How Do Absence Data Predict the Performance of Species Distribution Models?

By
Rebecca Pederson
Ecological and Evolutionary Biology, University of Colorado at Boulder

March 31, 2015

Thesis Advisor:
Robert Guralnick, Ecological and Evolutionary Biology

Defense Committee:
Barbara Demmig-Adams, Ecological and Evolutionary Biology
Rebecca Dickson, Program for Writing and Rhetoric

Abstract

Species distribution maps are a fundamental data source for ecologists and evolutionary biologist that connect more broadly into policy, management, and society. These maps are not unfamiliar to the general public, often found in field guides and park brochures. Today, species distribution mapping and modeling can be done at much finer-scale resolution than previously, facilitated by new Geographic Information System (GIS) tools and techniques and the availability of species presence data and environmental data. However, most broad-scale species distribution models only utilize species presence data. Here, I explore how absence data can help improve species maps. The goal for my work is to show how species inventories can provide a means to generate probabilistic assessment of absences. In order to generate absences, I helped develop a means to assess the completeness of an "area inventory" (e.g. park or protected area surveys) by capturing metadata of published inventories. For those inventories assessed to be "mostly" complete, I compared the list of species present with a broader-scale list generated from range maps (utilizing a species list tool already available in Map of Life). If species were present in the Map of Life species list but not the inventory list, this was considered an inferred absence. Finally, I selected 25 species presumed to be absent based on my comparisons, to assess the performance of habitat models. I was able to document that these habitat models were generally making errors of commission (predicting suitable habitat where they should predict absence). Further refining these maps can further help (i) our knowledge in biodiversity, (ii) in conservation strategy for endangered species, which can lead to management and policy, and in (iii) educating the public about species and their distribution.

Keywords: species distribution maps, absence, presences, completeness, area inventory, biodiversity, conservation, education

Introduction

For my Honors thesis, I explored the question of how absence data may provide critical information for building better, fine-grain species distribution maps. Better maps not only contribute to conservation of biodiversity, but also help with policy, management, and education of the public. To address this question, I broke down the problem into four parts of an overall workflow. Each of these parts was challenging, requiring innovation and development of approaches without many signposts in the literature. The first piece of the puzzle is how to document absences in the first place. One of the few sources, where such information resides, is "area inventories", for which researchers assay an area and generate a list of species based on their sampling methodology. Published accounts of inventories provide many details about survey methods utilized for sampling, and I assessed what type of data is needed to capture how complete the survey is. The second challenge was determining criteria for inventory "completeness". A complete survey is one that samples all (or at least most) species in an area. After the "mostly" complete inventories were separated from the rest, the third piece of the puzzle was figuring out how to convert these completeness assessments to a list of vetted absences. The last part of my project used these absences to test how well existing deductive habitat models are performing, i.e. whether they are accurately predicting absence where inventories also document absences, or whether they are overpredicting presences. Together, the four parts of this project shed light on the big picture question.

Background

Conducting regional or local inventories is not a new idea. Pliny the Elder's *Naturalis Historia* included the first formal account of organismal diversity that dates back to the first century (Kohler, 2006). Since then, more precise and accurate cataloging of species regionally and globally escalated through two processes. The first of those was Carl Von Linne's formal naming of biological diversity, providing a *lingua franca* for describing biological diversity (his Systema Naturae published in 1735). The second was driven by the strong need to catalog natural diversity discovered by European scholars and explorers during empire building that extended well past homeland political boundaries (Abernethy, 2000). One of the best-known examples is Charles Darwin's cataloging of modern and fossil species during his voyages on the Beagle, and his special fascination with island biodiversity (Darwin, 1839). Today inventories are still routinely performed, utilizing an array of old and new inventorying processes. Most inventories, albeit not all, have a goal to completely record all species of the focal taxonomic group sampled. How well this goal is met can often be determined by assessment of the published inventory account. Assessment of inventory completeness forms the basis for my Honors thesis.

Inventory processes and scope

As easy as recording a species name for a sighting may sound, actual inventorying is often a complex process. First of all, it requires carefully planning to perform an inventory, irrespective of the inventory's spatial and taxonomic extent. Here, I focus on "area species checklists", which I define as an inventory over a relatively broad, but well defined taxonomic scope (e.g., mammals, birds, or butterflies) and geospatial scope (e.g., state or national park, protected area, or county). Scientific field research in such fixed areas typically involves employing different sampling strategies, with different strengths and weaknesses. Such general descriptions of checklist processes have yet to be formally made, although this is in progress through the development of the Biocollections Ontology (Walls et al., 2014). Below, I present some of these broad descriptions of taxonomic inventory processes as described in the latter ontology and used later in definitions of different inventory subtypes.

One inventorying process subtype is a "restricted search", involving searches along a restricted path, such as transects, or using plots of a specific length, often with the intent of measuring effort along the restricted sampling. Another type of search is an "open search", involving a defined area, but no systematic search pattern. Trapping (e.g., "pitfall traps") is a variant of a "restricted search", but with the intent of passively capturing individuals and measuring effort in terms of trap hours or trap nights. A final broad type of survey technique is "advantageous/opportunistic", in that surveyors notice a species when or where they do not expect it. This type of search may also detect species outside the targeted taxonomic group. All of these terms are further defined formally in the Biocollections Ontology (Walls et al., 2014).

In any one inventory, researchers may use one or more of the inventory processes described above. A full accounting of the details of those processes is as important as the process itself. One component of this is time spent; the more effort is put in, the more likely it is that the inventory is complete. In this context, completeness implies that the inventory yields a full list of species actually occurring in the area surveyed. A survey extending over several years, including capturing of organisms in different seasons, would likely be more complete than one conducted over only a few days or one season. Also surveying diurnally (over the course of the light period), nocturnally (over the course of the dark period), or over the entire diel period (24 hours per day) also changes completeness.

Other factors related to the scope of the survey are also critical for completeness. A survey is more likely to be complete if the focal taxon is more limited, e.g., a large carnivore versus a vertebrate. Finally, the geospatial scope of the survey is also important. Since most surveys take place in a national park, it can be challenging to access all areas due to lack of road and trails. Surveys in large parks are more difficult to complete than ones in smaller areas with less ground to cover and fewer habitats with habitat-specialist groups.

Re-compiling published knowledge and Map of Life

Although area species inventories are usually made available through the peer review literature or as grey (e.g., non-peer reviewed) literature. Such surveys collectively represent a huge human investment in understanding our planet and where species live. However, these data are also scattered, often poorly described, and therefore cannot easily be used together to provide rigorous indications of species presence, and more importantly for my thesis work, species absence. Measuring completeness of these inventories across space, environment, and time is crucial (Ruggerio et al., 1992, Sohlgren et al., 1995, Leon-Cortes et al., 1998) because it can be presumed that any species not on this list is absent from the area. However, because absence data are harder to substantiate than presence data, species distributional modeling is usually based solely on presence data (Soberon & Peterson, 2005). These presence-only models have been a tremendous advance in species distribution maps, but are also extremely challenging to validate, and prone to overpredict (show as too broad) the distribution of a species.

Map of Life (MOL) is a biodiversity knowledge platform developed by Yale University and the University of Colorado Boulder (Jetz et al., 2012). The goal of the latter project is to reassemble scattered data about biodiversity and use that information in new ways to enhance human knowledge of biodiversity on the planet. On the website, one can either search by "species" or "area of interest". Searching by species name provides all distributional data products MOL has assembled for that species. Such products include point occurrences (e.g., records from single locations usually from natural history collections or citizen science efforts), expert opinion range maps as might come from a field guide or threat assessment process, and most importantly for my work, an initial set of area species checklists (list of species found during an inventory). Map of Life also provides a tool for showing a list of species in an area of interest, but only at very broad spatial extents. Because these lists are based solely on expert-opinion range maps, the scale of such "presence lists" is necessarily very coarse, at a resolution ranging from 50-300 km depending on user-specified selection of extent. In terms of taxonomic scope, the list tool has global coverage for all vertebrate groups but reptiles, and is rapidly growing to include resources for plants and insects.

Map of Life is continuing work on assembling area species checklists from the literature and compiling their contents for integration with other datasets. A team of graduate and undergraduate workers at Yale and Boulder are gathering data lists from published inventories for upload to MOL. The Yale team is in charge of reviewing, scraping, and staging the checklist data for uploading to the maps. The Boulder team is in charge of the more challenging task of ascertaining how the inventory was performed, which represents the collection of metadata (as defined by Duval, 2002) about the methodology of the inventory process (i.e., how, when, and where data were collected, and what species were included).

A workflow for utilizing absences in modeling

Although species inventorying is not new, what is new is the approach towards reassembling this knowledge and attempting to standardize the way information is reported from species checklists. My thesis tackles a full workflow, which I briefly describe here and explain in more detail in the Methods. The workflow starts with assembling consistent reporting from species inventories. Next, I utilize Map of Life functions to assess possible species presences from the larger area to compare the presences lists between the inventory and Map of Life to assemble a list of putative absences. These putative absences need to be validated more fully, again referencing the methodology and scope of the checklist. Finally, I can create a list of final "high probability" absences. With these data in hand, I can test if habitat models built solely from habitat classifications are either accurate or potentially overpredicting suitable habitats. I use new Map of Life tools to show how this can be done, but do not to actually derive new metrics, leaving that to future work.

Methods

Step 1: Metadata assembly

Capturing metadata

Assembling a qualitative estimate of completeness across surveys must be done in a way that is repeatable and usable by others. It must be thorough enough for the information captured to allow an estimate of completeness and general enough to work across the wide variety of reporting styles found in such inventory reports. In best-case scenarios, surveyors would quantitatively report completeness, e.g., species accumulation curves (Colwell & Coddington, 1994), but this is most often not the case. We, therefore, developed a "species inventory metadata" schema usable for assembling standard contents from a large number of inventory reports. The schema provides the means to capture standardized content from inventory reports. Terms in the schema relate to both scopes of survey and methodology, with the goal of capturing a qualitative (but meaningful) assessment of completeness. If the survey is "mostly" complete, it becomes possible to determine which species are thought to occur in the area, but were not directly reported in the survey. Those absences are inferred absences, and yet further vetting is then needed to confirm these inferred absences and justify their use in downstream analyses.

My mentor, three other undergraduates, and I spent the last year collaboratively developing a metadata capture process. We focused on key aspects of temporal, geospatial, habitat, taxonomy, inventory methodology, and measures of completeness (Figure 1). The full methodology is being developed as a separate paper for formal publication (Guralnick et al. in prep), but in a nutshell, when reading through inventories, we capture an extensive set of content with the goal of rating the inventory based on completeness. Appendix A provides a full term list we use for capturing inventory metadata. Those terms become column headers for a Google Spreadsheet (available for viewing: https://docs.google.com/spreadsheets/d/1sCsE64VPWwPLCJdgyvN561Xw5Wazk6O0M18e_VHf2K0/edit?usp=sharing). As is clear from the term list and Google Spreadsheet, metadata capture is relatively extensive although not exhaustive. The team capturing metadata assesses a publication describing an inventory, reads the contents, and from the written content enters information on methodology, scope, etc., to make an evaluation of completeness. Appendix B provides the manual used when capturing metadata.

