1

Digital Accessible Knowledge and well-inventoried sites for

birds in Mexico: baseline sites for measuring faunistic change

4

A. Townsend Peterson[1], Adolfo G. Navarro-Sigüenza[2], and Enrique Martínez-Meyer[3]

6

[1]*Biodiversity Institute, University of Kansas, Lawrence, Kansas 66045 USA*

8

[2]*Museo de Zoología, Facultad de Ciencias, Universidad Nacional Autónoma de México,*

Mexico City 04510, Mexico

11

[3]*Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City 04510,*

Mexico

14

15

Corresponding author: A. Townsend Peterson, town@ku.edu

Código de campo cambiado

Background Faunal change is a basic and fundamental element in ecology, biogeography, and conservation biology, yet vanishingly few detailed studies have documented such changes rigorously over decadal time scales. This study responds to that gap in knowledge, providing a detailed analysis of Digital Accessible Knowledge of the birds of Mexico, designed to marshal DAK to identify sites that were sampled and inventoried rigorously prior to the beginning of major global climate change (1980).

Methods We accumulated DAK records for Mexican birds from all relevant online biodiversity data portals. After extensive cleaning steps, we calculated completeness indices for each 0.05° pixel across the country; we also detected 'hotspots' of sampling, and calculated completeness indices for these broader areas as well. Sites were designated as well-sampled if they had completeness indices above 80% and >200 associated DAK records.

Results We identified 100 individual pixels and 20 broader 'hotspots' of sampling that were demonstrably well-inventoried prior to 1980. These sites are catalogued and documented to promote and enable resurvey efforts that can document events of avifaunal change (and non-change) across the country on decadal time scales.

Conclusions Development of repeated surveys for many sites across Mexico, and particularly for sites for which historical surveys document their avifaunas prior to major climate change processes, would pay off richly in information about distributional dynamics of Mexican birds.

Introduction

The temporal dynamics of geographic distributions of species and composition of local

biotic communities are central to much of biogeography, macroecology, and conservation

biology (Lavergne et al. 2010). That is, how species' distributions evolve through time and

how community composition changes as a result are key in determining essentially all

results in these areas of ecology and evolutionary biology. Although biogeographers invest

fundamentally in retracing the geography of evolving lineages over long periods of time,

strangely little information exists on short-term dynamics of species' distributions and

community composition (e.g., Nunes et al. 2007; Tingley & Beissinger 2009).

An important opportunity to understand these shorter-term dynamics of distributions

and communities by means of longitudinal comparisons of inventories of local faunas and

floras. That is, when a baseline of solid, complete, and well-documented knowledge exists

about a site (Colwell & Coddington 1994; Peterson & Slade 1998; Soberón & Llorente

1993), re-surveys over years, decades, and centuries can offer a fascinating view into the

natural dynamics of species' ranges and the effects of human presence and activities.

Drivers of these changes may act both on local scales (e.g., effects of land use change) and

on global scales (e.g., effects of climate change), and potentially may interact as well;

studies integrating and comparing effects of different such drivers are particularly rare (e.g.,

Peterson et al. 2015; Rubidge et al. 2011). Of course, detailed documentation of species

identifications and of the completeness of inventory efforts are necessary for both the

baseline and the resurvey, but the general paradigm has considerable potential.

Previous baseline/re-survey efforts have yielded fascinating information about

faunas and floras. For example, Nilsson et al. (2008) documented 40% extirpation of

butterfly species over a 90+ year span on a plot in southern Sweden, and found that species

65 disappearing from the fauna tended to be those with a short flight length period, narrow

66 habitat breadth, and small distributional area in Europe; however, only flight length period

67 was significant in multivariate analyses. Grixti & Packer (2006) studied bee communities at

68 a site in southern Ontario, and documented community changes and diversity increases

69 over a 40+ year span, likely in response to successional changes in the surrounding

70 landscapes. Scholes & Biggs (2005) proposed a biodiversity intactness index, and showed

71 widespread population declines, particularly among mammals, and ecosystem declines

72 concentrated in grasslands, across a large region of southern Africa. In California,

73 important efforts have been carried out to resurvey sites studied by Joseph Grinnell a

74 century ago, detecting fascinating distributional (Tingley & Beissinger 2009), phenotypic

75 (Leache et al. 2010), and genetic (Rubidge et al. 2011) changes in vertebrate species and

76 communities (Moritz et al. 2008).

77   Within Mexico, such before-and-after studies have been particularly scarce.

78 Peterson & Navarro-Sigüenza (2005) used nineteenth-century sources to reflect on changes

79 in the avifauna of the Valley of Mexico over the twentieth century. In a particularly

80 interesting example, Olvera-Vital (2012) re-surveyed the avifauna of Misantla, Veracruz—

81 of great interest is that the avifauna has been quite stable over the past 50-100 years, which

82 apparently reflects the very early mass-disturbance to the natural habitats of that region

83 (Sánchez G. 1998), such that the baseline inventory was itself already post-disturbance. On

84 broader and coarser spatial scales, Peterson et al. (2015) assessed countrywide changes in

85 Mexican endemic bird species' distributions over the second half of the twentieth century,

86 and found dominant effects of half-century changes in temperature (and not of human

87 impact on landscapes or changes in precipitation) in driving avifaunal change. Curiously,

88 however, to our knowledge at least, this short list includes all sites that have seen baseline

Comentado [AHA1]: It is unclear to the unfamiliar to the Peterson (2015) reference whether the changes (both distributional and climatic) cited occur within the second half of the XXth century or as respects to the previous period. Please rephrase slightly so as to dispel the possibility of misunderstanding.

89   and re-survey inventories of birds, such that the nature and pattern of avifaunal change

90   across Mexico remains very poorly characterized.

91   The purpose of this paper is to stimulate and enable a next generation of such re-

92   survey efforts for birds across Mexico by means of cataloguing sites for which solid

93   documentation exists for the original baseline inventory, and for which the baseline

94   inventory is demonstrably complete. We reviewed all existing Digital Accessible

95   Knowledge (DAK; Sousa-Baena et al. 2013) regarding the birds of Mexico: we probed four

96   major online data portals for relevant data (i.e., bird records from Mexico prior to 1980),

97   and evaluated inventory completeness at two spatial extents (0.05° grid squares, and

98   coarser hotspots of sampling). We present a simple list and catalog of sites that have seen

99   relatively complete (i.e., $\geq$80% of avifauna known and documented) inventories, as a

100   challenge and stimulus to the ornithological community in Mexico. Resurvey of these

101   localities would provide rich and informative rewards in understanding the dynamics of

102   bird populations and distributions across the country.

103

104   **Materials & Methods**

105   This analysis is based on a suite of assumptions about data quality and appropriate

106   resolutions inherent in biodiversity data. For instance, we used single days as the unit of

107   sampling effort throughout our analyses, as this resolution has proven an effective balance

108   between too much and too little resolution; temporal information finer than the level of day

109   is rarely available with older biodiversity data, whereas coarse temporal resolution can

110   underappreciate the efficacy of short-term, intensive inventory efforts. Hence, a basic

111   working data set was the combination of species identification, day (i.e., unique

112   combinations of year, month, and day), and place. This latter we defined in various ways,

113   described below, but we note that we used both geographic coordinates and textual locality

114   descriptions to avoid loss of information owing to lack of georeferencing, at least to all

115   extents possible.

