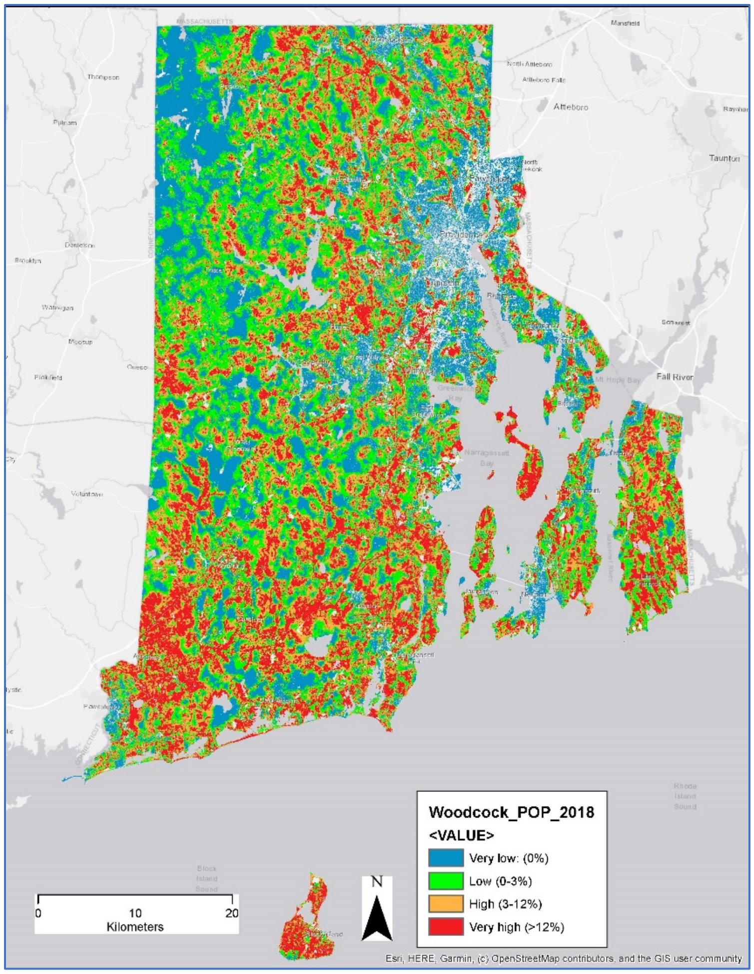
Species Distribution Modeling for American Woodcock in Rhode Island Based on 2018 Environmental Variables



Bill Buffum

University of Rhode Island
Department of Natural Resources Science
October 15, 2020

This report was prepared as part of a University of Rhode Island project entitled "Forest management for American Woodcock and associated wildlife: developing a state-wide spatially-explicit conservation plan (AWD 3695774)", with funding from the Rhode Island Department of Environmental Management and the University of Rhode Island.

Suggested citation: Buffum B. 2020. Species Distribution Modeling for American Woodcock in Rhode Island based on 2018 Environmental Variables. Kingston, Rhode Island: Department of Natural Resources Science, University of Rhode Island.

For questions about this report, contact Bill Buffum, Department of Natural Resources Science, University of Rhode Island, 13 Coastal Institute, Kingston RI 02881 (buffum@uri.edu).

Abstract

American woodcock (Scolopax minor) is a popular game species in the eastern U.S. whose populations have declined by roughly 1% per year since 1968. This decline has been linked to the loss of young forest habitat, so federal, state and private conservation agencies are implementing an active program to create young forest habitat for woodcock. The University of Rhode Island has conducted several detailed studies of woodcock habitat selection using species distribution modelling (SDM), an important ecological tool that predicts the probability of presence (POP) of a species based on an analysis of known occurrences and predictor variables such as topographic, edaphic, biogeographic and other ecological variables. The first SDM for woodcock in Rhode Island was based on two years of woodcock location data and spatial data that included the 2010 map of young forest habitat. In 2019 we decided to prepare a new SDM for woodcock in Rhode Island because several new spatial datasets were available, including a 2018 map of young forest habitat. We could also incorporate a third year of woodcock location data. This paper describes the methods used to create the 2018 SDM with Maxent modelling.

