Conservation and Management of Rare Wetland Butterflies:
Strategies for Monitoring, Modeling and Wetland Enhancement in the Mid-Atlantic Region

Project Director
Jennifer Selfridge, Invertebrate Ecologist
Maryland Department of Natural Resources, Natural Heritage Program
P.O. Box 68, 909 Wye Mills Road; Wye Mills, MD 21679
Phone: 410-827-8612 x102, Fax: 410-827-5186; jennifer.selfridge@maryland.gov

Principle Investigators
Betsy Leppo, Invertebrate Zoologist
Pennsylvania Natural Heritage Program, Western Pennsylvania Conservancy
bleppo@paconserve.org

Susan Olcott, Regional Wildlife Diversity Biologist
West Virginia Division of Natural Resources, Wildlife Diversity Program
susan.p.olcott@wv.gov

Robert Somes, Wildlife Biologist
New Jersey Department of Environmental Protection
Division of Fish and Wildlife, Endangered and Nongame Species Program
Robert.Somes@dep.nj.gov

Christopher Tracey, Conservation Planning Manager
Pennsylvania Natural Heritage Program, Western Pennsylvania Conservancy
ctracey@paconserve.org

Peter Woods, Inventory Ecologist
Pennsylvania Natural Heritage Program, Western Pennsylvania Conservancy
pwoods@paconserve.org

Suggested citation:
Selfridge, J., B. Leppo, S. Olcott, R. Somes, C. Tracey, and P. Woods. 2018. Conservation and Management of Rare Wetland Butterflies: Strategies for Monitoring, Modeling and Wetland Enhancement in the Mid-Atlantic Region. Final report for Regional Conservation Needs Grant 2015-02 to the Wildlife Management Institute.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

This project was funded in part by a Regional Conservation Needs Grant. Thanks to everyone who provided financial and administrative support for this work. The species information utilized in the project came from many sources as well as our own field surveys. We wish to acknowledge all those who carried out butterfly surveys over their years, and the landowners who granted permission to visit their properties. We are grateful for the contributions to the final report made by Laura Zielinski, a dedicated student at Wilkes University inspired by all things with six or more legs. We tip our nets to the many photographers who shared their photos of these charismatic wetland butterflies for the life history guide. *Maryland* would like to acknowledge help from multiple staff at MD DNR, most notably Siobhan Percey, Dana Limpert, Paula Becker, Scott Smith, Gwen Brewer, Jim McCann and Ed Thompson; Jim Brighton and Bill Hubick of the Maryland Biodiversity Project; Heather Cunningham of Chesapeake College; Richard Smith and numerous MD DNR volunteers who provided many hours of assistance with field surveys and habitat enhancement projects. *Pennsylvania* would like to thank Margaret Rohde, Kristy Morley, and their colleagues at Wissahickon Valley Watershed Association for their participation in this grant, and we recognize the fine work they are doing at Crossways Preserve. We thank other partners who worked with us on this project: Edith Dondero and Kevin Fryberger of Brandywine Conservancy; staff at Bureau of State Parks (Gifford Pinchot State Park and the Resource Management and Field Services Section), staff and students at the Diakon Wilderness Greenhouse and Native Plant Nursery, volunteers of the Friends of Pinchot State Park, Blaine Rothauser of GZA GeoEnvironmental; Dan Barringer, Gary Gimbert, Tom Kershner, Sean Quinn, Joshua Saltmer, Jarrod Shull, and Preston Wilson of the Natural Lands Trust; Mari-Beth DeLucia, Su Fanok, George Gress and Brad Maurer of The Nature Conservancy. Additionally, we would like to thank Molly Moore and David Yeany of the Pennsylvania Natural Heritage Program for assistance with the species distribution models. *West Virginia* gratefully acknowledges the staff and volunteers who helped with surveys and ArcGIS work, especially Jessica Perkins, Paula Waggy, and Pam Byrne; and to Cathy Johnson at the Monongahela National Forest for allowing survey work on the forest.

ABSTRACT

Fourteen species of wetland-inhabiting butterfly Species of Greatest Conservation Need (SGCN) status were surveyed in 2016 and 2017 at multiple sites across four states – Maryland, New Jersey, Pennsylvania and West Virginia. Survey data was used to evaluate the status of each species in all states where they occurred as well as refine the distribution data for each species across the region. All data points were mapped in ArcGIS and used to model species distribution in terms of both habitat and climate. Data collected prior to 2016 was also evaluated and added to the models, as was data from Delaware that was provided by DE NHP. The results are presented for each species and several examples are explored in greater depth. Best Management Practices (BMPs) were developed for both modeling procedures. A final goal of the project was to initiate habitat enhancement projects in a small number of survey areas in Maryland and Pennsylvania. The results of these projects are discussed as well as the BMPs that were developed for site selection and management.

INTRODUCTION

The status and distribution of many wetland butterfly species is uncertain in several Mid-Atlantic States. Many are considered Species of Greatest Conservation Need (SGCN) in many or all of the states in which they occur. For some species declines are well documented and may span a decade or more, while others are of concern based on low encounter rates and negative data during recent survey efforts. This may be in part due to threats impacting groundwater wetlands, including outright destruction, habitat degradation and the succession of open wetland habitats to forest or dense shrubland. Climate change and habitat fragmentation may further impact these species and leave them vulnerability to local extirpations.

In response to the concerns for wetland butterfly species across the region, we attempt to improve our understanding of 14 species of wetland butterfly SGCN by coordinating our efforts at a regional level. The primary objective of this effort is to enhance and expand populations of wetland butterfly SGCN through developing a greater understanding of the distribution and habitat requirements for these species, and by implementing habitat enhancement projects where needed. Our goals were (1) to update distribution data for 14 butterfly SGCN in the region, (2) model species distribution and climate conditions for each species; (3) identify and prioritize wetlands that support one or more of these 14 species, (4) implement wetland enhancement and improvement projects, and (5) develop Best Management Practices (BMPs) for both species distribution and climate modeling and for wetland enhancement projects. In the short-term our results should guide targeted survey work for these species as well as prioritize wetlands for enhancement projects. In the long-term our results may serve to improve habitats for these species, offering the potential to increase populations of butterfly SGCN and promote connectivity between populations through increased habitat availability.

Note that Objective (2) was not in our original proposal as a stand-alone objective, but is included now because it became a major component of this project, and the final results were well beyond our original expectations. We modeled species distribution and climate for all 14 species of wetland butterflies. A Species Distribution Model (SDM) was used to predict suitable habitat for each species. In recent years, SDMs have proven to be exceptional resources used to survey and manage for rare species and have been used many times to model butterflies for various functions. For example, Willis et al. (2009) used MaxEnt to predict suitable reintroduction sites for two species of butterflies in the United Kingdom in areas that were initially beyond the range of both species.

Climate Envelope Models (CEMs) were also used to delineate areas of climate suitability for each of the 14 target species by correlating georeferenced species occurrences with observed climate conditions (Watling et al. 2013). In contrast to SDMs, CEMs only include climate variables, so the models are only describing areas where climate is suitable for the species being modeled. In a strict sense, CEMs define minimum and maximum values of climate boundaries around species occurrences, thereby delimiting a 'climate envelope' within which species occur. Additionally, these climate models can be projected to future climates to understand the effects of climate change on the distribution of a species.

We had originally we proposed to use the results of species distribution and climate models to guide our survey efforts and help prioritize wetland sites for restoration during the grant period. However, the time investment in all other aspects of the project that were used to generate the model data did not allow us to obtain the final results until after the 2016-2017 field seasons had concluded. Intensive survey work in all participating states generated a wealth of data, including new element occurrence (EO) records, updated occurrence records, and records in which we failed to find species that had once been present, in some cases because the wetland had been degraded and was no longer suitable. It took a great deal of time to compile the data from multiple sources, enter the data, map each occurrence according to the specifications of each model and ensure data quality prior to running the models. As a result, even the initial runs for both species distribution and climate models were not completed until May 2017 and November 2017 respectively. However, the accuracy of the models was significantly improved by the addition of new and accurate data acquired during the grant period, as well as feedback from project partners, and the output of the models will aid in future surveys and assessments of these species and their habitats.

Physiographic regions targeted for wetland butterfly surveys included the Appalachian Plateau, Allegheny Mountains, Ridge and Valley, Highlands and the Piedmont (although data from Coastal Plain areas were also incorporated). All species considered are listed in Table 1, and all are considered SGCN in one or more participating states. Note that not every species occurs in every state nor is every species listed as SGCN in every state. NatureServe (2017) range maps for each species showing the rank in each state are included in Appendix A.

Table 1. Species of wetland butterfly SGCN targeted for survey and modeling and their global and state conservation status ranks. Global and state ranks follow a simple numerical scale (1-5) with the lower numbers reflecting increased rarity and risk. For *Poanes viator viator*, a taxon (T) rank follows the global rank to indicate the conservation status of the subspecies. If a species is absent from a given state, that state will not appear in State Ranks column. A rank of SNR in the State Ranks column indicates that the species does or may occur in the state but has not been formally ranked. The State SGCN column indicates each state in which the species is considered SGCN.

Species	**Global Rank**	**State Ranks**	**State SGCN**
Anatrytone logan (Delaware Skipper)	G5	S3-MD, WV; S4-NJ, PA	MD
Boloria selene (Silver-bordered Fritillary)	G5	S2-PA; S3-MD, WV	MD, NJ, PA, WV
Carterocephalus palaemon (Arctic Skipper)	G5	S3-PA	PA
Chlosyne harrisii (Harris' Checkerspot)	G5	S2-MD, WV; S3-NJ, PA	MD, NJ, PA, WV
Euphydryas phaeton (Baltimore Checkerspot)	G4	S2-MD; S3-PA, WV; S4-NJ	MD, NJ, PA, WV
Euphyes bimacula (Two-spotted Skipper)	G4	S1-MD, WV; S2-PA; S3-NJ	MD, NJ, PA, WV
Euphyes conspicua (Black Dash)	G4	S1-WV; S3-PA; S4-MD, NJ	MD, PA, WV
Euphyes dion (Dion Skipper)	G4	S3-MD, PA; S4-NJ	MD, PA
Lethe eurydice (Eyed Brown)	G5	S2-PA; SNR-NJ	PA
Lycaena epixanthe (Bog Copper)	G4G5	S1-MD, WV; S2-PA; S4-NJ	MD, PA, WV
Lycaena hyllus (Bronze Copper)	G5	S1-NJ; S2-WV; S3-PA; S4-MD	MD, NJ, PA, WV
Poanes massasoit (Mulberry Wing)	G4	S2-PA; S4-MD, NJ	MD, PA
Poanes viator viator (Broad-winged Skipper)	G5T4	S2-PA; SNR-MD, NJ	PA
Polites mystic (Long Dash)	G5	S3-MD, NJ, PA; S4-WV	MD, PA

With the exception of federally listed species, butterflies have not yet been evaluated as potential Regional SGCN. Therefore none of the focal species for this project are currently listed as Regional SGCN in the Northeast Region State Wildlife Action Plan Database (2017).

METHODS

SITE SELECTION AND DATA COLLECTION

Given the overall number of species present in any given state, Natural Heritage Programs (NHPs) generally track only those species that are designated as rare. Because not all 14 species are tracked in all states, data was necessarily limited for species that were considered watchlist or common in a given state (no historic database records). In 2016-2017 the four participating states collected data on all 14 species if they were encountered regardless of the species status in that state, but historical records likely under-represent the status of many species in some states due to a lack of databased observations. Some states were able to add data from additional sources for tracked and non-tracked species as outlined in the following sections.

Project survey sites were chosen for a variety of reasons, and this was generally consistent for all four states (although there were some differences). In general, sites were chosen if they were known to support habitat for any of the target species, if the target species were known to have historically occurred there, if they were accessible to staff and volunteers, or if they were in an area of high ecological functionality. All states had standardized wetland butterfly and habitat assessment forms available for use (Appendix B). All of the records were databased and mapped; shapefiles were shared with the Pennsylvania Natural Heritage Program (PNHP) for inclusion in the species distribution and climate models.

While all states shared the same overall objectives (i.e. data collection, prioritizing sites for survey work), representatives from all states exercised freedom in reaching those objectives. Additionally, because the project generated a wealth of data that will continue to be evaluated and utilized in months to come, priorities on data organization, data evaluation and species assessments also vary from state to state. To account for this the details of survey work are organized by state and the information presented differs between states.

Maryland

Maryland NHP focused survey work at 50 wetlands during the 2016-2017 survey period although data from an additional 96 wetlands and sites adjacent to those wetlands (roadsides and fields) was also included; this additional data came from records in the NHP database, records in the Maryland Biodiversity Project (MBP) database and records from local Lepidopterists and survey groups. NHP staff and volunteers surveyed the sites between 11 May and 10 August in 2016, and between 10 May and 6 October in 2017 with a total of approximately 105 survey days over two field seasons (about 40 survey days in 2016 and about 65 survey days in 2017). Dates are approximate as some volunteers did not report their hours. Most sites were surveyed on multiple occasions. While surveys were conducted all over the state, the 2016 field season focused on the Piedmont Region, which extends from the Catoctin Mountains east to the Coastal Plain. The 2017 field season focused on the highest elevations of the state in the Appalachian Plateau Region.

Maryland submitted 292 records for 12 of the 14 species present (Arctic Skipper and Eyed Brown do not occur in the state). For tracked species these came largely from existing records in the NHP database and from new data collected in 2016-2017 as a result of this project. For non-tracked species the records consisted of those compiled from the MBP database, from notes/databases of NHP staff and volunteers throughout the state, and from new data collected during the project duration. Of the 12 species present in Maryland, four are tracked elements (Baltimore Checkerspot, Bog Copper, Harris' Checkerspot and Two-spotted Skipper) and all are currently listed as SGCN with the exception of Broad-winged Skipper.

New Jersey
New Jersey NHP focused survey work at 26 wetlands during the grant period. Each of these was surveyed three times in 2016 and twice in 2017. Very few of the target SGCNs were located during the survey periods despite the fact that many of the species had been documented previously at these locations. Most wetlands targeted for surveys had been previously documented as Baltimore Checkerspot sites.

Pennsylvania
Pennsylvania NHP staff conducted field surveys at 47 wetland sites across the state for a total of 54 days during the grant period. Surveys occurred in both 2016 and 2017 primarily during the months of June, July and August. A list of sites visited and the criteria that was used to select each site is provided in Appendix C. Some of the sites were selected using the methods outlined above, others were selected based on the following additional criteria:

(1) Wetland butterfly site assessments: Most of these sites were selected based on recommendations by partner conservancies or landowners who are planning or implementing management in habitats that support or potentially support rare wetland butterflies. Surveys focused on assessing habitat conditions, recording the presence/absence of wetland butterfly SGCN targets, and noting the presence of host plants, nectar plants, and other habitat quality measures. Most of the conservancy partner sites were located in the Piedmont Province in the south-east corner of the state. A few additional sites scattered around the state were selected based on the criteria outlined previously in this section.

(2) High elevation peatland monitoring sites: High quality wetland sites at high elevations in PA were selected for a long term monitoring program to detect changes in the plant and animal communities related to habitat succession and climate change. Sites were evaluated using standardized methods to characterize the vegetation of occupied habitats. Staff also conducted surveys along standardized transects to evaluate the abundance of target groups of insects (including butterflies), host plants and nectar sources. Sites were scattered across the state in difference physiographic provinces, though most were located in the northern half of the state.

(3) Pennsylvania Bureau of Forestry habitat management monitoring: The Bureau of Forestry recommended a focused survey effort on sites actively managed by the Bureau to determine how their habitat management and restoration activities potentially impact pollinator species. These sites were evaluated using standardized transects to evaluate the abundance of butterflies and other pollinators, host plants, and nectar sources. These study sites were located Tiadaghton and Rothrock State Forests.

Pennsylvania submitted 420 records with representative from all 14 species present in the state. All of the wetland butterflies included in this project are tracked SGCN species in Pennsylvania, except for the Delaware Skipper. The 13 SGCN species have been tracked for different time periods. The Baltimore Checkerspot, Eyed Brown, and Long Dash have been tracked since 2002. The Long Dash was removed from the tracking list in 2008, but reinstated after the State Wildlife Action Plan update in 2015. The other ten species have been tracked for longer periods of time; the start date of their tracking is not recorded, but likely dates back to the late 1980s.
In addition to new survey data accumulated during the grant period, additional data was obtained from existing records in the PNHP database, published and gray literature, personal correspondence, data sets from PNHP programmatic surveys, records from individual collections and contributors, and data from museums including the Academy of Natural Science of Drexel University in Philadelphia, the Carnegie Museum of Natural History in Pittsburgh, and the Natural History Museum at the Tom Ridge Environmental Center. Other resources included a dataset from the Butterflies and Moths of North America website, the Atlas of Pennsylvania Butterflies, and other online datasets (e.g. iNaturalist).

West Virginia

West Virginia has over 600 documented wetlands in its Natural Heritage database. Most are at high elevations (above 2000 feet) and are primarily beaver dominated headwater stream wetlands. For survey purposes, that number was narrowed down to 60 sites that met the criteria outlined above, with the assumption that WV Department of Natural Resources (DNR) staff and volunteers would survey 30 sites each year. Visits occurred between 25 May and 9 August in 2016, and between 18 June and 30 July in 2017. A total of 42 wetlands were surveyed during the grant period, 32 of which had target SGCN present. Of the 14 wetland species targeted in this project, West Virginia hosts nine. Surveys conducted during the grant period documented six of the nine species. Additionally, because of the ongoing West Virginia Butterfly Atlas, substantial survey work had already been done since 2012; results from these surveys were added to the database.

SPECIES DISTRIBUTION MODELING

Acquisition and Preparation of Species Presence Data

Data for all 14 butterfly species was prepared prior to modeling. The known locations of each species and the extent of the study area were the basis for development of the presence and background points, respectively. These points, collectively called training data, were the input for the model. We used the most current species occurrence data from the five NHPs within the project area (the four states involved in the project and data submitted by DE NHP). Initial data was received in April 2017; additional records were added after the first draft of the SDMs and after the 2017 field season.

Data from NHPs are organized into Element Occurrences (EOs), with each unique EO approximating a species population. EOs may be single locations or comprised of multiple locations (called source features or, herein as ‘polygons’) in close proximity. To ensure appropriate tracking among data sets from different participating partners we assigned a project wide unique code to each EO consisting of the state abbreviation concatenated to EOID (e.g. “PA-123”). Any data whose source was not a state heritage program was assigned a unique EOID number for this project. Data received as either point or line representations were converted to polygon by buffering the feature using the locational uncertainty value. Data received as polygons were left as such.

The use of high quality occurrence data in the models is of utmost importance and, therefore, polygons used to train a model should represent only the extent of habitat appropriate for the focal species. Thus, we needed to review, cull, and edit the input data, using best professional judgment, to ensure that the polygons truly represented the habitats in which the species would occur. Since this project focused on SGCN, which have limited known locations, it was necessary, when screening the submitted polygons, to balance the preference for high-quality occurrence polygons against the demand for a sufficient distribution of known locations. If too many polygons were discarded (due to quality, age, or spatial accuracy issues), the model may not have predicted the presence of suitable habitat in areas where the species once occurred. Thus, the fewer polygons available in an area of the species’ range, the higher the tolerance for older observations and for lower accuracy of those observations. We developed standard guidelines for how to approach the edits, to further foster consistency among various editors (Appendix D). In general, the guidelines attempt to balance the desire to have high-quality occurrence polygons with the desire to have sufficient distributions of known locations. Each polygon was inspected using aerial imagery and attribute comments to discern if it should be included, included with editing, or excluded. In total, over 3,200 polygons (Table 2) were reviewed and, if needed, edited to better represent suitable habitat. For all species 1,689 total polygons were deemed suitable for inclusion in the model.

Table 2. The number of Element Occurrence (EO) records submitted by each state heritage program and reviewed by project partners. Also enumerated are the number of records accepted for use in the model (based on criteria such as age of record, removal of duplicate records, and accuracy of the representation). Some records enumerated here may represent more than one polygon (e.g. a multi-part EO).

Species	Source	Total Accepted 2016	Total Accepted 2017	Overall Total 2017
Anatrytone logan (Delaware Skipper)	MD	11	22	43
	PA	6	12	
	WV	5	9	
Boloria selene (Silver-bordered Fritillary)	MD	7	11	57
	NJ	13	13	
	PA	16	19	
	WV	24	28	
Carterocephalus palaemon (Arctic Skipper)	PA	7	7	7
Chlosyne harrisii (Harris' Checkerspot)	MD	15	19	55
	NJ	3	3	
	PA	22	24	
	WV	2	9	
Euphydryas phaeton (Baltimore Checkerspot)	MD	39	62	134
	PA	55	63	
	WV	8	9	
Euphyes bimacula (Two-spotted Skipper)	MD	3	4	27
	NJ	3	3	
	PA	7	14	
	WV	-	6	
Euphyes conspicua (Black Dash)	MD	16	31	76
	PA	36	42	
	WV	-	3	
Euphyes dion (Dion Skipper)	MD	1	8	17
	PA	10	9	
Lethe eurydice (Eyed Brown)	NJ	1	1	9
	PA	7	8	
Lycaena epixanthe (Bog Copper)	MD	5	15	51
	PA	44	36	
	WV	1	-	
Lycaena hyllus (Bronze Copper)	MD	11	17	68
	NJ	10	9	
	PA	34	38	
	WV	-	4	
Poanes massasoit (Mulberry Wing)	MD	17	18	28
	PA	10	10	
Poanes viator viator (Broad-winged Skipper)	PA	8	8	8
Polites mystic (Long Dash)	MD	2	9	51
	NJ	1	1	

	PA	20	29	
	WV	11	12	

EO Rank

Each heritage program assigns a qualitative rank to each EO, which reflects the viability (i.e., the probability of persistence 20 to 100 years in the future), of the species at that location. Ranks range from A (excellent estimated viability) to D (poor estimated viability), but EOs lacking viability information may be attributed (E = extant, F = failed to find, H = historical, and X = extirpated) otherwise (NatureServe 2002). We did not automatically exclude F, H, or X ranked EOs. Consideration was given to whether the current habitat is similar to what it was at the time of the last known observation. In most cases, F, H, and X ranked EOs were removed, but exceptions were allowed, particularly for a precise EO where habitat appears unchanged. For data from non-heritage sources, a rank of 'E' was presumed extant until other evidence suggested otherwise (for example, the date of the observation may reclassify it to an H).

Likely Suitable Habitat

Polygon review methods allowed for the addition of suitable habitat in the surrounding inferred extent of EO polygons in some cases. This was generally allowed in the following cases:

- For very small EO polygons (< 0.1 acre), usually based on GPS point locations, where clearly suitable habitat is larger than the polygon.
- For apparent mapping errors or generalizations (based on EO description data). For example: a small polygon located in uplands with nearby small wetlands, and with the observation notes "caterpillars found in three wetlands".
- For species in which adults are often found beyond appropriate larval habitat. Habitat that was clearly unsuitable habitat was removed from EO polygons, including non-natural areas (developed lands, roads, etc.) as well as unsuitable natural areas (e.g. uplands for an wetland species). If the entire record was considered unsuitable habitat (e.g. developed into urban area or a forest converted to pasture), it was not used in the model.

Representation Accuracy

Representation accuracy (RA) is a categorical estimation of the amount of locational uncertainty a particular EO might contain (NatureServe 2002). RA was used to weight input data for use in the modeling algorithm; the higher the RA (i.e. lower locational uncertainty) the higher the weight in the model. Each polygon and each EO was reviewed to ensure an RA was assigned. Since some EO sources were significantly modified for the project as outlined above, RA was sometimes modified (or assigned to new polygons), using the following guidelines:

- If a large low-precision EO polygon was reduced to smaller suitable habitat polygon(s), RA could be increased (from Low to Medium for example).
- If a small high-precision EO polygon was extrapolated to a larger suitable habitat polygon, RA might be decreased (from Very High to High or Medium for example).
- If an EO was improved by interpreting the observer's notes, the following guidelines were considered:
    - RA was assigned as High if there was high confidence that the edited polygon captured the actual location or feature intended by the source observer, and the original RA was Medium or higher.
    - RA was assigned as Medium if there were suitable features matching the source description but the exact feature/location could not be determined, and the original RA was Medium or higher.
    - RA was assigned as Low if there were suitable features, but no indication whether any of them were the one noted by the source observer, and the original RA was Medium or lower.

Overlapping Polygons
If multiple polygons for a single species with the same RA overlapped, they were dissolved into one polygon and further edited if necessary. If the overlapping polygons had different RA, overlap areas were split and assigned the appropriate RA, with higher RA values being preferentially maintained.

Development of Environmental Data Layers
Each species was modeled based on a suite of environmental predictor variables that could influence its distribution (Table 3). Environmental data for this project was initially created for two previous species distribution modeling projects across the eastern United States (Chazal et al. 2017a, 2017b). We produced and used a set of 88 environmental datasets (Appendix E). We used the Albers Conical Equal Area coordinate system, and a pixel size of 30m, consistent and perfectly aligned with many source datasets available for the continental United States, such as the National Land Cover Database (NLCD). The original source data and resulting derived environmental variables extended at least 5 km beyond the study area boundaries. This was to avoid introducing any boundary effects or artifacts in rasters developed using flowpath distance, or Euclidean distance, to polygons, or any kind of neighborhood analysis. Rasters were saved in the geo-referenced Tag Image File Format (GeoTIFF) for use in the SDM. A complete list of environmental variables is provided in Appendix E.

Table 3. Categories of environmental variables developed, primary data sources, and lead partners responsible for variable creation. Data initially developed for Chazal et al. (2017a, 2017b).

Category	# of Derived Environmental Variables	Source
Climate – Precipitation	11	PRISM[1]
Climate – Temperature	12	PRISM[1]
Elevation and Derivatives	15	NED[2]
Geology	13	TNC[3]
Hydrography	10	NHD[4], NHD+[5] , NLCD[6] , NED[2]
Land cover – NLCD	24	NLCD[6]
Land cover - NWI	3	NWI[7]

1 = PRISM climate data (http://www.prism.oregonstate.edu/); 2 = National Elevation Dataset (https://nationalmap.gov/elevation.html); 3= Eastern Geology data (combined bedrock/soils) developed by The Nature Conservancy, (A. Olivero, pers. comm.); 4 = National Hydrography Dataset (https://nhd.usgs.gov/); 5 = NHDPlus (http://www.horizon-systems.com/nhdplus/); 6 = National Land Cover Database (https://www.mrlc.gov/); 7 = National Wetlands Inventory (https://www.fws.gov/wetlands/).

Pre-processing of Background Points
Once the species data and environmental variables were finalized, additional preprocessing was needed to prepare the data for use in model development. A set of background sample points were created and then attributed with the values of each environmental variable at those locations.

A set of background points had previously been generated across the entire study regions for the Region 5 and SALCC projects using the Generalized Random Tesselation Stratified (GRTS) sampling scheme (Stevens and Olsen 2004, Kincaid and Olsen 2007) to generate spatially-balanced random points. Using the entire study region, from Maine to Northern Florida (1,070,500 km^2), as the bounding polygon, we generated a total of 53,000 spatially-balanced points (Chazal et al. 2017a, 2017b). We created a subset of this master background points file by clipping it to the smaller study area. The Mid-Atlantic Butterfly study area is 243,523 km^2 and has 11,473 background points. This 'master' set of attributed background points was used as the baseline for all models rather than creating a new set of points for each model.

Environmental data (such as elevation, distance to pond, etc.) was attributed to each background point. Values from each environmental variable raster (without interpolation) were extracted and stored in a corresponding field within the points file.

SDM Model Development and Execution

Because the network of NHPs oversees a robust data set of known locations for the species being modeled, we used an inductive approach (inferring general conditions based on specific examples). An inductive model uses known locations (specific) as training data to select areas with similar characteristics (general). In this case the characteristics assessed are the environmental variables that encompass the entire study area. A classification tree approach, Random Forest, was chosen as it can handle rare species (with few known locations), large predictor datasets (the environmental variables), and is capable of multiple solutions across the landscape. A resolution of $30m^2$ was selected to balance meaningful biological resolution with processing power needs over such large areas.

Random Forest and R Overview

Our primary analytic and modeling tool was the R statistical software (R Core Team 2015) with a suite of packages included to accomplish certain tasks. The primary packages include 'randomForest' (Liaw and Wiener 2002) for building the actual species-environment inductive model, 'RSQLite' (Wickham et al. 2014) for reading and writing information to our information systems database, 'raster' (Hijmans et al. 2014) for working with GIS-based raster data, and 'knitr' (Xie 2016) for writing out metadata in a readable format.

Random Forest is one of many modeling methods within the broad new category of Machine Learning classifiers and is used in a wide range of applications and fields (e.g. Díaz-Uriarte and Alvarez de Andrés 2006, Jones and Linder 2015). When pitted against other classification methods, Random Forest is regularly highly rated (e.g. Prasad et al. 2006, Cutler et al. 2007, Fernández-Delgado et al. 2014). The Random Forest algorithm can also be categorized as an ensemble model in that it depends upon many separate statistical models (an ensemble of models) to provide the final statistical inference about the relevant differences in environmental conditions between absence or pseudo-absence (background) locations and presence locations. Each model within the ensemble is a separate classification and regression tree (Breiman et al. 1984, Breiman 2001, Cutler et al. 2007), with each tree in the ensemble differing slightly due to two primary factors:

- a random subset of the total input points is drawn before building the tree, and those points are the only ones used for that individual tree
- at each split in the tree, a new random subset of environmental variables is drawn, from which the best variable is used to divide the points at that split.

Typically, a Random Forest consists of hundreds to thousands of trees. The ensemble of trees becomes the final "model," and while it is not possible to view each individual tree that makes up the model, the predictive power of the entire ensemble turns out to be quite robust. One the most important features of Random Forest is its ability to perform well even given a very large number of predictor variables, many of which may be highly correlated. This makes it able to handle situations in which more traditional modeling techniques (e.g., logistic regression) would be entirely inappropriate.

Modeling Procedure

To build the Random Forest model in R, we used the 'randomForest' package (Liaw and Wiener 2002). The number of trees in a Random Forest model is specified by the "ntrees" parameter. We specified 2000 trees to be included in our final, full ensemble models. The number of environmental variables drawn at each node in each tree depends on the "mtry" parameter. We used the "tuneRF" function twice in succession (with 300 trees in the forest) to determine the optimal value of mtry. The tuning function builds a forest with an initial mtry, estimates the error of the model (by testing records excluded from individual trees [out-of-bag records]), then repeats the process with higher and lower mtry values until it minimizes the model error. By default, for each tree in the Random Forest, a simple random sample is drawn from the entire combined set of presence and absence (i.e., background) points. This is problematic in our specific modeling situation for several reasons:

- Each observation in our dataset is polygon based, with points placed randomly inside the polygon(s) for model building. Thus, points falling within one EO are very closely related and need to be treated as such during the random point draw for building a tree.
- Different EOs have different RA, indicating different levels of confidences in the spatial accuracy of delineated polygons.
- When modeling rare species, because the number of background points is vastly larger than the number of presence points, it is likely that, for some trees, a random sample could draw only background points. Such trees would be entirely uninformative, effectively reducing the number of trees in the ensemble from the specified 2000 to some unknown lower number.

With these issues in mind, several customizations were necessary to assure the models were accurately informed with the input data. These customizations are highlighted and briefly explained in the appropriate modeling steps below (full details are available in Chazal et al. 2017a, 2017b):

1. <u>Defining the Input Data</u> – The first script ("0_pathsAndSettings.R") sets up for the modeling run by defining the locations for input data, the lookup database, and folders for writing results in the subsequent scripts. After setting these definitions here, the person doing the modeling has very few modifications to make in the subsequent scripts.
    a. The second script ("1_pointsInPolys_cleanBkgPts.R") places random points within each input polygon. The number of points assigned to each polygon varies by the size of each polygon. More specifically, we use a logistic function to increase the number of points as the size of the polygon grows but then to asymptote so the total number of points does not go over 400 (Howard and Schlesinger 2013). This script also checks the set of background points to see if any of these points fall within or nearby (30m) one of the input (presence) polygons. Any points that are coincident are removed from the background set.
    b. The next script ("2_attributePoints.R") attributes all the environmental data (such as elevation, distance to pond, etc.) to each presence point generated in the previous script.

