Encounter Rate Modelling of The Red-Headed Woodpeckers' (Melanerpes erythrocephalus) using Machine Learning.

Student No.: 100263294

Project Supervisor: James Gilroy

University of East Anglia, Norwich

Abstract:

There have been declines in the Red-headed Woodpecker since the last 50 years, and these implications are related towards declines in Savannas, Snags and dead trees that are effectively limiting the species distribution range. To understand these declines the study involved multiple BCR regions across North America along with multiple landcover classes to understand if the declines persist, and whether its range-dependent, or even landcover class dependent. Furthermore, the use of modern analytical techniques through machine learning and calibration models were implemented to determine recent trends. The use of recent Landsat data from Modis MCD12Q1, provide accurate estimates towards these trends. Given that, no research has yet been appended towards the woodpecker in grassland and croplands, the results were surprising and have shown high probabilities for encounters in these areas. However, it was also found that several BCR regions are continually declining and in need of further research and conservation to understand what the indicators for decline really represent. Recent literature has shown it is related to short burn-intervals, agricultural intensification, or even urbanisation.

Key words: Probability, Encounters, Landcover changes, Distribution models.

Table of Contents:

Introduction:

Landscape change has predominantly driven the decline of the Red-headed woodpeckers (Melanerpes erythrocephalus). However, minimal research appends to modeling the species distribution range relative to recent years and its changes to multiple habitats across the BCR regions.

While relevant research has proposed habitat suitability scores to measure the probability of the woodpecker to sustain in a preferred environment, these results have been insufficiently clear towards clarifying the species range expansion at multiple regional scales and whether habitat suitability is comparable to different habitats over time.
For example, Tirpak et al. (2009) show that habitat suitability index scores inversely correlate with Red-headed woodpecker abundance under the BBS-abundance dataset across BCR 24 and 25. Similarly, following results from Jones-Farrand et al. (2011) for BCR 25. The performance of using multiple citizen science datasets and recent landcover techniques that apply machine learning algorithms for greater accuracy and detection of landcover classes will provide new insights into species detection.

Although, minimal research effort directed at multiple BCR regions, combined with eBird citizen science data. The importance of studying the indicated species is the recent decline over the past 45 years (Rosenberg et al. 2016); this concern prompted research towards the cause for it designated by Partners in flight. As mentioned previously, most studies are derived using BBS data, hence the lack of research using eBird citizen science data. Implementing modern techniques of citizen science like eBird that uses complete checklists (Callaghan, 2017) combined with recent explorations of machine learning towards probability-encounters for the Red-headed Woodpecker across recent years may yield exciting results on their distribution range.
Furthermore, recent research interests from Keonig et al. (2017) proposed the constraints of Red-headed woodpecker success and range expansion are the lack of savanna habitat and low density of tall dead trees. These attribute to relevant studies that monitor the behavioral variants in how the Red-headed woodpecker use their nesting habitats (Vukovich, 2019). The effects of competition (Berl et al. 2015), on both inter-and intra- specific nesting grounds, and as Koenig (2017) mentioned, the preference for woodland areas with tall trees, high basal areas, low stem density in the understory (Rodewald et al. 2005; Frei et al. 2013), show that these habitat features are progressively diminishing in recent decades consequently limiting Red-headed woodpecker population (Kilo and Vukovich, 2014).
Although, Koenig et al. (2017) performed regression techniques for temporal-series across the sites for covariate analysis of habitat data. Along with mixed-effect models and structural equation models on BBS and CBC data to determine the correlation in the hypothesis that landcover affected Red-

headed Woodpecker populations. The issue with regression models is that several pre-specified conditions must work towards the transformation and normality of the data. Furthermore, Keonig must organize assumptions on collinearity and whether a correlation between the response and explanatory variables is linear before the models' development.

For example, Evans et al. (2011) mention that regression modeling requires pre-specified relationships between the response and explanatory variable, combined with the assumptions for normality and additivity. Arguably, Koenig has ln(n+1) transformed count data to stabilize the variance accordingly. However, BBS contains zeros that are not processed adequately through log-transformation nor through time-series regression. In support of this, O'hara and Kotze (2010) suggest that log-transformed count data perform poorly, and that Poisson or negative binomial models are suited to fit such data.

Taking this into consideration, unlike linear regression models, machine learning techniques such as Random forests are non-parametric and allow for non-linearities where no assumptions are necessary to pre-define the model. The methodology towards this is splitting data into a training and test set, where the models' performance is predicted based on an information criterion (Smith et al., 2013; Gomping, 2009). The advantages of using Random forests on eBird citizen data are that complex interactions model simplistically, and zero-inflations heavily prominent in eBird-data are handled effectively.

Literature review:

In recent years, the Red-headed Woodpecker have experienced range-wide declines across its distribution range in the past 50 years, with a 68% loss (Rosenberg et al. 2016). These reasons are attributed to landcover changes, for the woodpecker is an easily recognizable guild, reliant on woodlands, and stags (standing dead trees), for nesting, roosting, and foraging (Bull et a. 1997). The works of Anderson and LaMontagne (2016), show that habitat selection behaviours of the Red-headed Woodpecker, are characterised by selecting snag trees with large diameters and height. Additionally, the buffer for their selected tree would have a greater probability of surrounding dead trees. Suggestive that they are commonly associated with decisive habitat preferences, and open forest systems with selected open canopy characteristics (Pagel et al. 2020; Nickley and Bulluck, 2019).

Given the selection of habitat sites during nesting, the Red-headed woodpecker is known to use a range of other habitats, and these are not only limited to golf courses, forests edges, cropland, open agricultural areas, and floodplain forest (Smith et al. 2000). Given that much is already known on the habitat preferences for the Red-headed Woodpecker regarding nest-site selection at a minimal distribution range, usually local or regional. Minimal is appended to their entire geographic range across North America, and whether recent trends would show newer site preferences.

During the summer months, the Red-headed Woodpecker rely on habitats that provide space for aerial foraging (Kilgo and Vukovich, 2012), although absence of decaying wood material, savannas, and open canopy forests, will consequently deprive the red-headed woodpecker of aerial foraging opportunities. Nebel et al. (2010) found the aerial foraging birds were more likely to decline in North-eastern America and these link with insect diversity and biomass decline, which are resultingly caused by agricultural intensification and urbanisation (Sanches-Bayo and Wyckhuys, 2019).

<u>Prediction</u>

Proposed were 5 null hypotheses to test whether the multiple linear regression rejects the null hypothesis:

H_0 = Water bodies decreasing are correlating with encounter increase.

H_1 = Savannas decreasing are correlated with encounter increase.

H_2 = Urban increase is correlated with encounter increase.

H_3 = Grassland/Cropland decreasing are correlating with encounter increase.

H_4 = Woodlands decreasing are correlating with encounter increase.

As for the original hypothesis, we assume the following format across all hypotheses. If the null hypothesis stated as A is decreasing while encounter rates increase then the hypothesis, which is B would be that encounters are decreasing. Furthermore, as for landcovers with increasing habitat, it's supposed that encounters also increase.

<u>Methodology:</u>

<u>Data from eBird</u>

Ebird data are a powerful resource for driving scientific analysis on species distribution because of its extensive database, that contains over 600 million observations from every country in the world. From their data products, the chosen species of the Red-headed Woodpecker was downloaded for a specified region and time-range, in this case, North America and years 2010-2019 (https://ebird.org/data/download/ebd).

The structure of ebird data is facilitated through two critical aspects. Firstly, they are structured as checklists when submitted and contain a list of bird species recorded. They are considered a 'complete checklist' when all species on-site have been identified, this distinguishes between those species not detected, and lack of participant observations. Finally, most eBird checklists have associated metadata to help facilitate a further understanding of missing observations, and the effort of observation process (Kelling et al., 2018).

To model Red-headed Woodpecker encounter distribution it was decided to restrict this data geographically to Bird Conservation Regions (BCR), the restrictions were based on eBird breeding ranges collected from the earliest year, 2018, figure 1.

Figure 1. Breeding range map for Red-headed Woodpecker from June-August (eBird, 2005) [Accessed: 26/12/2020]

BCRs are ecologically distinct ecoregions in North America, that are similar in their biotic and abiotic characteristics, to facilitate recognition in management for bird conservation (Bird Studies Canada and NABCI, 2014). Shapefile data was collected from birds Canada (https://www.birdscanada.org/bird-science/nabci-bird-conservation-regions/) that contain all partitioned BCR regions, those areas of interest were derived from the breeding ranges from figure 1, there were 17 BCR regions in total, figure 2.

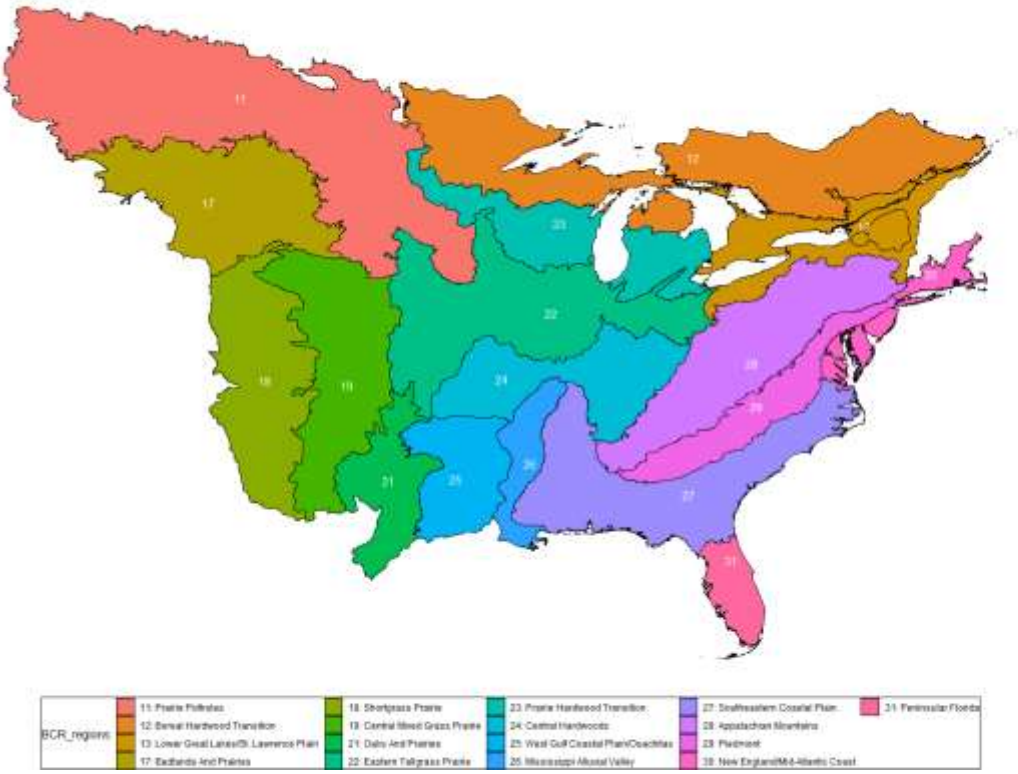

Figure 2. Map showing the BCR regions used in modelling Red-headed Woodpecker encounters (Rstudio, 2020)

The data was processed with the *auk* R package (Strimas-Mackey, Miller, & Hochachka, 2017). The checklists were filtered to those from 1st May to 31st August from 2010 to 2019. These were the selected breeding ranges (Frei et al. 2020). Further filtering was essential to impose consistent structure on the data for statistical inferences. Following Strimas-Mackey et al. (2020), the eBird data was restricted to checklists less than 5hours long and 5km in length along with 10 or less observers, this focused only on stationary and travelling protocol types.

<u>Environmental Covariate analysis</u>

To model species distribution for trends, the environmental covariates were a necessary parameter to model encounter rates across space and time. The MODIS MCD12Q1.006 land cover product was available (Friedl and Sulla-Menashe, 2015), which provides global maps of landcover at 500-m spatial resolution for years 2001-2018, whilst using the University of Maryland classification scheme (Hansen et al., 2000), in the Appendix.

MODIS data is provided as 1200 km by 1200km tiles as a raster consistent of 500-m spatial resolution. To extract the region of interest the MODIS package (MODIS, 2020) in R downloads and processes the data by providing automated access to the global online data archives (LP DAAC and LAADS). Whilst the landcover data spans from 2001-2018, only years 2010-2018 were needed, and data from 2018 was used to create data for 2019. To further facilitate filtering of habitat covariates with eBird data, a summarisation of landcover within a neighbourhood around the checklists was implemented to reduce spatial error, that may be attributed to extracting data directly from the checklist point-location.

To configure habitat metrics, Class area (CA) and percentage of landscape (PLAND) compositions were implemented to determine how much of a landscape is comprised of a patch type (McGarigal et al. 2002), using the following equation:

$$PLAND = P_I = \frac{\sum_{j=1}^{n} a_{ij}}{A}(100)$$

PLAND is equal to the sum of areas (m^2) of all patches corresponding the patch type (a_{ij}), divided by the total landscape area (m^2), and multiplied by 100 (to convert into percentages).

Firstly, a set of unique checklist locations for each year of the eBird data was gathered and then re-projected to the sinusoidal equal area used by MODIS, then these points are buffered to create a neighbourhood around each location. Further, for each square neighbourhood the raster values within them and their count of numbered cells are extracted to create landcover values for each unique checklist. Then by using equation 1, the data is summarised to the proportion of the neighbourhood within each landcover class (Strimas-Mackey et al. 2020)

Prediction surface

Habitat predictions for pland-coverage were developed as an indicator of habitat-dominance across the landscape, and as a framework to model encounter rates relative to the percentage of habitat coverage. The prediction surface map was developed by aggregating landcover by a factor of 5x5 modis cells, developed in the covariate analysis, and then using CRAN raster package (Hijmans, 2020), to load the function ‘rasterize’. Which transfers values associated with spatial data (points, lines, polygons) to raster cells, whilst retaining the field values that splits geometry relative to its respective BCR regions.

Finally, the CRAN package exactextractr (Baston, 2020) enables the extraction of cell values in a Raster, collecting polygonal geometries, as well as the fraction of each cell covered by the polygon. The result converts these values into a data frame containing the raw values and coverage fractions (when fun = NULL), which are indicated by numbers 0 to 15. However, this study chose to include 10 landcover classes in total, Water, Wetland, Deciduous Broadleaf, Mixed Forest, Woody Savanna, Savanna, Urban, Cropland and Grassland. These landcover classes were specifically picked from the rest because of literature that mentions these landcover classes as being most suitable towards encountering woodpeckers, or the Red-headed Woodpecker especially.

The occurrences of habitat type are counted, then the pland metric is calculated relative to count-occurrence, and finally these values are converted into descriptive names (Appendix, I). To effectively gather yearly data, the pland values were taken yearly, and each dataset was categorised by the habitats, so that yearly data of pland for each habitat could be calculated.

Topographical covariates

Topographic data collected from Amatulli (2018) (alternatively from earthenv), represents a range of patterns and processes especially important in ecology for determining trends in statistical ecology. The following dataset derives digital elevation models of global 250m GMTED2010 and near global 90m SRTM4.1dev, with 15 overall topographic variables, whilst only two are used, which are elevation and slope.

Slope and elevation were calculated using the GMTED median, for this is less sensitive to outliers and less affected by the potentially skewed distribution of elevation values (Amatulli, 2018).

The data collected from earthenv (https://www.earthenv.org/topography), was uploaded in R-studio as a raster, then cropped to the relevant regions in North-America. The elevation values were extracted within the neighbourhood of each checklist, and finally the median and standard deviation were calculated, and merged with the prediction surface dataset and the same method was followed with slope.

Encounter Rates

To estimate encounter rates proportional to Red-headed Woodpecker occupancy, then involving effort covariates into the model can take consistent measures against unaccounted detectability across all sites, for eBird data cannot typically account for ‘absolute’ detectability (Guillera-Arroita et al. 2015; Strimas Mackey et al. 2020).

Spatiotemporal subsampling was accounted for as a spatial bias, temporal bias, and class imbalance to increase the effectiveness of detections during the random forest model. Consequently, this increases the estimated probability of occurrence (Strimas Mackey et al. 2020).

Furthermore, Brieman (2001) defines using random forests as a collection of tree-structured weak learners comprised of identically distributed random vectors where each tree contributes to a prediction for the hypothesis. Liaw and Wiener (2002) further develops on this, mentioning that a selection of subset features is collected, representing nodes and each node is split whilst the best subset of predictors is randomly chosen by each node and used as predictor importance. The importance, namely ‘variable importance’, is decided by the prediction error, these are determined by “out-of-bag” (OOB) data, which are bootstrap samples taken from the data, where some observations do not make the bootstrap sample (Cutler et al. 2011). The calculation is based on training data, where the variable importance is obtained directly from the algorithm, then re-fit using predictions for the most important predictors during the prediction phase.

The model was developed for detection/non-detection of Red-headed Woodpecker relative to the habitat covariates (including elevation and slope), including the effort covariates to account for variability in detection. Firstly, the data was randomly split into 80% checklists for training and 20% for testing. An example of the random forest procedure is given in figure 3.

Furthermore, calibration models were fitted to estimate the predictive ability expected from a model in subsequent predictive applications (Vaughan and Ormerrod, 2005), this involves using test-data, that can be compared against the observed species distribution. The purpose was to increase the predictor force of the prediction model from ranger.

The calibration plot begins as a basic tool for assessing probabilistic classifier predictions, where the average predicted probability (x-axis) against species prevalence (y-axis) is plotted for determination of site-probability across the probability scale. A good classifier gathers positive examples near the upper right-corner (1) and negative examples near the lower left corner (0). To further assess the calibration model, AUC, Mean Square Error (MSE), Sensitivity, Specificity and Kappa are calculated to compare the predictions metrics for performance between the calibrated model and uncalibrated model. For the AUC, if all sites of known presence are predicted higher than all sites of known absence, the AUC is closer to 1 (Hijmans, 2012.) These are called accuracy measures, and Cutler et al (2007), state their purposes individually; sensitivity calculates the percentage of presences correctly classified, whilst specificity determines the percentages of absences correctly classified. Finally, kappa is the measure of agreement between both presence and absence predictions relative to the correction that might be due to chance alone.

Finally, predictions were created on the prediction surface dataset with a standard eBird checklists of 1km distance, 60 minutes travelling count at the peak time of day for detecting the following species, with a maximum number of observers, being 1 then rasterizing these predictions for a spatial analysis of encounter rates.

Summarizing predictions

Finally, to summarise these predictions in model format, then both values for encounter and habitat type were extracted at the cell-size corresponding to their rasters. For those rasters with increasingly high values, these were randomly sampled at 1000 points, to prevent over-dispersion during linear regression, whilst retaining a sample size effective as a predictor response. Finally, the purpose was to collect values corresponding to landcover changes since 2010, then splitting these into two models, both decreasing and increasing relative to encounter rates towards changes in pland for landcovers beyond years 2010. To finalise the model, a stepforward selection was produced to select the minimum covariates required to best fit significance for regression. Finally, the values belonging to changes in habitat, these where projected as points onto each encounter map to visualise the effect of landcover changes relative to encounter predictions.

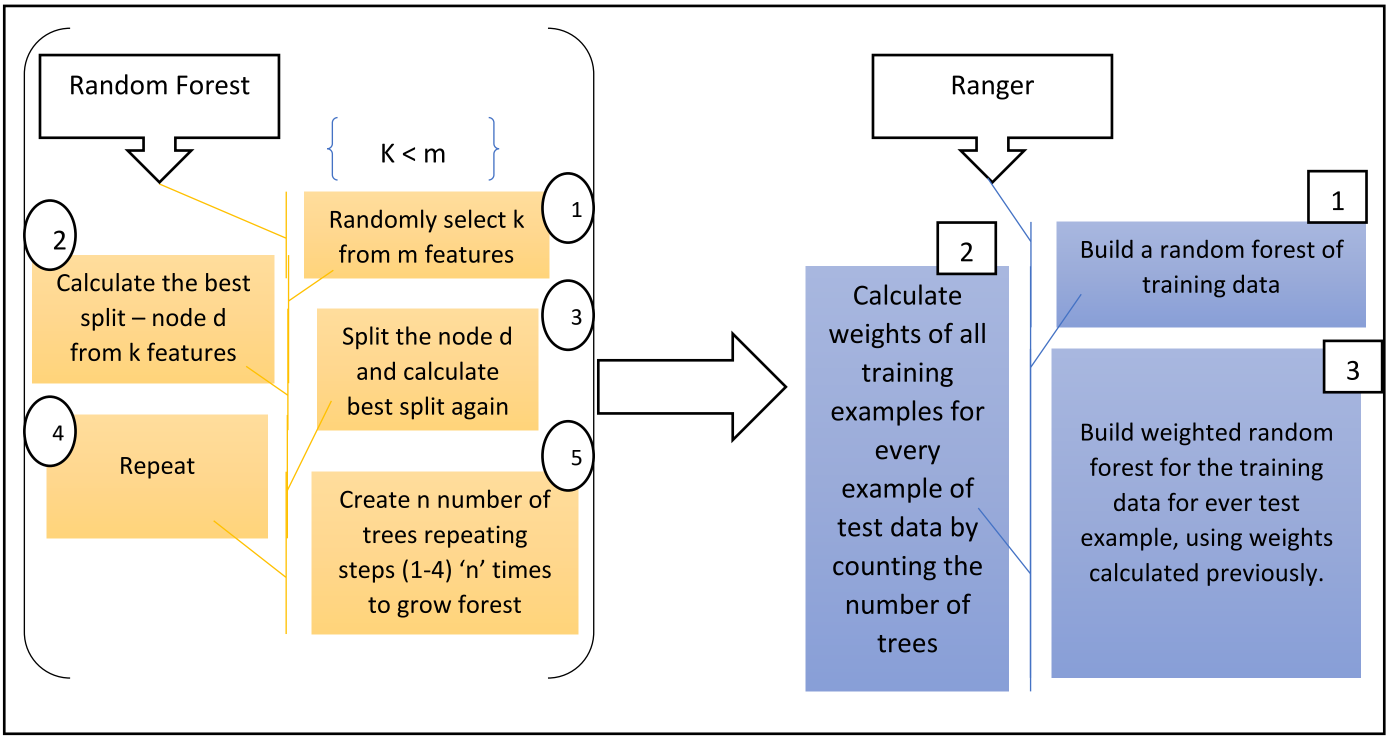

figure 3. Stages representing the algorithm reasoning of Random Forest and Ranger. Random forest splits the data, to accumulate trees by firstly randomly sub setting variables from the data, to train, then continuously training the algorithm towards it for a representation of the entire dataset. Ranger builds from that training data, using weights as a predictor force for importance of the variables.

<u>Results:</u>

Table 1 illustrates the calibrated model more effectively explained trends in comparison to the non-calibrated random forest (appendix), by the R-square value, the R^2 is the proportion of variance explain by the covariates relevant to encounter rates: comparing increases for urban: $R^2 = 0.0752_{RF}$ ~ $R^2 = 0.1151_{CRF}$. Differences in landcover increases between R^2 values were 0.03982 (Urban) and 0.0864 (Agriculture) and between 0.0026 (Grassland/Cropland) and 0.05785 (Savannas) in areas of landcover decrease.

Table 1. Displays the R-Square Value, F-statistics, and Degrees of Freedom of the Calibrated Random Forest models. Split into 4 groups, Water bodies (Wetland, Water), Savannas (Woody Savannas, Savannas), Urban, Grassland/Cropland, Woodland (Deciduous Broadleaf, Mixed Forest).

Outputs	**Water Bodies (Decrease)**	**Water Bodies (Increase)**	**Savannas (Increase)**	**Savannas (Decrease)**	**Urban (Increase)**	**Grassland/Cropland (Increase)**	**Grassland/Cropland (Decrease)**	**Woodland (Increase)**	**Woodland (Decrease)**
Multiple R^2	0.09053	0.09164	0.07867	0.07758	0.1151	0.1924	0.1364	0.1199	0.0547
Adjusted R^2	0.0854	0.08911	0.07324	0.07527	0.1077	0.1855	0.1307	0.1182	0.04897
F	17.63	36.26	14.48	33.61	15.52	27.71	23.92	69.48	9.552
DF	12, 2125	11, 3954	8, 1357	3, 1199	8, 954	10, 1163	9, 1362	5, 2550	12, 1981

Waterbodies:

The F-statistic (F = 17.63) for Waterbodies is much larger than the tail of $F_{0.95,\ 12,\ 2125}$ = 1.003832 for decreasing landcover, and for the increasing landcover of Waterbodies the F-statistic (F = 36.26) is also larger than the tail $F_{0.95,\ 11,\ 3954}$ = 1.791306, therefore we reject both Null Hypotheses. This indicates that as Water bodies decrease, encounters are expected to decrease, and as Water bodies increase, encounter rates are expected to increase.

In table 2, Water has a positive estimate of 0.021125 (P < 0.05) and a t-statistic of zero slope at t = 6.024, which is much larger than $t_{0.975,\ 2125}$ = 1.961081. Along with Wetlands, estimate = 0.011254 (P < 0.005), and a t-statistic of zero slope at t = 2.606, which is larger than $t_{0.975,\ 2125}$ = 1.961081.