1. Introduction

American woodcock (*Scolopax minor*; hereafter, woodcock) is a popular game species in the eastern U.S. whose populations have declined by roughly 1% per year since 1968 (Cooper and Rau 2013), and this decline has been linked to the loss of young forest habitat (McAuley et al. 2005, Kelley et al. 2008). In Rhode Island the extent of young forest habitat in non-coastal areas has been decreasing by at least 1.5% per year (Buffum et al. 2011), and federal, state and private conservation agencies in Rhode Island are implementing an active program to create young forest habitat for woodcock (Buffum et al. 2019). Woodcock also serve as an umbrella species in the sense that forest management focused on enhancing woodcock populations also benefits many other species of wildlife (Masse et al. 2014).

The University of Rhode Island has conducted several detailed studies of woodcock habitat selection as part of an ongoing research program with the Rhode Island Department of Environmental Management (Masse et al. 2013, Masse et al. 2014, 2015, Brenner et al. 2019, Masse et al. 2019). The complex habitat requirements of woodcock and their relatively low abundance make it difficult to predict where woodcock are most likely to be found. They are most often seen displaying in their singing grounds, where they use a variety of early successional cover types including recent clearcuts, herb-dominated forest openings, and abandoned hayfields. However, during the day they often commute to feeding coverts with dense vegetation, and then spend the nights in open fields where there is less risk of predation (Masse et al. 2013).

In order to analyze woodcock habitat selection, URI has used species distribution modelling (SDM), an important ecological tool that predicts the probability of presence (POP) of a species based on an analysis of known occurrences and predictor variables such as topographic, climatic, edaphic, biogeographic and other ecological variables (Phillips and Dudík 2008). The first SDM for woodcock in Rhode Island (Masse et al. 2014) was based on two years (2010-2011) of woodcock location data (N = 42 birds) and spatial data that included the 2010 map of young forest habitat prepared by URI and the most current other spatial data available at the time.

In 2019 we decided to prepare a new SDM for woodcock in Rhode Island. Several new spatial datasets were available, including a 2018 map of young forest habitat which included habitat created by conservation organizations and private landowners since 2010. We could also incorporate a third year of woodcock location data collected by Roger Masse in 2012 to supplement the 2010-2011 location data used in the previous SDM. This paper describes the methods used to create the 2018 SDM, which we shared with the Rhode Island Department of Environmental Management in January 2020.

2. Field-Site Description

The outputs of our woodcock SDM analysis cover the entire state, but we based the SDM on occurrence data from field studies in the Arcadia, Big River, and Great Swamp Wildlife Management Areas (WMA) in Kent and Washington counties. Upland forest (coniferous, deciduous, and mixed) are co-dominant at Arcadia WMA, whereas coniferous upland forest is dominant at Big River WMA and wetland forest is dominant at Great Swamp WMA. Eastern white pine (*Pinus strobus*) is the most common species in coniferous upland forests; various oaks (*Quercus spp*.), hickories (*Carya spp*.), and red maple (*Acer rubrum*) are common in deciduous upland forests; and red maple (*Acer rubrum*) is the most common tree species in wetland forests. The Rhode Island Department of Environmental Management has been creating patches of young forest in each area to help conserve woodcock and other young forest wildlife, and expanded this program in the Great Swamp WMA after a 40 Woodcock Habitat Demonstration Area was designated in 2008 (Buffum et al. 2019).

3. Methods

We used Maxent v. 3.4.0 (Phillips et al. 2017) to prepare the SDM, and ArcGIS Desktop v. 10.6 (Environmental Systems Research Institute, Redlands, CA) to prepare the environmental variables and maps. A recent comparison of four popular SDM models concluded that Maxent produced the best predictions and is recommended when absence data is not available (Grimmett et al. 2020). In addition, Maxent offers the advantage of being able to create the SDM based on location data and environmental variables from one time period, and easily projecting the SDM onto a more recent set of environmental variables from another time period. In our case, the woodcock location data was collected between 2010 and 2012, but we projected the SDM onto 2018 environmental data.