2. <u>Building the Model</u> – The Random Forest model of the relationship between known locations and environmental conditions is developed and validated in the next script ("3_createModel.R"). There are a few key customizations here as well.
    a. First, we removed all "distance-to" grids (such as distance to ocean or distance to calcareous rock) where the shortest distance for all presence points was greater than 10 km. These layers were developed to refine models when known locations were near a boundary (e.g. close to the ocean-land boundary or close to the calcareous rock-non calcareous rock boundary) and they should not influence the model when the boundaries are far away from known locations.
    b. Second, we created an initial Random Forest model to evaluate the relative importance of each environmental variable and then removed the least important variables. We explored many of the published approaches for dropping less important variables (e.g. Genuer et al. 2010, Hapfelmeier and Ulm 2013, Diaz-Uriarte 2014, Gregorutti et al. 2016) but these turned out to be too computational and time-intensive for this project. In the end we simply chose to drop the 25% least important variables.
    c. Third, after dropping the least important variables, the model was validated using a "leave-one-out" jackknife routine (Fielding 2002, Howard and Schlesinger 2013). To do this, a random forest

model of 1000 trees was built with representation from all but one EO, and then tested to determine if that model could predict the location of the excluded EO as suitable habitat. This was repeated for every EO, and summary statistics were generated including the overall accuracy, specificity, sensitivity, true skill statistic (TSS), Kappa, and the area under the receiver operating characteristic curve (AUC). These statistics were saved in an .Rdata file, and used for reporting in the metadata. We also customized the sampling scheme Random Forest used to draw the presence and background points to build each tree. The sampling scheme is specified by the "sampsize" parameter. First, we stratified by EO to ensure that each EO would contribute at least one sample point drawn for every tree. This avoided the possibility of very large (often spatially questionable) polygons carrying undue weight in the model, or of very small (but usually highly accurate) polygons being under sampled. Second, we define a sampling scheme for subsetting points within each EO based on its RA. We determined the number of points drawn from each EO based on the RA value (5 points for EOs with "very high" accuracy, 4 for "high", 3 for "medium", and 2 for "low"). This approach effectively weights the importance of each input EO based on its spatial reliability. Smaller polygons with fewer points might have the same points drawn repeatedly, while larger polygons, especially those with lower RA values, would have points drawn from throughout the polygon, but less frequently or possibly never repeated. Finally, the number of background points selected for each tree was equivalent to the number of presence points drawn for that tree. The background points, however, were drawn from a very large pool of available background points. This eliminated the possibility of including uninformative trees in the ensemble. After validation was completed, a final full model consisting of 2000 trees was run. Importance measures were calculated for the retained predictor variables, and data to produce partial plots were generated for the nine most important variables. Model run information was stored in an .RData file, with some information on the model run also written to an SQLite database.

3. <u>Model Prediction and Threshold Calculation</u>
    a. <u>Model Prediction</u> - After the model building phase was complete, we used script "4_predictModelToStudyArea.R" to generate predictions of the relative suitability of habitat throughout the study area. This script retrieves the model information stored in the .Rdata file, collects the values of the environmental predictor variables at each raster cell location, and runs these values through the saved random forest model to calculate the probability of suitable habitat for the target species. Each set of predictor variable values, for each location, is fed through every single classification tree in the forest. Some trees will classify a cell as suitable habitat, while others will classify it non-suitable. The final tally of 'votes' from all trees determines the predicted probability that a cell represents suitable habitat. The final output is a raster surface with cell values theoretically ranging from 0 to 1, where "1" equals 100% probability that the location represents suitable habitat.
    b. Threshold calculation - The continuous probability surface described above may not appropriately represent locations deemed suitable habitat for the target species. Most users will want a binary output in which each raster cell is classified as suitable or not suitable for the target species. A continuous probability surface can be reclassified into a binary map by setting any values above a specified threshold to "1" (habitat likely suitable) and any values below the threshold to "0" (habitat likely not suitable). We used the "4b_thresholdModel.R" script to create a suite of different thresholds customized to each individual species model. These thresholds were reported in the metadata that accompany each product, giving the user the ability to generate their own classified output if needed. The thresholds calculated are defined in Table 4. For additional discussion of most of these thresholds see Cantor et al. (1999), Liu et

al. (2005), and Pearson et al. (2007). Appendix F conceptualizes the different minimum training presence (MTP) thresholds.

4. Metadata – To summarize model inputs, outputs, and performance, we used the script "5_createMetadata.R" to produce metadata PDF for each model. This script retrieves data from the model, then automatically formats and writes out model information including: information about the data inputs, validation, importance measures for each environmental predictor variable, response curves for individual environmental gradients for the top nine most important predictor variables, the different thresholds calculated for the model, a coarse-resolution map of the model predictions, data sources, and any customized comments specific to the particular species or model. This PDF file should always accompany the model outputs and be available to any users using the model outputs.

Table 4. Thresholds used to convert a continuous probability surface to a binary output representing suitable versus non-suitable habitat.

Threshold	Description
Equal sensitivity and specificity	The probability value at which the absolute value of the difference between sensitivity and specificity is minimized.
Maximum sum of sensitivity and specificity (MSS)	The probability value at which the sum of sensitivity and specificity is maximized.
Minimum training presence (MTP)	The highest probability value at which 100% of the input presence points remain classified as suitable habitat.
Minimum training presence by polygon (MTPP)	The highest probability value at which 100% of the input polygons have at least one presence point classified as suitable habitat.
Minimum training presence by EO (MTPEO)	The highest probability value at which 100% of the input EOs have at least one presence point classified as suitable habitat.
Tenth percentile of training presence	The probability value at which 90% of the input presence points are classified as suitable habitat.
Maximum F-measure (alpha = 0.01)	The probability value at which the harmonic mean of precision and recall, with strong weighting towards recall, is maximized

Model review and refinement

Model development is an iterative process. As inputs are improved, the model can be rerun to reflect the changes. In the short term, comments from expert reviewers were incorporated into the model outputs before finalization. In the long term, updates to predictor data, or new occurrence data can be incorporated into future versions. As expected, our models evolved through several versions. The final methods described here benefitted from review. The first iteration (v1) of the models available for external review were posted at the end of May 2017. Project partners reviewed the 14 species' models. Based on initial feedback and updated field data from the 2017 field season all the models were rerun in the Fall 2017. The second iteration (v2) of the models were made available for review in December 2017.

CLIMATE ENVELOPE MODELING

We developed our CEM using the Maximum Entropy (MaxEnt) framework (Phillips et al. 2006). MaxEnt predicts the distribution of a species by inferring its environmental requirements (temperature, rainfall, etc.) from localities where it is currently known to occur (Hijmans and Graham 2006, Merow et al. 2013). MaxEnt predicts the distribution or niche of a species using species presence locations as input, often called presence-only data, and a set of environmental predictors (e.g. precipitation, temperature) across a user-defined landscape that is divided into grid cells. From this landscape, it extracts a sample of background locations that it contrasts against the presence locations (Pearson and Dawson 2003). The resulting model expresses a probability distribution for the species with each grid cell having a predicted suitability, for the environmental data used, for the species. These predictions can be interpreted as indices of climate suitability across the region. MaxEnt's predictive performance is consistently competitive with the highest performing methods (Elith et al. 2006). Compared to the previous section of this report on SDM based on Random Forest based models, we used MaxEnt as it is readily setup for climate envelope modeling.

Study Area

The study area for this part of the project was expanded to include a region around the mid-Atlantic United States including the four project states as well as all or portions of DE, CT, KY, MA, MI, NC, NH, NY, OH, TN, RI, VA, and VT. Bounding coordinates ranged from 35°N to 44°N and 71°W and 84°W.

Species Data

In addition to the data received from Natural Heritage programs in Delaware, Maryland, New Jersey, Pennsylvania, and West Virginia for the SDM part of the analysis (above), we obtained data for as much of the study area as possible. Online sources of presence data used in this study include: Biodiversity Information Serving Our Nation (BISON), and Global Biodiversity Information Facility (GBIF). GBIF data sources included Dmitriev (2015), Gall (2018), Harvard University (2018), iNaturalist.org (2017), Menard and King (2018), Natural History Museum of Utah (2017), Royal Belgian Institute of Natural Sciences (2017), Scholes (2015), University of Alberta Museums (2017), and University of Delaware (2017). We only used presence data since there was not sufficient absence data available. Original species data included 7,952 butterfly presence records.

Table 5. Counts of species data obtained for the Climate Envelope Modeling work.

Species	Training	Test	Total
Anatrytone logan (Delaware Skipper)	128	31	159
Boloria selene (Silver-bordered Fritillary)	169	42	211
Carterocephalus palaemon (Arctic Skipper)	106	26	132
Chlosyne harrisii (Harris' Checkerspot)	126	31	157
Euphydryas phaeton (Baltimore Checkerspot)	478	119	597
Euphyes bimacula (Two-spotted Skipper)	99	24	123
Euphyes conspicua (Black Dash)	218	54	272
Euphyes dion (Dion Skipper)	60	15	75
Lethe eurydice	107	26	133

(Eyed Brown)			
Lycaena epixanthe (Bog Copper)	19	4	23
Lycaena hyllus (Bronze Copper)	331	82	413
Poanes massasoit (Mulberry Wing)	180	44	224
Poanes viator viator (Broad-winged Skipper)	12	3	15
Polites mystic (Long Dash)	230	57	287

We converted polygon data into points based on the centroid. We standardized dataset attributes, as each data source has a different set of attributes for species record, and examined each dataset for potential issues such as coordinate errors.

In order to reduce spatial autocorrelation among occurrence points in the training dataset, and thus reducing overfitting in the final model, we used the 'Rarify Occurrence Data for SDMs' tool in the SDM Toolbox ArcGIS plugin (Brown et al. 2017). This tool reduced multiple occurrence records occurring within five kilometers of each other to a single record within the specified distance. From the 7,952 input points, we removed spatially redundant (e.g. duplicates) occurrence localities. Then we rarefied the occurrence data at five kilometers, removing spatially autocorrelated points removed, leaving 2,821 unique occurrence points in the final occurrence training point dataset (Table 5, Figure 1).

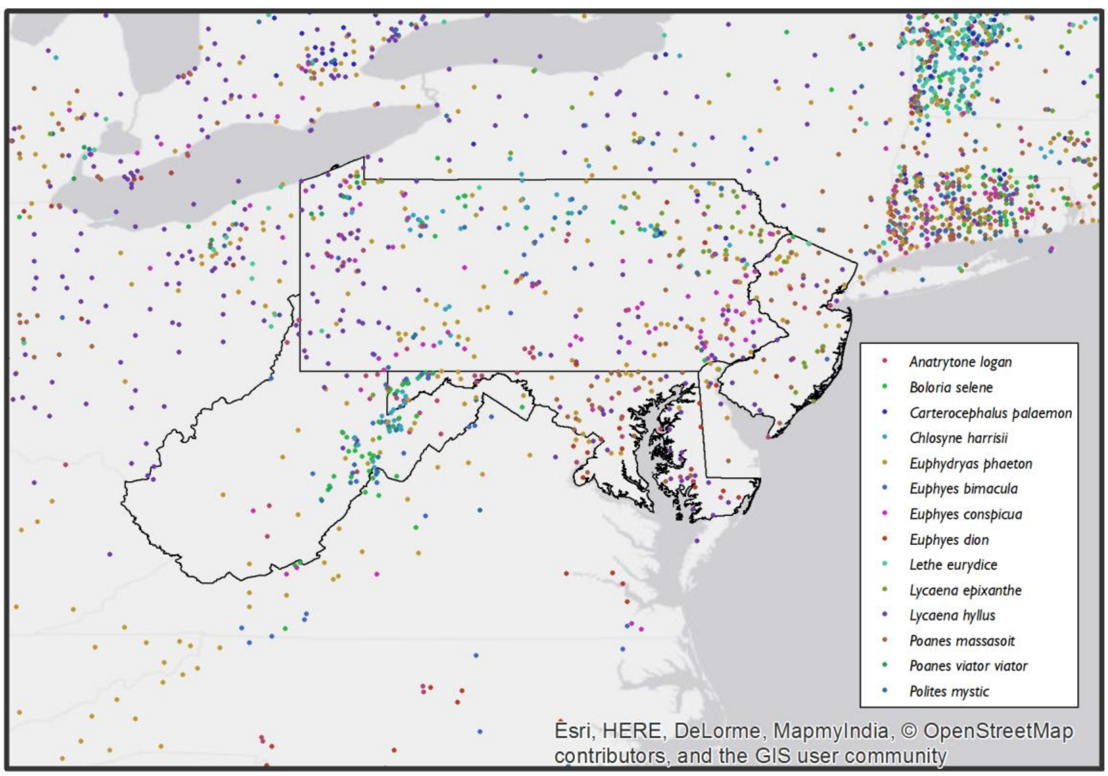


Figure 1. Map of training points for climate envelope models.

In this study, we limited the datasets combined across sources to the same date range for which historical climate data are available (1996 to 2017). For the climate modeling, we eliminated occurrences that fell outside the boundaries of species ranges maps (Cech and Tudor 2005) to reduce the number of records for which the species had been misidentified or were otherwise extreme geographic outliers.

All spatial data analyses were conducted using ArcGIS for Desktop 10.5 (ESRI 2017a) and ArcGIS Pro v2.1.0 (ESRI 2017b).

Climate Data

We obtained a set of 19 raster-based bioclimatic variables (Table 6) from among the WorldClim datasets (Hijmans et. al 2005) to describe present environmental conditions and explore the relationship between bioclimatic conditions and species distribution patterns. We downloaded climate data from WorldClim Version 1.4 (current/future projections; http://www.worldclim.org/version1). Current layers (Version 1.4) are based on 1960-2000 data. Bioclimatic variables are derived from the monthly temperature and rainfall values in order to generate more biologically meaningful variables. We removed one of each pair of highly correlated (r > 0.9) environmental variables from the bioclimatic-envelope models to avoid collinearity among variables (Gama et al. 2016, Dormann 2013). We used a qualitative assessment of the distribution of values of the variable at all presence points and of the relationship between the variable and species presence or pseudo-absence. Results of this correlation analysis are presented in Figure 2 and whether we used the variable in the models are presented in Table 6.

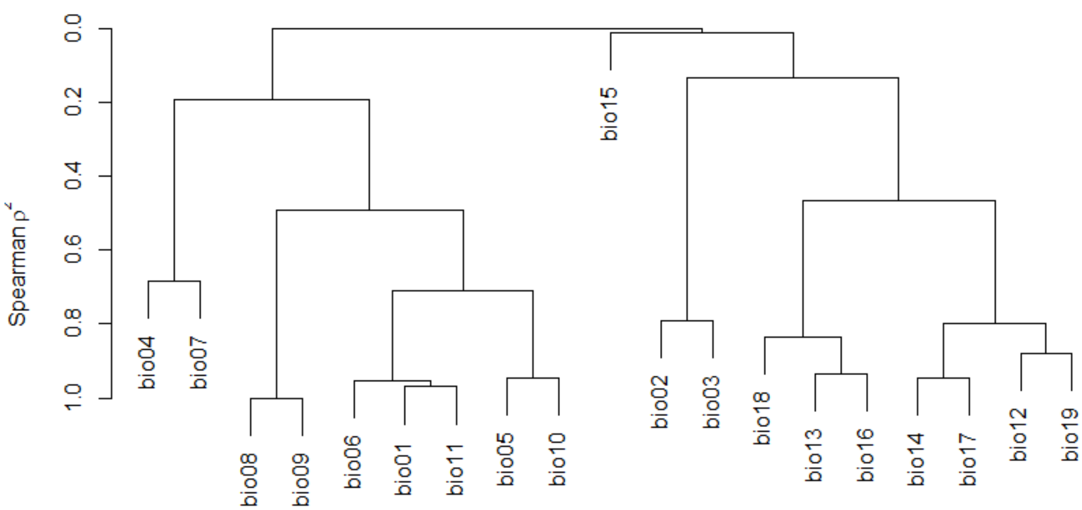


Figure 2. Results of the Spearman rank correlation analysis for 1,000 randomly sampled points across the project area.

Table 6. The 19 Bioclimatic Layers available through WorldClim. The 'Final Models' indicated if the variable was used in the models based on our correlation analysis. An 'N' in the Used column indicates that that it was not included in the final climate envelope model due to high correlations with other variables.

Variable	Description	Scale	Used
BIO01	Annual Mean Temperature	Annual	Y
BIO02	Mean Diurnal Range	Variation	Y
BIO03	Isothermality	Variation	Y
BIO04	Temperature Seasonality	Variation	Y
BIO05	Max Temperature of Warmest Month	Month	Y
BIO06	Min Temperature of Coldest Month	Month	Y
BIO07	Temperature Annual Range	Annual	Y
BIO08	Mean Temperature of Wettest Quarter	Quarter	Y
BIO09	Mean Temperature of Driest Quarter	Quarter	N
BIO10	Mean Temperature of Warmest Quarter	Quarter	N
BIO11	Mean Temperature of Coldest Quarter	Quarter	N
BIO12	Annual Precipitation	Annual	Y
BIO13	Precipitation of Wettest Month	Month	Y
BIO14	Precipitation of Driest Month	Month	Y
BIO15	Precipitation Seasonality	Variation	Y
BIO16	Precipitation of Wettest Quarter	Quarter	N
BIO17	Precipitation of Driest Quarter	Quarter	N
BIO18	Precipitation of Warmest Quarter	Quarter	Y
BIO19	Precipitation of Coldest Quarter	Quarter	N

The final cell size was 800m for the bioclimatic data, all data was entered in MaxEnt as continuous data. We ran the models using two both the 4.5 and 6.0 Representative Concentration Pathways (RCPs). The two pathways we used represented the two mid-level scenarios; the 4.5 RCP reflects a more optimistic change in temperature (an increase of 1.4°C between 2046-2065 with an increase to 1.8°C between 2081-2100) and sea level rise (an increase of 0.26m between 2046-2065 with an increase to 0.47m between 2081-2100). The 6.0 RCP reflects a more dramatic change in temperature (an increase of 1.3°C between 2046-2065 and an increase to 2.2°C between 2081-2100) and sea level rise (0.25m between 2046-2065 with an increase to .48m between 2081-2100). However, we only present results for the 4.5 RCP in this report for space considerations.

Model Run

We modeled climate using a Maximum Entropy (MaxEnt) Model Version 3.4.1 (Phillips et al. 2017). Twenty percent (20%) of the training/presence points were held back for testing. 10,000 background points were used for each model. We created a bias file (Figure 3) by using the "SDM Toolbox" ArcGIS extension (Brown et al. 2017). We did this because there was a high density of training points in the lower New England region of our study area. Additionally, we applied the 'maximum test sensitivity plus specificity' threshold rule to the final output. Infrequently (14 cases) data points very close to open water or the edges of the project area were omitted from the model as there was no climate data associated with them.

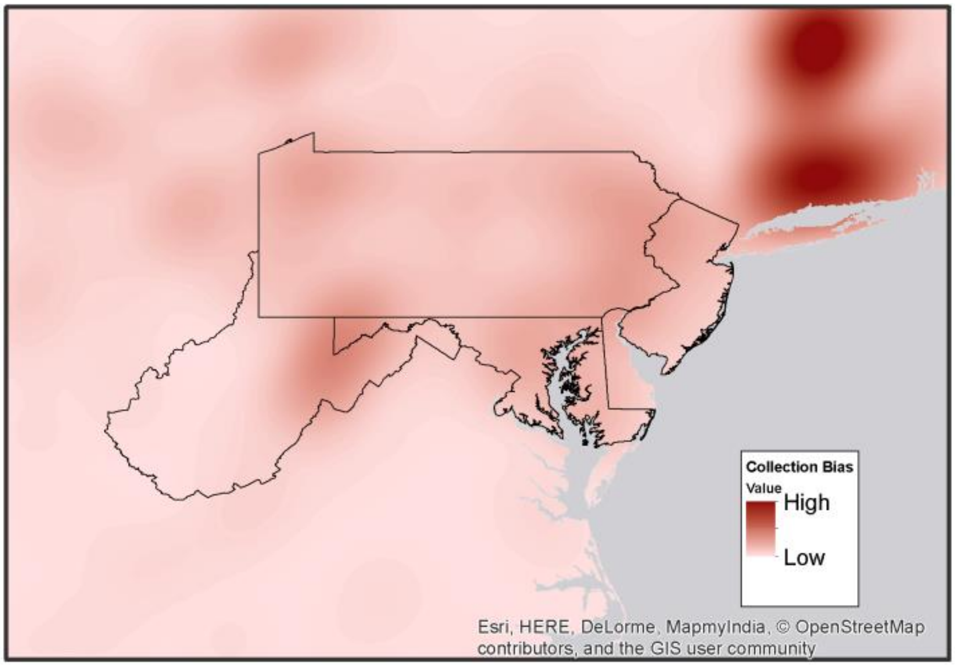


Figure 3. Bias file generated to deal with sampling intensity issues for the climate data. Dark red areas of the map indicate high sampling densities. Compare to Figure 1 for actual presence points.

Calculation of changes

Results of the climate envelope modeling were reclassified into a binary raster by the 'Minimum Training Presence' (MTP) threshold. We then combined the two binary maps in ArcGIS by the 'combine' command in spatial analyst. This produced a map with four categories: 'No prediction', 'Contracting', 'Stable', and 'Expanding'. We calculated the area of each category for each species.

HABITAT ENHANCEMENT

Large-scale wetland restoration projects were avoided in an effort to prevent "ground-disturbing" activities that would require NEPA or Historic Preservation Compliance. Instead, we focused our efforts on habitat enhancement projects, including but not limited to: mechanical clearing of trees and other woody vegetation, removal of invasive plant species by pulling, cutting, and/or herbicide treatments, and planting host and nectar plants for the benefit of the target species. Most projects were completed within 1-2 days by small groups of staff and volunteers using hand tools and small power tools (chainsaws, brush cutters, loppers, backpack sprayers, etc.). We created wetland butterfly habitat assessment forms (Appendix B) for use in evaluating sites where habitat enhancement was an option. The forms were created to be scored so wetlands could be ranked by condition and took into account plant species present, butterfly SGCN present, threats to wetland health, and other factors. Additional site selection criteria included an interest by the landowner or land manager, and on collaboration with other biologists (animal and plant experts) to ensure that habitat enhancement activities proposed for wetland butterflies would not harm other SGCN inhabiting the targeted wetland sites.

RESULTS

OVERVIEW

New element occurrence records were generated for multiple tracked SGCN species in each state. For each of the 14 target species distribution maps were populated with new occurrence data and many old records were updated. In some instances, the number of occurrence records that became apparent as a result of this project will result in reassessing species conservation status or inclusion on the SGCN list. In Maryland, at least two species will be reassessed as to whether they should remain as SGCN, Dion Skipper and Mulberry Wing, based on updated abundance and distribution within the state. In West Virginia, at least one species, the Baltimore Checkerspot, will be reassessed as to whether it should be included as an SGCN (it currently is not a listed SGCN).

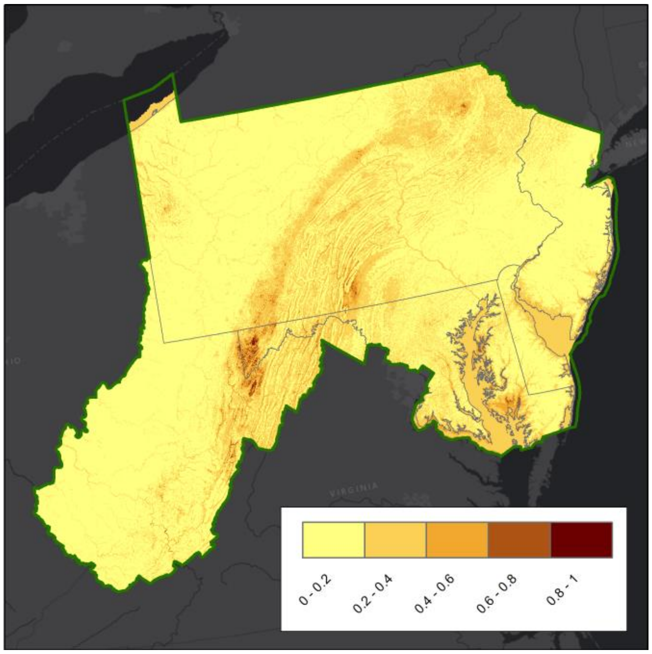
Pennsylvania generated the highest number of new occurrence records with a total of 54 representing all species in the study except for the Mulberry Wing. Thirty-three of these were found during field surveys conducted during the grant period by PNHP staff, and 21 were gathered from other sources. Two records from previously known sites were updated. A list of Pennsylvania records added or updated in 2016-2017 are listed in Appendix C. Maryland generated 16 new occurrence records for three species - Baltimore Checkerspot, Harris' Checkerspot and Bog Copper. The only other tracked species in Maryland is Two-Spotted Skipper, and this species was not documented during the grant period. New Jersey noted few occurrences for the target species even though many had been previously documented at many of the sites prior to the grant period, but did document 10 EO updates for Baltimore Checkerspots.

In addition to generating data for SGCN in all of the participating states, the project generated a wealth of modeling data for use in determining new survey areas and in prioritizing sites for enhancement, restoration and protection. The data also allow for a regional assessment of those areas that might be strongholds for a given species of interest given its distribution within the survey area and the future climate outlook in the region. Much of this data is still being evaluated. This section highlights SDM and climate model outputs for all species and reports how additional data and different assessment metrics impact model outputs.

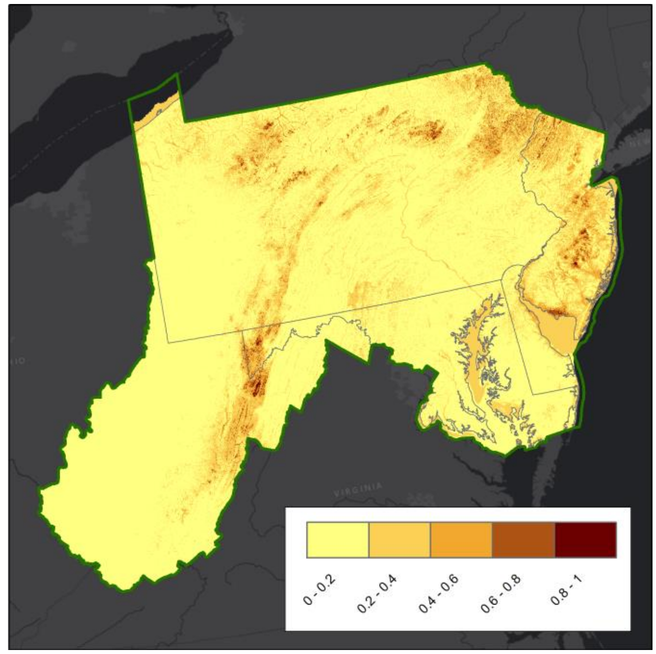
SPECIES DISTRIBUTION MODELS (SDM)

Wetland butterfly distribution models

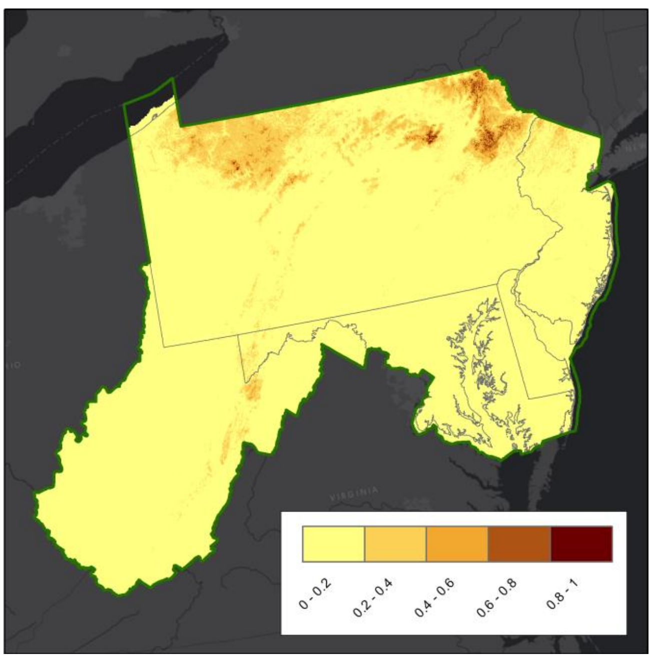
We created fourteen species distribution models, which are illustrated at the regional level in Figure 4. Models generally performed well and visual inspection by taxonomic experts verified that the models at least approximated the known distribution of the species. Nine of the fourteen models had True Skill Statistic (TSS), a measure of model performance, values above 0.8, which we categorized as 'good' (Table 7). Five models (Delaware Skipper, Dion Skipper, Baltimore Checkerspot, Bronze Copper, and Long Dash) had performance categorized as 'fair' based on TSS values and we recommend some caution in using these models.



