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

Emily Sandall[1], Claire Jones[2], Briana Ezray[3] , Andy Deans[1]

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

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

[3]Department 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
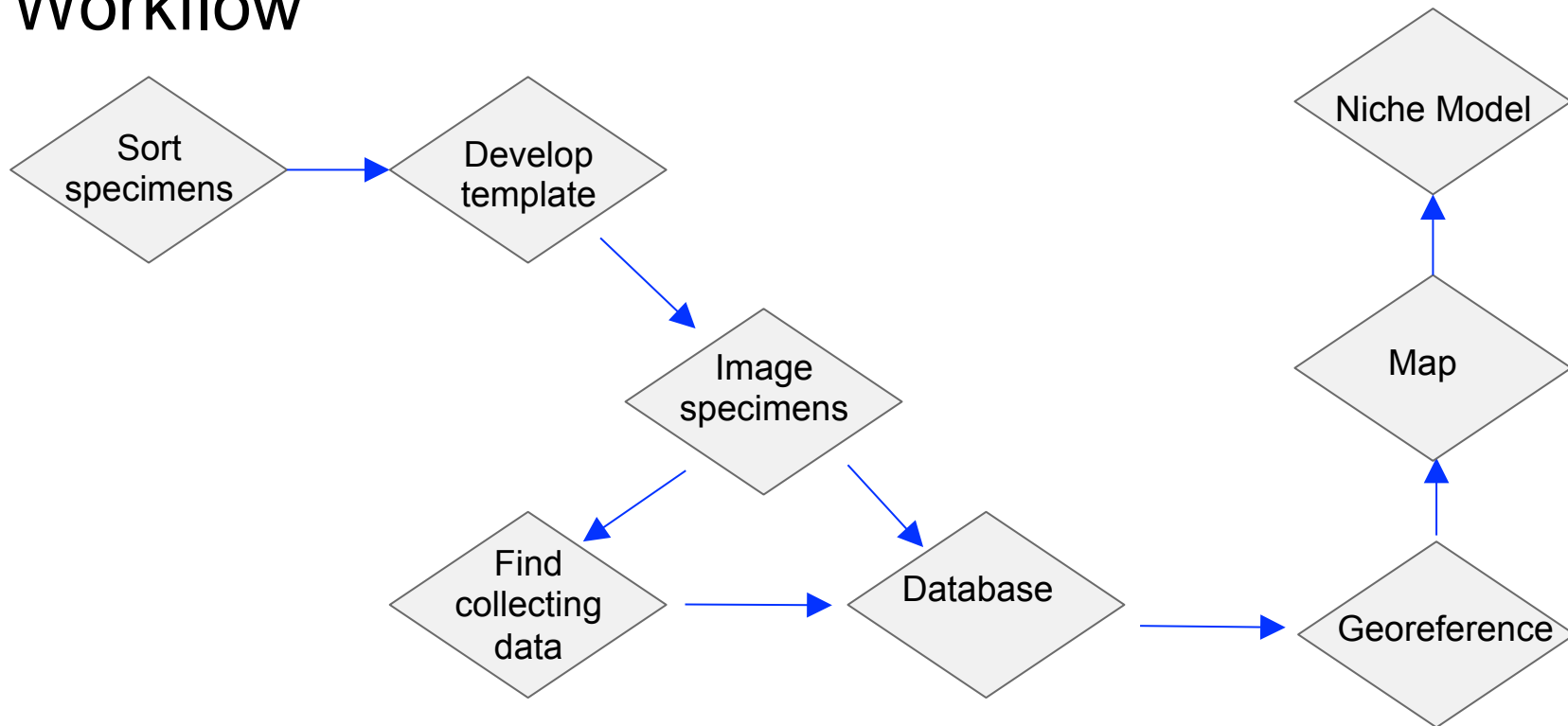
Extreme Collecting Events