Woodcock Location Data: Roger Masse collected location data for 68 birds between 2010 and 2012, with the number of location points per bird ranging from 14 to 51. The location data was obtained by catching woodcock in April/May, attaching VHF transmitters, and tracking their movements until the end of August (Masse et al. 2014). Maxent does not require a minimum number of points per bird, but we decided to (a) use the same number of points per bird, and (b) maximize the total number of location points. We achieved this by selecting birds with at least 28 location points (N=61) and then randomly selecting 28 points per bird, which resulted in 1,708 location points.

Available Area for Background Points: Maxent does not require absence data, and instead generates randomly selected ‘background’ points, also called pseudoabsences, that are contrasted against the presence points. The user must select the available area within which Maxent will generate the background points. The available area should include locations that the species could easily reach (Merow et al. 2013). We established the available area by using the ArcGIS near tool to measure the distance of each location point from the bird’s singing ground which it visits most every evening from March to June (McAuley et al. 2013). We calculated the average distance for all location points to singing ground (1,296 m) and used the ArcGIS buffer tool to delineate an area of 13,372 ha that was at least 1,296 m from any point (Figure 1).

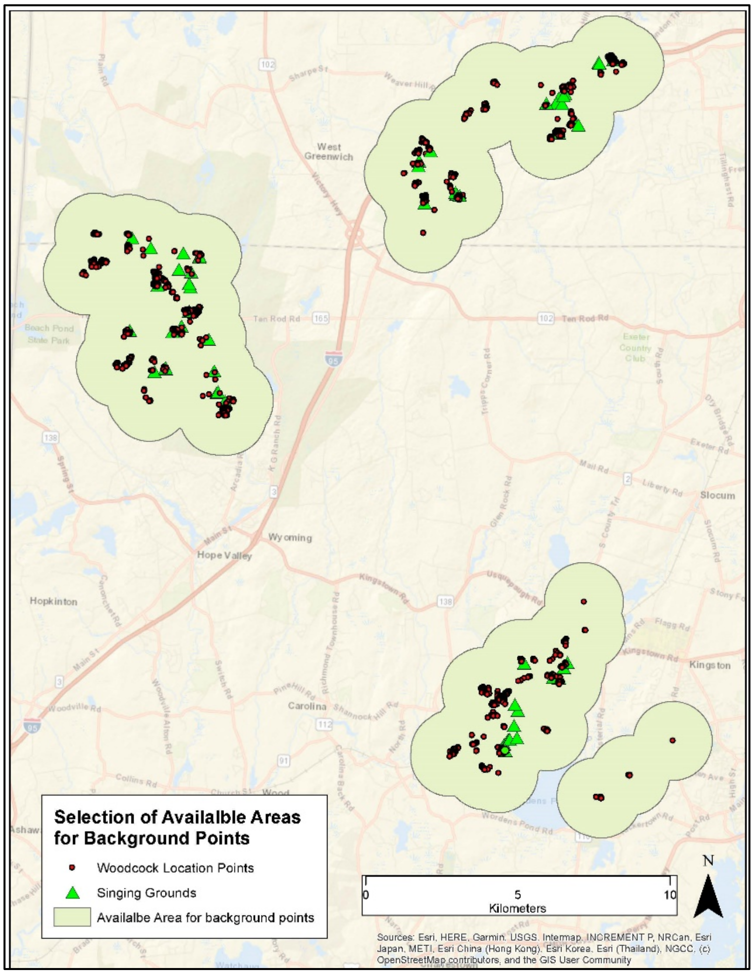


Figure 1. Selection of Available Areas for Background Points. Available areas (green-shaded clouds) selected within on average 1,296 m of woodcock singing grounds (triangles) given the recorded locations (black dots) of 61 woodcock outfitted with transmitters in three state-owned areas in Rhode Island, USA: Great Swamp Management Area (southeast corner), Arcadia Management Area (west), and Big River Management Area (northeast corner).