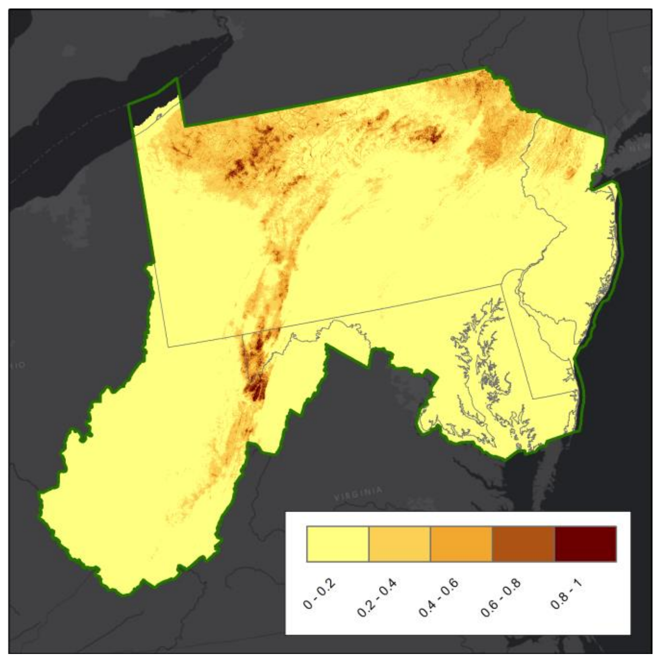
a) *Anatrytone logan*



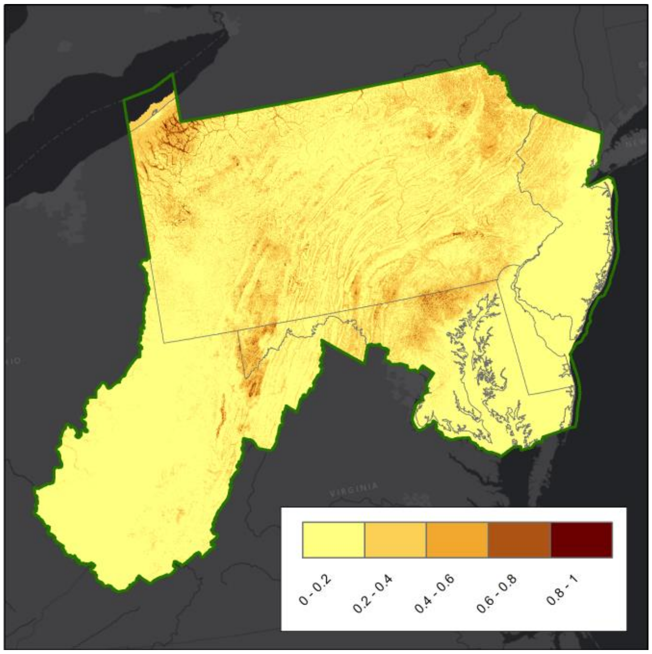
b) *Boloria selene*



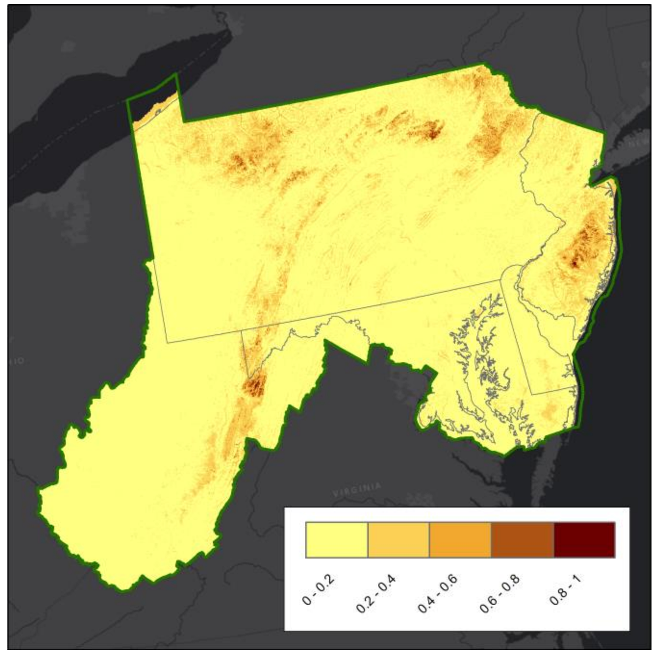
c) *Carterocephalus palaemon*



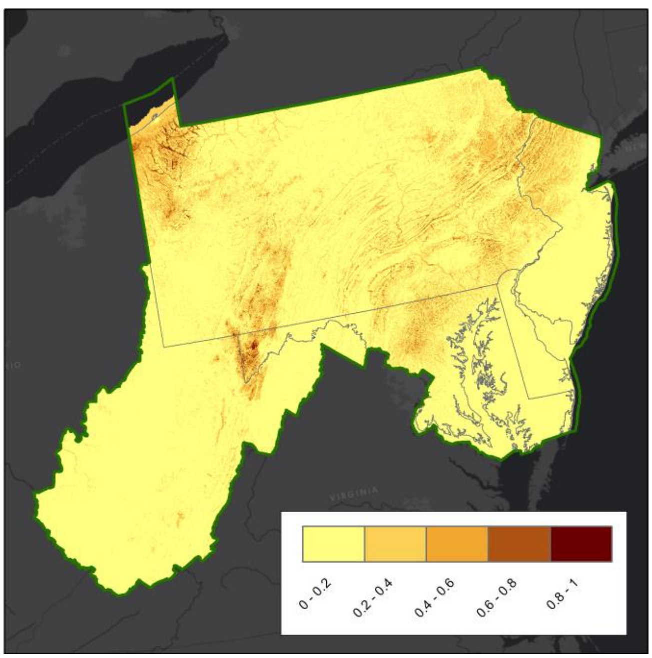
d) *Chlosyne harrisii*



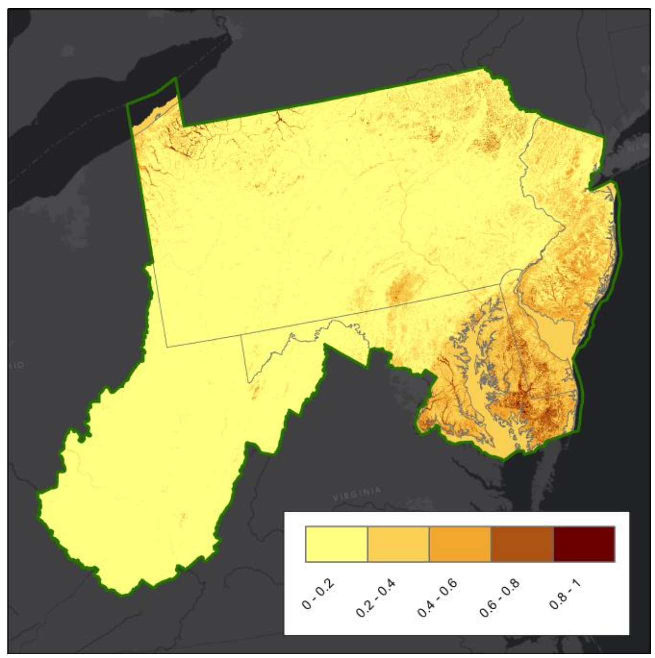
e) *Euphydryas phaeton*



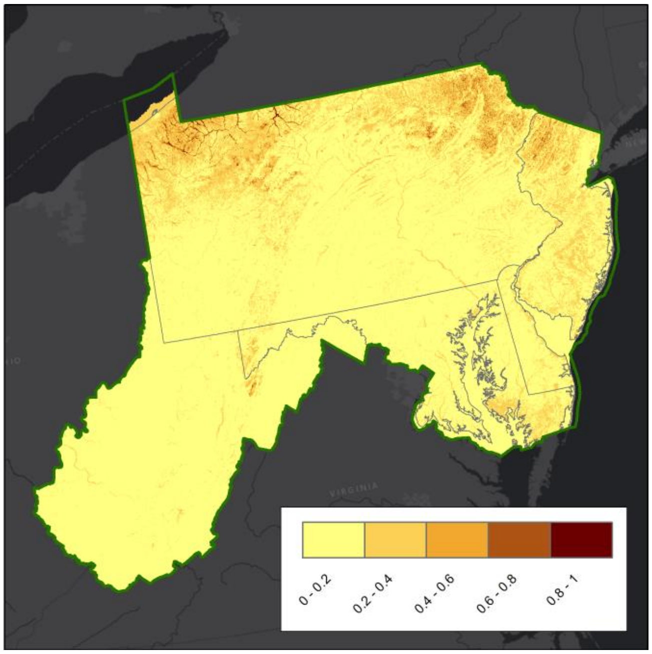
f) *Euphyes bimacula*



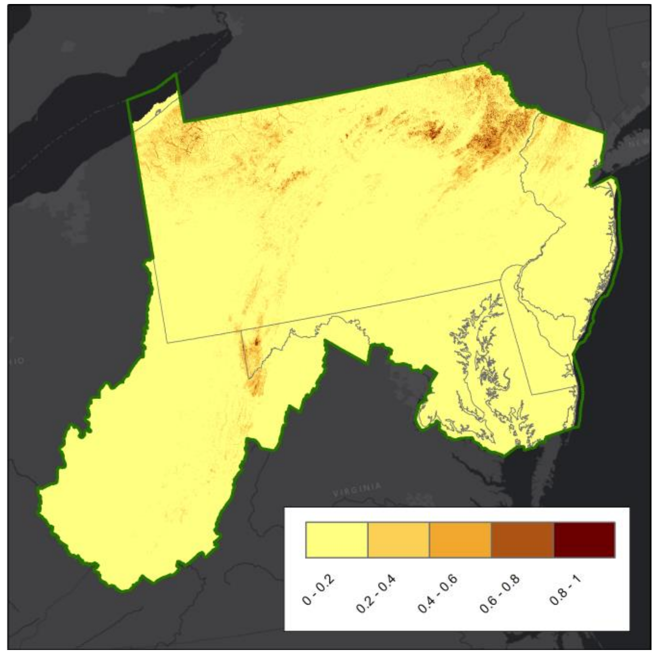
g) *Euphyes conspicua*



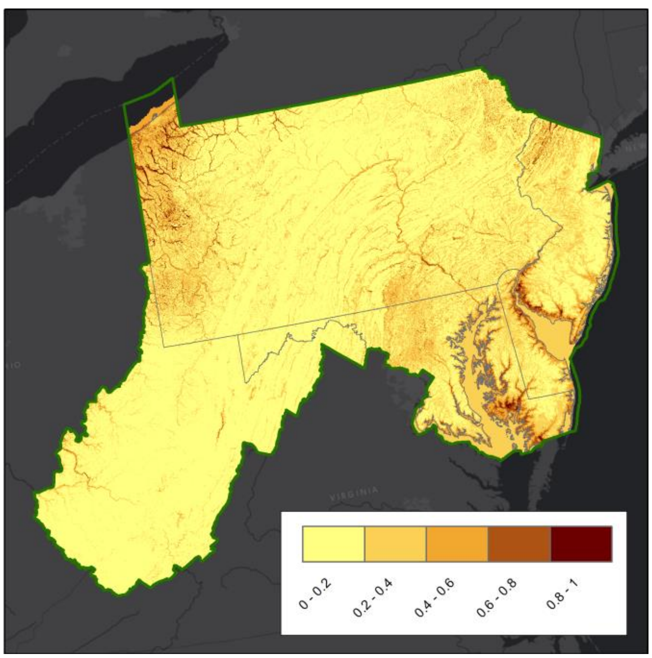
h) *Euphyes dion*



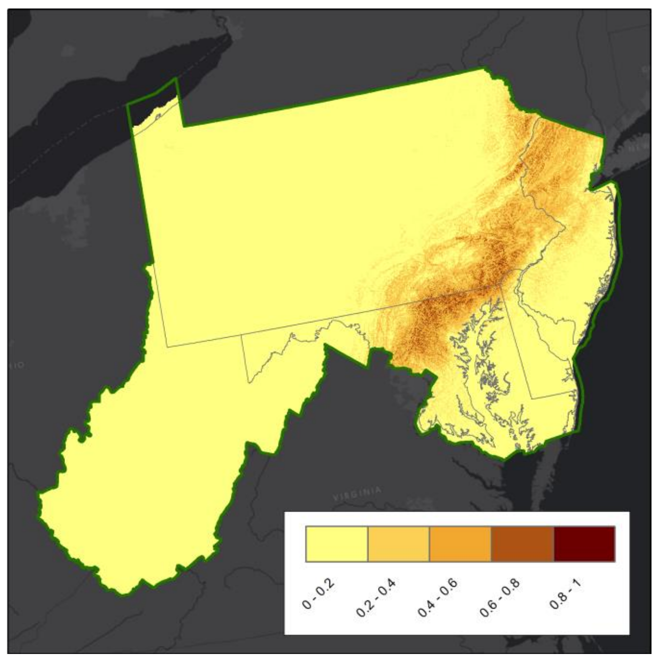
i) *Lethe eurydice*



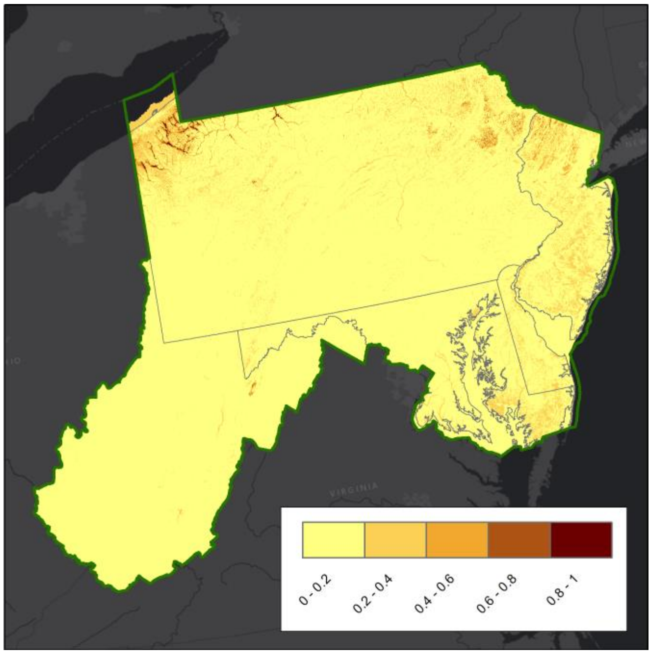
j) *Lycaena epixanthe*



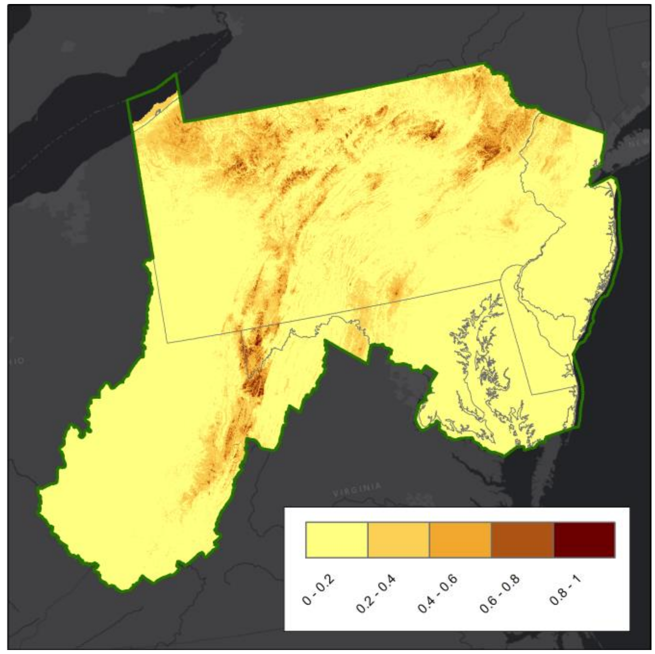
k) *Lycaena hyllus*



l) *Poanes massasoit*



m) *Poanes viator viator*



n) *Polites mystic*

Figure 4. Species distribution models for 14 wetland butterfly species.

We created at least two models for each species after each field season. For some species, we ran the models one to two additional times to improve them or correct errors. Model performance measures (i.e. TSS) generally improved with each model run.

Metadata sheets for each species model are located in Appendix J. We recommend that this PDF file should always accompany the model outputs and be available to any users using the model outputs.

Table 7. Compilation of calculated threshold values for each species model. The total number of polygons and element occurrences used in each model are shown in the '#polys/EOS' field. The True Skill Statistic (TSS) value indicates how well the overall model performed. The maximum of sensitivity plus specificity (MSS), and three minimum training presence (MTP) values are statistical values proposed to set model thresholds.

Species	# polys/EOs	TSS mean (SD)	MSS	MTP	MTP by Polygon	MTP by EO
Anatrytone logan (Delaware Skipper)	46/43	0.74 (0.33)	0.562	0.562	0.836	0.836
Boloria selene (Silver-bordered Fritillary)	211/57	0.85(0.20)	0.568	0.409	0.569	0.900
Carterocephalus palaemon (Arctic Skipper)	12/7	0.98 (0.01)	0.635	0.635	0.931	0.985
Chlosyne harrisii (Harris' Checkerspot)	78/55	0.81 (0.24)	0.685	0.529	0.697	0.929
Euphyes bimacula (Two-spotted Skipper)	28/27	0.93 (0.19)	0.672	0.577	0.964	0.964
Euphyes conspicua (Black Dash)	113/76	0.86 (0.21)	0.613	0.419	0.638	0.867
Euphyes dion (Dion Skipper)	22/17	0.77 (0.36)	0.625	0.407	0.894	0.899
Euphydryas phaeton (Baltimore Checkerspot)	186/134	0.78 (0.23)	0.485	0.226	0.622	0.651
Lethe eurydice (Eyed Brown)	12/9	0.91 (0.04)	0.683	0.603	0.960	0.960
Lycaena epixanthe (Bog Copper)	61/51	0.97 (0.05)	0.585	0.421	0.759	0.947
Lycaena hyllus (Bronze Copper)	92/68	0.76 (0.28)	0.674	0.291	0.684	0.880
Poanes massasoit (Mulberry Wing)	39/28	0.86 (0.21)	0.535	0.507	0.554	0.947
Poanes viator viator (Broad-winged Skipper)	18/8	0.98 (0.01)	0.543	0.543	0.965	0.997
Polites mystic (Long Dash)	69/51	0.78 (0.34)	0.582	0.279	0.513	0.836

The True Skill Statistic represents how well the model fits the data. TSS values were categorized as good (TSS = 0.8-1), fair (TSS = 0.5 - <0.8), and poor (TSS = 0.0 - <0.5) in terms of performance. SD = Standard deviation.

SDM Importance Variables

The model results also include a list of important variable for the models. We summarized these across all 14 species to calculate the average importance rank for each environmental variable (Figure 5). Generally speaking, environmental variables associated with wetlands (e.g. Wetland cover 1-cell mean, Distance to fresh marsh, Distance to woody wetland), topographic indices (e.g. Topographic position index, Elevation), and climate indices (Min temp of the coldest month, Max temp of the warmest month) were among the most important classes of variables across all species in this project. On the other hand, several environmental variables associated with land cover appeared to be the least important across all models. Appendix G presents a table of the nine most influential environmental variables illustrated in partial dependence plots for each species distribution model. The variables are placed in order from most to least frequently identified as important across all fourteen species. A complete list of environmental variables is provided in Appendix E.

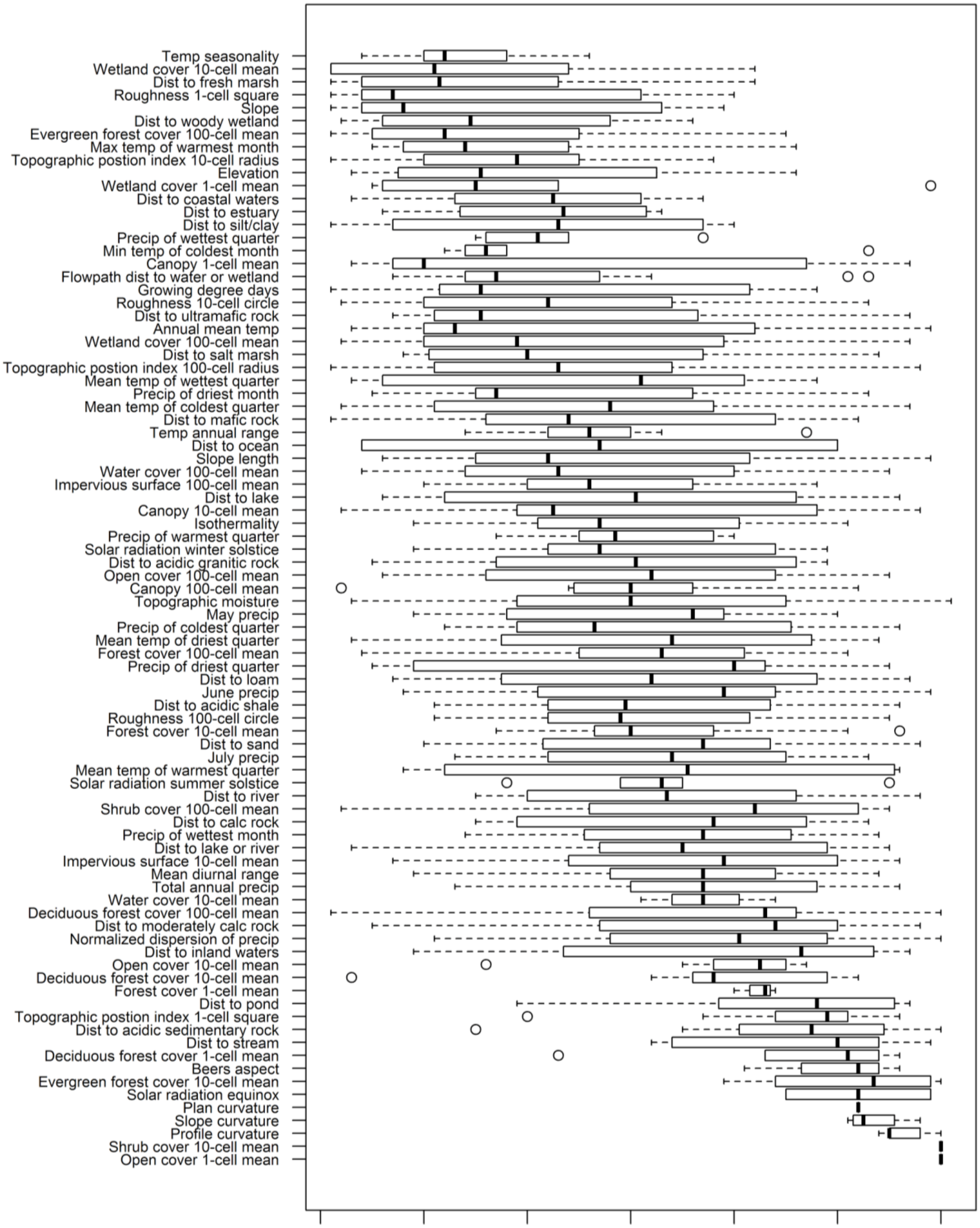


Figure 5. Mean rank importance of environmental variables across 14 wetland butterfly species. Lower rank values indicate that particular value was more important to predicting butterfly species presence.

Refining SDMs with Additional Data

The SDMs were run several times. The first run did not include the data collected during field season 2017. The final run was completed after all 2017 field data and other outstanding records were entered. The results are presented in Table 8. For all the species combined, the number of polygons increased from 612 in the initial model run to 987 in the final run (a 61% increase), and the number of EOs increased from 406 in the initial model run to 631 in the final run (a 56% increase). These majority of the added records are a compilation of older data that were aggregated and quality controlled for the final model run. A small percentage of added records were new localities of wetland butterflies documented during field season 2017. Between model runs, we continued screening and eliminating older, more imprecise records, and added newer, more precise ones. This allowed us to refine the model to better capture the probable distribution of a species at a regional scale.

Table 8. Results summary for comparing the initial and final runs of the Species Distribution Models.

Species	Model Run	#Poly-gons	# EORs	TSS mean (SD)	MSS	MTP	MTP by Poly-gon	MTP by EO
Anatrytone logan (Delaware Skipper)	Initial	23	22	0.62 (0.27)	0.705	0.705	nr	nr
	Final	46	43	0.74 (0.33)	0.562	0.562	0.836	0.836
Boloria selene (Silver-bordered Fritillary)	Initial	95	60	0.91 (0.13)	0.564	0.435	0.542	0.885
	Final	211	57	0.85 (0.20)	0.568	0.409	0.569	0.900
Carterocephalus palaemon (Arctic Skipper)	Initial	11	7	0.97 (0.03)	0.623	0.623	0.936	0.991
	Final	12	7	0.98 (0.01)	0.635	0.635	0.931	0.985
Chlosyne harrisii (Harris' Checkerspot)	Initial	61	42	0.81 (0.18)	0.657	0.586	0.761	0.937
	Final	78	55	0.81 (0.24)	0.685	0.529	0.697	0.929
Euphyes bimacula (Two-spotted Skipper)	Initial	13	13	0.98 (0.03)	0.592	0.545	0.982	0.982
	Final	28	27	0.93 (0.19)	0.672	0.577	0.964	0.964
Euphyes conspicua (Black Dash)	Initial	86	52	0.88 (0.20)	0.566	0.408	0.646	0.917
	Final	113	76	0.86 (0.21)	0.613	0.419	0.638	0.867
Euphyes dion	Initial	16	11	0.80 (0.40)	0.649	0.649	0.893	0.893

(Dion Skipper)	Final	22	17	0.77 (0.36)	0.625	0.407	0.894	0.899
Euphydryas phaeton (Baltimore Checkerspot)	Initial	55	29	0.83 (0.17)	0.498	0.246	nr	nr
	Final	186	134	0.78 (0.23)	0.485	0.226	0.622	0.651
Lethe eurydice (Eyed Brown)	Initial	11	8	0.92 (0.02)	0.545	0.545	0.964	0.964
	Final	12	9	0.91 (0.04)	0.683	0.603	0.961	0.960
Lycaena epixanthe (Bog Copper)	Initial	63	39	0.98 (0.03)	0.650	0.388	0.837	0.984
	Final	61	51	0.97 (0.05)	0.585	0.421	0.759	0.947
Lycaena hyllus (Bronze Copper)	Initial	76	54	0.79 (0.20)	0.670	0.444	0.656	0.908
	Final	92	68	0.76 (0.28)	0.674	0.291	0.684	0.880
Poanes massasoit (Mulberry Wing)	Initial	38	27	0.87 (0.21)	0.538	0.462	0.583	0.951
	Final	39	28	0.86 (0.21)	0.535	0.507	0.554	0.947
Poanes viator viator (Broad-winged Skipper)	Initial	18	8	0.97 (0.05)	0.591	0.591	0.948	0.990
	Final	18	8	0.98 (0.01)	0.543	0.543	0.965	0.997
Polites mystic (Long Dash)	Initial	46	34	0.76 (0.41)	0.562	0.316	0.600	0.877
	Final	69	51	0.78 (0.34)	0.582	0.279	0.514	0.836

Fourteen Species Overlay

A summary of all 14 models are presented in Figure 6. These areas indicate areas of the project region that may be important focal areas for wetland butterfly conservation. Numbers in the map indicate the number of species models (MTP by EO threshold) that overlap at each site. Overlaps range from 0 to 8 species at any particular site. Important areas include western Maryland south into Tucker Co. in West Virginia, northwest Pennsylvania, the Delmarva peninsula, and the Poconos.

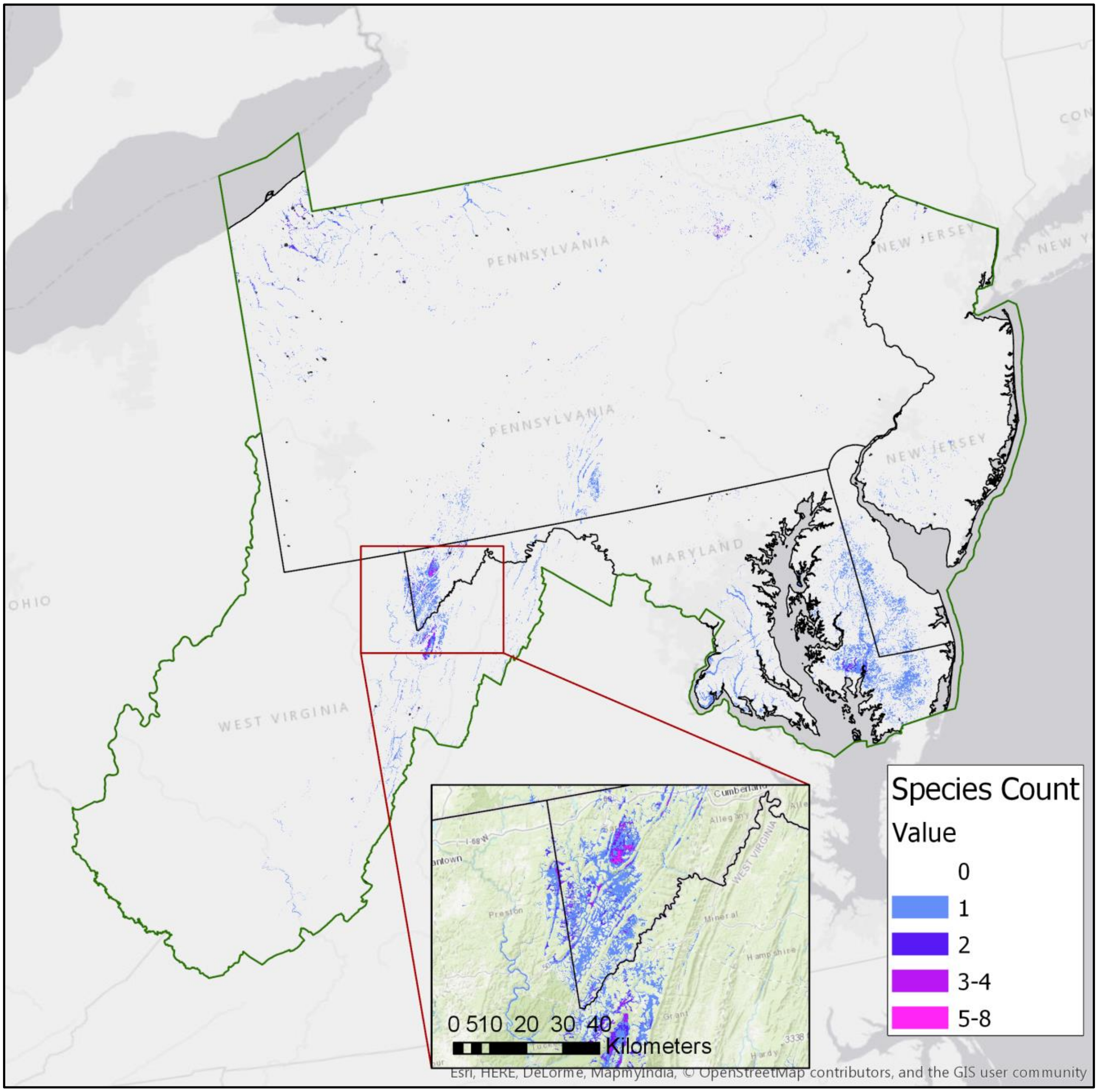


Figure 6. A summary of all 14 models is presented. These areas indicate areas of the project region that may be important focal areas for wetland butterfly conservation. The inset map shows the region of western Maryland where multiple species are predicted to overlap in the same wetland. Numbers in the map indicate the number of species models (MTP by EO threshold) that overlap at each site.

CLIMATE ENVELOPE MODELS

The output data shows the current scenario for each species (1970-2000), as well as climate space projections in 2050. Models were generally accurate, attaining a mean training AUC value of 0.826 (+/-0.108) across all 14 species Table 9. Average True Skill Statistic (TSS) values were 0.356 (+/-0.224). Copies of each model output are presented in Appendix K.

Table 9. Model statistics for the climate envelope models.

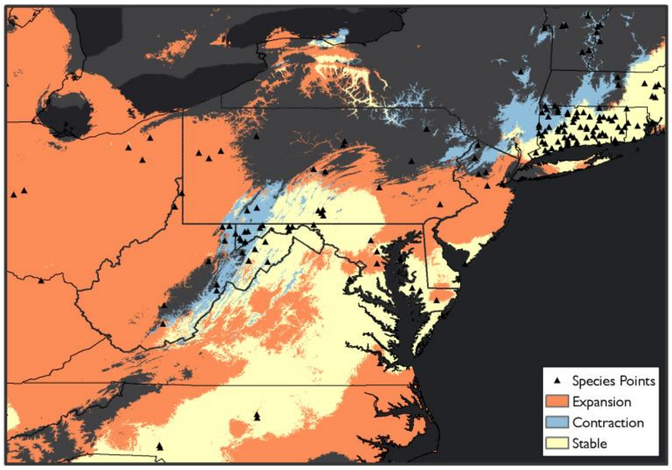
Species	Training AUC	Test AUC	TSS	MSS	MTP
Anatrytone logan (Delaware Skipper)	0.826	0.830	0.262	0.573	0.152
Boloria selene (Silver-bordered Fritillary)	0.892	0.896	0.211	0.486	0.054
Carterocephalus palaemon (Arctic Skipper)	0.908	0.849	0.609	0.124	0.042
Chlosyne harrisii (Harris' Checkerspot)	0.906	0.895	0.657	0.389	0.105
Euphydryas phaeton (Baltimore Checkerspot)	0.585	0.520	0.115	0.382	0.302
Euphyes bimacula (Two-spotted Skipper)	0.638	0.413	0.029	0.257	0.100
Euphyes conspicua (Black Dash)	0.841	0.813	0.191	0.536	0.070
Euphyes dion (Dion Skipper)	0.773	0.658	0.341	0.095	0.106
Lethe eurydice (Eyed Brown)	0.899	0.722	0.378	0.166	0.153
Lycaena epixanthe (Bog Copper)	0.869	0.838	0.513	0.398	0.120
Lycaena hyllus (Bronze Copper)	0.689	0.655	0.184	0.347	0.058
Poanes massasoit (Mulberry Wing)	0.890	0.863	0.450	0.620	0.130
Poanes viator viator (Broad-winged Skipper)	0.950	0.941	0.845	0.566	0.552
Polites mystic (Long Dash)	0.894	0.868	0.202	0.630	0.072

Present day bioclimatic envelope models

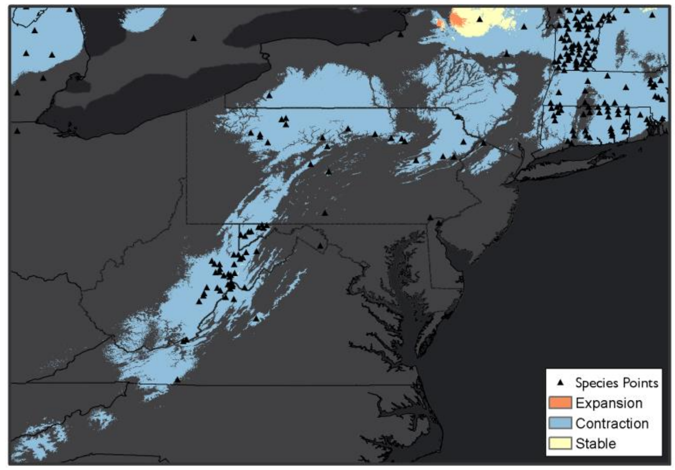
Present day (i.e. current climate) maps of suitable climate envelopes differed among all models, usually dependent on the range and ecology each species. Current distribution generally matched the general known patterns of distribution for the species based on visual comparisons of climate models to distribution maps (e.g. www.bamona.org). For example, the model for Arctic Skipper (*Carterocephalus palaemon*) showed high suitable conditions in the higher elevation regions of Pennsylvania and West Virginia within this species' range, in accordance with most of the reported species occurrence points in the region.

