# EDI Metadata Template (2020)[[1]](#footnote-1)

We would prefer data to be in csv text file. If starting with an Excel spreadsheet, please make sure it does not contain any formulas and comments on cells. If you need comments put them in their own column. If data were used in a database and major table linking is necessary to analyze, please de-normalize into a flat file, not just database table exports.

## Dataset Title

(be descriptive, more than 5 words):

*Wood-warbler (Parulidae) range overlap under climate change scenarios*

## Short name or nickname you use to refer to this dataset:

*Warbler Range Shifts*

## Abstract

(include what, why, where, when, and how)

*Anthropogenic climate change will dramatically alter species distributions. The rate and magnitude of range shifts, however, will differ among taxa, resulting in altered patterns of co-occurrence and interspecific interactions. We examined potential climate-mediated breeding range shifts among North American wood-warblers (Parulidae), a speciose avian family likely to be especially impacted by such changes. We used publicly available species distribution model (SDM) range outputs developed by Bateman et al. (2020) to compare current ranges and patterns of sympatry among warbler species to future ranges and sympatry under 1.5 °C, 2.0 °C, and 3.0 °C of average global warming. We used these outputs to calculate average breeding range area, range overlap among species, number of sympatric species, and distances of breeding range shifts. We additionally calculated the number gained and lost sympatric interactions under each warming scenario. We publish the data and code here for both transparency of methods and reproducibility of results, including specifications for code modifications required for running on user machines.*

Referenced dataset:

Bateman, B.L., C.B. Wilsey, L. Taylor, J. Wu, G.S. LeBaron, and G. Langham. 2020. North American birds require mitigation and adaptation to reduce vulnerability to climate change. Conservation Science and Practice, 2(8) e242 (https://conbio.onlinelibrary.wiley.com/doi/full/10.1111/csp2.242).

## Investigators

(list in order as for a paper with e-mail addresses, organization and preferably ORCID ID, if you don’t have one, get it, it’s easy and free: <http://orcid.org/>) add table rows as needed

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| First Name | Middle Initial | Last Name | Organization | e-mail address | ORCID ID (optional) |
| Cody | H | Pham | UC Davis | cody.pham1999@gmail.com | 0000-0001-6523-9387 |
| J. | Jordan | Price | St. Mary’s College of Maryland | jjprice@smcm.edu  |  |
| Jason | M. | Tallant | University of Michigan | jtallant@umich.edu  | 0000-0003-4790-8772 |
| David | N. | Karowe | Western Michigan University | karowe@wmich.edu  |  |

## Other personnel names and roles

(dataset creators & contact, field crew, data entry etc. with e-mail addresses, organization and ORCID ID)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| First Name | Middle Initial | Last Name | Organization | e-mail address | ORCID ID (optional) | Role in project |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

## License

(Select a license for release of your data. We have 2 recommendations: [CCO – most accommodating of data reuse](https://creativecommons.org/publicdomain/zero/1.0/), & [CCBY – requires attribution](https://creativecommons.org/licenses/by/4.0/))

## Keywords

(List keywords and separate with commas. Using keywords from a controlled vocabulary (CV) will improve the future discovery and reuse of your data. The LTER CV is effective at describing ecological and environmental data. [Access the LTER CV here](http://vocab.lternet.edu/vocab/vocab/index.php). [Try this text mining service to extract LTER CV keywords from your abstract or methods](http://vocab.lternet.edu/keywordDistiller/). Additionally, please determine one or two keywords that best describe your lab, station, and/or project (e.g., Trout Lake Station, NTL LTER). This will help others discover your data by site/project).

*University of Michigan Biological Station, UMBS, REU, climate change, interspecific, range shift, species distribution model, sympatric, wood-warbler*

## Funding of this work:

Add rows to table if several grants were involved, list only the main PI, start with main grant first:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| PI First Name | PI Middle Initial | PI Last Name | PI ORCID ID (optional) | Title of Grant | Funding Agency | Funding Identification Number |
| David | N. | Karowe |  | Climate Change in the Great LakesRegion | NSF, US Department of Defense | AGS-1659338 |

## Timeframe

* *All data in this project derived from model outputs produced by Bateman et al. (2020;* [*https://adaptwest.databasin.org/pages/audubon-survival-by-degrees/*](https://adaptwest.databasin.org/pages/audubon-survival-by-degrees/)*). These data utilize bird occurrence data described in the original paper and project species ranges up until the year 2100.*
* Begin date: *n/a*
* End date: *n/a*
* Data collection ongoing/completed: *completed*

