77	0.05	9	15	130	6	11	77	6	10	13	0	0
	LOS AJOS - LA MADERA	5,094	2091	0.41	2646	0.52	347	373	1330	35	6	2452	303	199	350	140	54
	LOS CABALLOS	600	52	0.09	157	0.26	3	6	31	2	10	157	7	2	40	0	0
	MARIQUITA - ELENITA	334	625	1.87	401	1.20	7	45	571	2	0	315	57	28	22	84	2
	MAZATÁN	174	61	0.35	158	0.91	15	45	0	1	0	137	50	24	20	0	21
	PELONCILLO - PAN DURO ^b	3,198	3069	0.96	4443	1.39	408	357	2185	14	105	3934	549	494	384	451	58
	SAN ANTONIO - AZUL	2,147	942	0.44	1776	0.83	55	113	733	35	6	1666	185	149	151	81	29
	SAN JAVIER	360	290	0.81	1007	2.80	26	173	91	0	0	969	143	33	160	0	38
	SAN JERONIMO - LA CABRERA	475	12	0.03	267	0.56	0	0	5	2	5	265	26	16	71	0	2
	SAN JOSE	85	51	0.60	105	1.24	1	0	50	0	0	95	8	8	8	9	1
	SAN JUAN	214	0	0.00	2	0.01	0	0	0	0	0	2	0	0	1	0	0
	SANTO NIÑO - LA CAMPANERIA	627	129	0.21	243	0.39	10	8	106	5	0	243	28	4	63	0	0

^aDensity was calculated by dividing the number of records by the area in square meters.^bIndicates that the range or complex has a direct oak woodland connection to the Sierra Madre.