All the BCR regions have a negative estimate, table 2. Inferring that as the estimate for Water and Wetland is positive in decreasing habitats, there is a positive relation with encounters decreasing with landcover decline. Additionally, the linear association for Water is greater than Wetland ($6.024_{t,water}$ > $2.606_{t,\ wetland}$), consequently inferring that significance for encounter rates declining in Water is proportionally larger than Wetlands by a difference of 0.009871 in the estimate. Given all of this, further factors such as BCR regions implicate the total variance, and association between Water and Wetland, which had the highest variable importance. It is shown that these two landcover habitats have encounter declines across the entire range, however, all BCR regions have a negative estimate, this infers that declines in encounters are decreasing in these individual regions by a marginal amount. This shows that BCR regions support the null hypothesis although water and wetland reject the null hypothesis.

Table 2. Water bodies with decreasing pland landscape-metric since 2010, from 2011 – 2019. The results are coefficients from a linear model, these include the Estimate, Standard Error, t-test, and P-value (P < α where α = 0.05) In R-studio. Signif. codes: 0 .0001'***', 0.001 '**', 0.01 '*'

Variables	**Estimate**	**Std.Error**	**t**	**Pr(>|t|)**
(Intercept)	0.033262***	0.001188	28.004	2.00E-16
Water	0.031358***	0.004577	6.851	9.54E-12
Peninsular Florida	-0.01495***	0.001688	-8.858	2.00E-16
Prairie Hardwood Transition	-0.02218***	0.00311	-7.13	1.37E-12
Central Hardwoods	-0.01638***	0.002175	-7.532	7.36E-14
Appalachian Mountains	-0.01989***	0.003063	-6.494	1.04E-10
Prairie Potholes	-0.01092***	0.001655	-6.599	5.22E-11
Piedmont	-0.01504***	0.002488	-6.044	1.77E-09
Badlands and Prairies	-0.01825***	0.003422	-5.334	1.06E-07
Boreal Hardwood Transition	-0.00749***	0.001419	-5.28	1.42E-07
Southeastern Coastal Plain	-0.00718**	0.00218	-3.292	0.00101
Wetland	0.011254**	0.004318	2.606	0.00922
Lower Great Lakes St. Lawrence Plain	-0.00673*	0.003208	-2.099	0.03596

The results of table 3 conveys that increases in encounter are positively correlated with Wetland (P<0.01) and a t-statistic of zero slope at t = 10.044, which is much larger than $t_{0.975,\ 3954}$ = 1.960564. However, Water has a negative estimate of -0.01819 (P<0.01), and a t statistic of zero slope at t = -

6.066, which is much larger than $t_{0.975, 3954}$ = 1.960564. It is apparent that encounters are increasing with wetland increase, however, declines are apparent with Water. Suggesting that encounter rates decline no matter whether Water landcover is increasing nor decreasing across the spatial range specified. Only three BCR regions are effectively positively correlated with increasing probabilities of encounter, these are Oaks and Prairies, Lower Great Lakes St.Lawrence Plain, and Mississippi Alluvial Valley. A total of 6 BCR regions are negative in their estimate towards encounter rates increasing, this infers potential declines in some regions and rejects the hypothesis, though the final model accepts the hypothesis.

Table 3. Water bodies with increasing pland landscape-metric since 2010, from 2011 – 2019. These results are coefficients from a linear model, these include the Estimate, Standard Error, t-test, and P-value ($P < \alpha$ where $\alpha = 0.05$) In R-studio. Signif. codes: 0 .0001'***', 0.001 '**', 0.01 '*'

Variables	**Estimate**	**Std.Error**	**t**	**Pr(>|t|)**
(Intercept)	0.021777***	0.000655	33.245	2.00E-16
Wetland	0.042542***	0.004236	10.044	2.00E-16
Water	-0.01819***	0.002999	-6.066	1.43E-09
Oaks and Prairies	0.009744***	0.001668	5.843	5.53E-09
Lower Great Lakes St. Lawrence Plain	0.005833***	0.001487	3.923	8.90E-05
Appalachian Mountains	-0.00971***	0.001872	-5.187	2.24E-07
Mississippi Alluvial Valley	0.010982**	0.003642	3.016	0.00258
Central Hardwoods	-0.00563**	0.00177	-3.178	0.00149
Badlands and Prairies	-0.00511**	0.001569	-3.255	0.00114
Prairie Potholes	-0.00232**	0.000831	-2.789	0.00532
Shortgrass Prairie	-0.00781*	0.00371	-2.106	0.03523
Eastern Tallgrass Prairie	-0.00616*	0.002871	-2.146	0.03192

Savannas:

The F-statistic (F = 33.61), table 1, for Savannas is much larger than the tail of $F_{0.95, 3, 1199}$ = 2.649979, for decreasing landcover, and for the increasing landcover of Savannas the F-statistics (F = 14.4) is also larger than the tail $F_{0.95, 8, 1357}$ = 1.945214, therefore the Null-hypothesis is rejected. This finding suggests that encounters decline when Savannas decrease, and they increase when Savannas increase. Savanna has a positive estimate of 0.038645 ($P < 0.05$) indicating that the probability of encounter increases across the landscape, and a t-statistic of zero slope at t = 6.269, which is much larger than $t_{0.975, 1357}$ = 1.961714. Table 4 shows, along with Woody savanna, estimate = 0.011283 ($P < 0.005$), and a t-statistic of zero slope at t = 2.272, which is larger than $t_{0.975, 1357}$ = 1.961714. Both landcovers show linearity with encounter increase relative to increasing landcover across the selected region of North America. Although, the estimate for all BCR regions included in the following model are negative, suggesting that there are declines in encounters for these regions that are significant ($P < 0.05$). Additionally, Appalachian Mountains has the highest negative value t = -7.091, this suggests that the standard error is smaller than the estimate, and the predicted values vary less relative to observed values for the estimate, hence comparingly, it is more significant (P < 0.01) towards declines than other BCR regions, which reject the hypothesis, though the final model accepts the hypothesis.

Table 4. Savannas with increasing pland landscape-metric since 2010, from 2011 – 2019. These results are coefficients from a linear model, these include the Estimate, Standard Error, t-test, and P-value (P < α where α = 0.05) In R-studio. Signif. codes: 0 .0001'***', 0.001 '**', 0.01 '*'

Variables	**Estimate**	**Std.Error**	**t**	**Pr(>|t|)**
(Intercept)	0.025581***	0.000973	26.291	2.00E-16
Savanna	0.038645***	0.006165	6.269	4.88E-10
Appalachian Mountains	-0.00938***	0.001322	-7.091	2.13E-12
Eastern Tallgrass Prairie	-0.01809***	0.003886	-4.656	3.54E-06
Prairie Potholes	-0.01589**	0.005015	-3.167	0.001572
Lower Great Lakes St. Lawrence Plain	-0.00747***	0.002021	-3.696	0.000228
Boreal Hardwood Transition	-0.00566***	0.001464	-3.868	0.000115
Central Hardwoods	-0.00592***	0.001581	-3.745	0.000188
Woody Savanna	0.011283*	0.004967	2.272	0.023252

The following table 5, displays the results for the decreasing pland metric of Savannas, which have significance (P<0.05), because this test rejects the null hypothesis (F = 33.61), and the intercept is negative for Savannas at t = -6.324, which is much smaller than $t_{0.975,\ 1199}$ = -1.961944. The t-test assumes that standard errors are lower relative to the slope in comparison to other variables.

The probability of encountering declines in Savanna for the Red-headed woodpecker is -0.03288, inferring that potential increases towards encountering the species may be occurring in some areas for the savanna. Furthermore, Appalachian Mountains have a negative estimate and relating back to table 4, this suggests that the final model is correlating with increasing probabilities of encounters. The south eastern coastal plain has a positive estimate, this assumes declines are prevalent in this region. However, whilst the estimates tend towards limiting encounter declines, the final model accepts the hypothesis.

Table 5. Savannas with decreasing pland landscape-metric since 2010, from 2011 – 2019. These results are coefficients from a linear model, these include the Estimate, Standard Error, t-test, and P-value (P < α where α = 0.05) In R-studio. Signif. codes: 0 .0001'***', 0.001 '**', 0.01 '*'

Variables	**Estimate**	**Std.Error**	**t**	**Pr(>|t|)**
(Intercept)	0.022494***	0.000683	32.925	2.00E-16
Appalachian Mountains	-0.0082***	0.001378	-5.95	3.52E-09
Savanna	-0.03288***	0.005198	-6.324	3.58E-10
Southeastern Coastal Plain	0.004556***	0.001238	3.679	0.000244

Urban:

The F-statistics (F = 15.52) for Urban is much larger than the tail of $F_{0.95,\ 8,\ 954}$ = 1.948092 for increasing landcover, this rejects the Null-hypothesis. The result suggests that as Urban areas increase, then encounters are expected to decrease. Table 6 shows that encounter rate declines are decreasing by an increment of -0.00595 in probability towards urban landcover, inferring that the Red-headed Woodpecker has expected increases of encounter for increasing urban extent. This is also consistent in three other BCR regions whilst four regions are declining in encounters with urban increase, which rejects the hypothesis that encounter rates decrease with urban areas, though the final model accepts it.

Table 6. Urban with increasing pland landscape-metric since 2010, from 2011 – 2019. These results are coefficients from a linear model, these include the Estimate, Standard Error, t-test, and P-value (P < α where α = 0.05) In R-studio. Signif. codes: 0 .0001'***', 0.001 '**', 0.01 '*'

Variables	**Estimate**	**Std.Error**	**t**	**Pr(>|t|)**
(Intercept)	0.017705***	0.000704	25.148	2.00E-16
Eastern Tallgrass Prairie	-0.00526***	0.001185	-4.435	1.03E-05
Southeastern Coastal Plain	-0.00933***	0.002107	-4.427	1.07E-05
Peninsular Florida	0.007399***	0.001717	4.309	1.81E-05
Oaks and Prairies	0.004967***	0.001084	4.583	5.19E-06
Mississippi Alluvial Valley	0.016068***	0.004495	3.575	0.000368
Lower Great Lakes St. Lawrence Plain	0.004914**	0.001576	3.117	0.00188
Urban	-0.00595*	0.002492	-2.388	0.017118
Appalachian Mountains	-0.00545*	0.002503	-2.178	0.029664

Grassland and Cropland:

The F-statistics, table 1, (F = 23.92) for encounter rates in grassland and cropland landscapes is much larger than the tail of $F_{0.95, 10, 1362} = 1.837639$ for decreasing landcover, and for increasing landcover of Agricultural landscapes the F-statistic (F = 27.71) is also larger than the tail $F_{0.95, 11, 1162} = 1.796876$. Therefore, we reject the null hypothesis, and suggest that Encounter rates decline with grassland and cropland decrease and increase when grassland and cropland both increase in percentage of area coverage. Table 7 shows grassland and cropland are both positively increasing (estimate = 0.037417, 0.007913), although greater significance is determined by grassland at t = 9.387, suggestive that the standard error is marked far lower than the estimate. Hence, the predicted and observed probabilities are similarly correlating. There is a mixture of increasing and decreasing BCR regions, notably Central Hardwoods and Appalachian Mountains having decreasing encounter rates, which individually show mixed support for the hypothesis, though the final model accepts the hypothesis.

Table 7. Grassland/Cropland with increasing pland landscape-metric since 2010, from 2011 – 2019. These results are coefficients from a linear model, these include the Estimate, Standard Error, t-test, and P-value (P < α where α = 0.05) In R-studio. Signif. codes: 0 .0001'***', 0.001 '**', 0.01 '*'

Variables	**Estimate**	**Std.Error**	**t**	**Pr(>|t|)**
(Intercept)	-1.74218***	0.481473	-3.618	0.000309
Grassland	0.037417***	0.003986	9.387	2.00E-16
Central Hardwoods	-0.00754***	0.001859	-4.055	5.34E-05
Appalachian Mountains	-0.01071***	0.002587	-4.139	3.73E-05
Central Mixed Grass Prairie	0.008127***	0.001578	5.15	3.05E-07

Eastern Tallgrass Prairie	0.00635***	0.00148	4.291	1.93E-05
Prairie Potholes	0.006031***	0.001449	4.161	3.40E-05
Badlands and Prairies	0.00611**	0.001857	3.291	0.001027
Cropland	0.007913*	0.003403	2.325	0.020246
Oaks and Prairies	0.004674*	0.002059	2.27	0.023372

The hypothesis for table 8, states that as grassland and cropland decreases so does Encounter rates. The following table, has Cropland with significance ($P < 0.01$) in the final regression model with a negative estimate of 0.008268, inferring similarly with table 5 that encounters are not effectively decreasing with declining Agricultural landscapes. Furthermore, Appalachian Mountains and Centrals Hardwoods have positive slopes, which correspond with table 7 that conveyed decreases of encounters in these BCR regions. This shows similarities in the support of the hypothesis between table 6 and 7, in declining cropland areas, showing an acceptance of the hypothesis.

Table 8. Grassland/Cropland with decreasing pland landscape-metric since 2010, from 2011 – 2019. These results are coefficients from a linear model, these include the Estimate, Standard Error, t-test, and P-value ($P < \alpha$ where $\alpha = 0.05$) In R-studio. Signif. codes: 0 .0001'***', 0.001 '**', 0.01 '*'

Variables	**Estimate**	**Std.Error**	**t**	**Pr(>|t|)**
(Intercept)	0.028591***	0.000687	41.648	2.00E-16
Appalachian Mountains	0.01441***	0.001522	9.472	2.00E-16
Oaks and Prairies	0.009541***	0.001957	4.876	1.21E-06
Southeastern Coastal Plain	-0.00918***	0.001335	-6.878	9.21E-12
Piedmont	-0.01513***	0.002498	-6.056	1.80E-09
Boreal Hardwood Transition	-0.01178***	0.001988	-5.924	3.98E-09
Central Hardwoods	0.00941***	0.001893	4.971	7.52E-07
Lower Great Lakes St. Lawrence Plain	-0.00666**	0.002039	-3.268	0.00111
Cropland	-0.008268**	0.002683	-3.082	0.0021
Peninsular Florida	-0.00697*	0.002801	-2.49	0.01288

Woodland:

The F-statistics, table 1, ($F = 9.552$) for encounter rates in Woodland landscapes is much larger than the tail of $F_{0.95, 12, 1982} = 1.79347$ for decreasing landcover, and for increasing landcover of Woodlands the F-statistics ($F = 69.48$) is also larger than the tail $F_{0.95, 5, 2551} = 2.375416$. Therefore, we reject the Null hypothesis, and suggest that Encounter rates decline with Woodland decrease, and increase with Woodland increase. Table 9 shows Mixed forests and Deciduous Broadleaf both have positive coefficients with an increasing slope ($P < 0.01$). Additionally, mixed forests have a t-statistics of zero slope at $t = 17.747$, this is much larger than $t_{0.975, 2551} = 1.960894$. The following illustrates that for the given sample size, the standard error is smaller than the estimate relative to the null hypothesis. Suggesting that there is linear association with encounter rates increasing with mixed forest, with a greater likelihood than deciduous broadleaf ($t = 17.747_{MF} > 7.961_{DB}$). Given this, Both Eastern tallgrass prairie, and New England mid Atlantic coast have negative estimates, suggestive of declines in this region, whilst, Southeastern coastal plain has a positive increase with encounters. This largely supports the hypothesis that encounter rates increased with increasing woodlands, though two regions individually reject this.

Table 9. Woodland with increasing pland landscape-metric since 2010, from 2011 – 2019. These results are coefficients from a linear model, these include the Estimate, Standard Error, t-test, and P-value (P < α where α = 0.05) In R-studio. Signif. codes: 0 .0001'***', 0.001 '**', 0.01 '*'

Variables	**Estimate**	**Std.Error**	**t**	**Pr(>|t|)**
(Intercept)	0.01932***	0.000779	24.79	2.00E-16
Mixed Forest	0.05542***	0.003123	17.747	2.00E-16
Deciduous Broadleaf	0.022354***	0.002808	7.961	2.55E-15
Southeastern Coastal Plain	0.00554***	0.001457	3.802	0.000147
Eastern Tallgrass Prairie	-0.01991***	0.005935	-3.355	0.000805
New England Mid-Atlantic Coast	-0.00869*	0.003993	-2.176	0.029632

Table 10 illustrates that both mixed forest and deciduous broadleaf supports the hypothesis, which states that the probability for encounter rate decline is related to woodland decline. This follows with several BCR regions, namely Eastern tallgrass prairie, New England mid Atlantic coast and Mississippi Alluvial valley. This infers that both Eastern tallgrass prairie and New England mid Atlantic coast are experiencing declines in encounters for declining woodlands, but also increases in areas with increasing woodland habitats. However, encounter rates are not declining relative to landcover decline for several regions, which shows most regions reject the hypothesis.

Table 10. Woodland with decreasing pland landscape-metric since 2010, from 2011 – 2019. These results are coefficients from a linear model, these include the Estimate, Standard Error, t-test, and P-value (P < α where α = 0.05) In R-studio. Signif. codes: 0 .0001'***', 0.001 '**', 0.01 '*'

Variables	**Estimate**	**Std.Error**	**t**	**Pr(>|t|)**
(Intercept)	0.025195***	0.001154	21.838	2.00E-16
Mixed Forest	-0.0219***	0.004123	-5.311	1.21E-07
Mississippi Alluvial Valley	0.016178**	0.004956	3.265	0.00112
Eastern Tallgrass Prairie	0.02047***	0.005117	4.001	6.53E-05
Piedmont	-0.01592***	0.00406	-3.92	9.16E-05
Deciduous Broadleaf	-0.00973**	0.003067	-3.171	0.00154
Central Hardwoods	-0.00965***	0.001834	-5.261	1.58E-07
Oaks and Prairies	-0.03006**	0.010908	-2.756	0.00591
Prairie Potholes	-0.02196*	0.009397	-2.337	0.01954
New England Mid-Atlantic Coast	0.012**	0.004253	2.822	0.00482
Appalachian Mountains	-0.00651***	0.001572	-4.14	3.62E-05
Boreal Hardwood Transition	-0.00562***	0.001311	-4.289	1.88E-05

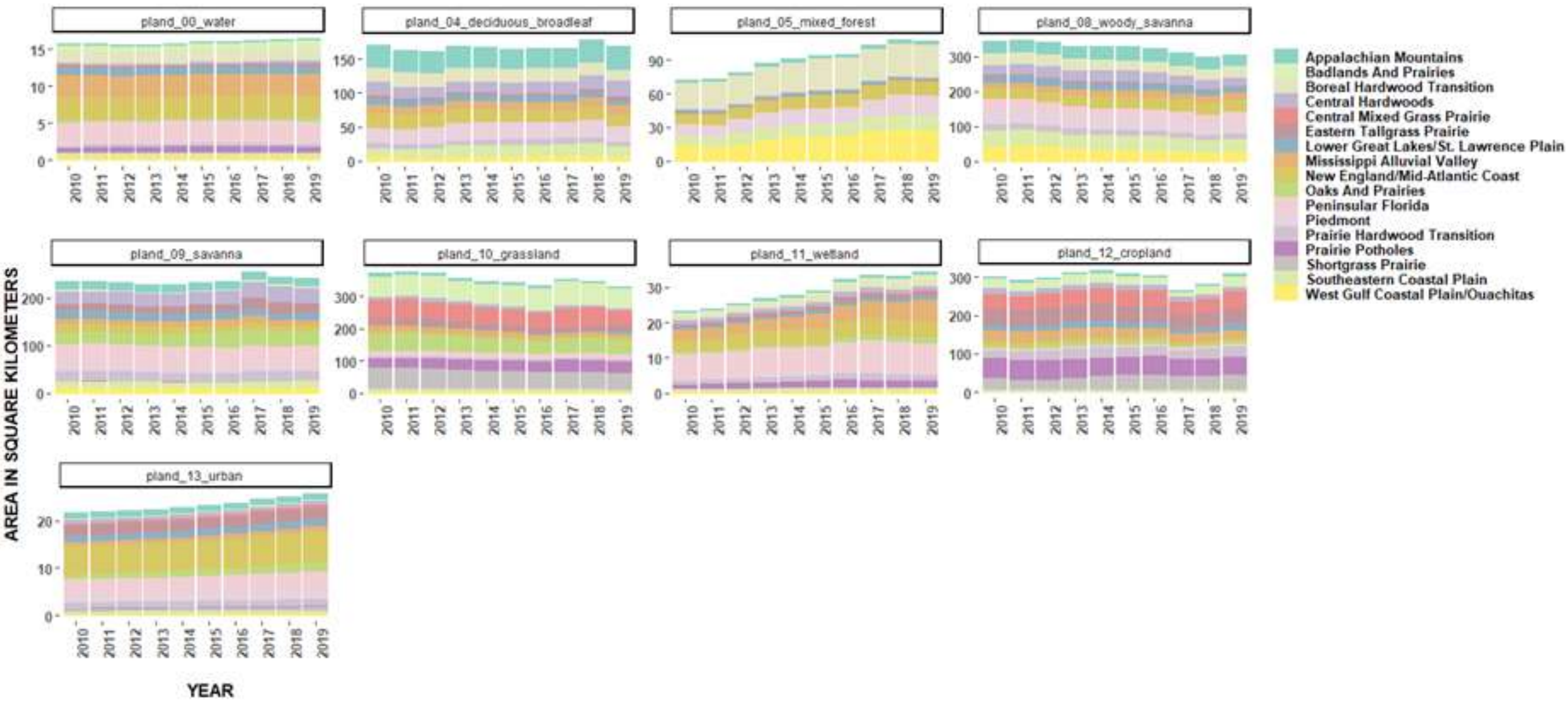

Figure 4. Multiple graphs of MODIS MCD12Q1 landcover types across the spatial zones of North America, that are given in the legend. The Y axis represents the area calculated, represented by the pland metric, such that for each region, all habitat types for that region accumulate to 100% pland coverage.

In figure 4, an overview of pland-percentage coverage across each BCR region for various landcover classes is shown with variable changes since years 2010-2019. Waterbodies, both water and wetland, are increasing, though wetland has a noticeable increase in percentage of pland coverage mostly in Piedmont and Peninsular Florida. Additionally, urban extent has gradually been increasing, distributed across several regions. However, there was no support for declines over recent years. Although, about table 6, whilst urban areas have shown increase in urban extent since 2010, it’s clear that there is not any coherent view that the Red-headed woodpecker is declining in urban areas. Hence, increases in urban extent supports the null hypothesis. As for the woodland habitats, these are deciduous broadleaf and mixed forests, each are showing increases in coverage since 2010, most significantly in mixed forests. This is predominantly increasing in the West Gulf Coastal Plain (BCR 25), and Badland and Prairies (BCR 17), with an overall increase of 30 km^2.

There are two landcover classes for savannas, these are considered top priority in understanding the Red-headed Woodpeckers distribution (Koenig et al. 2017), where the lack of savanna can impact their range expansion. It is shown that Woody Savanna has been decreasing since 2010, whilst Savannas have remained stable, with a slight increase in 2017 for Oaks and Praries, Eastern Tallgrass Prairie and Central Hardwoods. Furthermore in 2014, the Shortgrass prairie also increased by a marginal amount, then flattens thereafter.

Grasslands begin decreasing after 2012 up until 2016, then a gradual increase from 2017. These fluctuations in percentage coverage are mainly due to changes in Eastern Tallgrass Prairie. Compared to Croplands, there is a stable coverage until 2016, which it then begins to drop in 2017. This is resultant of a decrease in coverage across Prairie Potholes, with an increase in Shortgrass Prairie.

