Davis Courtney Lynn (Orcid ID: 0000-0002-6467-4288)
Robinson Orin Joseph (Orcid ID: 0000-0001-8935-1242)
Ruiz-Gutierrez Viviana (Orcid ID: 0000-0001-7116-1168)
Fink Daniel (Orcid ID: 0000-0002-8368-1248)

Journal: Ecology
Manuscript Type: Article
Handling Editor: James M. D. Speed

Deep learning with citizen science data enables estimation of species diversity and composition at continental extents

Courtney L. Davis[1*†], Yiwei Bai[2*], Di Chen[2], Orin Robinson[1], Viviana Ruiz-Gutierrez[1], Carla P. Gomes[2], Daniel Fink[1]

[1]Cornell Laboratory of Ornithology, Cornell University, Ithaca, NY 14850
[2]Department of Computer Science, Cornell University, Ithaca, NY 14850

Courtney L. Davis and Yiwei Bai contributed equally to this work.

Corresponding author: Courtney L Davis. Email: cld74@cornell.edu

Open Research: The complete list of data analyzed in this study is provided in Appendix S3. The eBird data used to conduct this study are fully described in Appendix S3 and freely available on the eBird website at https://ebird.org/science/use-ebird-data. Data related to the Checklist Calibration Index are sensitive because they pertain to the behavior and location of individual eBird users and cannot be made publicly available, however, Checklist Calibration Index data supporting this research can be directly requested by contacting ebird@cornell.edu and requesting access to the Checklist Calibration Index associated with the 2018 eBird Reference Dataset under a data sharing agreement. Model code and other non-sensitive data (Davis and Yiwei, 2023) are available on Zenodo at https://doi.org/10.5281/zenodo.8297796.

Abstract: Effective solutions to conserve biodiversity require accurate community and species-level information at relevant, actionable scales and across entire species' distributions. However, data and methodological constraints have limited our ability to provide such information in robust ways. Herein we employ DMVP-DRNets, an end-to-end deep neural network framework, to exploit large observational and environmental datasets together and estimate landscape-scale species diversity and composition at continental extents. We present results from a novel year-round analysis of North American avifauna using data from 9M eBird checklists and 72 environmental covariates. We highlight the utility of our information by identifying critical areas of high species diversity for a single group of conservation concern, the North American wood warblers, while capturing spatiotemporal variation in species' environmental associations and interspecific interactions. In so doing, we demonstrate the type of accurate, high-resolution information on biodiversity that deep learning approaches such as DMVP-DRNets can provide and that is needed to inform ecological research and conservation decision-making at multiple scales.

Keywords: artificial intelligence; biodiversity; critical areas; eBird; joint species distribution modeling; species richness

Introduction

Biodiversity loss is rapidly accelerating globally (Butchart et al. 2010), with significant implications for ecosystem function (Hooper et al. 2005, 2012) and human health (Cardinale et al. 2012). Several hundred international agreements focused on sustainability (e.g., United Nations Sustainable Development Goals (United Nations 2020) and conservation (e.g., Convention on Biological Diversity (CBD; Secretariat of the Convention on Biological Diversity 2020a) goals have recently been adopted, but with variable success. A marked example is the

recent failure to meet even one of the 20 Aichi Biodiversity Targets included in the CBD's Strategic Plan for Biodiversity 2011-2020 (Secretariat of the Convention on Biological Diversity 2020b). Pathways to success under the new post-2020 Global Biodiversity Framework (Secretariat of the Convention on Biological Diversity 2020a) will require that ratifying nations can assess the current state of biodiversity, quantify the impact of environmental change (e.g., climate change, changes in human land-use), and evaluate how present-day initiatives (e.g., national networks of protected areas) align with the protection of current and future biodiversity. With nearly one million plant and animal species facing extinction in the coming decades (Díaz et al. 2019), there is an increasingly urgent need for accurate, high-resolution information on biodiversity.

Unfortunately, information on species diversity (i.e., species richness, the number of species in a local community) and composition (i.e., the identity of species present in a local community) is notoriously difficult to estimate for large, diverse communities because of data limitations and sampling biases (Gotelli and Colwell 2001). Species richness is typically estimated across large spatial extents by overlaying expert range maps (Hurlbert and Jetz 2007), stacking of species distribution model predictions (Ferrier and Guisan 2006), or predicted directly via macroecological models (Francis and Currie 2003). In all cases, inferences about biodiversity are limited by the spatial and temporal resolution of available expert knowledge or modeling efforts (Hurlbert and Jetz 2007, Merow et al. 2017), and scale-dependent (Chase et al. 2019). Often, estimates are too coarse in spatial resolution and fail to capture seasonal changes in community composition (e.g., migratory species) and relative abundance, hindering the applicability of this information for conservation decision-making. Moreover, most methods are unable to account for the interspecific interactions that, in addition to environmental features,

influence what species can occur where and when (Ovaskainen et al. 2017). These challenges have impeded our ability to answer one of the most pressing questions in ecology today: how does biodiversity vary over space and time?

Recently developed deep learning joint species distribution models (e.g., Chen et al. 2018, Kong et al. 2020) provide a promising alternative to other statistical approaches for estimating species diversity and composition across a broad range of spatial and temporal scales. These methods use state-of-the-art artificial intelligence technologies (Chen et al. 2017) to decompose the spatial distributions of multiple species into shared environmental affinities and residual patterns of co-occurrence (Pollock et al. 2014, Warton et al. 2015, Ovaskainen et al. 2017), and are capable of scaling to the large numbers of species, locations, sample sizes, and environmental predictors necessary for broad-scale applications (Gomes et al. 2021). These methodological developments parallel the recent growth of digital citizen science platforms that can provide cost-effective and high-resolution information on entire ecological communities (Theobald et al. 2015, Thornhill et al. 2016, Chandler et al. 2017). Together, these advances greatly expand the scope of application, making it possible to predict, document, and study the spatiotemporal patterns of biodiversity at relevant, actionable scales and continental extents.

Here, we employ such an approach to make inferences on patterns of avian species diversity and composition at a high spatiotemporal resolution throughout the annual cycle. More specifically, we apply a Deep Reasoning Network (Chen et al. 2020, 2021) implementation of the Deep Multivariate Probit model (Chen et al. 2018; DMVP-DRNets) to citizen science data from 9,206,241 eBird checklists to examine the year-round spatiotemporal distributions, species-environment associations, and interspecific interactions for 500 species of North American avifauna while considering 72 environmental covariates. We highlight the utility of our results

by identifying areas of high species diversity throughout the annual cycle using the North American warblers (order Passeriformes, family Parulidae) as an exemplar group. Warblers are a subset of the Neotropical migrants that largely breed in temperate and northern forests and overwinter in tropical forests. Warblers are also a group of conservation concern, as 64% of species are currently experiencing widespread population declines (Rosenberg et al. 2019). Due to the migratory nature of this group, full annual cycle information is vital for identifying year-round critical areas for species diversity and key migration corridors to combat further declines. Moreover, understanding the seasonal habitat associations and interspecies interactions that structure species diversity can help prioritize research on the drivers of species declines across the annual cycle. Our results show how the information generated by the application of DMVP-DRNets to large-scale citizen science data can ultimately lead to more effective solutions for biodiversity conservation and improve our understanding of ecological communities.

Methods

Modeling Framework

We used a Deep Reasoning Network (Chen et al. 2020, 2021) implementation of the Deep Multivariate Probit (Chen et al. 2018) model (DMVP-DRNets; Chen et al. 2023), which employs a 3-layer-fully-connected network encoder to learn the relative importance of a large number of input features and generates a two-part structured latent space to express species' environmental associations as well as the interactions among species (Figure 1). A key advantage of deep learning is the ability to incorporate large, complex environmental data sets, allowing for a more accurate characterization of the high-dimensional processes that structure species' ecological niches and entire communities. Additionally, deep learning can isolate patterns that

are shared by multiple species, thereby improving predictions across all species but particularly those that are detected less frequently (Chen et al. 2017, Botella et al. 2018).

DMVP-DRNets integrates the multivariate probit model into an end-to-end deep learning framework with an interpretable latent space to produce three ecologically relevant outputs: 1) environmental association embeddings, which capture the multivariate associations of different environmental covariates, and interactions among these, on species' occurrences; 2) interactive association embeddings, which capture interactions among species via a residual correlation matrix; and 3) estimates of joint species occurrence probabilities across the study extent, which can be summarized at both the species- and community-level (e.g., to map species-specific distributions or species richness; Figure 1).

DMVP-DRNets uses a 3-layer-fully-connected network (denoted as MLP or multi-layer perceptron) to extract high-dimensional predictors h_i, from the raw environmental data x_i and encode a two-part latent space, which captures species-environment associations as $\mu_i = S^T \cdot h_i$ and the residual association among species as $\Sigma = I + \Lambda^T \Lambda$. We structured our MLP to have 1024, 1024, and 512 hidden units for each layer using the ReLU (Nair and Hinton 2010) activation function. The generative decoder then uses a multivariate probit distribution to map detection/non-detection data $y_{i,j}$ to a sequence of latent Gaussian random variables $r_{i,1}, \ldots, r_{i,m}$. Here, $\mathbf{r}_i$ is subject to a multivariate normal distribution with mean μ_i and covariance Σ: $\mathbf{r}_i \sim N(\mu_i, \Sigma)$, which captures species-specific environmental associations and residual interspecific interactions, respectively. A technical description of the model likelihood and implementation can be found in Appendix S1.

Model Evaluation

Before applying DMVP-DRNets at-scale, we first conducted a comprehensive comparison with the nearest neighbor approximation version of the spatial Gaussian Process HLR-S (Tikhonov et al. 2020). HLR-S is a hierarchical multivariate probit model with latent factors and explicit spatial correlation (Wilkinson et al. 2019, Niku et al. 2019) and is among the most widely applied joint species distribution model (JSDM) in ecology. We used 12 metrics (Norberg et al. 2019) to evaluate model performance with respect to accuracy, discrimination, calibration, and precision on three ecological levels: 1) species-specific occurrence, 2) species richness, and 3) community composition, which measured using three indices of pairwise community similarity: the Sørsenson-based dissimilarity index, the Simpson-based dissimilarity index, and the nestedness-resultant dissimilarity index. To aid in comparison, we modified the "discrimination" metrics to be 1 minus its original value and averaged the three similarity indices (Sorensen, Simpson and nestedness) to produce a singular community-level metric for each of the four performance categories. We refer readers to Norberg et al. (2019) for a more detailed description of these evaluation criteria.

We compared DMVP-DRNets and HLR-S using a variety of ecological datasets, including data from the 2011 Breeding Bird Survey (BBS; Pardieck et al. 2019), 5 other benchmark datasets used in recent JSDM comparisons (Wilkinson et al. 2019, Norberg et al. 2019), and random subsets of 1,000 and 10,000 eBird checklists from the full dataset described below. The BBS dataset contained detection/non-detection records from 2,752 sites, with observations of 370 species across North America and 8 climate-based site covariates. The other 5 datasets contained 1,200 presence-absence observations of 50 to 242 species with 3 to 5 environmental covariates. Additional information about each of these datasets can be found in Appendix S2.