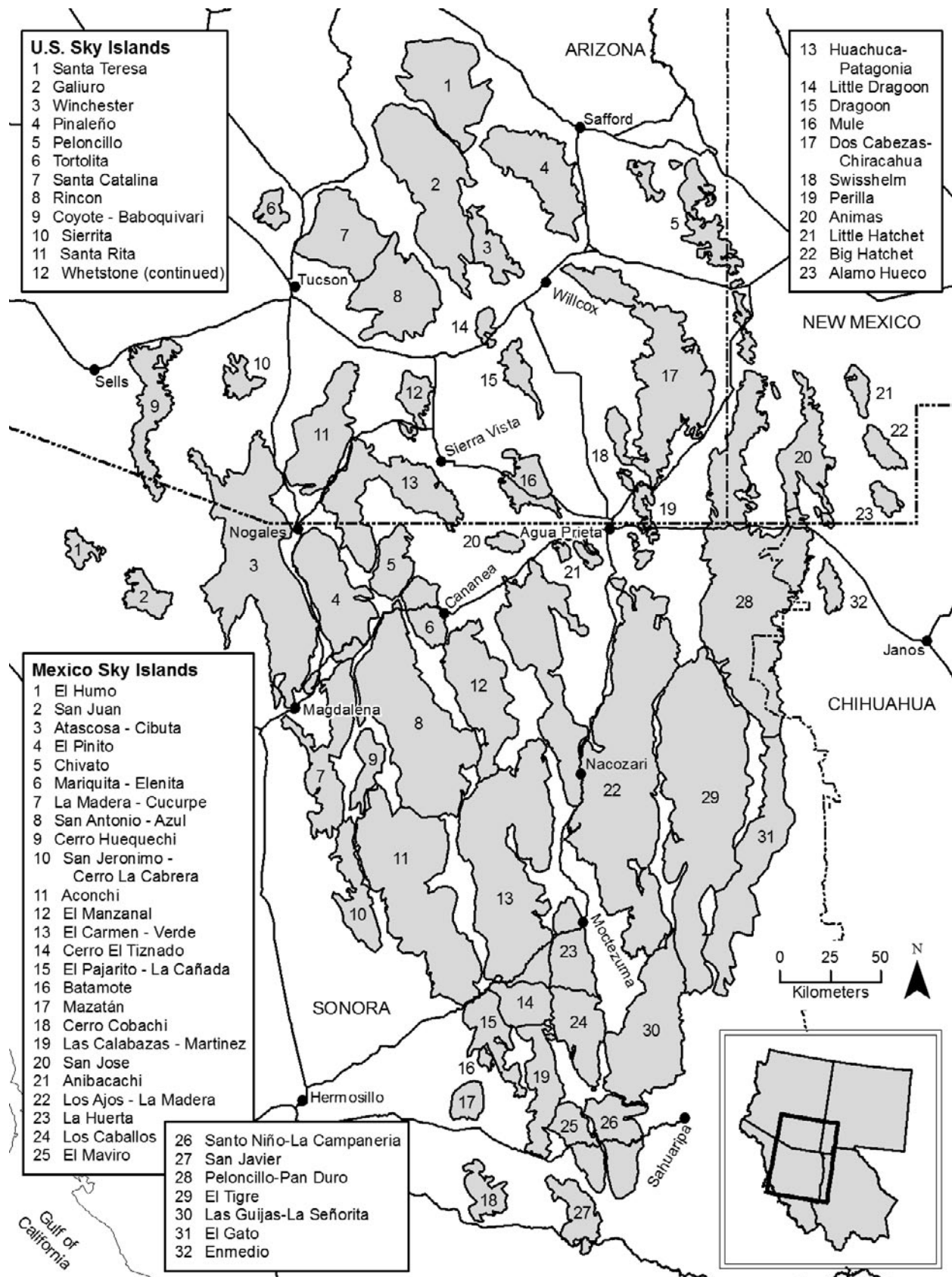


Figure 2—Map depicting the Sky Island ranges and complexes used to the distribution of species record data in the Madrean Archipelago.

and fern allies. Observations of members of the families Asteraceae, Poaceae, and Fabaceae were then isolated from the angiosperm data. Analysis of contributing institutions and record collection dates was performed using data downloads from separate SQL queries. This separate analysis contained some discrepancies between the two querying efforts. This may be attributed to the addition of new sets of data since the first queries were performed.

The Nature Conservancy's Apache Highlands Ecoregion conservation priorities analysis provided vegetation data for Arizona, New Mexico, and northern Sonora and Chihuahua (Marshall and others 2004). This data combined Gap Analysis Program (GAP) vegetation data for areas in the United States with Forest Inventory 2000 layers for areas in Mexico. As this data terminates in the middle of Sonora, data for more southern regions was acquired from an INEGI layer entitled "uso del suelo y vegetación" (2005). Additionally, a digitized layer of Brown and Lowe's Biotic Communities of the Southwest classification map (1980) was used to supplement the more modern, digitally mapped layers. Where oak woodland data differed between sources, the combined area of oak woodlands in all layers for the given region was used to define Sky Islands/complexes. Digital elevation model (DEM) data used to create the slope analysis layer is a stitched mosaic of 90 m resolution raster imagery from NASA's 2000 Shuttle Radar Topography Mission (USGS 2002).

Results

General Statistics

A total of 596,681 plant and 121,355 animal records with geographic coordinates were queried from the MABA flora and fauna databases for the States of Arizona, New Mexico, Sonora, and Chihuahua. Of these records, 87,134 plant and 13,278 animal records were identified within the Sky Island/complex polygons created for this analysis. These figures represent only 14 percent of all records within the states mentioned. However, as the Sky Islands/complexes only comprise 4 percent of the area of those states, these figures show a greater number of records within Sky Islands/complexes. Across all Sky Islands/complexes, the average sampling effort (the number of records/area in km²) for plants was 2.18 records/km² and 0.22 records/km² for animals, representing a tenfold difference in sampling effort between the MABA flora and fauna databases. When sampling effort was compared between Sky Islands/complexes in Mexico versus the United States, the results differed between the animal and plant records. Note that the country label of Sky Islands/complexes that straddled the border refers to the location of the polygon's centroid. Plants were better represented in the United States with an average sampling effort of 3.98 records/km² compared with 0.87 records/km² in Mexico. Animals were better represented in Mexico with an average sampling density of 0.35 records/km², while the United States had an average of 0.04 records/km². Table 1 provides an overview of the data generated in this analysis.

Contributing Institutions

The databases were analyzed to better understand where the majority of records originated. Twenty-two institutions contributed to the MABA flora database. The University of Arizona Herbarium contributed the most records, followed by the Arizona State University Herbarium, MABA observations, Sonoran Desert Plants: An Ecological Atlas, and the Desert Botanical Garden. Thirty-six institutions contributed to the MABA fauna database. MABA contributed 13,198 records, by

far the greatest number of observations, followed by the University of Arizona, Arizona State University, eBird, and the University of Texas at Arlington. Figure 3 shows the distribution of records among contributing institutions.