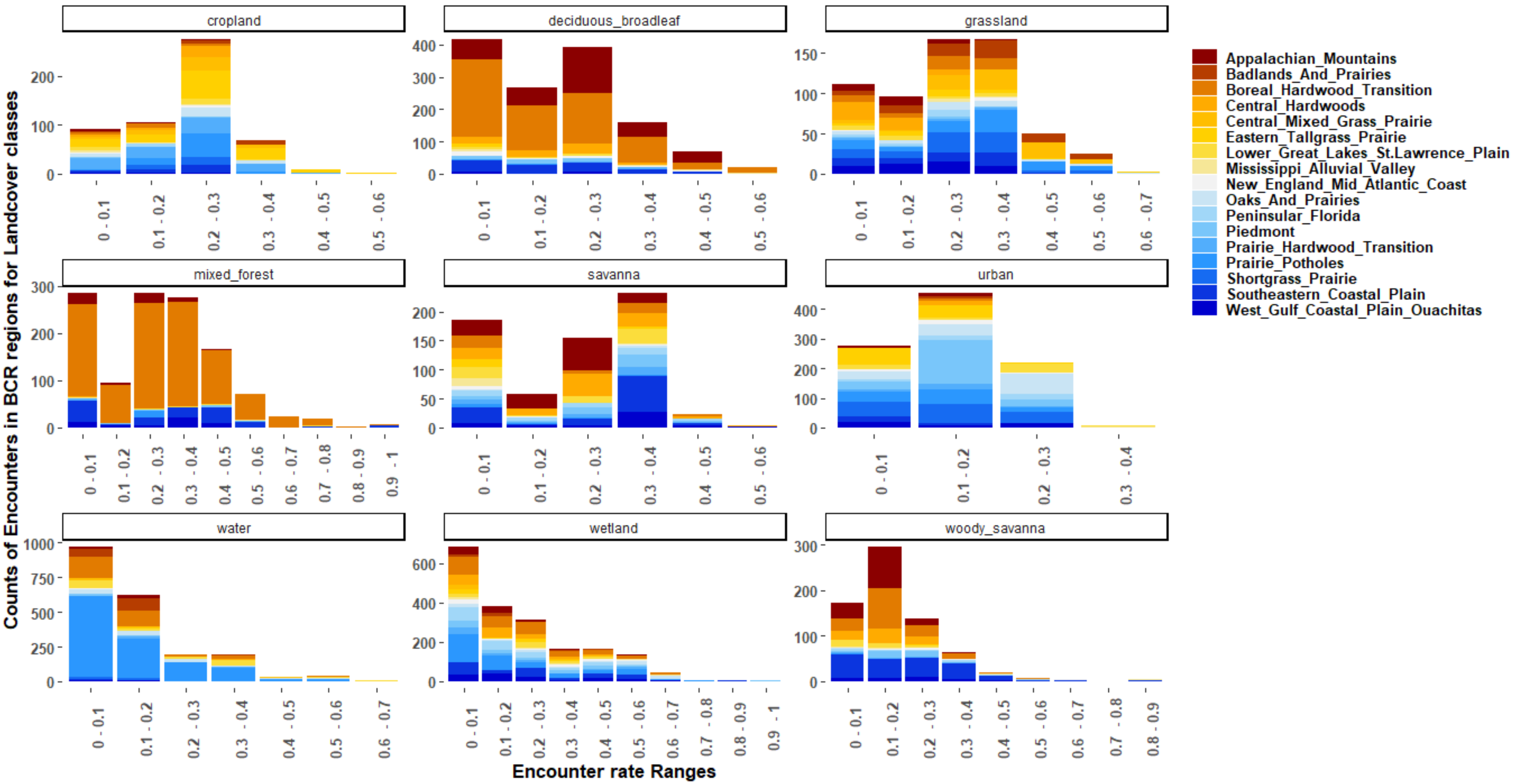

figure 5. Graph of landcover classes across the BCR regions with increasing landcover since 2010. This represents encounter rate ranges across a sequence of 10 ranges, these are: 0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, 0.4-0.5, 0.5-0.6, 0.6-0.7, 0.7-0.8, 0.8-0.9, 0.9-1. These are calculated as counts within each range.

Waterbodies both show higher probabilities of encounter within the 0-0.1 range. These counts suggest that probabilities are low, meaning that predicted observations are minimal although these observations are more prominent in Prairie potholes. It's noticeable that Appalachian Mountains, Badlands and Prairies, and Boreal hardwood transition, effectively decrease probabilities after the encounter range 0.1-0.2 in Water. As for Wetlands, a similar trend is consistent throughout until 0.5-0.6 like the first two probability ranges in Water. However, Wetlands are most prominent across multiple BCR regions, and its recognisable that Southeastern Coastal plain is prominent in range 0-0.1, however it is removed in the next encounter range and re-introduced in 0.2-0.3. Furthermore, Wetland have a minimal frequency for probability from 0.7-1, suggestive that in a 5 x 5 Modis cell, there's a very high probability in Piedmont. Although, the count is marginally small to derive an accurate interpretation.

Woodlands shows a similar trend, with a large probability increasing in 0-0.1 and a drop in ranges 0.1-0.2, then an increase from 0.2-0.3. Although, this increase remains consistent in mixed forests with probability ranges thereafter, until 0.5-0.6, whilst it drops in deciduous broadleaf after 0.2-0.3. As for the BCR regions, they both have consistency in high levels of encounter in Boreal hardwood transition, although the Appalachian Mountains have low probability of encounter in Mixed forests, whilst this amplifies in frequency for deciduous broadleaf. This is portrayed similarly in table 10, where encounter declines are negatively correlating with woodland decline in Appalachian Mountains and Boreal hardwood transition inferring that probabilities are more consistent with non-decreasing landscapes.

Urban landcover experiences the highest counts for probability in the 0.1-0.2 range, this evens out in close comparison to one another around 0-0.1, and 0.2-0.3. There's an exponential increase in encounter rate probabilities for Piedmont that transitions from 0-0.1 to 0.1-0.2, then decreases at the next probability range, however, encounters are visibly more frequent for Oaks and Prairies relative to previous ranges. Altogether, the probabilities of encounter are small.

As for the Savannas, woody Savannas have a high probability within ranges 0.1-0.2, whilst Savannas have greater probabilities in the range 0.3-0.4. Although, woody Savannas have probabilities as great as 0.8-0.9, their frequency is <5. Between the two, Savannas show very low probabilities in the range 0.1-0.2 comparative to woody Savannas. However, the lowest probability is high in both landscape types. It is noticeable that the southeastern coastal plain has highest probabilities of encounter in Savannas, whilst in Woody Savanna this is distributed between Apalachian mountain and the Boreal hardwood transition.

Finally, Cropland and Grassland, both show a strong distribution within the 0.2-0.3 probability range. Although, Grassland also has high values for ranges of 0.3-0.4, this is formed by the increased frequency of observed probabilities in Prairie Potholes. Grassland seem to be present in many BCR regions, whilst Cropland have their greater distribution in the Prairie landscapes.

Waterbodies have high frequency counts in the lower probability range, especially with Water which peaks at 80 counts for 0-0.1. Whilst wetlands have a higher spatial range, the counts are greater by comparison to Water. Although, the probabilities for Wetland are greater in the lowest range, with a greater frequency than increasing wetlands in figure 5, the accumulation of frequency and range probabilities are less than figure 5. However, wetlands seem to have a large proportion of frequency, up to 100 counts, within the range from 0.4 – 0.7.

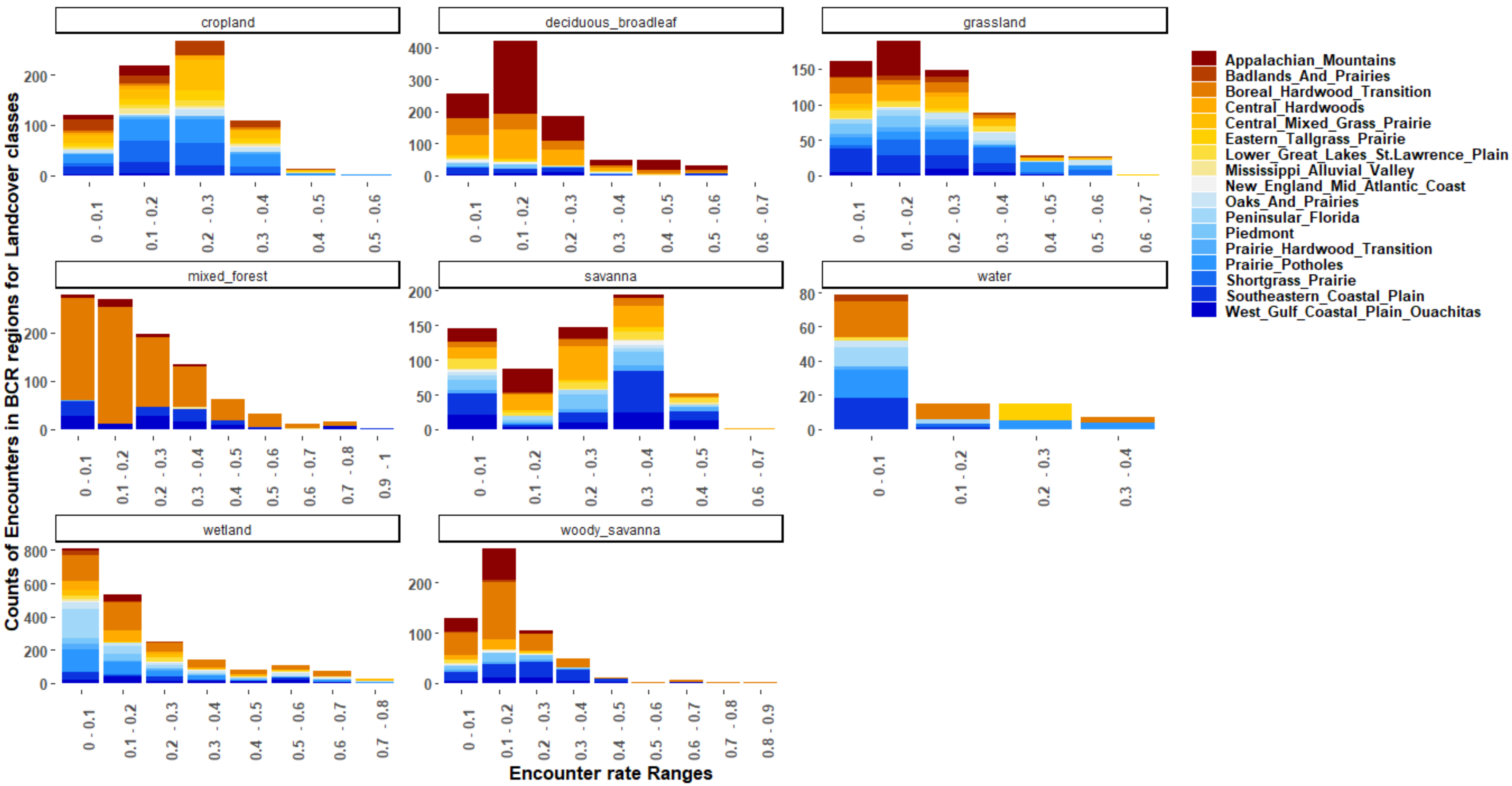

figure 6. Graph of landcover classes across the BCR regions with decreasing landcover since 2010. This represents encounter rate ranges across a sequence of 10 ranges, these are: : 0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, 0.4-0.5, 0.5-0.6, 0.6-0.7, 0.7-0.8, 0.8-0.9, 0.9-1These are calculated as counts within each range.

Mixed forests show a large range of probabilities for encounter from 0-0.4, however, it decreases from 0 – 0.1, though the majority of mixed forests that decline are in Boreal hardwood transition and Appalachian mountains, this is represented by the higher counts in lower probability ranges. Inferring that encounters are minimalised via encounter probability, hence if higher a higher frequency of counts is within a lower probability range, then it's apparent that declines in area of landcover decrease represent lower probabilities for encounter, for at least the majority of the landscape. This effectively supports the hypothesis, that areas of decline experience lower probabilities for encounter. Furthermore, the Deciduous broadleaf total area coverage for counts increases greatly in the Boreal Hardwood transition in figure 5, towards greater minimalised declines in figure 6, with an increase of landcover decline in the Appalachian mountains, but an increase in the probability ranges. Although, the probabilities of encounter are greater in the Appalachian mountains at higher ranges, from 0.1-0.2, whilst still relatively small, this suggests that areas of decline experience lower probabilities of encounter.

Croplands are increasing in probabilities and have a higher likelihood for encounters within the range 0.2-0.3. Compared with table 8, encounter rates declined with BCR regions, Central Hardwoods and Appalachian mountains, as these had ranges between 0.1-0.2, though Central Hardwoods had ranged to 0.2-0.3. There was higher common presence for encounters in grasslands as ranges were between 0 – 0.7. Encounter rates were largely distributed amongst lower probability ranges, suggesting higher counts with lower probabilities. This supports the hypothesis that birds have been declining in grasslands, however, there is a marginal distribution of probability encounters >0.4, suggesting encounter rates in BCR regions, Prairie Potholes and Shortgrass Prairie, have increased in these regions. Encounter rates in savannas were high, as they ranged between 0-0.7 and were distributed in 0.3-0.4. Although, bird encounters ranged lowest in BCR region, Appalachian mountains (mostly 0.1-0.3) and highest in Central Mixed Grass Prairie (0.0-0.7.) Woody savannas had a high range of encounter rates (0-0.9), however, they had lower encounter rates (mostly 0-0.3), including of BCR regions, Appalachian mountains and Boreal Hardwood Transition. Overall, figure 6 rejects the hypothesis that birds are declining in savannas.

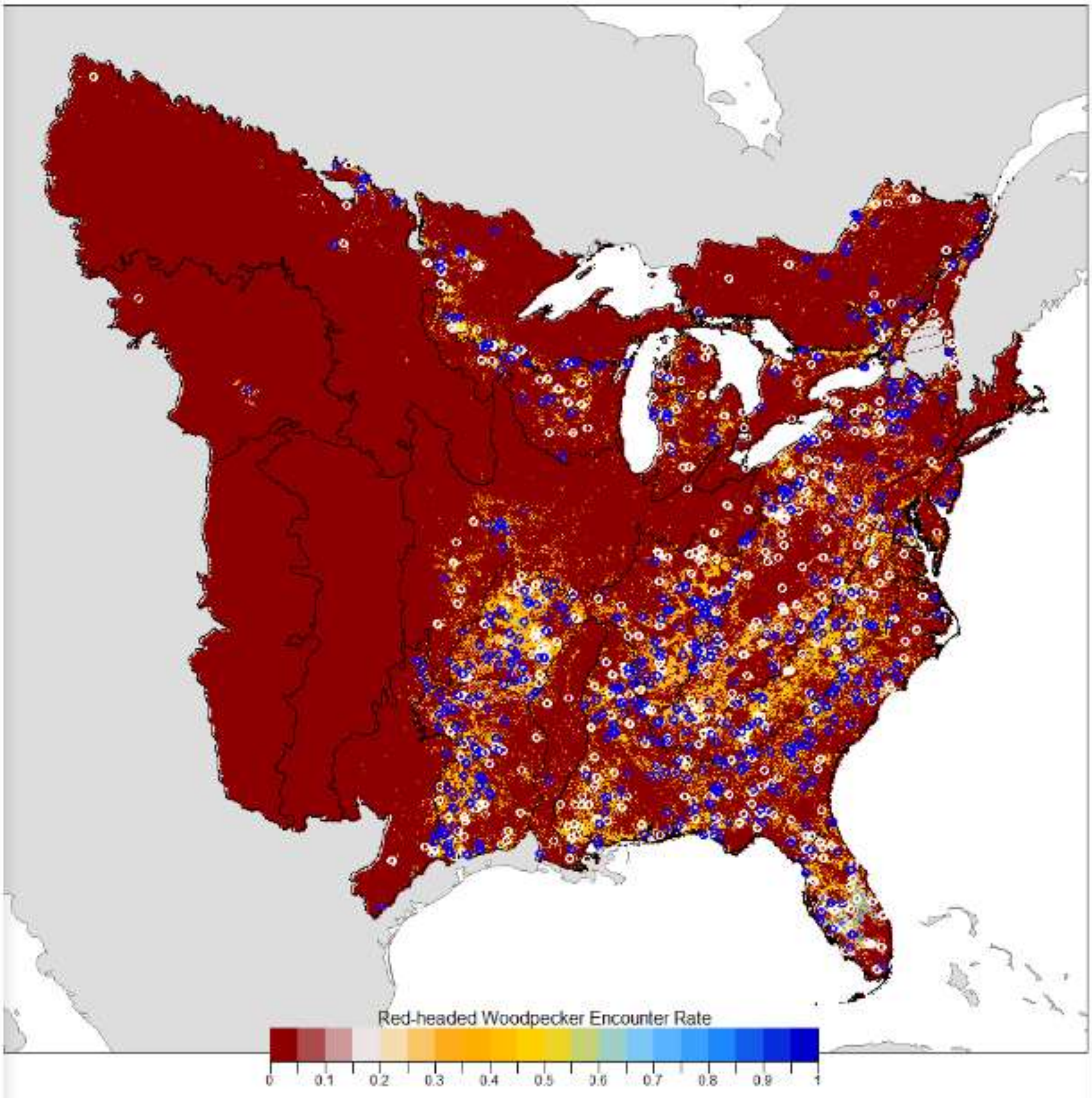

Figure 7. Probability of encountering the Red-headed Woodpecker across the BCR ranges for the Savanna Landcover class. The red range represents the lowest probability range (fixed at 0), This scale ranges from 0 to 1, and the white points represent the areas of encountering the Red-headed Woodpecker in areas decreasing in Savannas since 2010. Whilst the blue points represent the probability of encountering at areas increasing in Savannas since 2010.

In figure 7, encounter rates in savannas were distributed between Southern and Eastern USA. Including probability = 0.2 in Florida, although a greater distribution of probability for encounters are within the range 0.3 – 0.5, mainly across BCR regions: Peninsular Florida, South eastern coastal plain, Appalachian mountains, Piedmont, Central hardwoods and West gulf coastal plain. Blue areas indicate a wide distribution of increases in encounter probability across these BCR regions. Overall, this supports the hypothesis that encounters increased with savannas.

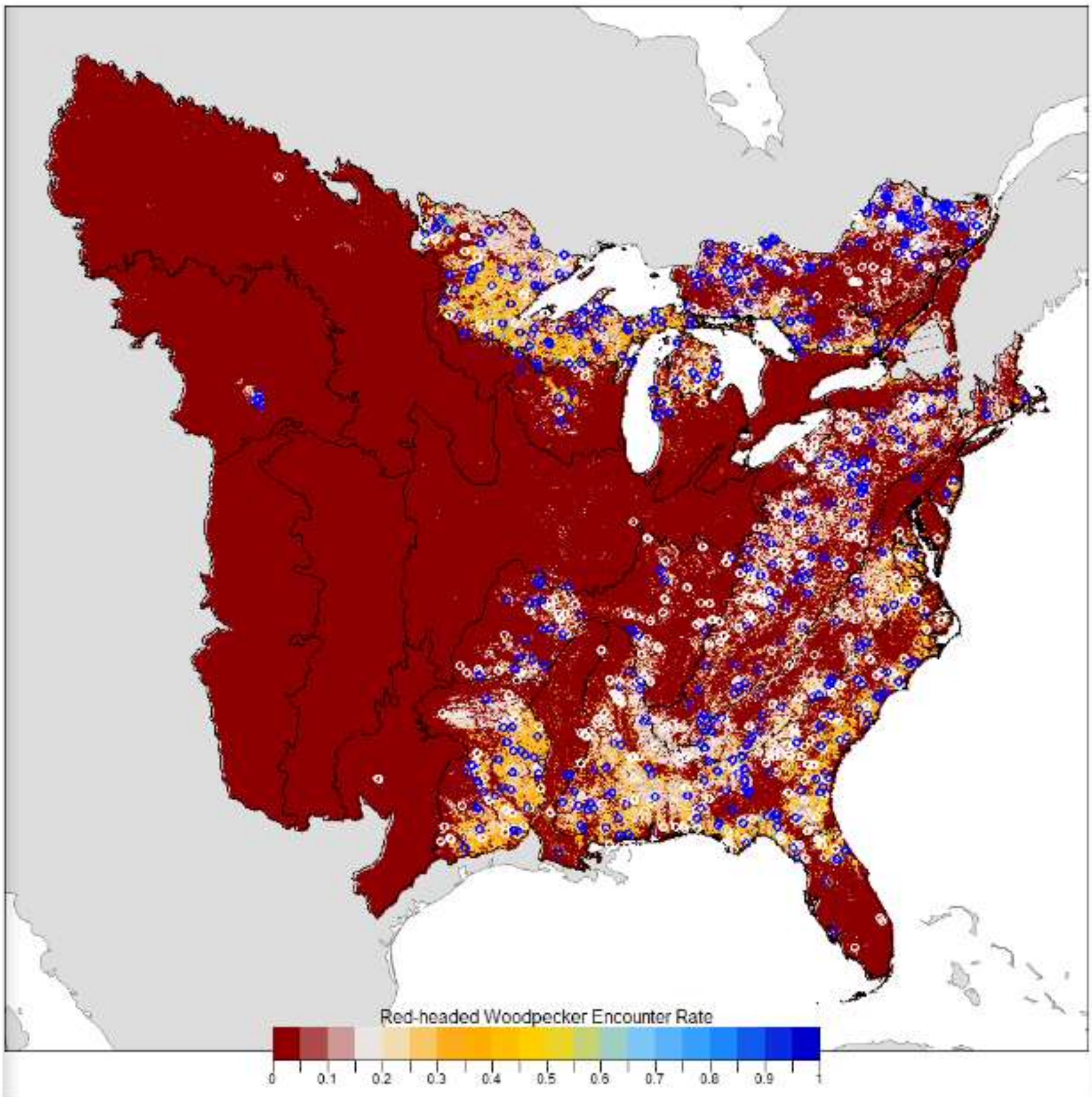

Figure 8. Probability of encountering the Red-headed Woodpecker across the BCR ranges for the Woody Savannas Landcover class. The red range represents the lowest probability range (fixed at 0), This scale ranges from 0 to 1, and the white points represent the areas of encountering the Red-headed Woodpecker in areas decreasing in Woody Savannas since 2010. Whilst the blue points represent the probability of encountering at areas increasing in Woody Savannas since 2010.

In figure 8, encounter rates in woody savannas were distributed along Northern, Eastern and Southern USA. Larger proportions of lower encounter probabilities were distributed amongst declining woody savannas in Boreal hardwood transition; South-eastern Coastal Plain; Mississippi, Alluvial Valley; Appalachian Mountains; Badlands and Prairies Highest encounter probabilities were in the Southern range of BCR regions: South Eastern Coastal Plain, Boreal Hardwood Transition and West Gulf Coastal Plain (0.5 in a range of $0 < x < 1$.) Landcover increases were distributed across the regions, although higher landcover declines occurred. Overall, the hypothesis that encounter rates decreased with woody savannas decline was accepted.

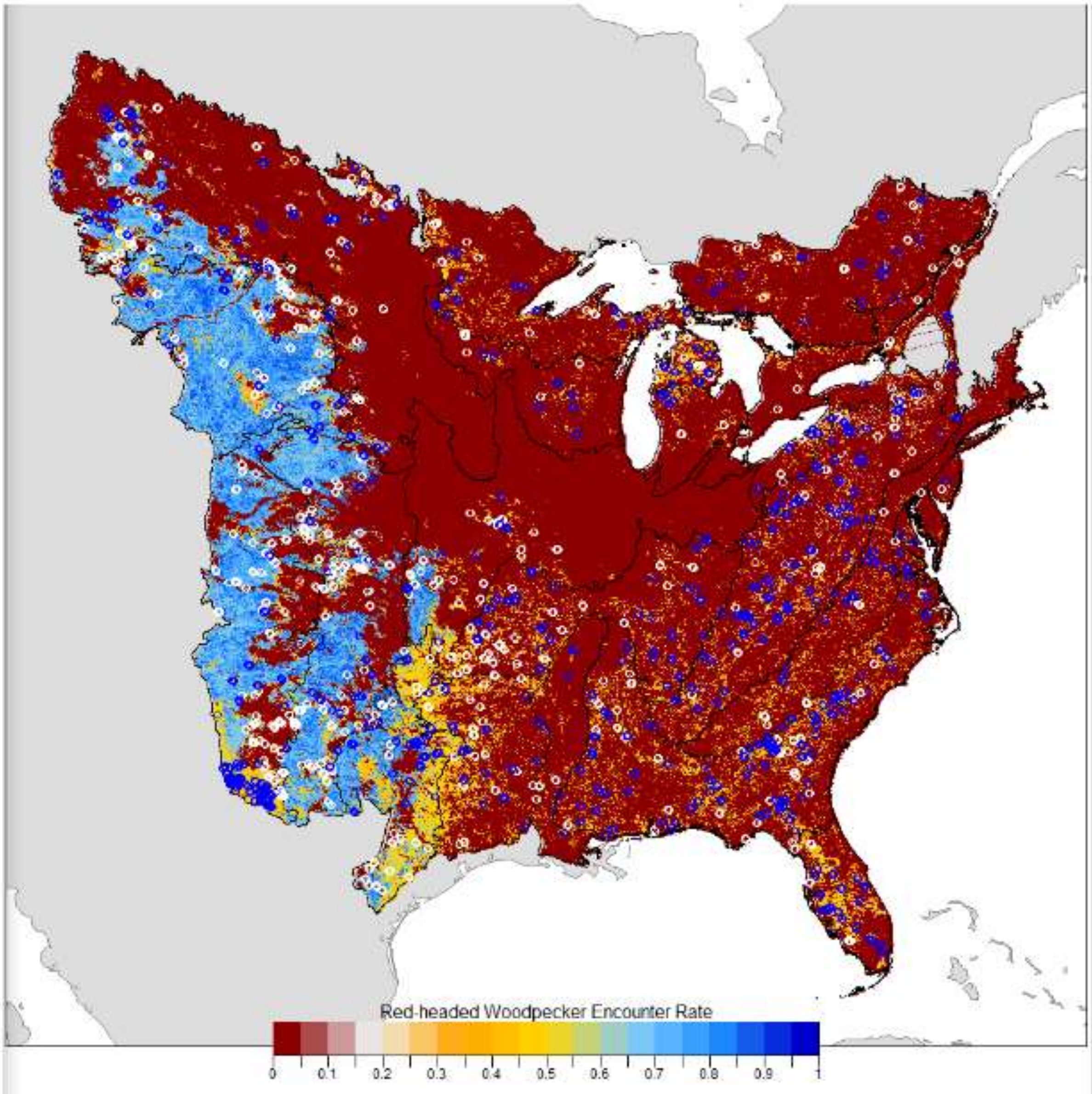

Figure 9. Probability of encountering the Red-headed Woodpecker across the BCR ranges for the Grassland Landcover class. The red range represents the lowest probability range (fixed at 0), This scale ranges from 0 to 1, and the white points represent the areas of encountering the Red-headed Woodpecker in areas decreasing in Grassland since 2010. Whilst the blue points represent the probability of encountering at areas increasing in Grassland since 2010.