We used the R-implementation of HLR-S (Tikhonov et al. 2020) with a similar MCMC configuration to conduct these model comparisons. We fitted HLR-S with 10,000 MCMC steps, discarding the first 2,000 samples as burn-in, and thinned to keep 1 out of every 80 iterations. Unlike deep learning models, Bayesian methods do not need a validation set to conduct model selection. Moreover, the training set was too small (only 600 data points) for each of the 5 benchmark datasets to split out a validation set. For these datasets, we trained DMVP-DRNets for a fixed number of epochs and then directly evaluated on the test set. For the BBS and eBird datasets, we randomly split 25% of the data to use as the test set, and 10% of the training data points to use as the validation set to perform model selection. Finally, we compared the computational efficiency of DMVP-DRNets to HLR-S by evaluating the wall-clock running time to train on each of the datasets. All models were trained and evaluated on one NVIDIA Tesla V100 GPU with 16GB memory. For the training process of our model, we selected a learning rate in (0:0001; 0:0005; 0:001) with Adam optimizer (Kingma and Ba 2017).

Finally, we compared DMVP-DRNets to a single-species approach that assumes species independence by constraining the shared covariance matrix to be zero. We evaluated differences in model performance associated with the estimation of these species-species associations for the BBS and eBird datasets. All model evaluation results can be found in Appendix S2.

eBird Data Description

We used a subset of eBird (Sullivan et al. 2009) data in which the time, date, and location of each checklist were reported and observers recorded all bird species detected and identified during the survey period, resulting in a "complete checklist" of species. The checklists used in this analysis were collected using the "stationary" or "traveling" protocols from January 1, 2004, to February 2, 2019, and within the spatial extent between 170° to 60° W longitude and between

20° to 60° latitude. We applied the best-practices for use of citizen science data from eBird (Johnston et al. 2021) to additionally filter checklists to those with durations of at most 1 hour and for traveling surveys at most 1km. We also removed duplicate records from group checklists where appropriate. The resultant dataset consisted of 9,206,241 eBird checklists.

We present an analysis of the 500 most frequently detected species on these filtered eBird checklists. This set of species includes a taxonomically diverse group of species, including both migrants and resident species that spanned a range of prevalence and range sizes, from common to infrequently detected species. To estimate the joint occurrence of these 500 species, we included 72 covariates of three general classes, including: 1) five observation-effort covariates to account for variation in detection rates; 2) three covariates to account for variation in detection rates at different temporal scales; and 3) 64 environmental covariates from remote sensing data to capture associations of birds with a variety of landscapes across the continent.

The observation-effort covariates capture heterogeneity in the observation process (i.e., effort expended and observer skill) and included: a) the duration spent searching for birds; b) whether the observer was stationary or traveling; c) the distance traveled during the search; d) the number of people in the search party; and e) the checklist calibration index, a standardized measure indexing differences in behavior among observers on checklists (Kelling et al. 2015, Johnston et al. 2018). We also included covariates to capture the time of day, day or year, and year of each observation. To account for differences in how time is recorded across time zones, the observation time of day was standardized as the difference from solar noon, the moment when the Sun crosses the local meridian and reaches its highest position in the sky at a given location. The day of the year (1–366) on which the search was conducted was used to capture intra-annual variation and the year of the observation was included to account for inter-annual

variation. Our initial data filtering combined with these observation-effort covariates helped to control for potential biases related to variable effort, spatial coverage, and preferential species reporting (Johnston et al. 2021).

The environmental descriptors included variables describing elevation, topography, shorelines, islands, land cover, land use, hydrology, and human development. To account for the effects of elevation and topography, each checklist location was associated with elevation (Becker et al. 2009), eastness, and northness. These latter two topographic variables combine slope and aspect to provide a continuous measure describing geographic orientation in combination with slope at 1km^2 resolution (Amatulli et al. 2018). Each checklist was also linked to a series of covariates derived from the NASA MODIS land cover, land use, and hydrology data (Carroll et al. 2017). We selected this data product for its moderately high spatial resolution, annual temporal resolution, and global coverage. We used the FAO-Land Cover Classification System which classifies each 500m pixel into land cover one of 21 vegetative cover classes, along additional classifications describing the land use and hydrology of each pixel. Checklists were linked to the MODIS data by year from 2001-2017, capturing inter-annual changes in land cover. The checklist data after 2017 were matched to the 2017 data, as MODIS data from after 2017 were unavailable at the time of analysis. To improve the classification of areas of human development, we also included the nighttime reflectance values from the 2016 NOAA VIIRS dataset (Cao et al. 2014). Additionally, to delineate the interface between terrestrial and marine environments we used NASA MODIS land water classification (Carroll et al. 2017) in conjunction with 30m shoreline and island data from GSV71 and the elevation data described above to classify each location into land, ocean, island and coastal areas. Finally, to identify

habitat for coastal species tidal mudflats were classified based on Murray's approach (Sayre et al. 2019).

To describe the composition and configuration of the local landscapes searched by participants, all cover classes were summarized within a 2.9km x 2.9km (877 hectare) neighborhood centered on the checklist location. In each neighborhood, we computed the composition as the proportion of each class in the neighborhood (PLAND). To describe the spatial configuration of each class we computed class level ED, an index of the edge density using the R package landscapemetrics (McGarigal et al. 2012). Because the elevation, eastness, northness, and VIIRS nighttime reflectance data are continuous measures, we computed the median and standard deviations of the values to capture the amount and variability of values within each neighborhood. A full list of the datasets analyzed in this study can be found in Appendix S3.

The STEM Wrapper

Our analysis uniquely captures spatiotemporal variation in both habitat and species associations using spatiotemporal exploratory models (STEM; Fink et al. 2010) as a wrapper around DMVP-DRNets to generate valuable insight into community-level processes across broad spatial and temporal extents. To do so, we repeatedly partitioned the study extent into randomly located grids of spatiotemporal blocks. Within each block, we trained DMVP-DRNets using a random split of 80% of data points falling within the block. The remaining data within the block were equally and randomly split into a validation (10% of data points), and test set (10% of data points). Within each spatiotemporal block, we also assumed the relationships between species' occurrence and the model covariates, and the relationships between species were stationary. This ensemble of partially overlapping local models was designed as a Monte Carlo sample of 484

randomly located spatial partitions of the study extent, with dimensions 15° longitude x 10° latitude, applied to each month of the year. This resulted in a uniformly distributed set of spatiotemporal blocks, and up to 24 local models covering each location in the study extent. Estimates at a given location and date were made by averaging across all the local models that contained that location and date and met a minimum sample size of 15,000 checklists. Spatiotemporal blocks that did not meet these minimum sample size requirements were removed from the ensemble.

Training over many smaller spatiotemporal extents with more localized datasets allows the ensemble to adapt to non-stationary species-species associations, i.e., seasaonal and regional changes in community structure across the study extent. Thus, the STEM ensemble functions to provide additional model flexibility, thereby reducing bias and providing better control for spatiotemporal extrapolation, compared to a single global model. Combining estimates across the ensemble provides control for inter-model variability associated with smaller sample sizes. Accounting for the spatiotemporal differences in these relationships not only provides novel insights into the factors structuring avian communities but also results in more accurate estimates of species' occurrences across broad spatial extents (Fink et al. 2010). The cumulative training time of all local models when using the STEM wrapper was 39.5 hours.

Critical Areas for Species Diversity

We used joint estimates of species occurrence to estimate richness of North American warbler species throughout the annual cycle. To calculate richness, we first applied a species-specific threshold that resulted in the maximum value of the kappa statistic (Monserud and Leemans 1992) to convert occurrence probabilities to a binary 0/1 scale. Occurrence probabilities greater than this threshold were set to 1, while probabilities less than this threshold

were set to 0. We then summed the binary predictions across all warbler species in a 2.9km x 2.9km pixel in each month to calculate species richness throughout the annual cycle. Our monthly estimates of warbler richness were reviewed by two world-renowned bird experts who oversee the data quality process in eBird. We also used the assessment of two co-authors of the paper (OR & VRG) who are well versed in the ecology and distribution of Neotropical migrants.

To identify critical areas of high warbler diversity, we first calculated the year-round maximum richness in each pixel. Pixels containing a high concentration of warbler species were then defined as those that exceeded the 95th percentile of year-round maximum richness across the entire study extent, excluding pixels with 0 warbler species. Medium and low concentration areas were defined as pixels that fell within the 90-95th and 80-90th percentiles, respectively. To better align our information with conservation decision-making, we overlapped these critical areas with existing Bird Conservation Regions (BCRs) that are composed of similar bird communities, habitats, and resource management issues (Bird Studies Canada and NABCI 2014). Within the BCRs that overlapped with year-round critical areas, we summarized maximum warbler richness as well as richness in the pre-breeding migration (i.e., May), breeding (i.e., June), and post-breeding migration (i.e., September) seasons. Finally, we identified critical migration areas using the estimated warbler richness in the months of May and September, respectively, where areas of highest concentration fell above the 95th percentile, areas of medium concentration fell between the 90th and 95th percentiles, and areas of low concentration fell between the 80th and 90th percentiles.

Shared Environmental Associations

We highlight year-round shared habitat relationships among warbler species across the Northeastern United States (i.e., within the spatial extent between 85° to 70°W longitude and

between 35° to 50°N latitude), while considering a broader community-level context. That is, we also examined shared habitat relationships among all species with a prevalence rate higher than 1% on all eBird checklists submitted in the Northeastern United States in each month.

To group species according to their similarity in environmental associations, we first z-standardized DMVP-DRNets' environmental association embedding and then used an average clustering algorithm on the correlation-based distance matrix. We then visualized patterns among species using hierarchical dendrograms and conducted 1000 bootstrap samples to calculate node support, defining relationships with >80% support as those that were strongly supported by the data. We also categorized species by primary breeding habitat and conservation status as defined by the Partners in Flight (PIF) database (Partners in Flight 2021). Categories of primary breeding habitat included: boreal forest, temperate forest, forest generalist, grassland, generalist, wetland, and coast. We used the range-wide population trends to define a species' conservation status, where a > 0.25% decline per year was defined as a negative population trend of concern. Species names listed are common names from the American Ornithological Society's checklist of North American birds (Chesser et al. 2020). We used the R packages pvclust (Suzuki and Shimodaira 2006), ggdendro (de Vries and Ripley 2020) and dendextend (Galili 2015), to compute and visualize dendrograms.

Interspecific Interactions

We examined the year-round residual correlations in occurrence between all pairs of warbler species, and between warblers and other 1) boreal forest species, 2) temperate eastern forest species, and 3) forest generalist species that also occur in the Northeastern United States. To do this, we first converted the residual covariance matrix Σ into a correlation matrix and then subset species groups according to their breeding habitat categories described previously.