116   *Data sources and quantities*. We drew data from four basic sources for this study:

117   the Global Biodiversity Information Facility (GBIF; http://www.gbif.org), VertNet

118   (http://www.vertnet.org/), the *Red Mundial de Información de la Biodiversidad* (REMIB;

119   http://www.conabio.gob.mx/remib/doctos/remib_esp.html), and UNIBIO

120   (http://unibio.unam.mx/). Each of these biodiversity information networks has its respective

121   strengths and weaknesses, and considerable overlaps exist among them in coverage of

122   biodiversity information sources. We drew data on bird occurrences in Mexico from all

123   four, and trusted that duplication would be removed in our data cleaning steps; for this

124   reason, we are unable to develop comparisons among the different data resources. For

125   GBIF and VertNet, automated download was possible; however, for REMIB and UNIBIO,

126   we requested and were provided with data 'dumps,' as some of the data were restricted

127   from full public access (REMIB) or not yet fully available to the public (UNIBIO). The full

128   set of sources contributing data to this analysis is provided in the Appendix.

129   *Data reduction and cleaning*.—Initial data downloads totaled hundreds of

130   thousands of records from Mexico from each of the data portals, reaching millions from

131   GBIF (although the GBIF numbers are largely from eBird and aVerAves, which come in

132   greatest part from post-1980; Fig. 1, Table 1). We expected considerable redundancy

133   between data sources, in the form of situations in which the same species was collected or

134   recorded at the same place on the same day, so we embarked on a lengthy process of data

135   reduction and cleaning. Without a doubt, some mistakes were made, and some information

136   was lost, but this effort aimed to detect and highlight the major features of Mexican bird

137 DAK, rather than all of the details. That is, we focused on sites that had the most

138 information available, and explicitly excluded information for less-well-known sites.

139   A first step was to concatenate all four datasets into a single, larger dataset, and to

140 reduce the set of fields to the essential three suites of fields mentioned above: species (we

141 retained order, family, genus, and species, to permit identification of the most difficult

142 names), date collected (year, month, and day), and place (latitude and longitude, when

143 available, and state, municipality, and specific locality). We filtered these records to

144 remove all those records from years after 1980. The initial total of 2,598,478 records

145 distilled to 845,658 that both (1) had dates and (2) were not from after 1980.

146   From this set, we extracted 10,762 unique combinations of order, family, genus, and

147 species, which included a wild diversity of name combinations, and required considerable

148 work to arrive at a consistent suite of names corresponding to the birds of Mexico. We

149 avoided the temptation to attempt to take the taxonomic treatment to a newer authority list

150 (e.g., Gill & Donsker 2014; Navarro-Sigüenza & Peterson 2004; Peterson & Navarro-

151 Sigüenza 2006) because taxonomic splitting, which has dominated recent years of

152 taxonomic work, would cause considerable confusion of names applied to older records.

153 Hence, we reduced the initial, highly redundant set of names to 1027 names that coincided

154 with the taxonomy of the American Ornithologists' Union (AOU 1998), except that we

155 merged *Empidonax traillii* and *E. alnorum*, in light of very frequent confusion in

156 identification (Heller et al. 2016). This step was achieved based on long years of experience

157 with Mexican bird taxonomy, plus occasional consultation of the literature (Monroe &

158 Sibley 1993; Peters 1931-1987) and online (http://avibase.bsc-eoc.org/) resources.

159   Dates were summarized as a concatenation of year, month, and day (e.g.,

160 1964_7_16). Records with dates for which all three time elements were missing were

161 removed from analysis, but records with partial dates (e.g., year and month, but not day)

162 were treated as a unique time event. All data management was carried out in OpenRefine

163 (http://openrefine.org), which permitted many important initial steps of combining similar

164 names, and in Microsoft Access, which permitted development of customized queries for

165 further refinement.

166 *Data analysis*.—Perhaps the most difficult-to-manage of the fields was that of

167 'place.' Here, we used a three-step process that aimed to retain a maximum of locality

168 information, yet avoid the massive and prohibitive task of full georeferencing of all records

169 for which no geographic coordinates were available. Hence, (1) we used the 502,935

170 records that had latitude-longitude data in a first-pass analysis that aimed to identify single

171 sites that were well inventoried, and to identify somewhat broader 'hotspots' of sampling

172 (details provided below). Next (2), we used locality names associated with the records

173 falling in the hotspots to probe the data lacking georeferences, and thereby rescued >54,000

174 records. Finally (3), we inspected the remaining data records—those lacking

175 georeferences—to identify additional sites that merited analysis (5 such additional, un-

176 georeferenced sites indeed proved to be relatively well-inventoried, such that this step was

177 important; see below). In this way, the number of localities that needed to be georeferenced

178 was minimized, and yet we managed to include the great bulk of Mexican DAK in our

179 analyses.

180 Step 1: Here, we aimed to develop first analyses of the 502,935 records (of which

181 481,409 were from 1980 or before) from the original data set that carried latitude-longitude

182 coordinates, both to identify individual pixels (0.05° resolution) that were well-inventoried,

183 and to identify concentrations ("hotspots") of sampling that may or may not prove to be

184 well-inventoried. We first filtered this data set to retain only the records that were within

185 0.05° (5-6 km) of the administrative outline of Mexico—this step left 499,794 records for

186 initial analysis.

187   We then created a shapefile with square elements of 0.05° for Mexico, and

188 eliminated pixels that were >0.05° from the administrative boundaries of Mexico. We then

189 used a spatial join to count numbers of records in each polygon; numbers of records per cell

190 ranged from nil to as high as 15,464 (a pixel centered on Chilpancingo, Guerrero). We used

191 optimized hot spot analysis (implemented in ArcGIS, version 10.2) to identify

192 concentrations of sampling effort across the country—we used the Getis-Ord Gi* statistic

193 (Hot Spot Analysis tool), and focused only on hotspots (0.05° grid squares) significant at

194 the highest (99%) confidence level. We isolated these well-sampled suites of pixels as a

195 separate shapefile, and then merged pixels in contiguous sets as preliminary hypotheses of

196 broader hotspots of sampling, which we enriched with more records in step 2 (see below).

197   We also used the initial data set to develop an identification of single 0.05° pixels

198 that were well-inventoried, as follows. The 481,409 records reduced to 277,249 unique

199 combinations of species x date x pixel. We calculated S_{obs} as the number of species that

200 have been recorded from each pixel, N as the total number of unique combinations of

201 species x date from each pixel, and a and b as the numbers of species recorded exactly once

202 and exactly twice, respectively, from each pixel. We then used these data to calculate, for

203 each pixel, the Chao2 estimator of expected species richness, with its associated adjustment

204 for small sample sizes (Colwell 1994–present), as

205

206
$$S_{exp} = S_{obs} + \left(\frac{N-1}{N}\right)\left(\frac{a(a-1)}{2(b+1)}\right)$$

207

Comentado [AHA2]: This is total. Couple of lines before it is mentioned that some 480K are pre-1980 and before the filter—so how many (number or %) are left after the clipping that are also pre-1980?