Projecting models to future climatic conditions

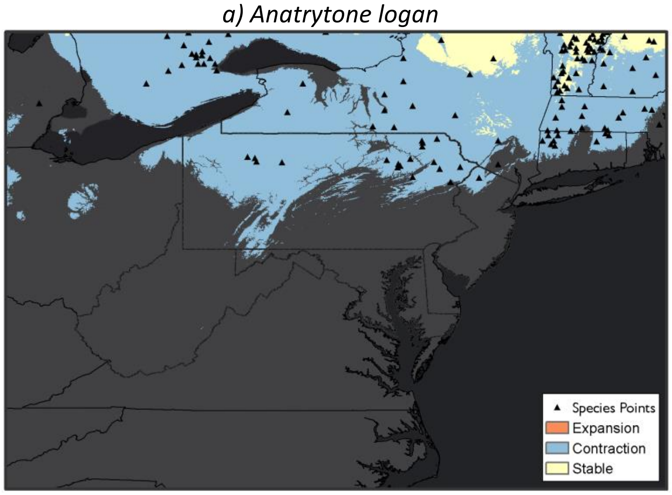
Figure 7 shows maps of the projected bioclimatic envelopes of the 14 butterfly species generated through collapsing thresholded maps of the suitable bioclimatic envelopes between present and projected future climatic conditions. These maps show areas where the models agree between present day and future (stable), future ensemble projects newly suitable conditions (expansion), present ensemble may be converted to unsuitable conditions in the future (contraction), and areas where conditions are unsuitable now and in the future (unsuitable).



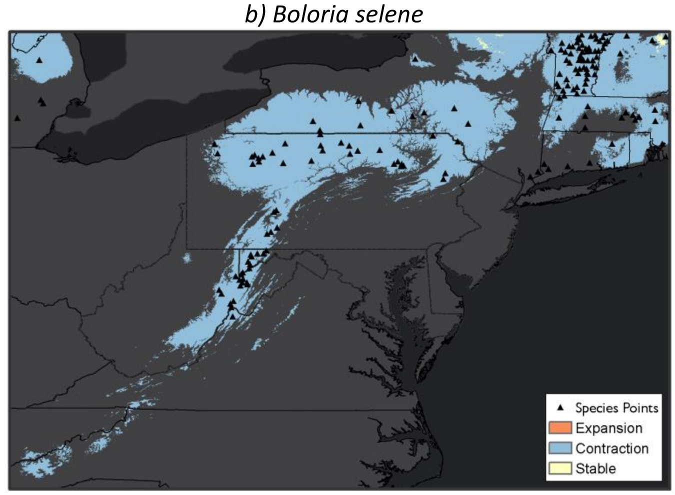
a) Anatrytone logan



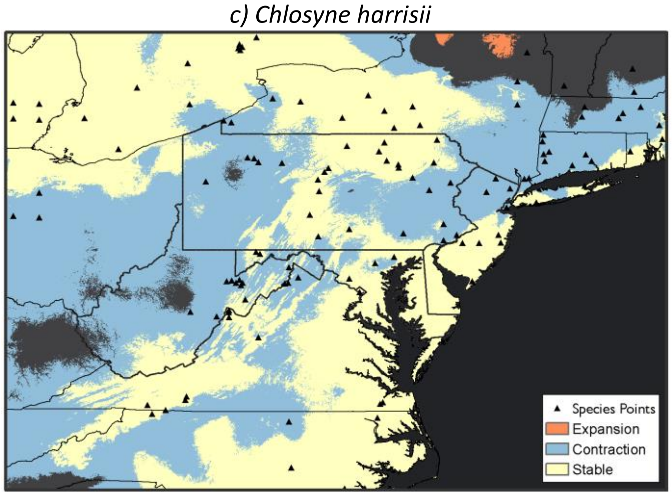
b) Boloria selene



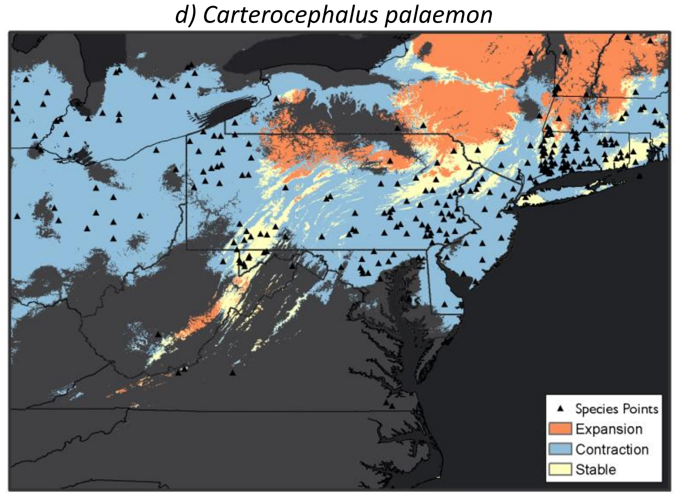
c) Chlosyne harrisii



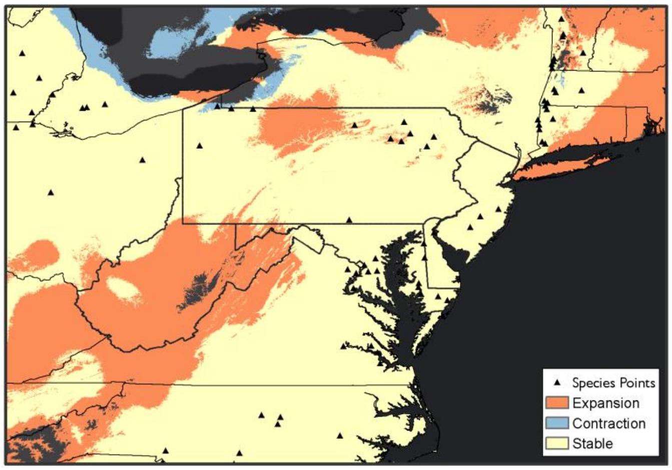
d) Carterocephalus palaemon



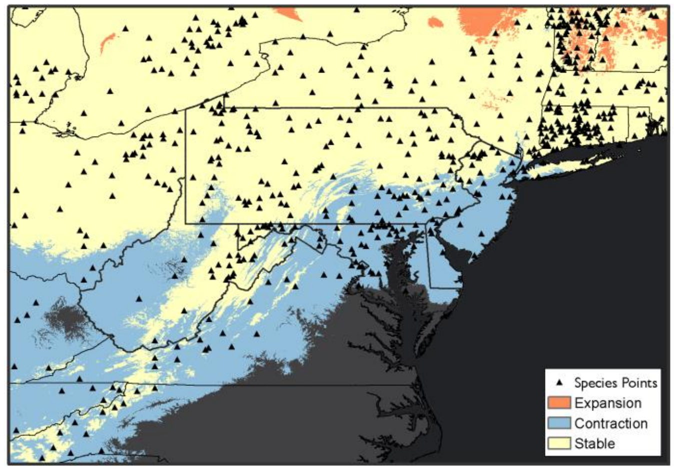
e) Euphyes bimacula



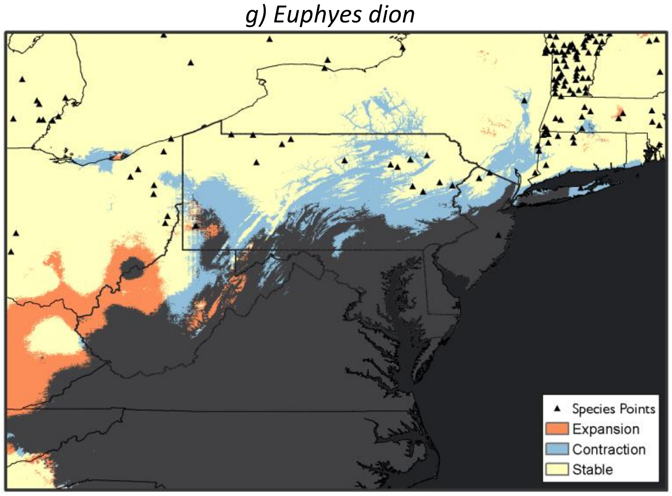
f) Euphyes conspicua



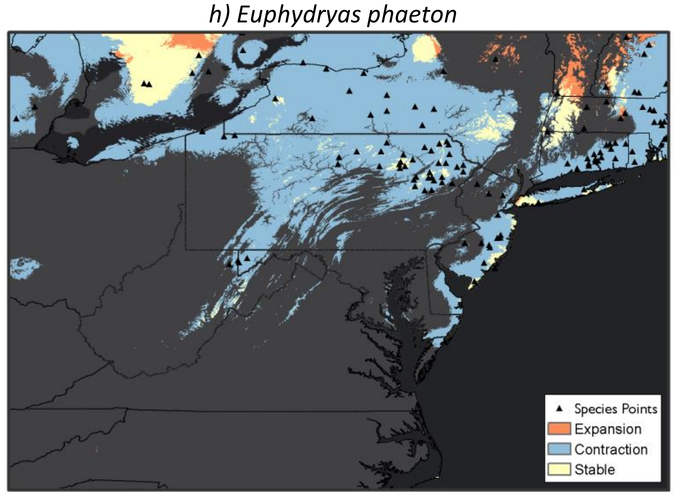
g) Euphyes dion



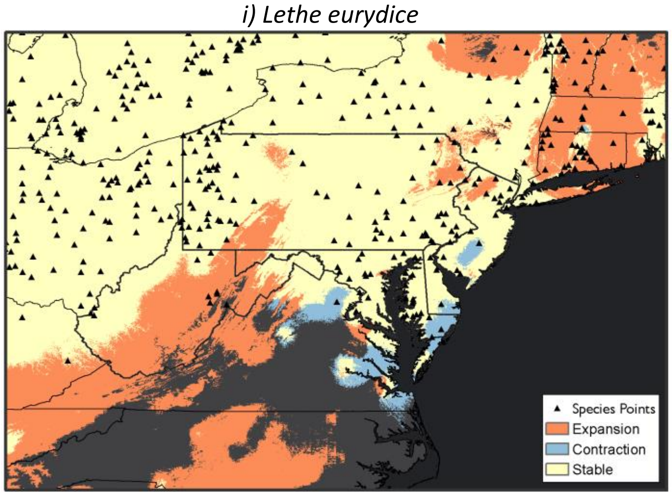
h) Euphydryas phaeton



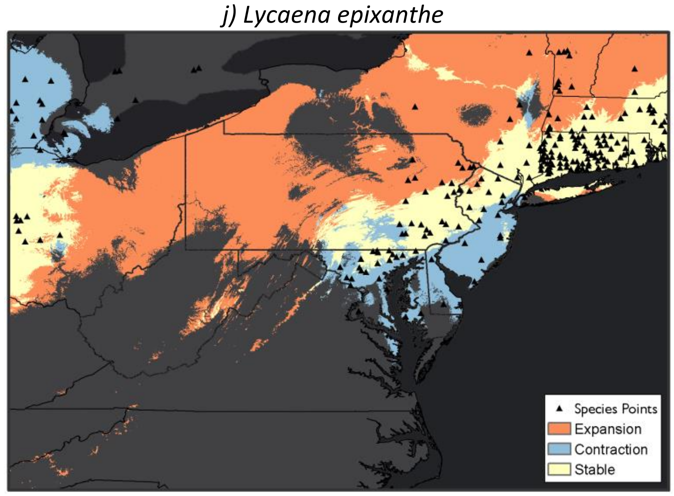
i) Lethe eurydice



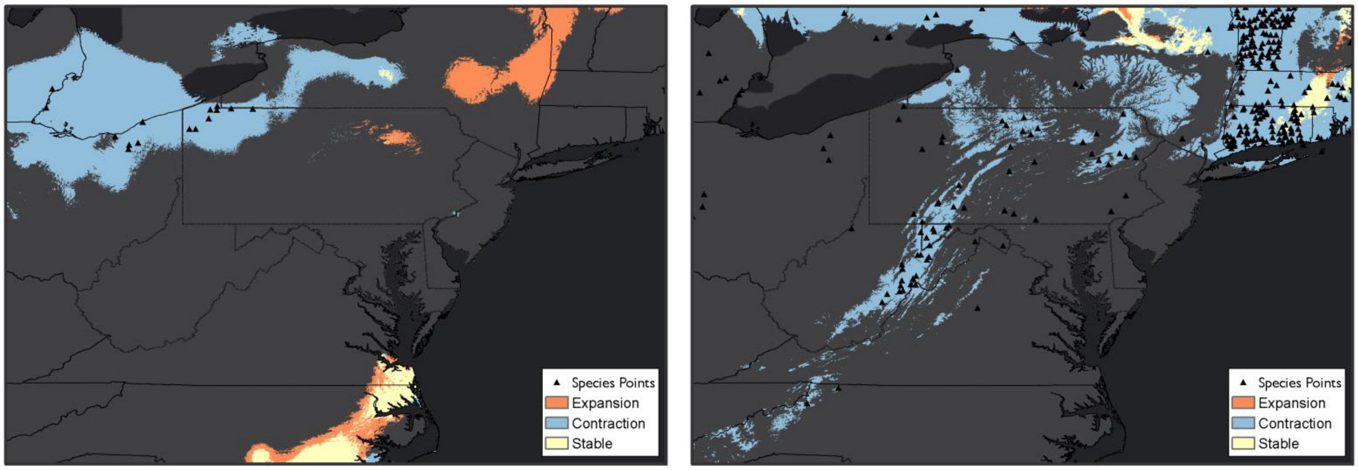
j) Lycaena epixanthe



k) Lycaena hyllus



l) Poanes massasoit



m) Poanes viator viator *n) Polites mystic*

Figure 7. Maps of expansion, stability, and contraction based on climate envelope models for 14 wetland butterfly species.

All of the 14 butterfly species exhibited changes in its bioclimatic envelope when projected from current conditions to the 2050 scenario (Figure 7). Dion Skipper (*Euphyes dion*) current area was stable in 2050, with significant expansion into the high elevations of West Virginia and Pennsylvania (Figure 7g). Three species, Harris' Checkerspot (*Chlosyne harrisii*)*,* Arctic Skipper (*C. palaemon*)*,* and Long Dash (*Polites mystic*) exhibited a significant reduction of its suitable climate envelope across the mid-Atlantic with suitable habitat moving into New York and New England. Interestingly, Bog Copper (*Lycaena epixanthe*), a species of cold peatlands, showed significant declines in its suitable climate envelope across much of the southern portion of its range (Figure 7j); stable pockets of suitable climate conditions are still present in the Poconos region of northeastern Pennsylvania.

For comparison purposes, we added the current day, suitable bioclimatic envelopes in the first column, for which the area (km^2) for each species is found in Table 10. Figure 8 shows the percentage of the project area that is predicted to be contracting, stable, or expanding for the 14 butterfly species.

Table 10. Quantitative summary of the bioclimatic envelope models for the fourteen focal butterfly species. The area (km^2) represents the amount of suitable bioclimatic conditions for each species in present day based on the agreement of at least three models. 'Present Area' is the sum of 'Contraction' and 'Stable', whereas 'Future Area' is the sum of 'Stable' and 'Expansion'.

Species	Present Area (km^2)	Future Area (km^2)	Contraction (km^2)	Stable (km^2)	Expansion (km^2)
Anatrytone logan (Delaware Skipper)	61,672	150,100	18,376	43,297	106,803
Boloria selene (Silver-bordered Fritillary)	61,564	36	61,527	36	0
Carterocephalus palaemon (Arctic Skipper)	58,196	0	58,196	0	0
Chlosyne harrisii (Harris' Checkerspot)	59,932	0	59,932	0	0
Euphydryas phaeton (Baltimore Checkerspot)	224,973	127,220	97,754	127,220	0
Euphyes bimacula (Two-spotted Skipper)	222,041	90,959	131,082	90,959	0
Euphyes conspicua (Black Dash)	139,014	38,310	115,175	2,384	14,471
Euphyes dion (Dion Skipper)	159,753	226,923	489	159,264	67,659
Lethe eurydice (Eyed Brown)	111,621	72,558	51,354	60,267	12,291
Lycaena epixanthe (Bog Copper)	75,505	3,528	71,977	3,528	0
Lycaena hyllus (Bronze Copper)	176,404	222,180	4,655	171,749	504,311
Poanes massasoit (Mulberry Wing)	64,477	111,721	32,886	31,591	80,130
Poanes viator viator (Broad-winged Skipper)	12,636	1,408	12,584	52	1,356
Polites mystic (Long Dash)	30,352	0	30,352	0	0

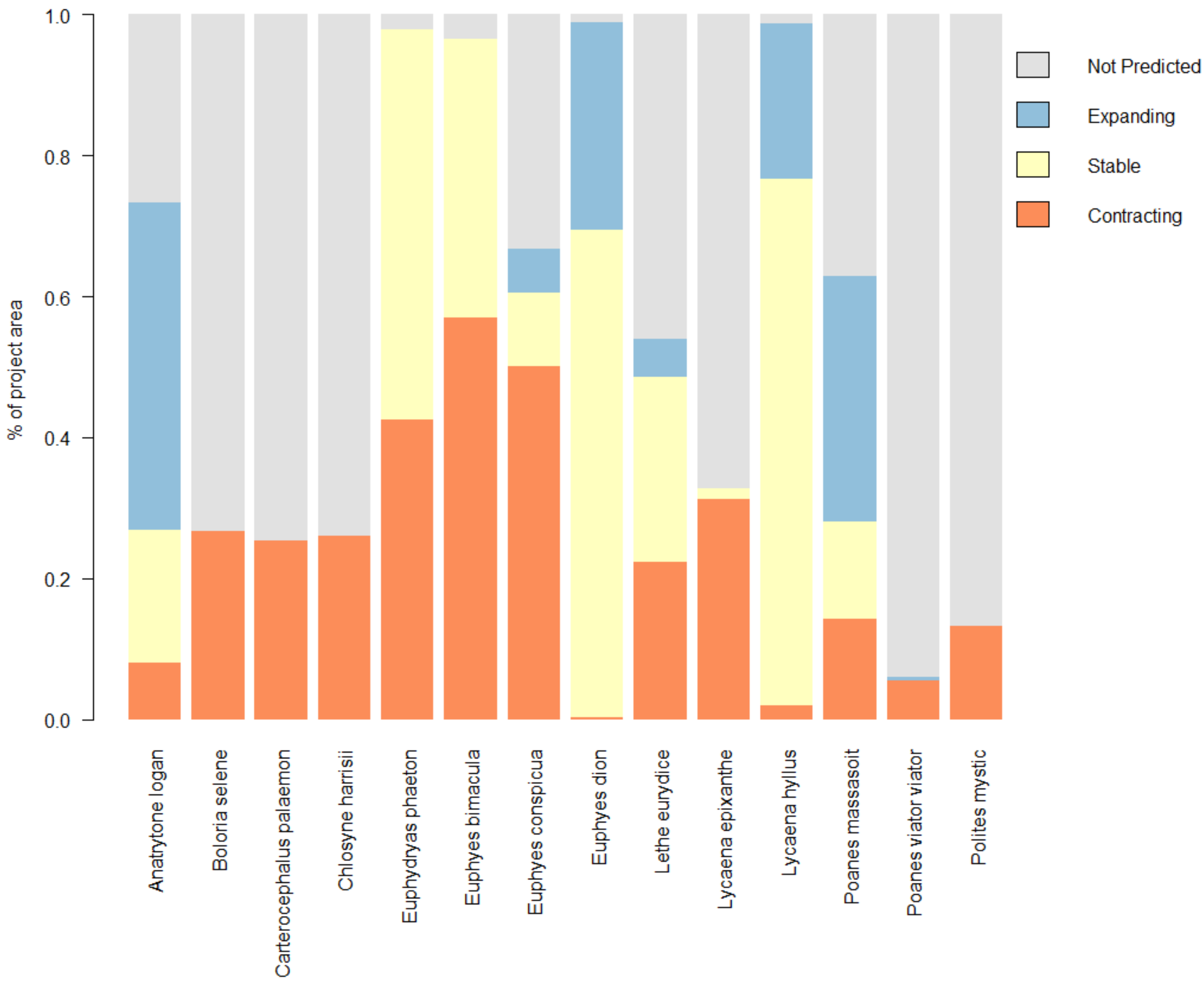


Figure 8. Percent of the project area predicted to be contracting, stable, or expanding for 14 wetland butterfly species.

Our results appear to line up with previous studies of shifting butterfly distributions due to climate change. For example, in a sample of 35 non-migratory European butterflies, 63% have ranges that have shifted to the north by 35–240 km during this century, and only 3% have shifted to the south (Parmesan et al. 1999). Beaumont and Hughes (2012) also showed similar patterns of range shifts across 24 butterfly species.

Climate and Species Distribution Modeling Overlay

There are many other factors in addition to climate that may limit species distributions, such as habitat availability, habitat fragmentation, competition with other species, and predators. Some of these factors can be included along with climate variables using the same modeling framework (variables such as land cover and elevation), but that would constitute a more general species distribution model. Here we focus specifically on climate variables and CEMs (Watling et al. 2013). In the CEM modeling literature, other non-climate predictors including land cover, were typically not found to be critical (Bucklin et al. 2015). Tracey (unpublished data) used a broad classification of geological types to predict future climate envelopes that may be limited by substrate affinity of 24 rare plant species. Additionally, CEMs do not explicitly incorporate species traits or additional information into models, although some species information is introduced into models implicitly.

In order to understand part of this issue, we overlaid the coarse future climate envelopes with the fine scale SDM data in order to determine potentially important areas for site-level conservation activities. We combined the SDM (a binary version based on the MTP threshold) and CEM (also a binary version MTP threshold) rasters for each species (resampling to the finer scale size). Results of this analysis are presented in Table 11.

Most species showed a minor to severe decline in suitable SDM-climate overlap area. Bronze Copper (*Lycaena hyllus*) showed a slight increase in area under the future climate. Three species - Arctic Skipper (*C. palaemon*), Harris' Checkerspot (*C. harrisii*), and Long Dash (*P. mystic*) - did not have any SDM area within their future climate envelope. Silver-bordered Fritillary (*Boloria selene*), had limited SDM area (~1 acre) in the future climate envelope, perhaps indicating the potential decline of these species highlighted in the CEM alone.

Table 11. Area in square meters of the overlap between the Species Distribution Model predictions and the Climate Envelope Models for 14 butterfly species in the Mid-Atlantic region. 'Present Area' is the sum of 'Contraction' and 'Stable', whereas 'Future Area' is the sum of 'Stable' and 'Expansion'. 'No Value' is the area that was not predicted under the climate envelope model.

Species	No Value Area (m^2)	Contraction Area (m^2)	Stable Area (m^2)	Expansion Area (m^2)	Present Area (m^2)	Future Area (m^2)
Anatrytone logan (Delaware Skipper)	133,245	928,572	522,811	90,792	1,451,383	61,3603
Boloria selene (Silver-bordered Fritillary)	5,158,196	7,423,298	4,510	0	7,427,808	4,510
Carterocephalus palaemon (Arctic Skipper)	471	76,3981	0	0	763,981	0
Chlosyne harrisii (Harris' Checkerspot)	189,929	5,765,756	0	0	5,765,756	0
Euphydryas phaeton (Baltimore Checkerspot)	10,157	16,449,187	41,150,367	0	57,599,554	41,150,367
Euphyes bimacula (Two-spotted Skipper)	39	792,586	946,506	0	1,739,092	946,506
Euphyes conspicua (Black Dash)	185,875	7,872,555	2,914,698	323,532	10,787,253	3,238,230
Euphyes dion (Dion Skipper)	29,950	70,259	18,240,657	416,070	18,310,916	1,865,6727
Lethe eurydice (Eyed Brown)	7,142	76,242	1,366,447	109	1,442,689	1,366,556
Lycaena epixanthe (Bog Copper)	500,439	3,610,346	574,300	0	4,184,646	574,300
Lycaena hyllus	119,463	1,580,498	42,891,848	2,529,246	44,472,346	45,421,094

(Bronze Copper)						
Poanes massasoit (Mulberry Wing)	51,506	1,528,723	4,967,840	485,367	6,496,563	5,453,207
Poanes viator viator (Broad-winged Skipper)	583,895	654,907	129	16,758	655,036	16,887
Polites mystic (Long Dash)	7,429,862	1,208,829	0	0	1,208,829	0

These overlay results should be used with caution, but may be useful as a guide for conservation and restoration activities.

HABITAT ENHANCEMENT

Maryland and Pennsylvania were involved in multiple wetland enhancement projects to improve quality for existing populations of wetland butterflies, or to create suitability for natural or facilitated reintroduction. There were no wetland enhancement projects in New Jersey or West Virginia associated with this project.

Sites in Maryland were selected for enhancement based on scored wetland butterfly habitat assessment forms (Appendix B) and on interest from conservation partners and private landowners. Pennsylvania sites were selected based solely on interest from partners and landowners. Pennsylvania staff adopted the wetland habitat survey forms and began evaluating sites using them during the course of this project. The wetland assessment forms provide a consistent, repeatable method for ranking wetland quality based on a number of factors including wetland vegetation composition, threats present, and SGCN species present. These forms have evolved as biologists and technicians have used them in the field and recommended improvements. The final score allows for ranking wetlands from higher to lower quality. Depending on funding, goals, and limitations at a site, the forms can assist biologists and land managers in finding the type of conservation project they are interested in. Projects can range from large scale wetland hydrology and plant community restoration, to smaller efforts such as targeted plantings of host plants, fencing of plants sensitive to deer browse, and localized invasive species control. These forms can also be used to identify high quality sites in need of land protection.

Maryland

Staff coordinated management activities on eight properties in Maryland during the project period to improve habitats for wetland butterflies. For reasons of data sensitivity, landowners and public lands are not named; please contact the Project Director for this information if required.

(1)Private Landowner, Baltimore County (Freeland, MD): we planted white turtlehead (*Chelone glabra*) in the wetland to enhance the habitat for Baltimore Checkerspot butterflies (white turtlehead is the primary host plant for this species). We will continue to monitor the success of the plantings in 2018.

(2) Private Landowner, Carroll County (Manchester, MD): we employed aquatic glyphosate treatment using a backpack sprayer to control invasive multiflora rose in the wetland; additional clearing planned during the post-grant period to clear encroaching red maples as well as downed trees around an existing electric fence with the hope of putting goats in next year to combat additional multiflora rose. Temporary fencing to protect the turtlehead from browsing by goats may be required.

(3) Private Landowner, Carroll County (Manchester, MD): treatment of multiflora rose using aquatic glyphosate was planned but the species was too dense and widespread within the wetland to treat with the limited staff and backpack sprayers available. The work was postponed and is now scheduled for some point during the 2018 growing season. This site is also scheduled for removal of downed trees around an existing electric fence with

the hope of putting goats in next year to combat woody invasives including the rose, which may limit or eliminate the need for glyphosate treatment.

(4) Private Landowner, Carroll County (Manchester, MD): this property contains relatively dense areas of white turtlehead and therefore one of the larger observed populations of Baltimore Checkerspots which uses the turtlehead as a host. We worked with the landowner to define an altered mowing regime in the areas of dense turtlehead that would both control multiflora rose and protect the turtlehead and Checkerspot populations during critical life stages as defined above. The landowner has committed to this new mowing regime in the designated areas.

(5) Private Landowner, Harford County (Darlington, MD): this was a Baltimore Checkerspot site in which we physically cleared small areas of dense vegetation and planted white turtlehead. This site had a moderate amount of invasive species including multiflora rose and Japanese stilt grass.

(6) Hopkins Branch Wildlife Management Area, Harford County (Darlington, MD): this site was targeted for removal of encroaching alder in the wetland; shrubs were mechanically removed using chainsaws and the stems treated with aquatic glyphosate. We also removed invasively Bradford pear trees around the wetland (also with chainsaws) and treated the stumps with glyphosate. Additional invasive species including multiflora rose will be treated during the 2018 growing season. The wetland was also planted with white turtlehead and arrowwood viburnum as there are known Baltimore Checkerspot populations in the vicinity of the wetland. This work had been scheduled for the 4th quarter of 2017 but due to inclement weather was rescheduled and completed in late January 2018.

(7) Private Landowner, Harford County (White Hall, MD): this was a Baltimore Checkerspot site in which we physically cleared small areas of dense vegetation and planted white turtlehead. We also hand pulled invasive species including purple loosestrife and Japanese stiltgrass.

(8) Private Landowner, Harford County (White Hall, MD): this site was targeted for invasive species and woody vegetation removal in the wetland during the winter of 2017. However, the site work was postponed at the request of the landowner and has not yet been rescheduled.

<u>Pennsylvania</u>
With the help of conservancy and agency partners, Pennsylvania identified 15 sites as good candidates for habitat enhancement and/or monitoring. PNHP provided technical support to partners as they developed and implemented management plans by visiting sites to look for target species and evaluate site conditions. PNHP completed site assessment and butterfly survey forms and provided habitat management recommendations that would benefit wetland butterflies and other wildlife. PNHP staff served as the lead on the enhancement activities at one site (Gifford Pinchot State Park in York County).

(1) Crossways Preserve, Montgomery County: The Wissahickon Valley Watershed Association (WVWA), recently completed a draft of a conservation management plan for this nature preserve. The Crossways Preserve includes approximately 40 acres of dry, mesic, and wet old fields and shrublands. This plan includes specific management recommendations to maintain and improve habitats for four target species of wetland butterflies: Delaware Skipper, Baltimore Checkerspots, Black Dash, and Long Dash. WVWA also conducted insect and plant inventory (in plots laid along transects) in areas targeted for habitat improvement for wetland butterflies. These surveys will help the site managers evaluate how the plant and associated insect communities are responding to the habitat management over time. WVWA was a partner and subcontractor in this RCN project. A summary of restoration work completed at the preserve is provided in Appendix H.

During the grant period, WVWA conducted invasive species removal and native planting workdays with volunteer assistance. Habitat management efforts at the Crossways Preserve have focused on removing invasive species to improve conditions for wetland butterflies and other wildlife utilizing the open wetlands and wet meadows. WVWA has installed a deer exclosure around a sensitive wet meadow area and hopes to expand the deer fencing to protect more habitat. White turtlehead is present in good numbers at Crossways Preserve and the population is targeted for protection and expansion with current restoration efforts. Baltimore Checkerspots have been documented at several locations in Montgomery and in adjacent counties but have not been found to date at the Crossways Preserve. One of the goals of the Crossways habitat restoration is to create suitable habitat for Baltimore Checkerspots so that they may either naturally colonize the site or be introduced as part of a captive breeding and release program.

(2) Acopian Preserve, Lancaster County: The Nature Conservancy (TNC) is working on a restoration plan for this site which may include wetland creation and stream channel restoration. Baltimore Checkerspots have been found at this site. We visited this site twice as The Nature Conservancy developed and significantly revised their plans based on input from regulatory agencies.

(3) Kerney Property, Adams County: This approximately 30 acre mesic to wet meadow site is impacted by a series of drainage ditches. The Nature Conservancy is working with the landowner to deactivate the ditches with the minimum amount of disturbance to the site. The site currently has good condition and diversity of native vegetation and should improve quickly with the restoration of hydrology. Additional inventory is needed at the Kerney property to assess the presence or potential for rare wetland butterflies.

(4) Waldman Property, York County. This property has approximately 55 acres of floodplain along the South Branch Codorus Creek and encompasses a series of vernal pools, floodplain forest, shrub and emergent wetland. The understory vegetation on the floodplain is dominated by invasive species in many areas. One shallow 10 acre vernal pool was dominated by reed canary grass in 2010, but with several years of focused control measures, the reed canary grass has successfully controlled and native wetland plants have regained dominance in this wetland. We conducted several site visits to determine the potential for wetland restoration on the floodplain with The Nature Conservancy, who had funding to plug wetland ditches that impair wetlands in the Chesapeake Bay watershed. Ultimately it was decided that the site did not meet the requirements in the grant in terms of total area that could be restored. Additional inventory is needed to assess the presence or potential for rare wetland butterflies, particularly Baltimore Checkerspots, as the host plant grows on site.

(5) Ricketts Glen State Park, Sullivan County: Bureau of State Park staff have begun implementation of an invasive species treatment and rotational burn program at an area called ‘The Hayfields’. This is a nearly 200 acre managed old field area that encompass approximately 25 acres of low wet meadow habitat. The site currently hosts several SGCN wetland butterflies including Harris’ Checkerspots, Baltimore Checkerspots and Eyed Browns.