Results

Critical Areas for Species Diversity

Using warblers as an exemplar species group, we generated monthly estimates of species richness at a spatial resolution of 2.9km x 2.9km. Warbler richness was incredibly dynamic throughout the annual cycle, with hotspots of species diversity that shifted drastically across the continent from one month to the next (Video S1). Critical areas with the highest number of warbler species year-round primarily overlapped with the Appalachian Mountains (Figure 2) but spanned 19 different Bird Conservation Regions (Bird Studies Canada and NABCI 2014) (BCRs; Appendix S4) established by the North American Bird Conservation Initiative. The Appalachian Mountains and Central Hardwoods BCRs consistently had the highest species richness of the regions that overlapped with year-round critical areas, but the distribution of species diversity across all other BCRs varied by season (Appendix S4: Figure S1).

Our results also allowed us to identify critical areas with a high concentration of warbler species in each of the migratory periods (Figure 3). Although these critical migration areas were similar between the two periods, we found key differences in the relative importance of areas based on the concentration of species diversity (Appendix S4: Figure S2). For example, the Appalachian Plateau of Ohio, West Virginia, and Pennsylvania hosts a higher concentration of species diversity during the pre-breeding migration season, while the northern Appalachians appear to be more important post-breeding.

Shared Environmental Associations

Our results on joint-species distributions also allow us to capture and disentangle the dynamic processes that structure species diversity and composition throughout the annual cycle. That is, our results not only provide information on the spatiotemporal distribution of species

diversity, but also retain information on the identity, habitat associations, and interspecific interactions of the individual species that make up that diversity across North America. Here, we use the Northeastern United States (i.e., within the spatial extent between 85° to 70° W longitude and between 35° to 50° latitude) to demonstrate how species' relationships with their environment vary over space and time.

At a coarse level, we were able to distinguish between species that are primarily associated with aquatic habitats from those associated with terrestrial environments in the Northeastern United States (Figure 4). Within these groups, our results allowed us to differentiate among species based on their primary habitat in the breeding season. For example, our results on habitat associations allowed us to accurately distinguish between warbler species that largely breed in boreal habitats from those that breed in temperate forests (Figure 4). These associations were further filtered by finer-scale habitat features. Warblers that are closely associated with riparian habitats or other water features, including the Prothonotary Warbler [*Protonotaria citrea*], Northern Parula [*Setophaga americana*], and Yellow-throated Warbler [*Setophaga dominica*], clustered strongly together, whereas the Pine Warbler [*Setophaga pinus*] clustered more strongly with other pine-associated species such as the Brown-headed Nuthatch [*Sitta pusilla*] and Red-headed Woodpecker [*Melanerpes erythrocephalus*]. As seasons change, our analysis captured the changing environmental associations of species as well (Appendix S4: Figure S3–S14). For example, most warblers clustered strongly together during the post-breeding migration season (Figure 5), except for the Palm Warbler [*Setophaga palmarum*] that tends to migrate later in the fall and the Pine Warbler and Yellow-rumped Warbler [*Setophaga coronata*] that remain prevalent in some parts of the Northeast throughout the year.

Interspecific Interactions

Our results on the co-occurrence patterns of species also discovered spatiotemporal variation in interspecific interactions, revealing drastically different patterns in residual correlation among species during the breeding, non-breeding, and migratory seasons. These unique insights allow us to demonstrate the ecological relevance of co-occurrence patterns for warblers in the Northeastern United States (Figure 4-5, Appendix S4: Figure S15). We found that warblers were negatively or weakly correlated in the breeding season when species are actively defending territories (Figure 6a). Among these were species pairs that have similar habitat preferences but exhibit interspecific aggression on the breeding grounds (e.g., Hooded Warbler [*Setophaga citrina*] and American Redstart [*Setophaga ruticilla*]). In general, warblers were more positively correlated with boreal species than with forest generalists or other species that breed in temperate forests but there was considerable interspecific variation in these relationships (Appendix S4: Figure S15a). The Yellow Warbler [*Septophaga petechia*], which is among the most abundant and widespread of warblers in the Northeast, was the only warbler strongly correlated with at least one species in each of the three different habitat types, including the Alder Flycatcher [*Empidonax alnorum*] (boreal forest), Cedar Waxwing [*Bombycilla cedrorum*] (forest generalist), Warbling Vireo [*Vireo gilvus*] (temperate eastern forest), and Willow Flycatcher [*Empidonax traillii*] (temperate eastern forest).

We further observed stark differences in interspecific interactions during migration, when most species were strongly correlated with each other (Figure 6b). Warblers were also positively correlated with most other forest species (Appendix S4: Figure S15b). These patterns align with anecdotal observations of warblers participating in mixed-species flocks with both migratory (e.g., Red-eyed Vireo [*Vireo olivaceus*] and resident passerines (e.g., Black-capped Chickadee [*Poecile atricapillus*]) during this time. Warblers were negatively correlated only with birds of

prey, including the Broad-winged Hawk [*Buteo platypterus*], Sharp-shinned Hawk [*Accipiter striatus*], Cooper's Hawk [*Accipiter cooperii*] and Red-shouldered Hawk [*Buteo lineatus*].

Discussion

Estimates of species diversity and composition are fundamental for studying processes that shape biological communities, and critically important for conservation planning (Myers et al. 2000, Fleishman et al. 2006). We applied a deep learning approach, the DMVP-DRNets, to uniquely generate continental estimates of species diversity and composition at relevant, actionable resolutions for decision-making, while also capturing year-round spatiotemporal variation in species' environmental associations and interspecific interactions. Our full annual cycle perspective reveals year-round critical areas for North American warblers, including key movement corridors during the pre-breeding and post-breeding migrations. This information can be used to coordinate adaptive conservation strategies (Reynolds et al. 2017) across regions and seasons (Sauer et al. 2003), and to identify and prioritize landscapes of high conservation value (Capmourteres and Anand 2016). Furthermore, the broad-scale community information provided by deep learning models allows us to integrate the influence of regional and seasonal processes across space throughout the year, making it possible to conduct accurate population-wide impact assessments. This is particularly important for studying various environmental and anthropogenic factors that contribute to or are affected by species declines, many of which are multi-scale processes, ranging from climate (Ådahl et al. 2006, Small-Lorenz et al. 2013) and land-use change (La Sorte et al. 2017) to ecosystem services (Birkhofer et al. 2015).

The fine-scale spatial structure shown in our maps of species richness arises from the environmental associations and interspecific interactions discovered through the deep learning process, and provides detailed, continental-extent information on the processes structuring

ecological communities. Using warblers in the Northeastern United States as an example, we show meaningful species-specific relationships with habitat and co-occurring species throughout the annual cycle. Communities were structured by shared environmental preferences (e.g., species associated with riparian habitats) and then further filtered by fine-scale biotic interactions (e.g., negative correlations indicative of spatial avoidance between predators and prey observed in the post-breeding migration season). These results not only provide basic ecological information about avian communities but may also help in the development of effective conservation strategies (i.e., targeted restoration to benefit an entire group of species) and identification of indicator groups that can be used to monitor species of conservation concern or the impacts of environmental change. Moreover, understanding the mechanisms underlying spatiotemporal patterns of species diversity can help prioritize research on the factors contributing to variable population trends across species' distributions.

A key contribution of deep learning approaches such as DMVP-DRNets is the increased computational power to handle more environmental covariates, species, and larger sample sizes than standard multi-species frameworks used to estimate richness. In addition, we show how end-to-end deep learning can allow us to characterize complex and high-dimensional relationships between entities, including species-environment associations and interspecific interactions. These advances greatly expand the scope of inference, making it feasible to predict, document, and study patterns of species diversity and community composition across a broader range of spatial and temporal scales. The application of deep learning methods that estimate relative abundance for multiple species (Kong et al. 2020), accommodate other data types (e.g., time series data), or that relax the assumption of symmetric, pairwise associations (Zhao et al. 2021) between species may also lead to more accurate predictions of species diversity and

composition that can better inform the prioritization of limited conservation resources (Johnston et al. 2015). Addressing challenges associated with interpretability of deep learning models and prediction uncertainty, more specifically, can further inform conservation decision-making and ensure that resources are more precisely directed to critical areas associated with high prediction certainty (Jansen et al. 2022, Wadoux et al. 2023). While our current application of DMVP-DRNets does not generate estimates of uncertainty, pixel-level estimates could be produced by assessing the variation among overlapping block-level estimates in the STEM ensemble.

Another promising avenue for future research includes extensions that explicitly estimate detection probabilities (e.g., Tobler et al. 2019). Our approach does not explicitly separate the observation and ecological processes by estimating species-specific detection probabilities (Dorazio and Royle 2005). Rather, we account for heterogeneity in the observation process (e.g., variation in detection rates) via model covariates. Estimates of occurrence should therefore be interpreted as a relative index of species occupancy probabilities (*sensu* Fink et al. 2020). Improvements in feature engineering to more fully describe known sources of variation in the observation process would provide inferential benefits without a need for the repeated sampling design needed to estimate detection probabilities. We also do not address challenges associated with the use of imbalanced data, which arise when the number of positive detections for a species is small compared to the number of locations where the species was not observed. Highly imbalanced data can be problematic for rare and infrequently detected species, but the performance of single-species distribution models can be improved using case-control sampling (Fithian and Hastie 2014, Robinson et al. 2017). Case-control sampling in the context of a multivariate (i.e., multispecies) problem is not trivial and remains an ongoing area of research (Tarekegn et al. 2021). Our joint model's structure likely provides some benefits for rare and

infrequently detected species (Ovaskainen and Soininen 2011, though see Erickson and Smith 2023) but additional work should be done to quantify the extent to which imbalanced data influence model performance and resulting inferences.

Amidst accelerating environmental changes, there is an increasingly urgent need for scalable ways to generate accurate, high-resolution information on biodiversity. Because of their ability to capture complex, community-level processes at multiple scales and across broad spatial extents, deep learning approaches such as DMVP-DRNets can serve as comprehensive, cost-effective, and adaptable methods by which to set biodiversity baselines and assess change in community composition at relevant spatiotemporal scales (Oliver et al. 2021). For example, estimates of species occurrence and biodiversity attributes (i.e., taxonomic, phylogenetic, and functional diversity) can inform several Essential Biodiversity Variables (Pereira et al. 2013, Jetz et al. 2019) proposed in the post-2020 Global Biodiversity Framework (Secretariat of the Convention on Biological Diversity. 2020a). While we highlight the utility of deep learning species distribution models as applied to eBird checklist data, broad-scale citizen science programs are just one of many growing sources of large ecological data streams (Farley et al. 2018). The integration of data across long-term monitoring networks (e.g., LTER, NEON) or automated sensor networks (e.g., wildlife camera traps), for instance, will provide a wealth of community-level information for other geographic regions and many other taxonomic groups (e.g., mammals; Ahumada et al. 2020) in the near future. In addition, the availability of new, high-resolution environmental data sources that better capture local habitat conditions (e.g., GEDI-derived data products on ecosystem structure; Dubayah et al. 2020) will further improve our ability to make robust ecological inference at high resolution and across broad spatiotemporal extents. Deep learning approaches can unlock the full potential of these large

observational and environmental datasets for both species- and community-level inference about biodiversity, and thus open a new and exciting frontier for data-driven conservation and ecology.