Temporal Distribution of Records

Examining collection dates for species records revealed that both the MABA flora and fauna databases were right skewed (fig. 4) toward more modern collection dates. Seventy-five percent of all plant records were collected after 1962 while the same percentage of animal records was collected after 1976. The earliest plant record in the database was an *Acacia texensis* collected by G. Thurber in 1851. The earliest animal record was a *Rana yavapaiensis* collected by Edgar A. Mearns in 1893.

Within the MABA flora database, most ranges/complexes show recent collection efforts; however, it appears that the Cerro Cobachi, Sierrita, Big Hatchet, El Maviro, Las Calabazas-Martinez, and Santo Niño ranges have not seen collection efforts since the late to mid-1990s. Fewer ranges within the MABA fauna database show recent collection efforts; the Sierrita, Tortolita, Pinaleño, Cerro El Tiznado, Mule, Cerro Cobachi, and Animas ranges do not show any new records since the late 1990s, and the Sierrita, Tortolita, and Pinaleño ranges do not show any records since the late 1970s. It should be noted that the Sierrita range only showed two animal records revealed in a later query of the database. The statistical mode of record collection dates was calculated for each individual Sky Island/complex for flora and fauna records separately. For plant records, 14 of the 33 Mexican Sky Islands/complexes had modes more recent than 2000, while only 4 of the 22 United States Sky Islands/complexes had modes more recent than 2000. This indicates that sampling efforts in Mexico have been more vigorous recently than those in the United States. For animal records, 12 of the 33 Mexican Sky Islands/complexes had modes more recent than 2000, while only 3 of the 22 United States Sky Islands/complexes had modes more recent than 2000. As with plant records, this shows that the collection effort for Mexico has been more vigorous recently.

Records by Taxa

The databases were queried based on major taxonomic groups to better understand which groups were better sampled; conversely, this analysis also revealed which Sky Island ranges/complexes have the fewest number of records. In terms of overall records (both animals and plants), the Huachuca-Patagonia complex yielded the most records with 11,104 observations. Not surprisingly, other ranges/complexes with large numbers of records were the Dos Cabezas-Chiracahuas (10,881 records), the Santa Catalinas (9,994 records), the Peloncillo-Pan Duros (7,512 records), and the Santa Ritas (6,748 records). The most poorly represented ranges/complexes were Batamote, Enmedio, San Juan, Cerro Cobachi, and Cerro El Tiznado. Of these, Batamote and Enmedio had no records and the remainder all had fewer than 50 records.

The animal taxa queried in the MABA fauna database were Arthropoda, Aves, Mammalia, Osteichthyes (bony fish), and Amphibia/Reptilia. There were more records for Aves than for any other class of animals with 8,690 records, followed by Amphibia/Reptilia (2,326 records), Arthropoda (1,683 records), Osteichthyes (362 records) and Mammalia (217 records). The top five ranges/complexes for Arthropoda were the Los Ajos-La Madera, Peloncillo-Pan Duro, El Gato, San Javier, and San Antonio-Azul. Of the 55 ranges/complexes identified