(6) Lackawanna State Park, Lackawanna County: The site has approximately 25 acres of mesic meadows surrounding a small 10 acre pond. The pond has a failing dam that needs to be either repaired or removed. If the Bureau of State Parks decide to remove the dam, they would like to create a wet meadow complex with additional habitat enhancement including deactivation of drainage tiles in the adjacent fields and invasive species control. Rare wetland butterflies using the peripheral wet meadow around the pond include Baltimore Checkerspots and Eyed Browns. We applied for and received grant funding to create a plan for this work, though implementation will depend upon the final decision of Bureau of State Parks on whether to keep or remove the dam.

(7) Ohiopyle State Park: Developed preliminary plan for wetland restoration work at Ohiopyle State Park including wet meadow enhancement for wetland butterflies based on the interest of Bureau of State Parks in conducting wetland restoration at this site. At Ohiopyle State Park, Bureau of State Parks will provide matching funds to conduct restoration activities that improve the health of a forest stream and associated small wetlands. Plans are still in development, but we intend to restore an impounded and isolated wetland into a more naturally functioning ephemeral wetland that connects to the stream's floodplain. We will repair a culvert that impairs stream connectivity. A nearby channelized spring run will be restored into a more natural habitat. And an old swimming pool facility will be converted into a pollinator meadow. We applied for and received grant funding to develop the plan and implement the work.

(8) Gifford Pinchot State Park (York County): Developed replanting recommendations for a wetland restoration site at Gifford Pinchot State Park in York County. Finalized replanting and fencing recommendations for a wetland restoration site at Gifford Pinchot State Park in York County. Prepared for volunteer workdays to implement recommendations in spring of 2017. Five different planting events at Gifford Pinchot State Park followed to engage volunteers in the installation of native pollinator plants around restored wetlands. PNHP also fenced the plantings to protect them from deer browse. PNHP conducted monitoring to evaluate hydrology, to watch for invasive species regrowth, and to check on the establishment of target native plants at wetland enhancement sites at Gifford Pinchot State Park. Finally, PNHP conducted a round of invasive species control in fall of 2017 at the five restoration sites with the help of staff with the Penn State Wildland Weed Management group, who provides technical assistance and field support to the Bureau of State Parks. There were several planting workdays in the fall of 2017 to install some wetland shrubs in the pools when the water levels were low, to allow the plants to establish before re-flooding occurs.

(9) Fulshaw Craeg (Natural Lands Trust): PNHP and NLT staff looked for wetland butterfly targets and assessed habitat management needs; no target wetland butterflies observed on this visit.

(10) Green Hills (Natural Lands Trust): PNHP and NLT staff looked for wetland butterfly targets and assessed habitat management needs. No target wetland butterflies observed on this visit, but Baltimore Checkerspots were documented by a local naturalist a few weeks after our visit.

(11) Crow's Nest (Natural Lands Trust): PNHP and NLT staff looked for wetland butterfly targets and assessed habitat management needs. Baltimore Checkerspots and Black Dash were documented at two different wetlands on this preserve.

(12) Stroud (Natural Lands Trust): PNHP and NLT staff looked for wetland butterfly targets and assessed habitat management needs; no target wetland butterflies observed on this visit.

(13) Bear Creek Preserve (Natural Lands Trust): PNHP staff conducted a site visit to look for wetland butterfly targets and assess habitat management needs. A Bog Copper record was updated at two large boggy vernal pools on this preserve.

(14) Waterloo Preserve (Brandywine Conservancy): PNHP and Brandywine Conservancy staff looked for wetland butterfly targets and assessed habitat management needs; no target wetland butterflies observed on this visit.

(15) Laurels Preserves (Brandywine Conservancy): PNHP and Brandywine Conservancy staff looked for wetland butterfly targets and assessed habitat management needs; no target wetland butterflies observed on this visit.

BEST MANAGEMENT PRACTICES

Best Management Practices (BMPs) for wetland enhancement projects were developed and refined during the grant period and are listed in Appendix I. Additionally, staff developed partnerships and engaged volunteers to expand the reach of this project. We provided technical assistance including data, field visits, and management recommendations to partners interested in the conservation of wetland butterflies and other pollinators. We developed resources for partner agencies, conservancies, land managers, and land owners. These resources provide detailed information related to wetland butterfly distribution, identification and life history, habitat management, and related conservation issues.

Partnerships

Collaboration with partners involved in SGCN butterfly conservation greatly expanded our opportunities for surveys, monitoring, and habitat enhancement. Our collective work engaged citizen scientists and landowners in conservation activities. We documented public participation by recording the number of volunteers and hours spent assisting with survey and habitat improvement efforts.

Maryland

- Completed 20 and scored 14 wetland assessment forms, all in the Piedmont where surveys were completed by staff.
- Coordinated a volunteer survey effort in the Appalachian Plateau region; volunteers utilized a simplified wetland assessment form with more general habitat information (versus the full wetland assessment form).
- Received 526.5 hours of reported time from volunteers participating in habitat enhancement and wetland butterfly survey efforts
- Collaborated with Tufts University, Susquehannock Wildlife Society and Rock Lodge Trust
- Engaged 7 private landowners during the course of this project for habitat enhancement projects, plus worked with 2 additional private landowners that have agreed to let us use their properties to conduct more extensive research on target species in 2018
- Identified one new landowner who is interested in a conservation easement (most sites were public land or already protected in some way)

Pennsylvania

- Completed and scored 15 wetland assessment forms for the Pennsylvania sites listed under habitat enhancement.
- Wissahickon Valley Watershed Association (WVWA), a partner and subcontractor on this project, utilized 117 individuals and 291 hours of volunteer time across 10 different service days on habitat improvement projects at Crossways Preserve.
- WVWA developed an in-depth conservation plan for the Crossways Preserve that addresses many plant and wildlife targets, including wetland butterflies.
- PA utilized 82 volunteers and 291 hours of volunteer time across 20 different service days on habitat improvement projects at Gifford Pinchot State Park
- Collaborated with Brandywine Conservancy, Bureau of State Parks, Bureau of Forestry, Diakon Wilderness Nursery, Gifford Pinchot State Park, Natural Lands Trust, GZA GeoEnvironmental, The Nature Conservancy, and Wissahickon Valley Watershed Association on site visits and habitat enhancement planning, and active management activities.
- Engaged two private landowners working on habitat enhancement projects.

West Virginia

- Completed 47 wetland habitat assessment forms
- Documented 129 hours of volunteer efforts in wetland butterfly surveys
- Collaborated with five partners, land managers, and/ or landowners

Resources for Partners

We developed data and resources that partners can use to better understand the distribution and habitat requirements of SGCN butterflies. We provide detailed recommendations for surveys, habitat management, and other conservation measures to protect pollinators in the region. These resources can help partners as they identify conservation and restoration priorities, develop conservation easements, and write habitat management plans. More information on these and additional pollinator resources are provided in the Discussion Section.

- Habitat management for pollinators (Pennsylvania) (Appendix L)
- Life history guide to fourteen rare wetland butterflies in the mid-Atlantic (Appendix M)
- Wetland habitat assessment and enhancement best management practices (Appendix I)
- Wetland butterfly habitat assessment and wetland butterfly survey forms (Appendix B)

DISCUSSION

The results of this project are still being reviewed by the various NHPs that participated in the projects. Some species will be reassessed as to their conservation status or inclusion as SGCN. Modeling outputs will be used to guide further survey work and prioritize sites for enhancement and protection. When appropriate outputs will be shared with partners to further promote habitat enhancement and species conservation. Specific examples illustrating how the data generated by this project will be utilized are highlighted later in this section.

SPECIES DISTRIBUTION MODELS

Future Steps and Limitations of SDMs

All models can be improved with better inputs, but limits to project timelines, personnel, and budgets dictate that the modeling process is best suited to be iterative. Below are some areas that could be improved with additional time, personnel, and/or funding:

Incorporate Temporal Considerations

Temporal considerations are not adequately addressed. Ideally, environmental variables would be temporally matched to species' observation dates. For example, if a species record is from 1965, the values of environmental variables attached to that record in the training data would ideally be from 1965 as well. This is particularly pertinent for land cover variables, which have certainly changed over time in many areas due to development and other human use, as well as for climate variables which are known to be in flux. Unfortunately, temporal matching was not possible with current resources. We used a single snapshot in time for all environmental variables. The older the species observation record, the more the training data could deviate from the true environmental conditions at the time the species was observed. Given this discrepancy, it would be best to use only very recent species observations. However, for rare species with few observations, discarding all historical occurrences from the training data is simply not an option. Older records were kept if habitat appeared intact, but this does not guarantee that current conditions are comparable to those at the time of the sighting, especially considering larger landscape extents.

Include Soils Data

Geology layers were used because high quality soil data is not available across the entire study region. Improving the quality and coverage of soil data could improve the models, especially for species and habitats closely related to soil type.

Address Environmental Variable Correlation

Some of our environmental variables are highly correlated with each other. For example, we included data layers that represent solar radiation at the summer solstice, winter solstice, and equinox. Similarly, we included data layers that represent precipitation in the driest month as well as driest quarter. While the variables within each of these examples are very closely correlated, different variables may perform better in different models. Our goal was to use only the highest performing variable within pre-specified groups of correlated variables. Unfortunately, we were unable to finalize this step and had to defer this for a later update. While removing strongly correlated variables is intuitively reasonable, we note that Random Forest is robust to handling large numbers of correlated variables (Hapfelmeier et al. 2014).

Direct Future Surveys to Fill Data Gaps

Multiple political boundaries cross the ranges of most of the modeled species. The personnel and financial resources available to inventory for species can vary considerably across its range as a result. While the Natural

Heritage methodology maintains consistency in how data are reported and analyzed, the effort and ability to conduct regular, thorough inventories for rare, threatened, and endangered species can suffer under budgetary constraints and changing political policies. Thus, the quality of input data often varies throughout the study area, potentially impacting the overall quality of the SDM. On a positive note, however, region-wide SDMs such as those produced with this project can help to “even the playing field” by using information from the states with higher quality location data to assess habitat in those with lower quality data. Even so, determining how to spend limited resources for a species across its range is difficult and problematic, but consideration should be given to strengthen the data where it is the least surveyed.

Improve Input Data
We have the opportunity to improve our species input data included in future SDM iterations. Data from non-Network sources are not attributed with an EO number even though they may biologically function as part of an EO tracked in Biotics. Standards for pre-processing non-Biotics observations should be developed that will allow the data to mimic Natural Heritage methodology and more seamlessly integrate into the modeling methods. Similarly, we may consider changing how presence points are sampled based on the RA of the EO to the RA of the polygon. Some EOs are multi-part and each polygon could potentially have a different RA. Assigning a single RA to such an occurrence could potentially misrepresent presence data in the model. Overall, using Random Forest and building thousands of trees within the model could mitigate this potential effect, but alternative methods of point sampling for the validation and final model development phases should be evaluated.

Additional Limitations of SDMs
There are additional natural and anthropogenic factors that influence the distribution of species and suitability of habitats that were not incorporated into our current modeling results. Many site level disturbances and threats were not captured by the environmental variables used in this analysis, such as loss of native plant diversity to invasive species, application of pesticides and herbicides, water pollution, and poorly timed vegetation management that removes host and nectar plants when then are needed by pollinators. More research is needed to spatially quantify additional threats before they can be incorporated into models like these. Programs like iMapInvasives, a spatial database storing locations of invasive species, and better documentation of locations, rates, and types of pesticide applications, will help researchers incorporate these important variables into species models. Researchers have already begun developing landscape models for honey bees that may have broader application to other wild pollinators. These models evaluate and compare bee forage quality, pesticide use, and microclimate data using a spatial habitat model, national land-cover data, and expert review (Koh et al. 2016).

Wetlands are dynamic systems ongoing change due to natural succession of open wetlands to shrub and forest types, or human activities such as conversion to development or agriculture uses by draining and impoundment. Even conservation oriented activities such as reforestation of riparian wet meadows to meet streamside tree planting quotas can reduce the suitability of a wetland for these wetland butterflies. Wetland butterfly populations are dynamic, too. Sometimes many visits to a site are required to determine with certainty whether or not a species is found there. Once a species distribution model has been generated, it becomes a static representation of potential suitable habitat based on known populations and environmental conditions at the time of the analysis. When new populations are discovered, especially in regions with few extant records, the models can be rerun to utilize the new data and will find additional suitable habitat. For practical reasons these models should not be rerun every time a new set of environmental data are released, or a new population of wetland butterfly is discovered. But periodic rerunning of the models to incorporate new data will keep the models current and improve them over time.

New Threshold Metrics

In the application for rare species modeling where classifying all currently known locations as suitable habitat is important, minimum training presence (MTP) is an appealing metric. However, with our polygon- and EO-based approach, we realized there may be room for improvement to MTP by evaluating habitat predictions at the polygon and EO level. Thus, we developed new threshold metrics (MTPP, MTPEO) to estimate the maximum threshold (probability value) that would still classify at least one point within the set of polygons or EOs, respectively. These methods were tested in Chazal et al. (2017a, 2017b) and appear to have value for this type of work. These thresholds are explained in detail in Appendix F.

Use of SDMs in Environmental Review

The final SDMs developed in this project have excellent potential to greatly improve the success of the environmental review process by lowering the number of times projects are asked to consider a species when habitat for that species is unlikely to occur on site. At the same time, sites with suitable habitat for these targeted species are more likely to be identified when these models are incorporated into project review. The region-wide aspect of the models developed here greatly increases their utility and robustness by leveraging information across political boundaries. The region-wide environmental variables developed as part of this project will continue to have exceptional utility and applications in the years to come. Now that these have been built, completing additional distribution models will require much less preparation time and effort and will allow us, in any future regional modeling efforts, to continue to improve our analytic methods and output.

Using SDM Output

SDM Example 1: Species Assessment and Survey

Case study: Bronze Copper assessments in Maryland

Bronze Copper, currently ranked S4 in Maryland, has recently been listed as SGCN in the most recent revision of Maryland's State Wildlife Action Plan. Its addition to the SGCN list was based upon recent survey data suggesting that the species is in decline on the Coastal Plain and Piedmont Regions of the State. While targeting historic sites for follow-up survey work is appropriate in many areas, choosing new sites for survey can be difficult, particularly on the Coastal Plain where they are may be found under differing site conditions and habitats including wet meadows, small streams, roadside ditches, and even near salt marsh edges. The host plants include species of docks (*Rumex* spp.) and knotweeds (*Persicaria* spp.) which are common and widely distributed.

Figure 9 illustrates the SDM Model output when the MTP is used as a threshold to capture the distribution of Bronze Copper throughout the study area. However, when using this threshold and focusing on smaller target areas within the region, the area of suitable habitat is very broad and not likely to be effective in guiding future survey efforts. Figure 10 zooms in on one such region captured in by the SDM; using the MTP as a threshold for this species results in an output that identifies all areas along the river and its tributaries indiscriminately as potential habitat. The areas for potential habitat are so broad that they include areas of obvious non-habitat such as farm fields and homes.

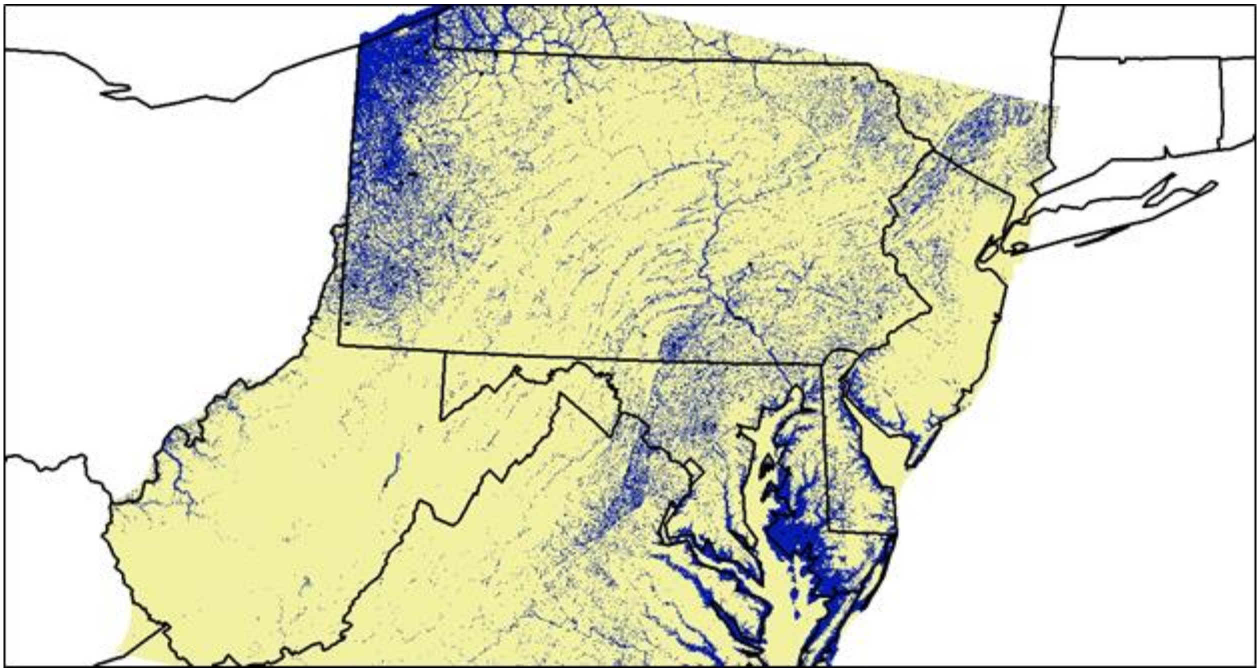


Figure 9. Map of the study area showing the projected distribution of Bronze Copper using the value of the Minimum Training Presence as the threshold. Blue areas indicate areas of suitable habitat where the species is likely to occur; yellow areas indicate areas of unsuitable habitat.

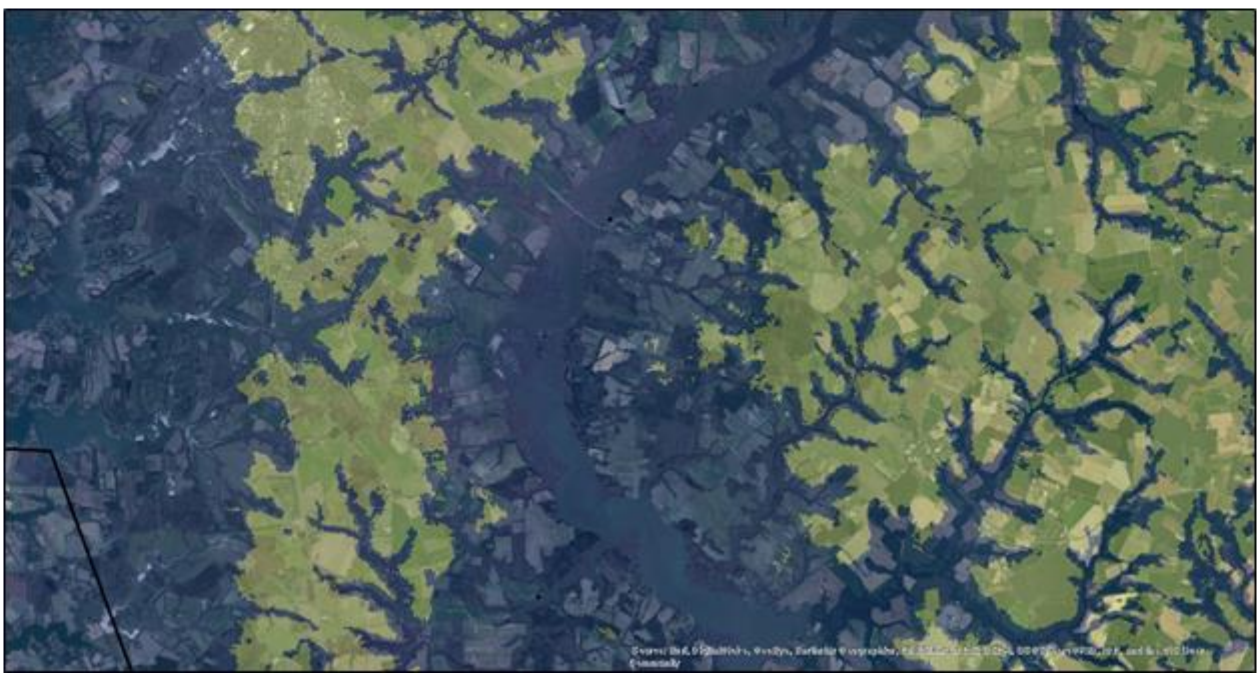


Figure 10. Aerial view of Coastal Plain area in Maryland and the resulting SDM outputs using the Minimum Training Presence value as a threshold for Bronze Copper. Blue areas indicate areas of suitable habitat where the species is likely to occur; yellow areas indicate areas of unsuitable habitat.

When the threshold value is increased to use the Minimum Training Presence by Element Occurrence (MTP-EO), it narrows the area of available habitat significantly (Figure 11).

Figure 11. Map of the study area showing the projected distribution of Bronze Copper using the value of the Minimum Training Presence by Element Occurrence as the threshold. Blue areas indicate areas of suitable habitat where the species is likely to occur; yellow areas indicate areas of unsuitable habitat.

Under this new threshold, the area of suitable habitat for Bronze Copper has narrowed considerably (Figure 12). The actual area of occupied habitat is probably somewhere in between what the maps in Figures 10 and 12 show, but the threshold level can be altered as needed to help aid in limiting survey areas in consideration of available time and resources.

Figure 12. Aerial view of Coastal Plain area in Maryland and the resulting SDM outputs using the Minimum Training Presence by Element Occurrence value as a threshold for Bronze Copper. Blue areas indicate areas of suitable habitat where the species is likely to occur; yellow areas indicate areas of unsuitable habitat.

When faced with surveying a species like Bronze Copper that may occur throughout different physiographic regions of the state, limiting focal areas allow staff and volunteers to target survey areas with the greatest potential.

SDM Example 2: Identifying Potential Species Targets for Surveys and Management

Case Study: How to prioritize species surveys and identify habitat management activities at a nature preserve in Pennsylvania.

A nature preserve in Bucks County is interested in wetland butterfly conservation. A review of records shows that Black Dash, Bronze Copper, Mulberry Wing and Long Dash were historically found within 10 km of the preserve. These old records are mapped very generally and the exact locations of the original observations are not known. An inspection of aerial imagery shows potential habitat for wetland butterflies at the preserve and at other wetlands in the vicinity. We can examine the SDM values for our preserve of interest to prioritize species surveys and identify habitat management activities that could make sites more suitable for colonization (natural or facilitated) by wetland butterflies.

The SDM values suggest how strongly conditions on the preserve match environmental characteristics of known occupied sites of the fourteen wetland butterflies. While the values that predict potential habitat for these species may not look very strong, we have evidence that within the relatively developed Piedmont of southeastern Pennsylvania, the habitats that support rare wetland butterflies can be very small and difficult to predict on a fine scale. Looking at multiple iterations of the SDM models can be a useful practice to determine thresholds on a local-regional basis. We examined the results from two iterations of the SDM model for a nearby preserve in Berks County. The first iteration of the model gave a low average value of 0.11 for the Baltimore Checkerspot at this preserve. Yet a field survey later discovered a population of Baltimore Checkerspots on site. A revised run of the model reflected this newly documented population, and as expected, the SDM value for the site increased to an average of range of 0.83. Another preserve in Chester County had a higher but still modest average SDM value of 0.39 for the Black Dash. After a field survey discovered this species at this preserve, the average value for the Black Dash increased to 0.96 in the next SDM model run that incorporated the new data. Given these examples for wetlands in the Piedmont of southeastern Pennsylvania, we decided to select species that attained an SDM value of 0.25 or higher within any cell in the focal preserve as targets for future surveys. We noted that the Long Dash has a historical record within 10 km of the preserve, but it did not make the 0.25 benchmark in the SDM model review. Conversely, the Baltimore Checkerspot, which is not known historically from within 10 km of the preserve, did attain our benchmark.

CLIMATE ENVELOPE MODELS

Future Steps and Limitations of CEMs

Climate Future Uncertainty

There are four RCPs for greenhouse gas concentration (not emissions) trajectories adopted by the IPCC fifth Assessment Report. The pathways are used for climate modeling and research and describe four possible climate futures, all of which are considered possible depending on how much greenhouse gases are emitted in the years to come. The projections are relative to temperatures and sea levels in the late-20th to early-21st centuries (1986-2005 average). We ran the models using two different Representative Concentration Pathways (RCPs), 4.5 and 6.0. These two pathways represent the two mid-level scenarios; the 4.5 RCP reflects a more optimistic change in temperature (an increase of 1.4°C between 2046-2065 with an increase to 1.8°C between 2081-2100) and sea level rise (an increase of 0.26m between 2046-2065 with an increase to .47m between 2081-2100). The 6.0 RCP reflects a more dramatic change in temperature (an increase of 1.3°C between 2046-2065 and an increase to 2.2°C between 2081-2100) and sea level rise (0.25m between 2046-2065 with an increase to .48m between 2081-2100).

Biotic Interactions

Species current and future ranges may be limited by biotic interactions, rather than by climatic factors. In addition to host plant interactions and phenological shifts discussed below, ranges might be limited by parasitism, predation, interspecific competition, or other factors. In other words, the realized niche of a species may not include parts of its fundamental niche, and projections based on the realized niche would then not include areas that may actually be suitable in the future (Araújo and Luoto 2007, Araújo and Peterson 2012). Keating et al. (2013) reviewed the literature for information on how climate change may affect biocontrol mechanisms such as insect pathogens, and diseases and incorporated their findings into a climate change assessment for 15 Lepidoptera species in Pennsylvania.

Evolutionary Change

Climate stress may cause adaptive responses in species (Hoffmann and Sgrò 2011, Parmesan 2006). We have not identified any studies have clearly documented genetic adaptation to a changing climate in lepidopteran species, but changes have been documented in other taxa groups. Despite this potential, genetic shifts may modulate the local effects of climate change, but there is currently little evidence that they will mitigate negative effects at the species level (Parmesan 2006). Much remains to be learned about preferred host plants and relationships of our wetland butterflies. Even in species with broad host plant usage, populations may prefer specific host plants regionally and avoid plants that are suitable to the same species in other parts of its range. Species may shift host plant preferences and tolerances during development or across broods (Keating et al. 2013). Different populations of a particular species may have different sensitivities to climate change based on food plant usage and other regional habitat adaptations. Interestingly, individuals in populations of the non-migratory, polyphagous imperial moth (*Eacles imperialis*) in the dry forests of Costa Rica share greater DNA sequence similarity with those from the Great Smoky Mountains in the U.S., compared with populations located much closer in Costa Rican rainforests (Janzen et al. 2005).

Host Plant Distribution
One major issue that remains unexplored for this project is the role of host plants under future butterfly distributions. Keating et al. (2013) evaluated the relationship between host plant distributions and climate change for 15 Lepidoptera species as an important factor in a climate vulnerability assessment for these species in Pennsylvania. Schweiger et al. (2008) found that host plants may not be able to track a butterfly's future range shift through dispersal limitation. Their findings strongly suggest that climate change has the potential to disrupt trophic interactions because co-occurring species do not necessarily react in a similar manner to global change, having important consequences at ecological and evolutionary time scales.

Phenology Shifts
Climate change causes phenological shifts, and if butterflies and their host and nectar plants do not shift synchronously, it may disrupt critical interactions (Parmesan 2006, Roy and Sparks 2000). For example, if eggs aren't laid on a host plant in the right stage of development, or it caterpillars don't finish development before a host plant senesces, survival is impacted (Posledovich et al. 2015, Posledovich et al. 2018). Likewise, if the flight window of adult butterflies does not overlap with the peak bloom of their preferred nectar sources, fitness would be affected (Kharouba and Vellend 2015). Keating et al. (2013) examined how phenological shifts by caterpillars and their host plants could increase or decrease climate change vulnerability for 15 Lepidoptera species in Pennsylvania. The possibility of phenological mismatch is not addressed by CEMs.

Sea–level Rise and Coastal Butterflies
Note that sea level rise is not one of the variables measured as it likely doesn't apply to many of the inland species. Reece et al. (2013) indicated that the Miami Blue Butterfly (*Cyclargus thomasi bethunbakeri*), which inhabits coastal areas of Florida, may be impacted by rising sea-levels. We recommend exploring additional models and analysis in the future for focal species in this study (e.g. Bronze Copper, Dion Skipper) which may also be impacted by sea level rise.

Using CEM Output

CEM Example 1: Prioritizing Sites for Restoration

Case study: Baltimore Checkerspot work in Maryland

The Baltimore Checkerspot is Maryland's state insect, and has experienced an apparent range contraction in recent decades. With the exception of Garrett County, most of the wetlands where it occurs are small and isolated, with low numbers of adults recorded during site visits in most instances. This has sparked an interest on the part of landowners and a variety of local organizations to create and improve habitat for the Baltimore Checkerspot. Threats to the habitat of this species include encroachment of woody vegetation, invasive species, and deer browse on white turtlehead, the primary host plant for the Baltimore Checkerspot.

Site management generally involves mechanical clearing of woody species and invasives species, as well as treatment with aquatic-use herbicides. Planting of host and nectar plants is also a component of habitat enhancement, and may also involve the construction of deer fencing around host plants where deer herbivory is an issue. Limitations of funding, equipment/supplies and in available staff and volunteers have made prioritizing sites for enhancement work necessary. Wetland habitat forms have been created for this purpose, but focus only on wetland characteristics and on populations of SGCNs present. Interest in restoration on the part of the landowner or land manager is also critical for enhancement work. Having access to climate data output now allows for an additional metric in site prioritization. The Baltimore Checkerspot map highlighted in Figure 7 indicates the projected climate space as contracting, stable or expanding and shows the locations for Baltimore Checkerspots within the four target states. The model output indicates the most stable climate space occurring in the western and northern portions of Maryland, showing a less positive outlook for Coastal Plain and Piedmont sites. Overlaying the habitat model adds another layer of complexity, although in the case of the Baltimore Checkerspot, the SDM output covers a broad area and may be less of a factor for consideration when looking at known sites for enhancement.

FUTURE DIRECTIONS

We will continue to use the data generated through this project to update species conservation statuses, guide survey work, and prioritize sites for enhancement, restoration and protection. We will continue to work together so that as new data becomes available, it can be added to the master dataset and used to re-run SDM and Climate models when needed. Maryland, Pennsylvania, West Virginia and New Jersey will continue collaborative efforts to conserve butterfly SGCN in the region and welcome the involvement of other states as well. We plan to meet periodically to continue the work we have completed thus far and determine next steps.

The methodology described in this report can also be used to model other butterfly SGCN that are of interest. Incorporating data from additional states will further improve model outputs.

Data sensitivity and data sharing issues must be considered when making the model outputs available to other NHPs and partners who want to use them for survey work and restoration. States that contribute data on rare species for SDM and climate models will need to have some control over how that data is shared. States may wish to put limitations on the thresholds that are used when viewing SDM outputs for example, which may vary depending on the species and the audience viewing the output. Designing a procedure to do this is beyond the scope of this project. Rather we suggest that this be something that is incorporated into data sharing agreements on a case by case basis. For the purposes of this report, GIS data and shapefiles have not been included as a product but can be made available if requested.

Upcoming Projects

Data generated as a result of this project led to additional opportunities that extend beyond the scope of this grant. The following three projects are currently underway and will utilize data and products generated in part by this RCN grant.

Pennsylvania Conservation Opportunity Areas Tool - The Pennsylvania Conservation Opportunity tool will provide site level mapping of SGCN and other information from the 2015 State Wildlife Action Plan into an easy to use web interface for conservation decision making. SDM output from this project was incorporated in the PA COA tool to represent potential distribution of the 13 SGCN butterflies present in Pennsylvania. Mapping was based on the MTPEO Threshold representing the 'Medium' probability of occurrence in the tool, and MTP and above for the low probability of occurrence. This Tool will be available in 2018.

Maryland NHP will continue to work with researchers at Tufts University to determine the phenology of Baltimore Checkerspots and the hosts and nectar plants present at several targeted sites in the Maryland Piedmont Region. Research goals include identifying limiting factors for Baltimore Checkerspot larvae and adult butterflies that prevent populations from thriving or cause individuals to abandon wetland habitats. While the grant period is now over, the information gathered in 2016 and 2017 allowed for the development of research questions and the selection of field sites for 2018-2020 projects, and will aid in refining existing BMPs and perhaps determining new BMPs that will aid in the enhancement of small wetland habitats.