Acknowledgments: We thank the thousands of eBird participants and organizations for their contributions, as well as the eBird Status and Trends team, and Chris Wood and Ian Davies for reviewing the eBird results. This work was funded by the National Science Foundation Awards CCF-1522054 (Expeditions in computing), CNS-1059284 (Infrastructure), and DBI-1939187 (ABI sustaining), the Air Force Office of Scientific Research (AFOSR/DURIP-FA9550-21-1-0316 and FA9550-23-1-0322), and the U.S. Department of Agriculture's National Institute for Food and Agriculture (USDA-NIFA-2023-67021-39829), as well as The Leon Levy Foundation, The Wolf Creek Foundation, and the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Futures program.

Author contributions: D.F. and C.P.G. conceived and managed the project. Y.B. ran the eBird analysis. D.C. implemented and ran baseline comparisons. C.L.D., D.F., O.R., and V.R.G. processed the eBird data and made substantial contributions to the analysis and interpretation of results. C.L.D. wrote the first draft of the manuscript with substantive contributions from V.R.G., and D.F. All authors provided comments on subsequent drafts and gave approval for submission.

Conflict of interest: The authors declare no conflict of interest.

References

Ådahl, E., P. Lundberg, and N. Jonzén. 2006. From climate change to population change: The need to consider annual life cycles. Global Change Biology 12:1627–1633.

Ahumada, J. A., E. Fegraus, T. Birch, N. Flores, R. Kays, T. G. O'Brien, J. Palmer, S. Schuttler, J. Y. Zhao, W. Jetz, M. Kinnaird, S. Kulkarni, A. Lyet, D. Thau, M. Duong, R. Oliver, and A. Dancer. 2020. Wildlife Insights: A platform to maximize the potential of camera trap and other passive sensor wildlife data for the planet. Environmental Conservation 47:1–6.

Amatulli, G., S. Domisch, M.-N. Tuanmu, B. Parmentier, A. Ranipeta, J. Malczyk, and W. Jetz. 2018. Data descriptor: A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. Scientific Data 5:180040. https://doi.org/10.1038/sdata.2018.40.

Becker, J. J., D. T. Sandwell, W. H. F. Smith, J. Braud, B. Binder, J. Depner, D. Fabre, J. Factor, S. Ingalls, S.-H. Kim, R. Ladner, K. Marks, S. Nelson, A. Pharaoh, R. Trimmer, J. Von Rosenberg, G. Wallace, and P. Weatherall. 2009. Global bathymetry and elevation data at 30 arc seconds resolution: Srtm30_plus. Marine Geodesy 32:355–371. https://doi.org/10.1080/01490410903297766.

Bird Studies Canada and NABCI. 2014. Bird Conservation Regions. Published by Bird Studies Canada on behalf of the North American Bird Conservation Initiative.

Birkhofer, K., E. Diehl, J. Andersson, J. Ekroos, A. Früh-Müller, F. Machnikowski, V. L. Mader, L. Nilsson, K. Sasaki, M. Rundlöf, V. Wolters, and H. G. Smith. 2015. Ecosystem services - current challenges and opportunities for ecological research. Frontiers in Ecology and Evolution 2: fevo.2014.00087.

Botella, C., A. Joly, P. Bonnet, P. Monestiez, and F. Munoz. 2018. A deep learning approach to species distribution modelling. Pages 169–199 *in* A. Joly, S. Vrochidis, K. Karatzas, A. Karppinen, and P. Bonnet, editors. Multimedia Tools and Applications for Environmental & Biodiversity Informatics. Springer International Publishing, Cham.

Butchart, S. H. M., M. Walpole, B. Collen, A. van Strien, J. P. W. Scharlemann, R. E. A. Almond, J. E. M. Baillie, B. Bomhard, C. Brown, J. Bruno, K. E. Carpenter, G. M. Carr, J. Chanson, A. M. Chenery, J. Csirke, N. C. Davidson, F. Dentener, M. Foster, A. Galli, J. N. Galloway, P. Genovesi, R. D. Gregory, M. Hockings, V. Kapos, J.-F. Lamarque, F. Leverington, J. Loh, M. A. McGeoch, L. McRae, A. Minasyan, M. H. Morcillo, T. E. E. Oldfield, D. Pauly, S. Quader, C. Revenga, J. R. Sauer, B. Skolnik, D. Spear, D. Stanwell-Smith, S. N. Stuart, A. Symes, M. Tierney, T. D. Tyrrell, J.-C. Vie, and R. Watson. 2010. Global biodiversity: Indicators of recent declines. Science 328:1164–1168.

Cao, C., F. J. De Luccia, X. Xiong, R. Wolfe, and F. Weng. 2014. Early on-orbit performance of the visible infrared imaging radiometer suite onboard the Suomi National Polar-orbiting Partnership (S-NPP) satellite. IEEE Transactions on Geoscience and Remote Sensing 52:1142–1156. https://doi.org/10.1109/TGRS.2013.2247768.

Capmourteres, V., and M. Anand. 2016. "Conservation value": A review of the concept and its quantification. Ecosphere 7:e01476. https://doi.org/10.1002/ecs2.1476.

Cardinale, B. J., J. E. Duffy, A. Gonzalez, D. U. Hooper, C. Perrings, P. Venail, A. Narwani, G. M. Mace, D. Tilman, D. A. Wardle, A. P. Kinzig, G. C. Daily, M. Loreau, J. B. Grace, A. Larigauderie, D. S. Srivastava, and S. Naeem. 2012. Biodiversity loss and its impact on humanity. Nature 486:59–67.

Carroll, M. L., C. M. DiMiceli, M. R. Wooten, A. B. Hubbard, R. A. Sohlberg, and J. R. G. Townshend. 2017. Mod44w modis/terra land water mask derived from modis and srtm 13 global 250m sin grid v006. NASA EOSDIS Land Process DAAC, Sioux Falls, SD, USA. https://doi.org/10.5067/MODIS/MOD44W.006.

Chandler, M., L. See, K. Copas, A. M. Z. Bonde, B. C. López, F. Danielsen, J. K. Legind, S. Masinde, A. J. Miller-Rushing, G. Newman, A. Rosemartin, and E. Turak. 2017. Contribution of citizen science towards international biodiversity monitoring. Biological Conservation 213:280–294.

Chase, J. M., B. J. McGill, P. L. Thompson, L. H. Antão, A. E. Bates, S. A. Blowes, M. Dornelas, A. Gonzalez, A. E. Magurran, S. R. Supp, M. Winter, A. D. Bjorkman, H. Bruelheide, J. E. K. Byrnes, J. S. Cabral, R. Elahi, C. Gomez, H. M. Guzman, F. Isbell, I. H. Myers-Smith, H. P. Jones, J. Hines, M. Vellend, C. Waldock, and M. O'Connor. 2019. Species richness change across spatial scales. Oikos 128:1079–1091.

Chen, D., Y. Bai, S. Ament, W. Zhao, D. Guevarra, L. Zhou, B. Selman, R. B. van Dover, J. M. Gregoire, and C. P. Gomes. 2021. Automating crystal-structure phase mapping by combining deep learning with constraint reasoning. Nature Machine Intelligence 3:812–822.

Chen, D., Y. Bai, D. Fink, and C.P. Gomes. 2023. Deep Reasoning Network Implementation of a Deep Multivariate Probit Model (DMVP-DRNets) Code (v1.0). Zenodo. doi.org/10.5281/zenodo.8297796.

Chen, D., Y. Bai, W. Zhao, S. Ament, J. M. Gregoire, and C. P. Gomes. 2020. Deep Reasoning Networks for unsupervised pattern de-mixing with constraint reasoning. Pages 1500–

1509 Proceedings of the 37th International Conference on Machine Learning. PMLR, Online.

Chen, D., Y. Xue, D. Fink, S. Chen, and C. P. Gomes. 2017. Deep multi-species embedding. Pages 3639–3646 Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, Melbourne, Australia.

Chen, D., Y. Xue, and C. Gomes. 2018. End-to-End learning for the Deep Multivariate Probit Model. Pages 932–941 Proceedings of the 35th International Conference on Machine Learning. PMLR, Stockholm, Sweden.

Chesser, R. T., S. M. Billerman, K. J. Burns, C. Cicero, J. L. Dunn, A. W. Kratter, I. J. Lovette, N. A. Mason, P. C. Rasmussen, J. V. Remsen, Jr., D. F. Stotz, and K. Winker. 2020. Check-list of North American Birds (online). American Ornithological Society. https://checklist.americanornithology.org/taxa/

Davis, C., and B. Yiwei. 2023. gomes-lab/DMVP-DRNets: Chen et al. 2023 - Deep Reasoning Network Implementation of a Deep Multivariate Probit Model (DMVP-DRNets) Code. (v1.0). Zenodo. https://doi.org/10.5281/zenodo.8297796

de Vries, A., and B. D. Ripley. 2020. ggdendro: Create Dendrograms and Tree Diagrams Using "ggplot2." R package version 0.1.23, https://CRAN.R-project.org/package=ggdendro

Díaz, S., J. Settele, E. S. Brondízio, H. T. Ngo, J. Agard, A. Arneth, P. Balvanera, K. A. Brauman, S. H. M. Butchart, K. M. A. Chan, L. A. Garibaldi, K. Ichii, J. Liu, S. M. Subramanian, G. F. Midgley, P. Miloslavich, Z. Molnár, D. Obura, A. Pfaff, S. Polasky, A. Purvis, J. Razzaque, B. Reyers, R. R. Chowdhury, Y.-J. Shin, I. Visseren-Hamakers, K. J. Willis, and C. N. Zayas. 2019. Pervasive human-driven decline of life on Earth

points to the need for transformative change. Science 366:eaax3100. http://dx.doi.org/10.1126/science.aax3100.

Dorazio, R.M., and J.A. Royle. 2005. Estimating size and composition of biological communities by modeling the occurrence of species. Journal of American Statistical Association 100: 389–398.

Dubayah, R., J. B. Blair, S. Goetz, L. Fatoyinbo, M. Hansen, S. Healey, M. Hofton, G. Hurtt, J. Kellner, S. Luthcke, J. Armston, H. Tang, L. Duncanson, S. Hancock, P. Jantz, S. Marselis, P. L. Patterson, W. Qi, and C. Silva. 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Science of Remote Sensing 1:100002. https://doi.org/10.1016/j.srs.2020.100002.

Erickson, K., and A.B. Smith. 2023. Modeling the rarest of the rare: a comparison between multi-species distribution models, ensembles of small models, and single-species models at extremely low sample sizes. Ecography e06500. https://doi.org/10.1111/ecog.06500.