Digitization Goals

- Image specimens & upgrade their storage

- Create database of images

- Associate collecting event data

- Utilize data to create ecological niche models

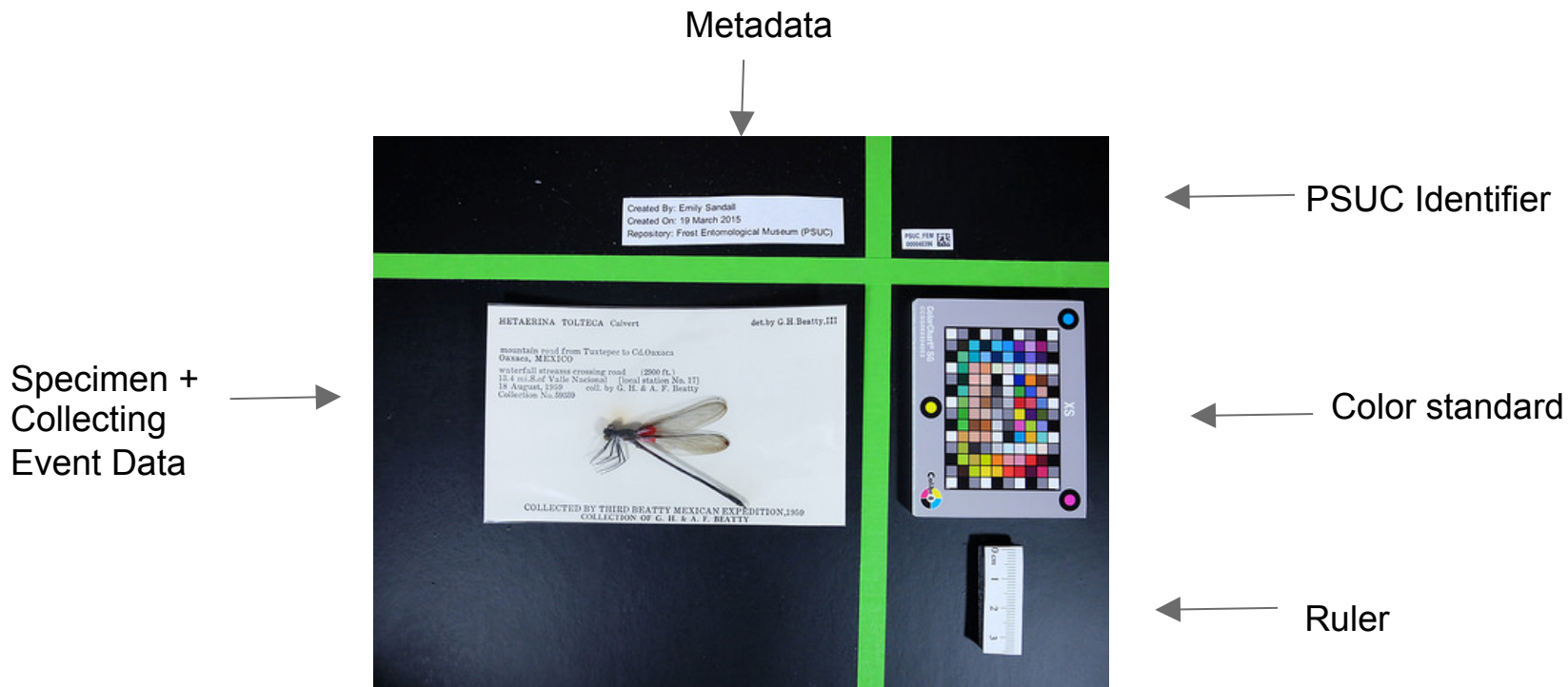- Locate any undescribed species in the collection
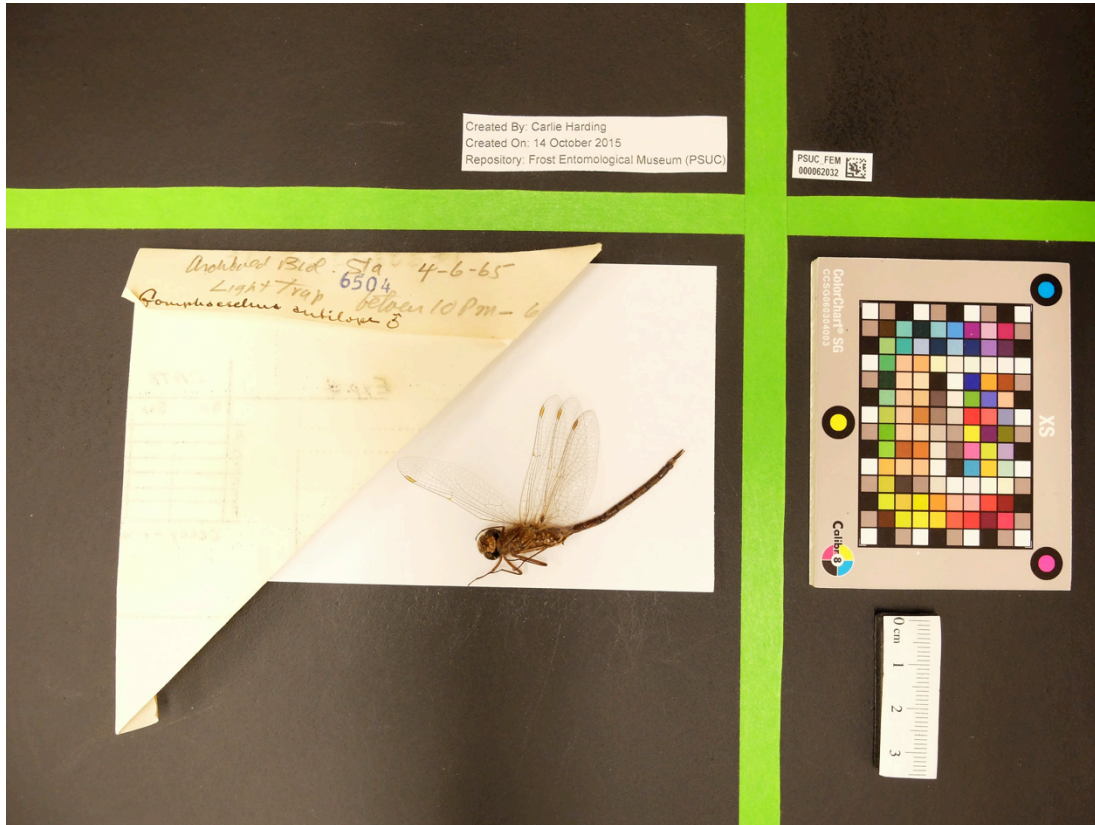
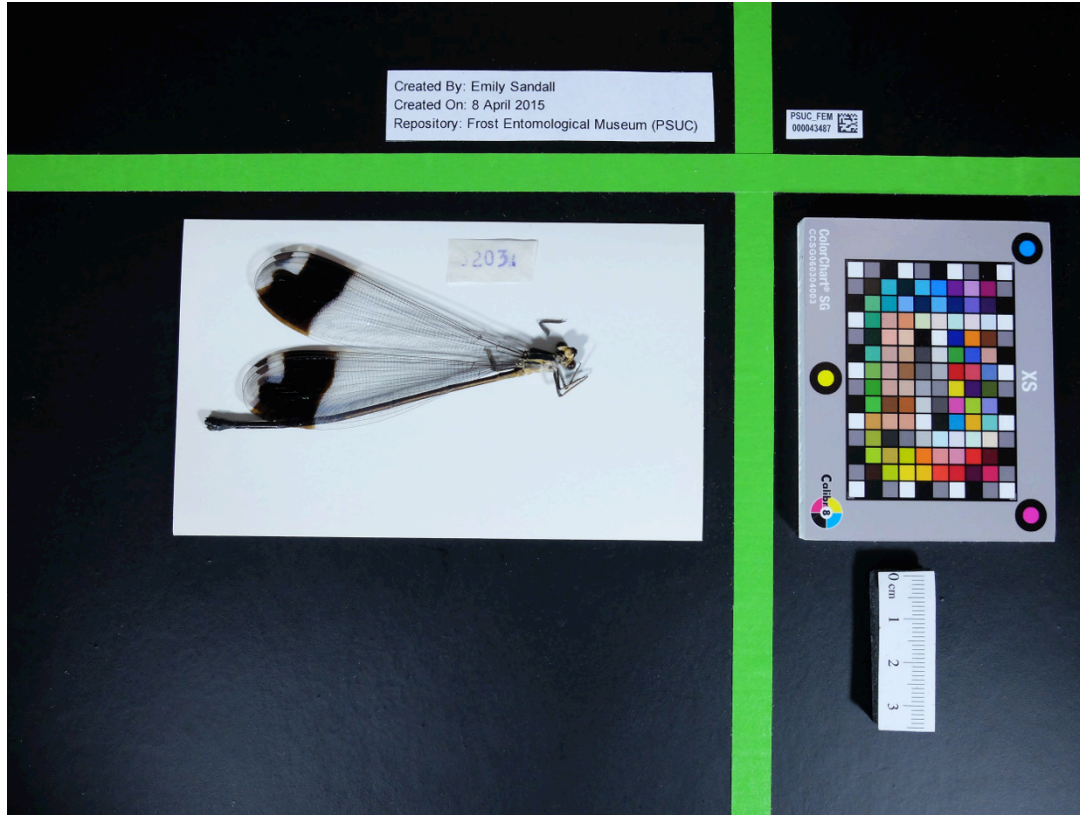