In figure 9, encounter rates in grasslands were distributed across the USA apart from central USA. Encounter rates were highest (0.6-1) along the West coast across BCR regions Badlands and Prairies and Shortgrass Prairies. Eastern and Southern BCR regions including Appalachian Mountains and Florida, had increases in landcover and declines were distributed in South Western and Western USA. Overall, the hypothesis regarding years was rejected.

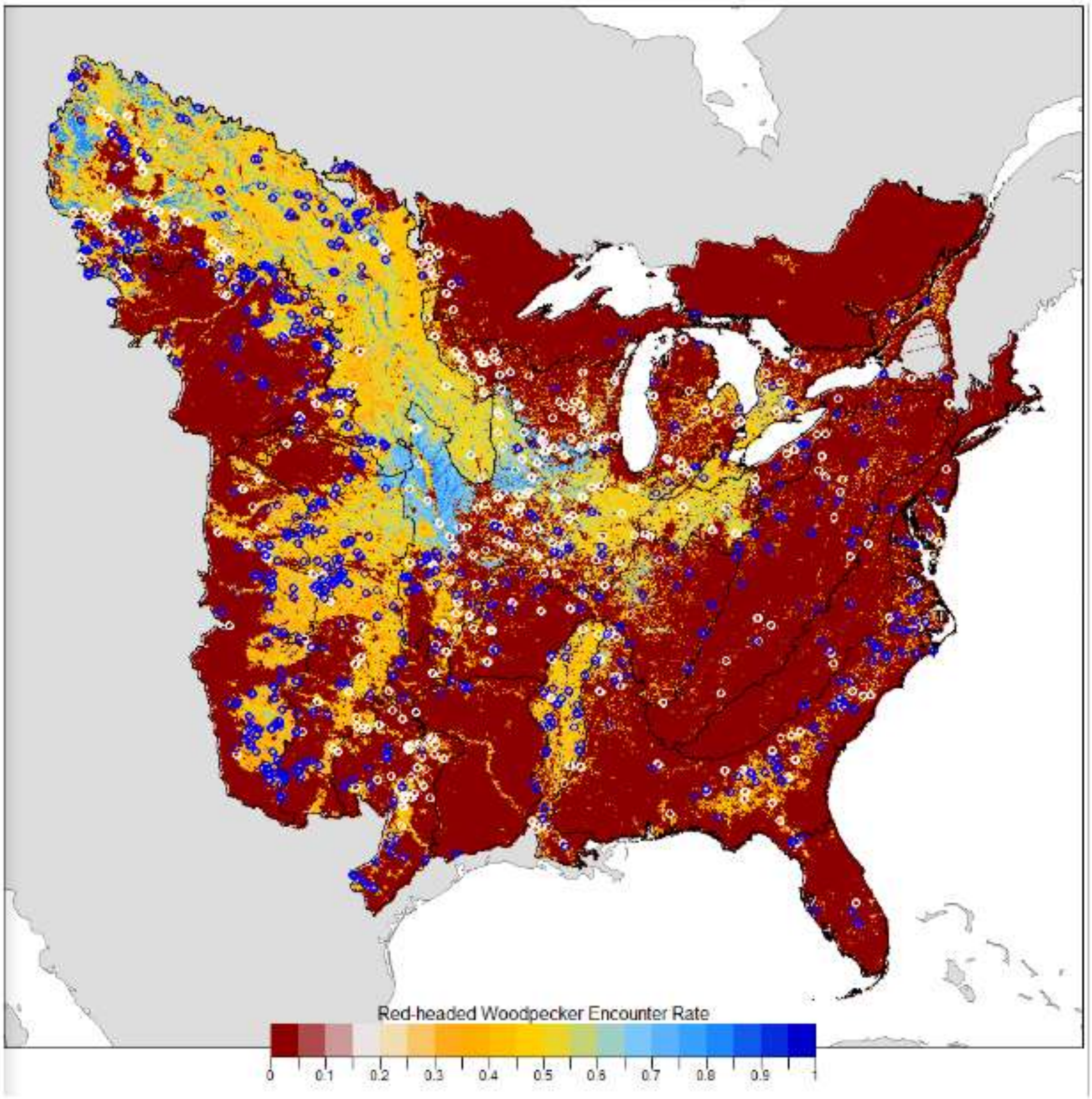

Figure 10. Probability of encountering the Red-headed Woodpecker across the BCR ranges for the Cropland Landcover class. The red range represents the lowest probability range (fixed at 0), This scale ranges from 0 to 1, and the white points represent the areas of encountering the Red-headed Woodpecker in areas decreasing in Cropland since 2010. Whilst the blue points represent the probability of encountering at areas increasing in Cropland since 2010.

In figure 10, encounter rates in croplands were distributed across Central, North Western and Western USA, the highest in BCR regions Eastern Tallgrass Prairie and Prairie Potholes (0.6-0.9.) There are a multitude of habitat increases towards the west, with greater declines towards the centre and distributed around Eastern Tallgrass Prairie, where higher probability values are present. Overall, this seems to accept the hypothesis.

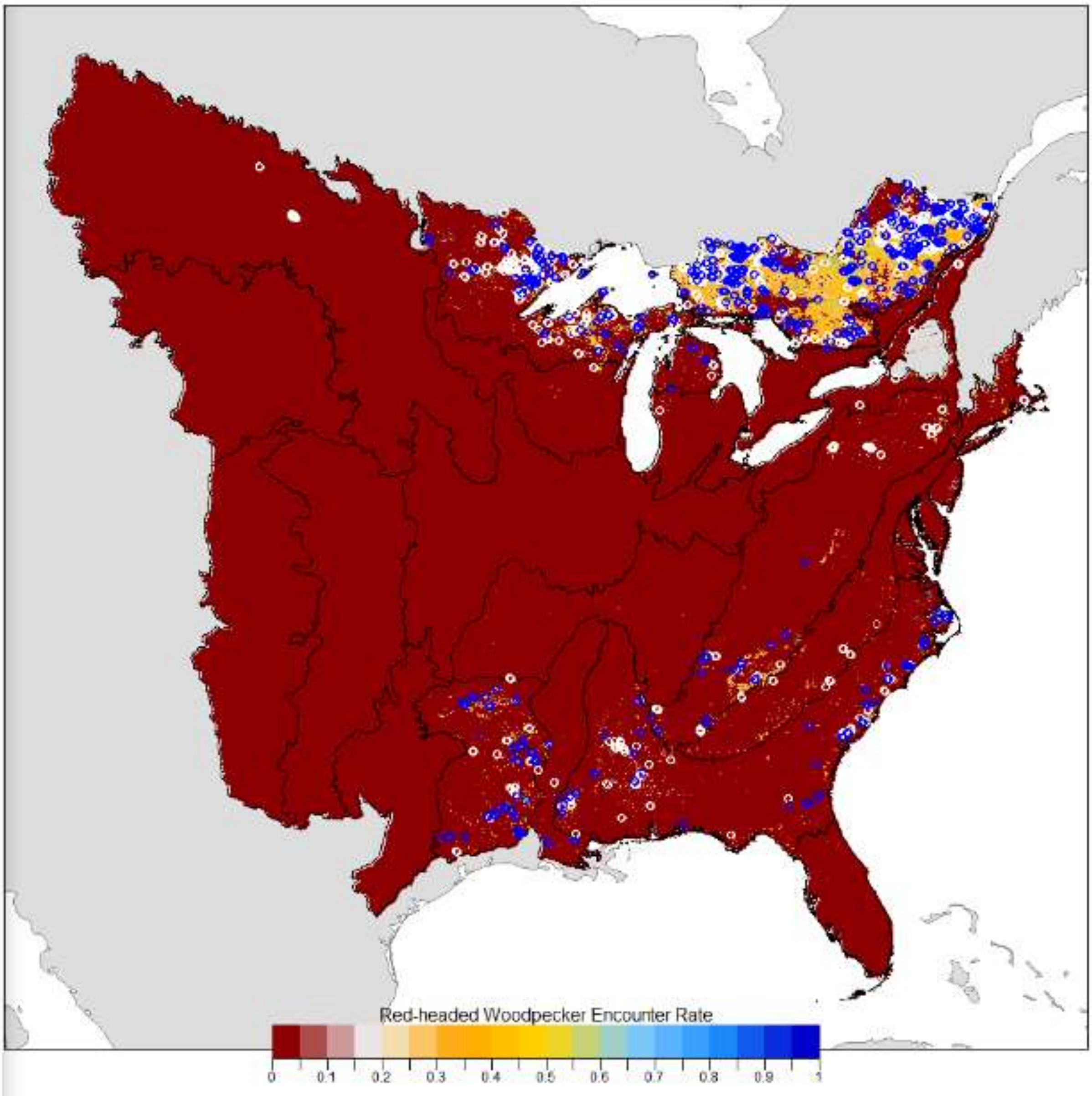

Figure 11. Probability of encountering the Red-headed Woodpecker across the BCR ranges for the Mixed Forest Landcover class. The red range represents the lowest probability range (fixed at 0), This scale ranges from 0 to 1, and the white points represent the areas of encountering the Red-headed Woodpecker in areas decreasing in Mixed Forest since 2010. Whilst the blue points represent the probability of encountering at areas increasing in Mixed Forest since 2010.

In figure 11, encounter rates in mixed forests were distributed in Northern, Eastern and Southern USA, the highest in BCR region Boreal Hard Wood Transition (0.2-0.5) which also had the most habitat increases and decreases. Towards the South-east there's a mixture of increasing and decreasing landcover, though more prominent towards increases since 2010. However, the probabilities for encounter here are lower, especially In the Southeastern Coastal Plain. This accepts the hypothesis when declines in habitat correlate with encounter declines, and inversely, when increasing in landcover then encounters increase.

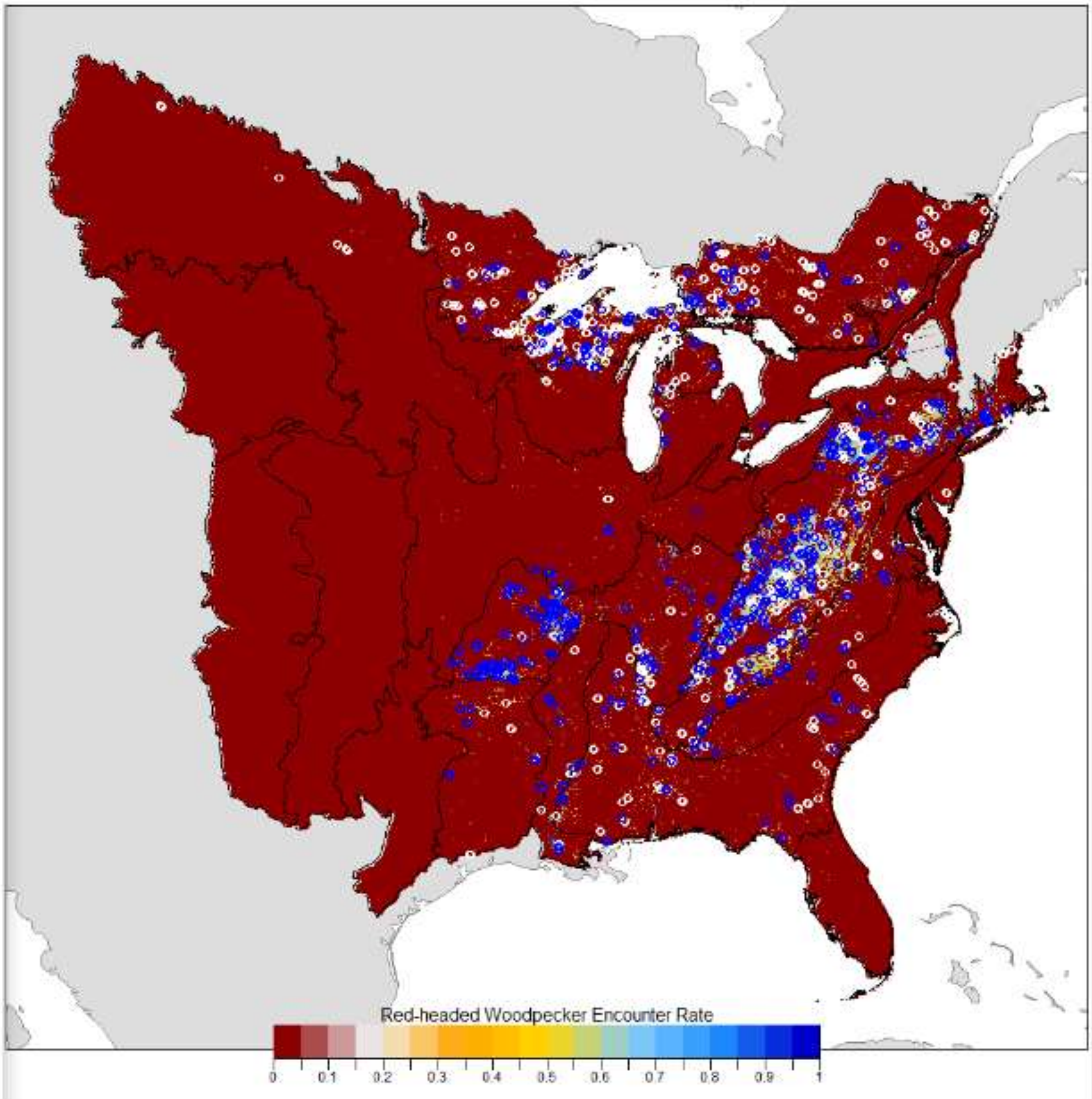

Figure 12. Probability of encountering the Red-headed Woodpecker across the BCR ranges for the Deciduous Broadleaf Landcover class. The red range represents the lowest probability range (fixed at 0), This scale ranges from 0 to 1, and the white points represent the areas of encountering the Red-headed Woodpecker in areas decreasing in Deciduous Broadleaf since 2010. Whilst the blue points represent the probability of encountering at areas increasing in Deciduous Broadleaf since 2010.

In figure 12, encounter rates in deciduous Broadleaf forests were distributed in Northern, Central, Eastern and Southern USA. Encounter rate probabilities were highest across the Appalachian Mountains as were habitat landcover increases, which were greater than decreases. There's a large density of habitat increase in Central Hardwoods with minimal decreasing landcover since 2010. However, a low frequency of probabilities is apparent here. Furthermore, a large decline in landcover is apparent in the Boreal Hardwood Transition. This supports the hypothesis that encounters increased with landcover.

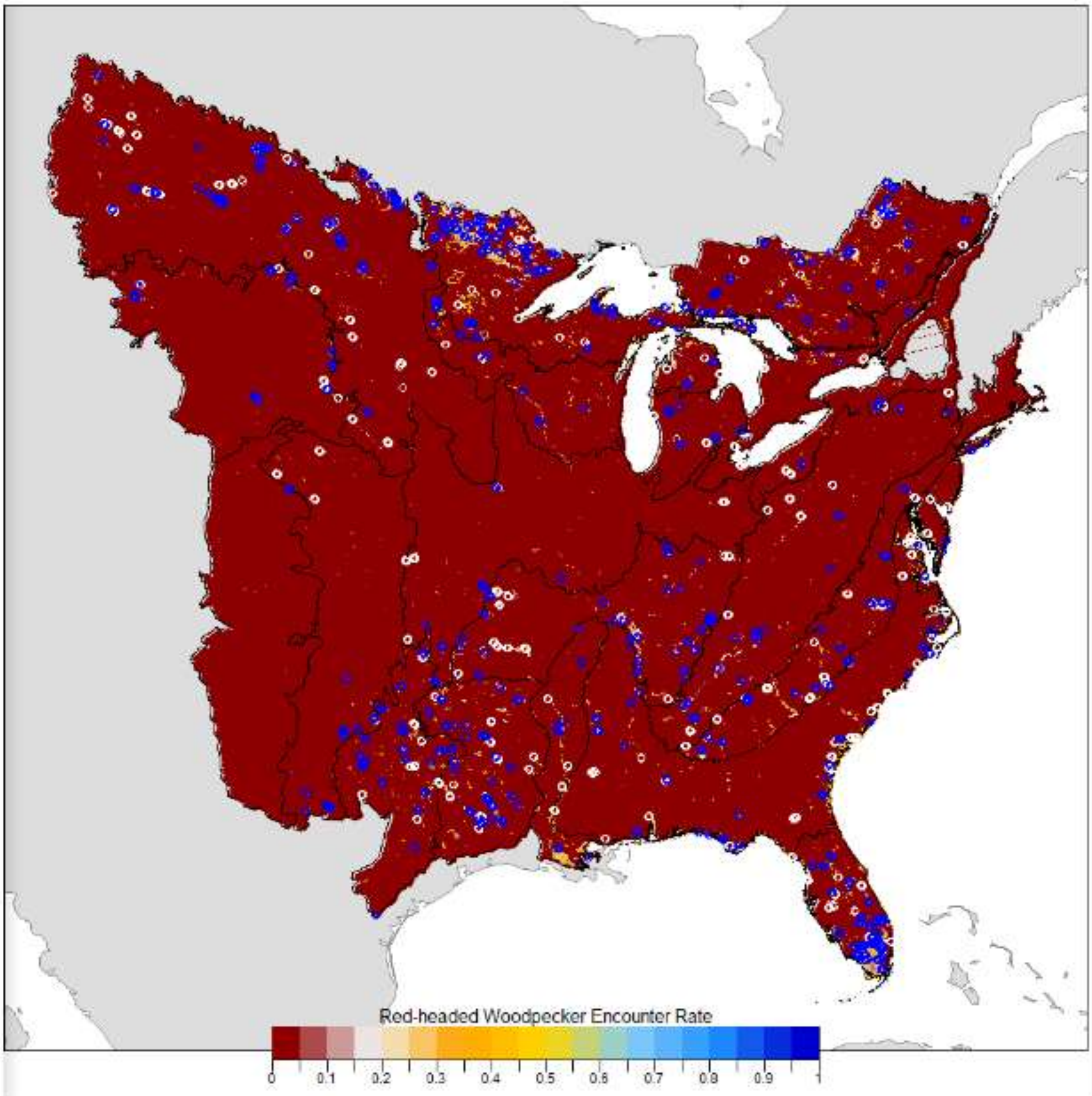

Figure 13. Probability of encountering the Red-headed Woodpecker across the BCR ranges for the Wetland Landcover class. The red range represents the lowest probability range (fixed at 0), This scale ranges from 0 to 1, and the white points represent the areas of encountering the Red-headed Woodpecker in areas decreasing in Wetland since 2010. Whilst the blue points represent the probability of encountering at areas increasing in Wetland since 2010.

Figure 13 shows relatively high proportion of encounters distributed across North America for increasing Wetland landcover since 2010, comparatively to decreases. These are densely populated in Peninsular Florida and Boreal Hardwood Transition, and widely scattered distributions elsewhere. This supports the hypothesis that encounter rates increased with wetland presence and with increasing landcover, through the illustration of probability ranges falling within 0.3-0.5, primarily with areas increasing in Wetlands.

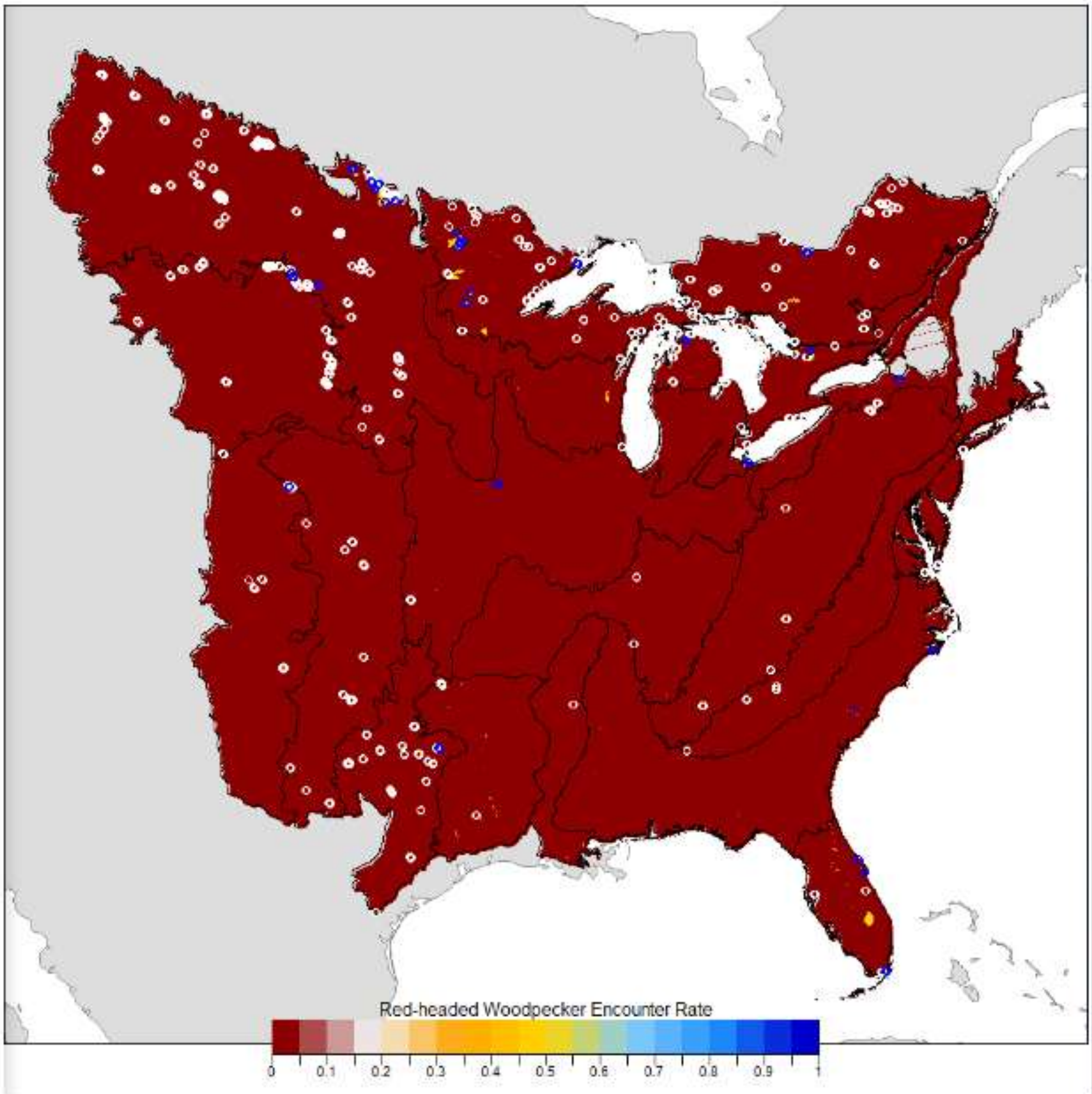

Figure 14. Probability of encountering the Red-headed Woodpecker across the BCR ranges for the Water Landcover class. The red range represents the lowest probability range (fixed at 0), This scale ranges from 0 to 1, and the white points represent the areas of encountering the Red-headed Woodpecker in areas decreasing in Water since 2010. Whilst the blue points represent the probability of encountering at areas increasing in Water since 2010.

In figure 14, encounter rates with water landcover were scattered across the USA, largely in the West. Encounter rate probabilities were very low and landcover decreases being high. This rejects the hypothesis that encounter rates increase with landcover.

