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

Target: Journal of Applied Ecology

*Introduction*

Animals frequently use different habitats in certain stages of their life or annual cycle. This principle has been a component of wildlife management since Leopold (1933), and multi-stage habitat use has been incorporated into the management of animals ranging from white-tailed deer (Voigt et al. 1997) to greater sage-grouse (Fedy et al. 2012). These insights have been a popular focus of studies on migratory birds (Norris 2005, Marra et al. 2015), which postulate that conserving migratory birds throughout their full annual cycle is necessary to slow population declines. These concerns are reflected in recent migratory bird management plans which advocate for international conservation of breeding, stopover, and wintering habitat (Rosenberg et al. 2016). This range wide approach to management is an important step forward in migratory bird conservation, but conserving habitat during the full annual cycle is a local issue as much as it is an international one. Much of a species’ migratory range, for example, can overlap with its breeding and wintering ranges, especially for short distance migrants or species with widespread breeding or non-breeding distributions. Most regions that are typically considered to be part of a bird’s breeding range, for example, might be more accurately considered to be in the joint breeding and migratory range, and full annual cycle conservation in these areas would require consideration of both the breeding and migratory habitat requirements of that species.

During the migratory season, bird habitat use occurs at migratory stopover sites, which Mehlman (2005) defines as any place where a bird can land and survive until the next migratory flight. However, those sites which have more resources can either improve or be less detrimental to a migratory bird’s condition and can increase the bird’s probability of successfully completing migration. As migratory survival is believed to be a limiting factor for many species of birds (Sillett and Holmes 2002, Rockwell et al. 2017, Robinson et al. 2020), conserving migratory stopover sites is assumed to be an important step in slowing bird declines (Faaborg et al. 2010). Habitat selection for stopover sites is frequently different from habitat selection during the breeding and wintering seasons (Allen et al. 2020, Stanley et al. 2021), which can result in birds using areas during migratory stopovers that they normally would not reside in during other portions of the year. These differences in seasonal space use can now be tracked with some ease due to recent technological advances in radar (Larkin and Diehl 2012), animal tracking (Bridge et al. 2011), and eBird data (Lin et al. 2020). These technologies can allow for the creation of migratory stopover habitat distribution models, which can be compared and used in conjunction with breeding and wintering habitat distribution models in much the same way that multi-season habitat use is examined for other wildlife species. Here we show an example of how this migratory stopover data can be used in combination with breeding season data to enable multi-seasonal management for a common migratory bird species in its breeding range.

Our case study focuses on American Woodcock (*Scolopax minor*) in the state of Pennsylvania. American Woodcock are a short distance migrant, most of whose migratory stopover habitat overlaps with their breeding and wintering range (eBird range figure tk). Pennsylvania provides both breeding habitat and stopover habitat for birds migrating to and from the northern extent of their range. However, Allen et al.(2020) suggests that woodcock stopover habitat often differs from habitat used during other seasons. Therefore, management for breeding season habitat alone may not be enough to protect woodcock stopover habitat. To determine whether additional management for stopover habitat is necessary, we first model migratory and breeding season woodcock habitat in Pennsylvania and examine the overlap between the two. If there is significant overlap between migratory and breeding season habitat, then existing management strategies for woodcock in Pennsylvania (which have been based on breeding season data) may be adequate to protect migratory stopover habitat. If not, then additional management may need to take place to additionally protect stopover habitat. We then demonstrate one possible method for incorporating stopover habitat into a habitat management framework by combining migratory and breeding season habitat suitability models into a single habitat prioritization tool. This tool allows users to manually decide their priorities for migratory and breeding season management and identify areas that provide woodcock habitat during both seasons, allowing users to incorporate full annual cycle conservation into a straightforward conservation prioritization framework.

