

Digitization of the Beatty Odonata Collection at the Frost Entomological Museum (PSUC): the Terrain of Ecological Niche Modeling

Emily Sandall¹, Claire Jones², Briana Ezray³ , Andy Deans¹

¹Frost Entomological Museum, Penn State University, University Park, PA

²Geographic Information Systems (GIS) Collaboratory, DePaul University, Chicago, IL

³Department of Entomology, Penn State University, University Park, PA

George H. & Alice F. Beatty Collection

- Dragonflies and damselflies
- Donated in 2005
- ~62,000 specimens
 - Northeastern US
 - Mexico
- 1940s to 1970s
- 900+ species



The Beattys

- Found over 20 undescribed species in Mexico
- Described 2 species
 - *Arigomphus maxwelli* (Gomphidae) (Ferguson 1950)
 - *Enallagma basidens* (Coenagrionidae) larvae
- Developed the papered storage standard used in Odonata collections today
- Published 14 papers in the Proceedings of the Pennsylvania Academy of Sciences



Extreme Collecting Events



Digitization Goals

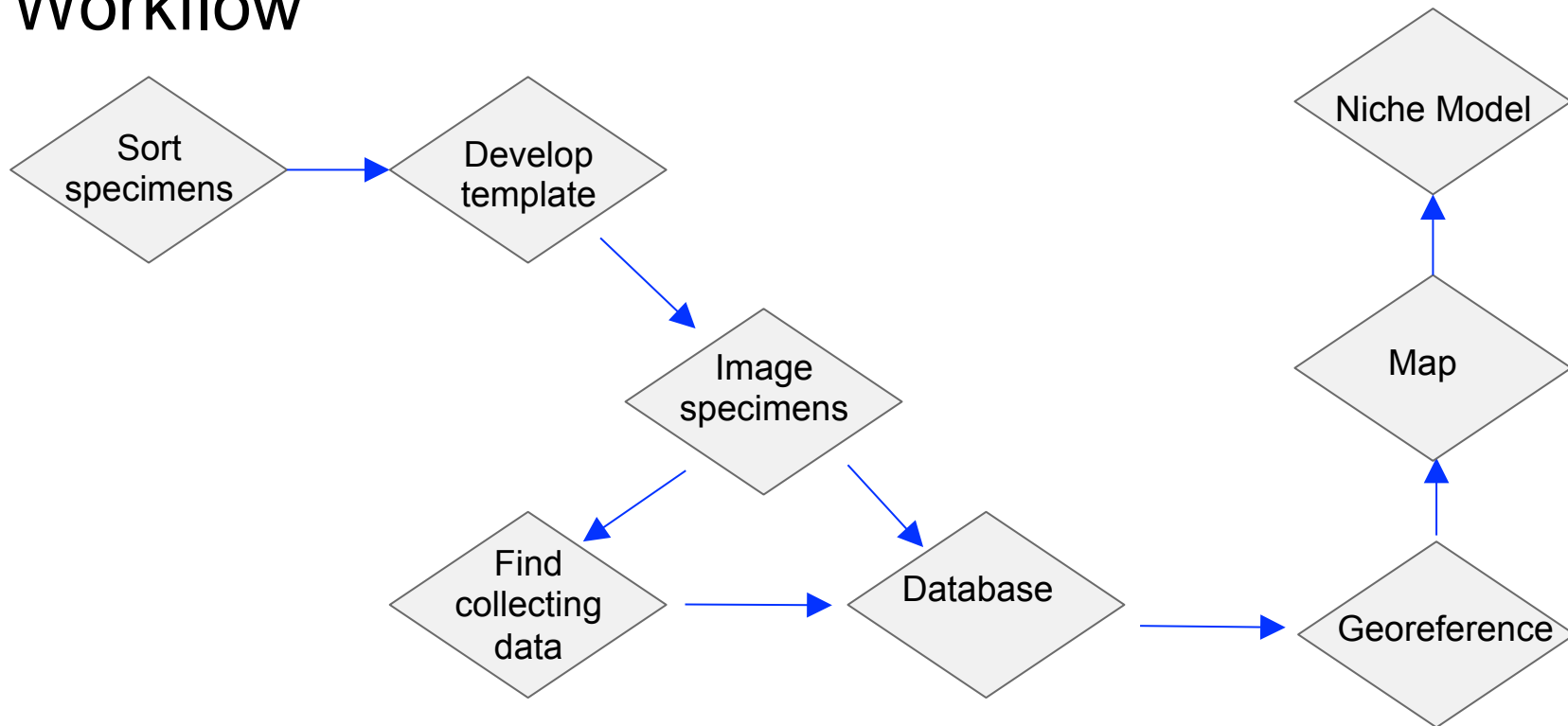
- Image specimens & upgrade their storage
- Create database of images
- Associate collecting event data
- Utilize data to create ecological niche models
- Locate any undescribed species in the collection



DBI-1349356
EF-1400993.