Farley, S. S., A. Dawson, S. J. Goring, and J. W. Williams. 2018. Situating ecology as a big-data science: Current advances, challenges, and solutions. BioScience 68:563–576.

Ferrier, S., and A. Guisan. 2006. Spatial modelling of biodiversity at the community level. Journal of Applied Ecology 43:393–404.

Fink, D., T. Auer, A. Johnston, V. Ruiz-Gutierrez, W. M. Hochachka, and S. Kelling. 2020. Modeling avian full annual cycle distribution and population trends with citizen science data. Ecology 30:e02056. https://doi.org/10.1002/eap.2056

Fink, D., W. M. Hochachka, B. Zuckerberg, D. W. Winkler, B. Shaby, M. A. Munson, G. Hooker, M. Riedewald, D. Sheldon, and S. Kelling. 2010. Spatiotemporal exploratory models for broad-scale survey data. Ecological Applications 20:2131–2147.

Fithian, W., and T. Hastie. 2014. Local case-control sampling: Efficient subsampling in imbalanced data sets. Annals of Statistics 42:1693–1724.

Fleishman, E., R. Noss, and B. Noon. 2006. Utility and limitations of species richness metrics for conservation planning. Ecological Indicators 6:543–553.

Francis, A. P., and D. J. Currie. 2003. A globally consistent richness-climate relationship for angiosperms. The American Naturalist 161:523–536.

Galili, T. 2015. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. Bioinformatics 31:3718–3720.

Gomes, C. P., D. Fink, R. B. van Dover, and J. M. Gregoire. 2021. Computational sustainability meets materials science. Nature Reviews Materials 6:645–647.

Gotelli, N. J., and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters 4:379–391.

Hooper, D. U., E. C. Adair, B. J. Cardinale, J. E. K. Byrnes, B. A. Hungate, K. L. Matulich, A. Gonzalez, J. E. Duffy, L. Gamfeldt, and M. I. O'Connor. 2012. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. Nature 486:105–108.

Hooper, D. U., F. S. Chapin, J. J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J. H. Lawton, D. M. Lodge, M. Loreau, S. Naeem, B. Schmid, H. Setälä, A. J. Symstad, J. Vandermeer, and D. A. Wardle. 2005. Effects of biodiversity on ecosystem function: A consensus of current knowledge. Ecological Monographs 75:3–35.

Hurlbert, A. H., and W. Jetz. 2007. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. Proceedings of the National Academy of Sciences 104:13384–13389.

Jetz, W., M. A. McGeoch, R. Guralnick, S. Ferrier, J. Beck, M. J. Costello, M. Fernandez, G. N. Geller, P. Keil, C. Merow, C. Meyer, F. E. Muller-Karger, H. M. Pereira, E. C. Regan, D. S. Schmeller, and E. Turak. 2019. Essential biodiversity variables for mapping and monitoring species populations. Nature Ecology & Evolution 3:539–551.

Johnston, A., D. Fink, W. M. Hochachka, and S. Kelling. 2018. Estimates of observer expertise improve species distributions from citizen science data. Methods in Ecology and Evolution 9:88–97.

Johnston, A., D. Fink, M. D. Reynolds, W. M. Hochachka, B. L. Sullivan, N. E. Bruns, E. Hallstein, M. S. Merrifield, S. Matsumoto, and S. Kelling. 2015. Abundance models improve spatial and temporal prioritization of conservation resources. Ecological Applications 25:1749–1756.

Johnston, A., W. M. Hochachka, M. E. Strimas-Macket, V. Ruiz-Gutierrez, O. J. Robinson, E. T. Miller, T. Auer, S. Kelling, and D. Fink. 2021. Analytical guidelines to increase the value of community science data: An example using eBird data to estimate species distributions. Diversity and Distributions 27:1265–1277.

Kelling, S., A. Johnston, W. M. Hochachka, M. Iliff, D. Fink, J. Gerbracht, C. Lagoze, F. A. La Sorte, T. Moore, A. Wiggins, W.-K. Wong, C. Wood, and J. Yu. 2015. Can observation skills of citizen scientists be estimated using species accumulation curves? PLOS ONE 10:e0139600. https://doi.org/10.1371/journal.pone.0139600.

Kingma, D. P., and J. Ba. 2017. Adam: A method for stochastic optimization. arXiv:1412.6980.

Kong, S., J. Bai, J. H. Lee, D. Chen, A. Allyn, M. Stuart, M. Pinsky, K. Mills, and C. Gomes. 2020. Deep Hurdle Networks for zero-inflated multi-target regression: Application to multiple species abundance estimation. Pages 4375–4381 Proceedings of the Twenty-

Ninth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, Yokohama, Japan.

La Sorte, F. A., D. Fink, P. J. Blancher, A. D. Rodewald, V. Ruiz-Gutierrez, K. V. Rosenberg, W. M. Hochachka, P. H. Verburg, and S. Kelling. 2017. Global change and the distributional dynamics of migratory bird populations wintering in Central America. Global Change Biology 23:5284–5296.

McGarigal, K., S. A. Cushman, and E. Ene. 2012. Fragstats v4: spatial pattern analysis program for categorical and continuous maps. Computer software program produced by the authors at the University of Massachusetts, Amherst.

Merow, C., A. M. Wilson, and W. Jetz. 2017. Integrating occurrence data and expert maps for improved species range predictions: Expert maps & point process models. Global Ecology and Biogeography 26:243–258.

Monserud, R. A., and R. Leemans. 1992. Comparing global vegetation maps with the Kappa statistic. Ecological Modelling 62:275–293.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853–858.

Nair, V., and G. E. Hinton. 2010. Rectified linear units improve restricted boltzmann machines. Pages 807–814 of the Proceedings of the 27th international conference on machine learning (ICML-10).

Niku, J., W. Brooks, R. Herliansyah, F. K. C. Hui, S. Taskinen, and D. I. Warton. 2019. Efficient estimation of generalized linear latent variable models. PLOS ONE 14:e0216129. https://doi.org/10.1371/journal.pone.0216129.

Norberg, A., N. Abrego, F. G. Blanchet, F. R. Adler, B. J. Anderson, J. Anttila, M. B. Araújo, T. Dallas, D. Dunson, J. Elith, S. D. Foster, R. Fox, J. Franklin, W. Godsoe, A. Guisan, B. O'Hara, N. A. Hill, R. D. Holt, F. K. C. Hui, M. Husby, J. A. Kålås, A. Lehikoinen, M. Luoto, H. K. Mod, G. Newell, I. Renner, T. Roslin, J. Soininen, W. Thuiller, J. Vanhatalo, D. Warton, M. White, N. E. Zimmermann, D. Gravel, and O. Ovaskainen. 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. Ecological Monographs 89:e01370. https://doi.org/10.1002/ecm.1370.

Oliver, R. Y., C. Meyer, A. Ranipeta, K. Winner, and W. Jetz. 2021. Global and national trends, gaps, and opportunities in documenting and monitoring species distributions. PLOS Biology 19:e3001336. https://doi.org/10.1371/journal.pbio.3001336.

Ovaskainen, O., and J. Soininen. 2011. Making more out of sparse data: Hierarchical modeling of species communities. Ecology 92:298–295.

Ovaskainen, O., G. Tikhonov, A. Norberg, F. Guillaume Blanchet, L. Duan, D. Dunson, T. Roslin, and N. Abrego. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. Ecology Letters 20:561–576.

Pardieck, K. L., D. J. Ziolkowski Jr., M. Lutmerding, and M. A. R. Hudson. 2019. North American breeding bird survey dataset 1966-2018, version 2018.0. U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland, USA.

Partners in Flight. 2021. Avian Conservation Assessment Database, version 2021. Available at http://pif.birdconservancy.org/ACAD.

Pereira, H. M., S. Ferrier, M. Walters, G. N. Geller, R. H. G. Jongman, R. J. Scholes, M. W. Bruford, N. Brummitt, S. H. M. Butchart, A. C. Cardoso, N. C. Coops, E. Dulloo, D. P.

Faith, J. Freyhof, R. D. Gregory, C. Heip, R. Hoft, G. Hurtt, W. Jetz, D. S. Karp, M. A. McGeoch, D. Obura, Y. Onoda, N. Pettorelli, B. Reyers, R. Sayre, J. P. W. Scharlemann, S. N. Stuart, E. Turak, M. Walpole, and M. Wegmann. 2013. Essential Biodiversity Variables. Science 339:277–278.

Pollock, L. J., R. Tingley, W. K. Morris, N. Golding, R. B. O'Hara, K. M. Parris, P. A. Vesk, and M. A. McCarthy. 2014. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). Methods in Ecology and Evolution 5:397–406.

Reynolds, M. D., B. L. Sullivan, E. Hallstein, S. Matsumoto, S. Kelling, M. Merrifield, D. Fink, A. Johnston, W. M. Hochachka, N. E. Bruns, M. E. Reiter, S. Veloz, C. Hickey, N. Elliott, L. Martin, J. W. Fitzpatrick, P. Spraycar, G. H. Golet, C. McColl, and S. A. Morrison. 2017. Dynamic conservation for migratory species. Science Advances 3:e1700707. https://www.doi.org/10.1126/sciadv.1700707.

Rosenberg, K. V., A. M. Dokter, P. J. Blancher, J. R. Sauer, A. C. Smith, P. A. Smith, J. C. Stanton, A. Panjabi, L. Helft, M. Parr, and P. P. Marra. 2019. Decline of the North American avifauna. Science 366:120–124.

Sauer, J. R., J. E. Fallon, and R. Johnson. 2003. Use of North American Breeding Bird Survey data to estimate population change for bird conservation regions. The Journal of Wildlife Management 67:372–389.

Sayre, R., S. Noble, S. Hamann, R. Smith, D. Wright, S. Breyer, K. Butler, K. Van Graafeiland, C. Frye, D. Karagulle, D. Hopkins, D. Stephens, K. Kelly, Z. Basher, D. Burton, J. Cress, K. Atkins, D. P. Van Sistine, B. Friesen, R. Allee, T. Allen, P. Aniello, I. Asaad, M. J. Costello, K. Goodin, P. Harris, M. Kavanaugh, H. Lillis, E. Manca, F. Muller-Karger, B.

Nyberg, R. Parsons, J. Saarinen, J. Steiner, and A. Reed. 2019. A new 30-meter resolution global shoreline vector and associated global islands database for the development of standardized ecological coastal units. Journal of Operational Oceanography 12:S47–S56. https://doi.org/10.1080/1755876X.2018.1529714.

Secretariat of the Convention on Biological Diversity. 2020a. Zero Draft of post-2020 Global Biodiversity framework. United Nations Environment Programme. https://www.cbd.int/doc/c/efb0/1f84/a892b98d2982a829962b6371/wg2020-02-03-en.pdf.

Secretariat of the Convention on Biological Diversity. 2020b. Global Biodiversity Outlook 5. Montreal. https://www.cbd.int/gbo/gbo5/publication/gbo-5-en.pdf.