DBI-1349356
EF-1400993.
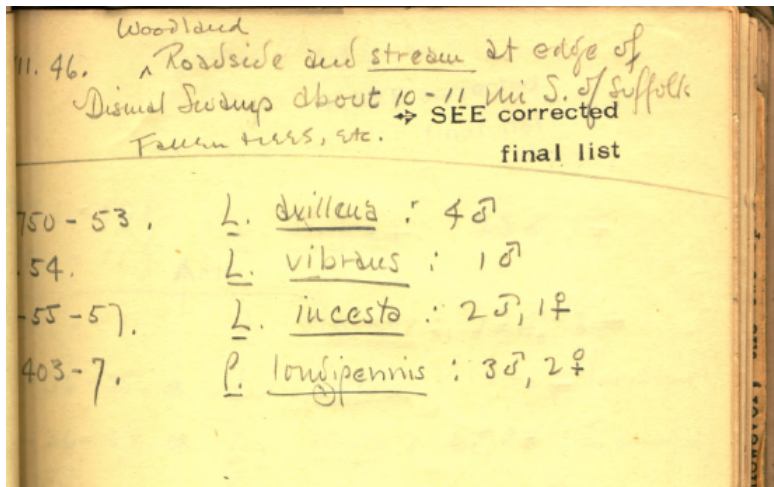
Workflow

Specimen Imaging

- 17,000+/60,000 specimens imaged

Metadata

PSUC Identifier

Specimen +
Collecting
Event Data

Color standard

Ruler

Variability in Specimen Data

Variability in Specimen Data

Associating Collecting Event Data

TaxonWorks

[1]

[2]

[3]

Drag box to ocr.

Append [q], Replace [w]

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'éigyiu'fl W^^^'' "I "MT

Identifiers

Taxon Determinations

Namespace | Identifier
PSUC_FEM | 000042397 ✗

Name
Select a otu

Namespace | Identifier
Select a namespace |

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 1   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
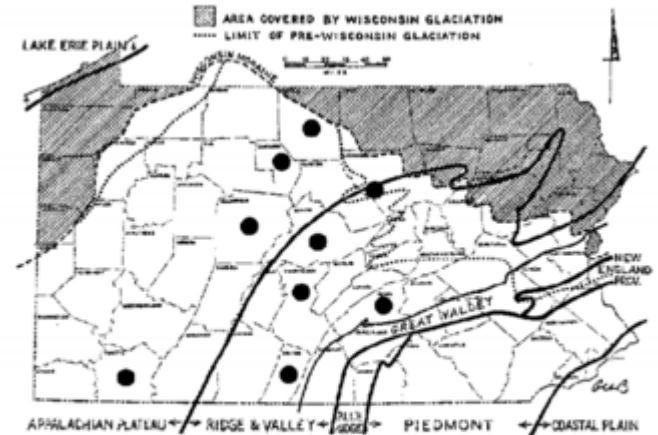
Figure 5. Distribution in Pennsylvania, by counties, of Boyeria grafiana.

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

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

Beatty & Beatty 1969

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

| 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.	Gomphaeschna antilope	S/F		AP			CP	O		D
2.	———— furcillata	S/F		AP+	RV		(CP)	O	S	(D)
3.	Basiaeschna janata	F		AP+	RV	P	CP	O	S	D
4.	Boyeria grafiana	F		AP	RV			O	S	
7.	Epiaeschna heros	S/F		AP+	RV	P	CP	O	S	D
8.	Aeshna canadensis	S		AP+	RV+	P		O	S	D
9.	———— clepsydra	S		AP					S	D
10.	———— constricta	S		AP+	(RV)		CP	O	S	D
11	———— mutata	S	LE		RV			(LE)	S	
12.	———— tuberculifera	S		AP	RV+		CP	O	S	D
14.	———— verticalis	S		AP+	RV+		CP	O	S	D
16.	Anax longipes	S		AP+	RV	P	CP	O	S	D
17.	Tachopteryx thoreyi	S		AP+	RV	P		O	S	D
18.	Progomphus obscurus	F		AP	RV			O	S	
19.	Hagenius brevistylus	F	LE	AP+	RV	P	CP	O	S	D
20.	Ophiogomphus carolus	F		AP+	RV+				S	D
22.	———— howei	F			RV				S	
23.	———— mainensis	F		AP+	RV			O	S	D
24.	———— rupinsulensis	F		AP+	RV			O	S	D
25.	Lanthus albistylus	F		AP+	RV	P		O	S	D
27.	Gomphus borealis	S		AP+	RV			O	S	D
28.	———— descriptus	F		AP+	RV+			O	S	D
31.	———— quadricolor	F		AP	RV	P		O	S	D
32.	———— spicatus	S	LE	AP+				O	S	D
33.	———— abbreviatus	F			RV	P			S	D
34.	———— brevis	F		AP+	RV+			O	S	D
35.	———— viridifrons	F		AP				O		
36.	———— fraternus	F		AP	(RV)			O	(S)	
37.	———— lineatifrons	F			RV				S	
38.	———— rogersi	F		AP	RV			O	S	
39.	———— vastus	F		AP	RV	P		O	S	D
40.	———— ventricosus	F			RV	P			S	