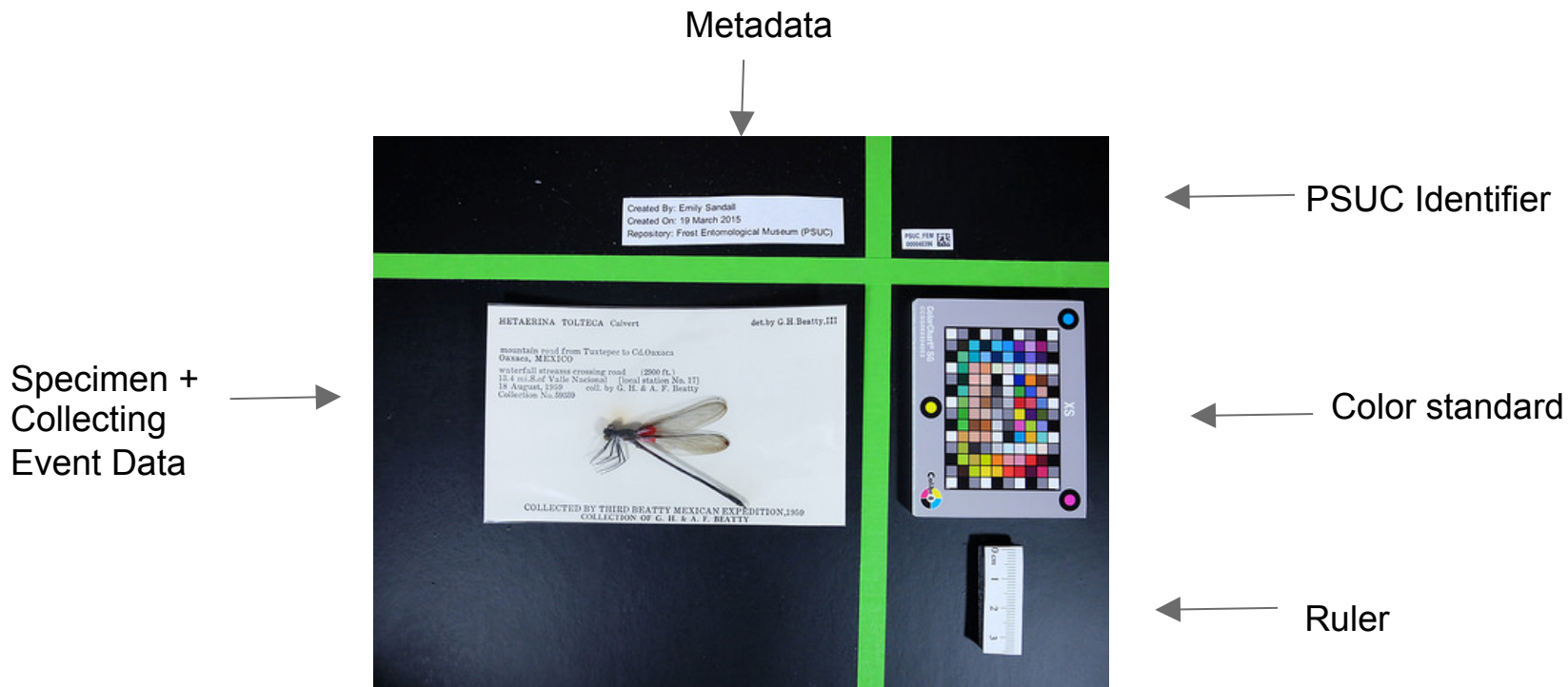


Workflow

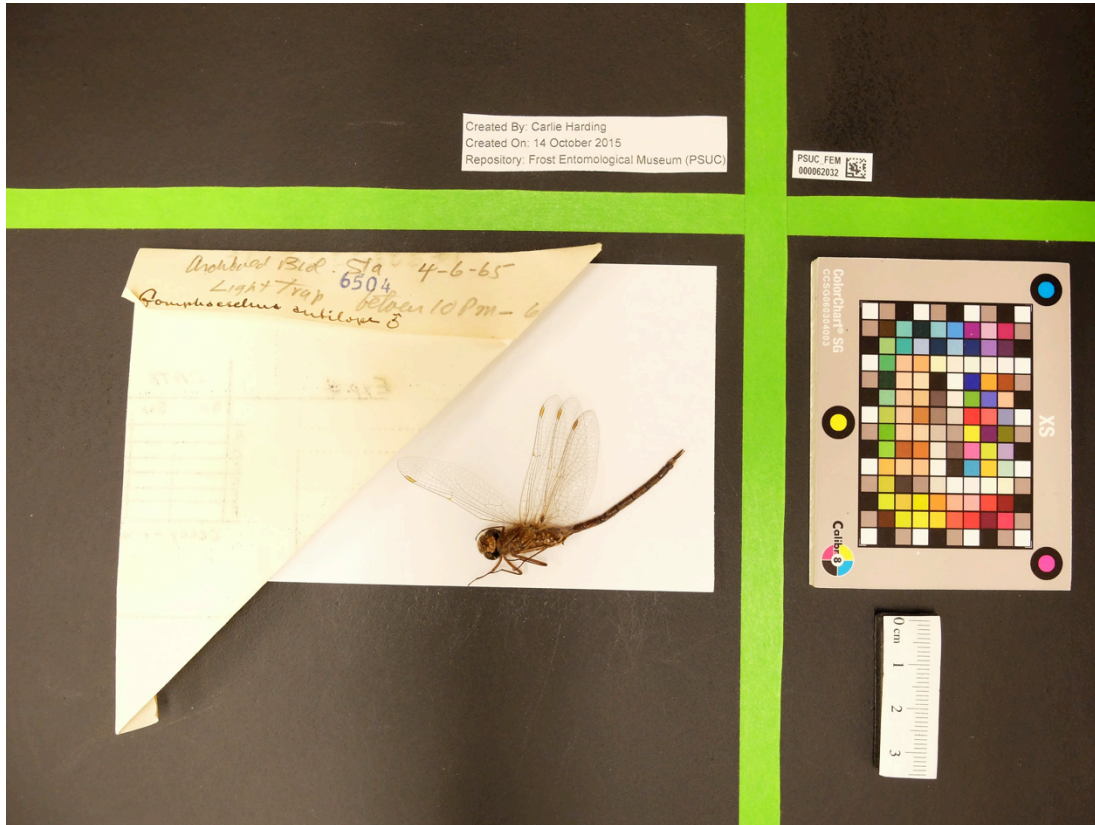


Specimen Imaging

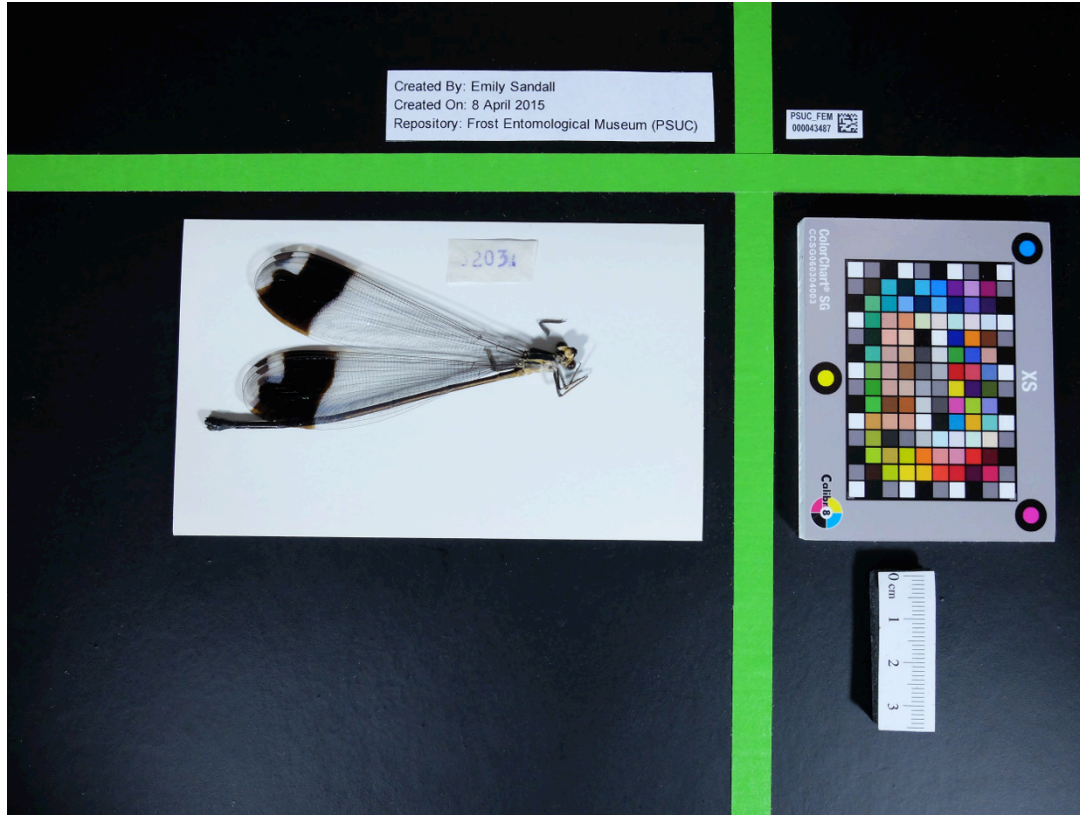
- 17,000+/60,000 specimens imaged