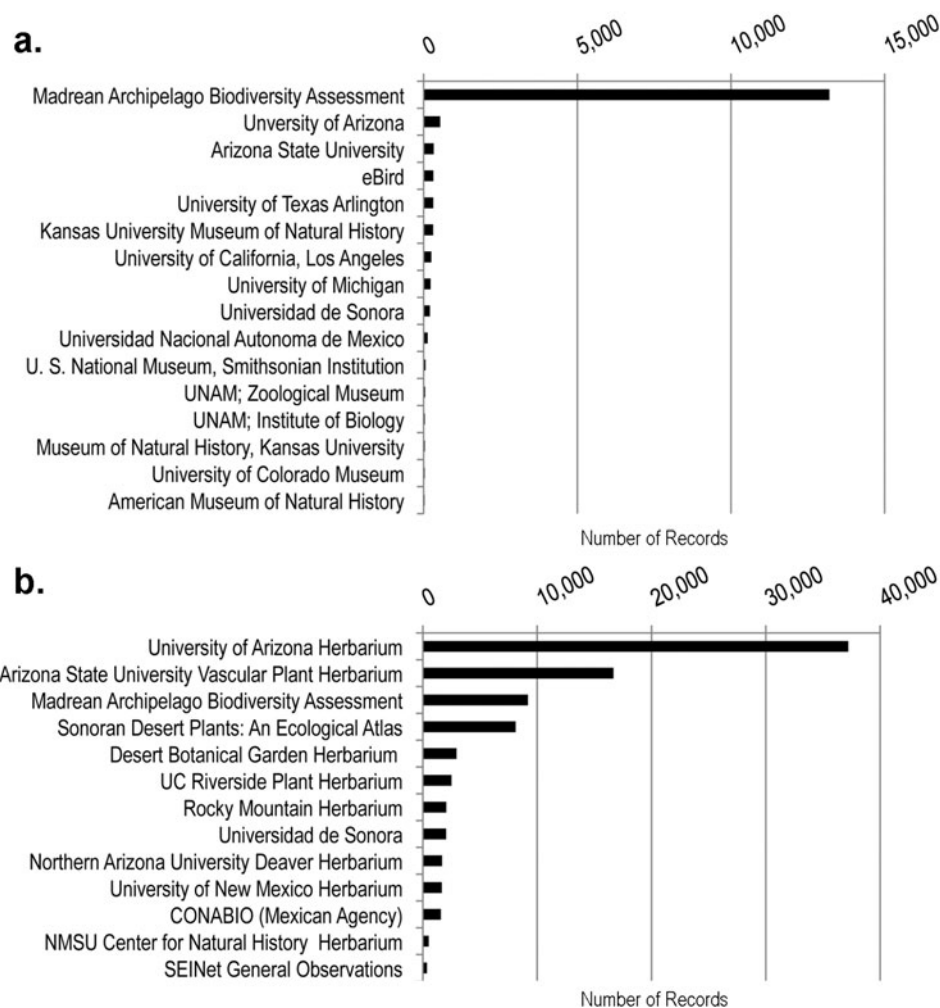


Figure 3—Major contributing institution for the MABA fauna database (a) and the MABA flora database (b). Note: not all contributing institutions are depicted here.

here, 40 have 10 records or fewer. The top five ranges/complexes for Aves were the Peloncillo-Pan Duro, Los Ajos-La Madera, El Tigre, San Antonio-Azul, and Mariquita-Elenita. While Aves was better represented than other taxa, 29 ranges/complexes remain without bird records. For Amphibia and Reptilia combined, the best-represented ranges/complexes were El Tigre, Peloncillo-Pan Duro, Los Ajos-La Madera, El Pinito, and La Huerta. Thirty-eight ranges/complexes had 10 or fewer records for reptiles and amphibians. Osteichthyes and Mammalia showed very few records, so comparisons between ranges are excluded here (see table 1).

The plant groups queried in the MABA flora database were the divisions for angiosperms, gymnosperms, and ferns and fern allies. Angiosperms were further divided into the families Asteraceae, Poaceae, and Fabaceae. All divisions had several thousand records each. Angiosperms were best represented with 81,861 records (expected since angiosperms include many more species than other groups), followed by ferns and fern allies (3,135 records) and gymnosperms (2,183 records). The angiosperm families were represented with 13,458 Asteraceae, 10,361 Poaceae, and 8,996 Fabaceae. Gymnosperms were least well represented with 34 ranges/complexes having fewer than 10 records, while ferns and fern allies reflected 31 ranges/complexes with fewer than 10 records. Angiosperms had at least 10

records in all ranges except 3: San Juan (2 records), Batamote, and Enmedio. Of the 55 Sky Island ranges and complexes, only 9 had records for each taxonomic group (both animals and plants); these were the Huachuca-Patagonias, Atascosa-Cibutas, Peloncillo-Pan Duros, Los-Ajos-La Maderas, Galiuros, San Antonio-Azuls, El Tigre, El Gato, and Aconchi.

Discussion

Biodiversity databases that are “extensive and exhaustive” are important because they allow for the application of biodiversity data to basic research and conservation planning purposes (Soberón and others 2004). This analysis was designed to help the Sky Island Alliance’s MABA program better understand the strengths and weaknesses of its flora and fauna biodiversity databases, and to inform the next steps to improve the coverage and quality of species data. It is also an opportunity to consider how these data can be applied for conservation planning purposes.