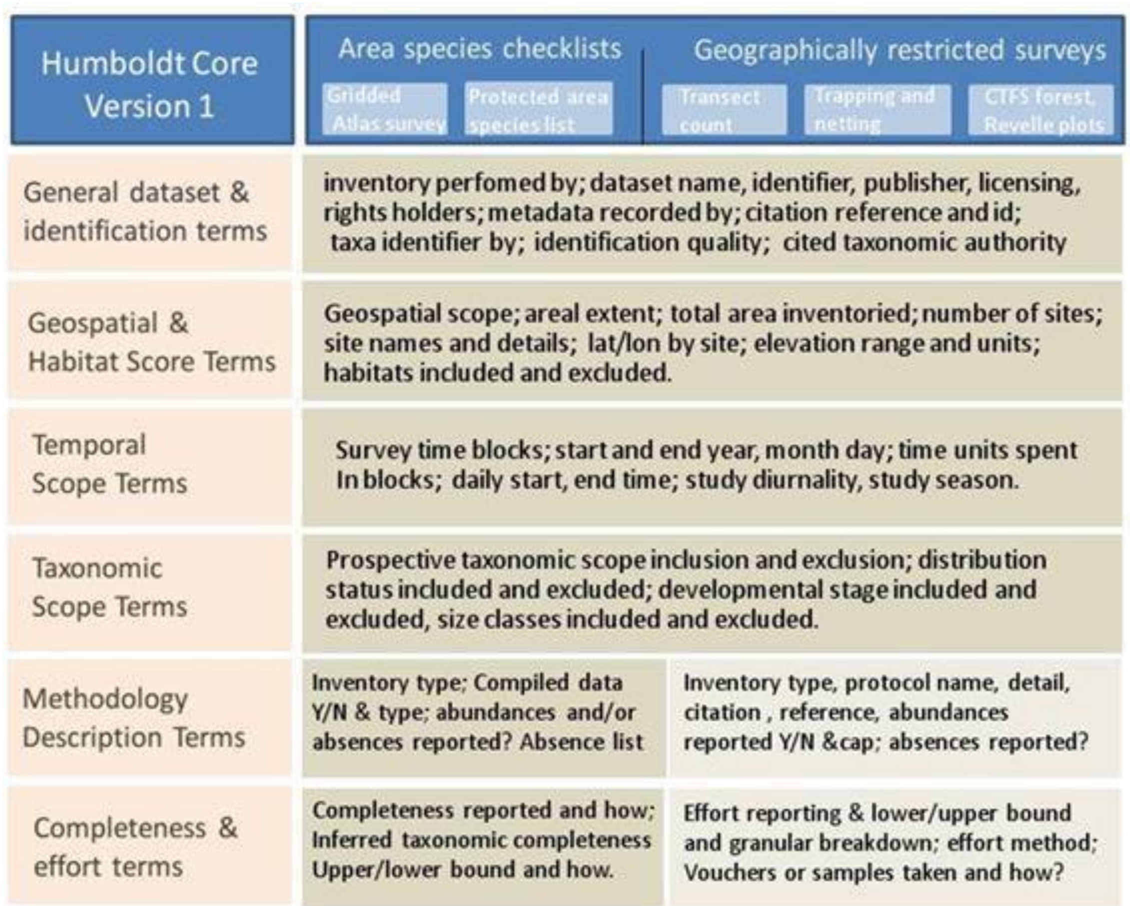

Figure 1. This figure summarizes the different aspects of capturing metadata.

Assessing inventory completeness

Perhaps the most critical and difficult assessment is that of survey completeness. Typically, inventories with higher completeness scores provide extensive information about methodology, including that inventorying processes conducted (e.g. restricted and opportunistic searches, etc.). Many inventories also bring together expert knowledge of the area separate from new inventory processes, including primary literatures, local experts such as park rangers, and data from museum collections. Taken together, it has been a major effort to create appropriate terms to capture and work through 143 individual area species checklists to capture data.

Survey completeness ranges from nearly incomplete (e.g., most focal taxa were missed) to 100% coverage of all target species. In the vast majority of cases in our test dataset of 143 surveys, authors do not directly quantify overall effort. In a few cases, authors provided a description of completeness themselves, based on other available evidence (e.g., previous work) and their own assessment of their methods. The more common case is for authors to provide no assessment of completeness at all, in which scenario some broad estimate can still be determined given the other metadata collected, e.g., the measures of scope of inventory and survey and compilation effort.

We chose to qualify completeness as a measure with lower and upper bounds that ranged between 0 and 100 in 25 percent increments. We recognize this is a rough estimate of completeness, but finer measures would necessitate a more formal model specification. The purpose here is rapid assessment of completeness for determining the likelihood of being able to document species absences. For every study, those entering data assigned a completeness range.

A range from 75-100 suggests that the survey yielded a nearly complete assessment of species presence for the region and taxon of interest. Such a score was given in cases, where the geospatial scope was relatively limited in areal extent; the inventory effort was well quantified, reported, and extensive enough given the biology of the focal taxon to credibly reach saturation in accumulation of species. Credibility increases, where habitats in the geographic scope were well sampled over different time periods, and compilation effort was high.

Scores between 0-25 indicated that the survey was clearly incomplete. Although such incomplete surveys are rare, we did find cases of rapid, short surveys, where the goal was to develop a first assessment of species in an area. In a few cases, reporting was so poor and the methods clearly inadequate to warrant classification as a nearly incomplete assessment. In many cases, values were broader, e.g., 25-75, which reflects higher uncertainty about assessing completeness due to paucity in reporting by the authors.

Assembling complete surveys, documenting initial absences and then further vetting absences

I first assembled species lists from mammal inventories that were nearly complete (e.g., 75-100 complete). I chose mammals because they are particularly well studied, have the fewest issues with out-of-date species names, and were generally more tractable in terms of number of species compared to, e.g., bird inventories. Since the absence assessment and use in modeling is meant to be a test case as opposed to a full production workflow, I limited myself to examining 10 near-complete mammal inventories.

To initially document absence, I either utilized digitized lists already scraped from text by the Yale team, or digitized those on my own. After digitization, these lists of species were imported into Excel. I also assembled species lists based on mammal range maps using Map of Life's species list tool. To do this, I selected the same location as the centroid of the "area inventory" and asked for mammal species with 50, 100, or 300 km radius. This creates a "coarse-grain" list of all species found within a 300 km buffer of the centroid (Figure 2).

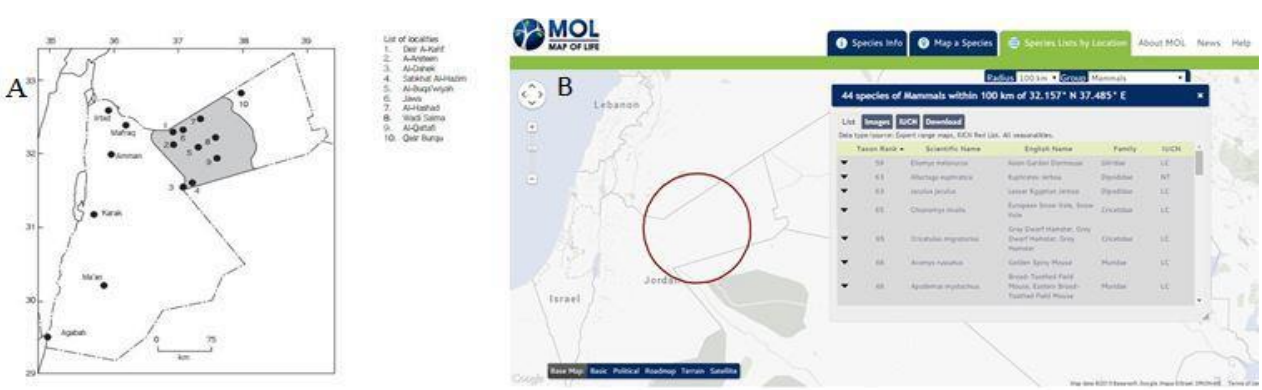

Figure 2. This figure shows the geospatial area, Badia, to generate both the Inventory list and the MOL list. A). Map provided in the Inventory (Bunaian et al. 2001). B). The process of generating the MOL list of species.

The next step was comparing the two lists with each other. If the MOL list showed that a species was present within that area, and a relatively complete survey did not have that species on its list, this was considered a presumed absence. In order to help automate this process, I set up a pivot table in Excel to compare the lists. There were three possible outcomes of these comparisons. The first outcome was that the species is present on both lists (Figure 3A). The second outcome was inferred absences. Inferred absence within the extent of the survey is defined as a species being on the MOL list, but not on the inventory list (Figure 3C). The third category was omission error. This is defined as the inventory list including a specific species that is not represented in MOL (Figure 3B). I usually went back and checked the range map on MOL. In such cases, I typically found a problem with Map of Life range products, which was important information in and by itself. Sometimes, it was a spelling error of the species name, or the species had two or more scientific names. Usually, the omission error was due to underprediction in the range map of MOL.

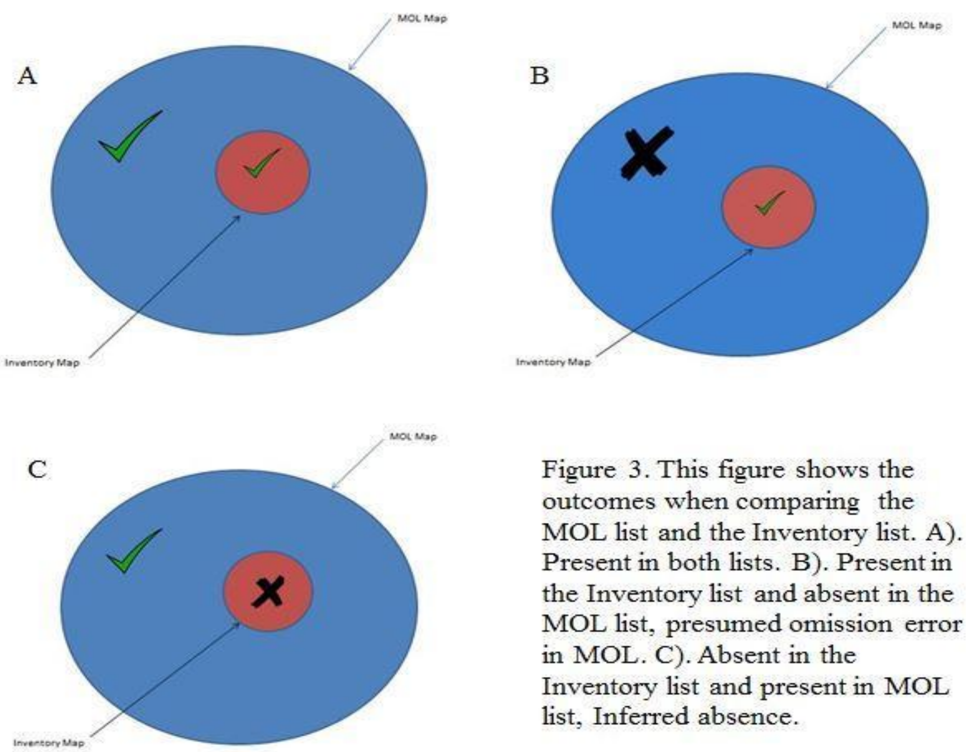

Figure 3. This figure shows the outcomes when comparing the MOL list and the Inventory list. A). Present in both lists. B). Present in the Inventory list and absent in the MOL list, presumed omission error in MOL. C). Absent in the Inventory list and present in MOL list, Inferred absence.

Further Vetting of Absence Data

After comparing these inferred absence lists, I further vetted the absences. This required learning more about the species on the putative absence list, especially their higher-classification affiliations (e.g., carnivore versus rodent, etc.), while also rechecking the inventory. In many cases, the species on the putative list were not part of the taxonomic (or body size, etc.) scope of the survey. Sometimes, initial metadata on methodology did not fully capture taxonomic scope. For example, bats can only be captured with specific methods like mist nets or equipment that records bat calls, and even if a survey claimed to capture “mammals”, many clearly implicitly excluded flying mammals. As another example, some camera trapping methodologies were biased towards larger animals. Yet another inventory was limited to ground searching and could not access canopies, which led me to eliminate animals known for being arboreal. In cases where I felt that the survey methods could not have captured a species, I removed those species from the inferred absence list. Those species that could potentially have been inventoried based on a finer reading of the methods, but were not found, was placed onto the final list of inferred absences. Some lists were easier than others to narrow down. I was able to narrow some lists down to having less than 10 absent species, while other lists exhibited absences in the hundreds. Appendix C presents a list of rationales behind every list of inferred absences that was vetted, and how I narrowed each list down.

Finally, I also went through both inferred absences and those that appeared to be omission errors in Map of Life, and checked for misspellings or different scientific names. Since misspellings usually occurred by only a few letters, it was not difficult to find the correct spelling. Sometimes, area inventories used a name for a species that had been synonymized or otherwise changed, versus the name used in Map of Life. In those cases, reconciling names was important.