Small-Lorenz, S. L., L. A. Culp, T. B. Ryder, T. C. Will, and P. P. Marra. 2013. A blind spot in climate change vulnerability assessments. Nature Climate Change 3:91–93.

Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling. 2009. eBird: A citizen-based bird observation network in the biological sciences. Biological Conservation 142:2282–2292.

Suzuki, R., and H. Shimodaira. 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics 22:1540–1542.

Tarekegn, A.N., M. Giacobini, and K. Michalak. 2021. A review of methods for imbalanced multi-label classification. Pattern Recognition 118: 107965. https://doi.org/10.1016/j.patcog.2021.107965.

Theobald, E. J., A. K. Ettinger, H. K. Burgess, L. B. DeBey, N. R. Schmidt, H. E. Froehlich, C. Wagner, J. HilleRisLambers, J. Tewksbury, M. A. Harsch, and J. K. Parrish. 2015. Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. Biological Conservation 181:236–244.

Thornhill, I., S. Loiselle, K. Lind, and D. Ophof. 2016. The citizen science opportunity for researchers and agencies. BioScience 66:720–721.

Tikhonov, G., L. Duan, N. Abrego, G. Newell, M. White, D. Dunson, and O. Ovaskainen. 2020. Computationally efficient joint species distribution modeling of big spatial data. Ecology 101:e02929. https://doi.org/10.1002/ecy.2929.

Tobler, M. W., M. Kéry, F. K. C. Hui, G. Guillera-Arroita, P. Knaus, and T. Sattler. 2019. Joint species distribution models with species correlations and imperfect detection. Ecology 100:e02754. https://doi.org/10.1002/ecy.2754.

United Nations. 2020. The Sustainable Development Goals Report. Available at https://sdgs.un.org/sites/default/files/2020-09/The-Sustainable-Development-Goals-Report-2020.pdf.

Warton, D. I., F. G. Blanchet, R. B. O'Hara, O. Ovaskainen, S. Taskinen, S. C. Walker, and F. K. C. Hui. 2015. So many variables: Joint modeling in community ecology. Trends in Ecology & Evolution 30:766–779.

Wilkinson, D. P., N. Golding, G. Guillera-Arroita, R. Tingley, and M. A. McCarthy. 2019. A comparison of joint species distribution models for presence–absence data. Methods in Ecology and Evolution 10:198–211.

Zhao, W., S. Kong, J. Bai, D. Fink, and C. Gomes. 2021. HOT-VAE: Learning high-order label correlation for multi-label classification via attention-based Variational Autoencoders. arXiv:2103.06375.

Figure 1. An overview of the modeling framework used in our study. DMVP-DRNets extracts high-dimensional predictors from the raw environmental data to encode the environmental association embeddings s_j and the interspecific association embeddings λ_j for each species j. These embeddings are used to compute species-specific habitat relationships and estimate the residual covariance among species, respectively. The generative-decoder then maximizes the joint likelihood of species j being present at location i under the multivariate probit distribution. DMVP-DRNets model outputs include: 1) environmental association embeddings s_j that capture high-dimensional representations of species-habitat relationships; 2) interspecific association embeddings λ_j, which are used to derive the pairwise covariance matrix Σ that describes residual species-species associations; and 3) species-specific occurrence predictions across the study extent, which can be summarized at either the species- or community-level (e.g., to calculate species richness).

Figure 2. Maps of year-round a) maximum warbler richness and b) critical areas of high warbler diversity across the study extent predicted at a spatial resolution of 2.9km. Full annual-cycle animation of warbler richness is included in Video S1. Critical areas are defined as the locations that fall above the 80th percentile of year-round maximum warbler richness, where areas of low concentration fall between the 80th and 90th percentiles, areas of medium concentration fall between the 90th and 95th percentiles, and areas of high concentration fall above the 95th percentile of maximum warbler richness.

Figure 3. Critical areas with high warbler diversity during a) pre-breeding and b) post-breeding migration seasons. Locations highlighted here fall above the 80th percentile of warbler richness

in the months of May and September, respectively, where areas of low concentration fall between the 80th and 90th percentiles, areas of medium concentration fall between the 90th and 95th percentiles, and areas of high concentration fall above the 95th percentile.

Figure 4. Hierarchical clustering dendrogram showing the similarity in environmental associations of species in the Northeastern United States during the breeding season. These dendrogram is a visual representation of the high-dimensional environmental association embeddings learned by DMVP-DRNets; additional months can be found in Appendix S3: Figure S3-S14. Solid lines indicate relationships with ≥ 80% bootstrap support, whereas dotted lines indicate relationships with < 80% bootstrap support.

Figure 5. Hierarchical clustering dendrogram showing the similarity in environmental associations of species in the Northeastern United States during the post-breeding migration season. This dendrogram is a visual representation of the high-dimensional environmental association embeddings learned by DMVP-DRNets; additional months can be found in Appendix S3: Figure S3-S14. Solid lines indicate relationships with ≥ 80% bootstrap support, whereas dotted lines indicate relationships with < 80% bootstrap support.

Figure 6. Residual pairwise correlation matrices for warbler species in the Northeastern United States during the a) breeding and b) post-breeding migration seasons. Species pairs that are negatively correlated (e.g., Yellow Warbler [*Setophaga petechia*] and Black-throated Green Warbler [*Setophaga virens*] in the breeding season) are less likely to co-occur, while positively correlated species pairs (e.g., Blackburnian Warbler [*Setophaga fusca*] and Bay-breasted

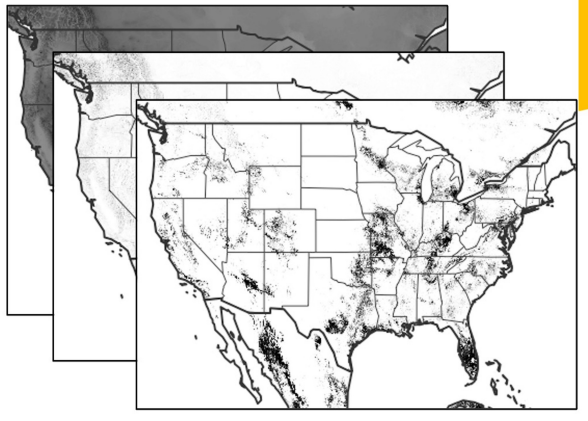
Warbler [*Setophaga castanea*] in the post-breeding migration season) are more likely to co-occur than expected after accounting for shared habitat preferences**.** Residual associations between species vary across the annual cycle, with drastically different patterns in the breeding, non-breeding, and migration seasons.

DMVP-DRNets' Model Structure

Observational & environmental data

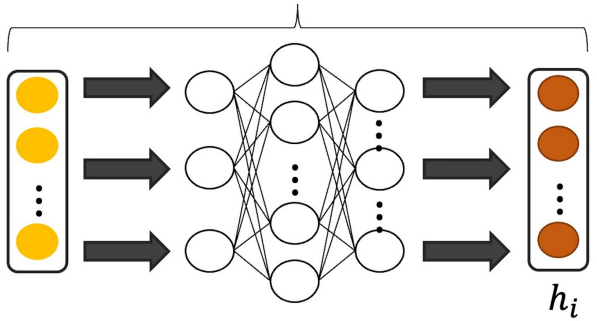
eBird

+



Encoder (MLP)

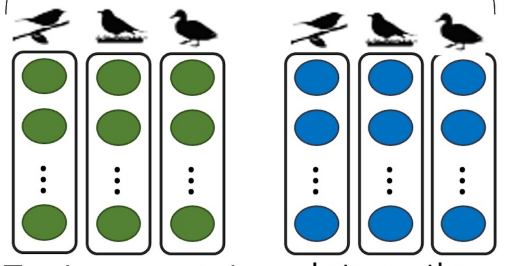
Environmental Feature Embedding



Deep Neural Network

h_i

Species Embedding



Environmental Preferences, s_j

Interactive Behavior, λ_j

Species-Habitat Relationships

$\mu_i = S^T h_i$

Residual Correlation

$\Sigma_{j,k} = \lambda_j^T \lambda_k$

Decoder (MVP)

Joint Species Distribution

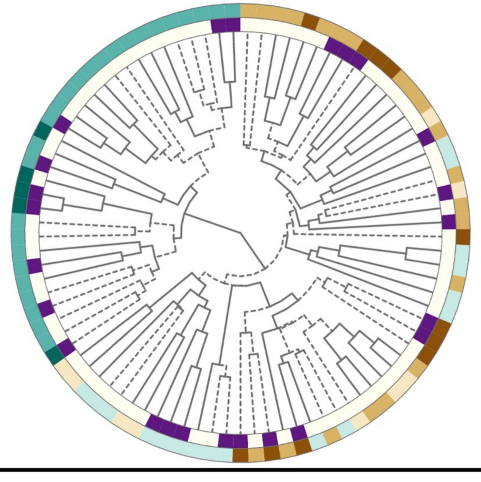
$\sim N(\mu_i, \Sigma)$

Where: i = location
j = species

DMVP-DRNets' Outputs & Summaries

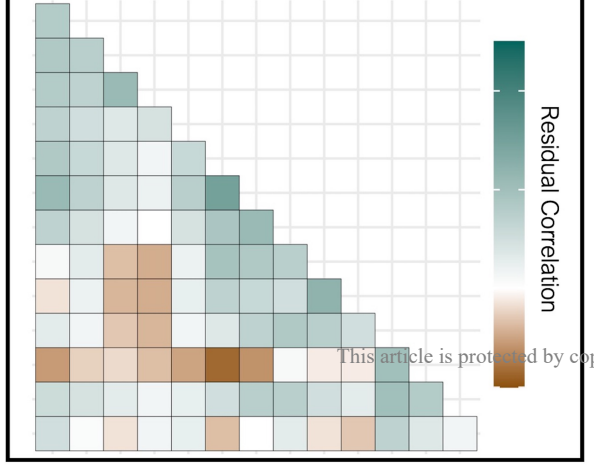
High-dimensional descriptions of species-habitat relationships

e.g., hierarchical clustering



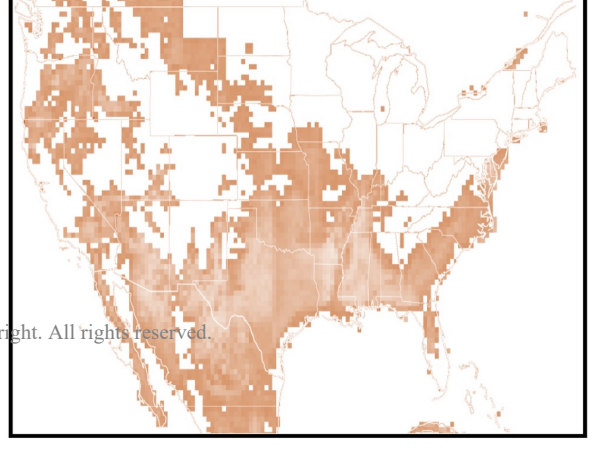
High-dimensional descriptions of species-species associations

e.g., pairwise correlation matrix



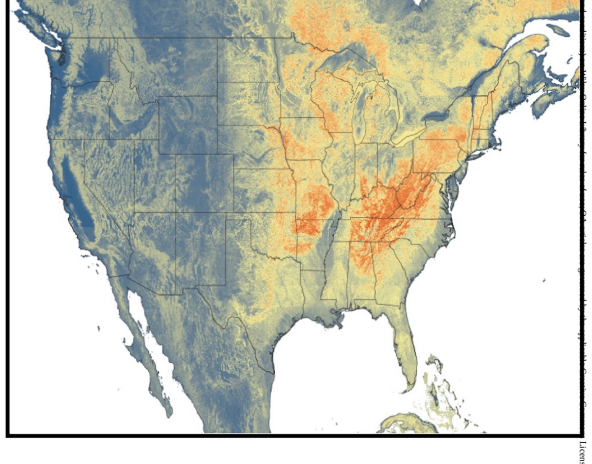
Species-level summaries of occurrence predictions