**Methods**

*Breeding season species distribution model*

We created a breeding season species distribution model for woodcock in the state of Pennsylvania using federal Woodcock Singing Ground Surveys (Seamans and Rau 2020) and similar state-level surveys conducted by the Pennsylvania Game Commission. Singing Ground Surveys, which were originally established in 1968, consist of 5.76 km survey routes consisting of 10 evenly spaced points, with presence-absence determined at each point based on whether male displays were visible during a 2 minute interval shortly after dusk. Singing Ground Survey routes are randomly distributed through the state, and the same routes are run annually when possible (Clark 1970). Pennsylvania Game Commission woodcock surveys are run using the same methodology, but their routes are placed near state gamelands or in areas where managers believe woodcock occupancy is likely. We converted state and federal survey data from 2016 — 2020 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 use/land cover (Jin et al. 2019), forest successional class (LANDFIRE 20tk), elevation (source tk?), slope (source tk?), EPA level 3 ecoregions (Omernik and Griffith 2014), soil drainage (source?), and topographic wetness index (cite methodology tk). We additionally added landscape metrics from the landscapemetrics R package (Hesselbarth et al. 2019) 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/non-forest layer to the extent of a circle of the given radius from each 90m pixel, and then ran the appropriate function from the landscapemetrics package on each cropped raster. We then assigned the output value from the function to the appropriate 90m pixel. These were then used as explanatory variables in the species distribution model. The species distribution model used a random forest classifier designed for clustered data (Wang and Chen 2016) to predict whether woodcock would be present or absent at survey points. The survey route id was used as the clustering variable to compensate for autocorrelation between points on the same survey route. We also included the survey type 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 (Genuer et al. 2019) to determine which set of variables should be used in the final model. This approach uses a three-step process, with each step producing a more parsimonious model. The first step eliminates variables that have little importance to prediction, the second step removes variables that have some relevance but are not critical for prediction, and third step eliminates variables that are redundant. We calculated the AUCs for the models produced by each of the three steps and used the set of variables produced by the step 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 placement of state survey locations near state gamelands.

*Migratory season species distribution model*

We used GPS data from the Eastern Woodcock Migration Research Cooperative (EWMRC) to designate woodcock migratory stopover sites throughout the state of Pennsylvania. The EWMRC is a collaboration of 34 federal, state, provincial, non-profit, and university partners throughout the United States and Canada that has been deploying transmitters on woodcock since 2017 throughout the eastern portion of their range (www.woodcockmigration.org). We captured birds at 34 sites in Quebec, Ontario, Nova Scotia, Maine, Vermont, New York, Rhode Island, Pennsylvania, Maryland, West Virginia, Virginia, North Carolina, South Carolina, Georgia, Alabama, and Florida. We captured woodcock using mist nets during their morning and evening flights (cite Sheldon 1960 tk), and on their night roosts using spotlights and dip nets (Rieffenberger and Kletzly 1966, McAuley et al. 1993). We attached 4g, 5g, and 6.3g PinPoint GPS Argos tags (Lotek tk) to captured woodcock. These tags record locations at 12 — 60m accuracy depending on cover type (Berigan, unpublished data), and were programmed to record locations every tk — tk days at tk pm or tk pm Eastern Standard Time, outside of the woodcock’s nocturnal flight period. Transmitters never exceeded tk% of a bird’s body weight, and all bird capture was conducted in accordance with best practices for wildlife research (Fair et al. 2010).

Transmitters were deployed on 463 woodcock from Fall 2017 — Spring 2021, and tk of those individuals recorded a total of tk GPS locations at migratory stopovers in Pennsylvania. Consecutive locations from the same individual that were within X km of each other were considered to be part of the same stopover, and all but the first of these locations were removed from the analysis. These stopover sites were used as present locations in the migratory season species distribution model. 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 (Vignali et al. 2020). 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 breeding season habitat, we created a Shiny application (cite RStudio tk) that would allow users to manually assign weights to each and combine them into a single layer (Figure tk: screenshot of the Shiny app). The user can choose the weighting of each layer in 10% increments (ex. 20% migratory and 80% breeding season. Because our application was targeted at users in the Pennsylvania Game Commission, the application also shows the comparative suitability of Pennsylvania state gamelands for each weighted layer. We used four metrics for comparing the habitat suitability of gamelands. The first was average pixel value, which favored small gamelands which were predominantly composed of woodcock habitat. The second was average pixel value multiplied by the acreage of the gameland, which we titled landscape suitability index. Landscape suitability index favored large gamelands which might not be entirely composed of woodcock habitat but might contain a large amount of woodcock habitat in aggregate. The final two metrics were the percent of the gameland which was of high quality, defined as all cells greater than the 33rd percentile of all pixels, or of medium quality, defined as all cells between the 66th and 33rd percentile. This layer is publicly accessible at woodcock.shinyapps.io/woodcock\_weighted\_habitat.