Using modeling for absences

After generating a refined list of absences, I tested models of species distributions generated using a deductive method based on species habitat characteristics. The deductive approach linked habitat preference data for species to habitat classifications from land-cover maps to delineate areas unsuitable for a given species. Within the known range borders of the taxon, those habitat classifications that did not match the known habitats of the species were "clipped" or recorded as presumed absences. Often used, this modeling approach differs from more inductive, correlative approaches (Elith & Leathwick, 2009), that rely on a series of assumptions about lack of bias and sufficient coverage of the input data.

Habitat maps produced by these deductive models are commonly tested using only presence records. In this case, presence records are plotted onto the map, thus assembling a simple matrix where the presence record and model either agree or disagree (disagreement are cases where the model predicts absence despite the presence record confirming presence). Critically missing is assessment of model "overcommission" or overpredicting presence.

Absence records promise to provide a means to assess model commission errors (e.g., overprediction as discussed in Václavík & Meentemeyer, 2009). I picked 25 species from my 10 inferred absence list to test the models. For these 25 species, I examined deductive "clipping" models, utilizing a new tool available on the Map of Life platform (currently in beta release (see: http://species.mol.org/pa for more details). I assessed whether or not areas predicted as absences agreed with the absence data I collected for those species. This allowed me to assess model overprediction or “commission”.

Results

Metadata assembly

Critical to all other steps of this project was establishing a common protocol for capturing inventory metadata. Appendix A shows all terms developed for capturing metadata, while I focus on the key findings in the text. One of the most critical aspects of data collection was establishing completeness metrics based on different components of scope and methodology of the inventory. Table 1 summarizes the final rubric we used to establish completeness, and is based on discussion of completeness in the Methods. There are “corner cases” for all these assessments, requiring some flexibility in application of the rubric, which is why each completeness assessment features notes explaining the rationale.

Table 1. A general rubric for assessing completeness on a scale of 25% percentiles.

Completeness	Geospatial	Temporal	Methodologies	Compiled	Taxonomy
0 - 25	Large area with only a few sites.	Short period of time.	Used only one time of methodology.	Use little or no primary literature.	Multiple taxa groups
25 - 50	Large area with multiple sites.	A short period of time, indication of days spent surveyed.	Used a few different types a methodology, but did not report their effort.	Used sources of primary resources.	Large taxa group
50 - 75	Small area with a few sites.	A longer period of time, with small amount of days spent. Did not spend both day and night surveying.	Used a few different types of methodology and reported some of their effort.	Only used one type of source other than primary literature.	Single taxa group
75 - 100	Small area, with multiple sites.	Surveyed over a longer period of time, with complete days of surveying. Surveyed both day and night and different seasons.	Multiple methods used. Specific methods for hard to find species. Calculated the amount of effort put into searching.	Used primary literature, expert knowledge, and/or museum collections.	Narrow taxa group

We captured metadata from 143 inventories from around the world and focused primarily on terrestrial vertebrates. By compiling data from this relatively significant sample size of inventories, we can assess some summary measures across all inventories. For example, one key aspect of inventories is whether the inventory used already compiled resources (e.g., expert opinion, literature, museum specimens) and the extent of compilation effort (measured as low, medium and high). While few inventories make significant compilation efforts (e.g., multiple sources), almost 40% did collate some already compiled information (low compilation effort). Similarly, over 40% of inventories we examined directly reported absences. Finally, we calculated the percent of surveys that were nearly complete (e.g. 75-100) versus other categories. Table 2 summarizes our findings from 129 terrestrial vertebrate surveys, and how these ranged in completeness. As can be seen, only 27% of surveys were categorized as nearly complete, and of the rest, only another 30% were 50-100 complete. This was a sobering result indicating the limits of inferring of absence from published inventory data.

Table 2. This table summarizes the distribution of 129 inventories.

Humbolt Core Term	Possible Values	Percents
Compilation effort	Low, medium, high, n.a	Low – 38.7%, Medium – 6.2%, High – 8.5%, n.a – 45.7%
Absences reported	Yes/No	Yes-41.8 , No-58.2%
Completeness assessment	Scale in 25 percent increments from 0-100	50-100% complete – 30.3%, 75-100% - 27.9%, other- 41.8%

We also sorted our inventories into regions, to give a sense of the distribution of our data sources (Table 3) and by taxon (Table 4). More effort was placed on accumulating inventories from areas, where other sources of data were lacking, which explained the concentration of inventories of Asia and Africa. These are areas, where biodiversity knowledge is most limited (Meyer et al., 2015).

Table 3. The distribution of inventories by geospatial focus. For example, of the 143 inventories from which we captured metadata, 49 had Africa as their geospatial foci.

Location	Inventory Count
Asia	35
Europe	9
Africa	49
Central America	2
South America	14
North America	9
Australia/Oceania	6
Middle East	19
Total	143

Tables 4. The distribution of inventories by taxonomic focus. For example, of the 145 inventories from which we captured metadata, 37 had mammals as the taxonomic foci. Some inventories had multiple taxa. Others focused on a more exclusive taxon unit (e.g. bats). We use the term "*sensu latu*" for broad taxonomic groups and "*sensu strictu*" for more exclusive units.

Taxonomy	Count
Mammals (sensu latu)	37
Reptiles	11
Amphibians	7
Herps (sensu latu)	39
Birds	30
exclusively Bats	3
Fish	11
Butterflies	2
exclusively Carnivores	1
Vertebrates (sensu latu)	1
Flora	1
All taxa	2
Total	145

Determining inventory absences

Table 5 provides a summary of initial and final vetted absences per the 10 focal mammal checklists that I used for absence assessment. The critical take home message here is that initial assessments of absences simply from list comparisons were clearly insufficient to provide a rational assessment of absences. Instead, further vetting was required to find high probability absences. For example, the list "Termessos" (De Marinis & Masserti, 2009), which refers to one of the inventories performed in Turkey, had 44 presumed absences based on comparison of Map of Life and inventory lists, but had to be winnowed down to 6 absences after further vetting the lists.

Table 5. This table shows the distribution of species in each specific list in each area. For example, Jabal had 12 species from the inventory list and 64 from the MOL list. From both of those lists, 54 were considered inferred absences, then after refined the list was 13.

List	# of Taxa in Inventory	# of Taxa in MOL	# of Presumed Omission	# of Initial absences	# of Refined absences
Horshedhden (Nader et al., 2011)	15	59	4	47	6
Gonja (Williams et al., 2000)	15	198	2	185	113
Cuzcoamazonico (Woodman et al., 1991)	135	234	57	154	44
Jabal (Abi-Said and Zuhair, 2012)	12	64	2	54	13
Seven Mile Camerong (Murphy, 1998)	19	50	5	36	26
Virginia Island Barrier (Dueser et al., 1979)	13	40	4	30	17
Hautniger (Ziegler et al., 2002)	97	135	29	65	52
Termessos (De Marinis and Masserti, 2009)	15	57	1	44	6
Badia (Bunaian et al., 2001)	10	44	8	42	1
Bale (Asefa, 2011)	61	113	15	65	63

Detecting overprediction in modeled habitat maps

When comparing the 25 exemplar absences to habitat maps on MOL, I looked up the area of each species to assess how well the map was predicating the absence, utilizing three categories, i.e. overprediction, “corner cases”, or true absence (described below). Overprediction was described as the map covering areas, where the species is supposed to be absent (Figure 4A). This overprediction can be broken into two different scenarios. The map was either fully covering the area of the absent species, or the map was partially covering the area of the absent species. The latter case involved some pixels missing in the area, but not excluding the whole area. The other category found was “corner cases”, which typically involved some issue with habitat maps range edges being right at the spot of a survey as shown in Figure 4B. The latter scenario represents true absence, where it is clear that the map is accounting for the absence because the area of the species is excluded from the map as in Figure 4C. Out of 25 range maps where we examined absences from inventories, 14 showed overpredictions, eight were corner cases, and three true absences. Table 6 summarizes all 25 species. By omitting the corner, 14 out of 17 (82.35%) maps were overpredicting the range of these absent species.

Figure 4. This figure shows the three possibilities of the performance of the habitat map through absence data. A). This habitat map is overpredicting the range of the Ursus arctos or the brown bear. B). The habitat map is considered a “corner case” because the habitat range was on the edge of the survey area. C). This habitat map is (mostly) correctly accounting for the absence of the Capreolus capreolus or the roe deer.

Table 6. This table shows the absent species and how it predicts the performance of the habitat map. For example, the habitat map is overpredicting the distribution of the Lynx lynx because the coverage is fully covering the geospatial scope.

Scientific name	Common name	Location	Type of Coverage
Vomela perejusna	Marble Pole Cat	Badia	Corner Case
Lepus Capensis	Hare	Jabal	Partial
Caracal caracal	Desert Lynx	Horshehden	Fully
Lynx lynx	Lynx	Termessos	Corner Case
Capreolus capreolus	Roe Deer	Horshehden	True absence
Heliophobious argenteocinereus	Mole rat	Gonja	Partial
Synaptomys cooperi	Bog Lemming	Virginia Barrier Island	Partial
Sminthopsis leucopus	Dunnart	Seven Mile Comeroong	Corner Case
Potorous tridactylus	Rat Kangaroo	Seven Mile Comeroong	Partial
Tragelaphus eurycerus	Bongo Antelope	Hautniger	Corner Case
Ursus arcatos	Brown Bear	Termessos	Partial
Hyaena hyena	Hyena	Termessos	True absence
Glauconycteris variegata	Bat	Hautniger	Partial
Choeropsis liberiensis	Hippo	Hautniger	Partial
Herpestes sanguienus	Mongoose	Bale	Fully
Sylvicapra grimmia	Duiker	Bale	Fully
Hystrix indica	Porcupine	Jabal	Partial
Dendrohyrax arboreus	Hyrax	Gonja	Fully
Heliosciurus undulatus	Squirrel	Gonja	Fully
Herpestes ichneumon	Mongoose	Horshehden	True absence
Canis lupus	Gray Wolf	Horshehden	Fully
Glaucomys volans	Flying Squirrel	Virginia Barrier Island	Corner Case
Didelphis virginiana	Opossum	Virginia Barrier Island	Corner Case
Neotomys ebriosus	Mouse	Cuzcoamazonico	Corner Case
Lagidium peruanum	Ciscacha	Cuzcoamazonico	Corner Case

Discussion

My work takes existing, but disaggregated knowledge about biodiversity and especially species distributions, and reassembles that knowledge into a new framework for use in species distribution modeling. Many aspects of this work are completely novel, from standardizing reporting of inventory scope, methods, and completeness to the use of absences in modeling frameworks. Reading through all inventories allowed me to create standardized reporting that works across nearly all published area checklists and that can be used to perform assessments of completeness. These metadata can be associated with species lists made available via Map of

Life. This research could also encourage inventory authors to utilize the same metadata form in their published work. With inventory standards increasing, the more complete surveys can be used to assess species absences.

The main take-home message of my work is that habitat models appear to be overpredicting suitability. Using both absence and presence data to assess map quality provides much better assessment capacity, even if absences are limited to only one or two records for a species. My project has shown that even though absence data are difficult to use due to their probabilistic element, where proving a negative is much harder than proving a positive, using absence data is not impossible. When taking the time to establish which species are truly absent, and when used properly, absence is a great additional factor to assess range maps because a species absent in an area is as important as a species present. Absence data can further our knowledge of species distributions because there are reasons why a species is not present, like habitat, human presence, or failure of the species to establish itself in that area. The end result is broadly useful for those interested in overcoming the Wallacean Shortfall (Whittaker et al., 2005), which describes the limitations arising from a lack of complete knowledge of species distribution.

Mapping species distributions is important work for many different reasons. One is to contribute to biodiversity knowledge. These maps are the basis for many assemblage-level characteristics, such as species richness (the number of different species in an area) and species turnover (Hurlbert & Jetz, 2007). Knowing biodiversity in an area itself relates to key functions of ecosystems; more biodiversity often means more functional types and more production, like primary production or decomposers (Schwartz et al., 2000).