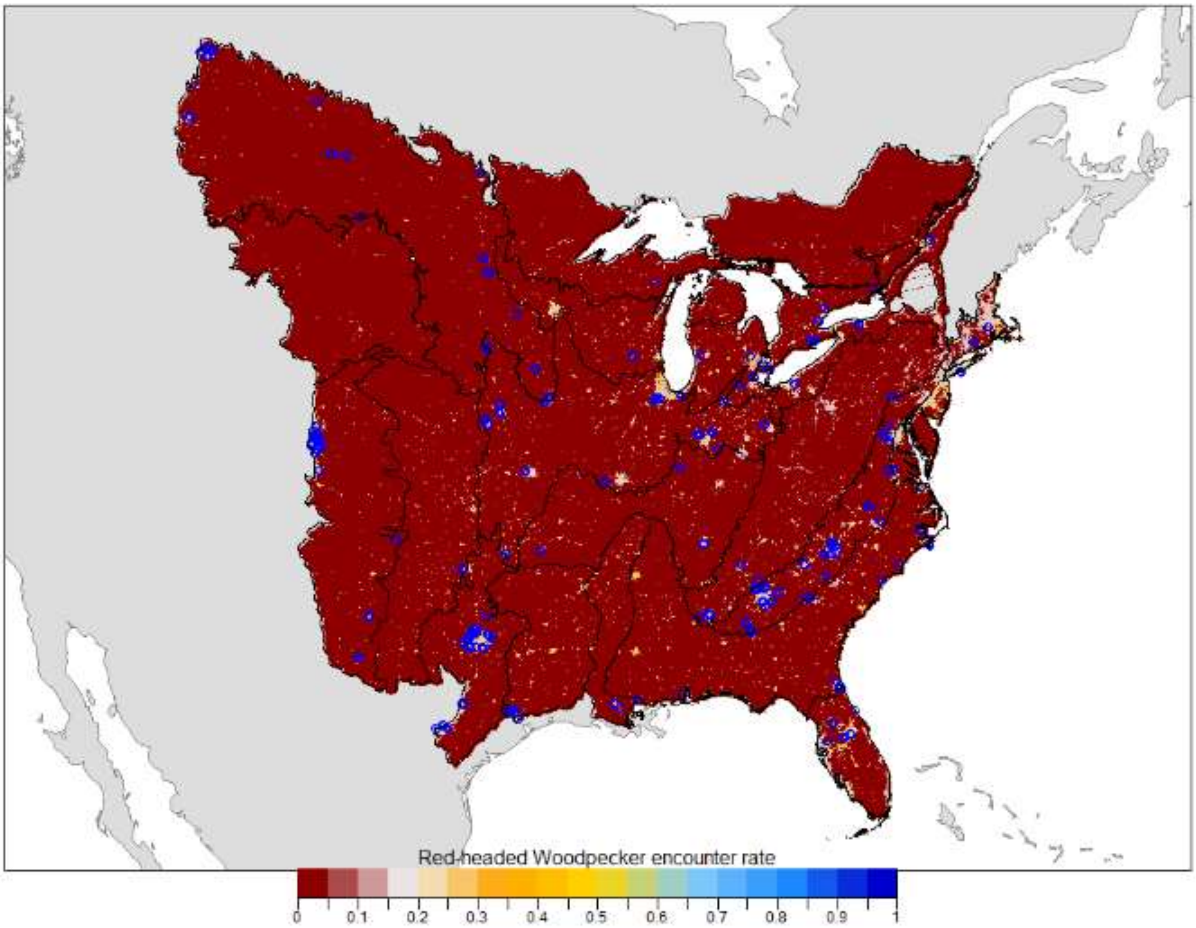

Figure 15. Probability of encountering the Red-headed Woodpecker across the BCR ranges for the urban Landcover class. The red range represents the lowest probability range (fixed at 0), This scale ranges from 0 to 1, and the white points represent the areas of encountering the Red-headed Woodpecker in areas decreasing in urban since 2010. Whilst the blue points represent the probability of encountering at areas increasing in urban since 2010.

In figure 15, encounter rates with urban landcover were scattered across the USA, largely in the centre of Oaks and Prairies, and distributions along Appalachian mountains. Encounter rate probabilities were very low and landcover increases being high, with an average range spanning from 0.15-0.3. Furthermore, urban extent only has increasing landcover distributed since 2010, with higher densities of changes in larger urban areas. This rejects the hypothesis that encounter rates increase with landcover.

<u>Discussion</u>:

Increasing landcover classes for the calibrated models are proportionally greater than random forest models; an explanation is that regression trees are subject to severe overfitting; hence the final tree for every node may not be the optimal solution (Evans et al. 2011). Therefore, the calibration builds towards correcting the error distribution within the regression, estimated by the coefficients, AUC, ROC, sensitivity, specificity, and Kappa (Bella et al. 2009). Additionally, these results are taken from table 2, conveying that the hypothesis for each model is accepted. However, multiple predictors for BCR regions may have negatively implicated the response variable because some regions have minimal values recorded as an explanation towards encounter rates. Therefore, selecting for encounter rates in regions above a threshold of counts present within a BCR region may improve the results; alternatively, using landcover changes as values in each BCR region and selecting threshold counts will provide a more significant prediction.

<u>Wetland</u>

Both Wetlands and Water landcover classes have decreasing encounter rates, given in table 2, compares to figure 6, where most probabilities are within the range of 0-0.1 chance of encountering the Red-headed Woodpecker. Inferring that within a given sample size, the probabilities of encounter are minor, and given that both Wetland and Water landcover classes are increasing in habitat coverage since 2010, figure 4 accepts the given hypothesis. However, Wetlands have probabilities ranging from 0.5-0.8 that relate to adverse estimates in table 2 for all BCR regions in the model. Inferring that while the entire range for both Water and Wetlands has declined, some selected BCR regions with declining land cover do not correlate with declines in encounters. Hence, those higher probabilities will attribute to these scenarios.

The explanation for the high frequency of low probabilities for encounter in wetland and water, figure 5. may be related to recent land-use changes, primarily in the rise for conservation towards preserving water bodies as biodiverse habitats in targeted regions. Hence, these changes may correlate with 'new' arrivals selecting nesting sites on newly formed waterbodies. Furthermore, the distinction between wetland and water is that the former is a combination of water and vegetation cover, while the latter is purely water cover. Therefore, wetlands provide enhanced habitat and resources and productivity towards biodiversity in flora and invertebrate abundance (Riffell et al. 2006); this is supported by table 3.

Furthermore, Kantrud et al. (1989) mention that many wetlands in the Prairie Pothole Region have been drained, with altered wetland productivity, although the recent Landsat data from MODIS shows an increase in wetland landcover since 2010, this includes the Prairie Pothole region, figure 4. Conservation strategies towards wetland formation in the conservation region intend to restore wetland habitats (Aronson and Galatowitsch, 2008). The increasing range of habitats for wetlands, improve diversity for various taxa (Leonard and Stout, 2006), and results from Nickley and Bulluck (2019), show that nest-sites in wetlands tended to be in high densities of medium/large snags, and these clusters provide a suitable site for the Red-headed Woodpecker when flooded. Currently correlated with table 3, the hypothesis shows that the woodpecker is increasing with wetlands; however, for the specific BCR region of Prairie Potholes, the adverse estimate rejects the hypothesis, table 2. Compared to figure 14, these results convey that not all decreasing landcover points in Prairie Potholes associate with declines in Encounter rates. Consequently, increasing habitats may provide different preferences and selection as opposed to declining wetlands in that region.

However, while the coefficients show a negative correlation for the specific region, it is unclear which landcover type the region correlates with towards the hypothesis. Finally, having multiple BCR predictors may implicate the distribution of predicted and observed errors.

Since 2011, Steven and Lowrance (2011) mention that approximately 60% of Piedmont and Coastal Plain East of Mississippi river has been forested, with most wetlands remaining in forested areas. In figure 4, since 2010, Piedmont has increased comparatively greater than other BCR regions. Figure 13 shows distributions within the probability range of 0.4 distributed across the center of Piedmont, with a greater proportion of observed encounter increases for increasing wetlands comparatively to decreasing. Table 2, Piedmont has a negative estimate, rejecting the hypothesis and inferring that encounter declines are not correlating with wetland declines. However, diversifying the types of wetlands available and distributed amongst North America, especially in specific BCR regions, may yield alternative results. Given this, Leonard and Stout's (2006) results showed that Woodpecker abundance correlated with spring-fed wetlands in Florida, and a similar correlation was considered by the results provided by table 2. However, it is challenging to maintain whether encounters are declining based on Wetlands and Water landcover changes or whether the dummy variables introduced by the BCR regions overfit the model.

Woodland

Woodland consists of mixed forests and deciduous broadleaf, and both are important indicators towards the presence of the Red-headed Woodpecker. Like the white-headed woodpeckers (Hollenbeck et al. 2011), their encounters are predominantly in sites with higher tree densities, with large diameters at breast height (Anderson and LaMontagne, 2016), which comprise open-canopy forests. Furthermore, in table 9, it is shown that mixed forests have a greater t distribution ($t = 17.747$) than deciduous broadleaf ($t = 7.961$), at $P < 0.001$. suggests that the dummy variables correlate with mixed forests, and the predictor strongly correlates with a positive slope. The following table accepts the hypothesis, although two BCR regions are inversely correlated, Eastern Tallgrass Prairie and New England mid-Atlantic coast. While observing both figure 5 and 13, it is noticeable that these two regions are predominantly in Deciduous broadleaf forests, though with small counts, hence it can be inferred that the smaller counts skew the distribution where a larger value would increase the standard error, and perhaps inversely correlate with the response. Therefore, a separate test of individual BCR regions with their respective landcover class while controlling for lower counts by either increasing the sample size or introducing a threshold would amplify the regression analysis.

Woodlands, chiefly mixed forests, have been increasing since 2010 to increase the population size of the Red-headed Woodpecker. In table 9, mixed forests positively support the hypothesis, although surprisingly, the Boreal hardwood transition is not apparent in the final model, while it represents most encounter rates for both decreasing and increasing land cover. Infers that deciduous forest, being also a positive predictor in table 9, skews the multiple regression towards regions with lower counts across the encounter ranges. A second inference indicates that habitat conversion through fires consistent with the estimate for the Southeastern Coastal Plain has an optimistic estimate in table 9. Because results from Vierling and Lentile (2006) reports that the Red-headed Woodpeckers nesting in recently burned pine-forests in the Black Hills, South Dakota, had a reproductive success of 47%, where fires create favorable habitats like open canopy forests (Nickley and Bulluck, 2019), that allow for aerial foraging (Kilgo and Vukovich, 2012), these complex habitats are also consistent in the Southeastern Coastal plain (Mitchell and Duncan, 2009), which provide higher metabolic energy saving, as burnt and dead-wood have softer wood, that enables less mechanical work for the woodpecker to produce cavities (Wesolowski, 2011). In turn, this would support the argument of increasing reproductive success in burnt pine forests for the species, increasing the probability of encounters.

However, Vierling and Lentile (2006) fail to mention the nesting success rate over time and whether it fluctuates or fluctuates depending on climate change and shifts of disturbances caused by fires over extended intervals. It suggested that the success rate may not be constant after every arrival of burned pine forests and whether it increases 47% from a smaller percentage, which may leave this increase as marginally significant under a total perspective. Given this, large-diameter trees provide further foraging opportunities and higher nest success prominent in the Boreal Hardwood transition in figure 11 and the Appalachian Mountains in figure 12. which support the higher range of probabilities towards encountering the Red-headed Woodpecker in figure 5 for both landcover types in increasing pland since 2010.

Overall, the hypothesis for woodlands is generally accepted. However, the R^2 value ($r^2 = 0.1199_{\text{increasing}}$, $0.0547_{\text{decreasing}}$) is small for increasing and decreasing landcover; hence only a minimal portion of the variance is explained, even though the significance is present by both $P < 0.001$ and F-statistics.

Savannas

Table 1 shows Savannas to have a consistent trend in their R2 value (R2 = 0.07758decreasing, 0.07867increasing), where both accept their given hypothesis. In table 4, multiple BCR regions had an adverse estimate correlating with the response variable, while Savanna and Woody Savannas were positive. Implying that the minimum model accepts the hypothesis, though multiple predictors overfit the data, the correlation is not a valid estimate of each explanatory variable's significance with the response—the approximate inference derived from the t-test where a higher t-test may imply a stronger correlation towards the slope. In comparison to figure 5, Woody Savannas show higher counts in low probabilities for encounter; when looking at figure 8, this is prevalent across BCR regions with lower probabilities consistent in many regions; South-eastern Coastal plain, Appalachian Mountains, and the Boreal Hardwood Transition.

These results infer similarly from Butler et al. (2004), which found snag-density thresholds in restored Savannas that impacted the probability of detecting a three-toed woodpecker. Consequently, going below this threshold will result in declining probabilities. It compared to figure 4, where Woody Savannas have been decreasing since 2010, suggestive of declines below a habitat threshold for probability detection of the Red-headed Woodpecker. Therefore, during those declines since 2010, a threshold is prevalent and determines the species' decline in several regions. Therefore, not agreeing with the hypothesis towards the model in table 4. However, it accepts the adverse estimates for several BCR regions.

Additionally, the increase in encounters across the North American range for both Savannas attributed to prescribed burns, which optimize the distribution of forest cover types at regional scales (Bergeron et al. 2004) by acting as a natural disturbance to the ecosystem. Woodpeckers can be considered sensitive disturbance species that invade rapidly after burning and effectively closing the canopy and competitively excluding the biologically rich understory (Mitchell and Duncan, 2009). Furthermore, Savanna restoration efforts in the late 1990s employed by Necedah national wildlife refuge (NNWR) showed that the Red-headed Woodpeckers were present in declining savannas, with a low probability of encounter. In the following year, the abundance of the species increased rapidly following burns to the Savannas ecosystem and remained high after other burns (King et al. 2010). Indicative that higher probability shown in figure 5 infers that landscapes with recent burns weaken the wood and amplify dead trees. Given this, testing for regions with burns and whether this phenomenon implicates encounter rates may improve the accuracy of detecting the Red-headed Woodpeckers under Savannas.

Urban

Table 6 shows the decrease in encounter rates for urban areas with landcover increase declined, and when comparing to figure 5, it is noticeable that most probabilities of encountering the Red-headed

Woodpecker fall within the lower probability range. Furthermore, both Southeastern Coastal Plain and Eastern Tallgrass Prairie have a negative slope towards increasing encounters, and figure 5 shows declines in their probability frequency were more significant in the range 0-0.1. While figure 4 shows urban area extent increasing since 2010, concluding that the decline of encounter rates is likely due to the removal of snags and dying trees from urban intensification (Morrison and Chapman, 2005), combined with fragmented landscapes and the absence of decaying wood material (Rodewald et al. 2005). Relevant to Rodewald et al. (2005), which investigated the breeding behaviors of the Red-headed Woodpeckers in golf courses. They found that the presence of the species was closely related to hard-mast trees, dead limbs, and snags, although they were more prevalent in older landscapes, and such features were lacking in the younger landscape. Hence, with figure 4 showing increases in urban landcover, it can be inferred that the decline in encounters is driven by newly formed mosaics in urban areas, such examples as golf courses.

Encounter rates are also dependent on the species suitability towards the habitat and whether the habitat can accommodate the resources required for nesting, foraging, and services for the Red-headed Woodpeckers to maintain residency. Works from Lerman et al. (2014) assessed habitat potential for several New York and Boston species. Their results showed that both cities failed to reach a suitable threshold for support species that required dead-wood, which are good indicators for Red-headed Woodpecker presence. Revealing itself in figure 15, where both urban areas have a low probability of encounter rate ranging from 0.15-0.25.

Additionally, the effort covariates that 8were used as predictions for the species distribution models imply that the probability of encountering a single observer at 1km total distance of surveying in urban landscapes may minimize the likelihood of encountering the Red-headed Woodpeckers. The higher anthropogenic disturbances caused by roads and buildings can fragment habitats for the red-headed woodpecker, making detection more difficult during eBird surveying.

While the hypothesis accepts the R2 ($r2 = 0.1151$), the overall variance explained by the predictors to determine a trend is slim. Although the distribution of probabilities across the range maps and bar plots provides a strong estimate towards the association that Red-headed Woodpeckers are still declining in some Urban areas, management towards New England/Mid-Atlantic Coast will be important for future population growth.

Grassland/Cropland

Grasslands and Croplands have optimistic estimates, as shown in table 7, towards the response variable, which accepts the hypothesis. However, only two BCR regions have an adverse estimate, hence rejecting the hypothesis. Surprisingly, both these BCR regions have optimistic estimates in table 8, therefore accepting the hypothesis of encounter declines with landcover decline. This model only has cropland as a predictor for landcover and an adverse estimate, inferring that the two BCR regions, Appalachian mountains and Central Hardwoods correlate with the response away from both landcover classes. Given that Cropland and Grasslands are pland changes across the entire North American region, this infers that the accumulation of landcover changes shows an increase in encounters across a larger spatial range. Though it seems further research effort towards examining the declines in Appalachian mountains and Central Hardwoods is required to conclude with unmistakable evidence about the drivers of decline in these regions.

However, evidence suggests that the probabilities for encountering the Red-headed Woodpecker have increased, and precise locations for their increase are appended to figures 9 and 10. While Figures 5 and 6 show the range of probabilities, which span from 0 to 0.7 for grassland and 0 to 0.6 for cropland. The frequency of counts towards encounter rates is primarily within the lower quartile of the probability range, inferring that detections are low in both landcover changes. Additionally, figure 9 surprisingly shows the highest probabilities in encountering the Red-headed Woodpecker for Grassland, distributed towards the Prairie conservation regions. A large proportion is within the range of 0.5-0.8, which infers that if a single observer is performing an eBird survey, traveling at least 1km in the distance, at peak detection time, there is a high probability of encountering the Red-headed Woodpecker within this landcover class. Unfortunately, no research to date describes trends towards the Red-headed Woodpecker population in grassland habitats and why such an extensive distribution is available in these areas. Inferring that both topographical covariates, elevation, and slope, increase detection within the Prairie regions, providing a higher probability than likely. Unless deadwood and snags along with Grasslands dominate these areas, figures 7, 8, 11, and 12 are landcover types prominent with deadwood and snags; however, their distribution is towards the east of North America. If a trend exists with the Red-headed Woodpecker at the grassland habitat, research attributed to their nesting behaviors and migrational patterns will better explain this connection.

For increasing landcover changes given in table 7, cropland has a fair estimate hence accepting the hypothesis. Firstly, its t-value is lower than grassland, assuming a lower predictor than grassland towards the response. However, it is significant ($P<0.001$) and correlating with multiple BCR regions with the same slope direction. Secondly, table 8 is the only landcover predictor in the final regression model. Therefore, it corresponds more to the response variable as a negation towards the hypothesis. Inferring that encounter declines are not apparent during cropland decrease, meaning that enough habitat structure is still provided to the woodpecker.

Furthermore, the model output for cropland shares a distributional comparison for encounter rates as grassland. Effectively, this is perhaps a consequence of grassland conversion to cropland in the Prairie regions (Rashford et al. 2016), which in part is inferred from figure 4, as grasslands have been decreasing since 2010 on the other hand, croplands have fluctuated between increase and decrease but finally remaining stable towards 2018 as in 2010. Rashford et al. (2016) results predict grassland to cropland conversion will increase on average by 60% from 2006 to 2011. While both habitats provide a suitable ecosystem for resident and migratory birds, it is not easy to interpret how the Red-headed Woodpecker's role will shift if further conversions continue and amplify in the years to come.

Conclusion:

Overall, Red-Headed Woodpecker encounter rates were diverse across landcover types, under several expectations of accepting the hypothesis. This expectation was met, although there were several inconsistencies towards the methodology. Primarily multiple linear regression, where several dummy variables failed to properly correlate between the explanatory variables and the response. This caused an overlap between the predicted and observed values. Given this, Encounter rates are apparent with increases of landcover over time, and especially in Woodlands, Grassland, Croplands, Savannas and Wetland. Though not enough research has been invested towards multiple BCR regions towards the western breeding range of the Red-headed Woodpecker.

When comparing between landcover types grassland have shown the highest probabilities for encounter in the species distribution maps. Although, wetlands and mixed forests have the highest probabilities ranges, that reach absolute certainty in encountering the Red-headed Woodpecker in some areas. However, the problem persists that most counts were within the lowest probability ranges. Ultimately, increases in landcover have explained that the red-headed Woodpecker have increased since 2010, however there are declines in several BCR regions that have not been investigated into.

Altogether, the species has increased across the entire range, but some BCR regions have been showing inconsistencies towards every model by implicating adverse trends. Carrying this study forward towards individual BCR regions and including covariates like burning cover to understand the impact of burns on landcover and the woodpecker’s regeneration after each burn, would support the understanding of these encounters more clearly.

References:

1. Amatulli, G., Domisch, S., Tuanmu, M-N., Parmentier, B., Ranipeta, A., Malczyk, J. and Jetz, W. 2018. A suite of global, cross-scale topographic variables for environmental and biodiversity modelling. *Scientific data*. **(5)**, 180040.
2. Anderson, E.C., and LaMontagne, J.M. 2016. Nest selection by red-headed woodpeckers across three spatial scales in an urban environment.*Urban Ecosystem*. **19**, 297-314.
3. Aronson, M.F.J., and Galatowitsch, S. 2008. Long-term vegetation developments of restored prairie pothole wetlands. *Wetlands*. **28**(4), 883-895.
4. Baston, D. 2021. Exactextractr: Fast extraction from Raster Datasets using Polygons. *R package version 0.6.0*. https://cran.r-project.org/web/packages/exactextractr/exactextractr.pdf

5. Bella et al. 2009, Calibration of Machine Learning Models, Universidad Politécnica de Valencia
6. Bergeron et al. 2004 Fire regimes at the transition between mixedwood and coniferous boreal forest in northwestern Québec. Ecology 85: 1916–1932
7. Berl, J.L., Edwards, J.W., and Bolsinger, J.S. 2015. Scale-dependent and multi-metric nest habitat thresholds for Red-headed Woodpeckers at the northern periphery of their range. The Condor. 117(2), 203-216.
8. Bird Studies Canada and NABCI. 2014. Bird Conservation Regions. Published by Bird Studies Canada on behalf of the North American Bird Conservation Initiative. https://www.birdscanada.org/bird-science/nabci-bird-conservation-regions Accessed: [26/12/2020]
9. Breiman, L. 2001. Random Forests. Machine Learning. 24, 123-140.
10. Bull, E.L., Parks, C.G., Torgersen, T.R. 1997. Trees and logs important to wildlife in the interior Columbian River basin. USDA Forest Service General Technical Report PNW-GTR-391. U.S. Department of Agriculture Forest Service, Pacific Northwest Research Station, Portland, Oregon.
11. Butler et al. 2014, Dead woodthreshold values for the three-toed woodpecker presence in boreal andsub-Alpine forest. Biological Conservation 119:305–318.
12. Callaghan, C.T., Lyons, M.B., Martin, J.M., Major, R.E., and Kingsford, R.T. 2017. Assessing the reliability of urban greenspace using eBird citizen science data. *Avian conservation and Ecology*. **12**(2)
13. Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., and Lawler, J.J. 2007. Random forests for classification in ecology. *Ecology*. **88**(11), 2783-2792.
14. eBird, 2005. Red-headed Woodpecker, Range map: Breeding. https://ebird.org/science/status-and-trends/rehwoo/range-map-breeding Accessed: [26/12/2020]
15. Evans J.S., Murphy M.A., Holden Z.A., Cushman S.A. (2011) Modeling Species Distribution and Change Using Random Forest. In: Drew C., Wiersma Y., Huettmann F. (eds) Predictive Species and Habitat Modeling in Landscape Ecology. Springer, New York, NY.
16. Frei, B., Fyles, J.W. and Nocera, J.J. 2013. Maladaptive Habitat Use of a North American Woodpecker in Population Decline. *Ethology*, 119, 377-388.
17. Frei B., et al. 2020, Birds of the World, Red-headed Woodpecker, version 1.0
18. Friedl, M., and Sulla-Menashe, D. 2015. MCD12Q1 MODIS/Terra+Aqua land cover type yearly L3 Global 500m SIN Grid V006. *NASA EOSDIS Land Processes DAAC.* ***10***
19. Gromping, U. 2009. Variable importance assessment in regression: Linear regression versus random forest. *Am. Statisticas.* **63**, 308-319.
20. Guillera-Arroita, G., Lahoz-Monfort, J.J., Elith, J., Gordon, A., Kujala, H., Lentini, P.E., McCarthy, M.A., Tingley, R., and Wintle, B.A. 2015. Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*. **24**(3), 276-292.
21. Hansen, M.C., Defries, R.S., Townshend, J.R., and Sohlberg, R. 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. International Journal of Remote Sensing. 21(6-7), 1331-1364.
22. Hijmans, R. 2012. Croass-validation of species distribution models: removing spatial sorting bias and calibration with a null-model. *Ecology*. **93**(3), 679-688.