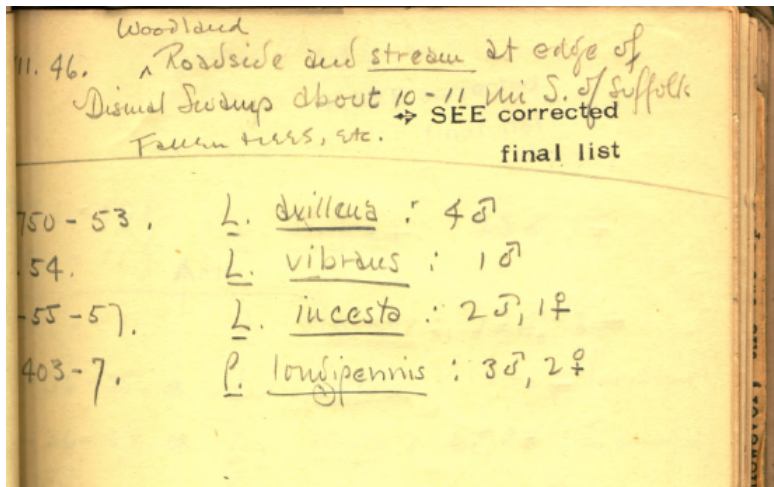
Variability in Specimen Data



Variability in Specimen Data



Associating Collecting Event Data



150' [Gaz] 63504 {LOC # 56 (1)
Rio Choapan, 14 mi. S. of Tuxtpec, OAX. 5.III.63.