Environmental Variables: To construct the SDM, we used 2011 versions of 22 environmental variables that were contemporary with the woodcock location points, and then projected the model onto 2018 versions of the same variables (Table 1). We used a number of existing datasets to create a set of statewide rasters (10 m pixels). Most of the source data was originally in vector format. Since the accuracy of the woodcock tracking was estimated at 18m (Masse et al. 2013), we used the ArcGIS focal statistics tool to calculate the average value of each environmental variable within 18m.

Table 1. Environmental Variables (2011 and 2018) used in 2018 Species Distribution Modelling for Woodcock in Rhode Island. We used a number of existing datasets to create a set of statewide rasters (10 m pixels). Since the accuracy of the woodcock tracking was estimated at 18m (Masse et al. 2013), we used the ArcGIS focal statistics tool to calculate the average value of each environmental variable within 18m.

Name of environmental variable	**Expressed as:**	**Source of Data**
Low young forest (<1m tall)	Percent cover within 18m	Forest Habitat *
High young forest (1-6m tall)	Percent cover within 18m	Young Forest in RI **
Low mature forest (6-10 m tall)	Percent cover within 18m	Young Forest in RI **
High mature forest (>10 m tall)	Percent cover within 18m	Young Forest in RI **
Impervious surface	Percent cover within 18m	Land Use*
Agriculture/pasture	Percent cover within 18m	Land Use*
Other open area	Percent cover within 18m	Land Use*
Extremely well-drained soil	Percent cover within 18m	Soil *
Well-drained soil	Percent cover within 18m	Soil *
Poorly-drained soil	Percent cover within 18m	Soil *
Coniferous forest	Percent cover within 18m	Forest Habitat *
Deciduous forest	Percent cover within 18m	Forest Habitat *
Mixed forest	Percent cover within 18m	Forest Habitat *
Fresh water	Percent cover within 18m	Land Use*
Distance to upland young forest	Average distance (m) within 18m	Young Forest in RI **
Distance to all young forest	Average distance (m) within 18m	Young Forest in RI **
Distance to stream	Average distance (m) within 18m	Freshwater Rivers and Streams *
Distance to moist soil	Average distance (m) within 18m	Soils *
Distance to agriculture/pasture	Average distance (m) within 18m	Land Use*
Distance to other open area open	Average distance (m) within 18m	Land Use*
Elevation	Average elevation (ft) within 18m	(Gesch et al. 2002)
Slope	Average slope (%) within 18m	(Masse et al. 2014)

Notes:

* Available on Rhode Island Geographic Information System: https://www.rigis.org/

** Available on ArcGIS Online: https://www.arcgis.com/index.html

The environmental variables included the 13 variables used in the previous woodcock SDM (Masse et al. 2013), plus nine additional variables. In 2017, a new statewide layer of canopy height was developed based on 2011 Light Detection and Ranging (Lidar) data (Lopes and Brenner 2017). We used this new layer to create two height categories of young forest (<1m, and 1-6m) and two categories of mature forest (6-10m, and 10m). We created a new layer for distance to all young forest (upland and wetland) because many shrubland birds were are more likely to use patches of upland young forest that are adjacent to patches of wetland young forest (Buffum and McKinney 2014). We also added one variable for impervious surfaces, and three variables for soil drainage: poorly drained, well drained, and extremely well drained). Some of the new variables were correlated with the previous variables, but Maxent offers a major advantage of being more stable with correlated variables than stepwise regression, so it is not necessary to remove correlated variables that are ecologically relevant (Elith et al. 2011).