Comentado [AHA3]: Yes, but, again—this figure is records before clipping for boundaries and pre-1980. As the step described here is after-clipping, at least the figure should be records post-clip (total) or, if specifically dealing with baseline records (pre-1980), then post-clip and pre-1980 (which does not appear in the text before but does not appear to be this figure). Please be consistent with the numbers throughout this section and make clear and unambiguous what subset of records the figures refer to.

Comentado [AHA4R3]:

Comentado [AHA5]: This is good and follows the methods description but may have a catch: N should be occurrences, and any 0.05º pixel having distinct occurrences (e.g. sufficiently different coordinates) for a given date and taxon but within the same pixel is counting as one occurrence. This is a consequence of reducing the data to the chosen spatial resolution, but for Sobs it does matter—N should be the number of occurrences rather than frequencies (to be consistent with a and b) even though Chao-2 deals with frequencies—but then it's frequencies throughout. So the main issue is: If a sample is considered a pixel, then the elements in the sample are the date-taxon combinations. This excludes co-occurrences. One may guess that such co-occurrences may be really few and not matter much in the overall analysis, but then an indication of the number (or percent) of co-occurrences (most likely tiny and very often zero for most pixels I presume, on account of using the much higher resolution afforded by dates) could help the reader to dismiss their potential impact.

Código de campo cambiado

208   Completeness was then calculated as $C = S_{obs} / S_{exp}$. To avoid including occasional

209   localities with small sample sizes and artificially 'complete' inventories, we further

210   removed all pixels for which $N < 200$. We summarized these calculations in a table, which

211   we then imported into ArcGIS and joined to the fishnet shapefile for visualization.

212   Step 2: This step aimed to assess the fairly large portion of the data that lacked

213   georeferences (195,371 records), and 'rescue' relevant data for further enrichment of the

214   hotspot analysis. That is, in Step 1, we used georeferenced occurrence data to identify

215   hotspots of sampling, but many more occurrences were documented in records lacking such

216   data. Hence, here, first, we used the merged hotspot shapefile to associate the original

217   occurrence data (i.e., records with georeferences) to hotspots, and identified key locality

218   names among those occurrence data. (e.g., the Comitán hotspot included localities such as

219   "La Trinitaria," "El Triunfo," "Lagunas de Montebello," "Las Margaritas," and "Santa

220   Rosa"). Finally, we used those names to probe the non-georeferenced data records

221   (checking, of course, that distance offsets were not >5 km), and thereby rescued 54,369

222   records that were added to the hotspot-based analyses. Inventory statistics and

223   completeness values were then recalculated for each hotspot just as they were calculated in

224   Step 1 for each pixel.

225   Step 3: This final step involved review of the remaining non-georeferenced records,

226   even after the rescue of Step 2. That is, Step 2 focused on non-georeferenced records that

227   corresponded to already-identified pixels and hotspots, and that could be used to enrich the

228   existing (georeferenced) data from those sites. This third step, however, involved a look at

229   the remaining data to see if additional well-known sites could be identified.

230   We assembled the raw locality descriptors, and tallied numbers of records

231   associated with each; we did some minor cleaning and synonymizing of minor variants on

Comentado [AHA6]: It's good to write down the actual numbers (it facilitates checking) but for the reader it's perhaps more compelling and meaningful to have a fraction or percentage.

232 locality descriptors to maximize numbers of records for each locality. Finally, we

233 developed completeness indices (as described above) for each of the non-georeferenced

234 localities that had a raw sample size (i.e., unique combinations of locality descriptor,

235 year/month/day, and species) of $\geq$200 records. All localities with $C \geq 0.8$ were then

236 georeferenced and included among well-known sites.

237   *Data availability*. All data managed in this study are openly available, or will be

238 shortly, as institutional permissions are finalized, from GBIF (http://www.gbif.org),

239 VertNet (http://www.vertnet.org/), REMIB

240 (http://www.conabio.gob.mx/remib/doctos/remib_esp.html), and UNIBIO

241 (http://unibio.unam.mx/). A full synonymy of locality names in relation to the hotspots

242 identified in this paper is available at http://hdl.handle.net/1808/20674. GIS shapefiles

243 showing the two sets of hotspots identified in this paper are available at

244 http://hdl.handle.net/1808/20673.

245

246 **Results**

247 Processing data available for birds of Mexico from raw DAK downloads into usable

248 records of species at localities on particular days involved considerable reduction in

249 numbers of records (Table 1). Simply progressing from raw to unique combinations of

250 species x locality x geographic coordinates x year x month x day involved a reduction of

251 20.1-26.2% in numbers of records (note that this reduction step was done prior to removing

252 records post-1980). Subsequent data cleaning and reduction steps (see Methods, above)

253 reduced redundancy both among data sources (GBIF, VertNet, REMIB, UNIBIO) and

254 among nearby localities, leaving a final number of 499,794 records for analysis.

Comentado [AHA7]: While I understand the rationale for rescuing as many data as possible, I'm not so sure this procedure as described does not incur into selective sampling. Under the hypothesis of uniform (or random) sampling, there would be a continuum in the number of occurrences between well-sampled and poorly-sampled sites, with the threshold at some arbitrary value (e.g. 200 in this case). However by selecting certain (abundant) records for inclusion while discarding other equally-valid but not-georeferenced records, a gap might be built by which potentially well-sampled sites are pushed further away from the threshold. I believe this last step should be discussed a little further, possibly with reference to potential bias in the results arising from this possible gap widening.

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Comentado [AHA8]: Is this consistent with the previous 481k figure?

255   A first analysis focused on 277,249 unique combinations of species x date collected

256   x 0.05° pixel. These records fell in pixels in numbers ranging from nil to 15,464 records

257   (Chilpancingo, Guerrero). Processing records in each pixel into estimates of expected

258   numbers of species and estimated inventory completeness (Table 2; Fig. 2), we found that

259   24 pixels were complete at the level of $C \geq 0.9$, and a further 71 pixels were complete at the

260   level of $C \geq 0.8$. The well-inventoried sites were well-distributed across the country, from

261   Baja California Sur and Sonora to Quintana Roo and Chiapas (Fig. 2). A further five

262   localities were rescued from among the pool of data lacking geographic coordinates, but

263   that were inventoried completely to the level of $C \geq 0.8$ (Table 3).

264   Seeking 'hotspots' of sampling (i.e., sets of contiguous 0.05° pixels), we identified

265   an initial large number of such hotspots, again well-distributed across the country. Our

266   inspection of the data lacking geographic coordinates 'rescued' 54,369 records, augmenting

267   the initial data set considerably. Of the initial sampling hotspots, only three made the $C \geq$

268   0.9 completeness criterion (Xalapa, Veracruz; El Triunfo, Chiapas; Isla Cozumel, Quintana

269   Roo); another 17 fell at the lower $C \geq 0.8$ criterion (Table 4). These hotspots covered areas

270   ranging from 60.5-4325.8 km^2, considerably larger than the ~30 km^2 of the 0.05° pixels.