*Results*

The most informative residential model was the “tk” step model, for which all autocorrelated variables had been removed. This produced a model with an AUC of 0.tk, which was heavily informed by landscape variables at the 5 and 10 kilometer scales (Table tk). No variables at the finest landscape scale (500m) or in the suite of moisture variables were included in the most informative model. While random forest models do not provide coefficients that can be used to determine the impact of each covariate on the model, graphs of habitat suitability for each covariate show strong, non-linear relationships with several of the most informative variables. Suitability was highest for landscapes at the 10km scale with 0 – 25% developed land area, 0 – 50% agricultural land area, and aggregation index values of 80 – 100. At the 5km scale, the residential model also showed high suitability for landscapes with 30 – 100% forest cover (Fig. tk).

The most informative migratory model was the full model, including all landscape, land cover, geographic, and moisture covariates (Table tk). This produced a model with an AUC of 0.tk. Likely due to the wide array of covariates influencing the model, individual covariate graphs do not show clear visual patterns between migratory habitat suitability and any one covariate. However, the migratory model showed greater tolerance for developed and dis-aggregated landscapes at a 10km scale than the residential model (Fig. tk). The two models are also distinguished by the scale at which covariates influence habitat suitability. While the most informative residential model was not influenced by any landscape covariates at the 500m scale, and only 1 landscape covariate at the 1km scale, the most informative migratory model included all available small-scale landscape covariates. This caused the migratory model to provide predictions at a finer spatial scale than the residential species distribution map (Fig. tk).

Despite the difference in the resolution of predictions, there were some regional differences between the two layers (Fig tk. Ecoregion bar chart and ecoregion map, with asterisks for p-values). Residential suitability was significantly lower than migratory suitability for woodcock in Ecoregions 1, 6, and 11 (add t-test tk). Conversely, suitability was significantly higher in the residential season than the migratory season in Ecoregion 2. In other Pennsylvania ecoregions (tk), migratory and residential suitability was not significantly different.

|  |  |  |
| --- | --- | --- |
| 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*

Under a traditional migratory bird conservation strategy, the Pennsylvania Game Commission would manage solely for breeding season habitat for this species. Here we show that such a strategy would have excluded most of the stopover habitat in the state. The differences between migratory and breeding habitat that we see here are likely to be replicated in other species, due to shared trends in tolerance for developed landscapes. Therefore, management strategies that account for both breeding and migratory habitat (as well as wintering habitat when appropriate) are likely to be necessarly to ensure that migratory birds are protected during all stages of their life cycle.

Modeling both breeding and migratory habitat is the first step

Multi-scalar, as the amount of overlap between migratory and breeding habitat is likely to change with scale

Second, incorporating these two models into a combined framework.

Tools like Shiny apps have the benefit of interactivity

Use of migratory data and tools in building partnerships

Wrap-up

These models produced migratory and residential layers that often predicted very different habitat use between the two seasons. The resolution of these predictions differed between the migratory and residential seasons, likely due to seasonal changes in the scale of selection. Someone (tk) has noticed changes in the scale of selection between the breeding and migratory seasons, which would cause an effect similar to the one we observed. We also noted that woodcock were far more likely to use developed landscapes during the migratory season than the residential season, and heavily utilized some ecoregions (tk) which are urbanized and has low residential habitat suitability. This fits with other research (tk) which suggests that many specialist bird species become generalists during migration and are more tolerant of certain habitat factors which would otherwise repel them during other seasons. Overall, our results show that woodcock use regions with a mix of forest and agricultural cover during the residential season (corroborating past research tk), but woodcock during the migratory season often choose sites based on a variety of local habitat characteristics.

The fact that woodcock habitat use changes drastically between seasons indicates that a “breeding only” strategy will not be enough to preserve woodcock stopover habitat. However, this does range the question of how best to conserve these habitats. To facilitate the incorporation of both migratory and residential habitat suitability into a single management framework, we used a Shiny tool to combine the two layers and left it up to the user to decide how these layers should be weighted. Leaving this decision to the user serves two purposes. First, it encourages discussion within the management agency regarding the agency’s priorities in conserving residential and stopover habitat, to meet both the objectives of their stakeholders and to achieve a stable woodcock population. Second, by allowing the user to make the weighting decision, we allow users to determine whether weighting should change based on regions. Some regions may be valuable as both stopover and residential habitat; in such a case, a balanced user weighting of the two layers may be a good way to determine which gamelands provide both types of habitat. Other regions (such as ecoregion tk) might provide only migratory habitat, and so a user weighting that favors migratory habitat might be best employed to determine where woodcock management would be best applied within that region. In addition, a map of migratory corridors for woodcock could be used to determine where woodcock are most likely to fly over the state and be used to determine the regions in which migratory habitat management would be most important (cite some component of my work tk).