upstream from bridge, for about 500' (probably much more, but not investigated) about 25' wide, entirely falling bed. Downstream bridge, which are usually vertical to about 3' or more) 100'. No perceptible flow (see below) but water in most places is 2-3' deep over a fairly flat bottom of sand & gravel; deeper holes under big trees & groves at water's edge. Extensive beds of "chara" up to 20 x 50" grow in mid up to the water's surface & breaking the surface to form thickets & narrow channels could sit on it. Ponds near the bridge, not so deep as those filled with chara, were choked with filamentous algae. Beds over deep (2-3') water often overgrown with ferns & vines as well as large trees here; margins of shallower, algae-filled pools with spruce ferns. Tree trunks & branches lying in water in several places. Sunlight penetration (at mid-day) to center of stream in most places, but banks are well shaded in most places.

Downstream from bridge, streambed is almost dry (for about 100' at least) with a few small shallow, algae-filled pools in a little trickling channel 11-2' wide with a flow of not more than 5 gal/min. This is the entire flow of Rio Choapan in this season; it is the same as it was in May, 1962. (Notice the drainage is in no way comparable with the exceedingly rich fauna of at least 4-1 spp. in May, 1962).

Woodland, including quite a few large trees, in the immediate vicinity of the stream - concentrated in cleared fields & populated areas beyond.

Temp. reached 118°F. about 3.00 P.M. - at end of collecting period 1.00-3.00 no clouds; slight breeze

ORTHEMIS FERRUGINEA - seen (lay AFB only)

3 PERITHEMIS "DOMITIA" : 3♂ } ♂♂ of Perithemis frequent on "big" (with long-
pendulous stem) red stream l. often in the shade
close to bank (Perithemis there were more abundant
"domitia") & also out in sun with Microthyria, etc.

1 P. "MOOMA" : 1♂

3 MICROTHYRIA AEGUALIS : 3♂ - frequent all along upper part of stream
(above bridge) especially where sunlight penetration was maximum
sitting on plants within 1' of water (unusual), on usual sticks, etc., or up in shrubs.

MICROTHYRIA sp. - a sp. larger than degenhousi seen (DIVINA?) HAGEN? see record sheet.

1 EGYPTRODIPLEX CONNATA : 1♂ - 5 or 6 other ♂♂ seen in place where tree-
trunks & branches nearly filled stream-bed.

3 HETAGRINA TITIA : 2♂ 1♀ - scarce; most of the Heteragrina seen (5-8 at most)
were concentrated on shrubs in shade overhanging the deep pool nearest the bridge
- but far above the point where the little trickle of flowing water began.

ARGIA - not seen! (see notes on p. 10)

* 1 TELEBASIS (ZETHEPINA) : 1♂ } telebasid common, with many pos. in cop., on surface
of "chara" beds.

5 T. SALVA : 4♂ 1♀

4 ENALLAGMA (N.) CULTELLATUM : 2♂, 2♀ - on "chara" beds - about a dozen seen
no other pos. in cop., in stream

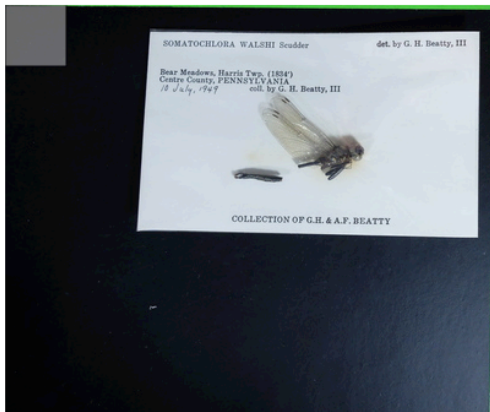
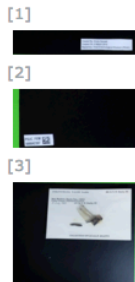
5 ACANTHAGRION [NIDUA] : 4♂ (1♀) } Acantagrion abundant on "chara" beds & on
shrubs, etc., along stream bank. The two spp.
* were not distinguished until being reported!

1 PROTONEURA (AWRANTIACA) : 1♂ (juv) oof.

30 spms. of 11 spp.