This brings up another application of knowing where a species is present, i.e. conservation efforts. Many species are threatened to extinction due to habitat loss, invasive species, overharvesting, pollution, disease and parasitism, and cryptic habitat (Hogan et al., 2010). Knowing where species are most common and in what kinds of designated protected areas, then it makes easier ways to establish to preserve them, especially if they are endangered. This can help make establishing management and policy easier. If an endangered species is known to be present in an area, and a company wants to build a new building there, evidence is available to show that they would be violating the Endangered Species Act if they build there. Policy can be applied to the area, where species are at most risk for extinctions.

Another importance is that maps like MOL can be used as an education tool. MOL is free on the internet, where anyone can access it. This knowledge is not limited to scientists, but anyone interested in the range of a certain species can find out. People using the website are not limited to looking at one species at a time. There are different functions, where the user can search for a list of species in a specific area or can view species ranges based on suitable habitat. MOL can also be used for education on species distributions for children to get them excited about different kinds of species and their ranges. Introducing children to species distributions increases the chance of them caring about species when they get older. It is a way to help educate future generations.

This project had many aspects that not only allowed me to assess inventory data to establish a norm for completeness, but also showed me the benefits of using absence data to test the performance of range maps. However, I did feel limited with my work. First of all, MOL was limited in a few different aspects. I was unable to gather absence data for reptiles because MOL only has data on these taxa for North America and not the rest of the world. This limitation prevented me from using herpetological (amphibians and reptiles) surveys, which were the most surveyed taxon. Inventory completeness also posed a challenge for me. Not only did it take a great amount of time to go through each inventory, collect the metadata, and assess completeness, not every inventory I read was complete. Of 143 surveys, only 39 could be considered complete. Due to time limitations, I was only able to compile 10 inferred species lists from the inventory and MOL, and had to focus only on 25 absent species to predict the performance of the habitat models. I may have been biased when I compiled the inferred absence lists because I only picked mammals. I did this because mammals were easier to narrow down in my search due to the methodology surrounding them. I am also more familiar with mammals than any other taxon groups I examined. Most of the mammal inventories also had narrow search criteria, like carnivores or small mammals, which made it easier to further narrow down my inferred absence lists.

If I were to continue my project, I would compile more complete species inventory lists. I would also produce more inferred species lists from species of other taxon. I did not have enough time to explore the diversity of the other taxa. In continuing my work, I would apply the knowledge of these absences to the range maps. By including my absence data with the presence data, range maps on MOL would be more precise. My project is only a sample of the kind of work that can be done with absence data creating more accurate species distribution maps. A key next step I did not have time to implement is utilizing the absence record to further score “map improvement” based on the initial habitat modeling. Map of Life has a metric for such improvements based only on presences. Cases where the models correctly predict absence would suggest no further improvement, while those where there is overprediction suggest a need for further map refinement. I also wanted to further explore how to reduce suitable habitats from the initial maps using more restrictive habitat classifications in order to see how much these values would need to be restricted to reduce overprediction.

The end result of my work is a full workflow for how to use species inventories and their ability to help improve model performance in the context of a large-scale modeling and infrastructure project. The work, although broad, leverages my long-term investment in the project and promises to show a new path for reusing knowledge for better documenting biodiversity.

References

Abernethy, D. B. (2000). The dynamics of global dominance: European overseas empires, 1415-1980. Yale University Press.

Abi-Said, M. & Amr, Z. (2012). Camera trapping in assessing diversity of mammals in Jabal Moussa Biosphere Reserve, Lebanon. *Vertebrate Zoology*,*62*(1), 145-152.

Asefa, A. (2011). Mammals of the Bale Mountains National Park, Ethiopia: A Compiled and Annotated Checklist, in Walia, special edition on Bale Mountains.

Bunaian, F., Hatough, A., Ababaneh, D., Mashaqbeh, S., Yousef, M., & Amr, Z. (2001). The carnivores of the northeastern Badia, Jordan. *Turkish Journal of Zoology, 25*(1), 19-26.

Colwell, R. K. & Coddington, J. A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *345*(1311), 101-118.

Darwin, C. (1839/2002) The Voyage of the Beagle. Dover Publications, New York.

De Marinis, A.M. & Masserti, M. (2009). Mammalian fauna of the Termessos National Park, Turkey. *ZooKeys 31*, 221-228.

Dueser, R.D., Brown, W.C., Hogue, G.S., McCaffrey C., McCuskey, S.A., & Hennessey, G.J. (1979). Mammals on the Virginia Barrier Islands. *Journal of Mammalogy, 60*(2), 425-429.

Duval, E., Hodgins, W., Sutton, S., & Weibel, S.L. (2002). Metadata principles and practicalities. *D-Lib Magazine*, *8*(4). Available: <doi:10.1045/april2002-weibel>.

Elith, J. & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677.

Hogan, M., Caley, K. J., & McGinley, M. (2010). Causes of extinction.*Encyclopedia of Earth, Washington, DC: Environmental Information Coalition, National Council for Science and the Environment,[Available at] http://www. eoearth. org/article/Causes of extinction*.

Hurlbert, A. H. & Jetz, W. (2007). Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proceedings of the National Academy of Sciences*, *104*(33), 13384-13389.

Jetz, W., McPherson, J. M., & Guralnick, R. P. (2012). Integrating biodiversity distribution knowledge: toward a global map of life. *Trends in Ecology & Evolution*, *27*(3), 151-159.

Kohler, R. E. (2006) All Creatures: Naturalists, Collectors and Biodiversity, 1850–1950, Princeton, NJ: Princeton University Press.

Linnæus, C. (1735). Systema naturæ, sive regna tria naturæ systematice proposita per classes, ordines, genera, & species. Lugduni Batavorum.

León-Cortés, J.L., Soberón-Mainero, J., & Llorente-Bousquets, J. (1998). Assessing completeness of Mexican sphinx moths inventories through species accumulation functions. *Diversity Distributions. 4,* 37-44.

Meyer C., Kreft H., Guralnick R. P., & Jetz W. (2015). Global priorities for an effective information basis of biodiversity distributions. *PeerJ PrePrints* 3:e1057.

Murphy, M. J. (1998). Mammal survey of Seven Mile Beach National Park and Comerong Island Nature Reserve on the south coast of New South Wales. *Australian Zoologist 30*(4), 419 - 425.

Nader, M. R., El Indary, S., Salloum, B. A., & Dagher, M. A. (2011). Combining non-invasive methods for the rapid assessment of mammalian richness in a transect-quadrat survey scheme–Case Study of the Horsh Ehden Nature Reserve, *North Lebanon ZooKeys*, (119), 63.

Ruggiero, M., Stohlgren, T., & Waggoner, G. (1992). Towards a biological survey of the U.S. national park system. In: Developments in Landscape Management and Urban

Planning, 7, Science and the Management of Protected Areas, J.H.M. Willison, et al. (eds.). Elsevier, Amsterdam. pp. 31-35.

Soberón, J., & Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics 2*, 1-10.

Stohlgren, T. J., Quinn, J. F., Ruggiero, M., & Waggoner, G. (1995). Status of biotic inventories in U.S. National Parks. *Biological Conservation, 71*, 97–106.

Schwartz, M. W., Brigham, C. A., Hoeksema, J. D., Lyons, K. G., Mills, M. H., & Van Mantgem, P. J. (2000). Linking biodiversity to ecosystem function: implications for conservation ecology. *Oecologia*, *122*(3), 297-305.

Václavík, T., & Meentemeyer, R. K. (2009). Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions?. *Ecological Modelling*, *220*(23), 3248-3258.

Walls, R., J. Deck, R. Guralnick & 22 other coauthors [listed in alphabetical order]. 2014. Semantics in Support of Biodiversity Knowledge Discovery: An Introduction to the Biological Collections Ontology and Related Ontologies. *PLOS ONE 9*(3): e89606.

Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R.J., Watson, J.E.M., & Willis, K.J. 2005: Conservation biogeography: assessment and prospect. *Diversity and Distributions 11*, 3-23.

Williams, S. T., Goodman S. M., Kihaule P. M., & Howell K. M. 2000. A Survey of The Small Mammals of the Gonja Forest Reserve, Tanzania. *Journal of East African Natural History 89*: 73-83.

Woodman, N., Timm, R. M., Arana C. R., Pacheco, V., Schmidt, C.A., Hooper, E.D., & Pacheco, A. C. (1991). Occasional Papers of the Museum of Natural History, The University of Kansas, *145*, 1-12.

Ziegler, S., Nikolaus, G., & Hutterer, R. (2002). High mammalian diversity in the newly established National Park of Upper Niger, Republic of Guinea. *Oryx*, *36*(01), 73-80.

Appendix A: Terminology

Area checklist usually involve a wide range of different inventorying processes. They also include not just direct			
observational methods but also compilation efforts from other authoritative sources The heterogeniety in			
methods makes it challenging to develop a single standardized form for either initial entry or retrospective data			
capture. If the intent is to capture some assessment of how complete an inventory may have been, it is possible to			
capture enough content to make that assessment repeatable and tractable. The terms below have been vetted for			
their use in just such an approach. This version is meant to be used for capturing data from published checklists.			

Metadata Profile Area Checklists

Item	Mapping to Existing Vocab	Description	Multiple Entries?
General			

Information			
Dataset Name	dwc:datasetName	Specified by data entry person. Title of paper in most cases.	
Dataset Identifier	dwc:datasetID	A unique identifer automatically filled out for the record when ingested to MOL	
Citation reference	eml:citation	The APA formatted reference to the published description of the checklist.	
Citation identifier	dc:identifier	A reference identifer. DOI is preferred but a URI or ISBN number is adequate	
CitationURL		A full URL to the cited document, if available.	
Inventory perfomed by	dwc:recordedBy	The person or people who performed the inventory, if reported	
Metadata recorders	dc:creator	Names of the data entry recorders and validators for this dataset	Y
Taxonomic Authority			
Taxa identified by	dc:identifiefBy	The agent(s) who performed taxonomic identifications in the field, if reported	Y
Citation reference of taxonomic authority	eml:taxonomicSystem	The published reference to the taxonomic authority used in the checklist	Y
Site details			
Site details		Author reported site information, for whole site or subsites	Y
Geospatial information			
Geospatial scope	dc:spatial, eml:studyAreaDescription	The region of the checklist as described in text form (e.g. "Kruger National Park" or "Alberta")	
Total area covered by geospatial scope		Author's reported area of the geospatial scope in km^2	
Temporal Scope			
Start year	dwc:year	The year surveying' began	
Start month	dwc:month	The month surveying began	
Start day	dwc:day	The day surveying began	
End year	dwc:year	The year surveying ended	
End month	dwc:month	The month surveying ended	
End day	dwc:day	The day surveying ended	
Number of (time units) spent surveying in each block	o&m:phenomenonTime	The number of (months, days, hour)s spent in each time block.	
Study diurnality		Author reported sampling during day, night, or both?	
Study seasons		Did the author report seasons surveyed? Which ones (from a pick list).	
Organismic information			

Prospective taxonomic scope	eml:TaxonomicCoverage	Set of searched species during inventory	
Methodology			
Was compiled data also collected?		Compiled data is ancillary data that was not directly collected by authors but assembled from other sources.	
(if yes), Type(s) of compilation?		Inventories can be compiled only, survey only, or both. The compilation types include expert and local knowledge, museum specimens, literature, or other sources.	
Inferred compilation effort		Measured as "low, medium and high" and discussed in more detail in paper text.	
Inventory type	bco:taxonomicInventoryProcess	Inventory type can be restrictired search, open search, opportunistic search, trap, incidental or adventitous or material sample	Y
Protocol detail		Any variations in the published protocol. For compilations that don't have a standard protocol or unpublished protocol, details presented here.	Y
Were absences reported?	dwc:occurrenceStatus	Did the authors report species expected in the survey that were not found?	
(If yes) Taxonomic list of reported absences	dwc:Taxon	List of taxa reported absence at whatever taxonomic level reported	
Completeness			
Was taxonomic completeness reported?		Choices for this field: "No, not reported", "Yes, reported as complete", "Yes reported as incomplete".	
(If reported incomplete), Did reporter assess how incomplete?		If author reported their assessment of incompleteness, capture author reported measurement.	
Inferred or reported taxonomic completeness , lower bound		Inferred lower bound completeness is captured in quartiles (0,25,50, 75,100). Lower values represent incomplete sampling.	
Inferred or reported taxonomic completeness, upper bound		Inferred upper bound completeness is captured in quartlies (0,25,50,75.100).	
How did you infer completeness?		A description of the rationale for the inferred completeness assessment often based on how well sampling was done across the geospatial scope, and the survey and compilation effort.	