271   The hotspots, once again, ranged across the country, from southern Sonora to Chiapas and

272   Quintana Roo (Fig. 3).

273

274   **Discussion**

275   This contribution is an exploration of the utility of the existing digital accessible knowledge

276   (Sousa-Baena et al. 2013) for Mexican birds in identifying well-inventoried sites across the

277   country. We chose our cut-off of 1980 to coincide roughly with initiation of large-scale

278   climate change worldwide (IPCC 2013). In this way, the sites that we have identified

Comentado [AHA9]: It is interesting to note that El Triunfo includes a Biosphere Reserve. By law such reserves must have a Management Plan (MP) in place, which in turn should at least list species and quite possibly occurrences (although it has been shown that the overlap between MP and DAK is far for complete; see e.g. works by Pino-del-Carpio et al.). In most cases, data from MP would not fit a stringent criterion for DAK as set forth by Sousa-Baena et al. (cited in the paper's literature) in that although they are digital knowledge that can be accessed, they are not so in a format that does not require some form of digitization or reformatting (let's say it's less-than-straightforward DAK, but it is not dark data either!). However, this is also true from records that required digitization (as described in the Methods section) from literal locality descriptors. It seems clear that all four DAK sources used here all share a common characteristic which is that they were already some type of database, as different from other DAK sources which existed as just documents. But then, perhaps the intention of the paper, and therefore the summary and possibly the title, should emphasize this distinction.
This is not limited to El Triunfo; the list of hotspots in Table 4 also lists other sites potentially within BRs (e.g. Los Tuxtlas).

Código de campo cambiado

Comentado [AHA10]: It's somewhat of a stretch to substantiate the claim of initiation of major climate change by 1980 on the IPCC's AR5. Most changes described in that report do not roughly coincide with the 1980 ballpark date: they are claimed to have happen generally earlier. AR5 (as did all previous Assessment Reports) offer a wide range of dates for major symptoms and drivers, but CC itself is set to have started in earnest at an earlier date (mid-XXth) and its drivers even before that in the cited reference. Perhaps the most conspicuous one that is assumed to have been initiated around that decade is the northern polar warming, but that does not seem to cascade well down to tropical areas. While I see absolutely no problem in setting one particular date to establish the before-and-after line (one has to do that in order to establish a baseline, and certainly there are multiple comparisons pre- and post-eighties), the statement remains unsupported by that reference. I'd suggest to base the selection in AR5 as stated, but not because it "roughly coincides with the initiation of large-scale climate change worldwide" –rather, because the selected date offers a convenient cutpoint between sets of "initial changes" and [possibly runaway?] future changes down the road according to most scenarios worked out by IPCC.

Código de campo cambiado

279   provide baseline points of reference for species composition, for comparison with species

280   composition later, decades into the processes of global climate change (Tingley et al. 2009).

281   It is of course tempting to map the species diversity that has been documented or

282   that we have estimated for 0.05° grid squares and/or hotspots across the country, to obtain a

283   picture of species diversity countrywide. We have avoided this temptation, however, in

284   view of the highly non-random and scattered distribution of the well-inventoried sites

285   across the country—the sites do not cover all regions of the country at all consistently, such

286   that any such picture would be incomplete and potentially misleading. That step was, quite

287   simply, not among the objectives of this study.

288   Rather, our aim was to compile a catalog of sites across Mexico that have been

289   inventoried historically in detail, to the point that the inventory is more or less complete.

290   This catalog, in and of itself, is not of great interest scientifically; however, to the extent

291   that new inventories can be developed for comparison with the old ones, the interest in the

292   comparisons grows considerably. Echoing our earlier contribution (Peterson et al. 2015),

293   we are fascinated by the long-term processes of population biology and biogeography that

294   are leading to turnover of species at sites.

295   We have improved on our earlier contribution (Peterson et al. 2015), however,

296   which had to aggregate occurrence data to coarse resolutions (1°, or ~110 km) until

297   inventories were sufficiently complete. Here, in contrast, we have sought single sites (0.05°

298   grid squares) that are completely inventoried—this difference has the major advantage of

299   not including as much beta diversity in single-site inventories as did our previous study.

300   Our consideration of hotspots, to some degree, began to coarsen the spatial extent of the

301   sites once again, but offered a somewhat more extensive list of sites that have seen

302   thorough inventories, yet still across extents much smaller than in our previous work.

303 One criticism that can be leveled at this work is that some important data may have

304 been left out of the analysis. That is, the entire concept of digital accessible knowledge is

305 that the data are (1) in digital format, (2) accessible readily via the Internet, and (3)

306 integrated with the remainder of DAK via common portals (i.e., making the transition from

307 individual data points to integrated "knowledge"), as was emphasized in the original

308 publication presenting the idea of DAK (Sousa-Baena et al. 2013). [We note that a

309 subsequent publication (Meyer et al. 2015) used "digital accessible information," we

310 believe unfortunately, as they provided no justification for or even notice of the change of

311 terminology.] In the case of Mexican birds, for example, the Natural History Museum (UK)

312 has very few data in digital format, and none has been made accessible, such that important

313 collections from Mexico, like those assembled by Godman & Salvin (1879-1915), have not

314 been analyzed in this contribution. That is the blessing and the curse of DAK: data that are

315 digital and shared on global portals are used broadly, whereas data that do not meet the

316 DAK criteria are frequently not used at all.

317 The purpose of this paper is to enable a broad suite of repeat avifaunal survey

318 efforts across Mexico. In effect, with the maps and tables of this paper, we challenge the

319 ornithological community interested in Mexican birds to focus attention on these sites (we

320 provide a complete compendium of the well-inventoried sites and hotspots documented in

321 this paper, as well as associated shapefiles, in a dataset made available permanently at

322 http://hdl.handle.net/1808/20673). Not only does work at these sites provide information

323 about the current community composition there, but also about the change in those

324 communities through time.

325 We suggest and advocate that resurvey efforts take the form of two inventories at or

326 near the site: one in as exactly the original site as is possible to determine, and the other in

Comentado [AHA12]: Again, this showcases the risk with stringent-DAK as commented earlier: That of selective sampling. It could be argued that selective sampling also arises from the fact that data that are digital is always a non-random subset of existing data. But as has been demonstrated (see references commented earlier) the amount of knowledge that is not being considered when a set of (already digital) sources is left aside can actually be measured--and it is not negligible. Therefore I think that this discussion on the utility and convenience of stringent-DAK (as opposed to "sensu lato" DAK, which I do not equate to Carsten's definition but to something like "efficiently harvestable data") could be expanded a bit to properly defend why the third requirement (integrated and in database form).

327   the closest and most comparable site that still retains the vegetation type that was

328   represented at the site at the time of the original inventory efforts; the former reflects

329   effects of local-scale processes (e.g., habitat destruction, aridification), whereas the latter

330   reflects more global processes (e.g., climate change), and comparisons of the two resurvey

331   inventories will yield a rich understanding of the relative magnitude of effects of the local

332   and global processes in changing avifaunas across the country. Once several such sites have

333   been re-surveyed, a rich picture of the dynamics of Mexican bird distributions will emerge,

334   in much greater detail than the picture presently available (Peterson et al. 2015).