TaxonWorks

sqed depiction: [334](#) | [original image](#)



Curator metadata [Use \[a\]](#)

Created By: Emily Sandall Created On: 6 March 2015 Repository: Frost Entomological Museum (PSUC)

Identifier [Use \[s\]](#)

Annotated specimen [Use \[d\]](#)

SOMATOCHLORA WALSHI Scudder det-b-V G' H' Beatty'm Bear Meadows, Harris Twp. (1834') Centre County, PENNSYLVANIA coll. by G. H. Beatty, III \\ M (/L/(7, /?%,7 COLLECTION OF G.H. & A.F. BEATTY i'égiyu'fl W'''''' "I "MT

Identifiers

Namespace Identifier ✕

Namespace Identifier

Taxon Determinations

Name

Drag box to ocr.

Append [q], Replace [w]

Buffered collecting event

Bear Meadows, Harris Twp. (1834') Centre County, PENNSYLVANIA coll. by G. H. Beatty, III

Buffered determinations

Buffered other labels

Created By: Emily Sandall Created On: 6 March 2015

Repository: Frost Entomological Museum

Total Note Tag with

[Save changes](#) [Save and next \[n\]](#) [Skip to next](#)

Beginning of Niche Modeling in the Collection

- Temperature (1968)
- *Edaphic* factors (1969)
 - substrate characteristics
 - water quality
 - hydraulics
 - physiographic/glacial features
 - river basins

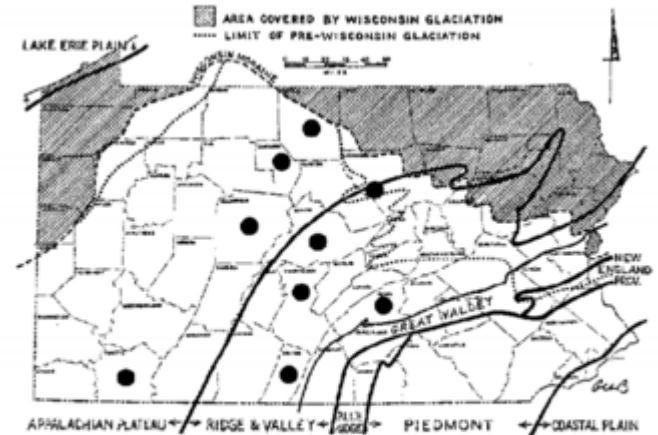


Figure 5. Distribution in Pennsylvania, by counties, of *Boyeria grafiana*.

Previous Methods

- Plot species by county
- Tabulate locality results
- Focus on species that seem to be confined to a particular province
- Overlay transparency of watersheds & drainage basins

116 *Calopteryx aquabilis* Say PENNSYLVANIA

PLACE	COUNTY	DATE	SPUNS.	COLLECTOR	COLL'N.	SOURCE
PENN CREEK, POE PADDY	MIFFLIN	1445, VII. 61.	1♀	HBW	HBW	HBW
"	CENTRE	29. VI. 63.	2♂, 2♀	RAK	RAK	RAK
McCONNELL'S NARROWS	UNION	4. VI. 60.	5♂, 1♀	"	"	"
" WEIKERT	"	4. VI. 60.	2♂, 1♀	HBW	HBW	HBW
PENN CR. AT POE PADDY	CENTRE	23. VI. 63.	4♂, 1♀, 1M.	" ^{1♀ reared} " D. S. WATSON "	"	"
"	"	25. VI. 57.	8♂, 4♀	D	D	D
"	"	7. VI. 57.	1♂	B	B	B
Penn Cr. at McConnell's Narrows	Union	9. IV. 61.	1♂, 1♀, 1(L)	CHB	CHB	CHB
		1. III. 69.	LARVAE	B et al	B	B
156♂ & 11♀ in cell	CNS?		2♀ reared	em. 3 & 9. IV. 69.		
Penn Cr., 0.5 mi. S. of Coburn	CENTRE	7. VI. 69.	1♂ SEEN			CNS

Beattys' Proposed Factors

- Temperature (aquatic and air)
- Edaphic factors considered:
 - Lotic/ Lentic
 - Lake Erie Plain
 - Appalachian Plateau
 - Ridge and Valley
 - Piedmont
 - Coastal Plain
 - Ohio River Basin
 - Susquehanna River Basin
 - Delaware River Basin

Table 4. ANALYSIS OF DISTRIBUTION OF 97 SPECIES OF PENNSYLVANIA ODONATA WITH RESPECT TO EDAPHIC FACTORS OF PHYSIOGRAPHY AND RIVER BASINS.¹