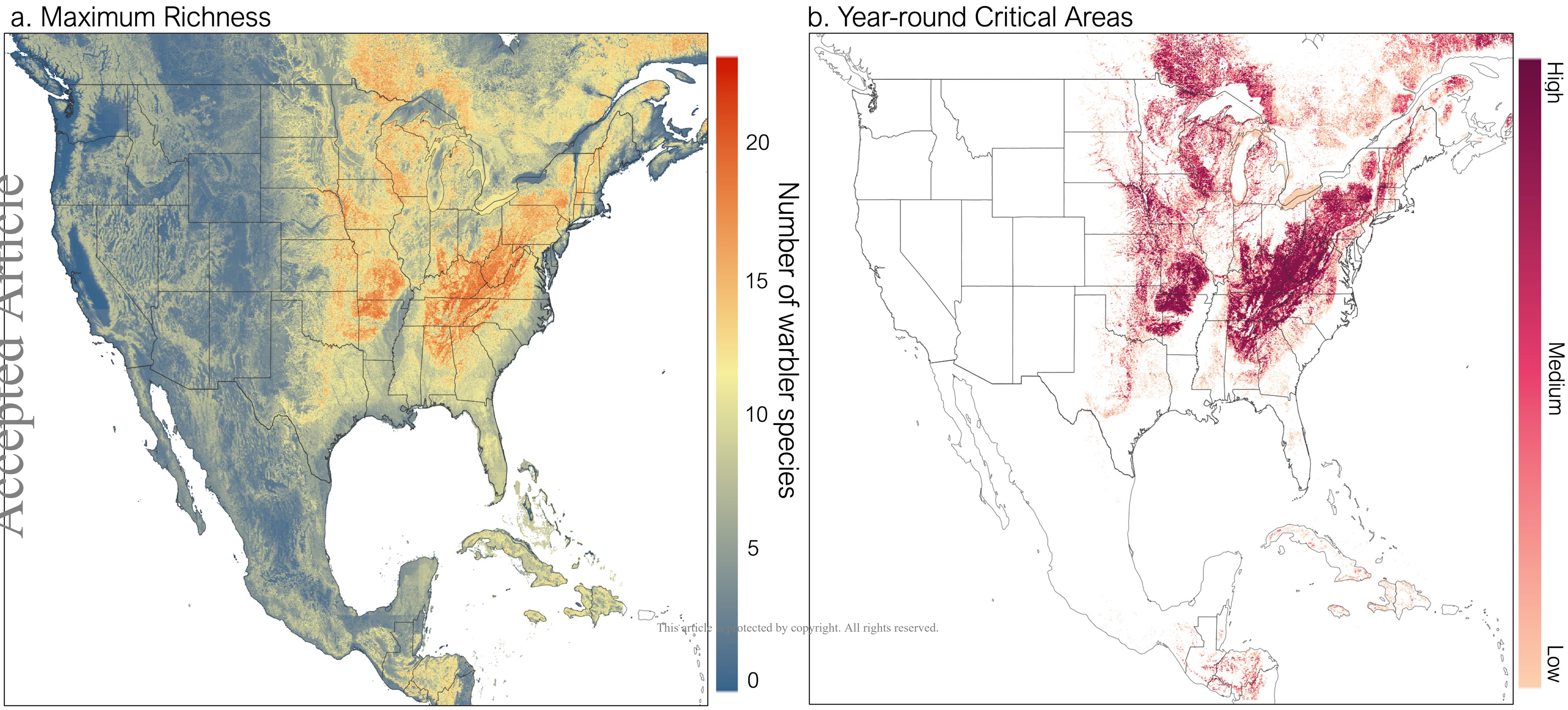
e.g., species distribution map

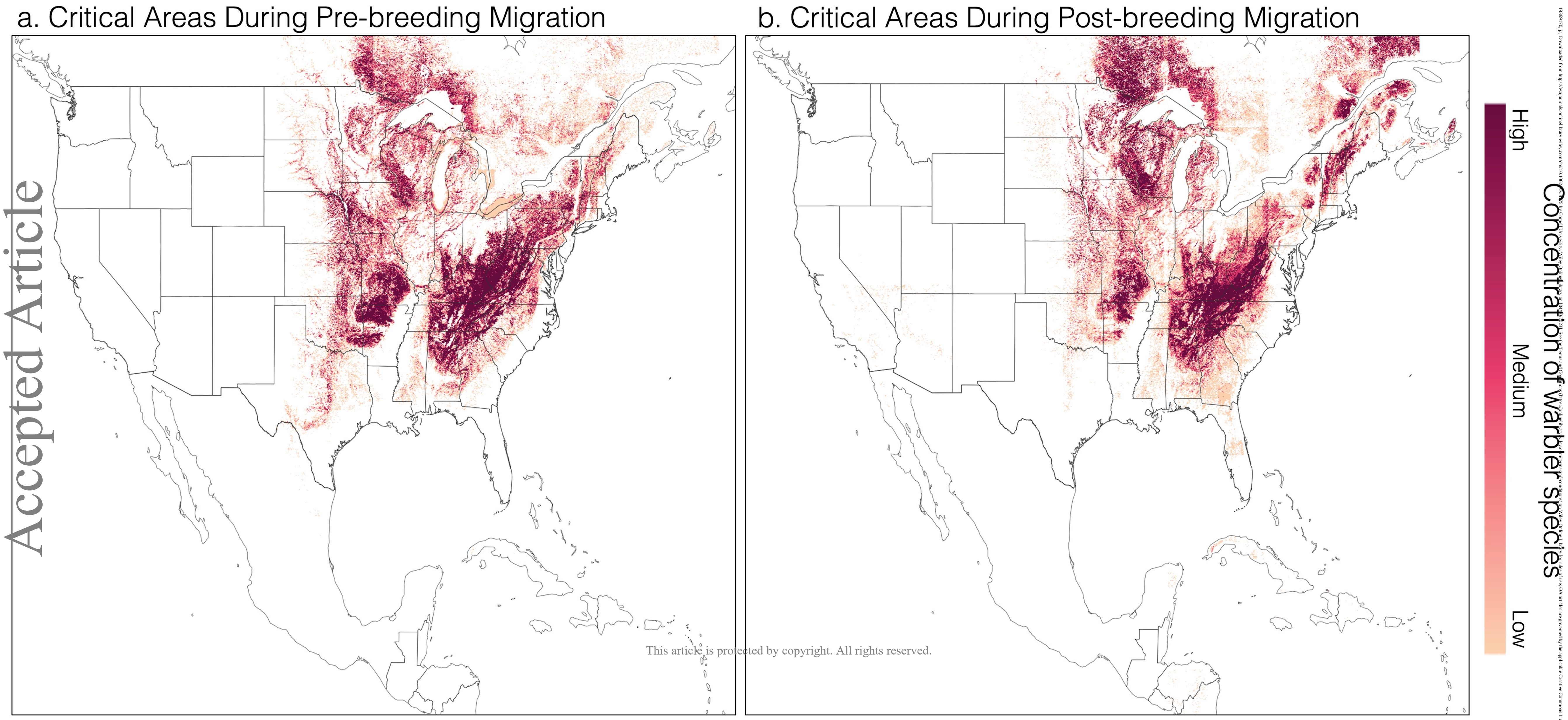


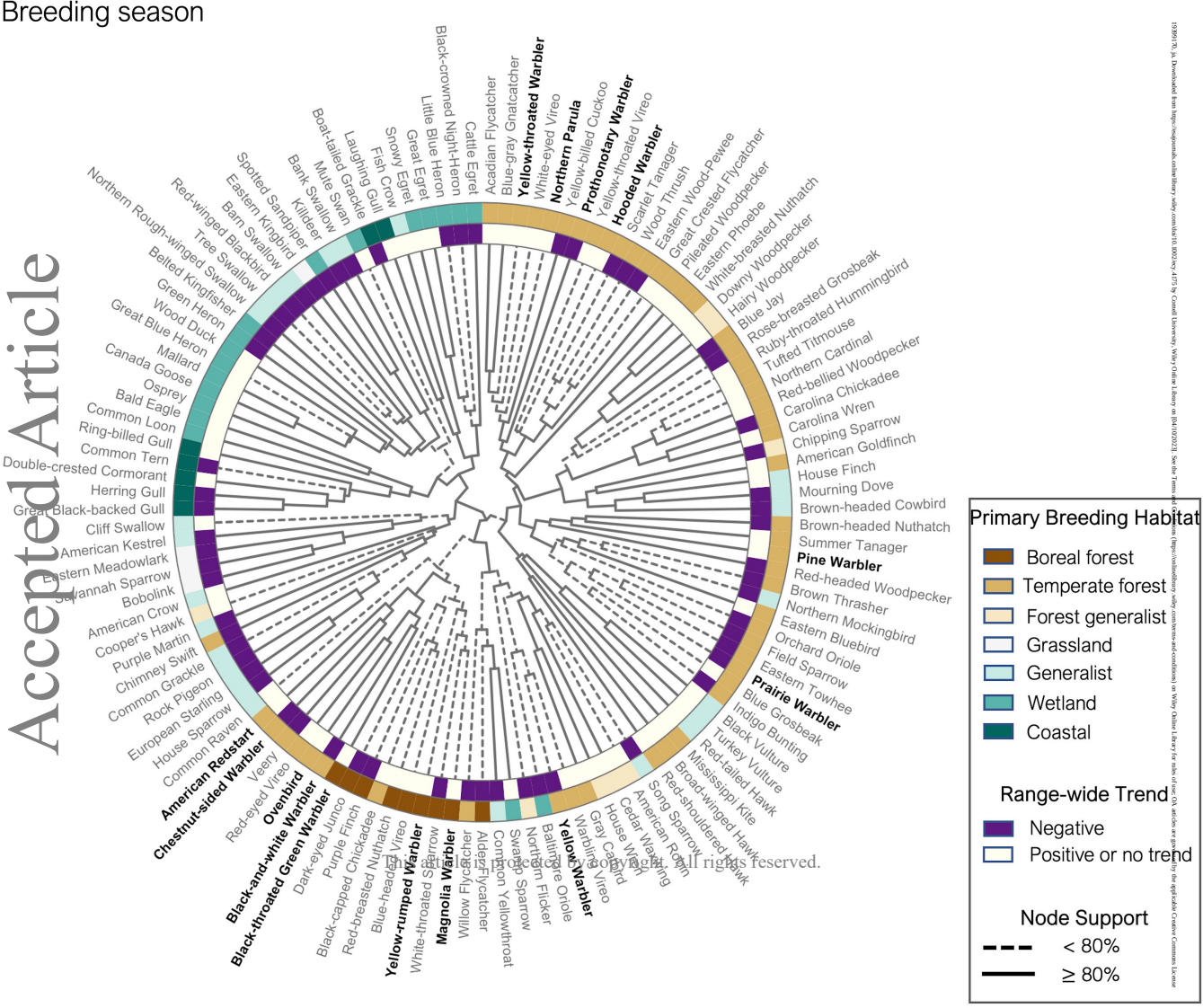
Community-level summaries of occurrence predictions