335   We are working to assemble what historical photographs exist for each of these

336   sites, in association with the original inventory efforts, to further enrich comparisons of

337   'before' and 'after.' For instance, the Nelson and Goldman expeditions across Mexico in

338   the late nineteenth and early twentieth centuries produced large quantities of images that

339   are well-documented in two summary volumes (Goldman 1951; Nelson 1921)—two

340   examples are shown in Fig. 4. A major complement to the repeat inventories that are

341   facilitated by the analyses in this study would be repeat photography to allow a clear view

342   of what sorts of landscape change have occurred at key sites (see, e.g., Spond et al. 2014).

343

Conclusions

345   Many rich insights can be gained from before-and-after comparisons to detect and

346   characterize patterns of change in biodiversity, yet such changes have been opaque to study

347   for lack either of the "before" or of the "after." This study explores a novel approach to

348   enabling such studies: we mine the existing DAK for Mexican birds to detect and document

349   well-inventoried sites, and provide a catalog of those well-known sites for others to use.

350   Developing repeated inventories at a series of sites across Mexico would yield a detailed,

Comentado [AHA13]: This is a highly suggestive idea. However—how can the vegetation continuity be assured? Changes in climate are expected to be reflected as changes in the environment—and therefore, the same baseline that has been set for birds should also be ensured for the vegetation. Does this baseline exist, or can it be obtained from DAK or other sources? Further down it is suggested to use archival photographs. A 2015 presentation at the SECEM conference by Escribano et al. explores, and puts to use, similar methods to observe historical vegetation changes that are of use to determine whether baselines depending on vegetation can be drawn (http://www.secem.es/wp-content/uploads/2016/01/Libro-XII-Congreso-SECEM.pdf, page 45; the work has been submitted as a paper to *Diversity & Distributions*.) But as discussed above, this type of information was not strictly-DAK although it was converted to that through some cost-effective digitization or, rather, transmogrification work. This is one thing that could enhance DAK and therefore gained knowledge.

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

351 controlled set of comparisons that would allow a view of avifaunal dynamics across the

352 country. Adding the dimension of repeated landscape photography to the repeated

353 inventories would enrich the insights still more.

354

355 **Acknowledgements**

358

359 **References**

360 **AOU. 1998.** *Check-list of North American Birds, 7th ed.* Washington, D.C.: American Ornithologists'

361   Union.

362 **Colwell RK. 1994–present.** *EstimateS*. Storrs: University of Connecticut.

363 **Colwell RK, and Coddington JA. 1994.** Estimating terrestrial biodiversity through extrapolation.

364   *Philosophical Transactions of the Royal Society of London B* 335:101-118.

365 **Gill F, and Donsker D.** 2014. IOC World Bird List, version 4.3; doi : 10.14344/IOC.ML.4.3.

366 **Godman FD, and Salvin O. 1879-1915.** *Biologia Centrali-Americana [published in 215 parts by*

367   *various authors]*. London: Taylor and Francis.

368 **Goldman EA. 1951.** Biological Investigations in Mexico. *Smithsonian Miscellaneous Collections*

369   115:1-476.

370 **Grixti JC, and Packer L. 2006.** Changes in the bee fauna (Hymenoptera: Apoidea) of an old field site

371   in southern Ontario, revisited after 34 years. *Canadian Entomologist* 138:147-164.

372 **Heller EL, Kerr KC, Dahlan NF, Dove CJ, and Walters EL. 2016.** Overcoming challenges to

373   morphological and molecular identification of *Empidonax* flycatchers: A case study with a

374   Dusky Flycatcher. *Journal of Field Ornithology*.

375   **IPCC. 2013.** *Climate Change 2013: The Physical Science Basis*. Cambridge: Cambridge University

376   Press.

377   **Lavergne S, Mouquet N, Thuiller W, and Ronce O. 2010.** Biodiversity and climate change:

378   integrating evolutionary and ecological responses of species and communities. *Annual*

379   *Review of Ecology, Evolution, and Systematics* 41:321-350.

380   **Leache AD, Helmer D-S, and Moritz C. 2010.** Phenotypic evolution in high-elevation populations of

381   western fence lizards (*Sceloporus occidentalis*) in the Sierra Nevada Mountains. *Biological*

382   *Journal of the Linnean Society* 100:630-641.

383   **Meyer C, Kreft H, Guralnick RP, and Jetz W. 2015.** Global priorities for an effective information

384   basis of biodiversity distributions. PeerJ.

385   **Monroe BLJ, and Sibley CG. 1993.** *A World Check-list of Birds.* New Haven: Yale University Press.

386   **Moritz C, Patton JL, Conroy CJ, Parra JL, White GC, and Beissinger SR. 2008.** Impact of a century of

387   climate change on small-mammal communities in Yosemite National Park, USA. *Science*

388   322:261-264. 10.1126/science.1163428

389   **Navarro-Sigüenza AG, and Peterson AT. 2004.** An alternative species taxonomy of Mexican birds.

390   *Biota Neotropica* 4:

391   http://www.biotaneotropica.org.br/v4n2/pt/abstract?article+BN02304022004.

392   **Nelson EW. 1921.** Lower California and its Natural Resources. *Memorias of the National Academy*

393   *of Sciences, Washington* 16:1-194.

394   **Nilsson SG, Franzén M, and Jönsson E. 2008.** Long-term land-use changes and extinction of

395   specialised butterflies. *Insect Conservation and Diversity* 1:197-207.

396   **Nunes MFC, Galetti M, Marsden S, Pereira RS, and Peterson AT. 2007.** Are large-scale

397   distributional shifts of the Blue-winged Macaw (*Primolius maracana*) related to climate

398   change? *Journal of Biogeography* 34:816-827.

Código de campo cambiado

399   **Olvera-Vital A. 2012.** *Avifauna del Municipio de Misantla, Veracruz. Tesis de Licenciatura*. Mexico:

400   Facultad de Ciencias, Universidad Nacional Autónoma de México.

401   **Peters JL. 1931-1987.** *Check-list of Birds of the World, vols. 1-16*. Cambridge: Harvard University

402   Press.

403   **Peterson AT, and Navarro-Sigüenza AG. 2005.** Hundred-year changes in the avifauna of the Valley

404   of Mexico, Distrito Federal, Mexico. *Huitzil* 7:4-14.

405   **Peterson AT, and Navarro-Sigüenza AG. 2006.** Consistency of taxonomic treatments: A response

406   to Remsen (2005). *Auk* 123:885-887.

407   **Peterson AT, Navarro-Sigüenza AG, Martínez-Meyer E, Cuervo-Robayo AP, Berlanga H, and**

408   **Soberón J. 2015.** Twentieth century turnover of Mexican endemic avifaunas: Landscape

409   change versus climate drivers. *Science Advances* 1:e1400071.

410   **Peterson AT, and Slade NA. 1998.** Extrapolating inventory results into biodiversity estimates and

411   the importance of stopping rules. *Diversity and Distributions* 4:95-105.

412   **Rubidge EM, Monahan WB, Parra JL, Cameron SE, and Brashares JS. 2011.** The role of climate,

413   habitat, and species co-occurrence as drivers of change in small mammal distributions

414   over the past century. *Global Change Biology* 17:696-708.