SPECIES # ON PA. CHECKLIST (2)	LOTIC (F), OR LENTIC (S)	PHYSIOGRAPHIC PROVINCES					RIVER BASINS		
		LAKE ERIE PLAIN	APPALACHIAN PLATEAU	RIDGE AND VALLEY	PIEDMONT	COASTAL PLAIN	OHIO R. BASIN	SUSQUEHANNA R. BASIN	DELAWARE R. BASIN
1.	<i>Gomphaeschna antilope</i>	S/F		AP			CP	0	D
2.	————— <i>fulviflata</i>	S/F		AP+	RV		(CP)	0	S
3.	<i>Baetiscaenia janata</i>	F		AP+	RV	P	CP	0	S
4.	<i>Boyeria graflana</i>	F		AP	RV		0	0	S
7.	<i>Eptaeschna heros</i>	S/F		AP+	RV	P	CP	0	S
8.	<i>Aeschna canadensis</i>	S		AP+	RV+	P	0	0	S
9.	————— <i>clepsydra</i>	S		AP				0	S
10.	————— <i>constricta</i>	S		AP+	(RV)		CP	0	S
11.	————— <i>mutata</i>	S	LE		RV		(LE)	0	S
12.	————— <i>tuberculifera</i>	S		AP	RV+		CP	0	S
14.	————— <i>verticalis</i>	S		AP+	RV+		CP	0	S
16.	<i>Anax longipes</i>	S		AP+	RV	P	CP	0	S
17.	<i>Zygopteryx thoreyi</i>	S		AP+	RV	P	0	0	S
18.	<i>Progomphus obscurus</i>	F		AP	RV		0	0	S
19.	<i>Hagenius brevistylus</i>	F	LE	AP+	RV	P	CP	0	S
20.	<i>Ophiogomphus carolinus</i>	F		AP+	RV+			0	S
22.	————— <i>howei</i>	F			RV			0	S
23.	————— <i>mainensis</i>	F		AP+	RV		0	0	S
24.	————— <i>rupinsulensis</i>	F		AP+	RV		0	0	S
25.	<i>Lanthus albitylus</i>	F		AP+	RV	P	0	0	S
27.	<i>Gomphus borealis</i>	S		AP+	RV		0	0	S
28.	————— <i>descriptus</i>	F		AP+	RV+		0	0	S
31.	————— <i>quadricolor</i>	F		AP	RV	P	0	0	S
32.	————— <i>spicatus</i>	S	LE	AP+			0	0	S
33.	————— <i>abbreviatus</i>	F			RV	P		0	S
34.	————— <i>brevis</i>	F		AP+	RV+		0	0	S
35.	————— <i>viridifrons</i>	F		AP			0	0	S
36.	————— <i>fraterculus</i>	F		AP	(RV)		0	(S)	
37.	————— <i>lineatifrons</i>	F			RV			0	S
38.	————— <i>rogersi</i>	F		AP	RV		0	0	S
39.	————— <i>vastus</i>	F		AP	RV	P	0	0	S
40.	————— <i>ventriosus</i>	F			RV	P		0	S

Our Question

What factors are influencing odonate distribution?

- Climatic?
- Edaphic?



Our Target Taxa

	% increase in available georeferenced data
<i>Argia tibialis</i> (Rambur, 1842) (Odonata: Coenagrionidae)	17.1%
<i>Boyeria grafiana</i> (Williamson, 1907) (Odonata: Aeshnidae)	14.8%
<i>Calopteryx aequabilis</i> (Say, 1839) (Odonata: Calopterygidae)	9.10%
<i>Erythrodiplax connata</i> (Burmeister, 1839) (Odonata: Libellulidae) *	194.7%
<i>Hetaerina cruentata</i> (Morse, 1895) (Odonata: Calopterygidae) *	160.4%



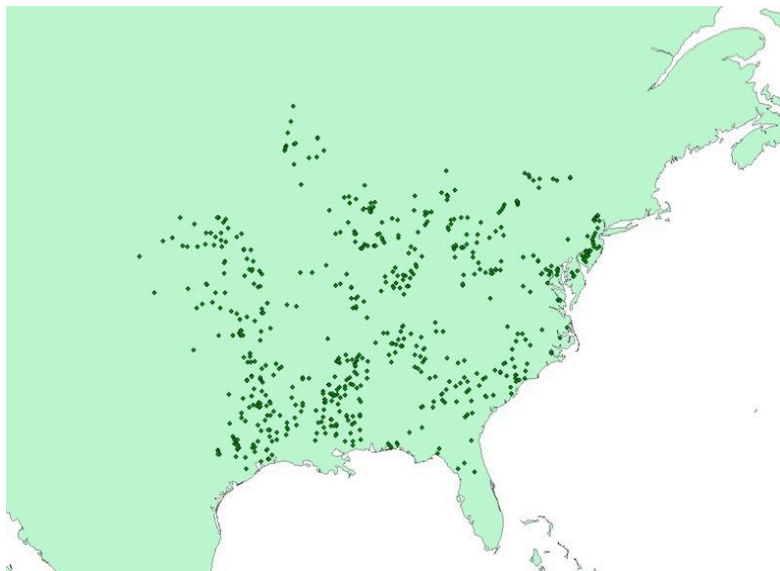
*=not in Beatty publications

Preparing Coordinate Data/Data Cleaning

- GEOLocate latitudes & longitudes
- CSV files of PSUC, GBIF, & Odonata Central collecting event data
- Total *Argia tibialis* records = 2089
 - coarse localities -0
 - coordinates georeferenced to centroid of county - 957
 - non-specific localities -126
- Resultant data = 1006

ArcGIS

- Shapefiles created based on specimen collecting event localities
 - BIOCLIM variables as ESRI grid files at 30' historical data from 1961-1990 layered over the specimen points enabled downloads of raster data at each locality