## Geographic location

* Verbal description: *North America*
* North bounding coordinates (decimals): *-49.0*
* South bounding coordinates (decimals): *83.3*
* East bounding coordinates (decimals): *-178.2*
* West bounding coordinates (decimals): -*6.6*

## Taxonomic species or groups

*Wood-warblers (Parulidae)*

## Methods

(please be specific, include instrument descriptions, or point to a protocol online, if this is a data compilation please specify datasets used, preferably their DOI or URL plus general citation information)

*Species distribution models*

*All data in this project derived from model outputs produced by Bateman et al. (2020;* [*https://adaptwest.databasin.org/pages/audubon-survival-by-degrees/*](https://adaptwest.databasin.org/pages/audubon-survival-by-degrees/)*). These model outputs included projections for the breeding ranges of 47 wood-warbler species under greenhouse gas Representative Concentration Pathways (RCPs) 4.5 and 8.5 in 2041-2070 (2050s) and 2071-2100 (2080s). To predict species ranges under each warming scenario, the SDMs incorporated current species occurrence, dispersal limitation, climate, vegetation, and land cover data to create continuous projections of suitable habitat for 1 km2 “pixels” across North America. We used thresholding approaches recommended by Bateman et al. (2020) to eliminate pixels that species are unlikely to occupy before creating predicted range maps for each model projection.*

*Range shift distance calculations*

*To calculate distances of species range shifts, we measured the distance between the centroid of a species’ current breeding range and the centroid of that range under each warming scenario. We weighted our centroid calculations in each warming scenario based on habitat suitability to place less emphasis on less climatically suitable areas where site fidelity will likely prevent occupation.*

*Species range overlap comparisons*

*To assess changes in sympatry among species, we compared range overlap between species’ current breeding ranges to future overlap under each warming scenario. For each possible pair of species (n = 2,162 pairs) in each warming scenario, we used Python for ArcGIS to overlay the range maps of each species and calculate the area of range overlap between the two species.*

*(see “Methodology” in the readme file for more detailed methods)*

## Data Table

* Column name: exactly as it appears in the dataset. Please avoid special characters, dashes and spaces.
* Description: please be specific, it can be lengthy
* Unit: please avoid special characters and describe units in this pattern: e.g. microSiemenPerCentimeter, microgramsPerLiter, absoptionPerMolePerCentimeter
* Code explanation: if you use codes in your column, please explain in this way: e.g. LR=Little Rock Lake, A=Sample suspect, J=Nonstandard routine followed
* Data format: please tell us exactly how the date and time is formatted: e.g. mm/dd/yyyy hh:mm:ss plus the time zone and whether or not daylight savings was observed.
* If a code for ‘no data’ is used, please specify: e.g. -99999

Please add rows as needed

**Naming conventions across columns in all files**

* *Scenarios: Present, RCP45\_2055 (1.5 C warming), RCP85\_2055 (2.0 C warming), RCP85\_2085 (3.0 C warming)*
* *Scenario comparisons columns: `SCENARIO1`\_`SCENARIO2` in column name indicates comparison between SCENARIO1 and SCENARIO2 (e.g. Present\_RCP45\_2055B\_Shift(km) indicates breeding range shift from present to 2.0 C warming)*
* *Habitat\_Group: determined based on NACBI 2010 State of the Birds Report, from Bateman et al. (2020)*

**RangeOverlapComparisons.csv:** overlap metrics for each possible pair of species in each scenario

* Column format: `SCENARIO `\_`SPP#`\_`METRIC`

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or code explanation or date format | Empty value code |
| `SCENARIO`\_Total\_Area | total breeding range area  | kilometers |  |
| `SCENARIO`\_Proportion\_Overlap | proportional overlap between breeding ranges |  |  |
| `SCENARIO`\_Total\_Area\_Overlap | total area overlap between breeding ranges | kilometers |  |
| Allopatric\_or\_Sympatric | allopatric or sympatric based on 10% threshold for determining sympatry |  |  |

**RangeOverlapData.csv:** summary of overlap metrics for each species (average overlap areas, number of sympatric species, etc.)