415   **Sánchez G. A. 1998.** *Misantla: Cultura, Tradición y Leyenda*. Veracruz: Asociación para el

416   Desarrollo Integral de la Región de Misantla, A.C.

417   **Scholes RJ, and Biggs R. 2005.** A biodiversity intactness index. *Nature* 434:45-49.

418   **Soberón J, and Llorente JE. 1993.** The use of species accumulation functions for the prediction of

419   species richness. *Conservation Biology* 7:480-488.

420   **Sousa-Baena MS, Garcia LC, and Peterson AT. 2013.** Completeness of Digital Accessible

421   Knowledge of the plants of Brazil and priorities for survey and inventory. *Diversity and*

422   *Distributions* 20:369-381.

423   **Spond MD, Grissino-Mayer HD, and Harley GL. 2014.** Comparing dynamics of mixed-conifer

424   woodlands on the Bandera Lava Flow, New Mexico. *Southwestern Naturalist* 59:235-243.

425   **Tingley MW, and Beissinger SR. 2009.** Detecting range shifts from historical species occurrences:

426   New perspectives on old data. *Trends in Ecology and Evolution* 24:625-633.

427   **Tingley MW, Monahan WB, Beissinger SR, and Moritz C. 2009.** Birds track their Grinnellian niche

428   through a century of climate change. *Proceedings of the National Academy of Sciences*

429   *USA* 106:19637-19643. 10.1073/pnas.0901562106

430

431

Table 1. Summary of initial data downloaded from each of four biodiversity data portals for Mexican vertebrate classes, and the relative redundancy of records in each, at the level of species x time (year, month, day) x place (geographic coordinates, textual descriptions). Note that subsequent data cleaning steps changed these initial tallies of redundancy, as they sought synonymies for taxon and place that may not have been visible in this initial step.

	GBIF	VertNet	REMIB	UNIBIO
Raw records	2,426,732	299,280	584,569	29,348
Unique records	1,917,800	226,004	431,240	23,446
Percent reduction	21.0	24.5	26.2	20.1

Table 2. Summary of individual 0.05° grid squares that are well inventoried ($C \geq 0.8$) across Mexico. Names of grid squares refer to the shapefile dataset summarizing the geographic distribution of these sites.

State	Name	Latitude	Longitude	N	S_{obs}	a	b	S_{exp}	C
Baja California	Isla Cedros	28.208	-115.244	213	45	12	10	51.0	0.88
	Isla San Benito	28.308	-115.594	398	60	19	11	74.2	0.81
Baja California Sur	El Triunfo 2	23.808	-110.094	267	56	15	9	66.5	0.84
	La Paz 2	24.158	-110.244	1167	137	35	34	154.0	0.89
	Laguna San Ignacio 2	27.308	-112.894	339	55	17	11	66.3	0.83
	San Jose de Cabo	23.058	-109.694	1339	174	47	25	215.5	0.81
	Sierra de la Laguna	23.458	-109.794	314	56	9	5	62.0	0.90
	Sierra de la Laguna 4	23.558	-109.944	813	55	14	6	68.0	0.81
Chiapas	El Fénix	16.658	-93.994	258	68	14	12	75.0	0.91
	Escuintla / Finca La Esperanza	15.358	-92.644	1362	229	45	31	259.9	0.88
	Laguna Ocotal	16.808	-91.444	531	125	21	30	131.8	0.95
	Pueblo Nuevo	17.158	-92.894	480	133	39	28	158.5	0.84
	San Cristóbal de las Casas	16.658	-92.594	339	69	17	11	80.3	0.86
	San Cristóbal de las Casas 3	16.708	-92.694	342	65	18	11	77.7	0.84
Chihuahua	Arroyo del Alamo	29.458	-106.794	214	76	21	17	87.6	0.87
	Arroyo el Mesteo	29.408	-106.894	208	76	22	21	86.4	0.88
	Colonia Pacheco	28.658	-106.194	283	63	17	9	76.6	0.82
	Colonia Pacheco 2	30.108	-108.294	945	103	21	9	124.0	0.83
	Rancho Pinos Altos	28.258	-108.294	259	49	8	5	53.6	0.91
	Rio El Gavilan	30.008	-108.394	325	83	19	15	93.7	0.89
Ciudad de México	Xochimilco	19.258	-99.094	229	83	30	21	102.7	0.81
Coahuila	Sabinas	27.858	-101.144	625	114	28	17	135.0	0.84
Colima	Cihuatlan	19.208	-104.544	354	89	24	15	106.2	0.84
Durango	El Salto	23.758	-105.544	205	55	12	16	58.9	0.93
	Neveros	23.758	-105.744	264	65	16	15	72.5	0.90
	Nievero	24.008	-104.744	215	52	12	12	57.1	0.91
	Santa Gertrudis	23.558	-104.394	254	86	27	17	105.4	0.82
	Santa Teresa	22.608	-104.844	264	68	16	10	78.9	0.86
	Tamazula	24.958	-106.944	225	99	37	29	121.1	0.82
Guanajuato	Rancho Enmedio	21.108	-101.194	401	82	22	14	97.4	0.84
Guerrero	Acahuizotla 2	17.358	-99.394	476	129	41	26	159.3	0.81
	Ajuchitlán	18.158	-100.494	219	40	9	4	47.2	0.85
	Chilpancingo	17.558	-99.494	6530	270	50	22	323.3	0.84
	Cuapongo	17.508	-99.644	1443	169	48	30	205.4	0.82
	Omiltemi 2	17.558	-99.644	2419	206	47	30	240.9	0.86