BIOCLIM Parameters-Original Variables

Variable Number	Variable	Minimum temp (°C)	Maximum temp (°C)	Rainfall (mm month ⁻¹)	Radiation (W m ⁻² d ⁻¹)	Pan evaporation (mm d ⁻¹)	Ref.
Bio01	Annual mean temperature (°C)	×	×				1
Bio02	Mean diurnal temperature range (mean(period max-min)) (°C)	×	×				1
Bio03	Isothermality (Bio02 ÷ Bio07)	×	×				1
Bio04	Temperature seasonality (C of V)	×	×				1
Bio05	Max temperature of warmest week (°C)		×				1
Bio06	Min temperature of coldest week (°C)	×					1
Bio07	Temperature annual range (Bio05-Bio06) (°C)	×	×				1
Bio08	Mean temperature of wettest quarter (°C)	×	×	×			1
Bio09	Mean temperature of driest quarter (°C)	×	×	×			1
Bio10	Mean temperature of warmest quarter (°C)	×	×				1
Bio11	Mean temperature of coldest quarter (°C)	×	×				1
Bio12	Annual precipitation (mm)			×			1
Bio13	Precipitation of wettest week (mm)			×			1
Bio14	Precipitation of driest week (mm)			×			1
Bio15	Precipitation seasonality (C of V)			×			1
Bio16	Precipitation of wettest quarter (mm)			×			1
Bio17	Precipitation of driest quarter (mm)			×			1
Bio18	Precipitation of warmest quarter (mm)	×	×	×			1
Bio19	Precipitation of coldest quarter (mm)	×	×	×			1
Bio20	Annual mean radiation (W m ⁻²)				×		1

BIOCLIM Parameters-Added Variables

Bio20	Annual mean radiation ($W m^{-2}$)					x		1
Bio21	Highest weekly radiation ($W m^{-2}$)					x		1
Bio22	Lowest weekly radiation ($W m^{-2}$)					x		1
Bio23	Radiation seasonality (C of V)					x		1
Bio24	Radiation of wettest quarter ($W m^{-2}$)			x		x		1
Bio25	Radiation of driest quarter ($W m^{-2}$)			x		x		1
Bio26	Radiation of warmest quarter ($W m^{-2}$)	x	x			x		1
Bio27	Radiation of coldest quarter ($W m^{-2}$)	x	x			x		1
Bio28	Annual mean moisture index			x			x	1
Bio29	Highest weekly moisture index			x			x	1
Bio30	Lowest weekly moisture index			x			x	1
Bio31	Moisture index seasonality (C of V)			x			x	1
Bio32	Mean moisture index of wettest quarter			x			x	1
Bio33	Mean moisture index of driest quarter			x			x	1
Bio34	Mean moisture index of warmest quarter	x	x	x			x	1
Bio35	Mean moisture index of coldest quarter	x	x	x			x	1
Bio36	First principal component of the first 35 Bioclim variables	x	x	x		x	x	2
Bio37	Second principal component of the first 35 Bioclim variables	x	x	x		x	x	2
Bio38	Third principal component of the first 35 Bioclim variables	x	x	x		x	x	2
Bio39	Fourth principal component of the first 35 Bioclim variables	x	x	x		x	x	2
Bio40	Fifth principal component of the first 35 Bioclim variables	x	x	x		x	x	2