Maxent allowed us to project the results onto current conditions based on 2018 versions of the same 22 environmental variables. The primary output of the Maxent SDM analysis was a 10m pixel raster dataset of the 2018 woodcock POP throughout Rhode Island.

Estimated Predictive accuracy of the SDM: Maxent estimates the predictive accuracy of the SDM by calculating the area under the curve (AUC) of the receiver-operator characteristic (ROC), which expresses the probability that a randomly chosen presence location is ranked higher than a randomly chosen background point (Merow et al. 2013). AUC scores range from 0.5 to 1.0, and the predictive accuracy has been categorized as "fair" for values of 0.7–0.8, "good" for values of 0.8–0.9, and "excellent" for values over 0.9 (Swets 1988, Wei et al. 2018).

For the replicated run type, we selected cross-validation, in which the occurrence data is randomly split into a number of equal-size groups called "folds", and models are created leaving out each fold in turn so that the left-out folds can be used for evaluation. Cross-validation has an advantage over using a single training/test split: it uses all of the data for validation (Phillips et al. 2017). We set the number of cross-validation replications at 10.

Maxent also estimates the relative contributions of each environmental variable to the Maxent model. These estimates are based on the "gain", which is closely related to deviance, a measure of goodness of fit used in linear models (Phillips 2017). In each iteration of the training algorithm, the increase in gain is added to the contribution of the corresponding variable or subtracted from it if the change is negative (Phillips 2017).

4. Results

Woodcock Probability of Presence 2018: This primary output of the SDM was a map that presents the woodcock POP in any location in RI based on 2018 environmental data. The data is stored in an ArcGIS raster that provides the POP value for each 10m pixel. You can request access to a geodatabase containing the raster by contacting Bill Buffum (buffum@uri.edu). Figure 2 presents a simpler version of the map, in which the POP values are grouped into four quantiles with each class containing 25% of the pixels in the state.