23. Hijmans, R.J., Etten, J.V., Cheng, J., Mattiuzzi, M., Summer, M., Greenberg, J.A., Lamigueiro, O.P., Bevan, A., Racine, E.B. and Shortridge, A. 2015. Geographic data analysis and modelling package 'raster' version 2.5-2. *CRAN R Repositry*. 1-242.
24. Hollenbeck et al. 2011, Habitat Suitability and Nest Survival of White-Headed Woodpeckers in Unburned Forests of Oregon. Journal of Wildelife Management. 75(5), 1061-1071.
25. Jones-Farrand, D.T., Fearer, T.M., Thogmartin, W.E., Thomson, F.R., Nelson, M.D., and Tirpak, J.M. 2011. Comparison of statistical and theoretical habitat models for conservation planning: the benefit of ensemble prediction. *Ecological Applications*. **21**(6), 2269-2282.
26. Kantrud et al. 1989, Prairie basin wetlands of the dakotas: a community profile. Biological report 85, U.S. Department of the Interior, U.S. Fish and Wildlife Service, Jamestown, ND, USA
27. Kelling, S., Johnston, A., Fink, D., Ruiz-Gutierrez, V., Bonney, R., Bonn, A., Fernandez, M., Hochachka, W.M., Julliard, R., Kraemer, R., and Guralnick, R. (2018). Finding the signal in the noise of Citizen Science Observations. *bioRxiv.*
28. Kilgo, J.C., and Vukovich, M. 2012. Factors affecting breeding season survival of Red-headed woodpeckers in South Carolina. *The Journal of Wildlife Management*. **76**, 328-335.
29. Kilgo, J.C., and Vukowich, M.A. 2014. Can snag creation benefit a primary cavity nester: Response to an experimental pulse in snag abundance. *Biological Conservation*. **171**, 21-28.
30. King et al. 2010, Red-headed Woodpecker Nest-Habitat Thresholds in Restored Savannas. Journal of Wildlife Management. 71(1), 30-35.
31. Koenig, W.D., Walters, E.L., and Rodewald, P.G. 2017. Testing alternative hypotheses for the cause of population declines: The case of the Red-headed Woodpecker. *The Condor*. **119**(1), 143-154.
32. Liaw, A., Wiener, M. 2002. Classification and regression by randomForest. *R News*. **2**, 18-22.
33. Leonard and Stout, 2006, Woodpecker use of Forested Wetlands in Central Peninsular Florida. Southeastern Naturalist. 5(4), 621-636
34. Lerman et al. 2014, Using Urban Forest Assessment Tools to Model Bird Habitat Potential. Landscape and Urban Planning. 122, 29-40.
35. McGarigal, K., Cushman, S.A., Neel, M.C., and Ene, E. 2002. FRAGSTATS: Spatial pattern analysis program for categorical maps. *University of Massachusetts, Amherst, Ma*: http://www.umass.edu/landeco/research/fragstats/fragstats.html. Accessed: [20/02/2021]
36. Mitchell and Duncan, 2009, Old Forests and Endangered Woodpeckers: Old-Growth in the Southern Coastal Plain. Natural Areas Journal. 29, 301-310.
37. Morrison and Chapman, 2005, Can Urban Parks Provide Habitat for Woodpeckers? Northeastern Naturalist. 12(3), 253-262.
38. Nebel, S., Mills, A., McCracken, J.D. and Taylor, P.D. 2010. Declines of aerial insectivores in North America follow a geographic gradient. *Avian and Conservation Ecology*. **5**(1)
39. Nickley, B., and Bulluck, L.P. 2019. Spanning the habitat gradient: Red-headed woodpecker nest-site selection in three distinct cover types. *Forest Ecology and Management*. **444**, 115-126.
40. O'Hara, R., and Kotze, J. 2010. Do not log-transform count data. *Methods in Ecology and Evolution*. **1**(2), 118-122.

41. Pagel, R., West, E., Jones, A., and Streby, H. 2020. Variation in individual autumn migration and winter paths of Great Lakes red-headed woodpeckers (*Melanerpes erythrocephalus*). *Animal Migration*. **7**(1), 9-18.
42. Rashford et al. 2016, Impacts of climate change on land-use and wetland productivity in the Prairie Pothole Region of North America. Regional Environmental Change. 16, 515-526.
43. Riffell et al. 2006, Birds in depressional forested wetlands: area and habitat requirements and model uncertainty. Wetlands. 26(1), 107-118.
44. Rodewald, P.G., Santiago, M.J., and Rodewald, A.D. 2005. Habitat use of breeding Red-headed Woodpeckers on golf courses in Ohio. *Wildlife Society Bulletin*. **33**(2), 448-453.
45. Rosenberg,K. V., Kennedy, J. A. , Dettmers, R.., Ford, R. P., Reynolds, D., Beardmore, C. J., Blancher, P. J., Bogart, R., Butcher, G. S., Camfield, A. F. , Couturieret A., Demarest, D.W., Easton, W.E., Giocomo, J.J., Keller, R.H., Mini, A.E., Panjabi, A.O., Pashley, D.N., Rich, T.D., Ruth, J.M., Stabins, H., Stanton, J. and Will, T. (2016). Partners in Flight Landbird Conservation Plan: 2016 Revision for Canada and Continental United States. Partners in Flight Science Committee. https://www.partnersinflight.org/resources/the-plan/
46. Sanchez-Bayo, F., and Wyckhuys, K.A.G. 2019. Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation*. **232**, 8-27.
47. Smith, K. G., Withgott, J. H. and Rodewald, P. G. 2000. Red-headed woodpecker (Melanerpes erythrocephalus). The Birds of North America Online (A. Poole, Ed.). Cornell Lab of Ornithology, Ithaca, NY.
48. Smith, P.F., Ganesh, S., and Liu, P. 2013. A comparison of random forest regression and multiple linear regression for prediction in neuroscience. *Journal of Neuroscience Methods*. **220**(1), 85-91.
49. Steven and Lowrance 2011, Agricultural conservation practices and wetland ecosystem services in the wetland-rich Piedmont-Coastal Plain region. Ecological Applications. 21(3), 3-17.
50. Strimas-Mackey, M., W.M. Hochachka, V. Ruiz-Gutierrez, O.J. Robinson, E.T. Miller, T. Auer, S. Kelling, D. Fink, A. Johnston. 2020. Best Practices for Using eBird Data. Version 1.0. https://cornelllabofornithology.github.io/ebird-best-practices/ . Cornell Lab of Ornithology, Ithaca, New York. https://doi.org/10.5281/zenodo.3620739
51. Tirpak, J.M., Jones-Farrand, D.T., Thomson, F.R., Twedt, D.J., Baxter, C.R., Fitzgerald, J.A., and Uihlein, W.B. 2009. Assessing Ecoregional-scale habitat suitability index models for priority landbirds. *The journal of wildlife management*. **73**, 1307-1315.
52. Vaughan, I.P., and Ormerod, S.J. 2005. The continuing challenges of testing species distribution models. *Journal of Applied Ecology*. **42**(4), 720-730.
53. Valiela, I., and Martinetto, P. 2007. Changes in bird abundance in eastern North America: Urban sprawl and global footprint? *BioScience.* **57**(4)
54. Vierling and Lentile 2006, Red-headed woodpecker nest-site selection and production in mixed ponderosa pine and aspen woodland following fire. Condor. 108, 957-962.
55. Vukovich, M., and Kilgo, J.C. 2019. Diet of Nestling Red-Headed Woodpeckers in South Carolina. *Southeastern Naturalist* **18**(2), 173-182
56. Wesolowski, 2011, “lifespan” of woodpecker-made holes in a primeval temperate forest: A thirty year study. Forest Ecology and Management. 262(9), 1846-1852.

Appendix:

Name	Value	Description
Water bodies	0	At least 60% of area is covered by permanent water bodies
Deciduous Broadleaf Forests	4	Dominated by deciduous broadleaf trees (canopy >2m). Tree cover >60%
Mixed Forests	5	Dominated by neither deciduous nor evergreen (40-60% of each) tree type (canopy >2m). Tree cover >60%
Woody Savannas	8	Tree cover 30-60% (canopy >2m)
Savannas	9	Tree cover 10-30% (canopy >2m)
Grasslands	10	Dominated by herbaceous annuals (<2m)
Permanent Wetlands	11	Permanently inundated lands with 30-60% water cover and >10% vegetated cover
Croplands	12	At least 60% of area is cultivated cropland
Urban and Built-up Lands	13	At least 30% impervious surface area including building materials, asphalt, and vehicles
Non-Vegetated Lands	15	At least 60% of area is non-vegetated barren (sand, rock, soil) or permanent snow and ice with less than 10% vegetation

