Digitizing the Frost Entomological Museum: Lessons learned and given

Andrew R. Deans & Emily L. Sandall
Department of Entomology, Penn State
~1,000,000(?) specimens
Welcome to InvertEBase: reaching back to see the future: species-rich invertebrate faunas document causes and consequences of biodiversity shifts

The rapid biodiversity change in North America has significant effects on essential ecosystem services, from impact on soil health and nutrient cycling, to agriculture, forestry and water quality. Effective monitoring of changes in biodiversity requires easy electronic access to historical specimen baseline information for temporal and regional species diversity comparisons, which can facilitate informed land management decisions. Vast amounts of specimen data are housed within the nation’s natural history collections, but most of these data are not yet readily accessible as digital resources. The TCN "InvertEBase" is a 4-year collaborative effort to digitize specimen records from ten arthropod and mollusk collections housed at six major US museums in six states, three of them ranking among the top 10 collections in the world. They include the Field Museum of Natural History (Chicago, IL), Cleveland Museum of Natural History (CMNH, OH), Auburn University Museum of Natural History (AUMNH, AL), University of Michigan Museum of Zoology (UM), Delaware Museum of Natural History (DMNH, DE); and Frost Entomological Museum at Pennsylvania State University (PA). InvertEBase will digitize de novo ~2.4 million georeferenced specimen records as well as integrate and mobilize data for 3.9 million terrestrial and aquatic arthropod and invertebrate specimens with special focus on the United States fauna. InvertEBase will greatly expand the taxon and geographic coverage of existing TCNs, and include the phylum Mollusca for the first time; DMNH, AUMNH, and CMNH will serve all of their invertebrate specimen data online for the first time. This project will significantly automate specimen data capture by utilizing optical character and voice-recognition technologies. The digitized data from this project will be immediately deployed for habitat-based distribution modeling and analyses.

This project made possible by the National Science Foundation awards EF-14-02667 to P. Sterwald and R. Bieler (Field Museum), EF-14-00993 to A. Deans (Penn State), EF-14-02697 to E. Shea (Delaware Museum of Natural History), EF-14-04984 to D. O’Foighil (Museum of Zoology, UM), EF-14-01176 to J. Bond (Auburn University), and EF-14-02785 to G. Svenson (Cleveland Museum of Natural History), and EF-14-01450 to J. Hanken (Harvard University). These awards are made as part of the National Resource for Digitization of Biological Collections through the Advancing Digitization of Biological Collections program and all data resulting from this award will be available through the national resource (IDigBio.org).
Lesson 1:

Digitization and collection improvement are best done *together*
Lesson 2:

Digitization of insects is *hard*
Lesson 3:

Digitization of insects can be easy!
<table>
<thead>
<tr>
<th>catalogNumber</th>
<th>institutionCode</th>
<th>basisOfRecord</th>
<th>countryCode</th>
<th>stateProvince</th>
<th>county</th>
<th>municipality</th>
<th>locality</th>
<th>decimal_latitude</th>
<th>decimal_longitude</th>
<th>geocodeDatum</th>
<th>samplingProtocol</th>
<th>eventDate</th>
<th>recordedBy</th>
<th>individualCount</th>
<th>identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENT432_ARB_001</td>
<td>PSUC</td>
<td>PreservedSpecimen</td>
<td>US</td>
<td>Pennsylvania</td>
<td>Centre</td>
<td>University Park</td>
<td>leading dock area of Headhouse III</td>
<td>40.803</td>
<td>-77.882</td>
<td>WGS84</td>
<td>hand collected</td>
<td>2018-09-13</td>
<td>Andrew R. Dears</td>
<td>2 adult</td>
<td>Andrew R.</td>
</tr>
<tr>
<td>ENT432_ARB_002</td>
<td>PSUC</td>
<td>PreservedSpecimen</td>
<td>US</td>
<td>Pennsylvania</td>
<td>Centre</td>
<td>Rockspring</td>
<td>pond near Phipps Agri. Museum</td>
<td>40.701</td>
<td>-77.992</td>
<td>WGS84</td>
<td>D rel</td>
<td>2019-08-24</td>
<td>Andrew R. Dears</td>
<td>1 larva</td>
<td>Andrew R.</td>
</tr>
</tbody>
</table>
Pinned specimens $\Rightarrow$ Darwin Core + spreadsheets

>127,000 occurrences @
Wet specimens => point and shoot + people power
Wet specimens => point and shoot + people power

>48,000 lots/specimens
Slide specimens => scan + slice + people power

>15,000 slides
Lesson 4:

Digitization is the *best* way to learn about a collection and to set priorities
Pennsylvania!
Pennsylvania!
Odonata!
Odonata!
Araneae?!
Tarantula from Lurifico Peru. Said to have caused the death of 4 men.
Lesson 5: Digitization inspires and catalyzes research
ARGIA VIOLACEA Hagen