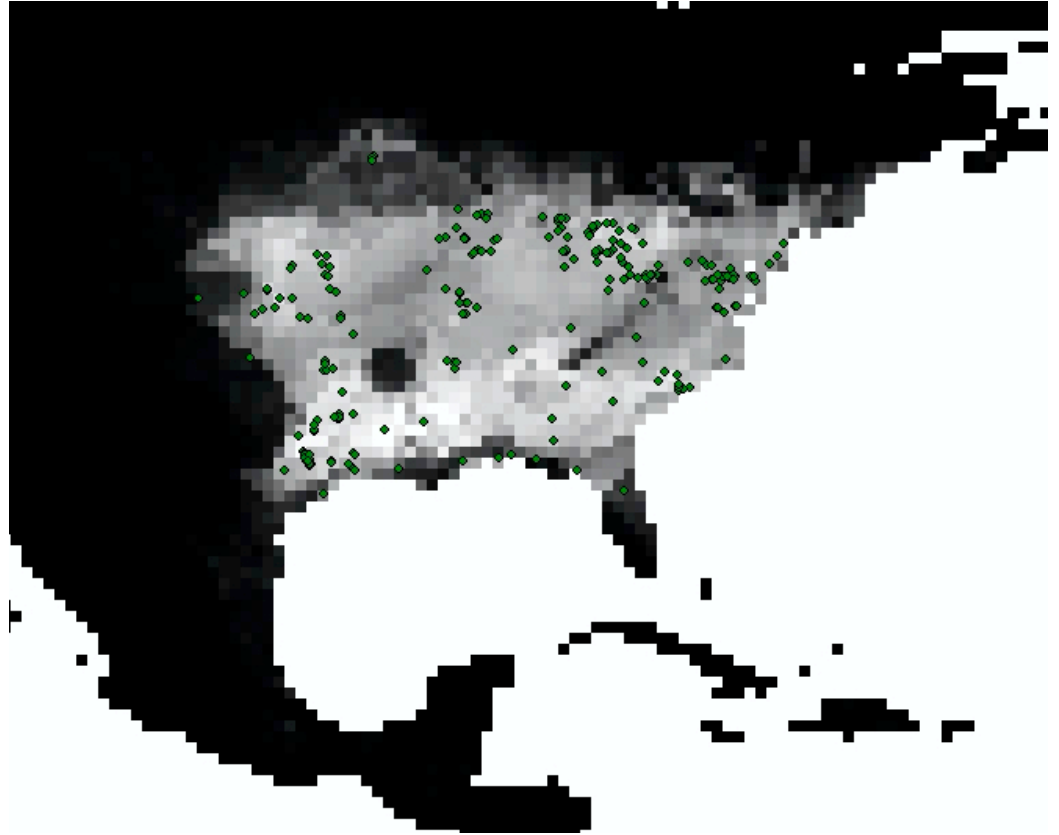
SDMToolbox

- Optimizes extracting data from each parameter in one step
- Enables testing for factor correlation prior to modeling
- Estimates biodiversity metrics
- Creates bias files to use as background points
- Chooses best model

MaxEnt

- Models with presence-only data
- CSV of species
- Utilize ASCII files of BIOCLIM variables at each set of coordinates as the parameters
- Output includes:
 - Omission & Predicted Area
 - Sensitivity vs Specificity
 - Analysis of Variable Contributions

MaxEnt-Map Representation



MaxEnt-Analysis of Variable Contributions

- Percent contribution-depends on path of the training algorithm to get to the model
- Permutation importance-depends on solely the final MaxEnt model
- Jackknifing examines each variable in isolation as well as the effect on the model without the isolated variable

Variable	Percent contribution	Permutation importance
bio10clip	42.4	11.7
bio16clip	10.8	0.7
bio20clip	7.4	6
bio01clip	6	6
bio12clip	5.1	0

Preliminary Results

- Temperature, solar radiation as major factors for all 5 species



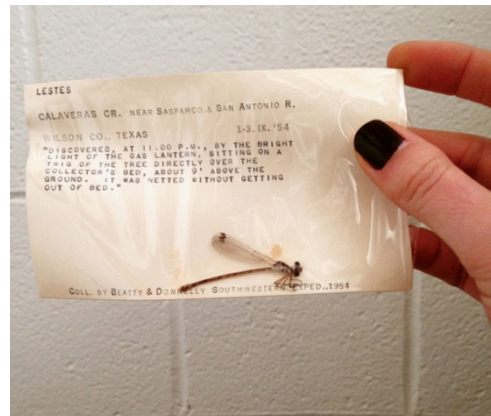
Back to the Beattys

- Precipitation and temperature (climatic) factors play a greater role
- How can more collecting event data be incorporated?
- Would pinned specimen data enhance our understanding?



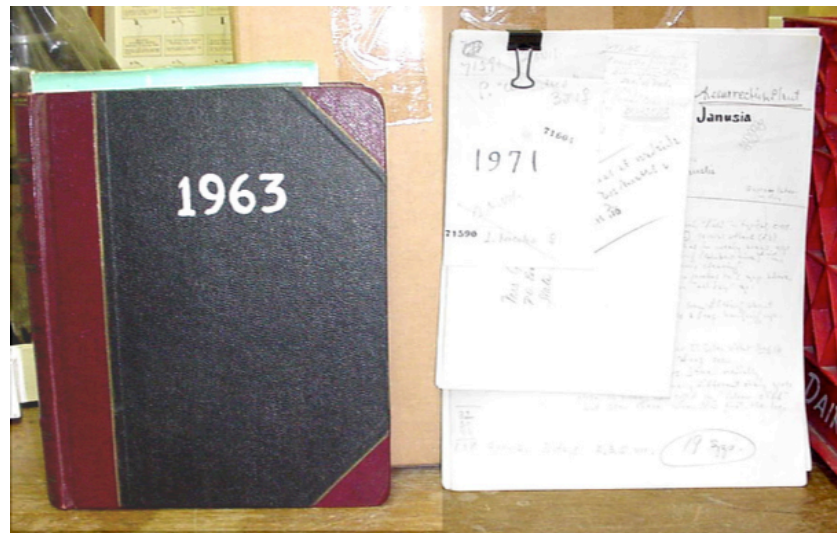
Adjusting Our Focus

- Prioritize digitization of specimens with easily associated collecting event data
- Transparency in collecting event data
 - Ensure that all uploaders understand best practices for georeferencing
 - Vetting records for quality locality data
- Incorporate absence data (available for some species)
- Narrow parameters for modeling
 - add edaphic parameters



Future Projects

- Digitizing field notes to free collecting event data
- Resampling localities for spatiotemporal study of Pennsylvania Odonata
- Sampling bias in the collection



Opportunities

- Development of an internship to work on digitization and curation of the Beatty Collection-stay tuned!



Acknowledgments



<http://dx.doi.org/10.6084/m9.figshare.1602235>



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.

FROST ENTOMOLOGICAL MUSEUM



Carlie Harding



Marcus Thaw



Michael Przybys



István Mikó



Sam Kurth



PennState



Hal White

TaxonWorks



Matt Yoder