Table 2. Random forest algorithm output of linear models

	Water Bodies (Decrease)	**Water Bodies (Increase)**	**Savannas (Increase)**	**Savannas (Decrease)**	**Urban (Increase)**	**Agriculture (Increase)**	**Agriculture (Decrease)**	**Woodland (Increase)**	**Woodland (Decrease**
Multiple R-squared	0.0525	0.0332	0.03318	0.01971	0.07528	0.106	0.1338	0.04321	0.02917
Adjusted R-squared	0.08142	0.03052	0.02774	0.01634	0.07142	0.09881	0.1286	0.0388	0.02733
F-statistics	22.25	12.39	6.097	5.852	19.5	14.79	25.57	12.05	15.81
DF	9, 2149	11, 3968	8, 1421	4, 1164	4, 958	9, 1123	8, 1324	9, 2455	4, 2105

```
#Creating the BCR map used in my methodology:
library(tidyverse)
library(broom)
library(rgdal)
#load the geopackage
p <- readOGR(dsn = "bcr.gpkg")
#turn the geopackage into a tibble
tidy_bcr <- tidy(p)
#create a plot to see what it looks like
ggplot(p, aes(x = long, y = lat, group = group)) +
  geom_polygon(color = "black", size = 0.1, fill = "lightgrey") +
  coord_equal() +
  theme_minimal()

p$id <- row.names(p)
#add bcr_name to the dataframe
tidy_bcr <- left_join(tidy_bcr, p@data)
#select bcr regions
bcrcop <- data.frame(bcr_name = sort(p@data$bcr_name),
                     bcr_code =
c(28,17,12,24,19,22,13,26,30,21,31,29,23,11,18,27,25))
```

```
tidy_bcr <- left_join(tidy_bcr, bcrcop)

#label the regions so they centre on the map
bcrLabel <- tidy_bcr %>%
  group_by(bcr_name) %>%
  summarise(label_long = mean(range(long)), label_lat = mean(range(lat)),
bcr_code = mean(bcr_code)) %>%
mutate(label_long = replace(label_long, bcr_name %in% c("Badlands And
Prairies", "Shortgrass Prairie ", " Central Mixed Grass Prairie ", "Oaks
And Prairies ", "West Gulf Coastal Plain/Ouachitas ", "Peninsular Florida
", "Central Hardwoods","Eastern Tallgrass Prairie","Prairie Hardwood
Transition","Boreal Hardwood Transition","Lower Great Lakes/St. Lawrence
Plain","New England/Mid-Atlantic Coast"),
                            values=c(-105.1,-103.1,-100.4,-97.1,-94.1,-82,-
90,-89.7,-89.4,-83.8,-77,-73.7)),
       label_lat = replace(label_lat, bcr_name %in% c("Prairie Potholes",
"Southeastern Coastal Plain ", "Piedmont  ", "Appalachian Mountains
","Mississippi Alluvial Valley "),
                           values=c(48.5,32.5 ,36.3 ,38.3,33.9 )))
#create plot
map <- ggplot(tidy_bcr, aes(x = long, y = lat, group = group, fill =
bcr_name)) +
  geom_polygon(color = "black", size = 0.1) +
  coord_equal() +
  theme_void()  +
  theme(plot.title = element_text(margin = margin(t = 40, b = -40)))
#plot wit textx
c <- map + geom_text(data = bcrLabel, mapping = aes(x=label_long, y =
label_lat, label = bcr_code, group = NA), cex = 4, col = "white") +
theme(legend.box.background = element_rect(colour = "black"),
legend.position = "bottom")

####

#Begin the filtering of the eBird data
library(auk)
library(lubridate)
library(sf)
library(gridExtra)
library(tidyverse)
# resolve namespace conflicts
select <- dplyr::select

# setup data directory
dir.create("data", showWarnings = FALSE)
#select the red-headed woodpecker txt file and the sampling text file
ebd <- auk_ebd("rehwoo.txt",
               file_sampling = "sampling.txt")
#Filter the data
ebd_filters <- ebd %>%
  auk_species("Red-headed Woodpecker") %>%
  # june, use * to get data from any year
  auk_date(date = c("*-05-01", "*-08-31")) %>%
  # restrict to the standard traveling and stationary count protocols
  auk_protocol(protocol = c("Stationary", "Traveling")) %>%
  auk_complete()

#create a file to store the values in
data_dir <- "data"
if (!dir.exists(data_dir)) {
  dir.create(data_dir)
```

```
}
#name the files
f_ebd <- file.path(data_dir, "rehwoo.txt")
f_sampling <- file.path(data_dir, "esampling.txt")

# only run if the files don't already exist
if (!file.exists(f_ebd)) {
  auk_filter(ebd_filters, file = f_ebd, file_sampling = f_sampling)
}

#fill the data with zeros
ebd_zf <- auk_zerofill(f_ebd, f_sampling, collapse = TRUE)

# function to convert time observation to hours since midnight
time_to_decimal <- function(x) {
  x <- hms(x, quiet = TRUE)
  hour(x) + minute(x) / 60 + second(x) / 3600
}

# clean up variables
ebd_zf <- ebd_zf %>%
  mutate(
    # convert X to NA
    observation_count = if_else(observation_count == "X",
                                NA_character_, observation_count),
    observation_count = as.integer(observation_count),
    # effort_distance_km to 0 for non-travelling counts
    effort_distance_km = if_else(protocol_type != "Traveling",
                                 0, effort_distance_km),
    # convert time to decimal hours since midnight
    time_observations_started = time_to_decimal(time_observations_started),
    # split date into year and day of year
    year = year(observation_date),
    day_of_year = yday(observation_date)
  )

# additional filtering
ebd_zf_filtered <- ebd_zf %>%
  filter(
    # effort filters
    duration_minutes <= 5 * 60,
    effort_distance_km <= 5,
    # last 10 years of data
    year >= 2010,
    # 10 or fewer observers
    number_observers <= 10)

ebird <- ebd_zf_filtered %>%
  select(checklist_id, observer_id, sampling_event_identifier,
         scientific_name,
         observation_count, species_observed,
         state_code, locality_id, latitude, longitude,
         protocol_type, all_species_reported,
         observation_date, year, day_of_year,
         time_observations_started,
         duration_minutes, effort_distance_km,
         number_observers)
write_csv(ebird, "data/rehwoo.csv", na = "")

####
#begin covariate analysis
```

```
library(sf)
library(raster)
library(MODIS)
library(exactextractr)
library(viridis)
library(tidyverse)
# resolve namespace conflicts
select <- dplyr::select
map <- purrr::map
projection <- raster::projection

bcr <- read_sf("data/gis-data.gpkg", "bcr") %>% filter(bcr_code %in% c(11,
12, 13, 17, 18, 19,21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31)) %>%
st_transform(crs = paste("+proj=sinu +lon_0=0 +x_0=0 +y_0=0",

"+a=6371007.181 +b=6371007.181 +units=m +no_defs"))
ebird <- read_csv("data/rehwoo.csv")
begin_year <- format(min(ebird$observation_date), "%Y.01.01")
# end date for ebird data
end_year <- format(max(ebird$observation_date), "%Y.12.31")
tifs <- runGdal(product = "MCD12Q1", collection = "006", SDSstring = "01",
                extent = bcr %>% st_buffer(dist = 10000),
                begin = begin_year, end = end_year,
                outDirPath = "data", job = "modis",
                MODISserverOrder = "LPDAAC") %>%
  pluck("MCD12Q1.006") %>%
  unlist()

new_names <- format(as.Date(names(tifs)), "%Y") %>%
  sprintf("modis_mcd12q1_umd_%s.tif", .) %>%
  file.path(dirname(tifs), .)
file.rename(tifs, new_names)

landcover <- list.files("data/modis", "^modis_mcd12q1_umd",
                        full.names = TRUE) %>%
  stack()
# label layers with year
landcover <- names(landcover) %>%
  str_extract("(?<=modis_mcd12q1_umd_)[0-9]{4}") %>%
  paste0("y", .) %>%
  setNames(landcover, .)
landcover

max_lc_year <- names(landcover) %>%
  str_extract("[0-9]{4}") %>%
  as.integer() %>%
  max()

neighborhood_radius <- 5 * ceiling(max(res(landcover))) / 2
ebird_buff <- ebird %>%
  distinct(year = format(observation_date, "%Y"),
           locality_id, latitude, longitude) %>%
  # for 2019 use 2018 landcover data
  mutate(year_lc = if_else(as.integer(year) > max_lc_year,
                           as.character(max_lc_year), year),
         year_lc = paste0("y", year_lc)) %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
  # transform to modis projection
  st_transform(crs = projection(landcover)) %>%
  # buffer to create neighborhood around each point
```

```
  st_buffer(dist = neighborhood_radius) %>%
  # nest by year
  nest(data = c(year, locality_id, geometry))

calculate_pland <- function(yr, regions, lc) {
  locs <- st_set_geometry(regions, NULL)
  exact_extract(lc[[yr]], regions, progress = FALSE) %>%
    map(~ count(., landcover = value)) %>%
    tibble(locs, data = .) %>%
    unnest(data)
}
# iterate over all years extracting landcover for all checklists in each
lc_extract <- ebird_buff %>%
  mutate(pland = map2(year_lc, data, calculate_pland, lc = landcover)) %>%
  select(pland) %>%
  unnest(cols = pland)

pland <- lc_extract %>%
  # calculate proporiton
  group_by(locality_id, year) %>%
  mutate(pland = n / sum(n)) %>%
  ungroup() %>%
  select(-n) %>%
  # remove NAs after tallying so pland is relative to total number of cells
  filter(!is.na(landcover))

lc_names <- tibble(landcover = 0:15,
                   lc_name = c("pland_00_water",
                               "pland_01_evergreen_needleleaf",
                               "pland_02_evergreen_broadleaf",
                               "pland_03_deciduous_needleleaf",
                               "pland_04_deciduous_broadleaf",
                               "pland_05_mixed_forest",
                               "pland_06_closed_shrubland",
                               "pland_07_open_shrubland",
                               "pland_08_woody_savanna",
                               "pland_09_savanna",
                               "pland_10_grassland",
                               "pland_11_wetland",
                               "pland_12_cropland",
                               "pland_13_urban",
                               "pland_14_mosiac",
                               "pland_15_barren"))
pland <- pland %>%
  inner_join(lc_names, by = "landcover") %>%
  arrange(landcover) %>%
  select(-landcover)

# tranform to wide format, filling in implicit missing values with 0s%>%
pland <- pland %>% group_by(year) %>% mutate(nrow = row_number()) %>%
  pivot_wider(names_from = year,
              values_from = pland,
              values_fill = list(pland = 0)) %>% select(-nrow)

# save
write_csv(pland, "data/modis_pland_location-year.csv")
#aggregate the resolution to get 5x5 modis cells
agg_factor <- round(2 * neighborhood_radius / res(landcover))
r <- raster(landcover) %>%
  aggregate(agg_factor)
r <- bcr %>%
```

```
  st_transform(crs = projection(r)) %>%
  rasterize(r) %>%
  # remove any empty cells at edges
  trim()
r <- writeRaster(r, filename = "data/prediction-surface.tif", overwrite =
TRUE)
#extract points
r_centers <- rasterToPoints(r, spatial = TRUE) %>%
  st_as_sf() %>%
  mutate(id = row_number())
#convert to polygon with neighbourhood radius
r_cells <- st_buffer(r_centers, dist = neighborhood_radius)
#extract values at every radius
lc_extract_pred <- landcover %>%
  exact_extract(r_cells, progress = FALSE) %>%
  map(~ count(., y2010, y2011, y2012, y2013, y2014, y2015, y2016, y2017,
y2018, y2019)) %>%
  tibble(id = r_cells$id, layer = r_cells$layer, data = .) %>%
  unnest(data)
#remove nas
lc_extract_pred <- lc_extract_pred[complete.cases(lc_extract_pred),]
#calculat epland
pland_pred <- lc_extract_pred %>%
  count(id,layer, y2010, y2011, y2012, y2013, y2014, y2015, y2016, y2017,
y2018, y2019, landcover) %>%
  group_by(id) %>%
  mutate(pland = n / sum(n)) %>%
  ungroup() %>%
  select(-n) %>%
  # remove NAs after tallying so pland is relative to total number of cells
  filter(!is.na(.))

lc_names <- tibble(y2010 = 0:15,y2011 = 0:15,y2012 = 0:15,y2013 =
0:15,y2014 = 0:15,y2015 = 0:15,y2016 = 0:15,y2017 = 0:15,y2018 = 0:15,y2019
= 0:15,
                   lc_name = c("pland_00_water",
                               "pland_01_evergreen_needleleaf",
                               "pland_02_evergreen_broadleaf",
                               "pland_03_deciduous_needleleaf",
                               "pland_04_deciduous_broadleaf",
                               "pland_05_mixed_forest",
                               "pland_06_closed_shrubland",
                               "pland_07_open_shrubland",
                               "pland_08_woody_savanna",
                               "pland_09_savanna",
                               "pland_10_grassland",
                               "pland_11_wetland",
                               "pland_12_cropland",
                               "pland_13_urban",
                               "pland_14_mosiac",
                               "pland_15_barren"))
#take the names of years
cols <- grep('y\\d+', names(pland_pred), value = TRUE)
#replace landcover values by names in lc_names
pland_pred[cols] <- Map(function(x, y) lc_names$lc_name[match(y, x)],
                        lc_names[cols], pland_pred[cols])


#time series
pland_pred.water <- pland_pred
pland_pred.deciduous_broadleaf <- pland_pred
```

```
pland_pred.mixed_forest <- pland_pred
pland_pred.woody_savanna <- pland_pred
pland_pred.savanna <- pland_pred
pland_pred.grassland <- pland_pred
pland_pred.wetland  <- pland_pred
pland_pred.cropland <- pland_pred
pland_pred.urban  <- pland_pred

#Replace those beloning to the landcover name by pland, and those not
belonging to it the value 0
pland_pred.water[, 2:11] <- ifelse(pland_pred[, 2:11] == "pland_00_water",
pland_pred$pland, 0)
pland_pred.deciduous_broadleaf[, 2:11] <-  ifelse(pland_pred[, 2:11] ==
"pland_04_deciduous_broadleaf", pland_pred$pland, 0)
pland_pred.mixed_forest[, 2:11] <-  ifelse(pland_pred[, 2:11] ==
"pland_05_mixed_forest", pland_pred$pland, 0)
pland_pred.woody_savanna[, 2:11] <-  ifelse(pland_pred[, 2:11] ==
"pland_08_woody_savanna", pland_pred$pland, 0)
pland_pred.savanna[, 2:11] <-  ifelse(pland_pred[, 2:11] ==
"pland_09_savanna", pland_pred$pland, 0)
pland_pred.grassland[, 2:11] <-  ifelse(pland_pred[, 2:11] ==
"pland_10_grassland", pland_pred$pland, 0)
pland_pred.wetland[, 2:11]  <-  ifelse(pland_pred[, 2:11] ==
"pland_11_wetland", pland_pred$pland, 0)
pland_pred.cropland[, 2:11] <- ifelse(pland_pred[, 2:11] ==
"pland_12_cropland", pland_pred$pland, 0)
pland_pred.urban[, 2:11]  <- ifelse(pland_pred[, 2:11] == "pland_13_urban",
pland_pred$pland, 0)

#filter
pland_pred.water <- pland_pred.water[, -12]
pland_pred.deciduous_broadleaf <- pland_pred.deciduous_broadleaf[, -12]
pland_pred.mixed_forest <- pland_pred.mixed_forest[, -12]
pland_pred.woody_savanna <- pland_pred.woody_savanna[, -12]
pland_pred.savanna <- pland_pred.savanna[, -12]
pland_pred.grassland <- pland_pred.grassland[,-12]
pland_pred.wetland <- pland_pred.wetland[, -12]
pland_pred.cropland <- pland_pred.cropland[, -12]
pland_pred.urban <- pland_pred.urban[, -12]

#join the landcover dataframe with the coordinates from the bcr range
pland_coords.water <- st_transform(r_centers, crs = 4326) %>%
  st_coordinates() %>%
  as.data.frame() %>%
  cbind(id = r_centers$id, .) %>%
  rename(longitude = X, latitude = Y) %>%
  inner_join(pland_pred.water, by = "id")

pland_coords.deciduous_broadleaf <- st_transform(r_centers, crs = 4326) %>%
  st_coordinates() %>%
  as.data.frame() %>%
  cbind(id = r_centers$id, .) %>%
  rename(longitude = X, latitude = Y) %>%
  inner_join(pland_pred.deciduous_broadleaf, by = "id")

pland_coords.mixed_forest <- st_transform(r_centers, crs = 4326) %>%
  st_coordinates() %>%
  as.data.frame() %>%
  cbind(id = r_centers$id, .) %>%
  rename(longitude = X, latitude = Y) %>%
  inner_join(pland_pred.mixed_forest, by = "id")
```

```
pland_coords.woody_savanna <- st_transform(r_centers, crs = 4326) %>%
  st_coordinates() %>%
  as.data.frame() %>%
  cbind(id = r_centers$id, .) %>%
  rename(longitude = X, latitude = Y) %>%
  inner_join(pland_pred.woody_savanna, by = "id")

pland_coords.savanna <- st_transform(r_centers, crs = 4326) %>%
  st_coordinates() %>%
  as.data.frame() %>%
  cbind(id = r_centers$id, .) %>%
  rename(longitude = X, latitude = Y) %>%
  inner_join(pland_pred.savanna, by = "id")

pland_coords.grassland <- st_transform(r_centers, crs = 4326) %>%
  st_coordinates() %>%
  as.data.frame() %>%
  cbind(id = r_centers$id, .) %>%
  rename(longitude = X, latitude = Y) %>%
  inner_join(pland_pred.grassland, by = "id")

pland_coords.wetland <- st_transform(r_centers, crs = 4326) %>%
  st_coordinates() %>%
  as.data.frame() %>%
  cbind(id = r_centers$id, .) %>%
  rename(longitude = X, latitude = Y) %>%
  inner_join(pland_pred.wetland, by = "id")

pland_coords.cropland <- st_transform(r_centers, crs = 4326) %>%
  st_coordinates() %>%
  as.data.frame() %>%
  cbind(id = r_centers$id, .) %>%
  rename(longitude = X, latitude = Y) %>%
  inner_join(pland_pred.cropland, by = "id")

pland_coords.urban <- st_transform(r_centers, crs = 4326) %>%
  st_coordinates() %>%
  as.data.frame() %>%
  cbind(id = r_centers$id, .) %>%
  rename(longitude = X, latitude = Y) %>%
  inner_join(pland_pred.urban, by = "id")



#create landcover maps
forest_cover.water <- pland_coords.water %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
  st_transform(crs = projection(r)) %>%
  # rasterize points
  rasterize(r) %>%
  # project to albers equal-area for mapping
  projectRaster(crs = st_crs("ESRI:102003")$proj4string, method = "ngb")
%>%
  # trim off empty edges of raster
  trim()

forest_cover.deciduous_broadleaf <- pland_coords.deciduous_broadleaf %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
```

```
  st_transform(crs = projection(r)) %>%
  # rasterize points
  rasterize(r) %>%
  # project to albers equal-area for mapping
  projectRaster(crs = st_crs("ESRI:102003")$proj4string, method = "ngb")
%>%
  # trim off empty edges of raster
  trim()

forest_cover.mixed_forest <- pland_coords.mixed_forest %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
  st_transform(crs = projection(r)) %>%
  # rasterize points
  rasterize(r) %>%
  # project to albers equal-area for mapping
  projectRaster(crs = st_crs("ESRI:102003")$proj4string, method = "ngb")
%>%
  # trim off empty edges of raster
  trim()


forest_cover.woody_savanna <- pland_coords.woody_savanna %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
  st_transform(crs = projection(r)) %>%
  # rasterize points
  rasterize(r) %>%
  # project to albers equal-area for mapping
  projectRaster(crs = st_crs("ESRI:102003")$proj4string, method = "ngb")
%>%
  # trim off empty edges of raster
  trim()

forest_cover.savanna <- pland_coords.savanna %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
  st_transform(crs = projection(r)) %>%
  # rasterize points
  rasterize(r) %>%
  # project to albers equal-area for mapping
  projectRaster(crs = st_crs("ESRI:102003")$proj4string, method = "ngb")
%>%
  # trim off empty edges of raster
  trim()

forest_cover.grassland <- pland_coords.grassland %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
  st_transform(crs = projection(r)) %>%
  # rasterize points
  rasterize(r) %>%
  # project to albers equal-area for mapping
  projectRaster(crs = st_crs("ESRI:102003")$proj4string, method = "ngb")
%>%
  # trim off empty edges of raster
  trim()

forest_cover.wetland <- pland_coords.wetland %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
```

```
  st_transform(crs = projection(r)) %>%
  # rasterize points
  rasterize(r) %>%
  # project to albers equal-area for mapping
  projectRaster(crs = st_crs("ESRI:102003")$proj4string, method = "ngb")
%>%
  # trim off empty edges of raster
  trim()

forest_cover.cropland <- pland_coords.cropland %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
  st_transform(crs = projection(r)) %>%
  # rasterize points
  rasterize(r) %>%
  # project to albers equal-area for mapping
  projectRaster(crs = st_crs("ESRI:102003")$proj4string, method = "ngb")
%>%
  # trim off empty edges of raster
  trim()

forest_cover.urban <- pland_coords.urban %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
  st_transform(crs = projection(r)) %>%
  # rasterize points
  rasterize(r) %>%
  # project to albers equal-area for mapping
  projectRaster(crs = st_crs("ESRI:102003")$proj4string, method = "ngb")
%>%
  # trim off empty edges of raster
  trim()



#write landcover maps
writeRaster(forest_cover.water, "forest_cover.water.tif")
writeRaster(forest_cover.deciduous_broadleaf,
"forest_cover.deciduous_broadleaf.tif")
writeRaster(forest_cover.mixed_forest, "forest_cover.mixed_forest.tif")
writeRaster(forest_cover.woody_savanna, "forest_cover.woody_savanna.tif")
writeRaster(forest_cover.savanna, "forest_cover.savanna.tif")
writeRaster(forest_cover.grassland, "forest_cover.grassland.tif")
writeRaster(forest_cover.wetland, "forest_cover.wetland.tif")
writeRaster(forest_cover.cropland, "forest_cover.cropland.tif")
writeRaster(forest_cover.urban, "forest_cover.urban.tif")


#Create elevation and soil maps
#download the file; replace elevation with soil to get soil and repeat the
code for soil
f_dem <- "elevation_01_05_1km_uint16.tif"
if (!file.exists(file.path("data", f_dem))) {
  download.file(paste0("https://data.earthenv.org/texture/", f_dem),
                file.path("data", f_dem), mode = "wb")
}

elev <- raster("data/elevation_01_05_1km_uint16.tif")
# crop, buffer bcr by 10 km to provide a little wiggly room
elev <- bcr %>%
  st_buffer(dist = 10000) %>%
```

```
  st_transform(crs = projection(elev)) %>%
  crop(elev, .) %>%
  projectRaster(crs = projection(landcover))

ebird_buff_noyear <- ebird %>%
  distinct(locality_id, latitude, longitude) %>%
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
  st_transform(crs = projection(elev)) %>%
  st_buffer(dist = neighborhood_radius)
# extract elevation values and calculate median and sd
locs <- st_set_geometry(ebird_buff_noyear, NULL) %>%
  mutate(id = row_number())
elev_checklists <- exact_extract(elev, ebird_buff_noyear, progress = FALSE)
%>%
  map_dfr(~ tibble(elevation_median = exp(mean(log(.$value),na.rm = TRUE)),
                   elevation_sd = exp(sd(log(.$value),na.rm = TRUE )))) %>%
  # join to lookup table to get locality_id
  bind_cols(locs, .)

elev_pred <- exact_extract(elev, r_cells, progress = FALSE) %>%
  map_dfr(~ tibble(elevation_median = exp(mean(log(.$value), na.rm =
TRUE)),
                   elevation_sd = exp(sd(log(.$value), na.rm = TRUE)))) %>%
  # join to lookup table to get locality_id
  bind_cols(st_drop_geometry(r_cells), .)


pland_elev_checklist <- inner_join(pland, elev_checklists, by =
"locality_id")
write_csv(pland_elev_checklist, "data/year.pland-elevation_location-
year.csv")
# prediction surface covariates
pland_elev_pred <- inner_join(pland_coords.water, elev_pred, by = "id")
write_csv(pland_elev_pred, "data/l.water-elev_prediction-surface.csv")
pland_elev_pred <- inner_join(pland_coords.deciduous_broadleaf, elev_pred,
by = "id")
write_csv(pland_elev_pred, "data/l.decidous_broadleaf-elev_prediction-
surface.csv")
pland_elev_pred <- inner_join(pland_coords.mixed_fored, elev_pred, by =
"id")
write_csv(pland_elev_pred, "data/l.mixed_forest-elev_prediction-
surface.csv")
pland_elev_pred <- inner_join(pland_coords.woody_savanna, elev_pred, by =
"id")
write_csv(pland_elev_pred, "data/l.woody_savanna-elev_prediction-
surface.csv")
pland_elev_pred <- inner_join(pland_coords.savanna, elev_pred, by = "id")
write_csv(pland_elev_pred, "data/l.savanna-elev_prediction-surface.csv")
pland_elev_pred <- inner_join(pland_coords.cropland, elev_pred, by = "id")
write_csv(pland_elev_pred, "data/l.cropland-elev_prediction-surface.csv")
pland_elev_pred <- inner_join(pland_coords.grassland, elev_pred, by = "id")
write_csv(pland_elev_pred, "data/l.grassland-elev_prediction-surface.csv")
pland_elev_pred <- inner_join(pland_coords.wetland, elev_pred, by = "id")
write_csv(pland_elev_pred, "data/l.wetland-elev_prediction-surface.csv")
pland_elev_pred <- inner_join(pland_coords.urban, elev_pred, by = "id")
write_csv(pland_elev_pred, "data/l.urban-elev_prediction-surface.csv")
pland_elev_pred <- inner_join(pland_coords10, elev_pred, by = "id")
write_csv(pland_elev_pred, "data/pland2019-elev_prediction-surface.csv")

#create encounter rate models; repeat this for each different landcover
class
```

```
library(sf)
library(raster)
library(dggridR)
library(lubridate)
library(ranger)
library(scam)
library(PresenceAbsence)
library(verification)
library(ebirdst)
library(fields)
library(gridExtra)
library(tidyverse)
# resolve namespace conflicts
select <- dplyr::select
map <- purrr::map
projection <- raster::projection
# set random number seed to insure fully repeatable results
set.seed(1)
# setup output directory for saved results
if (!dir.exists("output")) {
  dir.create("output")
}
###########################################################################
#########################################################
###########################################################################
#########################################################
ebird <- read_csv("data/rehwoo.csv") %>%
  # year required to join to habitat data
  mutate(year = year(observation_date))
# modis habitat covariates


k <- read_csv("data/year.pland-elev_location-year.csv") %>% arrange(id)
l <- read_csv("data/year.pland-slope_location-year.csv") %>% arrange(id)

habitat <- cbind(l, k[,15:16])

cols <- grep('y\\d+', names(habitat))
habitat$elevation_median[rowSums(habitat[cols] == 0) == length(cols)] <- 0
habitat$elevation_sd[rowSums(habitat[cols] == 0) == length(cols)] <- 0
habitat$slope_median[rowSums(habitat[cols] == 0) == length(cols)] <- 0
habitat$slope_sd[rowSums(habitat[cols] == 0) == length(cols)] <- 0

[1] "pland_00_water"                "pland_04_deciduous_broadleaf"
"pland_05_mixed_forest"         "pland_08_woody_savanna"
[5] "pland_09_savanna"              "pland_11_wetland"
"pland_12_cropland"             "pland_14_mosiac"
[9] "pland_13_urban"                "pland_10_grassland"
"pland_15_barren"               "pland_01_evergreen_needleleaf"
[13] "pland_02_evergreen_broadleaf"  "pland_03_deciduous_needleleaf"
"pland_07_open_shrubland"       "pland_06_closed_shrubland"

habitat1 <- habitat[habitat$lc_name %in% "pland_00_water",]


habitat1$year <- as.integer(habitat1$year)
rm(k, l)
# combine ebird and habitat data

ebird_habitat <- inner_join(ebird, habitat1, by = c("locality_id", "year"))
# prediction surface
```

```
j <- read_csv("data/l.water-elev_prediction-surface.csv") %>% arrange(id)
h <- read_csv("data/l.water-slope_prediction-surface.csv") %>% arrange(id)

pred_surface_2010 <- cbind(j, h[,15:16])

cols <- grep('y\\d+', names(pred_surface_2010))
pred_surface_2010$elevation_median[rowSums(pred_surface_2010[cols] == 0) ==
length(cols)] <- 0
pred_surface_2010$elevation_sd[rowSums(pred_surface_2010[cols] == 0) ==
length(cols)] <- 0
pred_surface_2010$slope_median[rowSums(pred_surface_2010[cols] == 0) ==
length(cols)] <- 0
pred_surface_2010$slope_sd[rowSums(pred_surface_2010[cols] == 0) ==
length(cols)] <- 0



rm(j, h)
# latest year of landcover data
max_lc_year <- max(pred_surface$year)
r <- raster("data/layer.prediction-surface.tif")
# load gis data for making maps
map_proj <- st_crs("ESRI:102003")
ne_land <- read_sf("data/gis-data.gpkg", "ne_land") %>%
  st_transform(crs = map_proj) %>%
  st_geometry()
bcr <- read_sf("data/gis-data.gpkg", "bcr") %>% filter(bcr_code %in% c(11,
12, 13, 17, 18, 19,21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31)) %>%
  st_transform(crs = map_proj) %>%
  st_geometry()
ne_country_lines <- read_sf("data/gis-data.gpkg", "ne_country_lines") %>%
  st_transform(crs = map_proj) %>%
  st_geometry()
ne_state_lines <- read_sf("data/gis-data.gpkg", "ne_state_lines") %>%
  st_transform(crs = map_proj) %>%
  st_geometry()

bb <- st_bbox(c(xmin = -0.1, xmax = 0.1, ymin = -0.1, ymax = 0.1),
              crs = 4326) %>%
  st_as_sfc() %>%
  st_sf()
# random points
pts <- st_sample(bb, 500) %>%
  st_sf(as.data.frame(st_coordinates(.)), geometry = .) %>%
  rename(lat = Y, lon = X)
# contruct a hexagonal grid with ~ 5 km between cells
dggs <- dgconstruct(spacing = 5)
# for each point, get the grid cell
pts$cell <- dgGEO_to_SEQNUM(dggs, pts$lon, pts$lat)$seqnum
# sample one checklist per grid cell
pts_ss <- pts %>%
  group_by(cell) %>%
  sample_n(size = 1) %>%
  ungroup()
# generate polygons for the grid cells
hexagons <- dgcellstogrid(dggs, unique(pts$cell), frame = FALSE) %>%
  st_as_sf()
ggplot() +
  geom_sf(data = hexagons) +
```

```
  geom_sf(data = pts, size = 0.5) +
  geom_sf(data = pts_ss, col = "red") +
  theme_bw()
set.seed(1)

dggs <- dgconstruct(spacing = 5)
# get hexagonal cell id and week number for each checklist
checklist_cell <- ebird_habitat %>%
  mutate(cell = dgGEO_to_SEQNUM(dggs, longitude, latitude)$seqnum,
         year = year(observation_date),
         week = week(observation_date))
# sample one checklist per grid cell per week
# sample detection/non-detection independently
ebird_ss <- checklist_cell %>%
  group_by(species_observed, year, week, cell) %>%
  sample_n(size = 1) %>%
  ungroup()

nrow(ebird_habitat)
count(ebird_habitat, species_observed) %>%
  mutate(percent = n / sum(n))
# after sampling
nrow(ebird_ss)
count(ebird_ss, species_observed) %>%
  mutate(percent = n / sum(n))

pct_before <- count(ebird_habitat, species_observed) %>%
  mutate(proportion = n / sum(n)) %>%
  filter(species_observed) %>%
  pull(proportion) %>%
  round(3) %>%
  scales::percent()
pct_after <- count(ebird_ss, species_observed) %>%
  mutate(proportion = n / sum(n)) %>%
  filter(species_observed) %>%
  pull(proportion) %>%
  round(3) %>%
  scales::percent()

ebird_split <- ebird_ss %>%
  # select only the columns to be used in the model
  select(species_observed,
         time_observations_started, duration_minutes,
         effort_distance_km, number_observers,
         starts_with("y"),
         starts_with("elevation_"),