Our Question

What factors are influencing odonate distribution?

- Climatic?

- Edaphic?

Our Target Taxa

% increase in available georeferenced data

Argia tibialis (Rambur, 1842) (Odonata: Coenagrionidae)   17.1%

Boyeria grafiana (Williamson, 1907) (Odonata: Aeshnidae)   14.8%

Calopteryx aequabilis (Say, 1839) (Odonata: Calopterygidae   9.10%

Erythrodiplax connata (Burmeister, 1839) (Odonata:Libellulidae) *   194.7%

Hetaerina cruentata (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
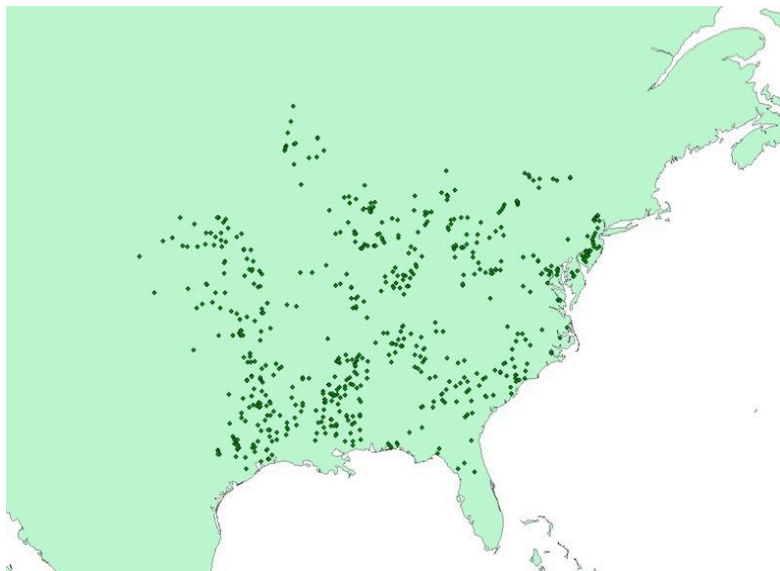
BIOCLIM Parameters-Original Variables

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

Hutchinson et al. 2009

BIOCLIM Parameters-Added Variables

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

Kriticos et al. 2014.