Appendix B: Manual

The species inventory metadata capture form V2.0

The goal of the Google Doc is to capture key information that is reported by surveyors when they perform a survey for species. Such reports provide critical information on not only a list of species found, but the taxonomic, geospatial, habitat and temporal scope of the survey, the methodologies employed, and ultimately an assessment of completeness of the survey. The form has been tuned to capture these essential details and is not meant to be an exhaustive capture device. It also attempts to standardize something that is quite heterogeneous in terms of reporting quality, style and content. Thus we expect that you will have a lot of fields filled in with “n/r’s” - which stands for non-reported. The best surveys will have many fields filled in.

What role should guessing or assumptions play in the data entry process? In developing this form, we decided to not complicate data entry tasks by asking those entering data to also analyze and infer too much from the reports. This would require an even further level of training and would conflate data capture and downstream analysis in bad ways. We therefore came to agreement that the right approach here is to is to make sure that we treat this form as a verbatim capture device except in two places, and make little to no inferences except in some key places (e.g. total area covered, measures of effort) .

The two places where data entry people are expected to make assessments are in inferring compilation effort and overall completeness. Those are discussed in more detail below. Finally, Version 2.0 is a much smaller set of fields to capture compared to the first set of versions. This second set might be considered a more focused capture process.

Please enter something in every column of the form. If something was not reported then put n/r or if something was not available, i.e. DOI numbers, then put n/a.

General Information

Dataset name

- Use the title from the study as it was published

Dataset Identifier

- For this one we just need to put the name of the place where the study was done, and the taxon. For ex: “South Jordan Herps”
- You should get this one from the Priority Reserve Checklist (that way we can keep track of it).
- Make sure you don’t just put the country because there could be more studies done in that country. Be as specific as possible.

Citation Reference

- This is the citation for the published paper.
- We need to be consistent when entering. So APA style has been chosen to use. 1) the first author has first initials, and then last name, and the coauthors are all formatted "last name, initials" 2) Note the formats on date (e.g. no parentheticals) and the Volume(Issue):Page numbers formatting.
    - Ex: L. Belbin, Daly J., Hirsch T., Hobern D., LaSalle J. 2013. A specialist’s audit of aggregated occurrence records: An ‘aggregator’s’ perspective. ZooKeys 305(1): 67–76.

Citation identifier (DOI etc.)

- This is where we provide the DOI, ISSN, ISBN, or URI numbers that reference the survey. Usually one is provided in the study. DOI is the prefered one to use.
- If there is no DOI, etc., provided in the study, then find the number through a google search with the study's title. Crossref.org or doi.org can also be used to look up the DOI number.
- Sometimes the publication is considered a book or a serial and will have an ISSN or ISBN instead.
- In some rare cases, no identifier can be found and just put "none found"
- We need to use the right syntax for this section, i.e. ISBN:xxxx or DOI:xxxx
    - DOI:10.3897/zookeys.305.5438
    - ISBN: 0553804677
    - ISSN:0036-8075

Citation URL

- This is the web address for the study.
    - Ex: http://faculty.washington.edu/leache/wordpress/wp-content/uploads/2011/02/2 005HerpReview.pdf
- This is not expected to necessarily be permanent but provides a mechanism to allow another (quick) link between the metadata, the data and the original source.
- Not always present so use "none found" when it isn't.

Metadata recorder first name/Metadata recorder last name

- You put your first and last name in the appropriate boxes.

Metadata editor first name/Metadata editor last name

- You put your first and last name in the appropriate boxes.

Taxonomic Authorities Used

This field provides critical information on the taxonomic authority used and thus the "concepts" behind the list of names.

Citation reference of taxonomic authority

- The document that specifies the published taxonomic authority used for species names described in the checklist. Provide full citation in APA format.
- This needs to be cited consistently by using APA style.
- There are multiple columns to capture this data. Add only a few if needed. We don't want too many columns.

Geospatial information

Geospatial scope

- This is the location that the checklist represents (you can usually pull it directly from the title).
- For example, if the checklist is for the "Herps of Acadia National Park", then the

geospatial scope is “Acadia National Park”.

Total reported area covered by geospatial scope (km2)

- In the case of the example above, the total area covered by geospatial scope would be the area of Acadia National Park.
- If the author reports the size of their geographic scope, we’ll capture it. If they don’t, we’ll just put n/r.
- This question is to capture the habitats that the author(s) purposefully excluded during the survey for any reason.
- It should only be filled out if the paper explicitly provides information about the habitats that were not surveyed. For example, the surveyors could have purposely avoided agriculture habitats in their survey.

Temporal scope

Start date and End Date

Here we can fill out when surveys began and ended. Many surveys happening in distinct, multiple sessions. Please use start and end date to bracket the FULL length of the survey from the first day it began to the very last survey day. We are no longer capturing “time-blocks”. If the author only puts down the month time block or year time block then we can presume they searched from the beginning of that time until the end. For example, if they searched in September, then we would report their search from September 1 - 30. If they only reported a year then it would be January 1 until December 31 of that year.

Total Days Surveyed

This is the length of the survey in days. If the survey has multiple time blocks, count the days of each time-block and add all together for a final assessment of days surveyed. For example, if the inventory report says “We surveyed from 9/1/2007 to 9/10/2007 and 11/03/2008 to 11/06/2008”, you would count 10 days for the first block and 4 for the second, for a total of 14 days.

Season(s)

- Use the following terms, and report what the paper said if mentioned. If it isn’t mentioned and is completely clear from start and end dates, do “infer” season. You can use multiple terms (e.g., “wet dry” or “spring summer”)
    - Spring
    - Summer
    - Fall
    - Winter
    - Dry
    - Wet
    - All year

Was the study conducted during the day/night or both?

More coverage over the whole of day and night is likely to yield a better assessment of species than either just daytime or just nighttime surveys.

- Please use term “day”, “night”, or “both”.
    - In cases where surveying includes all the way through “dusk”, consider this to be “both”.

Methodology

Inventory type

- The definitions for the choices are here: https://docs.google.com/spreadsheet/ccc?key=0AhrY0qRdO4budC1mSUNWTDlkOXBVMjcza2Y3aV84SkE&usp=drive_web#gid=2
- Some loose working definitions of the choices:
    - Restricted Search - Use if the author mentions transects or plots.
    - Open Search - Less restrictive than restricted search - a particular search area is defined, but effort is presumed to be standardized across those areas. “*We searched sites for a length of time proportional to their area*”.
    - Opportunistic Search - This is a search such as often performed in efforts such as eBird where a surveyor goes to an area and does a search that doesn’t have a pre-defined shape and size of the survey effort, nor necessarily a clear effort measurement associated with it. It may still be relatively complete.
    - Trap - Select this one if the authors used stationary traps, camera traps, or nets. Also use if the authors used secondary evidence such as scat.
    - Adventitious - This refers to any survey data the authors obtain from other sources besides live sightings, such as scat, roadkill, specimens presented by hunters, or animals sold at markets.
    - Incidental- This is a search that happens with minimal effort. This is when authors report that they saw a species while they were eating or not actively looking. This may also refer to organisms which are outside the scope of the survey but which the authors specifically note.

Inventory and/or Compilation?

The values for this field should be “inventory-only”, “compilation-only” or “both”. “Inventory” refers to some on-the-ground survey effort. “Compilation” refers to using other sources of data, information and knowledge such as expert opinion, literature, museum specimens or other sources. In almost all cases we have examined, there is some form of on-the-ground surveying done. In rare cases, a new list is assembled purely from an assembly of other sources; use “compilation-only” here. In many cases, compiled data was not used at all,

- Examples of “inventory-only”:
    - “*A generalized search for amphibians and reptiles was undertaken at this site by a pair of researchers…*”
    - “*Data were collected from random field observations during a number of visits*

Compilation type

If compiled data was used, what types of resources were consulted? We use the following controlled vocabulary for this field:

- Expert knowledge
    - Park rangers or scientists are more reliable than villagers.
    - Unless the authors used a method to make sure the villagers knew the species in the area, i.e. vouchers or pictures, then we do not count villagers as expert knowledge.
- Primary literature
- Museum specimens
- Other

Compilation Effort

The amount of effort in compiling sources of ancillary data and information. This is critical for ultimately assessing completeness. For those sources that have compiled data, how much effort was spent? We use a categorical scale ranging from "low" to "medium" to "high" as defined below.

- *Low Effort.* Low effort is where only type of resource is moderately or poorly consulted, or that resource quality is relatively low. In cases where claims are made about multiple resources but there is NO information on how those were used, you may also consider this "low effort" (it is certainly low reporting effort). One common example is where authors consult a few previous checklists plus a few other papers which describe individual species not described elsewhere. Even if the authors use multiple categories of sources, there are very few sources overall (e.g. Consulted two checklists, one set of museum specimens, and one expert).
- *Medium Effort*: This is probably the hardest category --- a goldilocks problem. I think we define medium effort as follows: EITHER cases where one and only one type of source is well consulted (e.g. museum records) or where multiple sources are consulted but with superficial description of how they were used.
- *High effort*: 1) Multiple type of sources consulted (e.g. two or more sources such as museum records, literature, experts, etc). 2) Each source type has more than one and preferably many resources consulted within that type. If a surveyor consulted 5 museum collections and 5 experts, this is better than consulting just one type. 3) Consider quality of sources if possible. Consulting a significant and well curated collection is better than "Fred's attic full of pressed flowers'.

Were abundances reported for species inventoried? (Y/N)

- Often, in addition to a list of species, author(s) will included the number of individuals of each species. This is abundance information.
- If the author(s) only provides a checklist of species with +/- and not numbers then the answer for this question is n/r.

Were absences reported? (Y/N)

- Oftentimes, the researcher(s) may provide information about species that were expected to be found in the surveyed area, but were not observed during the survey.

- Here are some search words that will help when looking if the author reported absences
    - absent/absence
    - present/presence
    - record(ed)
    - observe(d)
    - find/found
    - not
    - lack
    - overlook(ed)
    - miss
    - detect(ed)
    - disappear(ed)
    - except
    - elude(d)
    - loss/lost
    - only
    - possible
    - nothing
    - indicate

If absences were reported, what was reported absent?

- Please use this field to record what was absent, at the level reported by the authors of the survey report. Preferred are actual species or higher level taxonomic units reported absent.

Prospective Taxonomic Scope

Surveyors always have a particular set of taxa in mind when they inventory an area. These taxa might be assembled purely by clade, but sometimes they also focus on a subset of a clade that are organized by size, volancy, etc. This field captures what the surveyor reports a the prospective taxonomic scope.

- The idea here is to check all information that an author give us, and not assume any hierarchy. So if an author reports that they searched for "small mammals", we just check "small mammals", with the knowledge that we don't know exactly what the author considers the taxonomic scope for small mammals.
- If the author reports they search for small mammals and then later mentions they searched for rodents and shrews, we put "small mammals", "rodents" and "shrews". The list can also be “herps, birds, mammals”. Here is a list of possible entries:
    - Mammals
    - Large mammals
    - Medium mammals
    - Small mammals
    - Rodents

- Bats
- Primates
- Birds
- Herps
- Reptiles
- Amphibians
- Plants
- Trees
- Shrubs
- Herbs
- Insects
- Molluscs
- Other

Geospatial Site Details

The current spreadsheet has 8 fields that are meant to capture content about site details as described in written survey reports. Each site can be described separately in one field. If more fields are needed to describe sites (e.g. a report includes 10 sites), you may add a column to the spreadsheet to capture this content. These fields can include habitat descriptions of the sites, vegetation, or other details reported. This provides a more granular look at the sites explored on the survey and can be used to determine how many different habitats across the whole of the area were examined by surveyors.