* Column format: `SCENARIO `\_`METRIC`

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or code explanation or date format | Empty value code |
| `SCENARIO`\_Number\_Species\_Overlapping | number of sympatric species based on 10% threshold for determining sympatry |  |  |
| `SCENARIO`\_Total\_Area | total breeding range area | kilometers |  |
| `SCENARIO`\_Mean\_Proportion\_Overlap | average proportional overlap with present sympatric species |  |  |
| `SCENARIO`\_Mean\_Total\_Area\_Overlap | average total area overlap with present sympatric species | kilometers |  |
| `SCENARIO`\_Sample Size | number of sympatric species in the present based on 10% threshold used to calculate averages |  |  |
| `SCENARIO`\_Overlapping Species | 4 letter codes of overlapping species |  |  |
| AlloToSym\_`SCENARIO` | number of gained sympatry in given scenario |  |  |
| SpeciesAlloToSym\_`SCENARIO` | 4 letter codes of gained sympatric species in given scenario |  |  |
| SymToAllo\_`SCENARIO` | number of lost sympatry in given scenario |  |  |
| SpeciesSymToAllo\_`SCENARIO` | 4 letter codes of lost sympatric species in given scenario |  |  |

**BreedingRangeShiftData.csv:** breeding range shift distances in each warming scenario

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or code explanation or date format | Empty value code |
| `SCENARIO1`\_`SCENARIO2`\_Shift(km) | breeding range shift distance | kilometers |  |

**ScenarioComparisons\_pValues.csv:** p-values and t-statistics from paired t-tests comparing metrics in each scenario

* Column format: `TESTMETRIC`\_`SCENARIO1`\_`SCENARIO2`\_`METRIC`

|  |  |  |  |
| --- | --- | --- | --- |
| Column name | Description | Unit or code explanation or date format | Empty value code |
| `TESTMETRIC`\_``SCENARIO1`\_`SCENARIO2`\_NumSppOlap | t-test values for comparison of number of sympatric species |  |  |
| `TESTMETRIC`\_``SCENARIO1`\_`SCENARIO2`\_MeanPropOlap | t-test values for comparison of mean proportional overlap |  |  |
| `TESTMETRIC`\_``SCENARIO1`\_`SCENARIO2`\_MeanAreaOlap | t-test values for comparison of mean total area of overlap |  |  |
| `TESTMETRIC`\_``SCENARIO1`\_`SCENARIO2`\_TotalArea | t-test values for comparison of total area |  |  |
| `TESTMETRIC`\_``SCENARIO1`\_`SCENARIO2`\_AlloToSym | t-test values for comparison of number of gained sympatric species |  |  |
| `TESTMETRIC`\_``SCENARIO1`\_`SCENARIO2`\_SymToAllo | t-test values for comparison of number of lost sympatric species |  |  |
| `TESTMETRIC`\_``SCENARIO1`\_`SCENARIO2`\_AlloToSymVSSymToAllo | t-test values for comparison of gained to lost sympatric species |  |  |
| `TESTMETRIC`\_``SCENARIO1`\_`SCENARIO2`Shift | t-test values for comparison of breeding range shifts (measured by shift of weighted centroid) |  |  |

## Articles

(List articles citing this dataset)

|  |  |  |
| --- | --- | --- |
| Article DOI or URL (DOI is preferred) | Article title | Journal title |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

## Scripts/code (software)

(List any software scripts/code you would like to archive along with your data. These may include processing scripts you wrote to create, clean, or analyze the data.)

|  |  |  |
| --- | --- | --- |
| File name | Description | Scripting language |
| OverlapCalculations.ipynb | calculates overlap metrics between each possible pair of species | Python |
| DataSummary.ipynb | calculates summary of overlap metrics for each species | Python |
| ShiftDistanceCalcs.ipynb | gets location of weighted centroids, calculates distance of range shifts | Python |
| GetPvalues.ipynb | runs paired t-tests comparing metrics in across warming scenarios, outputs statistical test values | Python |
| SppWarmingCompareDataClean.Rmd | extracts most useful metrics to compare what is happening to each species under each warming scenario | R |

## Data provenance

(Were these data derived from other data? If so, you will want to document this information so users know where these data come from.)

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset title | Dataset DOI or URL | Creator (name & email) | Contact (name & email) |
| Audubon Climate-Based Bird Distribution Models for North America | https://adaptwest.databasin.org/pages/audubon-survival-by-degrees/ | Brooke L. Bateman (Brooke.Bateman@audubon.org) | Brooke L. Bateman (Brooke.Bateman@audubon.org) |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

## Notes and Comments

1. This document liberally borrows from similar documents at SBC and GCE [↑](#footnote-ref-1)