Full annual cycle habitat prioritization improves local habitat prioritization decisions

*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) can only 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 present a case study of woodcock management by the Pennsylvania Game Commission demonstrating 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.
4. *Policy implications:* The Pennsylvania Game Commission used this technique to prioritize new woodcock habitat management in northeastern Pennsylvania based on both migratory and breeding season habitat use. Our study exemplifies how multi-season habitat usage can be applied to a single decision-making framework to assist in full annual cycle conservation of migratory animals.

*Introduction*

* Migratory bird species pose a particular challenge for management, in that birds depend on geographically separate areas for different stages of their life cycle
* Birds are known to select different habitat during the breeding, winter, and migratory stages
* While the migratory stopover habitat typically differs from breeding or wintering habitat, most migratory birds have some overlap between the migratory range and the other two ranges
* This presents a unique challenge for managers in overlapping areas, as the scale and type of habitat selected for can differ based on life stage
* Managers in overlapping areas will need to take differing habitat requirements into account when designing their management strategies to ensure that migratory birds can find habitat during multiple life stages
* Need for full annual cycle management goes in here somewhere???
* We demonstrate the merits of this strategy through a demonstration of the Pennsylvania Game Commission’s prioritization of state gamelands for woodcock management
* The PGC’s interests were, not just to promote breeding habitat in the state of Pennsylvania, but also to preserve the connectivity of woodcock migration throughout the eastern management region

*Methods*

* *Breeding season species distribution model*
  + Breeding season species distribution models are the most frequently used tool for planning bird conservation (cite past work)
  + 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 XX mile long survey routes consisting of XX evenly-spaced points, with presence-absence determined at each point based on whether male displays were visible during a XX minute interval at dusk?
  + SGS survey routes were randomly distributed through the state? In 19XX? And each route has been run annually since
  + 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 20XX-20XX 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.
  + Land cover: Forest (NLCD 20XX), Successional class (LANDFIRE 20XX)
  + Geography: Elevation (source?), Slope (source?), Ecoregions (EPA level 3)
  + Moisture: Drainage (source?), Topographic wetness index (cite methodology)
  + Landscape: metrics from the landscapemetrics R package representing landscape composition (% forest, % agri, %dev) 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) 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
    - I 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
  + We used GPS data from the EWMRC to designate woodcock migratory stopover sites throughout the state of Pennsylvania
    - Include the information from the metadata file
    - These sites were used as present locations
  + To create pseudo-absence locations, we randomly allocated 10,000? 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.
  + Interactive applications (e.g. shinyapps) are an excellent way to distribute this information, although still images work too

*Results*

* How does this model differ from a single season model?
  + How much do gameland rankings differ between the two models?
  + How much do regional rankings (eg epa level 3 ecoregions) differ between the two models?
  + Developed areas are highlighted in the migratory model but not the residential model
    - Lower “habitat threshold”
    - Southeast region is a management priority in the migratory model, but not the residential model
    - Same deal with Pittsburgh in southwest
    - However, north central region is avoided in both models and northeast is selected for in both models
  + Takeaways:
    - Raises the prospect of urban habitat conservation
    - Adding migratory data isn’t going to change the regional calculus as much (most migratory stopover decisions aren’t regional).
      * However, it is likely to change local (eg gameland prioritization) decisions
      * Those are the kind of decisions where this type of combined migratory/breeding approach would be most useful

*Discussion*