As in any habitat analysis, we should be cautious about the spatial scale at which these results are applicable. For example, we found here that there was limited overlap between migratory and residential habitat at a landscape scale. However, Allen et al. (20tk) found considerably more overlap between migratory and residential habitat at a local scale. In this case, we may be able to identify areas of overlap between residential and migratory woodcock habitat at a landscape scale and assume that habitat managed for woodcock in these areas will support both breeding and migratory birds. However, this will not be the case for all migratory bird species, and multi-scalar studies into how selection differs between migratory and residential seasons will be necessary before appropriate stopover management can be implemented.

Agencies should also be prepared for the possibility that landscape-scale habitat suitability models for migratory habitat may show that some migratory stopover sites are focused in areas that aren’t traditionally managed for wildlife habitat. Our woodcock migratory stopover habitat model, for example, shows that several urban areas (including the greater Philadelphia area) are hotspots for migratory stopover activity. However, there may still be conservation opportunities in these areas through collaboration with other entities that maintain urban greenspace, such as parks departments and private landowners (cite tk). Partnerships between these organizations may provide the opportunity for wildlife management agencies to extend their reach into areas that are not traditionally managed for wildlife but may be important migratory stopover hotspots.

We hope that this first step into incorporating stopover habitat management into woodcock management plans for Pennsylvania provides an example of the type of methods that can be used for incorporating stopover habitat management into migratory bird conservation. This type of approach should be especially useful for species that are presumed to be limited by migratory survival.

Literature cited:

Allen, B. B., D. G. McAuley, and E. J. Blomberg. 2020. Migratory status determines resource selection by American Woodcock at an important fall stopover, Cape May, New Jersey. The Condor 122:1–16.

Bridge, E. S., K. Thorup, M. S. Bowlin, P. B. Chilson, R. H. Diehl, R. W. Fléron, P. Hartl, R. Kays, J. F. Kelly, W. D. Robinson, and M. Wikelski. 2011. Technology on the Move: Recent and Forthcoming Innovations for Tracking Migratory Birds. BioScience 61:689–698. Oxford Academic.

Clark, E. R. 1970. Woodcock status report, 1969.

Faaborg, J., R. T. Holmes, A. D. Anders, K. L. Bildstein, K. M. Dugger, S. A. Gauthreaux, P. Heglund, K. A. Hobson, A. E. Jahn, D. H. Johnson, S. C. Latta, D. J. Levey, P. P. Marra, C. L. Merkord, E. Nol, S. I. Rothstein, T. W. Sherry, T. Scott Sillett, F. R. Thompson, N. Warnock, S. A. Gauthreaux Jr, P. Heglund, K. A. Hobson, A. E. Jahn, D. H. Johnson, others, S. A. Gauthreaux, P. Heglund, K. A. Hobson, A. E. Jahn, D. H. Johnson, S. C. Latta, D. J. Levey, P. P. Marra, C. L. Merkord, E. Nol, S. I. Rothstein, T. W. Sherry, T. Scott Sillett, F. R. Thompson, and N. Warnock. 2010. Conserving migratory land birds in the New World: Do we know enough? Ecological Applications 20:398–418. Wiley Online Library.

Fair, J. M., E. Paul, J. Jones, A. B. Clark, C. Davie, and G. Kaiser. 2010. Guidelines to the use of wild birds in research. Washington, D.C.

Fedy, B. C., C. L. Aldridge, K. E. Doherty, J. L. Beck, B. Bedrosian, M. J. Holloran, G. D. Johnson, N. W. Kaczor, C. P. Kirol, C. A. Mandich, D. Marshall, G. Mckee, C. C. Swanson, and B. L. Walker. 2012. Interseasonal Movements of Greater Sage-Grouse, Migratory Behavior, and an Assessment of the Core Regions Concept in Wyoming. The Journal of Wildlife Management 76:1062–1071.

Genuer, R., J.-M. Poggi, and C. Tuleau-Malot. 2019. VSURF: Variable Selection Using Random Forests.