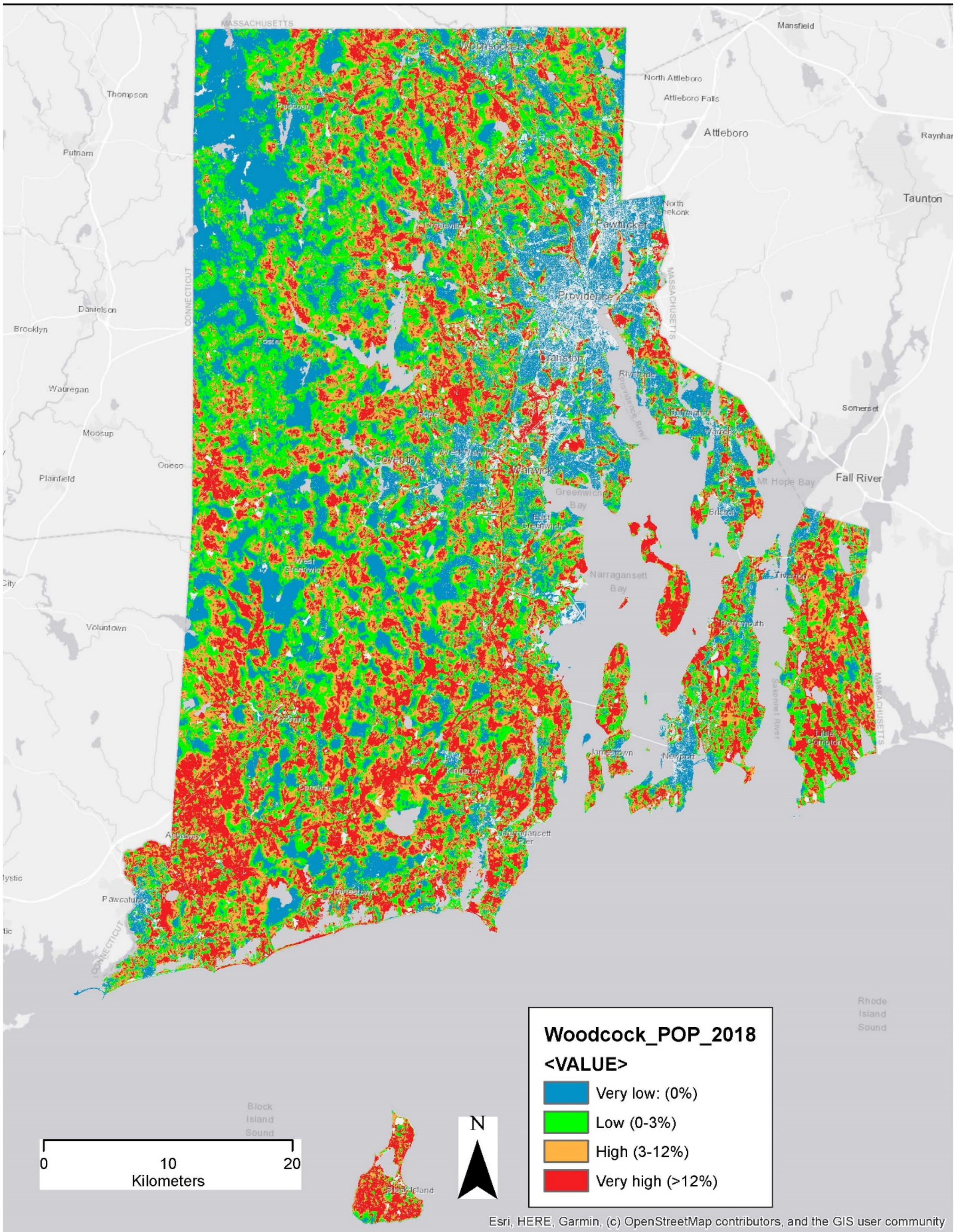


Figure 2. Woodcock Probability of Presence in Rhode Island based on 2018 Environmental Variables

Estimated Predictive accuracy of the SDM: Maxent calculated the average test AUC for the ten replicate runs at 0.881. This AUC result is at the top end of the "good" category (0.8–0.9) for SDMs, and just below the excellent category (Wei et al. 2018).

Maxent also estimated the relative contribution of each environmental variable to the model predicting POP. The three environmental variables with the highest contributions were high young forest, distance to all young forest, and distance to upland young forest (Table 2).

Table 2. Contribution of Environmental Variables to Model

Name of environmental layer	Percent Contribution to Maxent Model
High young forest (1-6m)	23.1%
Distance to all young forest	17.3%
Distance to upland young forest	8.9%
Distance to agriculture/pasture	5.2%
Distance to stream	4.9%
Elevation	4.5%
Impervious surface	4.3%
Well-drained soil	3.9%
Distance to other open area open	3.7%
Distance to moist soil	3.5%
Poorly-drained soil	3.4%
Low young forest (<1m)	3.1%
Coniferous forest	3%
High mature forest (>10 m)	2.7%
Other open area	2.4%
Extremely well-drained soil	1.7%
Mixed forest	1.4%
Deciduous forest	0.9%
Fresh water	0.9%
Low mature forest (6-10 m)	0.5%
Slope	0.5%
Agriculture/pasture	0.3%

