*Incorporating migratory data into species distribution models improves conservation of American Woodcock habitat*

Target: Journal of Applied Ecology

*Abstract*

1. Recent studies (esp 3 billion birds) have made it clear that we need to be expanding migratory bird conservation in the United States. Bird declines are due not only to the loss of breeding and winter habitat, but also migratory stopover habitat, which often has different characteristics.
2. Our most frequent tool for prioritizing conservation areas (species distribution models) is typically only used to model habitat for a single migratory stage at a time. However, decisionmakers will usually only develop a single conservation prioritization plan for a species, and so they need a way to combine migratory and breeding season habitat use into a single prioritization framework.
3. We use a joint residential/migratory species distribution model developed for managing American Woodcock in Pennsylvania to demonstrate how decision makers can incorporate migratory and breeding season habitat models into a single decision-making framework, allowing local land managers to decide whether to prioritize migratory or breeding season habitat management based on an area’s suitability.

*Introduction*

Migration is generally acknowledged to be among the most dangerous parts of the year for most migratory birds (cite tk) and ensuring bird survival during the migratory stage is increasingly becoming a priority for conservation managers (cite tk). One driver of migratory survival is access to stopover habitat, which provides birds with the opportunity to safely rest and refuel between migratory flights (cite tk). Because stopover habitat often has different habitat characteristics than breeding or wintering season habitat (cite tk), and stopovers can occur outside the traditional breeding and wintering ranges, management for stopover habitat is frequently conducted separately from residential habitat management, often as a part of a multi-species management strategy (cite tk). Certain bird species, however, have specific migratory habitat requirements that are not well represented in multi-species habitat management (cite Brian Allen tk). In these cases, single-species management for migratory stopover sites is likely to be required to ensure survival throughout the migratory cycle.

Despite this, migratory stopover habitat management is an infrequent component of single-species management plans (cite tk). Instead, managing agencies prioritize their land management based on breeding or winter season occupancy by that species. While these methods are effective for improving single-season demographic rates, and may increase the populations of birds in those states, this methodology ignores the importance of these regions as migratory corridors, and potential demographic benefits of stopover habitat for birds which might spend the breeding or wintering season in another region.

Here we offer a method for integrating stopover habitat management into an agency’s habitat management framework, through the combination of two migratory and breeding season species distribution models. The exact degree to which habitat in each season should be weighted when considering habitat prioritization is likely an agency-level decision, and this framework is designed to allow the user to designate the weighting that is used in the habitat prioritization. This tool can be used to prioritize areas like wildlife management areas where…

This method, demonstrated through a model of American Woodcock occupancy in the state of Pennsylvania, provides a way for agencies to integrate migratory habitat into their land management decisions.

New outline:

* There is an increasing realization that bird habitat use is important year-round
* In addition to depending on breeding habitat, birds also rely on winter habitat and migratory stopover habitat
* These “off-season” habitats can have implications for bird population stability
  + Survival
  + Carryover effect
* The quality of this habitat, especially stopover habitat, can differ
  + Fire escape
  + Motel
  + 5 star hotel
* Good quality stopover habitat can increase migratory survival, which is low for most migratory birds

----------------------------------------------------------------------------------------------------------------------------

* As we learn more about stopover habitat, there are increasing opportunities to incorporate it into management
* As the need to provide stopover habitat for birds becomes better represented in the literature, there is an increasing need for improved methodologies to incorporate stopover habitat into management plans
* Traditionally, most migratory bird management at the state/provincial agency level has been based on breeding season data
  + Based on SGS and similar datasets
  + Modeled in SDMs
* As more migratory stopover data becomes available, we can now incorporate this new data into management prioritization frameworks
* Here, we demonstrate a methodology to do so through a blended migratory and residential SDM tool that allows users to manually decide their seasonal priorities for management

----------------------------------------------------------------------------------------------------------------------------

* We demonstrate the use of this methodology to reconcile migratory and residential datasets for American Woodcock in the state of Pennsylvania
* American Woodcock are a short distance migrant, the majority of whose migratory stopover habitat overlaps with their breeding and wintering range (eBird range figure)
* Therefore, almost all American Woodcock managers are going to be concerned with two or more seasons of woodcock use
* Some differences have been noted between woodcock breeding and migratory habitat in past studies (Brian Allen citation goes here)
* Drier, more urbanized habitats
* Habitat during the other two seasons is different enough that management for breeding habitat alone likely won’t be enough to protect stopover habitat
* To allow for the joint modeling of breeding and migratory stopover habitat of woodcock in Pennsylvania, we designed separate breeding season and migratory species distribution models
* The breeding season species distribution model used a more traditional data source (SGS and state surveys), while the migratory stopover species distribution model used stopover locations gained from an extensive collaborative effort to deploy satellite transmitters throughout the Eastern Woodcock management region (cite management region tk)
* By combining these products into an app with user controlled weighting, we demonstrate how migratory and residential models can be used to drive dialogue about the prioritization of migratory and residential habitat