Survey Methods Details

The following 4 fields are used to capture more granular descriptions of the survey methods used, including quotes descriptions of survey approach, measures of effort from those different approaches, etc. The reporting here is meant to help provide details needed to assess completeness. As for the "Geospatial Site Details" section, if more fields are needed, these can be added to the spreadsheet (e.g. Insert Column). A variety of well reported survey methods with associated effort provide evidence of thoroughness.

Completeness Assessment

Completeness lower and upper bound

This assessment of completeness lower and upper bound is made by each recorder, and are assessed as a range in 25 percent increments. The highest possible score is thus 75-100 and the lowest 0-25. Many inventories may be hard to gauge and the range can therefore be greater e.g. 50-100, or 25-75, in such cases. We use the following rubric If the inventory included survey work and is not a compilation.

  - How many sites were surveyed across the full geospatial scope? Was coverage across space adequate? Did it cover different habitats?
  - How well did sampling cover time periods? Did it cover different time periods in the season, and during the course of a single day (e.g. day and night)? The more coverage is adequate, the more complete the survey was likely to be.
  - Was effort adequate? Effort adequacy really depends on a lot of factors. 100 trap hours might be great for common small mammals, but not so good for capturing insect diversity.
  - Were abundances or common-rare species noted? If so, you can use information on rare and common species along with effort to make a better estimate of completeness. Lots of reports of single individuals or just two individuals per species is often an indication that more sampling will yield more taxa
  - If the inventory also included compilations/expert knowledge, can you assess effort associated of the compilation efforts. Give more weight to multiple sources of compilation and adjust completeness estimate upward as a result.
  - Were there other factors that contributed to sampling deficiency or completeness? Give a quick explanation in the "completeness assessment notes field".
- If the inventory was ONLY compilation:
  - If the authors report completeness, use that information first.
  - If there is no reported completeness, then... How many different type of resources were consulted? Multiple sources (e.g. Museums AND expert opinion) should yield higher levels of completeness.
  - Within types of sources, how many different items were consulted. For example, how many different museum collections were examined, or how many different local experts. The more items, the higher completeness.
  - Can you assess the effort the authors have put into compilation beyond the type or resources and number of items used to create the compilation? This could be data quality effort, measures of time spent per interview as a finer grain view of effort, etc.

Completeness assessment rationale.

This is mandatory field and captures content about how data entry folks made their assessment. This should be detailed enough to provide the core thinking about the assessment made but avoid making this “voluminous” or “tedious”.

Other Notes

Inventory reporting data should be examined by at least two people, an initial data entry person and a "reviewer". Either may use notes to leave any further comments not covered elsewhere. Please do be clear who is leaving a note, since these are going to be stored along with all other data, into perpetuity.

Appendix C: Inferred Absence Rational

Horshehden

This inventory states in its methodology that they only searched for species that were medium to small. The authors also state that 14 of this classification are in the area of study and they only found 12; however, they did not state which species they did not find. In the list that I compiled of inferred absences, I was able to scrape the list into a list of only 5 species that I would count as true absences. These species, I would classify as medium to large. I was able to omit species that were small, like rodents. I was also able to omit bats because they did not search for them.

Badia

This inventory only searched for species that were carnivores, which was stated in their methodology section. With this list, it was easy to eliminate all the species that were not carnivores. I was also able to omit bats because they did not search for them.

Gonja

This inventory only searched for mammals that were small, which is stated in their methodology section. It was straight forward to eliminate the mammals that were classified as medium or large. I defined medium to be no greater than a hedgehog. The inventory also searched for bats, so I was not able to omit them.

Jabal

This inventory dealt with searching for medium and large mammals. Even though the authors did not directly say that these are the classification of mammals, they used camera trapping for their method of surveying. I omitted smaller mammals because the cameras are more likely to pick up medium to large mammals. I also was able to omit bats because they did not state they were searched for or used methods to survey them. Also the camera traps would not be able to catch them on film.

Cuzoamazonico

This inventory taxonomic scope was all mammals. However, the authors mentioned in their methodology section that they were limited with their searching and could not search most of the canopy. With this in mind, I omitted mammals that could reach the canopy, so mammals found in trees or ones that were defined as climbing or arboreal. Even though they searched for bats, I believe that I could not make the assumption that bat were absent because they might have been present in the canopy.

Seven Mile Comerong

In this inventory the authors state that they searched for all mammals, including bats. However, they only searched for bats that were insectivores because they only had equipment that picked up the vocalization of insectivorous bats. I was also able to refine the list by misspellings or species with more than one scientific name. I also was able to find species through a table that stated whether each one was presumed to live in the reserve or presumed extinct.

Termessos

All mammals were the taxonomic scope of this study. However, at the end of the inventory the authors state that more research can be done with specific groups of mammals. This allowed me to omit mammals of shrews, hedgehogs, rodents, and bats. A few species were found in small numbers, caracal and fallow deer. While the hyena and brown bear were considered extinct from that area for a couple of decades. The hyena was not found with the MOL data.

Virginia Barrier Island

This inventory was done of all mammals, except bats. The methodology did not suggest any species to omit based on their sampling techniques. I was also able to refine the list based on the species in stated in the inventory but not on their checklist. The last refining I did was through species with two scientific names.

Hautniger

This inventory was done of all mammals. The methodology did not give any indication of using tools that would help me refine the list. The authors surveyed with many different types of traps that caught a variety of mammals, including mist nets for bats. They also surveyed scat, tracks, and carcasses. The authors also went to local markets to see what mammals locals caught to sell for food. With these methods in mind, I was only able refine the list based on mammals with two scientific names.

Bale

This inventory was a more challenging one to refine. This is because it is a compilation study of mammals in this reserve, so there were no methods at all. I was only able to refine the list due to species that had another scientific name.

Appendix D: Metadata References

Afsar, M. & Tok, C.V. (2011). The herpetofauna of the Sultan Mountains (Afyon-Konya-Isparta), Turkey. *Turkey Journal of Zoology*, *35*(4): 491-501.

Al-Quran, S. (2009). The herpetofauna of the Southern Jordan. *American-Eurasian Journal of Agricultural & Environmental Sciences, 6*: 385-391.

Albayrak, T., Giannatos, G., & Kabasakal, B. (2012). Carnivore and ungulate populations in the Beydaglari Mountains (Antalya, Turkey): border region between Asia and Europe. *Polish Journal of Ecology, 60*(2): 419-428.

Aleixo, A., Poletto, F., Lima, M. F. C., Poretes, M. C. E., & Miranda, L. S. (2011). Notes on the Vertebrates of northern Pará, Brazil: a forgotten part of the Guianan Region, II. Avifauna. Boletim do Museu Paraense Emílio Goeldi. *Ciências Naturais 6*(1): 11-65.

Almaw, R. (2012). A Checklist of the Birds of the Abijatta-Shalla Lakes National Park. Ethiopian Wildlife Conservation Authority.

Alonso, L.E., Lauginie F., & Rondeau, G. (eds.). (2005). A Rapid Biological Assessment of two Classified Forests in South-Western Côte d'Ivoire. RAP Bulletin of Biological Assessment No. 34. Conservation International. Washington, D.C.

Amr, Z. S., Kalishaw, G., Yosef, M., Chilcot, B. J., & Al-Budari, A. (1996). Carnivores of Dana Nature Reserve (Carnivora: Canidae, Hyaenidae and Felidae), Jordan. *Zoology in the Middle East, 13*(1), 5-16.

Angelici, F.M., Laurenti, A., & Nappi, A. (2009). A checklist of the mammals of small Italian islands. *Hystrix, the Italian Journal of Mammalogy, 20*(1), 3-27.

Ararat, K., Abid I. M., & Abdulrahman, S. (2008). Key Biodiversity Survey of Kurdistan, Northern Iraq. Nature Iraq. 150 pages.

Asefa, A. (2011). Mammals of the Bale Mountains National Park, Ethiopia: A Compiled and Annotated Checklist, in Walia, special edition on Bale Mountains.

Austin, C.C., Hayden, C.J., Bigilale, I., Dahl, C., & Anaminiato, J. (2008). Checklist and Comments on the Terrestrial Amphibian and Reptile Fauna from Utai, Northwestern Papua New Guinea. *Herpetological Review, 39*(1), 40-46.

Aynalem, S. & Bekele, A. (2008). Species composition, relative abundance and distribution of bird fauna of riverine and wetland habitats of Infranz and Yiganda at southern tip of Lake Tana, Ethiopia. *Tropical Ecology, 49*(2), 199-209.

Baker, M. A. & Amr, Z. (2003). Rodent diversity in the northeastern desert of Jordan, with special reference on the ecology of Gerbillus cheesmani (Mammalia: Rodentia). *Casopsis Narodniho Muzea Rada Prirodovedna 172*(1-4), 141-152.

Baker, M. (2001). A survey of the avifauna of Chome Forest Reserve. East African Cross-Border Biodiversity Project "Reducing biodiversity loss at crossborder sites in East Africa". Tanzania component-URT/97/631. Tanzania.

Baker, M. A., Qarqaz, M., Rifai, L., Hamidan, N., Omari, K. A., Modry, D., & Zuhair, A. (2004). Results of herpetofaunal inventory of Wadi Ramm Protected Area, with notes on some relict species. *Russian Journal of Herpetology, 11*(1), 1-5.

Baran, I., Kumlutas, Y., Olgun, K., Ilgaz, C., & Kaska, Y. (2001). The Herpetofauna of the Vicinity of Silifke. *Turkish Journal of Zoology, 25*, 245-249.

Barata, M., Perera, A., Harris, D. J., Van der Meijden, A., Carranza, S., Ceacero, F., Garcí-a-Muñoz E., Gonçalves D., Henriques, S., Jorge, F., Marshall, J. C., Pedrajas, L., & Sousa, P. (2011). New observations of amphibians and reptiles in Morocco, with a special emphasis on the Eastern Region. *Herpetological Bulletin, 116*, 4-13.

Barnett, L. & Emms, C. (2005). Common reptiles of the Gambia. Rare Repro, Hailsham, East Sussex, UK.

Behangana, M. (2004). The diversity and status of amphibians and reptiles in the Kyoga Lake Basin. *African Journal of Ecology, 42*, 51-56.

Benda, P. (2010.) On a small collection of bats (Chiroptera) from western Sabah (North Borneo, East Malaysia.) *Vespertillo, 13/14*, 45-76.

Bene, J-C. K., Bitty, E. A., Bohoussou, K. H., Abedi-Lartey, M., Gamys, J., & Prince Soribah, A. J. (2013). Current Conservation Status of Large Mammals in Sime Darby Oil Palm Concession in Liberia. *G.J.B.A.H.S., 2*(3), 93-102.

Blaber, S. J. M. (1992).A checklist of the fishes of Groote Eylandt, north-western Gulf of Carpentaria,CSIRO Marine Laboratories Report 218.

Branch, W.R., Rödel, M.-O., & Marais, J. (2005). Herpetological survey of the Niassa Game Reserve, northern Mozambique - Part I: Reptiles. *Salamandra, 41* (4), 195-214.

Bunaian, F., Hatough, A., Ababaneh, D., Mashaqbeh, S., Yousef, M., & Amr, Z. (2001). The carnivores of the northeastern Badia, Jordan. *Turkish Journal of Zoology, 25*(1), 19-26.

Chabanet, P. & Durville, P. (2005). Reef fish inventoryof Juan de Nova's Natural Park (Western IndianOcean). *Western Indian Ocean Journal of MarineScience 4*, 145-162.

Coudrat, C. N. Z., Nanthavong, C., Sayavong, S., Johnson, A., Johnston, J., & Robichaud, W. G. (2014). Conservation importance of Nakai-Nam Theun National Protected Area, Laos for small carnivores based on camera trap data. *Raffles Bulletin of Zoology, 62*, 31-49.

De Marinis, A. M. & Masserti, M. (2009). Mammalian fauna of the Termessos National Park, Turkey. *ZooKeys 31*, 221-228.

Decher, J., Norris, R.W., & Fahr, J. (2010.) Small mammal survey in the upper Seli River valley, Sierra Leone. *Mammalia, 74*, 163-176.