The Pennsylvania Natural Heritage Program will continue to provide technical assistance to conservancy and landowner partners as they develop and implement habitat restoration and management plans. We will begin to utilize the wetland butterfly SDM and climate model results within our programmatic and with our partners. We recently received word that a grant proposal we submitted to the Pennsylvania Department of Conservation and Natural Resources - Bureau of Recreation and Conservation was funded. Through this project, we will work private landowners to document the wildlife using their vernal pool and wet meadow habitats and evaluate the health of their wetlands. We will provide recommendations on how to protect water quality, increase habitat connectivity, retain sources of food and shelter for wildlife, control invasive species, promote native trees and understory vegetation, and best management practices for logging. We will connect landowners to conservancy partners and provide information that can be used in conservation planning and development of easements. This grant will also provide funds to restore and enhance wetlands in Lackawanna and/or Ohiopyle State Parks, and host workshops to train participants in the steps for designing and implementing a wetland restoration or enhancement plan. This RCN wetland butterfly grant helped us develop data, resources, and connections that will be advanced through the DCNR grant.

BEST MANAGEMENT PRACTICES

We developed or identified the following resources for partners to use to better understand the distribution and habitat requirements of SGCN butterflies. These documents provide detailed recommendations for surveys, habitat management, and other conservation measures to protect pollinators in the region. These resources can help partners as they identify conservation and restoration priorities, develop conservation easements, and write habitat management plans.

Featured Resources

Life History Guide to Fourteen Rare Wetland Butterflies in the Mid-Atlantic
A guide to the life history of the fourteen wetland butterflies studied under this grant is provided in Appendix M. This guide features photos of adults, caterpillars, and host plants, along with notes on species identification, range, habitat, phenology, and reproduction.

Pennsylvania Habitat Management Guide for Pollinators
This guide developed for Pennsylvania land managers and property owners is provided in Appendix L. We provide information on best management practices for:

- promoting habitat variety to support all life stages of pollinators including adults and immatures
- maintaining open habitats through management practices such as rotation mowing
- controlling invasive plants
- protecting pollinator diversity and rare species
- selecting native and local plants for pollinators

Wetland Butterfly Habitat Enhancement Best Management Practices
A list of recommendations for selecting sites for wetland enhancement, and minimizing impacts to wetland butterflies and other wildlife when conducting management activities (Appendix I).

Wissahickon Valley Watershed Association - Wetland Restoration Report
The WVWA conservancy is working to increase the suitability of the Crossways Preserve in Montgomery County Pennsylvania for a variety of wildlife, including wetland butterflies. This restoration and habitat management summary provides an example activities that can be employed to reduce invasive species and increase native plant diversity (Appendix H).

Additional Regional Pollinator Resources

Maryland Department of Natural Resources Pollinator Habitat Management Plan

A blueprint for the conservation of pollinators and pollinator habitat on the natural lands managed by the Maryland Department of Natural Resources for the benefit of Maryland's citizens. Available online at: http://dnr.maryland.gov/wildlife/Documents/PollinatorHabitatPlan_June2017.pdf.

Maryland State Wildlife Action Plan

The MD-SWAP was developed to assess the health of wildlife species in and the habitats on which they depend, to identify threats to species' survival, and provide conservation actions to maintain sustainable populations of SGCN and common species in the state. Invertebrate species are addressed in Chapter 3 on pages 64-73. This chapter is available online at: http://dnr.maryland.gov/wildlife/Documents/SWAP/SWAP_Chapter3.pdf. A special case study on SGCN pollinators describing their importance, threats, conservation actions is presented on pages 67-69.

New Jersey Monarch Conservation Guide

Available online at: http://www.nj.gov/dep/docs/monarch-guide.pdf

New Jersey State Wildlife Action Plan

Available online at: http://www.state.nj.us/dep/fgw/ensp/wap/pdf/wap_plan17.pdf

Pennsylvania Pollinator Protection Plan (P4)

The P4 (2018) outlines the current status of pollinators in Pennsylvania, and provides recommendations for best practices and resources to support and expand pollinator populations. The P4 was developed as a collaborative effort with contributions from 36 individuals representing 28 state and national organizations and stakeholder groups. There are chapters on Best Practices for Forage and Habitat, Best Practices for Pesticide Use, and Best Practices for Beekeepers, and Recommendations for Research, Policy and Communication. The Recommendations Chapter (Chapter 5) provides a comprehensive list of recommendations for research, policy, and communication. These recommendations are a good reference for researchers, policy makers, funding programs, land managers, and conservation groups to consider when designing programs and selecting projects to benefit pollinators. The P4 is extensively hyperlinked to other online resources and can be viewed and downloaded at Pennsylvania Pollinator Protection Plan Website at: http://ento.psu.edu/pollinators/research/the-pennsylvania-pollinator-protection-plan-p4

Pennsylvania State Wildlife Action Plan

The purpose of the PA-SWAP (2015) is "to conserve Pennsylvania's native wildlife, maintain viable habitat, and protect and enhance Species of Greatest Conservation Need (SGCN), in order to conserve our Commonwealth's rich natural heritage for future generation". This plan details urgent conservation and management issues that need to be addressed to keep common species common, conserve species of global and regional importance, reduce knowledge gaps to better assess the conservation status of species, and promote partnerships for wildlife conservation. The conservation status of 67 pollinator species and 383 other invertebrates were evaluated for this plan. The invertebrate assessment report located in Chapter 1, Appendices 1.1, which is available online at http://www.fishandboat.com/Resource/Documents/SWAP-CHAPTER-1-apx11-12.pdf. The full invertebrate assessment is located on pages 76-149, with a special section on pollinators on pages 118-120.

West Virginia Monarch Summit

West Virginia is holding a Monarch Summit in early March 2018, the first in the Northeast, to provide an opportunity and venue for stakeholders in the state to gather and learn about monarch butterfly natural history and habitat. Through Focus Sessions at the Summit, facilitators will solicit goals and objectives from stakeholders to generate a framework from which a collaborative and voluntary West Virginia monarch conservation plan can be developed by the West Virginia DNR, hopefully with collaboration from stakeholders. The DNR anticipates that monarch conservation will also benefit pollinators in general, including the wetland species targeted in this grant.

West Virginia Pollinator Protection Plan

The West Virginia Department of Agriculture developed a Pollinator Protection Plan to address threats to managed hives on agricultural lands. The main focus of the document is to provide best management practices to growers in relation to managed hives and pesticide use on the properties. The plan mentions the value of native pollinator species, and that stakeholders should strive to reduce pesticide use and provide habitat for all pollinators. Available online at:
http://agriculture.wv.gov/divisions/plantindustries/Documents/PID%20pdfs/WV%20Managed%20Pollinator%20Protection%20Plan_Final.pdf

West Virginia State Parks Mowing Reduction Program

WV State Parks, with technical assistance form the WV DNR, has instituted a mowing reduction strategy on its properties. The goals for this program include: reduced financial expenditure in manpower and equipment, beautification, increased visitor experiences, and enhancement of pollinator and monarch butterfly habitat. Several parks have participated in the two years of the program and have enhanced pollinator habitat mostly through simply not mowing. Very little additional management has been done initially to reap significant habitat improvement. Additionally, visitor appreciation of the new habitats has been wholly positive. An expansion of the program is anticipated in the coming year.

West Virginia State Wildlife Action Plan

West Virginia's State Wildlife Action Plan (2015) provides a general list of the state's native pollinators, basic natural history, and threats to populations. Pollinators are addressed in section 3.2.1.15. Available online at:
http://www.wvdnr.gov/2015%20West%20Virginia%20State%20Wildlife%20Action%20Plan%20Submittal.pdf

LITERATURE CITED

Araújo, M.B. and M. Luoto, 2007. The importance of biotic interactions for modelling species distributions under climate change. Global Ecology and Biogeography, 16:743-753.

Araújo, M.B., and A.T. Peterson, 2012. Uses and misuses of bioclimatic envelope models. Ecology 93(7):1527-39.

Breiman, L. 2001, January. Random Forests [Online]. Available: https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf. [Accessed: 30-Jun-2017].

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. Classification and regression trees. Chapman & Hall, New York.

Brown et al. 2017. SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. PeerJ 5:e4095; DOI 10.7717/peerj.4095

Bucklin, D.N., M. Basille, A.M. Benscoter, L.A. Brandt, F.J. Mazzotti, S.S. Romanach, C. Speroterra, and J.I. Watling. 2015. Comparing species distribution models constructed with different subsets of environmental predictors. Diversity and distributions, 21:23-35.

Cantor, S. B., C. C. Sun, G. Tortolero-Luna, R. Richards-Kortum, and M. Follen. 1999. A Comparison of C/B Ratios from Studies Using Receiver Operating Characteristic Curve Analysis. Journal of Clinical Epidemiology 52:885-892.

Cech, R. and Tudor, G., 2005. Butterflies of the East Coast: an observer's guide (p. 345). Princeton, NJ: Princeton University Press.

Chazal, A. C., J. Oetting, A. Knight, K. Hazler, R. Gilb, and N. Pasco. 2017a. Integration of At-risk and Range-Restricted Species Models and Strategic Conservation Information into the SALCC Conservation Blueprint. Natural Heritage Technical Report 17-14. Virginia Department of Conservation and Recreation, Division of Natural Heritage, Richmond VA. 71 pages plus appendices.

Chazal, A. C., T.G. Howard, K. Hazler, R. Gilb, A. Conley, and C. Tracey. 2017b. Species Distribution Modeling of Threatened and Endangered Species throughout Region 5. Natural Heritage Technical Report 17-12. Virginia Department of Conservation and Recreation, Division of Natural Heritage, Richmond VA. 30 pages plus appendices.

Cutler, D. R., T. C. Edwards Jr, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. Ecology 88:2783–2792.

Diaz-Uriarte, R. 2014. varSelRF: Variable Selection using Random Forests. R package version 0.7-8, https://cran.r-project.org/web/packages/varSelRF/varSelRF.pdf.

Díaz-Uriarte, R. and De Andres, S.A., 2006. Gene selection and classification of microarray data using random forest. *BMC bioinformatics*, *7*:3.

Dmitriev, D. 2015. Illinois Natural History Survey Insect Collection. Illinois Natural History Survey. Occurrence Dataset accessed via GBIF.org on 2018-01-08. doi:10.15468/eol0pe

Dormann CF. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27-46.

Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 1:129-51.

Environmental Systems Research Institute (ESRI). 2017a. ArcGIS 10.5. Redlands, CA.

Environmental Systems Research Institute (ESRI). 2017b. ArcGIS Pro 2.1.0. Redlands, CA.

Fernández-Delgado, M., E. Cernadas, S. Barro, and D. Amorim. 2014. Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? Journal of Machine Learning Research 15:3133–3181.

Fielding, A. H. 2002. What are the appropriate characteristics of an accuracy measure. Pages 271–280 in in J. M. Scott, P. Heglund, M. L. Morrison, J. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington D.C.

Gall, L. 2018. Entomology Division, Yale Peabody Museum. Yale University Peabody Museum. Occurrence Dataset. accessed via GBIF.org on 2018-01-08. doi:10.15468/95waq3

Gama, M., D. Crespo, M. Dolbeth, and P. Anastácio. 2016. Predicting global habitat suitability for Corbicula fluminea using species distribution models: The importance of different environmental datasets. Ecological Modelling 319:163-169.

Genuer, R., J.-M. Poggi, and C. Tuleau-Malot. 2010. Variable selection using random forests. Pattern Recognition Letters 31:2225–2236.

Gregorutti, B., B. Michel, and P. Saint-Pierre. 2016. Correlation and variable importance in random forests. Statistics and Computing:1–20.

Hapfelmeier, A., and K. Ulm. 2013. A new variable selection approach using Random Forests. Computational Statistics & Data Analysis 60:50–69.

Hapfelmeier, A., Hothorn, T., Ulm, K. and C. Strobl. 2014. A new variable importance measure for random forests with missing data. Statistics and Computing 24:21-24.

Harvard University MCZ, P.J. Morris. 2018. Museum of Comparative Zoology, Harvard University. Version 162.81. Museum of Comparative Zoology, Harvard University. Occurrence Dataset https://doi.org/10.15468/p5rupv accessed via GBIF.org on 2018-01-08., doi:10.15468/p5rupv doi:10.15468/p5rupv

Hijmans, RJ & CH Graham. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology* 12(12):2272-2281.

Hijmans, R.J., S.E. Cameron, J.L. Parra, et al. 2005. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25:1965-1978.

Hijmans, R. J., J. van Etten, M. Mattiuzzi, M. Sumner, J. A. Greenberg, O. P. Lamigueiro, A. Bevan, E. B. Racine, and A. Shortridge. 2014. raster: Geographic data analysis and modeling. R package version 2.5-8

Hoffmann, A.A. and C.M. Sgrò, 2011. Climate change and evolutionary adaptation. Nature, 470(7335), p.479.

Howard, T. G., and M. D. Schlesinger. 2013. Wildlife habitat connectivity in the changing climate of New York's Hudson Valley. Annals of the New York Academy of Sciences 1298:103–109.

iNaturalist.org. 2017. iNaturalist Research-grade Observations. Occurrence Dataset https://doi.org/10.15468/ab3s5x accessed via GBIF.org on 2018-01-08. doi:10.15468/ab3s5x

Janzen, D., M. Hajibabaei, J. Burns, W. Hallwachs, E. Remigio, and P. Hebert. 2005. Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. Philosophical Transactions of the Royal Society 360 (1462):1835-1845; doi:10.1098/rstb.2005.1715.

Jones, Z. and Linder, F., 2015, April. Exploratory data analysis using random forests. In Prepared for the 73rd annual MPSA conference.

Keating, R., E. Yerger, and T. Nuttle. 2013. Impacts of Climate Change on Commonly Encountered Forest Lepidoptera of Pennsylvania. Final report to the Pennsylvania Department of Conservation and Natural Resources, Wild Resource Conservation Program (WRCP-010383).

Kharouba, H.M and M. Vellend. 2015. Flowering time of butterfly nectar food plants is more sensitive to temperature than the timing of butterfly adult flight. Journal of Animal Ecology 84(5).

Kincaid, T., and T. Olsen. 2007. Spsurvey: spatial survey design and analysis. R package version 3.3.

Kow, I. et al. 2016. Modeling the status, trends, and impacts of wild bee abundance in the United States. PNAS 113:140-145. Available from: http://www.pnas.org/content/pnas/113/1/140.full.pdf. Accessed 14 February 2018.

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. R News 2:18–22.

Liu, C., P.M. Berry, T.P. Dawson, et al. 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28:385-393.

Menard, K., and P. King. 2018. Recent Invertebrates Specimens. Version 101.55. Sam Noble Oklahoma Museum of Natural History. Occurrence Dataset accessed via GBIF.org on 2018-01-08. doi:10.15468/glxcep

Merow, C., M. Smith, and J. Silander. 2013. Practical guide to MaxEnt for modeling species distributions: what it does and why inputs and settings matter. Ecography 36:1058-1069

Natural History Museum of Utah. 2017. Entomology Collection at the Natural History Museum of Utah. Occurrence Dataset accessed via GBIF.org on 2018-01-08. doi:10.15468/iyehoe

NatureServe. 2002. Element occurrence data standard. In cooperation with the Network of Natural Heritage Programs and Conservation Data Centers. 201 pages.

NatureServe Explorer. 2017. Available from: http://explorer.natureserve.org/.

Northeast Region State Wildlife Action Plan Database. 2017. Developed by Terwilliger Consulting, Inc. with oversight from the Northeast Fish and Wildlife Diversity Technical Committee. Last Updated 14 June 2017. Last Accessed 14 February 2018.

Parmesan, C. 2006. Ecological and Evolutionary Responses to Recent Climate Change. Annual Review of Ecology, Evolution, and Systematics 37:637-669.

Pearson, R.G., and T.P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? Global ecology and biogeography. 12:361-371.

Pearson, R. G., C. J. Raxworthy, M. Nakamura, and A.T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography 34:102–117.

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecol Model 190:231-259.

Phillips, S.J., M. Dudík, and R.E. Schapire. 2017. [Internet] MaxEnt software for modeling species niches and distributions (Version 3.4.1). Available from url: http://biodiversityinformatics.amnh.org/open_source/maxent/. Accessed on 2017-12-20.

Poseldovich, D., T. Toftegaard, C. Wiklund, J. Ehrlén, and L.K. Gotthard. 2015. The developmental race between maturing host plants and their butterfly herbivore – the influence of phenological matching and temperature. Journal of Animal Ecology 84(6).

Poseldovich, D., T. Toftegaard, C. Wiklund, J. Ehrlén, and L.K. Gotthard. 2018. Phenological synchrony between a butterfly and its host plants: Experimental test of effects of spring temperature. Journal of Animal Ecology. 87(1):150-161.

Prasad, A. M., L. R. Iverson, A. Liaw, and A. Liaw. 2006. Newer classification and regression tree techniques: Bagging and Random Forests for ecological prediction.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reece, J.S., R.F. Noss, J. Oetting, T. Hoctor, M. Volk M. 2013. A Vulnerability Assessment of 300 Species in Florida: Threats from Sea Level Rise, Land Use, and Climate Change. PLoS ONE 8(11): e80658. https://doi.org/10.1371/journal.pone.0080658

Roy, D.B. and T.H. Sparks. 2000. Phenology of British butterflies and climate change. Global change biology, 6:407-416.

Royal Belgian Institute of Natural Sciences. 2017. RBINS DaRWIN. Occurrence Dataset accessed via GBIF.org on 2018-01-08. doi:10.15468/qxy4mc

Scholes III, Ph.D. E (2015). Macaulay Library Audio and Video Collection. Cornell Lab of Ornithology. Occurrence Dataset accessed via GBIF.org on 2018-01-08. doi:10.15468/ckcdpy

Schweiger, O., J. Settele, O. Kudrna, S. Klotz, and I. Kühn. 2008. Climate change can cause spatial mismatch of trophically interacting species. Ecology, 89:3472-3479.

Stevens, D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262–278.

University of Alberta Museums. 2017. University of Alberta E. H. Strickland Entomological Museum (UASM). Occurrence Dataset accessed via GBIF.org on 2018-01-08., doi:10.18165/9enbmt

University of Delaware. 2017. University of Delaware Insect Research Collection. Occurrence Dataset accessed via GBIF.org on 2018-01-08. doi:10.15468/1hrpha

Watling, J.I., L.A. Brandt, F.J. Mazzotti, and S.S. Romañach. 2013. Use and Interpretation of Climate Envelope Models: A Practical Guide. University of Florida, 43 pp.

Wickham, H., D. A. James, and S. Falcon. 2014. RSQLite: SQLite Interface for R. R package version 1.0.0.

Willis, SG, JK Hill, CD Thomas, DB Roy, R Fox, DS Blakeley & B Huntley. 2009. Assisted colonization in a changing climate: a test-study using two UK butterflies. *Conservation Letters* 2:45-51.

Xie, Y. 2016. knitr: A general-purpose package for dynamic report generation in R. R package version 1.14.

APPENDIX SUMMARY

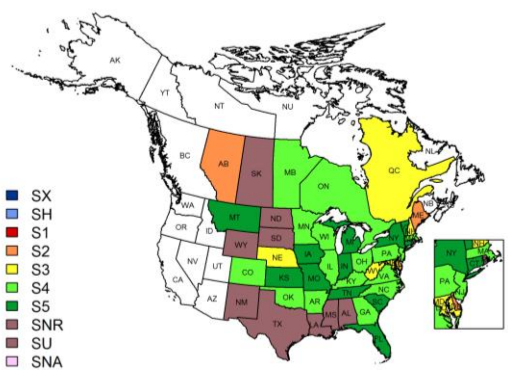
Appendix A. NatureServe range and status maps
Appendix B. Wetland butterfly habitat assessment and wetland butterfly survey forms
Appendix C. Additional field survey information
Appendix D. Species distribution modeling data development guidelines
Appendix E. Species distribution modeling environmental data
Appendix F. Conceptualized description of minimum training presence variants
Appendix G. Summary of important environmental variables
Appendix H. Wissahickon Valley Watershed Association wetland restoration report
Appendix I. Wetland butterfly habitat assessment and enhancement best management practices

STAND ALONE APPENDICES
Appendix J. Species distribution modeling metadata for fourteen wetland butterflies
Appendix K. Climate envelope modeling metadata for fourteen wetland butterflies
Appendix L. Habitat management for pollinators (Pennsylvania)
Appendix M. Life history guide to fourteen rare wetland butterflies in the mid-Atlantic

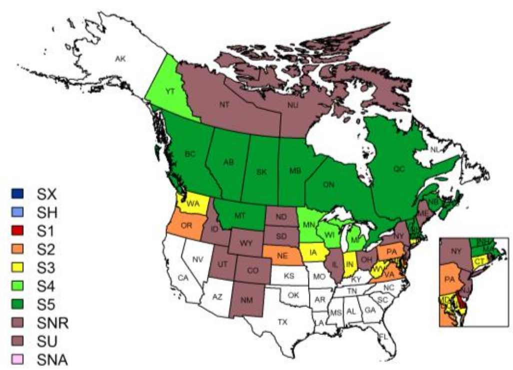
APPENDIX A. NATURESERVE RANGE AND STATUS MAPS

http://explorer.natureserve.org/nsranks.htm

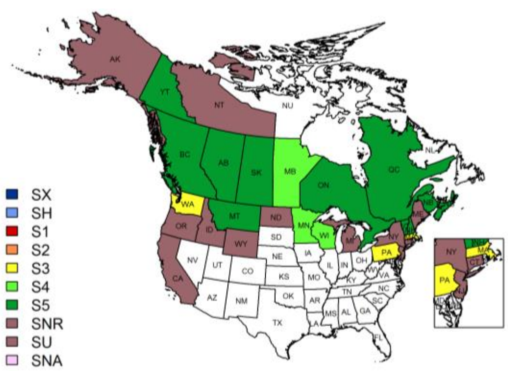
Status	Definition
SX	**Presumed Extirpated**—Species or community is believed to be extirpated from the nation or state/province. Not located despite intensive searches of historical sites and other appropriate habitat, and virtually no likelihood that it will be rediscovered.
SH	**Possibly Extirpated (Historical)**—Species or community occurred historically in the nation or state/province, and there is some possibility that it may be rediscovered. Its presence may not have been verified in the past 20-40 years. A species or community could become SH without such a 20-40 year delay if the only known occurrences in a nation or state/province were destroyed or if it had been extensively and unsuccessfully looked for. The SH rank is reserved for species or communities for which some effort has been made to relocate occurrences, rather than simply using this status for all elements not known from verified extant occurrences.
S1	**Critically Imperiled**—Critically imperiled in the nation or state/province because of extreme rarity (often 5 or fewer occurrences) or because of some factor(s) such as very steep declines making it especially vulnerable to extirpation from the state/province.
S2	**Imperiled**—Imperiled in the nation or state/province because of rarity due to very restricted range, very few populations (often 20 or fewer), steep declines, or other factors making it very vulnerable to extirpation from the nation or state/province.
S3	**Vulnerable**—Vulnerable in the nation or state/province due to a restricted range, relatively few populations (often 80 or fewer), recent and widespread declines, or other factors making it vulnerable to extirpation.
S4	**Apparently Secure**—Uncommon but not rare; some cause for long-term concern due to declines or other factors.
S5	**Secure**—Common, widespread, and abundant in the nation or state/province.
SNR	**Unranked**—State/province conservation status not yet assessed.
SU	**Unrankable**—Currently unrankable due to lack of information or due to substantially conflicting information about status or trends.
SNA	**Not Applicable** —A conservation status rank is not applicable because the species is not a suitable target for conservation activities.
S#S#	**Range Rank** —A numeric range rank (e.g., S2S3) is used to indicate any range of uncertainty about the status of the species or community. Ranges cannot skip more than one rank (e.g., SU is used rather than S1S4).



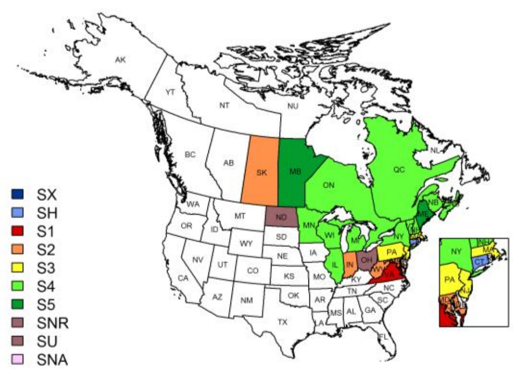
a) *Anatrytone logan*



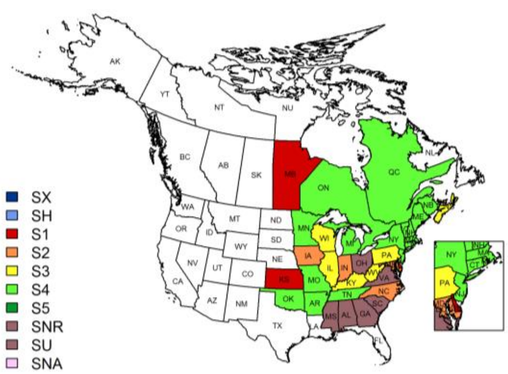
b) *Boloria selene*



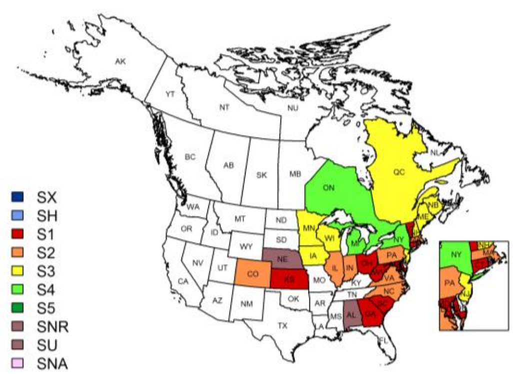
c) *Carterocephalus palaemon*



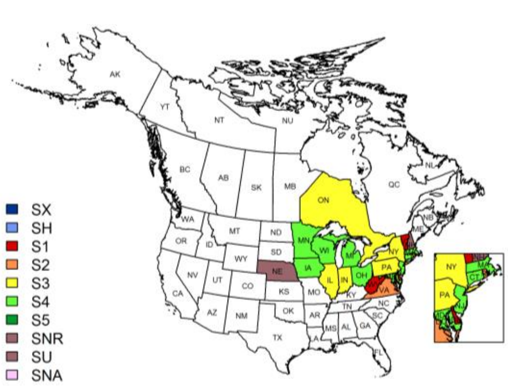
d) *Chlosyne harrisii*



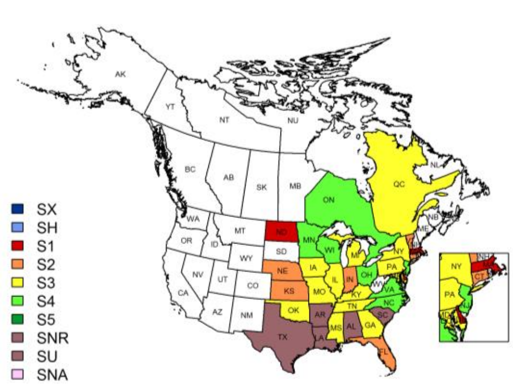
e) *Euphydryas phaeton*



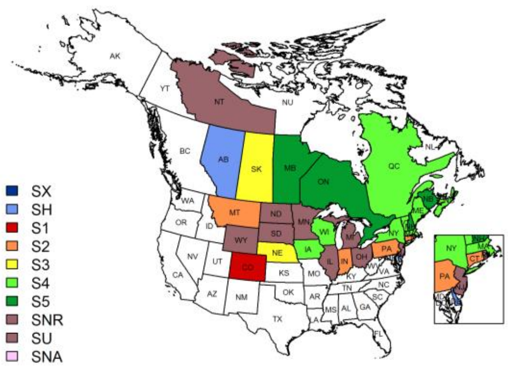
f) *Euphyes bimacula*



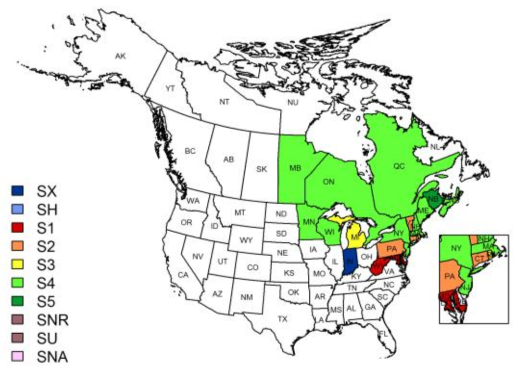
g) *Euphyes conspicua*



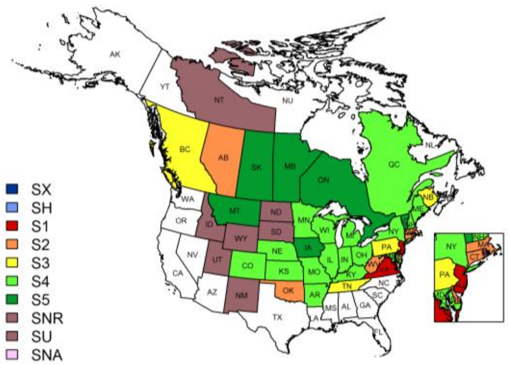
h) *Euphyes dion*



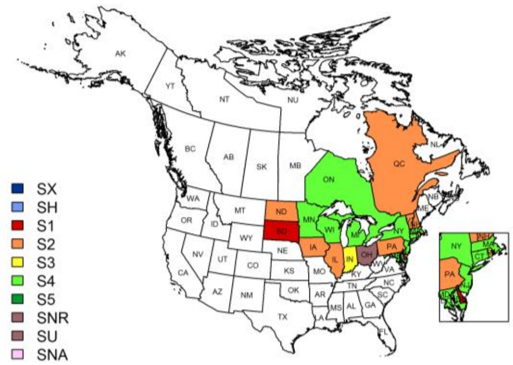
i) *Lethe eurydice*



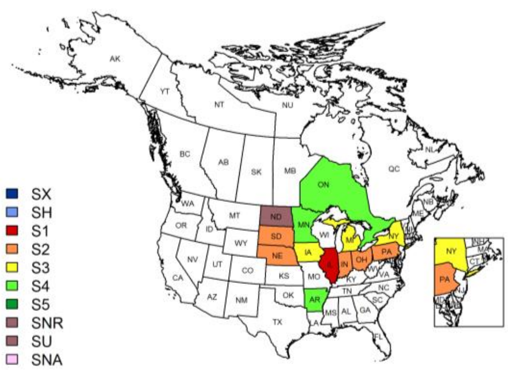
j) *Lycaena epixanthe*



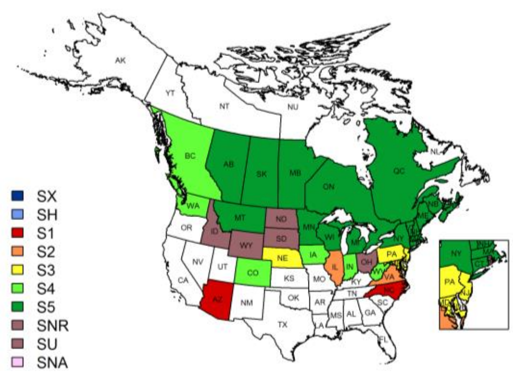
k) *Lycaena hyllus*



l) *Poanes massasoit*



m) *Poanes viator viator*



n) *Polites mystic*

APPENDIX B. WETLAND BUTTERFLY HABITAT ASSESSMENT AND WETLAND BUTTERFLY SURVEY FORMS

WETLAND BUTTERFLY HABITAT ASSESSMENT FORM

Please fill in as much information as possible keeping in mind that accuracy is important. If you are uncertain about a plant species or other data, note that on the form.

Date __________

Examiner Name__________________

SITE NAME:

GENERAL DESCRIPTION: (Describe the ecological and landscape setting)

IS THE SITE IN THE VICINITY OF ANOTHER WETLAND THAT HAS GCN BUTTERFLIES? (Please be specific):
(Less than 2km = 3pts., 2-5km = 2pts., 5+km = 1pt.)

WETLAND SIZE (also note size of adjacent uplands where nectar plants may be present):
(5+acres = 3pts., 2-5 acres = 2pts., less than 2 acres = 1pt.)