Specimen records are typically biased by the spatial and temporal distribution of collection efforts (Soberón and Peterson 2004; Soberón and others 2004). Biases in species location data are often attributed

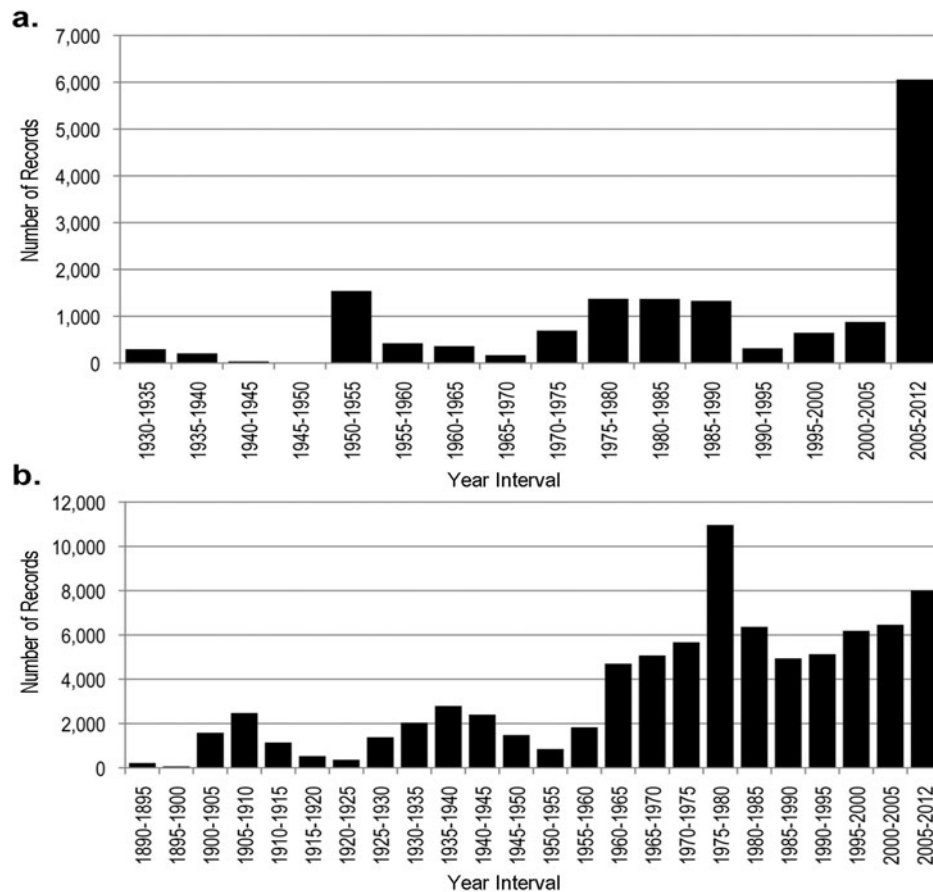


Figure 4—The distribution of species records collection dates for 5-year intervals for the MABA fauna database (a) and the MABA flora database (b).

to proximity to roads, research stations, educational institutions, land tenure, and cities (Soberón and others 2000). This analysis revealed that many Sky Islands and complexes have no or very little species data; only nine ranges contained records for all of the taxonomic groups queried here. Additionally, there remains a large disparity in sampling effort between the United States and Mexico. Plants are better represented in the United States by a factor of nearly five to one, which is likely due to actual sampling effort because of the large network of herbaria that have provided data. Animal records are better represented in Mexico by a factor of eight to one. This discrepancy along with the overall better representation of plants in the database indicates that a large number of existing animal records in museum collections are not included in the MABA database. Therefore, adding additional animal collections from museums and other sources to the MABA database should be a priority. Additionally, any Sky Islands/complexes lacking in species data across multiple groups of taxa should be considered for future field collection efforts.

Biodiversity databases rely on the species collections contained in museums and herbaria. Natural history data may become antiquated and less useful for contemporary applications if collections are not continually updated (Peterson and others 1998; Winker 1996). At first glance, the distributions of record collection dates appeared to be relatively contemporary, with 75 percent of records being collected in the last 50 years. However, this does not account for biases in the collection of individual taxa, which may occur when institutions receive funding to study specific taxonomic groups or when certain

taxa are treated with particular interest in a given period (Soberón and others 2004). Individual Sky Islands/complexes that do not contain records from recent years are due for new collection efforts.