Dimaki, M. & Legakis, A. (1999). The reptile fauna of the Fourni Archipelago (Eastern Aegean, Greece). *Herpetozoa*, *12*(3/4), 129-133.

Doggart, N., Loader, S., Perkin, A., & Mwangoka, M. (2002). Ngaramia Riverine Forest: A biodiversity survey. Tanzania Forest Conservation Group Technical Paper 14.

Donnelly, M.A., Chen, M.H., & Watkins, G.G. (2005). Sampling amphibians and reptiles in the Iwokrama Forest ecosystem. *Proceedings of the Academy of Natural Sciences of Philadelphia, 154*, 55-69.

Dowsett-Lemaire, F. & Dowsett. R.J. (2000). Birds of the Lobéké Faunal Reserve, Cameroon, and its regional importance for conservation. *Bird Conservation International, 10*, 67-87.

Drew J.A., Buxman C.L., Holmes D.D., Mandecki J.L., Mungkaje A.J., Richardson A.C., & Westneat M.W. (2012). Biodiversity inventories and conservation of the marine fishes of Bootless Bay, Papua New Guinea. *BMC Ecol 12*, 15.

Drewes, R. C. & Vindum, J. V. (1994). Amphibians of the Impenetrable Forest, Southwest Uganda. J. *Afr. Zool., 108*, 55-70.

Dueser, R.D., Brown, W.C., Hogue, G.S., McCaffrey C., McCuskey, S.A., & Hennessey, G.J. (1979). Mammals on the Virginia Barrier Islands. *Journal of Mammalogy, 60*(2), 425-429.

Fricke, R., Allen, G. R., Andréfouët, S., Chen, W. J., Hamel, M. A., Laboute, P., ... & Uyeno, D. (2014). Checklist of the marine and estuarine fishes of Madang District, Papua New Guinea, western Pacific Ocean, with 820 new records. *Zootaxa*, *3832*(1), 1-247.

Germano, J. M., Sander, J. M., Henderson, R. W., & Powell, R. (2003). Herpetofaunal communities in Grenada: a comparison of altered sites, with an annotated checklist of Grenadian amphibians and reptiles. *Caribbean Journal of Science, 39*(1), 68-76.

Ghalib, S.A., Jabbar, A., Khan, A.R., & Zehra, A. (2007). Current Status of the Mammals of Balochistan. *Pakistan Journal of Zoology, 39*(2), 117-122.

Gokturk, T., Bucak, F., & Artvinli, T. (2011). Mammalian Fauna of Artvin. *African Journal of Agricultural Research, 6*(6), 1418-1425.

Grismer, L. L., Youmans, T. M., Wood Jr., P. L., Ponce, A., Wright, S. B., Jones, B. S., Johnson, R., Sanders, K. L., Gower, D. J., Norsham, S. Y., & Lim, K. K. P. (2006). Checklist on the herpetofauna of Pulau Langkawi, Malaysia, with comments on taxonomy. *Hamadryad 30*, 61-74.

Habib, B. (2006). Status of Large Mammals in Proposed Big Pamir Wildlife Reserve, Wakhan, Afghanistan: Annual Report 01. Retrieved from http://pdf.usaid.gov/pdf_docs/pdacj928.pdf.

Hardesty, R. L., & Groothuis, D. R. (1993). Butterflies of the Laramie Mountains, Wyoming (Lepidoptera: Rhopalocera). *Journal of Research on the Lepidoptera, 32*, 107-123.

Hartmann, T., Ihlow, F., Edwards, S., Sothanin, S., Handschuh, M., & Bohme, W. (2013). A Preliminary Annotated Checklist of the Amphibians and Reptiles of the Kulen Promtep Wildlife Sanctuary in Northern Cambodia. *Asian Herpetological Research, 4*(1), 36–55.

Hoke, P., Demey R., & Peal, A. (2007). A rapid biological assessment of North Lorma, Gola and Grebo National Forests, Liberia. RAP Bulletin of Biological Assessment 44. *Conservation International*, 29-59.

Hollowell, T. & Reynolds, R.P. (Eds.) (2005). Checklist of the terrestrial vertebrates of the Guiana Shield. *Bulletin of the Biological Society of Washington, 13*, 1-98.

Honório, P. P. F., Ramos, R. T. C., & Feitoza, B. M. (2010). Composition and structure of reef fish communities in Paraíba State, north-eastern Brazil. *Journal of Fish Biology, 77*: 907-926.

Ibáñez, D., R. (2005). A note on amphibians and reptiles in the upper Rí-o Chagres basin, Panama In Harmon, R.S. (Ed.), The Rí-o Chagres, Panama (pp. 237-242). Springer, Berlin.

Ibrahim, A. A. (2001). The reptile community of the Zaranik Protected area, North Sinai, Egypt with special reference to their ecology and conservation. *Egyptian Journal of Natural History, 3*, 81-92.

Jackson, K., Zassi-Boulou, A. G., Mavoungou, L. B., & Pangou, S. (2007). Amphibians and Reptiles of the Lac Télé Community Reserve, Likouala Region, Republic of Congo (Brazzaville). H*erpetological Conservation and Biology, 2*(2), 75-86.

Janzen, P. & Bopage, M. (2011). The Herpetofauna of a Small and Unprotected Patch of Tropical Rainforest in Morningside, Sri Lanka. *Amphibian and Reptile Conservation 5*(2), 1-13.

Karunarathna, D. & Amarasinghe, A. (2011). Reptile diversity of a fragmented lowland rain forest patch in Kukulugala, Ratnapura district, Sri Lanka. *Taprobanica*: *The Journal Of Asian Biodiversity, 2*(2), 86-94.

Keller, G. S. & Yahner, R. H. (2002). Butterfly communities in two Pennsylvania national parks. *Northeastern Naturalist, 9*(2), 235-242.

Khonsue, W. & Thirakhupt, K. (2001). A Checklist of the Amphibians in Thailand. *The Natural History Journal of Chulalongkorn University, 1*(1), 69-82.

Klop, E., Lindsell, J. A., & Siaka, A. M. (2010). The birds of Gola Forest and Tiwai Island, Sierra Leone. *Malimbus, 32*, 33-58.

Kofron, C. P. & Chapman, A. (1995). Deforestation and bird species composition in Liberia, West Africa. *Tropical Zoology, 8*(2): 239-256.

Kumara, D. M. N. P., & Ukuwela, K. D. B. (2009). A survey on the amphibians of Ambagamuwa, a tropical wet mid-land area in Sri Lanka. *Herpetology Notes, 2*, 81-85.

Lares, R. V., Martínez R. M., Gadsden, H., León, G. A., Gaytán, G. C., & Trápaga, R.G. (2013). Checklist of amphibians and reptiles of the state of Durango, México. *Check List 9*(4), 714-724.

Leache, A.D. (2005). Results of Herpetological Survey in Ghana and a New Country Record. *Herpetological Review, 36*(1), 16-19.

Lebanon, A. R. A Study into the Importance of Jabal Moussa for Birds in Lebanon.

Murphy, M. J. (1998). Mammal survey of Seven Mile Beach National Park and Comerong Island Nature Reserve on the south coast of New South Wales. *Australian Zoologist 30*(4), 419-425.

Matveev, V. A. (2005). Checklist of Cambodian bats (Chiroptera), with new records and remarks on taxonomy. *Russian Journal of Theriology, 4*(1), 43-62.

McCarthy, J. L., McCarthy, K. P., Fuller, T. K., & McCarthy, T. M. (2010). Assessing Variation in Wildlife Biodiversity in the Tien Shan Mountains of Kyrgyzstan Using Ancillary Camera-trap Photos. *Mountain Research and Development, 30*(3), 295-301.

Menegon, M., Doggart, N., & Owen, N. (2008). The Nguru Mountains of Tanzania, an outstanding hotspot of herpetofaunal diversity. *Acta Herpetologica, 3*, 107-127.

Morin, M. P. (1996.) Bird inventory of Pu'ukohola Heiau National Historic Site, South Kohala, Hawai'i Island. *Cooperate National Park Resources Study Unit Technical Report, 102*, 1-13.

Mugerwa, B., Sheil, D., Ssekiranda, P., van Heist, M., & Ezuma, P. (2013). A camera trap assessment of terrestrial vertebrates in Bwindi Impenetrable National Park, Uganda. *African Journal of Ecology, 51*, 21-31.

Nader, M. R., El Indary, S., Salloum, B. A., & Dagher, M. A. (2011). Combining non-invasive methods for the rapid assessment of mammalian richness in a transect-quadrat survey scheme–Case Study of the Horsh Ehden Nature Reserve, North Lebanon. *ZooKeys, 119*, 63.

Nago, S. G. A., Grell, O., Sinsin, B., & Rodel, M. (2006). The amphibian fauna of Pendjari National Park and surroundings, northern Benin. *Salamandra-Bonn-*, *42*(2/3), 93.

Nande, R.. & Deshmukh, S. (2007). Snakes of Amravati district including Melghat, Maharashtra, with important records of the Indian Egg-eater, Montane Trinket snake and Indian Smooth snake. *Zoos Print J.*, *22*(12), 2920-2924. doi:10.11609/jott.zpj.1653.2920-4.

Newton, S. F. & Newton, A. V. (1996). Seasonal changes in the abundance and diversity of birds in threatened juniper forest in the southern Asir mountains, Saudi Arabia. *Bird Conservation International, 6*, 371-392.

O'Brien, T. G., & Kinnaird, M. F. (2013). Checklist of Birds from Laikipia County, Kenya. Retrieved from http://www.mpala.org/documents/index_40_360744062.pdf.

Ohler, A., Swan, S.R., & Daltry, J. C. (2002). A recent survey of the amphibian fauna of the Cardamom Mountains, Southwest Cambodia with descriptions of three new species. *The Raffles Bulletin of Zoology 50*(2), 465-481.

Onn, C. K., van Rooijen, J., Grismer, L. K., Belabut, D., Akil, M. A. M. M., Jamaludin, H., Gregory, R., Ahmad, & Norhayati, A. (2009). First report on the herpetofauna of Pulau Pangkor, Perak, Malaysia. *Russian Journal of Herpetology, 17*(2), 139-146.

Ostrowski, S. (2007). An annotated list of bird species observed by the Ecosystem Health Team in Wakhan in November–December 2006.Unpublished survey report. USAID-funded Biodiversity and Natural Resource Management Program of the Wildlife Conservation Society, Kabul, Afghanistan.

Owen, J. G. & Giron, L. (2012). Revised Checklist and Distributions of Land Mammals of El Salvador. *Museum of Texas Tech University, 310*, 1-32.

Pacheco, V., Patterson, B. D., Patton, J. L., Emmons, L. H., Solari, S., & Ascorra, C. F. (1993). Publicaciones del Museo de Historia Natural Universidad Nacional Mayor de San Marcos, *44*, 1-12.

Padial, J. M. & de la Riva, I. (2004). Annotated checklist of the amphibians of Mauritania (West Africa). *Rev. Esp. Herp. 18*, 89-99.

Patel, C. D. & Patel, M. I. (2011). Checklist of mammals in Taranga Hill Forest, Gujarat, India. *Zoo's Print, 26*(5), 30-32.

Persons, T. B. & Nowak, E. M. (2004). Inventory of amphibians and reptiles at Hovenweep National Monument, 2001–2003 – Final Report. USGS.

Pietersen, D. W. & Pietersen E. W. (2010). Annotated checklist of the birds of Banhine National Park, Southern Mozambique. *Ornithological Observations, 1*, 7-37.

Pietersen, D. W., Pietersen, E. W., & Haacke, W. D. (2013). First herpetological appraisal of the Parque Nacional de Banhine, Gaza Province, southern Mozambique. *Annals of the Ditsong National Museum of Natural History, 3*, 153-163.

Pinault, M., Loiseau, N., Chabanet, P., Durville, P., Magalon, H., Quod, J. P., & Galzin, R. (2013). Marine fish communities in shallow volcanic habitats. *Journal of Fish Biology, 82*, 1821-1847.