♂ in cop. det. by G. H. Beatty, III

Darby Creek at Earle Estate
Radnor Twp., Delaware Co., PENNA.
17 August, 1946

coll. by G. H. Beatty, III

461635 COLLECTION OF G. H. & A. F. BEATTY
EDAPHIC FACTORS IN THE DISTRIBUTION OF PENNSYLVANIA ODONATA

G. H. & A. F. BEATTY
P. O. Box 281
State College, Pennsylvania 16801

ABSTRACT
Pennsylvania is examined as a habitat for Odonata from the standpoint of physiography, glacial characteristics, and watersheds of river systems, all being features distinct from climate. Known distributions of most of the 170 species comprising the Pennsylvania Odonata fauna were plotted with respect to these features, and the results are summarized and discussed.
Temporal differentiation in environmental niche modeling of Nearctic narrow-winged damselflies (Odonata: Coenagrionidae)

Emily L. Sandall1 and Andrew R. Deans1

1Frost Entomological Museum, Department of Entomology, Pennsylvania State University, University Park, PA

Corresponding author:
Emily L. Sandall
Email address: els22@psu.edu

ABSTRACT

Narrow-winged damselflies (Odonata: Coenagrionidae) can be observed in a variety of habitats, by both professional collectors and amateur entomologists. Their abundance and ease of recognition has resulted in a large amount of occurrence data, which can be used to establish species distribution maps through environmental niche modeling. Distributional data often aim to maximize the quantity of occurrence data available, but it is known that these datasets are not independent. We examined two datasets of Nearctic narrow-winged damselflies, and found that these two datasets were not independent of one another. We used occurrence data for twelve species of Coenagrionidae damselflies to generate niche models separated by time periods of specimen collection. Our study examined environmental niche models generated for four time periods for each of these coenagrionid species: Amphionion alpestris (Selys, 1874), Enallagma civile (Hagen, 1861), Chromagrion quadrimaculatum (Selys, 1874), Ischnura optima (Hagen, 1870), Enallagma civile (Hagen, 1861), Agris venusta (Hagen, 1861), Ischnura attenuata (Hagen, 1861), Ischnura elegans (Hagen, 1861), Ischnura annulata (Hagen, 1861), and Ischnura helvola (Hagen, 1861). The best supported models in each analysis were generated with occurrences collected from the 1970s to 2000s, and we used occurrence data outside of this range, from the 1860s to 2017, to compare the consistency of model predictions based on specimens of different time periods. In this approach, combining traditional environmental niche modeling and analysis of the specimen records themselves, we have found that ranges for narrow-winged damselflies expand over time, with increase in distributional coverage and decrease in model strength without temporal overlap between occurrences and environmental variables.

INTRODUCTION

Linking current and past occurrences on a large scale, to document species distribution over time requires that specimens be identified and accessible. The digitization of natural history collections (Graham et al., 2004; Page et al., 2015) has leveraged the occurrence and taxonomic determination data of millions of taxon specimens. Some of these efforts have captured the attention of taxonomists for centuries, resulting in extensive datasets that allow for a thorough analysis of shifting distributions, outbreaks, or other population changes (Pimm et al., 2000; Estabrook-Peck et al., 2015). Odonata, commonly known as dragonflies and damselflies, is one such order with extensive representation in natural history collections throughout time (Rother et al., 2016). Maximal amounts of occurrence data are used to establish species distribution typically without thorough examination of individual specimen’s temporal data. For many taxa, the quantity of digital data available may not be great enough to carry out such analyses. Thus, with widespread collection and digitization efforts enabling analysis of data associated with the specimen occurrence in order to build environmental niche models of distribution. In doing so, we can identify how concept of species distribution changes as specimens occurrences are documented.

The addition of recent observations and citizen science projects to digital repositories outside the reach of occurrence data beyond the walls of a single museum or individual researcher (Graham et al., 2004).
Lesson 6:

Digitization connects people to collections and to science
What can collection data tell us about mimicry?

- Batesian
- Müllerian
- imperfect
- aggressive
Spilomyia fusca Loew, 1864

Dolichovespula maculata (L., 1763)
Epistrophe grossulariae (Meigen, 1822)
Vouchering
Lesson 7:

Digitization connects people to urgent global problems
Lesson 8:

Digitization is *transformative*
Collections research @ PennState College of Agricultural Sciences

- Pollinator ID tools
- Documenting pollinators
- Monitoring and understanding invasive species
- Diversity of beneficial insects in applied systems
- Roles of microbes in mediating plant-insect interactions
Acknowledgments

This material is based upon work supported by the U. S. National Science Foundation, under Grant Numbers DBI-1349356 and EF-1400993. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

We thank Sal Anzaldo, Karah Roof, Brian Reily, Angela Hoover, Rachel Davis, Isa Betancourt, and Rafa Toro for their substantial contributions to collection health and digitization.
Questions?

DOI: 10.26207/86jq-hq45