	Teotepec	17.458	-100.194	254	47	10	6	53.4	0.88
Jalisco	Atoyac	20.008	-103.594	316	107	31	24	125.5	0.85
	Atoyac 2	20.008	-103.544	223	72	20	15	83.8	0.86
	Lagos de Moreno 2	21.508	-101.694	202	70	24	15	87.2	0.80
	Tapalpa	19.958	-103.744	288	102	34	22	126.3	0.81
Michoacán	Apatzingan	19.158	-102.444	244	68	18	12	79.7	0.85
	Pátzcuaro	19.458	-101.594	361	111	35	23	135.7	0.82
	Rancho El Bonete	18.958	-101.894	297	95	30	20	115.6	0.82
	Tzitzio	19.608	-100.944	285	78	19	13	90.2	0.87
	Tzitzio 2	19.658	-100.894	299	97	32	22	118.5	0.82
	Uruapan	19.408	-101.994	391	94	26	14	115.6	0.81
	Zacapu	19.808	-101.794	475	126	34	30	144.1	0.87
Morelos	Cuernavaca	18.908	-99.244	541	159	50	32	196.1	0.81
México	East of Zitácuaro	19.408	-100.194	250	77	22	14	92.3	0.83
	Puerto Lengua de Vaca	19.258	-99.894	224	63	14	14	69.0	0.91
	Temascaltepec	19.058	-100.044	632	128	29	21	146.4	0.87
Nayarit	Islas Tres Marías	21.458	-106.444	394	64	15	15	70.5	0.91
	Islas Tres Marías 3	21.458	-106.394	475	68	20	12	82.6	0.82
	San Blas 3	21.558	-105.294	701	235	75	50	289.3	0.81
	Sauta	21.708	-105.144	370	107	34	20	133.6	0.80
	Tepic	21.258	-104.644	207	69	18	20	76.3	0.90
	Tepic 2	21.508	-104.894	1126	178	43	22	217.2	0.82
Oaxaca	Cerro San Felipe	17.158	-96.694	236	66	16	10	76.9	0.86
	Chivela	16.708	-94.994	251	73	20	16	84.1	0.87
	Palomares	17.108	-95.044	782	213	67	44	262.1	0.81
	Rancho San Carlos	17.208	-94.944	340	127	42	28	156.6	0.81
	Rio Molino	16.058	-96.444	374	84	17	16	92.0	0.91
	Totontepec	17.258	-96.044	533	107	15	30	110.4	0.97
Quintana Roo	Felipe Carrillo Puerto	19.558	-88.044	212	86	26	31	96.1	0.89
	Felipe Carrillo Puerto 2	19.608	-88.044	342	131	45	37	157.0	0.83
	Isla Cozumel 4	20.508	-86.944	650	154	17	49	156.7	0.98
Sinaloa	Babizos	25.758	-107.444	472	81	20	10	98.2	0.82
	Rancho Liebre	23.558	-105.844	601	126	26	21	140.7	0.90
Sonora	Babizos 2	27.008	-108.394	539	106	29	18	127.3	0.83
	Chinobampo	26.958	-109.294	215	74	22	22	84.0	0.88
	Hacienda de San Rafael	27.108	-108.694	220	61	18	14	71.2	0.86
	Huasa	28.608	-109.794	221	66	21	12	82.1	0.80
	La Chumata	29.908	-110.594	241	68	15	17	73.8	0.92
	Oposura	29.808	-109.694	494	124	27	24	138.0	0.90
	Rancho Guirocoba 2	26.958	-108.694	879	168	39	27	194.4	0.86
	Tecoripa	28.608	-109.944	242	75	21	18	86.0	0.87
Tamaulipas	Above Ciudad Victoria	23.708	-99.244	202	75	2	49	75.0	1.00
	Ciudad Victoria	23.708	-99.144	370	177	52	33	215.9	0.82

	Gomez Farías 3	23.058	-99.094	354	123	44	32	151.6	0.81
	Matamoros	25.858	-97.494	628	207	60	42	248.1	0.83
	Tampico	22.258	-97.844	1038	211	47	41	236.7	0.89
Veracruz	Balzapote 4	18.608	-95.044	3491	286	56	43	321.0	0.89
	Cerro Guzman	19.208	-96.394	542	141	43	29	171.0	0.82
	Coatepec	19.458	-96.944	377	122	16	49	124.4	0.98
	Coyame	18.408	-94.994	338	129	30	41	139.3	0.93
	Huatusco	19.158	-96.944	282	95	26	30	105.4	0.90
	Lago de Catemaco	18.408	-95.144	205	79	28	21	96.1	0.82
	Misantla	19.908	-96.844	307	108	3	62	108.0	1.00
	Perote	19.558	-97.244	263	88	14	44	90.0	0.98
	Presidio 2	18.658	-96.744	1000	219	54	36	257.6	0.85
	Rancho Caracol 3	18.458	-96.644	358	127	41	26	157.3	0.81
	Xalapa	19.558	-96.944	887	192	34	54	202.2	0.95
Yucatán	Chichén-Itzá	20.658	-88.594	1727	193	26	23	206.5	0.93
	Mérida	20.958	-89.644	242	118	38	25	144.9	0.81
	Xocompich	20.758	-88.544	270	115	44	40	138.0	0.83

Table 3. Summary of additional sites that were 'rescued' from among digital data lacking geographic coordinates, but that were detected based on unique locality descriptors.

State	Locality name	Latitude	Longitude	N	S_{obs}	a	b	S_{exp}	C
Nuevo León	Monterrey	25.687	-100.316	418	145	52	38	178.9	0.81
Tabasco	1 mi E Teapa	17.563	-92.948	361	158	57	41	195.9	0.81
San Luis Potosí	[Ciudad] Valles	21.997	-99.011	227	93	34	26	113.7	0.82
Chiapas	26 km N by road Ocozocuautla	16.995	-93.379	400	145	55	40	181.1	0.80
Oaxaca	1 mi SW Valle Nacional	17.757	-96.320	485	169	54	46	199.4	0.85

Table 4. Hotspots of sampling that are relatively completely inventoried (i.e., $C \geq 0.8$). Hotspot names correspond to the shapefile dataset summarizing the geographic distribution of these sites.

State	Hotspot	N	S_{obs}	a	b	S_{exp}	C
Chiapas	Arriaga Tonalá	2162	337	88	53	407.86	0.83
	Comitán	3821	422	79	62	470.9	0.89
	El Triunfo	4884	396	41	34	419.4	0.95
	Tuxtla Gutiérrez	4058	463	93	54	540.8	0.86
Ciudad de México	Valle de México	6818	349	71	28	434.7	0.80
Guerrero	Chilpancingo / Omiltemi	15230	380	51	24	431.0	0.88
Michoacán	Patzcuaro / Morelia / Lagos de Michoacán	2144	269	60	26	334.5	0.80
	Tancítaro / Uruapan	1396	253	50	28	295.2	0.86
Nayarit	Tepic / San Blas	5079	405	66	34	466.3	0.87
Oaxaca	Matías Romero to north	3994	404	84	42	485.0	0.83
	Miahuatlán / Mixtepec	1054	205	54	34	245.9	0.83
	Tapanatepec / Zanatepec / Cerro El Baúl	4752	355	63	28	422.3	0.84
Quintana Roo	Isla Cozumel	2159	216	39	46	231.8	0.93
Sinaloa	Durango Hwy / Espinazo del Diablo	3857	293	57	24	356.8	0.82
	Rosario	1912	247	55	35	288.2	0.86
Sonora	Álamos	3403	273	42	23	308.9	0.88
Veracruz	Los Tuxtlas	10499	457	78	41	528.5	0.87
	Orizaba / Córdoba	2741	402	94	52	484.4	0.83
	Xalapa	2135	303	54	49	331.6	0.91
Yucatán	Chichén-Itzá	2612	220	37	20	251.7	0.87