Purcell, K. L., Drynan, D. A., & Mazzocco, K. M. (2007). Vertebrate Fauna of the San Joaquin Experimental Range, California: an annotated checklist based on 70 years of observations. Pacific Southwest Research Station, US Department of Agriculture - Forest Service.

Qumsiyeh, M. B., Zavala, S. S., & Amr, Z. S. (2014). Decline in Vertebrate Biodiversity in Bethlehem, Palestine. *Jordan Journal of Biological Science 7*(2), 101-107.

Rabinowitz, A., Schaller, G. B., & Uga, U. (1995). A survey to assess the status of Sumatran rhinoceros and other large mammal species in Tamanthi Wildlife Sanctuary, Myanmar. *Oryx, 29*(2), 123-128.

Rabou, A. F. N. A., Yassin, M. M., Al-Agha, M. R., Hamad, D. M., & Ali, A. K. S. (2007). The herpetofauna of the Gaza Strip with particular emphasis on the vicinity of Wadi Gaza. *The Islamic University Journal, 15*(1), 111-135.

Rakotondravony, H.A. & Goodman, S.M. (2011). Rapid Herpetofaunal Surveys Within Five Isolated Forests on Sedimentary Rock in Western Madagascar. *Herpetological Conservation and Biology, 6*(2), 297-311.

Ramos, M. H., Mendoza, E. M, Santos, A. G., Reyes, R. B., Jr., Capuli, E. C., & Bimbao, M. P. (2014). Atlas of Common Fishes of Tayabas Bay, Quezon Province, Philippines. 67 p., Department of Agriculture. Bureau of Fisheries and Aquatic Resources Region IV-A.

Rao, K. T., Ghate, H. V., Sudhakar, M., Javed, S. M. M., & Krishna, I.S.R. (2004). Herpetofauna of Nallamalai Hills with eleven new records from the region including ten new records for Andhra Pradesh. *Zoos' Print Journal, 20*(1), 1737-1740.

Rappole, J. H., Aung, T., Rasmussen, P. C., & Renner, S. C. (2011). Ornithological Exploration in the Southeastern Sub Himalayan Region of Myanmar. *Ornithological Monographs, 70*, 10-29.

Rasmussen, P. C., Aung, T., & Rappole, J. H. (2011). The Breeding Avifauna of the Sub-Himalayan Zone of Northern Kachin State, Myanmar. *Ornithological Monographs, 70*, 95-108.

Rastegar-Pouyani, N., Kami, H. G., Rajabzadeh, M., Shafiei, S., & Anderson, S. C. (2008). Annotated Checklist of Amphibians and Reptiles of Iran. *Iranian Journal of Animal Biosystematics. 4*(1), 7-30.

Raxworthy, C. J. & Attuquayefio, D. K. (2000). Herpetofaunal communities at Muni Lagoon in Ghana. *Biodiversity and Conservation, 9*, 501-510.

Razzetti, E., Sindaco, R., Grieco, C., Pella, F., Ziliani, U., Pupin, F., Riservato, E., Pellitteri-Rosa, D., Butikofer, L., Suleiman, A.S., Al-Aseily, B.A., Carugati, C., Boncompagni, E., & Fasola, M. (2011). Annotated checklist and distribution of the Socotran Archipelago Herpetofauna (Reptilia). *Zootaxa, 2826,* 1-44.

Read, J. L. & Moseby, K. (2006). Vertebrates of Tetepare Island, Solomon Islands. *Pacific Science, 60*(1), 69-79

Reynolds, R. P. & MacCulloch, R. D. (2012). Preliminary Checklist of Amphibians and Reptiles from Baramita, Guyana. *Check List, 8*(2), 211-214.

Rocha, L. A. & Rosa, I. L. (2001). Baseline assessment of reef fish assemblages of Parcel Manuel Luiz Marine State Park, Maranhão, north-east Brazil. *Journal of Fish Biology, 58*, 985-998.

Rodgers, K. S. (2003). Effects of habitat, wave exposure, and marine protected area status on coral reef fish assemblages in the Hawaiian archipelago. *Coral Reefs 22*, 291-305.

Roelke, C. E. & Smith, E. N. (2010) Herpetofauna, Parc National des Volcans, North Province, Republic of Rwanda. Check List: *Journal of Species Lists and Distribution, 6*, 525-531.

Rovero, F. & De Luca, D. W. (2007). Checklist of mammals of the Udzungwa Mountains of Tanzania. *Mammalia, 71*, 47-55.

Rufford Small Grants Foundation. (2011). Tugay forests in Tajikistan: the last remainders of globally endangered ecosystem (Small Grants Project No 9304-1). Moscow Russia: Pavel Kvartalnov. Retrieved from http://www.rufford.org/files/9304-1%20Detailed%20Final%20Report.pdf.

Salim, M., Porter, R., & Rubec, C. (2009). A summary of birds recorded in the marshes of southern Iraq, 2005-2008. *BioRisk, 3*, 205-219.

Seddon, P. J., Heezik, Y. V., & Nader, I. A. (1997). Mammals of the Harrat al-Harrah protected area, Saudi Arabia. *Zoology in the Middle East, 14*(1), 37-46.

Shaw, P. & Shewry, M. (2001). Population density and habitat associations of restricted-range bird species at Ruhija, Bwindi Impenetrable Forest, Uganda. *Bird Conservation International, 11(*3), 161-174.

Shehab, A., Karatas, A., Amr, Z., Mamkhair, I., & Sözen, M. (2007). The distribution of bats (Mammalia: Chiroptera) in Syria. *Vertebrate Zoology*, *57*, 103-132.

Smith-Vaniz, W. F. & Jelks H. L. (2014). Marine and inland fishes of St. Croix, U. S. Virgin Islands: an annotated checklist. *Zootaxa 3803*(1): 1-120.

Soorae, P. S. (2004). A herpetological survey of some islands in the Arabian Gulf, Abu Dhabi Emirate, United Arab Emirates. *Zoology in the Middle East, 32*, 33-38.

Soorae, P. S., Quarqz, M. A., & Gardner, A. S. (2010). An Overview and Checklist of the Native and Alien Herpetofauna of the United Arab Emirates. *Herpetological Conservation and Biology, 5*(3), 529-536.

Srinivasan, U., Dalvi, S., Naniwadekar, R., Anand, M. O., & Datta, A. (2010). The birds of Namdapha National Park and surrounding areas: recent significant records and a checklist of the species. *Forktail*, *26*, 92-116.

Stalmans, M. (2006). Tinley's plant species list for the Greater Gorongosa Ecosystem, Moçambique. Unpublished report by International Conservation Services to the Carr Foundation and the Ministry of Tourism.

Stanley, W. T., Gunn, J., & Kihuale, P.M. (2005). Results of a preliminary small mammal survey of Malundwe Mountain, Mikumi National Park, Tanzania. *Journal of East African Natural History, 94*(1), 213-230.

Stiassny, M. L. J. (1991). Report on a small collection of fishes from the Wologizi Mountains of Liberia, West Africa, with a description of two new species of Barbus (Ostariophysi, Cyprinidae). *American Museum Novitates, 3015*, 1-12.

Stuart, B. (1998). A survey of amphibians and reptiles in Phou Louey National Biodiversity Conservation Area, Houaphanh Province, Lao PDR, Vientiane, Lao PDR: Wildlife Conservation Society.

Suranjan Karunarathna, D. M. S., Thasun Amarasinghe A. A., Gabadage D. E., Bahir M. M., & Harding L. E. (2010). Current Status of Faunal Diverity in Bellanwila - Attidiya Sancturary, Colombo District - Sri Lanka. *Taprobanica 2*(1), 48-63.

Tawiri (2010). Aerial Census in the Selous-Niassa Corridor, Dry Season, 2009.

Terborgh, J.W., Fitzpatrick, J.W., Emmons, L. (1984). Annotated checklist of bird and mammal species of Cocha Cashu Biological Station, Manu National Park, Peru. *Fieldiana Zoology, 21*, 1-29.

Teynie, A. (2004). Notes on Reptiles of Nam Lan Conservation Area in Phongsaly Province of Lao PDR. Société d'Histoire Naturelle.

Timmins, R. J. & Duckworth, J. W. (2012). A survey of gibbons and other wildlife in the Bokeo section of Nam Kan National Protected Area, Lao, PDR. Fauna & Flora International.

Timmins, R. J. & Khounboline, K. (1996). A Preliminary Wildlife and Habitat Survey of Hin Namno National Biodiversity Conservation Area, Khammoune Province, Lao PDR. Wildlife Conservation Society, New York.

Tordoff, A.W., Hung, L. M., Truong, N. Q., & Swan, S. R. (2002). A Rapid Field Survey of Van Ban District, Lao Cai Province, Vietnam. BirdLife International Vietnam Program.

Trainor, C.R. (2005). Birds of Tapuafu penninsula, Roti island, Lesser Sundas, Indonesia. *Forktail*, *21,* 121-131.

Tsetan, C. & Ramanibai, R. (2011). Reptilian fauna of agricultural landscapes of Chembarambakkam Lake, Chennai, Tamil Nadu. *Reptile Rap, (13)*, 2-8.

Tuyisingize, D. (2010). Terrestrial small mammal community composition in the Volcanoes National Park, Rwanda. (Master's thesis.) Percy FitzPatrick Institute of African Ornithology, University of Capt Town.

Vaz-Silva, W., Guedes, A. G., Azevedo-Silva, P. D., Gontijo, F. F., Barbosa, R. S., Aloísio, G. R., & Oliveira, F. C. G. (2007). Herpetofauna, Espora hydroelectric power plant, state of Goiás, Brazil. *Check List, 3*(4), 338-345.

Venugopal, P. D. (2010). An updated and annotated list of Indian lizards (Reptilia: Sauria) based on a review of distribution records and checklists of Indian reptiles. *Journal of Threatened Taxa, 2*(3), 725-738.

Von May, R., Siu-Ting, K., Jacobs, J. M., Medina-Muller, M., Gagliardi, G., Rodriguez, L. O. & Donnelly, M. A. (2009) Species diversity and conservation status of amphibians in Madre de Dios, Southern Peru. *Herpetological Conservation and Biology, 4*(1), 14-29.

Wanger, T.C., Motzke, I., Saleh, S., & Iskandar, D. (2011). The amphibians and reptiles of the Lore Lindu National Park area, Central Sulawesi, Indonesia. *Salamandra 47*(1), 17-29.

Werner, Y. L. & Avital, E. (1980). The herpetofauna of Mt. Hermon and its altitudinal distribution. *Israel Journal of Zoology, 29*(4), 192-193.

Wesołowski, T., Mitrus, C., Czeszczewik, D., & Rowiński, P. (2010). Breeding bird dynamics in a primeval temperate forest over thirty-five years: variation and stability in the changing world. *Acta Ornithologica*, *45*(2), 209-232.

Williams, S. T., Goodman, S. M., Kihaule, P. M., & Howell, K. M. (2000). A Survey of The Small Mammals of the Gonja Forest Reserve, Tanzania. *Journal of East African Natural History 89*, 73-83.

Woodman, N., Timm, R.M., Arana C., R., Pacheco, V., Schmidt, C.A., Hooper, E.D., & Pacheco A., C. (1991). Occasional Papers of the Museum of Natural History, The University of Kansas, *145*, 1-12.

Yalden, D. W., Largen, M. J., Kock, D., & Hillman, J. C. (1996). Catalogue of the mammals of Ethiopia and Eritrea. 7. Revised checklist, zoogeography and conservation. *Tropical Zoology, 9*(1), 73-164.

Yousif, R.A. (2010). Abundance and distribution of Birds in Dinder National Park, Sudan. *Pakistan Journal of Wildlife, 1*(2), 67-72.

Ziegler, S., Nikolaus, G., & Hutterer, R. (2002). High mammalian diversity in the newly established National Park of Upper Niger, Republic of Guinea. *Oryx, 36*(01), 73-80.

Zug, G. R., Win, H., Thin, T., Min, T. Z., Lhon, W. Z. & Kyaw, K. (1998). Herpterofauna of the Chatthin Wildlife Sanctuary, North-Central Myanmar with Preliminary Observations of their Natural History. *Hamadryad, 23*(2), 111-120.