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

Brown J.L. 2014

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

Phillips et al 2004, Wisz et al 2008

MaxEnt-Map Representation
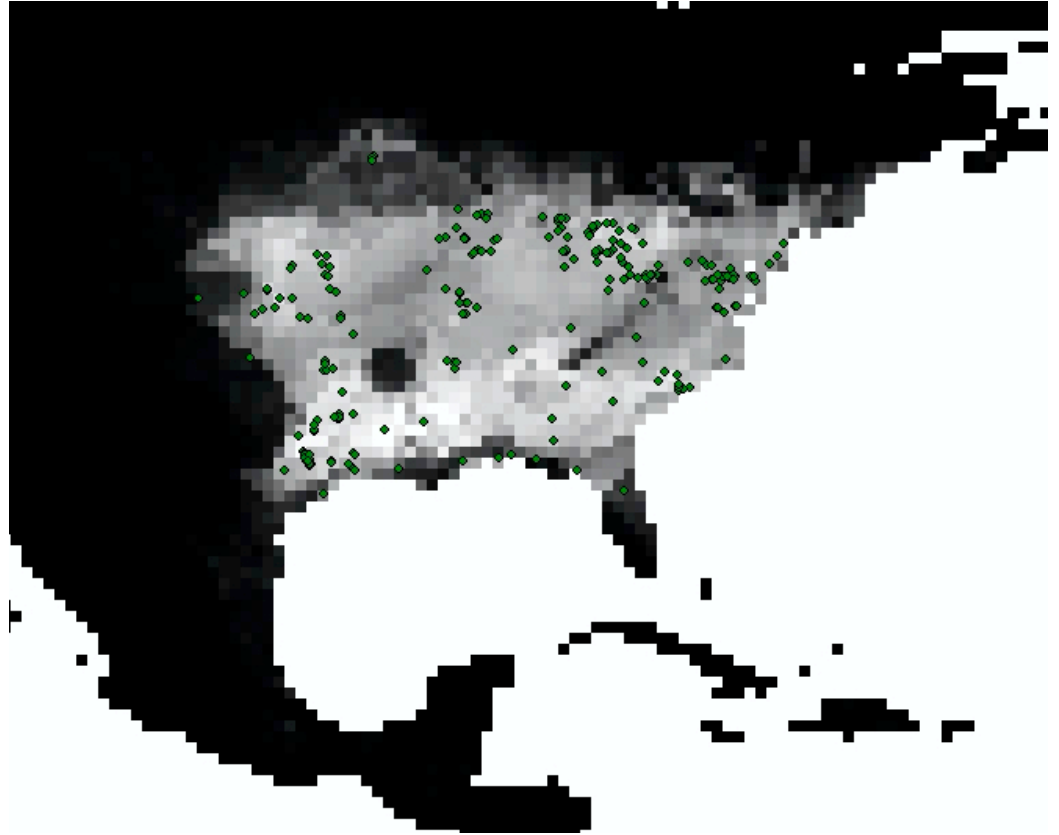
MaxEnt-Analysis of Variable Contributions

- 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
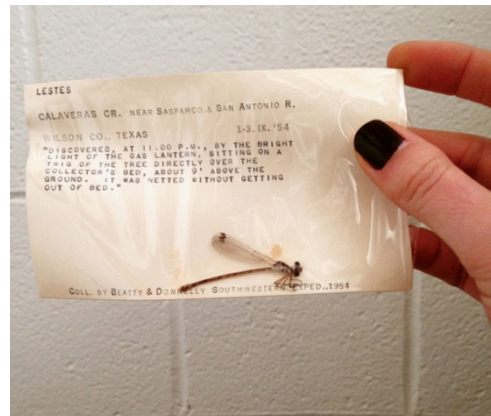
Future Projects

- Digitizing field notes to free collecting event data

- Resampling localities for spatiotemporal study of Pennsylvania Odonata

- Sampling bias in the collection
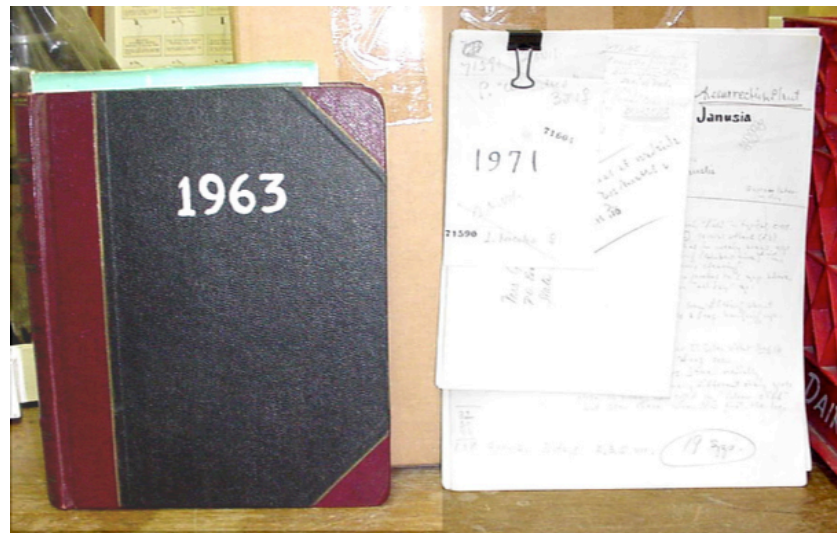
Opportunities

- 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

Carlie Harding

Marcus Thaw

Michael Przybys

István Mikó

Sam Kurth

Hal White

Matt Yoder