The importance of continuing the collection and compilation of species data cannot be understated. Species records from herbaria, museum collections, and other sources are the foundation of our understanding of the biodiversity of the planet and how to conserve it (Winker 1996). While there are remote sensing methods for assessing areas for high biodiversity and conservation priority, the failure to include species data in these analyses may leave unique and vulnerable species unprotected (Brooks and others 2004). Furthermore, increased access to species data through online databases like MABA's, as well as the huge increase in other sources of environmental data, has led to innovative methods of modeling species distributions for researching biodiversity and conservation planning purposes (Lobo and others 1997; Jones and Gladkov 1999; Soberón and others 2004; Rebelo and Siegfried 1992; Csuti and others 1997; Godown & Peterson 2000; Peterson and others 2000; Kelley and others 2002; Burgess and others 2002; Chen & Peterson 2002). Combining species data with data on climate, physiography, and other characteristics allows the modeling of ecological niches (Soberón and Peterson 2004). The use of these methods would help overcome some of the gaps and biases in species data now available for the Sky Island region and contribute to a better understanding of conservation priorities in the region.

However, before data can be used for the applications mentioned above its quality must be assessed. Large mixed source databases

such as MABA's often face difficulties in maintaining data quality. The errors in most need of review are those related to taxonomy and incorrect georeferencing. Taxonomic errors are somewhat mitigated by taxonomic thesauruses that check for antiquated scientific names, a continual challenge in a world of rapidly changing taxonomy. The correction of georeferencing errors is more laborious and requires cross-referencing species' locality descriptions with their geographic coordinates (Soberón and others 2002); although there are some statistical tools available to help with this process (Chapman and Busby 1994). Some assurance of quality control would improve the confidence with which researchers and lay people could use MABA data. Without quality control measures,

The MABA project has made great strides in compiling biodiversity data within the Madrean Archipelago, both from existing collections and from field collection efforts. There still remain numerous gaps in the biodiversity data available for the region, further emphasizing the need for more collection efforts. Despite the lack of data for specific areas, there remains a great deal of data that is available for conservation and research efforts. Understanding the biodiversity of the Madrean Archipelago will help highlight the uniqueness of the region and protect its species for future generations.