SITE OWNERSHIP (If privately owned, does landowner have an easement on the property? Would they consider one?):
(Yes = 2 pts., No = 1 pt.)

WETLAND SOURCE (Is source vulnerable to contamination (i.e. fertilizers), drainage (i.e. new housing developments) or other factors?):
(No contamination concern = 3 pts., moderate or potential concern = 2 pts., significant concern = 0 pts.)

OFFSITE INFLUENCES (Is site well buffered? Does it fall within a natural landscape? Is the surrounding area developed? In agriculture? Specify.):
(Well buffered, greater than 500ft = 3 pts., moderate buffer, up to 500ft = 2 pts., urban landscape, less than 150ft buffer = 0 pts.)

POTENTIAL THREATS (HUMAN-INDUCED) (ATV use, pesticide spraying, etc.):
(No known threats = 3 pts., moderate or potential threats = 2 pts., significant threats = 1 pt.)

POTENTIAL THREATS (NATURAL) (Forest succession, deer browse, high density of invasive plants, etc.):
(No known threats = 3 pts., moderate or potential threats = 2 pts., significant threats = 1 pt.)

GCN BUTTERFLY HOST PLANTS PRESENT? (Include secondary host plants, when applicable. For estimated number of host plants for each species use: (1) <10 stems, (2) <50 stems; (3) <100 stems; (4) hundreds of stems; or (5) thousands of stems).
(Points equal category numbers 1-5)

NECTAR PLANTS PRESENT? (For estimated number of nectar plants for each species use: (1) <10 stems, (2) <50 stems; (3) <100 stems; (4) hundreds of stems; or (5) thousands of stems).
(Points equal category numbers 1-5)

INVASIVE PLANTS PRESENT? For estimated number of invasive plants for each species use: (5) <10 stems, (4) <50 stems; (3) <100 stems; (2) hundreds of stems; or (1) thousands of stems).
(Points equal category numbers 1-5)

pH LEVEL:
(should be close to neutral or slightly acidic)

LIGHT LEVELS: (Full sun, partial shade, filtered or dappled sunlight, closed canopy, etc.)
(Full-Partial Sun = 2 pts., mostly shaded = 1 pt., closed canopy = 0 pts.)

SITE CONDITION: (Land use history, anthropogenic disturbance, exotic species, alterations of natural processes, etc.).
(Excellent = 3 pts., Good = 2 pts., Poor = 1 pt.; there is some subjectivity in this scoring so please justify your response with an explanation).

RESTORATION OR MANAGEMENT NEEDS:
(Minor = 6pts., Moderate = 3 pts., Major = 0 pt.)

WETLAND BUTTERFLY SURVEYS - DATA COLLECTION FIELD FORM

Please fill in as much information as possible keeping in mind that accuracy is important. If you are uncertain about a butterfly species or other data, note that on the form.

STATE ________________DATE _____________________ TIME (Start) _______________ (End) _____________

Observer(s) Name____________________________________ County__________________________________

SITE NAME AND LOCATION:

ESTIMATED SURVEY AREA: (Approximate size of the property or wetland that was surveyed, e.g. walked through most of it, explored only one section, etc.):

WEATHER CONDITIONS:
(Temperature, wind, percent cloud cover, precipitation):

SPECIES OBSERVED (ADULT SURVEYS):
(Include number of adults, condition, distribution within site (concentrated, widespread, sparse), behavior, etc. If nectaring, note plant species):

Species & Voucher Number	# Obs	Condition	Distribution	Behavior	Nectar Plant

If no or few individuals observed, explain (e.g., clouds moved in, area recently burned, no obvious explanation):

Note if any specimens were collected or photographed (both can be useful if ID is uncertain) – use WV butterfly Atlas protocol.

SPECIES INFORMATION (LARVAE)
(Number of larvae or webs observed, plants they were observed on, instars, distribution):

Species & Voucher Number	# Obs	Host Plant	Distribution

Host Plant availability and distribution (abundant, common, spotty, concentrated in one area):

Other Observations:

IMMEDIATE THREAT OR MANAGEMENT NEEDS (e.g., extreme deer browse of host plants, fallen tree creating a dangerous situation, outbreak of invasive plants that you may not have noticed before, ATV damage, etc.):

APPENDIX C. ADDITIONAL FIELD SURVEY INFORMATION

The following is a list of Pennsylvania sites visited for wetland butterfly surveys, and the criteria that were used to select each site.

1. Acopian Preserve, Lancaster County, 2 visits [2]
2. Algerine Swamp, Tioga / Lycoming Counties [1]
3. Allegheny National Forest near Buzzard Swamp, Forest County [2]
4. Asaph Swamp, Tioga County [1]
5. Bear Creek Preserve, Natural Lands Trust, Luzerne County, 2 visits [2]
6. Bear Meadows Natural Area, Centre County, 3 visits [1]
7. Beaver Meadows, Forest County [2]
8. Blacklog Creek, Tuscarora State Forest, Juniata County [3]
9. Bruce Lake Natural Area, Pike County [1]
10. Black Ash Swamp, Tioga County [2]
11. Black Moshannon State Park, Centre County [1]
12. Chalk Hill Bog, Fayette County [1]
13. Chartiers Creek, Crawford County [2]
14. Christner Bog, Somerset County [1]
15. Crow's Nest, Natural Lands Trust, Chester County [2]
16. East Licking Creek, Tuscarora State Forest, Juniata County [3]
17. Fern Ridge Bog, Monroe County [1]
18. Fulshaw Craeg, Natural Lands Trust, Montgomery County [2]
19. Green Hills, Natural Lands Trust, Berks County [2]
20. Hells Half Acre, Erie County [1]
21. Hesselgessel Swamp, Tioga County [2]
22. Kerney Property, Adams County [2]
23. Lackawanna State Park, Trostler Tract, Lackawanna County [2]
24. Lake Lacawac, Wayne County [1]
25. Lake Leigh, Ricketts Glen State Park, Sullivan County [1]
26. Laurels Preserves, Brandywine Conservancy, Chester County [2]
27. Lowville Fen, Erie County [2]
28. Muddy Creek, Erie National Wildlife Refuge, Crawford County, 2 visits [1]
29. Reynolds Spring Natural Area, Tioga County [2]
30. Ricketts Glen State Park, The Hayfields, Sullivan County, 2 visits [2]
31. Ricketts Glen State Park, Big Run Headwaters, Sullivan County [1]
32. Two Mile Run, Monroe County [1]
33. Pine Swamp, Mercer County [1]
34. Scotch Pine Hollow, Tioga County [2]
35. Spruce Flats Bog, Westmoreland County [1]
36. Stroud, Natural Lands Trust, Chester County [2]
37. Swamp Branch Swamp, Clinton County [1]
38. Tamarack Run, Sullivan County [1]
39. Tamarack Swamp, Clinton County [1]
40. The Hook, Centre County [1]
41. The Meadows, State Game Lands 13/57, Luzerne County [2]
42. Titus Bog, Erie County, 2 visits [1]
43. Toplovich Bog, Warren County [1]
44. Treaster Kettle, Rothrock State Forest, Centre County [3]
45. Waldman Property, York County, 2 visits [2]
46. Waterloo, Brandywine Conservancy, Chester/Delaware Counties [2]

47. Wattsburg Fen, Erie County [1]

[1] High elevation peatland monitoring sites
[2] Wetland butterfly site assessments
[3] PA Bureau of Forestry habitat management monitoring

The following is a list of Pennsylvania wetland butterfly records added or updated in 2016-2017. All are new to the Pennsylvania heritage database except for one PNHP Baltimore Checkerspot record and one PNHP Bog Copper record. Thirty-five occurrences were found during field surveys by PNHP staff, and twenty-one were gathered from other sources.

Common Name	**Scientific Name**	**PNHP Staff**	**Other Sources**	**Total**
Delaware Skipper	*Anatrytone logan*	4	4	8
Silver-bordered Fritillary	*Boloria selene*	2	0	2
Arctic Skipper	*Carterocephalus palaemon*	1	0	1
Harris' Checkerspot	*Chlosyne harrisii*	3	0	3
Baltimore Checkerspot	*Euphydryas phaeton*	3	6	9
Two-spotted Skipper	*Euphyes bimacula*	3	1	4
Black Dash	*Euphyes conspicua*	2	4	6
Dion Skipper	*Euphyes dion*	1	0	1
Eyed Brown	*Lethe eurydice*	2	0	2
Bog Copper	*Lycaena epixanthe*	4	1	5
Bronze Copper	*Lycaena hyllus*	0	2	2
Mulberry Wing	*Poanes massasoit*	0	0	0
Broad-winged Skipper	*Poanes viator viator*	2	0	2
Long Dash	*Polites mystic*	8	3	11
	TOTAL	**35**	**21**	**56**

APPENDIX D. GUIDELINES FOR PREPARING EO DATA FOR SPECIES DISTRIBUTION MODELS

Guidelines for preparing EO data for Species Distribution Models for the Mid-Atlantic Wetland Butterfly Modeling Project

Note: These guidelines are based on the "Version 3" guidelines developed by eastern region Natural Heritage programs (VNHP, NYNHP, PNHP, and FNAI)

Introduction

The use of high quality occurrence data in the models is of utmost importance. The ultimate goal is to have polygons that represent only the extant, appropriate *habitat* for that species, relying on the best judgment of the biologist doing the review. Thus we need to review, cull, and edit the data as best as possible. The following are guidelines to help think through the preparation of the data so that multiple reviewers can, as much as possible, be following the same guidance.

What we need are current EO spatial data, as polygon features that:

- Include suitable habitat in the immediate area of an observation (i.e. EO). This needs to be 'tightly' drawn to minimize unsuitable habitat, but to still include likely suitable habitat.
- Exclude unsuitable habitat in the immediate observed area such as roads, buildings, etc.
- Have a reasonable expectation of the species being present there. That is, historical observation might be usable, if the habitat is still suitable in your expert opinion.

Overall Guidance

<u>F, H, and X ranked EOs</u> – We are trying to balance the need for have high-quality occurrence polygons with the desire to have sufficient distributions of known locations. That is, if we discard too many polygons because they are not the highest quality (e.g., very old or low RA) the model may not predict areas where it once occurred. Thus, the fewer polygons in an area of the species' range, the higher the tolerance for older observations and for lower accuracy of those observations. Therefore, we should not exclude, by default, F, H, and X ranked EOs. We must think more about whether the current habitat is similar to what it was at the time of the observation, based on that assessment. F, H, and X ranked EOs may often be removed, but there will be exceptions. If it is a precise EO and the habitat appears unchanged, there's not much reason to say that the species is gone (or the habitat has changed) if only time has passed.

<u>Likely Suitable Habitat</u> – With regard to 'likely suitable habitat' and inferred extent, a review of the occurrence polygons to determine if additional suitable habitat can be added, is needed. Remember, this is a habitat suitability model, and thus we are really asking for polygons inclusive of species locations *and associated habitats*. The model will populate these polygons with training points. The larger the polygon, the more training points will be used; however it will cut-off at 400 points which is a polygon about 40 acres. Polygons larger than 40 acres will have the 400 points more spread-out. If the occurrence polygon is small, less than 0.1 acre, then please take time to review it to see if additional suitable habitat can be added. If the occurrence polygon is greater than 40 acres, this review step is not as necessary, but it could improve the model to capture additional habitat, particularly if it is clearly suitable (i.e. extending to include an entire wetland). *See below – 'A little more about editing'.*

<u>Representational Accuracy</u> – Representational accuracy (RA) is a qualitative estimation of the amount of location error a spatial representation might contain. So a GPS point right on top of a single plant, or walked around the exact outer boundary of a population (no matter how large) would both have negligible LU and RA = Very High. Directions from an old museum specimen that are vague may lead to a large mapped polygon encompassing

1,000s of Ha – this would be Very Low RA. Given that larger polygons receive more training points, yet our Very High RA (i.e. best mapped, least amount of uncertainty) occurrences tend to be very small (typically GPSed point locations with negligible uncertainty, therefore a 9m diameter circle) we are working to develop a way to weight the training points according to the RA assigned to the SF (not the EO). In many cases, this may have to be assigned by the reviewer. A SFRA, should be assigned prior to making any edits and changes to the polygon size/quality should be assessed based in part on taxa/mobility. *See below – Assigning RA to training polygons*.

<u>Overlapping Polygons</u> – If there are multiple overlapping polygons for a single species, these will need to be dissolved into one polygon and edited from there. Be careful to not merge polygons with multiple RAs. RAs will be used to weight the training points, thus donuts within the largest polygon should be created for smaller, intersecting polygons with different RAs to maintain their RA value.

Examples

The following are some examples that we anticipate, which will require some pre-model development data clean up. Staff at the source Programs for these data have the most expertise about these species and habitat data. If source Programs were not able to conduct this data clean-up prior to sharing data, we can do so, but seek guidance as needed:

- A bog turtle is found along the edge of a larger wetland but you know it occurs throughout. If the EO does not already reflect inferred extent, please include the entire wetland in the polygon
- A bog turtle is mapped on a road (if that is EO-worthy). We would not want the model to pick up roadways as suitable habitat, so eliminating that polygon would be warranted. If there is a wetland immediately adjacent and in your expert opinion the turtle uses that wetland, please map the wetland instead.
- Sensitive joint-vetch is mapped in a marsh. Based on knowledge about the patchy distribution of this species, inferring the entire marsh may not be a good idea, so keep the polygons as mapped.
- A point location is given for a southern hognose snake located in a very large area of suitable habitat. Buffer the point out to 100m diameter (about 0.78Ha) being careful that the habitat captured is suitable.
- You have several source features for a single EO. One SF is very old and you don't think it is suitable habitat any more – that polygon should be removed. (See Key)
- Any mapped PFs with RA of 'Very Low' should be eliminated (this should be accomplished by the query, if used). Any with 'Low' should be reviewed. If it is possible to discern the very best potential habitat in that polygon, only that portion should used. If it is not discernible because it is all appropriate habitat, then leave as is. If it is not discernible because it is all questionable habitat, then remove it. (See Key)
- Any obvious unsuitable habitat should be edited out of the final polygons. Examples: roads, rivers (if terrestrial species), buildings, golf courses, etc. (see * note on Key)

Key of suggested guidelines for keeping/editing EOs for species distribution models (SDM)

We are trying to balance the desire to have high-quality occurrence polygons with the desire to have sufficient distributions of known locations. That is, if we discard too many polygons b/c they are not the highest quality (very old, or low RA) the model may not predict areas where it once occurred. Thus, the fewer polygons in an area of the species' range, the higher the tolerance for older observations and for lower accuracy of those observations.

The following key is to be used to help make decisions regarding when to keep and when to discard occurrence polygons for used in SDM. Biological expertise may override these suggestions. Note, after the first couplet evaluating EO Rank, all decisions should be made for each polygon comprising an EO (i.e., the procedural feature or PF).

1a.	EO Rank is A, B, C, D, or E	Edit*
1b.	EO Rank is F, H, or X	2
2a.	Ten or fewer PFs in your state	3
2b.	More than ten PFs in your state	4
3a.	Last observation of PF was before 1940	Discard
3b.	Last observation of PF made in 1940 to present..	5
4a.	Last observation of PF was before 1955	Discard
4b.	Last observation of PF made in 1955 to present	5
5a.	Representational Accuracy of PF = Very Low	Discard
5b.	Representational Accuracy of PF = Low or better	6
6a.	Representational Accuracy of PF = Medium or better..	Edit*
6b.	Representational Accuracy of PF = Low	7
7a.	Suitable habitat is present...	Edit*
7b.	Suitable habitat is not present	Discard

* Edit each PF to remove roads, buildings, etc., and unsuitable habitat patches larger than a 60m square. You may also determine that you can include more habitat.

A little more about editing

We are creating polygons for *training data* - not EOs (though some of our resulting polygons might be better than EO as mapped)

EOs inherently have added buffer already (LU) so with the exception of the negligible LU, a buffer is already capturing some habitat where the EO was not actually observed (and does so indiscriminately).

EOs are mapped in different manners. A biologist may have placed a point on a map as a 'seconds' record, or provided coordinates, or drawn a polygon using ArcGIS etc. Or, someone interpreted written directions from a label on a museum specimen/literature record and mapped it, hopefully, in a biologically meaningful way. Or taken coordinates from non-Heritage source and mapped. The point is this: the manner in which a record was mapped *may* have introduced poor habitat information. Always – try to understand how it got mapped (added buffer etc.), the information about the habitat given and make decisions from there.

One of the bigger decisions of the editing process is when to add additional habitat. After much discussion, and determining that the RA weighting process will help highlight the importance of the best mapped locations, we have settled on the following key:

Key when considering adding inferred habitat

This key treats EO features differently depending on the type of observation. It should be used with the original key whenever it advises to edit, particularly adding habitat. This key gives guidance about whether or not it is appropriate to add inferred habitat. It is a given that you would always remove unsuitable habitat. Note, this was handled differently for the amphibians with discrete habitat patches and additional habitat added in buffers – RA was assigned after editing – see below.

1a.	Conceptual feature is a polygon	Do not add additional habitat.
1b.	Conceptual feature is a point or line	2
2a.	Species is a plant or sedentary animal	Do not add additional habitat.
2b.	Species is a vagile animal#	3
3a.	Locational uncertainty type is negligible or	Buffer by [X*m minus procedural buffer]; judiciously add habitat within buffer zone only.
3b.	Locational uncertainty type is delimited or estimated	4
4a.	LU distance is greater than or equal to 50-m	Do not add additional habitat.
4b.	LU distance is less than 50-m	Buffer by [X*m minus LU distance]; judiciously add habitat within buffer zone only.

For a vagile animal where the habitat is usually discernable from an aerial image (e.g. a wetland), it is ok to capture that entire habitat.
X* - buffer to be determined by species. Programs are talking to their biologists about this.
We don't have conceptual feature and locational uncertainty for all data sets we have received. If poly is 9m diameter circle, might be able to assume point/negligible. Otherwise, assume polygon/delimited.

Assigning RA to training polygons

We will be trying to use Representational Accuracy (RA) for each polygon (SF) as a way to weight the model. Thus each *polygon* needs to have an RA assigned to it.

RA is a qualitative estimation of the amount of location error a spatial representation might contain. So a GPS point right on top of a single plant, or walked around the exact outer boundary of a population (no matter how large) would both have negligible LU and RA = Very High. Directions from an old museum specimen that are vague may lead to a large mapped polygon encompassing 1,000s of Ha – this would be Very Low RA. In loose terms, polygons with Very High RA are our best mapped occurrences, i.e. with the least uncertainty attached.

Note an RA can be assigned by a NHP at both the EO and the SF level. Where an EO is multi-part an 'average' of the SFRA is assigned. Thus, when SFRA is not assigned, assuming EORA is equal is not always the case.

A SFRA should be calculated prior to making any edits. Calculating SFRA for amphibians was handled differently. See below.

There are two (subtly) different editing procedures that may occur: 1) removing unsuitable habitat such as roads or development etc., and 2) keeping suitable habitat/improving the mapping, such as isolating the wetland habitat for a salamander.

The first scenario, removing unsuitable habitat, is likely being done on upland plants and animal species, where looking at aerial photography may not definitively inform better on-the-ground habitat. In this case, use the following suggestions for assigning RA to the polygon:

- If a *SF* RA has been assigned by the data source, use that regardless of edits made.
- If an RA is only assigned at the EO level, and the EO is a single feature then = SFRA regardless of edits made.
- If there is an EORA available and it is a multi-part EO, use the RA KEY* to assign SFRA prior to making edits.
- If no EORA, use the RA Key* to assign SFRA prior to making edits.

The thinking is this: if prior to any edits the RA is Medium, then no amount of removing roads/buildings etc., or resulting numbers of polygons, will imply better mapping of the occurrence and hence the uncertainty remains the same.

The second scenario, keeping suitable habitat/improving the mapping, is likely being done for animals or plants that feature discrete habitat patches that can (typically) be more clearly defined from aerial imagery. In this case, we can often improve the spatial representation of large aerial-estimated polygons (circles) by removing clearly unsuitable habitat and in some cases identifying the specific habitat patch noted by the source documentation. Thus an old “minutes precision” circle polygon for larvae of a pond-breeding amphibian, which captures both upland forest and an isolated pond, can be mapped better by removing the upland forest. In this case, the resulting RA may actually be higher than the original polygon. In other cases, multiple suitable habitat patches may be present within or very near the original polygon. In this case, keeping all patches is recommended but RA may not be improved.

Guidelines for Determining Representation Accuracy (RA), in Dichotomous Key Form (NatureServe)

1. *Locational Uncertainty Type* is **Negligible**...RA is **Very high**
1. *Locational Uncertainty Type* is **Estimated, Delimited or Linear**...2
    2. Source Feature is 1 hectare (ca. 2.5 acres) or less...RA is **High**
    2. Source Feature is larger than 1 hectare...3
        3. *Conceptual Feature Type* is **Point**...4
            4. Source Feature is 50 hectares (ca. 125 acres) or less...RA is **Medium**
            4. Source Feature is larger than 50 hectares...5
                5. Source Feature is 2500 hectares (ca. 6178 acres) or less...RA is **Low**
                5. Source Feature is larger than 2500 hectares...RA is **Very Low**
        3. *Conceptual Feature Type* is **Line** or **Polygon**...6
            6. More than 80% of the Source Feature is comprised of the observed area (i.e., 20% or less of the Source Feature is comprised of area added for locational uncertainty) or the Source Feature is 1 hectare or less...RA is **High**
            6. 80% or less of the Source Feature is comprised of the observed area (i.e., 20% or more of the Source Feature is comprised of area added for locational uncertainty) or the observed area is unknown (Source Feature is greater than 1 hectare)...7
                7. 20% or more of the Source Feature is comprised of the observed area (i.e., less than 80% of the Source Feature is comprised of area added for locational uncertainty)...RA is **Medium**
                7. Less than 20% of the Source Feature is comprised of the observed area (i.e., 80% or more of the Source Feature is comprised of area added for locational uncertainty) or the observed area is unknown...8
                    8. Source Feature is 50 hectares (ca. 125 acres) or less.....................RA is **Medium**
                    8. Source Feature is larger than 50 hectares...9
                        9. Source Feature is 2500 hectares (ca. 6178 acres) or less...........RA is **Low**
                        9. Source Feature is larger than 2500 hectares......................RA is **Very Low**

This key was developed to assess RA for individual observations, i.e., Source Features. However, because Element Occurrences are utilized in conservation planning and analyses, an overall RA value is needed for each EO:

- When the EO is comprised of a single Source Feature, the RA value for the EO would equal the RA assigned to the underlying Source Feature;
- When the EO is comprised of more than one Source Feature, the RA of the EO would be developed using the RA values assigned to each of its component Source Features. In this case, a comment should be entered in the RA Comments field of the EO record describing how the overall RA value was derived.
- When the RA values assigned to individual Source Features that comprise a single EO differ significantly (e.g., High RA versus Low), as would be the case when a historical observation with a large amount of associated locational uncertainty is combined with more recent observations with much less associated uncertainty, consider treating the historical observation as a separate principal EO. In such cases, use the Separation Comments field to explain the rationale for creating separate principal EOs from observations that have been grouped into a single EO according to the EO specifications for the Element.

Comments relating to modeling work:

- Each SF needs an RA.
- When CF and LU are available run the key. If you get past couplet 6, you may have to assume observed area is unknown which yields: <50Ha = medium, 50-2500 = Low and >2500 = Very Low.

- When CF and LU are not available, assume CF = polygon and LU = delimited. This can lead to ‘high’ when should be ‘very high’ in some cases. But overall is a conservative approach. Exception: if the polygon is a 9m diameter circle, assume CF = point and LU = negligible. A-18
- In cases where a polygon with a higher RA intersects or is completely within another polygon with a lower RA, create a ‘donut’ in the lower RA poly where the higher RA poly intersects. If the higher RA subsumes the lower RA, no donut is needed.

APPENDIX E. SDM ENVIRONMENTAL DATA

This data was initially created for the USFWS Region 5 (Chazal et al. 2017b) and SALCC (Chazal et al. 2017a) projects. Detailed information regarding the methods used to produce each layer is available upon request. To ensure consistency among raster datasets prepared by different partners, we established a set of specifications that had to be met. It was critical for the modeling inputs to be in a common coordinate system and to have the exact same spatial extent, pixel size, and pixel alignment. We used the Albers Conical Equal Area coordinate system, and a pixel size of 30m, consistent and perfectly aligned with many source datasets available for the continental United States, such as the National Land Cover Database. We created a designated template raster, available to all partners, which was used as a snap raster to set the pixel alignment, and as a mask to set the spatial extent and zone of analysis. Rasters were saved in the geo-referenced Tag Image File Format (GeoTIFF).

The original source data and resulting derived environmental variables extended at least 5 km beyond the study area boundaries. This was to avoid introducing any boundary effects or artifacts in rasters developed using flowpath distance or, Euclidean distance to polygons, or any kind of neighborhood analysis. Prior to use in the modeling process, rasters were clipped to the boundary of the study. Typically, initial raster outputs were in floating point format. To greatly reduce raster storage size on disk, all rasters were converted to integer format, after multiplying by a suitable factor to maintain precision.

Category	Raster Name	Description	Source[1]
Climate – Precipitation	JulyPrecip	July precipitation	PRISM
Climate – Precipitation	JunePrecip	June precipitation	PRISM
Climate – Precipitation	MayPrecip	May precipitation	PRISM
Climate – Precipitation	NrmDspPrcp	Normalized dispersion (CV) of precipitation	PRISM
Climate – Precipitation	PrcpCldQtr	Precipitation of coldest quarter	PRISM
Climate – Precipitation	PrcpDryMth	Precipitation of driest month	PRISM
Climate – Precipitation	PrcpDryQtr	Precipitation of driest quarter	PRISM
Climate – Precipitation	PrcpWetMth	Precipitation of wettest month	PRISM
Climate – Precipitation	PrcpWetQtr	Precipitation of wettest quarter	PRISM
Climate – Precipitation	PrcpWrmQtr	Precipitation of warmest quarter	PRISM
Climate – Precipitation	TtlAnnPrcp	Total annual precipitation	PRISM
Climate - Temperature	AnnMnTemp	Annual mean temperature	PRISM
Climate – Temperature	gddays	Growing degree days	PRISM
Climate – Temperature	Isotherm	Comparison of day-to-night and summer-to-winter temperature oscillations	PRISM
Climate – Temperature	MnDiurnRng	Mean diurnal temperature range (mean of the differences between monthly maximum and monthly minimum temperatures	PRISM
Climate – Temperature	MnTpCldMth	Minimum temperature of coldest month	PRISM
Climate – Temperature	MnTpCldQtr	Mean temperature of coldest quarter	PRISM
Climate – Temperature	MnTpDryQtr	Mean temperature of driest quarter	PRISM
Climate – Temperature	MnTpWetQtr	Mean temperature of wettest quarter	PRISM
Climate – Temperature	MnTpWrmQtr	Mean temperature of warmest quarter	PRISM
Climate – Temperature	MxTpWrmMth	Maximum temperature of warmest month	PRISM
Climate – Temperature	TempAnnRng	Annual temperature range (difference between the maximum temperature of the warmest month and the minimum temperature of the coldest month)	PRISM
Climate – Temperature	TempSeason	Temperature seasonality (standard deviation of temperatures)	PRISM

Elevation and Derivatives	beersx1000	Beers et al. (1966) transformation of aspect (slope direction). The scale ranges from 0 (most exposed, southwest-facing slopes) to 2 (most sheltered, northeast-facing slopes), with values grading equivalently in both directions between the extremes.	USGS NED
Elevation and Derivatives	crvplax100	Plan curvature (the curvature of the slope as fitted through the focal cell and its neighbors perpendicular to the direction of the maximum slope)	USGS NED
Elevation and Derivatives	crvprox100	Profile curvature (the curvature of the slope as fitted through the focal cell and its neighbors in the direction of the maximum slope)	USGS NED
Elevation and Derivatives	crvslpx100	Slope curvature (the curvature of the slope as fitted through the focal cell and its neighbors)	USGS NED
Elevation and Derivatives	elevx10	Elevation	USGS NED
Elevation and Derivatives	radequinx	Total insolation derived from direct and diffuse, but not reflected, radiation for the equinox	USGS NED
Elevation and Derivatives	radsumsol	Total insolation derived from direct and diffuse, but not reflected, radiation for the summer solstice	USGS NED
Elevation and Derivatives	radwinsol	Total insolation derived from direct and diffuse, but not reflected, radiation for the winter solstice	USGS NED
Elevation and Derivatives	rgh100x100	Roughness within 100-cell radius (the standard deviation of elevation values within a circular neighborhood with a radius of 100 cells.	USGS NED
Elevation and Derivatives	rgh1cx100	Roughness within 1-cell radius (the standard deviation of elevation values within the neighborhood immediately surrounding the focal cell (9-cell square))	USGS NED
Elevation and Derivatives	slopex100	Slope in degrees.	USGS NED
Elevation and Derivatives	tp001x1000	Topographic position index within 1-cell radius (difference between the elevation of the focal cell and the mean elevation within the immediate neighborhood (9-cell square))	USGS NED
Elevation and Derivatives	tp010x1000	Topographic position index within 10-cell radius (difference between the elevation of the focal cell and the mean elevation within a radius of 10 cells	USGS NED
Elevation and Derivatives	tp100x1000	Topographic position index within 100-cell radius (difference between the elevation of the focal cell and the mean elevation within a radius of 100 cells	USGS NED
Elevation and Derivatives	rgh10cx100	Roughness within 10-cell radius (the standard deviation of elevation values within a circular neighborhood with a radius of 10 cells.	USGS NED
Geology	geo001	Euclidean distance to sand	SoilGeo-TNC
Geology	geo002	Euclidean distance to loam	SoilGeo-TNC
Geology	geo003	Euclidean distance to silt/clay	SoilGeo-TNC
Geology	geo031	Euclidean distance to coastal plain sand over limestone	SoilGeo-TNC
Geology	geo032	Euclidean distance to coastal plain loam over limestone	SoilGeo-TNC
Geology	geo033	Euclidean distance to coastal plain silt and clay over limestone	SoilGeo-TNC
Geology	geo100	Euclidean distance to acidic sedimentary bedrock	SoilGeo-TNC
Geology	geo200	Euclidean distance to acidic shale bedrock	SoilGeo-TNC
Geology	geo300	Euclidean distance to calcareous bedrock	SoilGeo-TNC
Geology	geo400	Euclidean distance to moderately calcareous bedrock	SoilGeo-TNC
Geology	geo500	Euclidean distance to acidic granitic bedrock	SoilGeo-TNC
Geology	geo600	Euclidean distance to mafic bedrock	SoilGeo-TNC

Geology	geo700	Euclidean distance to ultramafic bedrock	SoilGeo-TNC
Hydrography	distcstwat	Euclidean distance to nearest estuary or sea/ocean	NHD; NLCD; USGS NED
Hydrography	distestury	Euclidean distance to nearest estuary	EPA/USGS NHD+V2; NHDH
Hydrography	distinlwat	Euclidean distance to nearest stream, river, or other inland waterbody (excluding estuaries)	NHD; NLCD; USGS NED
Hydrography	distlake	Euclidean distance to nearest NHD lake/pond/reservoir > 1 ha	NHD; NLCD; USGS NED
Hydrography	distocean	Euclidean distance to nearest sea/ocean	NHD; NLCD; USGS NED
Hydrography	distpond	Euclidean distance to nearest NHD lake/pond/reservoir <= 1 ha	NHD; NLCD; USGS NED
Hydrography	distriver	Euclidean distance to nearest NHD stream/river feature represented by polygons	NHD; NLCD; USGS NED
Hydrography	diststrm	Euclidean distance to nearest NHD stream feature represented by lines	NHD; NLCD; USGS NED
Hydrography	downdist	The downslope distance along the flow path to a water or wetland feature.	EPA/USGS NHD+V2; NHDH, NED, NLCD
Hydrography	flowacc	Flow accumulation is used as a proxy for topographic moisture. For each cell, this is determined by summing the weights of all cells flowing into it. This does not account for flow differences over different soil types.	EPA/USGS NHD+V2; NHDH, NED, NLCD
Land Cover - NLCD	canopy1	Mean percent canopy cover in 1-cell radius	NLCD 2011
Land Cover - NLCD	canopy10	Mean percent canopy cover in 10-cell radius	NLCD 2011
Land Cover - NLCD	canopy100	Mean percent canopy cover in 100-cell radius	NLCD 2011
Land Cover - NLCD	impsur1	Mean percent impervious cover in 1-cell radius	NLCD 2011
Land Cover - NLCD	impsur10	Mean percent impervious cover in 10-cell radius	NLCD 2011
Land Cover - NLCD	impsur100	Mean percent impervious cover in 100-cell radius	NLCD 2011
Land Cover - NLCD	nlcddfr1	Percent deciduous forest cover within 1-cell radius	NLCD 2011
Land Cover - NLCD	nlcddfr10	Percent deciduous forest cover within 10-cell radius	NLCD 2011
Land Cover - NLCD	nlcddfr100	Percent deciduous forest cover within 100-cell radius	NLCD 2011
Land Cover - NLCD	nlcdefr1	Percent evergreen forest cover within 1-cell radius	NLCD 2011
Land Cover - NLCD	nlcdefr10	Percent evergreen forest cover within 10-cell radius	NLCD 2011
Land Cover - NLCD	nlcdefr100	Percent evergreen forest cover within 100-cell radius	NLCD 2011
Land Cover - NLCD	nlcdopn1	Percent open cover within 1-cell radius	NLCD 2011
Land Cover - NLCD	nlcdopn10	Percent open cover within 10-cell radius	NLCD 2011
Land Cover - NLCD	nlcdopn100	Percent open cover within 100 cell radius	NLCD 2011
Land Cover - NLCD	nlcdshb1	Percent shrub cover within 1-cell radius	NLCD 2011
Land Cover - NLCD	nlcdshb10	Percent shrub cover within 10-cell radius	NLCD 2011
Land Cover - NLCD	nlcdshb100	Percent shrub cover within 100 cell radius	NLCD 2011
Land Cover - NLCD	nlcdwat1	Percent open water cover within 1-cell radius	NLCD 2011
Land Cover - NLCD	nlcdwat10	Percent open water cover within 10-cell radius	NLCD 2011
Land Cover - NLCD	nlcdwat100	Percent open water cover within 100 cell radius	NLCD 2011
Land Cover - NLCD	nlcdwet1	Percent wetland cover within 1-cell radius	NLCD 2011
Land Cover - NLCD	nlcdwet10	Percent wetland cover within 10-cell radius	NLCD 2011
Land Cover - NLCD	nlcdwet100	Percent wetland cover within 100 cell radius	NLCD 2011
Land Cover - NWI	dnwifemw	Euclidean distance to freshwater emergent wetland	NWI
Land Cover - NWI	dnwiffw	Euclidean distance to forested palustrine wetland	NWI
Land Cover - NWI	dnwisemw	Euclidean distance to saltwater emergent wetland	NWI

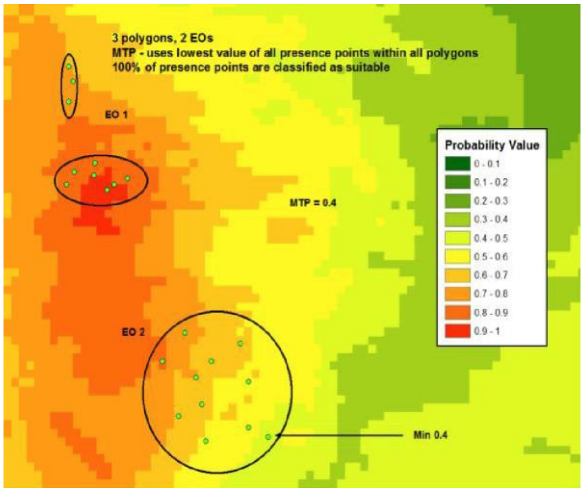
1 = PRISM = PRISM climate data (http://www.prism.oregonstate.edu/); NED = USGS National Elevation Dataset (https://nationalmap.gov/elevation.html); SoilGeo-TNC = Anderson, M.G., Barnett, A., Clark, M., Prince, J., Olivero Sheldon, A. and Vickery B. 2016. Resilient and Connected Landscapes for

Terrestrial Conservation. The Nature Conservancy, Eastern Conservation Science, Eastern Regional Office. Boston, MA.; NHD = National Hydrography Dataset (https://nhd.usgs.gov/NHD_High_Resolution.html); NHD+ = NHDPlus (http://www.horizon-systems.com/nhdplus/); NLCD = National Land Cover Database (https://www.mrlc.gov/index.php); NWI = National Wetlands Inventory (https://www.fws.gov/wetlands/)

APPENDIX F. CONCEPTUALIZED DESCRIPTION OF MINIMUM TRAINING PRESENCE VARIANTS

Three Minimum Training Presence (MTP) threshold values are conceptualized below. For each pair of figures, the first represents the same ‘full spectrum’ raster with two EOs and three polygons labeled. The second figure show how a binary output, using the chosen threshold, would look in comparison to the known locations.