Hesselbarth, M., M. Sciaini, K. With, … K. W.-, and undefined 2019. 2019. landscapemetrics: an open‐source R tool to calculate landscape metrics. Wiley Online Library 42:1648–1657. Blackwell Publishing Ltd.

Jin, S., C. Homer, L. Yang, P. Danielson, J. Dewitz, and C. Li. 2019. Overall methodology design for the United States national land cover database 2016 products. Remote Sensing.

Larkin, R. P., and R. H. Diehl. 2012. Radar Techniques for Wildlife Management. Pages 319–348 *in* N. J. Silvy, editor. The Wildlife Techniques Manual: Volume 1: Research. Seventh edition. John Hopkins University Press, Baltimore, MD, USA.

Leopold, A. 1933. Game management. Scribner, New York, New York, USA.

Lin, H. Y., R. Schuster, S. Wilson, S. J. Cooke, A. D. Rodewald, and J. R. Bennett. 2020. Integrating season-specific needs of migratory and resident birds in conservation planning. Biological Conservation 252:108826. Elsevier.

Marra, P. P., E. B. Cohen, S. R. Loss, J. E. Rutter, and C. M. Tonra. 2015. A call for full annual cycle research in animal ecology. Biology Letters 11. The Royal Society.

McAuley, D. G., J. R. Longcore, and G. F. Sepik. 1993. Techniques for research into woodcocks: experiences and recommendations. Pages 5–11 *in*. Proceedings of the Eighth American Woodcock Symposium, US Fish and Wildlife Service Biological Rep. Volume 16.

Mehlman, D., S. Mabey, D. Ewert, C. Duncan, B. Abel, D. Cimprinch, R. Sutter, and M. Woodrey. 2005. Conserving stopover sites for forest-dwelling migratory landbirds. The Auk 122:1281–1290.

Norris, D. R. 2005. Carry-over effects and habitat quality in migratory populations. Oikos 109:178–186. John Wiley & Sons, Ltd.

Omernik, J. M., and G. E. Griffith. 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environmental Management 54:1249–1266.

Rieffenberger, J. C., and R. C. Kletzly. 1966. Woodcock night-lighting techniques and equipment. WH Goudy, compiler. Woodcock research and management 33–35.

Robinson, R. A., C. M. Meier, W. Witvliet, M. Kéry, and M. Schaub. 2020. Survival varies seasonally in a migratory bird: Linkages between breeding and non-breeding periods. Journal of Animal Ecology 89:2111–2121.

Rockwell, S. M., J. M. Wunderle, · T Scott Sillett, C. I. Bocetti, D. N. Ewert, D. Currie, J. D. White, · Peter, P. Marra, and S. Org. 2017. Seasonal survival estimation for a long-distance migratory bird and the influence of winter precipitation. Oecologia 183:715–726.

Rosenberg, K. V, J. A. Kennedy, R. Dettmers, R. P. Ford, D. Reynolds, J. D. Alexander, C. J. Beardmore, P. J. Blancher, R. E. Bogart, G. S. Butcher, and others. 2016. Partners in Flight landbird conservation plan: 2016 revision for Canada and continental United States. Partners in Flight Science Committee.

Seamans, M. E., and R. D. Rau. 2020. American woodcock population status, 2020. Laurel, MD, USA.

Sillett, T., and R. T. Holmes. 2002. Variation in survivorship of a migratory songbird throughout its annual cycle. Journal of Animal Ecology 71:296–308.

Stanley, C. Q., M. R. Dudash, T. B. Ryder, W. G. Shriver, K. Serno, S. Adalsteinsson, and P. P. Marra. 2021. Seasonal variation in habitat selection for a Neotropical migratory songbird using high‐resolution GPS tracking. Ecosphere 12:e03421. John Wiley & Sons, Ltd.

Vignali, S., A. G. Barras, R. Arlettaz, and V. Braunisch. 2020. SDMtune: An R package to tune and evaluate species distribution models. Ecology and Evolution 00:1–18.

Voigt, D. R., J. D. Broadfoot, and J. A. Baker. 1997. Forest management guidelines for the provision of white-tailed deer habitat. Ontario Ministry of Natural Resources.

Wang, J., and L. S. Chen. 2016. MixRF: A Random-Forest-Based Approach for Imputing Clustered Incomplete Data.