References

- Brooks, Thomas; Fonseca, Gustavo; Rodrigues, A. 2004. Species, data, and conservation planning. *Conservation Biology* 18 (6): 1682-1688.
- Brown, David E.; Lowe, Charles H. 1980. Biotic communities of the Southwest. Gen. Tech. Rep. RM-78 (map) [Revision 1983]. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experimental Station.
- Burgess, N.D.; Rahbek, C.; Williams, P.; Larsen, F.W.; and Balmford, A. 2002. How much of the vertebrate diversity of sub-Saharan Africa is represented by recent conservation proposals? *Biological Conservation*. 107: 327-339.
- Chapman, A.; Busby J.R. 1994. Linking plant species information to continental biodiversity inventory, climate modeling and environmental monitoring. In: R.I. Miller (editor) *Mapping the diversity of nature*. Chapman & Hall, London.
- Chen, G.; Peterson, A.T. 2002. Prioritization of areas in China for biodiversity conservation based on the distribution of endangered bird species. *Bird Conservation International*. 12: 197-209.
- Conservation International. 2011. Madrean pine-oak woodlands. Conservation International Website. Online: http://www.conservation.org/where/priority_areas/hotspots/north_central_america/madrean-pine-oak-woodlands/Pages/default.aspx. [2011].
- Csuti, B.; [and others]. 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation*. 80: 83-97.
- DeBano, Leonard H.; Ffolliott, Peter F.; Ortega-Rubio, Alfredo.; [and others], tech. coords. 1995. Biodiversity and management of the Madrean Archipelago: The sky islands of southwestern United States and northwestern Mexico. 1994 Sept. 19-23; Tucson, AZ. Gen. Tech. Rep. RM-GTR-264. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station. 669 p.
- Gilbert, Edward. 2011. Symbiota: Specimen-based virtual flora software platform. Unpublished presentation on file at: Sky Island Alliance office, Tucson, AZ. 45 p.
- Godown, M.E.; Peterson, A.T. 2000. Preliminary distributional analysis of U.S. endangered bird species. *Biodiversity Conservation*. 9: 1313-1322.
- Gottfried, Gerald J.; Gebow, Brooke S.; Eskew, Lane G.; Edminster, Carleton B., comps. 2005. Connecting mountain islands and desert seas: Biodiversity and management of the Madrean Archipelago II; 2004 May 11-15; Tucson, AZ. Proceedings RMRS-P-36. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 631 p.
- Heidorn, P.B. 2004. Publishing virtual floras and faunas. *Bulletin of the American Society for Information and Technology*. 30(2): 8-11.
- Higgins, S.I.; Richardson, D.M.; Cowling, R.M.; Trinder-Smith, T.H. 1999. Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. *Conservation Biology*. 13: 303-313.
- Jones, P.G.; Gladkov, A. 1999. FLORAMAP: A computer tool for predicting the distribution of plants and other organisms in the wild. Cali, Colombia: Centro Internacional de Agricultura Tropical.
- Kelly, C.; Garison, J.; Aggarwal, A.; Sarkar, S. 2002. Place prioritization for biodiversity reserve network design: A comparison of SITES and ResNet software packages for coverage and efficiency. *Diversity and Distributions*. 8: 297-306.
- Lobo, J.M.; Lumaret, J.P.; Jay-Robert, P. 1997. Taxonomic databases as tools in spatial biodiversity research. *Annals of the Society of Entomology of France*. 33: 129-138.
- Marshall, J. 1957. The birds of the pine oak woodland in southern Arizona and adjacent Mexico. *Pacific Coast. Avifauna* No.32.
- Marshall, R.M.; Turner D.; Gondor A.; [and others]. 2004. An ecological analysis of conservation priorities in the Apache Highlands Ecoregion. Prepared by The Nature Conservancy of Arizona, Instituto del Medio Ambiente y el Desarrollo Sustentable del Estado de Sonora, agency and institutional partners. 152 p.
- Mexican National Institute of Statistic and Geography (INEGI). 2005. *Uso del Suelo y Vegetación Dataset*. Scale: 1:250 000. Mexico City, Mexico.
- Peterson, A.T.; Navarro-Sigüenza, Adolfo G.; Hesiquio, B. 1998. The need for continued scientific collecting; a geographic analysis of Mexican bird specimens. *Ibis*. 140: 288-294.
- Peterson, A.T. 2004. Predictability of the geography of species' invasions via ecological niche modeling. *The Quarterly Review of Biology*. 78: 419-433.
- Peterson, A.T.; Egbert, S.L.; Sanchez-Cordero, V.; Price, K.P. 2000. Geographic analysis of conservation priorities using distributional modelling and complementarity: Endemic birds and mammals in Veracruz, Mexico. *Biological Conservation*. 93: 85-94.
- Peterson, A.T.; Soberon, J.; Sanchez-Cordero, V. 1999. Conservatism of ecological niches in evolutionary time. *Science*. 285: 1265-1267.
- Rebelo, A.; Siegfried, W.R. 1992. Where should nature reserves be located in the Cape Floristic Region, South Africa? Models for the spatial configuration of a reserve network aimed at maximizing the protection of floral diversity. *Conservation Biology*. 6: 243-252.
- Schnase, John L.; Kama, Deborah L.; Tomlinson, Kay L.; [and others]. 1997. The Flora of North America digital library: A case study in biodiversity database publishing. *Journal of Network and Computer Applications*. 1997(20): 87-103.
- Soberón, Jorge; 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(1444): 689-98.
- Soberón, Jorge; Arriaga, Laura; Lara, Liliana. 2002. Issues of quality control in large, mixed-origin entomological databases. In: *Towards a global biological information infrastructure*. European Environment Agency. 17-22.
- Soberón, Jorge; Davila, M.P.; Golubov, J. 2004. Targeting sites for biological collections. In: *Seed storage: Turning science into practice* (ed. H. Pritchard & M. Way). Richmond: Kew Royal Botanic Gardens.
- Soberón, Jorge; Llorente, Jorge B.; Oñate, Leonor. 2000. The use of specimen-label databases for conservation purposes: An example using Mexican Papilionid and Pierid butterflies. *Biodiversity and Conservation*. 9: 1441-1466.
- U.S. Geological Survey (USGS). 2002. Shuttle Radar Topography Mission (SRTM) Elevation Dataset [BIL raster images]. 90m spatial resolution. Sioux Falls, SD.
- Winker, K. 1996. The crumbling infrastructure of biodiversity: The avian example. *Conservation Biology*. 10: 703-707.
- Zhong, Y.; Luo, Y.; Pramanik, S.; Beaman, J.H. 1999. HICLAS: A taxonomic database system for displaying and comparing biological classification and phylogenetic trees. *Bioinformatics*. 15(2): 149-156.

The content of this paper reflects the views of the authors, who are responsible for the facts and accuracy of the information presented herein.