Fig. 1. Digital accessible knowledge of bird distributions (481,409 unique combinations of species x place x time) across Mexico prior to 1980, drawn from GBIF, VertNet, REMIB, and UNIBIO (records are not coded by source owing to frequent overlaps among sources in serving copies of the same record).
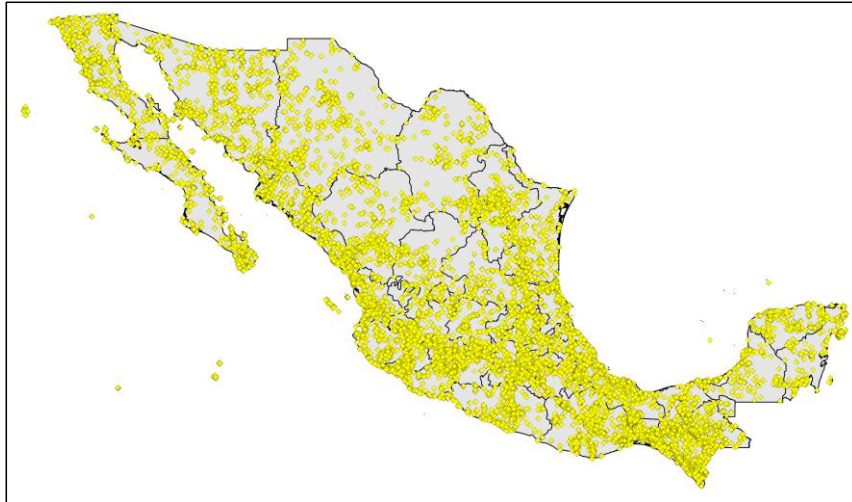
Fig. 2. Distribution of single 0.05° grid squares for which >200 records were available and completeness (*C*) was 0.8 ≤ *C* <0.9 (pink) or *C* > 0.9 (brown). Sites detected in Step 3 (i.e., single sites that are relatively complete, but that were not georeferenced prior to this study) are shown in blue (all had 0.8 < *C* < 0.9).
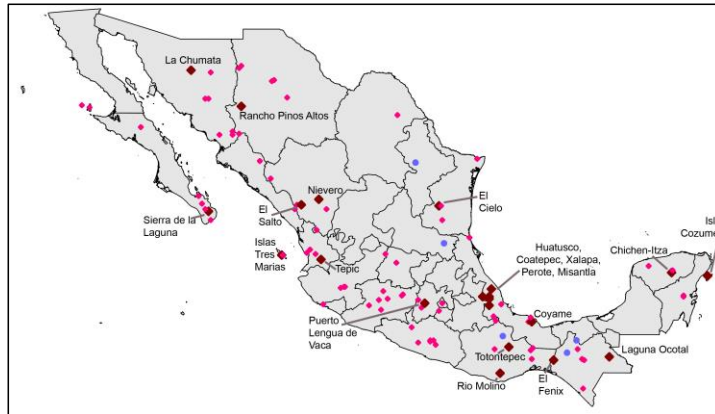
Fig. 3. Summary of "hotspots" of sampling of Mexican avifaunas based on the Getis-Ord Gi* statistic. Gray areas have $0.8 \leq C < 0.9$; red areas have $C > 0.9$.

Fig. 4. Photos of landscapes of two of the hotspots identified in this study from the Nelson-Goldman expeditions across Mexico in the late nineteenth and early twentieth centuries: above Xalapa, Veracruz (SIA2014-03203), Mt. Tancítaro, Michoacán (SIA2011-1908). Photos reproduced with permission of the Smithsonian Institution, Washington, D.C.

Appendix

Institutions contributing data to the analyses reported in this paper: Academy of Natural Sciences of Philadelphia; American Museum of Natural History; Angelo State University; Australian National Wildlife Collection; Bell Museum of Natural History; Biologiezentrum Linz; Bird Studies Canada; Borror Laboratory of Bioacoustics; Burke Museum, University of Washington; California Academy of Sciences; California State University, Chico; Canadian Museum of Nature; Carnegie Museum of Natural History; Chicago Academy of Sciences; CIIDIR-Oaxaca; Comisión Nacional para el Uso y Conocimiento de la Biodiversidad; Cornell Laboratory of Ornithology; Delaware Museum of Natural History; Denver Museum of Nature & Science; ECOSUR Chetumal; ECOSUR San Cristóbal; Emporia State University; ~~Estacion~~ Estación ~~Biologica~~ Biológica de Doñana; FES Zaragoza UNAM; Field Museum of Natural History; Florida Museum of Natural History; Fort Hays State University; HawkCount; Humboldt State University; Illinois State University; iNaturalist; INIFAP; Instituto de Biología, Universidad Nacional Autónoma de México; Instituto de ~~Ecologia~~Ecología, A.C., Xalapa; Instituto de Historia Natural de Chiapas; Instituto Humboldt, Colombia; Louisiana State University; Lund Museum; Michigan State University; Moore Laboratory of Zoology; ~~Musee~~ Musée de la ~~Vallee~~Vallée, Barcelonette, France; Musée George Sand et de la Vallée Noire; Museo de las Aves de México; Museo de Zoología "Alfonso L. Herrera," Facultad de Ciencias, Universidad Nacional Autónoma de México; Museo Nacional de Ciencias Naturales; Museu Paraense Emilio Goeldi; Museum für Naturkunde, Berlin; Museum Heineanum Halberstadt; Museum National d'Histoire Naturelle; Museum of Comparative Zoology, Harvard University; Museum of Evolution, Uppsala; Museum of Nature and Human Activities, Hyogo, Japan; Museum of Southwestern Biology; Museum of Vertebrate

Zoology; Museum Victoria, Australia; Natural History Museum (Bird Group, Tring); Natural History Museum of Los Angeles County; Naturalis, Amsterdam; Neotropical Ornithological Foundation; New York State Museum; North Carolina Museum of Natural Sciences; Ocean Biodiversity Information System; Ohio State University; Orcutt Trust Collection; Perot Museum of Nature and Science; Polish Academy of Sciences; Provincial Museum of Alberta; Queensland Museum; Royal Belgian Institute of Natural Sciences; Royal Ontario Museum; San Diego Natural History Museum; Santa Barbara Museum of Natural History; Senckenberg Museum~~Senkenberg Musum~~; Slater Museum of Natural History; South Australian Museum; Staatliche Naturhistorische ~~Sammlung~~ Sammlungen Dresden, Museum für Tierkunde; Staatliches Museum für Naturkunde, Stuttgart; Tall Timbers Research Station and Land Conservancy; Texas A&M University; Tulane University; U.S. National Museum of Natural History; Uberseemuseum, Bremen; Universidad Autónoma de Baja California; Universidad Autónoma de Baja California Sur; Universidad Autónoma de Campeche; Universidad Autónoma de Nuevo León; Universidad Autónoma de San Luis Potosí; Universidad Autónoma de Tamaulipas; Universidad Autónoma del Estado de Hidalgo; Universidad Autónoma del Estado de México; Universidad Autónoma del Estado de Morelos; Universidad de Ciencias y Artes de Chiapas; Universidad de Guanajuato; Universidad de Navarra; Universidad Juárez Autónoma de Tabasco; Universidad Juárez del Estado de Durango; Universidad Michoacana de San Nicolás de Hidalgo; University Museum of Zoology, Cambridge; University of Alabama; University of Alaska Museum; University of Alberta; University of Arizona; University of British Columbia; University of California, Davis; University of California, Los Angeles; University of Colorado; University of Iowa; University of Kansas; University of Michigan Museum of Zoology; University of Nebraska State Museum;

University of Oklahoma; University of Oslo; University of Texas El Paso; University of Wyoming; Utah Museum of Natural History; Washington State University; Western Foundation of Vertebrate Zoology; Western New Mexico University; Yale Peabody Museum; Yamashina Institute of Ornithology; Zoologische Staatssammlung München; and Zoologischen Sammlung der Universitat Rostock.