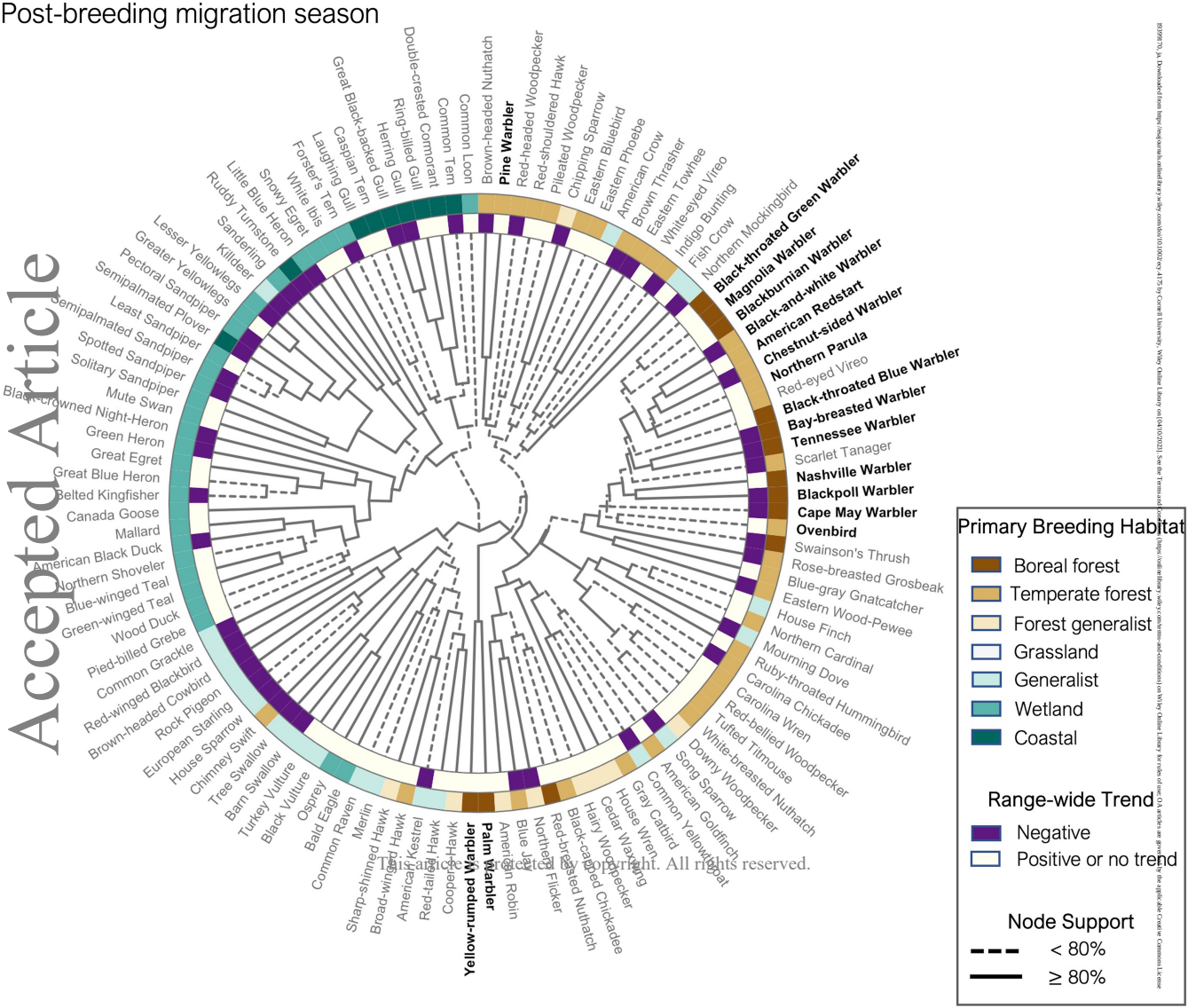
e.g., richness map

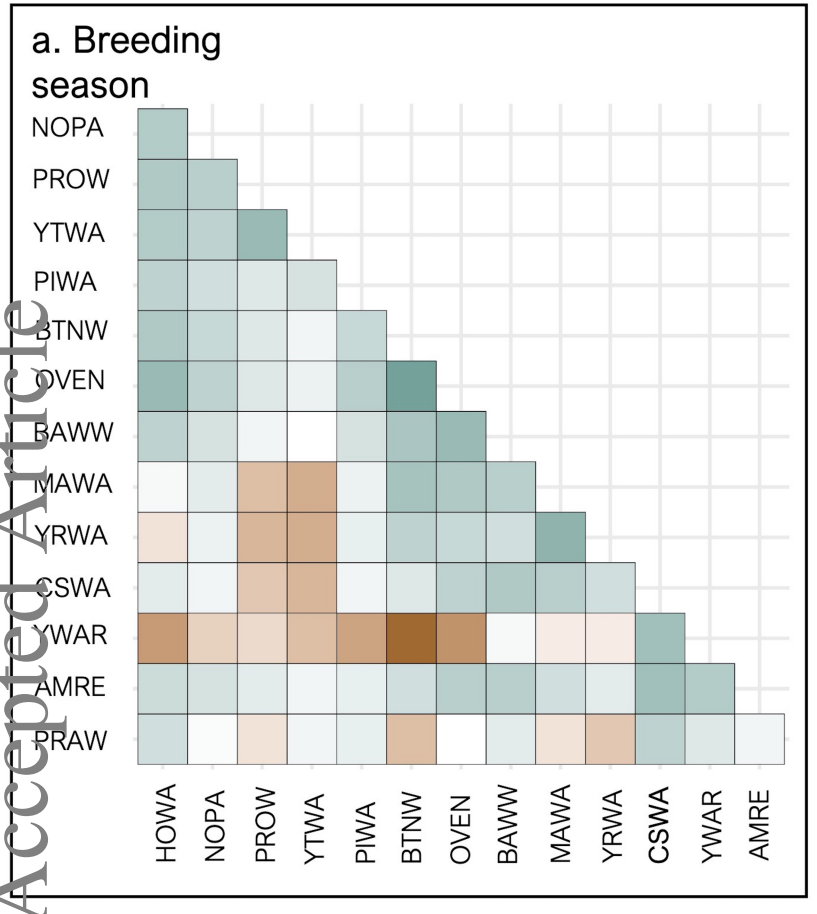


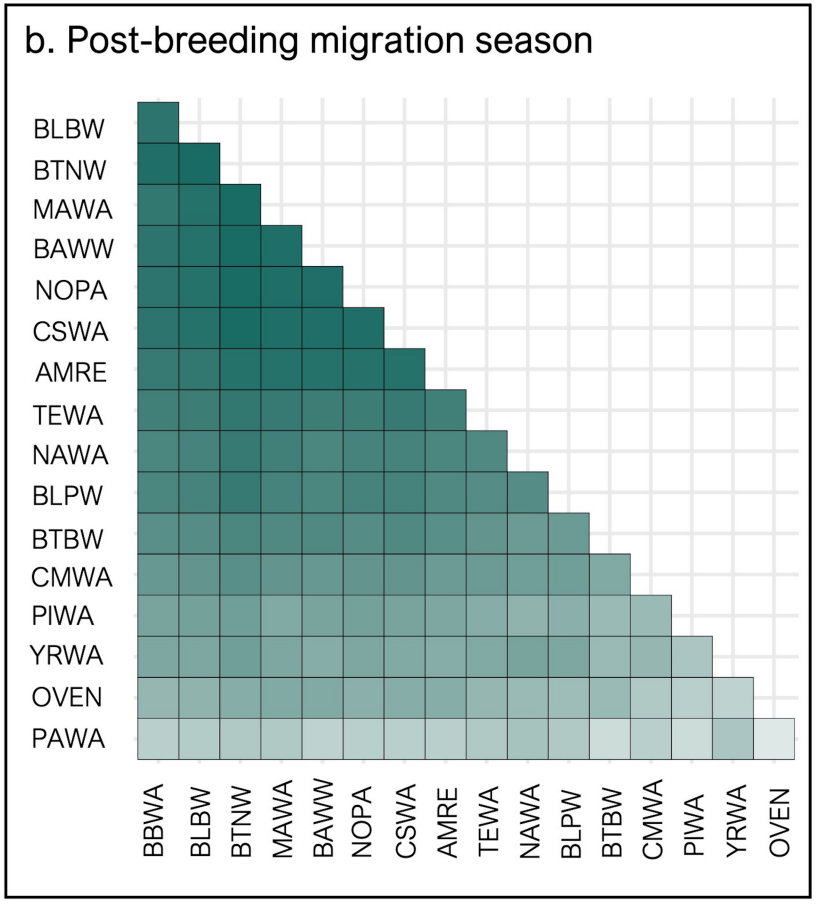


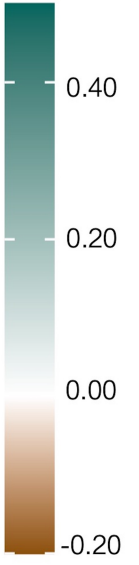












Species Code Key

AMRE	American Redstart
BAWW	Black-and-white Warbler
BBWA	Bay-breasted Warbler
BLBW	Blackburnian Warbler
BLPW	Blackpoll Warbler
BTBW	Black-throated Blue Warbler
BTNW	Black-throated Green Warbler
CMWA	Cape May Warbler
CSWA	Chestnut-sided Warbler
HOWA	Hooded Warbler
MAWA	Magnolia Warbler
NAWA	Nashville Warbler
NOPA	Northern Parula
OVEN	Ovenbird
PAWA	Palm Warbler
PIWA	Pine Warbler
PROW	Prothonotary Warbler
PRWA	Prairie Warbler
TEWA	Tennessee Warbler
YWAR	Yellow Warbler
YTWA	Yellow-throated Warbler
YRWA	Yellow-rumped Warbler