**Methods**

*Breeding season species distribution model*

Breeding season species distribution models are the most frequently used tool for planning bird conservation (cite past work tk). As such, they can be created using most traditional survey datasets. For the Pennsylvania example, we used federal Woodcock Singing Ground Surveys and similar state-level surveys conducted by the Pennsylvania Game Commission. SGS surveys consist of tk mile long survey routes consisting of tk evenly-spaced points, with presence-absence determined at each point based on whether male displays were visible during a tk minute interval at dusk. SGS survey routes were randomly distributed through the state? In 19XX? And each route has been run annually since tk. Pennsylvania Game Commission woodcock surveys were run using the same methodology, but their routes were intentionally placed near state gamelands or in areas where managers believed woodcock occupancy was likely. We converted state and federal survey data from 20tk-20tk to a presence-absence dataset by marking each survey point as present if there were woodcock observed at that point at least once during the 5 year interval, and absent if they were not. These presence-absence locations were then used as the response variable in the breeding season species distribution model. The explanatory variables in the species distribution model included several suites of variables presumed to be relevant to woodcock habitat. These included variables representing land cover (land use/land cover (NLCD 20tk), forest successional class (LANDFIRE 20tk)), geography (Elevation (source tk?), Slope (source tk?), Ecoregions (EPA level 3 tk)), and moisture (Drainage (source?), Topographic wetness index (cite methodology tk)). We additionally added landscape metrics from the landscapemetrics R package representing landscape composition (% forest, % agricultural, % developed) and configuration (aggregation index, cohesion, edge density). We ran each of these landscape metrics at multiple scales, represented by radii from 90m pixels. The radii used were 500m, 1km, 5km, and 10km. To generate these landscape metrics, we cropped a binary forest/nonforest layer to the extent of a circle of the given radius from each 90m pixel, and then ran the appropriate landscapemetrics function on each cropped raster. We then assigned the output value from the landscapemetrics function to the appropriate 90m pixel. These explanatory variables were then used in the species distribution model. The species distribution model used a random forest classifier designed for clustered data (MixRF package) to predict whether survey points would be present or absent. The survey route id was used as the clustering variable to compensate for autocorrelation between points on the same survey route. We also included federal/state as an explanatory variable in the analysis to account for bias in the state survey route allocation. To avoid overwhelming the model with highly correlated variables, we elected to use a backwards variable-selection approach (VSURF package tk) to determine which set of variables should be used in the final model. This approach uses a three-step process, first to eliminate variables that have little importance to prediction, second to remove variables that are not important for prediction, and third to eliminate variables that are redundant. We calculated the AUCs for each of the three steps, and used the set of variables with the highest AUC in the final model. We used a k fold cross validation approach with 10 folds to evaluate our final model, using 90% of the data in each fold as a training dataset and the remaining 10% as a testing dataset. We averaged together the AUCs calculated for each of the 10 folds to find the AUC for the final model. We then calculated a predictive layer using the models calculated for each of the 10 folds and averaged those layers together to create a final predictive layer for the breeding season model. For the predictive layer, the survey type predictive variable was set to “federal” to exclude bias resulting from the state survey locations

*Migratory season species distribution model*

While breeding season models can be created using traditional survey methods, if those surveys are conducted during the breeding season, getting migratory stopover information is a little bit harder. Thankfully, the emergence of widespread GPS technology is making it easier than ever to track migratory stopover sites (cite tk). We used GPS data from the EWMRC to designate woodcock migratory stopover sites throughout the state of Pennsylvania. Sentence tk. These sites were used as present locations. To create pseudo-absence locations, we randomly allocated 10,000?tk locations throughout the state. These two sets of data were combined to form the response variable in the model. The explanatory variables were the same as used in the breeding season model. We then used these data to create a species distribution model using a random forest classifier (SDMTune package). We used a k fold cross validation approach with 10 folds to evaluate our final model, using 90% of the data in each fold as a training dataset and the remaining 10% as a testing dataset. We averaged together the AUCs calculated for each of the 10 folds to find the AUC for the final model. We then calculated a predictive layer using the models calculated for each of the 10 folds, and averaged those layers together to create a final predictive layer for the migratory season model.