5. References

- Brenner, S. J., B. Buffum, B. C. Tefft, and S. R. McWilliams. 2019. Landscape context matters when American Woodcock select singing grounds: Results from a reciprocal transplant experiment. The Condor: Ornithological Applications **121**:duy005.
- Buffum, B., and R. A. McKinney. 2014. Does Proximity to Wetland Shrubland Increase the Habitat Value for Shrubland Birds of Small Patches of Upland Shrubland in the Northeastern United States? International Journal of Forestry Research **2014**.
- Buffum, B., S. R. McWilliams, and P. V. August. 2011. A spatial analysis of forest management and its contribution to maintaining the extent of shrubland habitat in southern New England, United States. Forest Ecology and Management **262**:1775-1785.
- Buffum, B., B. C. Tefft, R. J. Masse, and S. R. McWilliams. 2019. Mobilizing Private Landowners to Create American Woodcock Habitat: Lessons Learined in Rhode Island. Pages 85-92 *in* D. G. Krementz, T. R. Cooper, and D. E. Andersen, editors. Proceedings of the Eleventh American Woodcock Symposium. University of Minnesota Libraries Publishing, Minneapolis, Minnesota.
- Cooper, T. R., and R. D. Rau. 2013. American woodcock population status - 2013. U.S. Fish and Wildlife Service, Laurel, Maryland.
- Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions **17**:43-57.
- Gesch, D., M. Oimoen, S. Greenlee, C. Nelson, M. Steuck, and D. Tyler. 2002. The national elevation dataset. Photogrammetric Engineering and Remote Sensing **68**:5-32.
- Grimmett, L., R. Whitsed, and A. Horta. 2020. Presence-only species distribution models are sensitive to sample prevalence: Evaluating models using spatial prediction stability and accuracy metrics. Ecological Modelling **431**:109194.
- Kelley, J., S. Williamson, and T. R. Cooper, editors. 2008. American Woodcock Conservation Plan. Wildlife Management Institute. , Cabot, Vermont. Available at: http://timberdoodle.org/
- Lopes, C., and A. Brenner. 2017. Rhode Island Statewide Canopy Mapping. Quantum Spatial, Inc.
- Masse, R. J., B. C. Tefft, J. A. Amador, and S. R. McWilliams. 2013. Why woodcock commute: testing the foraging-benefit and predation-risk hypotheses. Behavioral Ecology **24**:1348-1355.
- Masse, R. J., B. C. Tefft, B. Buffum, and S. R. McWilliams. 2019. Habitat selection of American Woodcock and its implications for habitat management where young forests are rare. Pages 168-177 *in* D. G. Krementz, T. R. Cooper, and D. E. Andersen, editors. Proceedings of the Eleventh American Woodcock Symposium. University of Minnesota Libraries Publishing Minneapolis, Minnesota.
- Masse, R. J., B. C. Tefft, and S. R. McWilliams. 2014. Multiscale habitat selection by a forest-dwelling shorebird, the American woodcock: Implications for forest management in southern New England, USA. Forest Ecology and Management **325**:37-48.
- Masse, R. J., B. C. Tefft, and S. R. McWilliams. 2015. Higher bird abundance and diversity where American woodcock sing: Fringe benefits of managing forests for woodcock. The Journal of Wildlife Management **79**:1378-1384.
- McAuley, D., D. Keppie, and R. Whiting Jr. 2013. American woodcock (Scolopax minor), version 2.0 *in* A. Poole, editor. The Birds of North America. Cornell Lab of Ornithology, Ithaca, New York.
- McAuley, D. G., J. R. Longcore, D. A. Clugston, R. B. Allen, A. Weik, S. Williamson, J. Dunn, B. Palmer, K. Evans, W. Staats, G. F. Sepik, and W. W. Halteman. 2005. Effects of hunting on

survival of American woodcock in the Northeast. Northeast Journal of Wildlife Management **69**:1565–1577.

- Merow, C., M. J. Smith, and J. A. Silander Jr. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography **36**:1058-1069.
- Phillips, S. J. 2017. A brief tutorial on Maxent. Available from url: http://biodiversityinformatics.amnh.org/open_source/maxent/. Accessed on 10/29/2019.
- Phillips, S. J., R. P. Anderson, M. Dudík, R. E. Schapire, and M. E. Blair. 2017. Opening the black box: An open-source release of Maxent. Ecography **40**:887-893.
- Phillips, S. J., and M. Dudík. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography **31**:161-175.
- Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. Science **240**:1285-1293.
- Wei, B., R. Wang, K. Hou, X. Wang, and W. Wu. 2018. Predicting the current and future cultivation regions of Carthamus tinctorius L. using MaxEnt model under climate change in China. Glob Ecol Conserv 16: e00477.