1a



1b

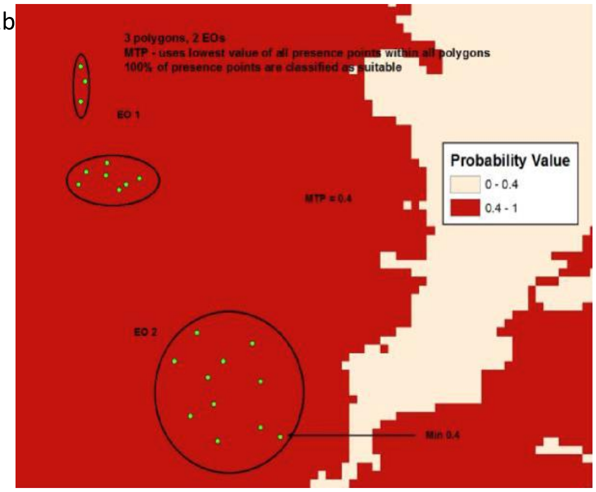
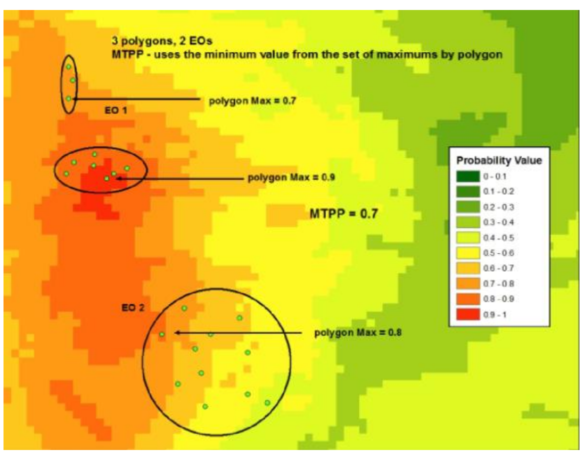


Figure 1. Minimum Training Presence – The lowest probability value (raster cell) assigned to any of the input presence points (green points) is used as the threshold. In this scenario, 0.4 is the lowest probability and is used as the threshold value in Figure 1b. Note, all the presence points are classified as ‘suitable’

2a



2b

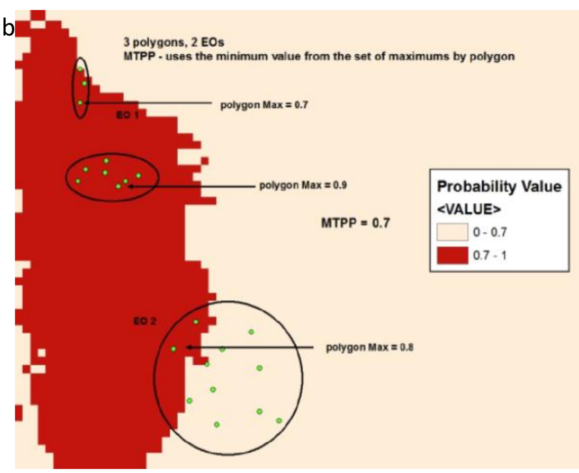


Figure 2. MTP by polygon (MTPP) – The calculation first takes the maximum value of presence points from each polygon, then uses the lowest value from the set of maximums. In this scenario, the maximums from the three polygons are 0.7, 0.9, and 0.8, thus 0.7 is used as the threshold. This is represented in Figure 2b. Note the higher threshold value decreases the area classified as ‘suitable’ compared to Figure 1b. In this example, portions of some polygons are not classified as suitable, yet at least some part of every polygon is classified as suitable. If only part of a polygon is classified as suitable, the portion classified represents the portion with the highest habitat suitability.

3a 3b

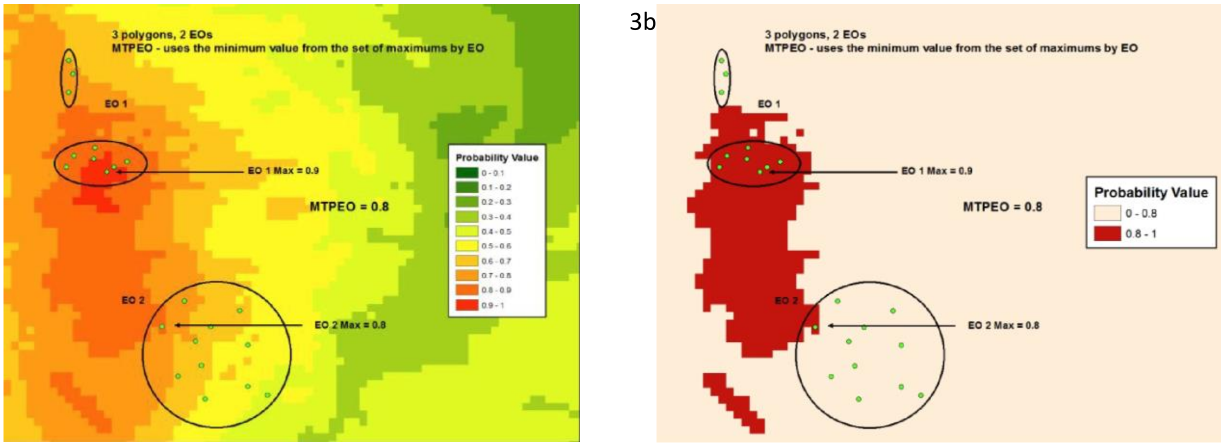


Figure 3. MTP by element occurrence (MTPEO) – The calculation first takes the maximum value of presence points from each EO, then uses the lowest value from the set of maximums. In this scenario, the maximums from the two EOs are 0.9 and 0.8, thus 0.8 is used as the threshold. This is represented in Figure 3b. Note the higher threshold value decreases the area indicated as 'likely suitable' compared to Figures 1b and 2b. In this example, some known locations are not classified as suitable, yet at least a portion of every EO is classified as suitable. The portion that is classified as suitable represents the highest probability of suitable habitat for that EO. Recall that an EO is intended to represent a single population of the targeted species, which may or may not occur in multiple patches.

APPENDIX G. SUMMARY OF IMPORTANT ENVIRONMENTAL VARIABLES IN WETLAND BUTTERFLY SPECIES DISTRIBUTION MODELS

This table consolidates the nine most influential environmental variables illustrated in partial dependence plots for each species distribution model. The variables are placed in order from most to least frequently identified as important across all fourteen species. A complete list of environmental variables is provided in Appendix E.

Environmental Variable	anat loga	bolo sele	cart pala	chlo harr	euph bima	euph cons	euph dion	euph phae	leth eury	lyca epix	lyca hyll	poan mass	poan via	poli myst	TOTAL
Roughness 1-cell square	X	X		X	X	X				X	X			X	8
Dist to fresh marsh	X	X			X		X		X		X	X			7
Slope		X			X	X				X	X		X	X	7
Dist to silt/clay				X		X					X	X	X		5
Wetland cover 10-cell mean		X	X				X		X				X		5
Canopy 1-cell mean				X	X	X				X				X	5
Max temp of warmest month		X	X	X					X						4
Elevation						X		X			X			X	4
Dist to woody wetland		X		X		X								X	4
Evergreen forest cover 100-cell mean	X			X	X									X	4
Roughness 10-cell circle					X	X				X					3
Dist to lake	X		X					X							3
Topographic positions index 100-cell radius										X	X	X			3
Growing degree days			X	X					X						3
Topographic positions index 10-cell radius					X	X						X			3
Wetland cover 100-cell mean							X		X				X		3
Annual mean temp			X	X					X						3
Canopy 10-cell mean	X			X										X	3
Open cover 100-cell mean	X				X					X					3
Mean temp of wettest quarter			X											X	2
Mean temp of coldest quarter			X					X							2
Isothermality	X							X							2
Mean temp of driest quarter			X						X						2
Flowpath dist to water or wetland		X					X								2
Water cover 100-cell mean	X						X								2
June precip			X						X						2
Dist to lake or river	X										X				2
Dist to loam							X					X			2
Dist to salt marsh											X	X			2
Topographic moisture							X						X		2
Shrub cover 100-cell mean							X		X						2
Wetland cover 1-cell mean							X						X		2
Precip of driest quarter	X										X				2

Slope length		X											X		2
Mean temp of warmest quarter					X									X	2
Mean diurnal range										X					1
Dist to estuary												X			1
Dist to ocean		X													1
Dist to coastal waters												X			1
Dist to inland waters												X			1
Dist to moderately calc rock								X							1
Canopy 100-cell mean								X							1
Deciduous forest cover 100-cell mean								X							1
Forest cover 100-cell mean								X							1
May precip						X									1
Precip of driest month										X					1
Impervious surface 10-cell mean								X							1
Temp seasonality													X		1
Solar radiation winter solstice													X		1

APPENDIX H. WISSAHICKON VALLEY WATERSHED ASSOCIATION - WETLAND RESTORATION REPORT

WISSAHICKON valley WATERSHED ASSOCIATION

60 Years of Watershed Moments

Restoration and Habitat Management
Crossways Preserve
2016-2017

Below is a summary of habitat restoration and management activities completed over two years and multiple seasons at WVWA's Crossways Preserve. This work was completed as part of the Regional Conservation Needs Grant, "Conservation and Management of Rare Wetland Butterflies," which WVWA is subcontracted to complete in partnership with the Western Pennsylvania Conservancy. It was also completed in accordance with the preserve's Conservation Management Plan, a main goal of which is to conserve, restore, and enhance natural habitats vital to native species which are threatened, endangered, rare, or imperiled.

Invertebrate surveys to determine the presence or absence of butterfly Species of Greatest Conservation Need (SGCN) and vegetation surveys to determine habitat diversity were completed in 2016; these surveys also helped inform restoration and management activities.

Please refer to the map (Figure 1.), below, which shows the management units referenced herein.

Invasive Species Removal & Native Plant Installation

2016

In August of 2016, non-native, invasive species were removed along the border between the Hardwood Forest and the Shrub-scrub Meadow in an area of approximately 2 acres. Several species were targeted, but predominantly included obtuse-leaved and Chinese privet, multiflora rose, Japanese Angelica tree, tree of heaven, Japanese stiltgrass, wineberry, and Japanese barberry. These species were flagged and removed using mechanical methods of pulling and cutting – some larger individuals were cut at approximately 3ft above ground and treated with a 2% solution of glyphosate to prevent re-growth of cut stems. Invasive removal was accomplished with the help of more than 60 volunteers over 3 days of service. Replanting with 97 native trees and shrubs (Table 1) was completed in October and November, again with the help of volunteers.

Table 1. Native Species Installed at Crossways Preserve, Fall 2016

TREES

Scientific Name	Common Name	Quantity
Acer saccharum	sugar maple	6

Betula lenta	sweet birch	7
Carpinus caroliniana	American hornbeam	5
Cornus florida	flowering dogwood	6
Diospyros virginiana	common persimmon	4
Hamamelis virginiana	witch-hazel	7
Nyssa sylvatica	black gum	7
	Total	42

SHRUBS

Scientific Name	Common Name	Quantity
Cornus racemosa	gray dogwood	10
Ilex glabra	inkberry holly	10
Morella pensylvanica	northern bayberry	12
Sambucus canadensis	elderberry	18
Vaccinium corymbosum	highbush blueberry	5
	Total	55

2017

In March of 2017, invasive removal continued within the Scrub-shrub Meadow. Efforts were focused on controlling non-native Callery (Bradford) pear trees, privet, and multiflora rose. Another 187 more trees and shrubs were added in April and May of 2017 (Table 2).

Table 2. Native Species Installed at Crossways Preserve, Spring 2017

TREES

Scientific Name	Common Name	Quantity
Cornus alternifolia	alternate-leaf dogwood	14
Crataegus phaenopyrum	Washington hawthorn	25
Diospyros virginiana	common persimmon	13
Juniperus virginiana	eastern red cedar	5
Prunus americana	American plum	10
Prunus serotina	black cherry	13
	Total	80

SHRUBS

Scientific Name	Common Name	Quantity
Alnus serrulata	smooth alder	25
Cornus amomum	silky dogwood	15
Hamamelis virginiana	witch-hazel	2
Physocarpus opulifolius	common ninebark	20
Rosa palustris	swamp rose	10
Rosa virginiana	Virginia rose	10
Vaccinium corymbosum	highbush blueberry	20
Viburnum lentago	nannyberry	5
	Total	107


Trees and shrubs planted in the Shrub-scrub Meadow after mowing, April, 2017.

Approximately 2 acres of the Managed Meadow were mowed in April and early May (prior to songbird nesting season). Following mowing, the area was treated with a 2% solution of glyphosate to control non-native, invasive Canada thistle, which had spread and was present in several large patches; the Shrub-scrub meadow was also mowed, and the thistle patches present treated.

Following herbicide treatment, the mowed and treated areas were seeded with a native wildflower and grass mix (Table 3), chosen for its variety of host and nectar plants necessary to several target SGCN butterflies.

Table 3. Native Wildflowers And Grasses Seeded at Crossways Preserve, Spring 2017

Scientific Name	Common Name	Percentage of Seed Mix
Aquilegia canadensis	eastern columbine	0.10%
Asclepias tuberosa	butterfly milkweed	2.00%
Aster laevis	smooth blue aster	1.00%
Aster novae-angliae	New England aster	0.50%
Aster prenanthoides	zigzag aster	0.60%
Baptisia australis	blue false indigo	0.40%
Baptisia tinctoria	yellow false indigo	0.20%
Bouteloua curtipendula	sideoats grama	23.40%
Chamaecrista fasciculata	partridge pea	2.60%
Coreopsis lanceolata	lanceleaf coreopsis	3.00%
Echinacea purpurea	purple coneflower	3.50%
Elymus riparius	riverbank wildrye	14.00%
Liatris spicata	blazing star	1.00%
Monarda fistulosa	wild bergamot	0.40%
Penstemon digitalis	tall white beardtongue	2.00%
Penstemon hirsutus	hairy beardtongue	0.10%
Rudbeckia fulgida var. fulgida	orange coneflower	0.50%
Rudbeckia hirta	black-eyed Susan	2.20%

Schizachyrium scoparium	little bluestem	40.00%
Senna hebecarpa	wild senna	0.40%
Senna marilandica	Maryland senna	0.10%
Solidago juncea	early goldenrod	0.20%
Solidago nemoralis	gray goldenrod	0.30%
Tradescantia ohiensis	Ohio spiderwort	0.50%
Zizia aurea	golden Alexanders	0.50%

Habitat Creation - Vernal Pool

In July, a 40' X 60' vernal pool was installed in the Hardwood Forest to create additional water resources. The pool was fenced in with 8' deer fencing and will be monitored throughout the year.

Projected Management and Restoration Activities to Begin 2018

__Vernal Pool__

The vernal pool will be assessed in the spring and summer of 2018 to identify what species emerge from the seed bank surrounding the pool in the absence of deer predation. Once presence and absence is determined, additional or missing species will be planted to increase host and nectar plants for SGCN butterfly species, including Buttonbush, Tussock Sedges, Jewelweed, Viburnum, and other similarly beneficial plants associated with vernal habitat.

__Wetland Enhancement__

Below are highlighted species for habitat restoration to the Right-of-Way (ROW) wetland (part of the unit indicated by the magenta line on the map (Figure 1). These suggested species are the result of invertebrate and vegetation surveys completed in 2016.

Planting of native species will occur only after **two to three full years** of mechanical and chemical control of the invasive species, *Phragmites australis*, of which there are a few invasive patches along the ROW, to ensure good conditions for their establishment. The wetland will be assessed after the first and second years to determine the effectiveness of initial control and readiness for native plant restoration.

The ROW has been divided into 6 Quadrants; refer to Figure 2, below. These are further divided into north edge, middle area, and southern edge by symbols, lower-case n, m, s. The plants in Table 4., below, all have been numbered. Recommended areas include all designated zones in that sub-area. Example: soft rush - En = to be planted anywhere along the northern limits of Quadrant E.

Table 4. Native Plant Species Recommendations for ROW Wetland and Forest Edge

Scientific Name	Common Name	Habitat preference	Quad Zone Recommendation	Attributes
Andropogon gerardii	giant broomsedge	Wet Areas	Bm, Em, Bs, En CS, Fn	Songbird winter food seed
Andropogon virginicus	broomsedge	Drier edges	An, Am, Dn, Bn, Es, Cn, Fs	Songbird – seeds / warm season
Carex lurida	shallow sedge	Wet areas	Bm, Em, Bs, En CS, Fn	Songbird winter seed source and micro-lep larvae host plants
Carex vulpinoidea	fox sedge	Moist areas- sedge meadow	Bm, Em, Bs, En CS, Fn, As, Dn, Fn, Cs	Micro lep host and songbird seed source
Chelone glabra	white turtlehead	Wet and borders of wet areas	Bm, Bs, En, Em, Es, Bn, Cs, Fn	Baltimore Checkerspot larvae host plant, pollinator plant
Eupatorium purpureum	purple stem boneset	Mesic-wet-upland areas	All Zones	Pollinator plant late summer/early fall
Euthamia graminifolia	flat-top goldenrod	Wet and mesic areas	Bm, Bs, En, Em, Es, Bn, Cs, Fn	Pollinator source summer
Heliopsis helianthoides	smooth oxeye	Mesic sites	Bm, Em, Bs, En CS, Fn, As, Dn, Fn, Cs, Am, Bm, Cm, Fm, Em, Dm	Late season color and height, seed source birds, pollinator plant
Hibiscus moscheutos	swamp hibiscus	Wet areas	Bm, Em, Bs, En CS, Fn	High plant, beautiful summer bloom, pollinator host plant
Ilex verticillata	winterberry holly	Scattered in wet open areas, ~6 shrubs	Bm, Bs, En, Em, Cs, Fn	Late winter song bird food source, winter color, spring pollinator source
Juncus effusus	soft rush	Wet open areas	Bs, Em, Cs, Fn, En,	Micro leps larvae host, Grass skippers
Liatris spicata	blazing star	Moist areas -sedge meadow	Bm, Em, Bs, En CS, Fn, As, Dn, Fn, Cs	Hummingbird nectar and pollinator plant
Lobelia cardinalis	cardinal flower	Wet Open Areas	Bm, Em, Bs, En CS, Fn	Pollinator plant, summer color, hummingbird plant
Lobelia siphilitica	blue lobelia	Wet Areas	Bm, Em, Bs, En CS, Fn	Pollinator plant
Monarda punctata	beebalm	Mesic areas	Am, As, Dm, Dn, Cm, Cs, Fn, Fm, Fs	Mid-season pollinator source, hummingbird nectar source
Penstemon digitalis	beardtongue	Mesic and xeric areas	All Quads All areas	Spring pollinator plant
Rosa palustris	swamp rose	Scattered throughout, ~6 shrubs	Bm, Em, Bs, En CS, Fn	Pollinator source spring, high quality migrating bird food source
Rudbeckia hirta	black-eyed Susan	Dry and mesic macro sites	An, Am, As, Dn, Dm, Ds, Cm	Summer and fall blooming season, pollinator plant
Sambucus canadensis	elderberry	Ecotone along ROW	All Quads along open forested edges	Pollinator and song bird food source
Scirpus cyperinus	wool grass	Wet open Areas	Bs, Em, Cs, Fn, En,	Micro leps larvae host, Grass

				skippers
Sorghastrum nutans	Indian grass	Drier Edges	An, Am, As, Dn, Dm, Ds	Songbird food source and micro lep larvae food.
Symphyotrichum novae-angliae	New England aster	Moist areas -sedge meadow	Bm, Em, Bs, En CS, Fn, As, Dn, Fn, Cs	Pollinator plant
Vaccinium corymbosum	highbush blueberry	All Forested edge areas	All Quad Edges – into forest interior	Spring nectar source, bird fruit source
Vernonia noveboracensis	New York ironweed	Mesic to hydric macrosites	Bm, Em, Bs, En CS, Fn, As, Dn, Fn, Cs, Am, Bm, Cm, Fm, Em, Dm	Pollinator plant, late season color, song bird seed source
Zizania aquatica	wild rice	Wet areas	Bm, Em, Bs, En CS, Fn	Wildlife food source and micro lep larvae host plant

Figure 1. Crossways Management Units

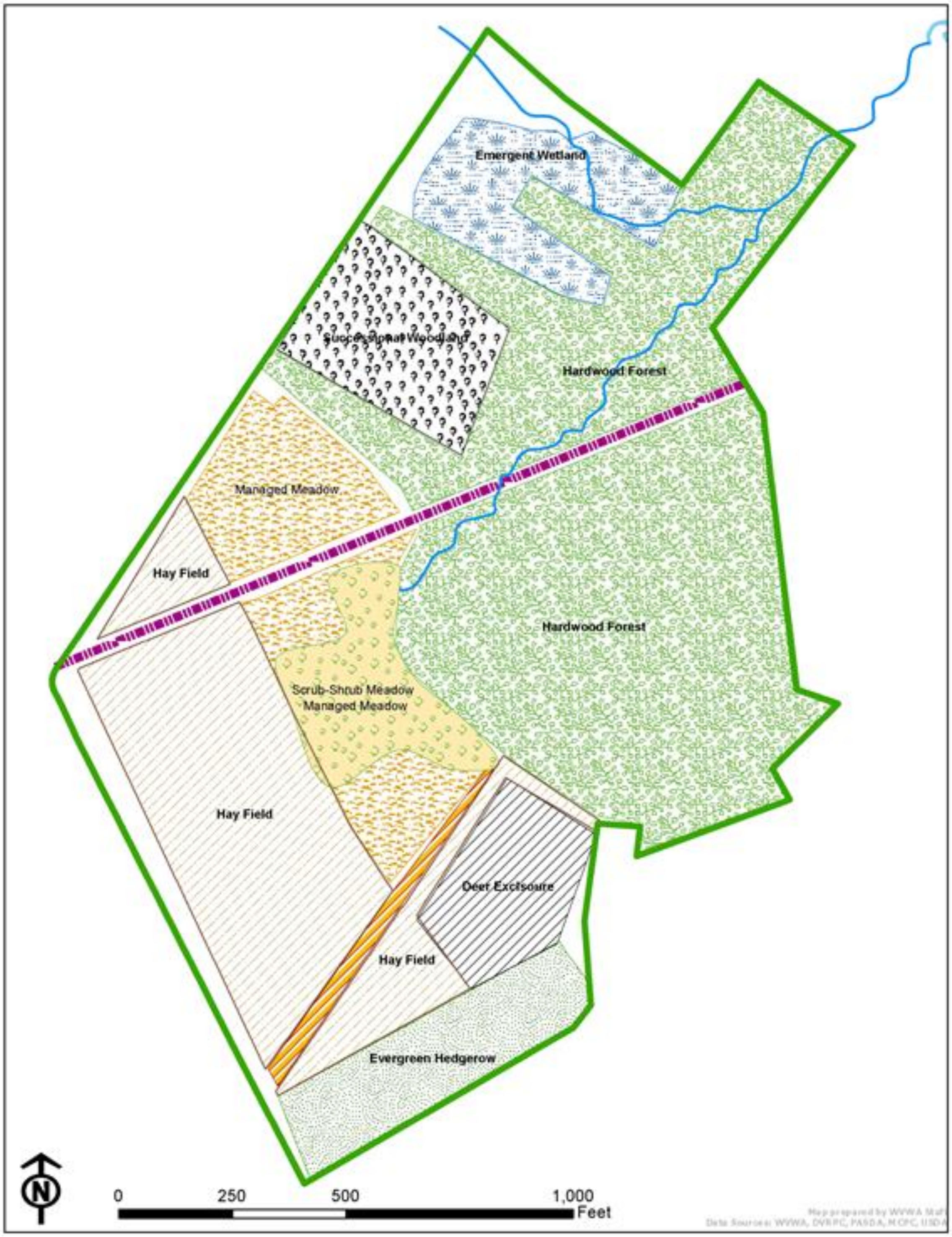


Figure 2. Pipeline ROW Quadrants

APPENDIX I. WETLAND BUTTERFLY HABITAT ASSESSMENT AND ENHANCEMENT BMPS

Utilize a standardized assessment form to prioritize wetland sites for enhancement or restoration: Our wetland butterfly habitat assessment and butterfly survey forms (Appendix B) provided a consistent, repeatable method for ranking wetland quality based on a number of factors including wetland vegetation composition, threats present, and SGCN species present. The forms can be modified to assess different variables or accommodate different types of sites; even our original forms were modified in 2016 based on suggestions from the biologists and technicians using them for assessments. They can be scored allowing sites to be ranked from high to low quality. Since our grant allowed for enhancement projects and not full scale restoration projects, it allowed us to identify and avoid work at lower quality wetlands that required intensive management and focus instead on maintaining higher quality wetlands with reasonable management goals.

Use climate model data as an additional assessment tool when prioritizing wetland sites for enhancement or restoration: Climate models generate output data for each species showing areas of declining, expanding, or stable climate. Using these models in addition to Wetland Site Assessment Forms allows for the prioritization of wetland enhancement projects when long-term wetland health is desired.

Use aquatic formulation glyphosate with LI-700 non-ionic surfactant for control of invasive plant species in wetlands: Many of the wetlands surveyed are also used by other wildlife, including bog turtles. The aquatic formulation glyphosate and LI-700 surfactants are the only products currently approved for use in bog turtle wetlands, would presumably be relatively safe for many other wildlife species as well, and have generally been effective in controlling woody vegetation.

In areas where mowing in or around wetlands is an option, avoid this activity during the growing season: whenever possible, vegetation that is used by wetland butterflies as host or nectar sources should be mowed once in the fall and after the first frost to avoid direct harm to active larvae and to maintain food resources.

For sites with Baltimore Checkerspots present, raise the mowing blade several inches during fall or winter mowing to prevent mortality of overwintering larvae present in the leaf litter.

Construct temporary fencing around recent plantings of host and nectar sources in areas where deer herbivory is a problem or in areas managed with grazing animals (i.e. goats): this will also prevent trampling of young plants. The use of welded wire fencing to form a circular exclosure with a diameter of less than 3m has shown to be effective in excluding deer and protecting plants from grazing animals.

Plant host plants and nectar plants in locations within the wetland where they already occur to assure suitable growing conditions: we do not have enough experience with wetland plantings to know if this is true, but some anecdotal evidence in Maryland suggests that white turtlehead plants planted in areas near other naturally-occurring white turtlehead plants are more successful than those planted randomly throughout the wetland. This is likely because naturally-occurring plants are already present in areas with favorable water tables, pH levels, and other physical characteristics, and are likely adjacent to vegetation that is not invasive or dominant.

STAND ALONE APPENDICES

The following documents are provided as stand-alone reports due to their large size.

APPENDIX J. SPECIES DISTRIBUTION MODELING METADATA

Metadata sheets for each species distribution model are located in this appendix. We recommend that this PDF file should always accompany the model outputs and be available to users.

APPENDIX K. CLIMATE ENVELOPE MODELING METADATA

Metadata sheets for each species climate envelope model are located in this appendix. We recommend that this PDF file should always accompany the model outputs and be available to users.

APPENDIX L. HABITAT MANAGEMENT FOR POLLINATORS (PA GUIDE)

This guide developed for Pennsylvania land managers and property owners is provided in this appendix. We provide information on best management practices for:

- promoting habitat variety to support all life stages of pollinators including adults and immatures
- maintaining open habitats through management practices such as rotation mowing
- controlling invasive plants
- protecting pollinator diversity and rare species
- selecting native and local plants for pollinators

APPENDIX M. LIFE HISTORY GUIDE TO 14 RARE WETLAND BUTTERFLIES IN THE MID-ATLANTIC

A guide to the life history of the fourteen wetland butterflies studied under this grant is provided in this appendix. This guide features photos of adults, caterpillars, and host plants, along with notes on species identification, range, habitat, phenology, and reproduction.