*Multi-season predictive layer*

To facilitate user choice in how and where to prioritize migratory and residential habitat, we generated a series of combined layers with different weights assigned to the migratory and breeding season layers. 0% migratory 100% breeding, 10% migratory 90% breeding, etc. We used an interactive application (e.g. shinyapps) to distribute this information and facilitate user weighting.

*Results*

The residential model achieved an AUC of tk, after getting to the tk stage of variable selection. The migratory model had a slightly lower AUC (tk), with the full model being the most informative stage (table tk). For the residential model, selection for landscape variables primarily occurred at 5km and 10km scales; in the most informative residential model, there were no important variables in either the landscape (500m) or moisture suite. The residential model had strong, non-linear relationships with a few landscape variables, while the migratory model had weak relationships with many variables (Figure tk). During the migratory season, woodcock were far more tolerant of developed and fragmented landscapes (within 10 km). These two models produced results that were largely uncorrelated (Fig. tk), with the residential model producing a much smoother surface (Fig. tk) than the migratory model (Fig. tk).

New results outline:

* AUCs from each model
* Significant covariates for each model
  + Significant variable table
* Model relationships with covariates
  + The best migratory model had weak relationships with many covariates, while the best residential model had strong relationships with a few covariates
  + Fig: model relationships
* Lack of correlation between the models
  + Fig: Comparative model maps
  + Different “scale of smoothness”
  + For discussion: manage for residential at a regional scale, migratory at a site scale
* Regional differences between models
  + Ecoregions 1 and 6 had way more migratory than residential habitat
    - For discussion: Maybe this means that we deprioritize those regions for woodcock management, or maybe we just focus on migratory habitat alone when managing in those regions
  + Fig: Ecoregion box & whiskers

|  |  |  |
| --- | --- | --- |
| Suite | Migratory | Residential |
| Landscape (500m) | Aggregation Index, Cohesion, Edge Density, % Forest, % Agricultural, % Developed |  |
| Landscape (1km) | Aggregation Index, Cohesion, Edge Density, % Forest, % Agricultural, % Developed | % Agricultural |
| Landscape (5km) | Aggregation Index, Cohesion, Edge Density, % Forest, % Agricultural, % Developed | Cohesion, % Forest, % Agricultural, % Developed |
| Landscape (10km) | Aggregation Index, Cohesion, Edge Density, % Forest, % Agricultural, % Developed | Aggregation Index, Cohesion, % Agricultural, % Developed |
| Land Cover | Forest, Successional Class |  |
| Geography | Elevation, Slope, Ecoregions | Elevation, Ecoregions |
| Moisture | Drainage, Topographic Wetness Index |  |

Table tk. Variables selected via backwards variable selection in VSURF for the migratory and residential models. The migratory model employs the full set of variables, while the residential model uses a subset of variables inclined towards coarse resolution landscape variables.

![Chart, scatter chart

Description automatically generated]()

Figure tk. Comparison of relationships between landscape variables and habitat suitability for migratory and residential models. During the residential season, woodcock habitat suitability is highest in highly aggregated landscapes with ~75% forest and ~25% agricultural cover. During the migratory season, however, woodcock become far more tolerant of landscapes that are unsuitable during the residential season, including landscapes with higher proportions of developed cover.

Map

Description automatically generated

Figure tk. The lack of correlation between residential and migratory habitat suitability (measured on a percentile scale) indicates that single-season species distribution models will likely not reflect the suitability of a landscape for woodcock during other portions of the year.

*Discussion*

New discussion outline:

* The models produced
  + Wide scale residential, finer scale migratory
  + Tolerance of developed areas
  + Ecological implications of differing selection during these seasons
* Incorporation into a tool
  + Taken on their own, these tools have interesting ecological implications
    - However, our goal was to take them together and determine whether, when combined, they would provide more valuable insights for management
  + We blended the model using a Shiny application (screenshot as a figure), which allowed the user to determine in 10% intervals how much they wanted to value residential vs. migratory habitat
  + In addition to the blending, the Shiny app enabled users (presumably managers) to examine how the layers overlayed with their gamelands, and the percent of their gamelands which fell into high (upper 33rd percentile of all pixels) or medium (66th to 33rd percentile of all cells) quality categories
  + By tailoring the application to the decision-making processes that we believe managers are going to make, we hope to increase use of this tool
  + By increasing use of this tool, we hope to drive folks towards making more decisions that consider whether migratory data should be incorporated
* The weighting decision
  + Could be used to prioritize regions, or prioritize within-region focuses
  + Migratory corridors
  + Breeding season population goals
* Tool limitations
  + This tool was designed primarily for agency managers, and in early testing we noticed some important stipulations for our results
* In what circumstances should someone else do this?
  + Migratory birds for whom stopover habitat is presumed to be a limiting factor