         starts_with("slope_"))%>%
  drop_na()
# split 80/20
ebird_split <- ebird_split %>%
  split(if_else(runif(nrow(.)) <= 0.8, "train", "test"))
map_int(ebird_split, nrow)

detection_freq <- mean(ebird_split$train$species_observed)

ebird_split$train$species_observed <-
factor(ebird_split$train$species_observed)
# grow random forest

ebird_split$train$year <- NULL
```

```
rf <- ranger(formula =  species_observed ~ .,
             data = ebird_split$train,
             importance = "impurity",
             probability = TRUE,
             replace = TRUE,
             sample.fraction = c(detection_freq, detection_freq),)

occ_pred <- rf$predictions[, 2]
# convert the observered response back to a numeric value from factor
occ_obs <- ebird_split$train$species_observed %>%
  as.logical() %>%
  as.integer()
rf_pred_train <- tibble(obs = occ_obs, pred = occ_pred) %>%
  drop_na()
# fit calibration model
calibration_model <- scam(obs ~ s(pred, k = 5, bs = "mpi"),
                          gamma = 1.4,
                          data = rf_pred_train)

search_hours <- ebird_split$train %>%
  mutate(hour = floor(time_observations_started)) %>%
  count(hour) %>%
  mutate(pct = n / sum(n)) %>%
  filter(pct >= 0.01)
# constrained peak time
t_peak <- pd_time %>%
  filter(floor(time_observations_started) %in% search_hours$hour) %>%
  top_n(1, wt = desc(time_observations_started)) %>%
  pull(time_observations_started)
t_peak

human_time <- str_glue("{h}:{m} {ap}",
                       h = floor(t_peak),
                       m = str_pad(round((t_peak %% 1) * 60), 2, pad =
"0"),
                       ap = ifelse(t_peak > 12, "PM", "AM"))
#prepare checklist data for prediction using missRanger to fill missing
values
#library(missRanger)
#pred_surface <- missRanger(pred_surface, pmm.k = 3, num.trees = 100)
pred_surface_eff_2010 <- pred_surface_2010 %>%
  mutate(time_observations_started = t_peak,
         duration_minutes = 60,
         effort_distance_km = 1,
         number_observers = 1)


# predict
pred_rf_2010 <- predict(rf, data = pred_surface_eff_2010, type =
"response")
pred_rf_2010 <- pred_rf_2010$predictions[, 2]
# apply calibration models
pred_rf_cal_2010 <- predict(calibration_model,
                            data.frame(pred = pred_rf_2010),
                            type = "response")
# add to prediction surface
pred_er_2010 <- bind_cols(pred_surface_2010, encounter_rate =
pred_rf_2010_cal_2010) %>%
  select(latitude, longitude, encounter_rate, id, layer) %>%
  mutate(encounter_rate = pmin(pmax(encounter_rate, 0), 1))
#select the names of the columns
```

```
cols <- grep('y\\d+', names(pred_surface_2010))
#add those columns to the dataframe with predictions
pred_er_2010 <- pred_er_2010 %>% mutate(pred_surface_2010[, c(cols)])
cols <- grep('y\\d+', names(pred_er_2010))
#select only those predictions with landcover values otherwise turn them to
0
pred_er_2010$encounter_rate[rowSums(pred_er_2010[cols] == 0) ==
length(cols)] <- 0

pred_er_2010 <- pred_er_2010[, -c(cols)]

r_pred_2010 <- pred_er_2010 %>%
  # convert to spatial features
  st_as_sf(coords = c("longitude", "latitude"), crs = 4326) %>%
  st_transform(crs = projection(r)) %>%
  # rasterize
  rasterize(r)
r_pred_2010 <- r_pred_2010[[-1]]
# save the raster

writeRaster(r_pred_2010, "output/p.barren.tif")

##gather both increasing and decreasing landcover changes relative to
regions and encounter rate
library(raster)
library(sf)


habs <- stack("l.forest_cover.water.tif")
rf <- stack("output/p.water.tif")

#To make a colour ramp centered on zero:
rf <- rf[[c(1, 3)]]
map_proj <- st_crs("+proj=sinu +lon_0=0 +x_0=0 +y_0=0 +R=6371007.181
+units=m +no_defs")
habs <- projectRaster(habs, crs = map_proj$proj4string, method = "ngb")

habs <- habs[[3:13]]
minus <- function(a, b){
  a-b
}
habs_change <- overlay(habs[[3:11]], habs[[2]], fun=minus)
habs_change <- stack(habs[[1]], habs_change)
names(habs_change) <- c("layer", "y2011","y2012", "y2013", "y2014",
"y2015","y2016", "y2017", "y2018", "y2019")

############
# Below is a pipeline to extract random points from within species range
and compare these to habitat change
df<-getValues(habs_change)
rpoints <- data.frame(xyFromCell(habs_change,
1:ncell(habs_change)),getValues(habs_change))
rpoints <- subset(rpoints,!is.na(layer)) #limit to non-na points
y2011 <- rpoints[rpoints$y2011 != 0, c(1, 2)]
y2012 <- rpoints[rpoints$y2012 != 0, c(1, 2)]
y2013 <- rpoints[rpoints$y2013 != 0, c(1, 2)]
y2014 <- rpoints[rpoints$y2014 != 0, c(1, 2)]
y2015 <- rpoints[rpoints$y2015 != 0, c(1, 2)]
y2016 <- rpoints[rpoints$y2016 != 0, c(1, 2)]
y2017 <- rpoints[rpoints$y2017 != 0, c(1, 2)]
y2018 <- rpoints[rpoints$y2018 != 0, c(1, 2)]
```

```
y2019 <- rpoints[rpoints$y2019 != 0, c(1, 2)]


#randomly select 1000 points:
y2011 <- y2011[sample(1:nrow(y2011),1000,replace = F),]
y2012 <- y2011[sample(1:nrow(y2012),1000,replace = F),]
y2013 <- y2011[sample(1:nrow(y2013),1000,replace = F),]
y2014 <- y2011[sample(1:nrow(y2014),1000,replace = F),]
y2015 <- y2011[sample(1:nrow(y2015),1000,replace = F),]
y2016 <- y2011[sample(1:nrow(y2016),1000,replace = F),]
y2017 <- y2011[sample(1:nrow(y2017),1000,replace = F),]
y2018 <- y2011[sample(1:nrow(y2018),1000,replace = F),]
y2019 <- y2011[sample(1:nrow(y2019),1000,replace = F),]


#now extract data for bird encounter rate change and habitat change
# Example habitat change raster:

e_change.1 <- extract(rf, y2011) #extract cell values for bird change
h_change.1 <- data.frame(extract(habs_change[[c(1, 2)]],y2011))#extract
cell values for habitat change
h_change.1$year <- 2011
colnames(h_change.1)[2]<- "water"
y2011$id <- 1:nrow(y2011)
h_change.1$id <- 1:nrow(h_change.1)
h_change.1 <- merge(h_change.1, y2011)

e_change.2 <- extract(rf, y2012) #extract cell values for bird change
h_change.2 <- data.frame(extract(habs_change[[c(1, 3)]],y2012))#extract
cell values for habitat change
h_change.2$year <- 2012
colnames(h_change.2)[2]<- "water"
y2012$id <- 1:nrow(y2012)
h_change.2$id <- 1:nrow(h_change.2)
h_change.2 <- merge(y2012, h_change.2)

e_change.3 <- extract(rf, y2013) #extract cell values for bird change
h_change.3 <- data.frame(extract(habs_change[[c(1, 4)]],y2013))#extract
cell values for habitat change
h_change.3$year <- 2013
colnames(h_change.3)[2]<- "water"
y2013$id <- 1:nrow(y2013)
h_change.3$id <- 1:nrow(h_change.3)
h_change.3 <- merge(y2013, h_change.3)

e_change.4 <- extract(rf, y2014) #extract cell values for bird change
h_change.4 <- data.frame(extract(habs_change[[c(1, 5)]],y2014))#extract
cell values for habitat change
h_change.4$year <- 2014
colnames(h_change.4)[2]<- "water"
y2014$id <- 1:nrow(y2014)
h_change.4$id <- 1:nrow(h_change.4)
h_change.4 <- merge(y2014, h_change.4)

e_change.5 <- extract(rf, y2015) #extract cell values for bird change
h_change.5 <- data.frame(extract(habs_change[[c(1, 6)]],y2015))#extract
cell values for habitat change
h_change.5$year <- 2015
colnames(h_change.5)[2]<- "water"
y2015$id <- 1:nrow(y2015)
h_change.5$id <- 1:nrow(h_change.5)
```

```
h_change.5 <- merge(y2015, h_change.5)

e_change.6 <- extract(rf, y2016) #extract cell values for bird change
h_change.6 <- data.frame(extract(habs_change[[c(1, 7)]],y2016))#extract
cell values for habitat change
h_change.6$year <- 2016
colnames(h_change.6)[2]<- "water"
y2016$id <- 1:nrow(y2016)
h_change.6$id <- 1:nrow(h_change.6)
h_change.6 <- merge(y2016, h_change.6)

e_change.7 <- extract(rf, y2017) #extract cell values for bird change
h_change.7 <- data.frame(extract(habs_change[[c(1, 8)]],y2017))#extract
cell values for habitat change
h_change.7$year <- 2017
colnames(h_change.7)[2]<- "water"
y2017$id <- 1:nrow(y2017)
h_change.7$id <- 1:nrow(h_change.7)
h_change.7 <- merge(y2017, h_change.7)

e_change.8 <- extract(rf, y2018) #extract cell values for bird change
h_change.8 <- data.frame(extract(habs_change[[c(1, 9)]],y2018))#extract
cell values for habitat change
h_change.8$year <- 2018
colnames(h_change.8)[2]<- "water"
y2018$id <- 1:nrow(y2018)
h_change.8$id <- 1:nrow(h_change.8)
h_change.8 <- merge(y2018, h_change.8)

e_change.9 <- extract(rf, y2019) #extract cell values for bird change
h_change.9 <- data.frame(extract(habs_change[[c(1, 10)]],y2019))#extract
cell values for habitat change
h_change.9$year <- 2019
colnames(h_change.9)[2]<- "water"
y2019$id <- 1:nrow(y2019)
h_change.9$id <- 1:nrow(h_change.9)
h_change.9 <- merge(y2019, h_change.9)

r.all.change <-
rbind(e_change.1,e_change.2,e_change.3,e_change.4,e_change.5,e_change.6,e_c
hange.7,e_change.8,e_change.9)
h.all.change <-
rbind(h_change.1,h_change.2,h_change.3,h_change.4,h_change.5,h_change.6,h_c
hange.7,h_change.8,h_change.9)
c_change <- data.frame(r.all.change, h.all.change)
c_change <- c_change[, -2]
names(c_change) <- c("encounter","id", "layer", "water", "year", "X", "Y")


write.csv(c_change, "t.water.csv")


#filter into groups and then select for changes in pland
library(tidyverse)
t_cropland <- read.csv("q.cropland.csv")
t_deciduous_broadleaf <- read.csv("q.deciduous_broadleaf.csv")
t_grassland <- read.csv("q.grassland.csv")
t_mixed_forest <- read.csv("q.mixed_forest.csv")
t_mosaic <- read.csv("q.mosaic.csv")
t_urban <- read.csv("q.urban.csv")
t_wetland <- read.csv("q.wetland.csv")
```

```
t_savanna <- read.csv("q.savanna.csv")
t_woody_savanna <- read.csv("q.woody_savanna.csv")
t_water <- read.csv("q.water.csv")

names(t_cropland)[5] <- "cropland"
names(t_deciduous_broadleaf)[5] <- "deciduous_broadleaf"
names(t_grassland)[5] <- "grassland"
names(t_mixed_forest)[5] <- "mixed_forest"
names(t_mosaic)[5] <- "mosaic"
names(t_urban)[5] <- "urban"
names(t_wetland)[5] <- "wetland"
names(t_savanna)[5] <- "savanna"
names(t_woody_savanna)[5] <- "woody_savanna"

t_cropland <- t_cropland[, -1]
t_deciduous_broadleaf <- t_deciduous_broadleaf[, -1]
t_grassland <- t_grassland[, -1]
t_mixed_forest <- t_mixed_forest[, -1]
t_mosaic <- t_mosaic[, -1]
t_urban <- t_urban[, -1]
t_wetland <-t_wetland [, -1]
t_savanna <- t_savanna[, -1]
t_woody_savanna <- t_woody_savanna[, -1]
t_water <- t_water[, -1]


t_cropland <- t_cropland%>% pivot_longer(-c(1, 2, 3, 5, 6, 7))
t_deciduous_broadleaf <- t_deciduous_broadleaf%>% pivot_longer(-c(1, 2, 3,
5, 6, 7))
t_grassland <- t_grassland%>% pivot_longer(-c(1, 2, 3, 5, 6, 7))
t_mixed_forest <- t_mixed_forest%>% pivot_longer(-c(1, 2, 3, 5, 6, 7))
t_urban <- t_urban%>% pivot_longer(-c(1, 2, 3, 5, 6, 7))
t_wetland <- t_wetland%>% pivot_longer(-c(1, 2, 3, 5, 6, 7))
t_savanna <- t_savanna%>% pivot_longer(-c(1, 2, 3, 5, 6, 7))
t_woody_savanna <-t_woody_savanna%>% pivot_longer(-c(1, 2, 3, 5, 6, 7))
t_water <- t_water %>% pivot_longer(-c(1, 2, 3, 5, 6, 7))


ly_names <- tibble(layer = 1:17,
                   ly_name = c("Prairie_Potholes"     ,
                               "Boreal_Hardwood_Transition"   ,
                               "Lower_Great_Lakes_St.Lawrence_Plain",
                               "Badlands_And_Prairies"    ,
                               "Shortgrass_Prairie"    ,
                               "Central_Mixed_Grass_Prairie"    ,
                               "Oaks_And_Prairies"       ,
                               "Eastern_Tallgrass_Prairie"  ,
                               "Prairie_Hardwood_Transition" ,
                               "Central_Hardwoods"      ,
                               "West_Gulf_Coastal_Plain_Ouachitas",
                               "Mississippi_Alluvial_Valley" ,
                               "Southeastern_Coastal_Plain" ,
                               "Appalachian_Mountains" ,
                               "Piedmont"       ,
                               "New_England_Mid_Atlantic_Coast" ,
                               "Peninsular_Florida" ))

Woodland <- rbind(t_deciduous_broadleaf, t_mixed_forest)
Woodland <- Woodland[, -2]
Woodland <- Woodland[, c(4, 5, 3, 2, 6, 7, 1)]
```

```
colnames(Woodland) <- c("X", "Y", "year", "layer", "landcover",
"pland_change", "encounter")
Woodland <- Woodland %>%
  inner_join(ly_names, by = "layer") %>%
  arrange(layer) %>%
  select(-layer)

Woodland <- Woodland[complete.cases(Woodland),]
Woodland.2 <- Woodland[Woodland$pland_change <=0,]
Woodland.3 <- Woodland[Woodland$pland_change >=0,]


#landcover
Woodland.2$landcover <- as.factor(Woodland.2$landcover)
Woodland.2$landcover <- as.numeric(Woodland.2$landcover)
Woodland.3$landcover <- as.factor(Woodland.3$landcover)
Woodland.3$landcover <- as.numeric(Woodland.3$landcover)

#ly_name
Woodland.2$ly_name <- as.factor(Woodland.2$ly_name)
Woodland.2$ly_name <- as.numeric(Woodland.2$ly_name)
Woodland.3$ly_name <- as.factor(Woodland.3$ly_name)
Woodland.3$ly_name <- as.numeric(Woodland.3$ly_name)

#correlation test
cor.test(Woodland.3$ly_name, Woodland.3$encounter)
cor.test(Woodland.3$landcover, Woodland.3$encounter)
cor.test(Woodland.2$ly_name, Woodland.2$encounter)
cor.test(Woodland.2$landcover, Woodland.2$encounter)

#
savannas <- rbind(t_woody_savanna, t_savanna)
savannas <- savannas[, -2]
savannas <- savannas[, c(4, 5, 3, 2, 6, 7, 1)]
colnames(savannas) <- c("X", "Y", "year", "layer", "landcover",
"pland_change", "encounter")
savannas <- savannas %>%
  inner_join(ly_names, by = "layer") %>%
  arrange(layer) %>%
  select(-layer)

savannas <- savannas[complete.cases(savannas),]
savannas.2 <- savannas[savannas$pland_change <=0,]
savannas.3 <- savannas[savannas$pland_change >=0,]

#landcover
savannas.2$landcover <- as.factor(savannas.2$landcover)
savannas.2$landcover <- as.numeric(savannas.2$landcover)
savannas.3$landcover <- as.factor(savannas.3$landcover)
savannas.3$landcover <- as.numeric(savannas.3$landcover)

#ly_name
savannas.2$ly_name <- as.factor(savannas.2$ly_name)
savannas.2$ly_name <- as.numeric(savannas.2$ly_name)
savannas.3$ly_name <- as.factor(savannas.3$ly_name)
savannas.3$ly_name <- as.numeric(savannas.3$ly_name)

#correlation test
cor.test(savannas.3$ly_name, savannas.3$encounter)
cor.test(savannas.3$landcover, savannas.3$encounter)
cor.test(savannas.2$ly_name, savannas.2$encounter)
```

```
cor.test(savannas.2$landcover, savannas.2$encounter)

agricultural <- rbind(t_cropland, t_grassland)
agricultural <- agricultural[, -2]
agricultural <- agricultural[, c(4, 5, 3, 2, 6, 7, 1)]
colnames(agricultural) <- c("X", "Y", "year", "layer", "landcover",
"pland_change", "encounter")
agricultural <- agricultural %>%
  inner_join(ly_names, by = "layer") %>%
  arrange(layer) %>%
  select(-layer)

agricultural <- agricultural[complete.cases(agricultural),]
agricultural.2 <- agricultural[agricultural$pland_change <=0,]
agricultural.3 <- agricultural[agricultural$pland_change >=0,]

#landcover
agricultural.2$landcover <- as.factor(agricultural.2$landcover)
agricultural.2$landcover <- as.numeric(agricultural.2$landcover)
agricultural.3$landcover <- as.factor(agricultural.3$landcover)
agricultural.3$landcover <- as.numeric(agricultural.3$landcover)

#ly_name
agricultural.2$ly_name <- as.factor(agricultural.2$ly_name)
agricultural.2$ly_name <- as.numeric(agricultural.2$ly_name)
agricultural.3$ly_name <- as.factor(agricultural.3$ly_name)
agricultural.3$ly_name <- as.numeric(agricultural.3$ly_name)

#correlation test
cor.test(agricultural.3$ly_name, agricultural.3$encounter)
cor.test(agricultural.3$landcover, agricultural.3$encounter)
cor.test(agricultural.2$ly_name, agricultural.2$encounter)
cor.test(agricultural.2$landcover, agricultural.2$encounter)

urban <- rbind(t_urban)
urban <- urban[, -2]
urban <- urban[, c(4, 5, 3, 2, 6, 7, 1)]
colnames(urban) <- c("X", "Y", "year", "layer", "landcover",
"pland_change", "encounter")
urban <- urban %>%
  inner_join(ly_names, by = "layer") %>%
  arrange(layer) %>%
  select(-layer)

urban <- urban[complete.cases(urban),]
urban.2 <- urban[urban$pland_change <=0,]
urban.3 <- urban[urban$pland_change >=0,]

#landcover
urban.2$landcover <- as.factor(urban.2$landcover)
urban.2$landcover <- as.numeric(urban.2$landcover)
urban.3$landcover <- as.factor(urban.3$landcover)
urban.3$landcover <- as.numeric(urban.3$landcover)

#ly_name
urban.2$ly_name <- as.factor(urban.2$ly_name)
urban.2$ly_name <- as.numeric(urban.2$ly_name)
urban.3$ly_name <- as.factor(urban.3$ly_name)
urban.3$ly_name <- as.numeric(urban.3$ly_name)

#correlation test
```

```
cor.test(urban.3$ly_name, urban.3$encounter)
cor.test(urban.3$landcover, urban.3$encounter)
cor.test(urban.2$ly_name, urban.2$encounter)
cor.test(urban.2$landcover, urban.2$encounter)
cor.test(urban.3$pland_change, urban.3$encounter)


water_bodies <- rbind(t_water, t_wetland)
water_bodies <- water_bodies[, -2]
water_bodies <- water_bodies[, c(4, 5, 3, 2, 6, 7, 1)]
colnames(water_bodies) <- c("X", "Y", "year", "layer", "landcover",
"pland_change", "encounter")
water_bodies <- water_bodies %>%
  inner_join(ly_names, by = "layer") %>%
  arrange(layer) %>%
  select(-layer)

water_bodies <- water_bodies[complete.cases(water_bodies),]
water_bodies.2 <- water_bodies[water_bodies$pland_change <=0,]
water_bodies.3 <- water_bodies[water_bodies$pland_change >=0,]

#landcover
water_bodies.2$landcover <- as.factor(water_bodies.2$landcover)
water_bodies.2$landcover <- as.numeric(water_bodies.2$landcover)
water_bodies.3$landcover <- as.factor(water_bodies.3$landcover)
water_bodies.3$landcover <- as.numeric(water_bodies.3$landcover)

#ly_name
water_bodies.3$ly_name <- as.factor(water_bodies.3$ly_name)
water_bodies.3$ly_name <- as.numeric(water_bodies.3$ly_name)
water_bodies.3$ly_name <- as.factor(water_bodies.3$ly_name)
water_bodies.3$ly_name <- as.numeric(water_bodies.3$ly_name)

#correlation test
cor.test(water_bodies.3$ly_name, water_bodies.3$encounter)
cor.test(water_bodies.3$landcover, water_bodies.3$encounter)
cor.test(water_bodies.2$ly_name, water_bodies.2$encounter)
cor.test(water_bodies.2$landcover, water_bodies.2$encounter)

#create a big dataframe containing both decreases and increases
test.all <-
rbind(t_cropland,t_deciduous_broadleaf,t_grassland,t_mixed_forest,t_urban,t
_wetland,t_savanna,t_woody_savanna,t_water)
test.all <- test.all[, -2]
test.all <- test.all[, c(4, 5, 3, 2, 6, 7, 1)]

colnames(test.all) <- c("X", "Y", "year", "layer", "landcover",
"pland_change", "encounter")

#library(missRanger)
#test.all.1<- missRanger(test.all, pmm.k = 3, num.trees = 100)

ly_names <- tibble(layer = 1:17,
                   ly_name = c("Prairie_Potholes"     ,
                               "Boreal_Hardwood_Transition"    ,
                               "Lower_Great_Lakes_St.Lawrence_Plain",
                               "Badlands_And_Prairies"     ,
                               "Shortgrass_Prairie"     ,
                               "Central_Mixed_Grass_Prairie"    ,
                               "Oaks_And_Prairies"       ,
                               "Eastern_Tallgrass_Prairie"   ,
```

```
                               "Prairie_Hardwood_Transition" ,
                               "Central_Hardwoods"     ,
                               "West_Gulf_Coastal_Plain_Ouachitas",
                               "Mississippi_Alluvial_Valley" ,
                               "Southeastern_Coastal_Plain" ,
                               "Appalachian_Mountains" ,
                               "Piedmont"      ,
                               "New_England_Mid_Atlantic_Coast" ,
                               "Peninsular_Florida" ))

test.all <- test.all %>%
  inner_join(ly_names, by = "layer") %>%
  arrange(layer) %>%
  select(-layer)
#split this dataframe
test.all <- test.all[complete.cases(test.all),]
test.all.2 <- test.all[test.all$pland_change <=0,]
test.all.3 <- test.all[test.all$pland_change >=0,]

round(max(test.all.3$encounter), 2)/10

#calculate ranges of probability
myIntervals <- c("0 - 0.1", "0.1 - 0.2 ", "0.2  - 0.3", "0.3  - 0.4","0.4
- 0.5","0.5  - 0.6","0.6  - 0.6","0.7  - 0.8","0.8 - 0.8","0.9  - 1")
test.all.3$encounter_range<- myIntervals[findInterval(test.all.3$encounter,
c(0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1))]

tt3.l <- test.all.3 %>% count(encounter_range, landcover, ly_name) %>%
group_by(ly_name)

#plot these ranges but first make sure the colours match the number of bcr
regions
library(RColorBrewer)
brks <- seq(0, 1, 0.057)
cols <-
colorRampPalette(c("red4","gray95","orange","gold","lightskyblue","dodgerbl
ue","blue3"))(length(brks)-1) #this is the colour ramp - change the colours
as you like, bu make sure they are symetrical around the "gray95", which
represents the zero point!

forestlossline <- ggplot(tt3.l, aes(x=factor(encounter_range), y=n,
group=ly_name, label=round(n, 1))) +
  geom_bar(aes(fill=ly_name), stat='identity') +
  scale_colour_manual(values=cols) +
  scale_fill_manual(values=cols)+
  ylab("Counts of Encounters in BCR regions for Landcover classes") +
  xlab("Encounter rate Ranges") +
  facet_wrap(~landcover, scales="free") +
  theme_classic() +
  scale_x_discrete(label = function(x) stringr::str_trunc(x, 12))+
  theme(
    axis.title.x=element_text(face="bold", size=13, family="TT Times New
Roman"),
    axis.title.y=element_text(face="bold", size=13, family="TT Times New
Roman"),
    axis.text.x=element_text(face="bold", size =10, family="TT Times New
Roman", angle=90),
    axis.text.y=element_text(face="bold",  size =10, family="TT Times New
Roman"),
    legend.title=element_blank(),
```

```
    legend.text=element_text(color="black", size =10, face="bold",
family="TT Times New Roman"),
    legend.justification=c(1.2,1),
    plot.title=element_text(face="bold", size = 18, hjust=0.5, colour =
"black"),
    axis.line=element_blank(),
    legend.key.height=unit(.1, "cm")
  )


#calculate the stepforward regression using the stepfor function built
library(stepfor)

tbest1 <- stepfor(savannas.2$encounter, savannas.2[, -c(1, 4, 2, 6)],
alpha=0.2)



#get points from landcover changes
#calculate a buffer radius around each point of observation
neighborhood_radius <- 5 * ceiling(max(res(landcover))) / 2
ebird_buff <- Woodland.3[Woodland.3$landcover %in% "deciduous_broadleaf",]
%>%
  distinct(year, X, Y,encounter) %>%
  # for 2019 use 2018 landcover data
  mutate(year_lc = if_else(as.integer(year) > max_lc_year,
                           as.character(max_lc_year), as.character(year)),
         year_lc = paste0("y", year_lc)) %>%
  # convert to spatial features
  st_as_sf(coords = c("X", "Y")) %>%
  # buffer to create neighborhood around each point
  st_buffer(dist = neighborhood_radius) %>%
  # nest by year
  nest(data = c(year, encounter, geometry))

#raster
agg_factor <- round(2 * neighborhood_radius / res(landcover))
r <- raster(landcover) %>%
  aggregate(agg_factor)
########################################################

#conver to dataframe with geometry points
ebird_buff <- ebird_buff %>% arrange(year_lc)
#2010
bird_buff <- NULL
for(i in 1:9){
  bird_buff[[i]] <- ebird_buff$data[i] %>%
    bind_cols %>%
    st_cast(to = "POINT") %>%
    dplyr::mutate(
      X =  sf::st_coordinates(geometry)[,1], #retrieve X coord
      Y =  sf::st_coordinates(geometry)[,2]  #retrieve Y coord
    ) %>%
    sf::st_drop_geometry()}

#rbind the lists
bird_buff <- do.call(rbind.data.frame, bird_buff)


bird_buff$year <- as.integer(bird_buff$year)
bird_buff$X <- as.numeric(bird_buff$X)
```

```
bird_buff$Y <- as.numeric(bird_buff$Y)

#convert into raster
bird.raster <- bird_buff %>% st_as_sf(coords = c("X", "Y")) %>%
rasterize(r)
map_proj <- st_crs("ESRI:102003")
r_bird<- projectRaster(bird.raster, crs = map_proj$proj4string, method =
"ngb")

#extract points and create a spatial points
rpoints <- rasterToPoints(r_bird) %>% data.frame()
xy <- SpatialPointsDataFrame(coords = rpoints[, 1:2], rpoints, proj4string
= CRS("+proj=aea +lat_0=37.5 +lon_0=-96 +lat_1=29.5 +lat_2=45.5 +x_0=0
+y_0=0 +datum=NAD83 +units=m +no_defs"))


#Load the maps into R and filter the coordinates points as you load them
also

#
map_proj <- st_crs("ESRI:102003")
##########################
p.water <- stack("output/p.water.tif")
p.water <- projectRaster(p.water[[1]], crs = map_proj$proj4string, method =
"ngb")
##########################
p.evergreen_needleleaf <- stack("output/p.evergreen_needleleaf.tif")
p.evergreen_needleleaf <- projectRaster(p.evergreen_needleleaf[[1]], crs =
map_proj$proj4string, method = "ngb")
##########################
p.evergreen_broadleaf <- stack("output/p.evergreen_broadleaf.tif")
p.evergreen_broadleaf <- projectRaster(p.evergreen_broadleaf[[1]], crs =
map_proj$proj4string, method = "ngb")
##########################
p.deciduous_broadleaf <- stack("output/p.deciduous_broadleaf.tif")
p.deciduous_broadleaf <- projectRaster(p.deciduous_broadleaf[[1]], crs =
map_proj$proj4string, method = "ngb")
##########################
p.deciduous_needleleaf <- stack("output/p.deciduous_needleleaf.tif")
p.deciduous_needleleaf <- projectRaster(p.deciduous_needleleaf[[1]], crs =
map_proj$proj4string, method = "ngb")
##########################
p.mixed_forest <- stack("output/p.mixed_forest.tif")
p.mixed_forest <- projectRaster(p.mixed_forest[[1]], crs =
map_proj$proj4string, method = "ngb")
##########################
p.closed_shrubland <- stack("output/p.closed_shrubland.tif")
p.closed_shrubland <- projectRaster(p.closed_shrubland[[1]], crs =
map_proj$proj4string, method = "ngb")
##########################
p.open_shrubland <- stack("output/p.open_shrubland.tif")
p.open_shrubland <- projectRaster(p.open_shrubland[[1]], crs =
map_proj$proj4string, method = "ngb")
##########################
p.woody_savanna <- stack("output/p.woody_savanna.tif")
p.woody_savanna <- projectRaster(p.woody_savanna[[1]], crs =
map_proj$proj4string, method = "ngb")
##########################
p.savanna <- stack("output/p.savanna.tif")
p.savanna <- projectRaster(p.savanna[[1]], crs = map_proj$proj4string,
method = "ngb")
```

```
##########################
p.grassland <- stack("output/p.grassland.tif")
p.grassland <- projectRaster(p.grassland[[1]], crs = map_proj$proj4string,
method = "ngb")
##########################
p.cropland <- stack("output/p.cropland.tif")
p.cropland <- projectRaster(p.cropland[[1]], crs = map_proj$proj4string,
method = "ngb")
##########################
p.urban <- stack("output/p.urban.tif")
p.urban <- projectRaster(p.urban[[1]], crs = map_proj$proj4string, method =
"ngb")
##########################
p.wetland <- stack("output/p.wetland.tif")
p.wetland <- projectRaster(p.wetland[[1]], crs = map_proj$proj4string,
method = "ngb")
##########################
p.mosaic <- stack("output/p.mosaic.tif")
p.mosaic <- projectRaster(p.mosaic[[1]], crs = map_proj$proj4string, method
= "ngb")
##########################
p.barren <- stack("output/p.barren.tif")
p.barren <- projectRaster(p.barren[[1]], crs = map_proj$proj4string, method
= "ngb")


###########################################################################
###########
xy_barren <- xy_barren[,-c(1, 7:9)]
xy_water <- xy_water[,-c(1, 7:9)]
xy_evergreen_needleleaf <- xy_evergreen_needleleaf[,-c(1, 7:9)]
xy_evergreen_broadleaf <- xy_evergreen_broadleaf[,-c(1, 7:9)]
xy_deciduous_broadleaf <- xy_deciduous_broadleaf[,-c(1, 7:9)]
xy_mixed_forest <- xy_mixed_forest[,-c(1, 7:9)]
xy_open_shrubland <- xy_open_shrubland[,-c(1, 7:9)]
xy_savanna <- xy_savanna[,-c(1, 7:9)]
xy_grassland <- xy_grassland[,-c(1, 7:9)]
xy_wetland <- xy_wetland[,-c(1, 7:9)]
xy_cropland <- xy_cropland[,-c(1, 7:9)]
xy_mosaic <- xy_mosaic[,-c(1, 7:9)]
xy_woody_savanna <- xy_woody_savanna[,-c(1, 7:9)]
ixy_barren <- ixy_barren[,-c(1, 7:9)]
ixy_water <- ixy_water[,-c(1, 7:9)]
ixy_evergreen_needleleaf <- ixy_evergreen_needleleaf[,-c(1, 7:9)]
ixy_evergreen_broadleaf <- ixy_evergreen_broadleaf[,-c(1, 7:9)]
ixy_deciduous_needleleaf <- ixy_deciduous_needleleaf[,-c(1, 7:9)]
ixy_deciduous_broadleaf <- ixy_deciduous_broadleaf[,-c(1, 7:9)]
ixy_mixed_forest <- ixy_mixed_forest[,-c(1, 7:9)]
ixy_closed_shrubland <- ixy_closed_shrubland[,-c(1, 7:9)]
ixy_open_shrubland <- ixy_open_shrubland[,-c(1, 7:9)]
ixy_woody_savanna <- ixy_woody_savanna[,-c(1, 7:9)]
ixy_savanna <- ixy_savanna[,-c(1, 7:9)]
ixy_grassland <- ixy_grassland[,-c(1, 7:9)]
ixy_wetland <- ixy_wetland[,-c(1, 7:9)]
ixy_cropland <- ixy_cropland[,-c(1, 7:9)]
ixy_mosaic <- ixy_mosaic[,-c(1, 7:9)]
ixy_urban <- ixy_urban[,-c(1, 7:9)]

#create maps
windows(10, 8)
```

```
par(mar = c(3.5, 0.25, 0.25, 0.25))
# set up plot area
plot(bcr, col = NA, border = NA)
plot(ne_land, col = "#dddddd", border = "#888888", lwd = 0.5, add = TRUE)

# encounter rate
r_max <- ceiling(10 * cellStats(p.water, max)) / 10
brks <- seq(0, r_max, by = 0.006)

cols <-
colorRampPalette(c("red4","gray95","orange","gold","lightskyblue","dodgerbl
ue","blue3"))(length(brks)-1) #this is the colour ramp - change the colours
as you like, bu make sure they are symetrical around the "gray95", which
represents the zero point!

plot(p.water,
     col = cols, breaks = brks,
     maxpixels = ncell(p.water),
     legend = FALSE, add = TRUE)
points(xy_water[, 1:2], add=TRUE, pch=".", cex=2, col="darkred")
points(ixy_water[, 1:2], add=TRUE, pch=1, cex=1, col="blue")

# borders
plot(bcr, border = "#000000", col = NA, lwd = 1, add = TRUE)

box()

brks <- seq(0, 1, by = 0.05)
cols <-
colorRampPalette(c("red4","gray95","orange","gold","lightskyblue","dodgerbl
ue","blue3"))(length(brks)-1) #this is the colour ramp - change the colours
as you like, bu make sure they are symetrical around the "gray95", which
represents the zero point!

lbl_brks <- seq(0, 1, by = 0.05)
# legend
par(new = TRUE, mar = c(0, 0, 0, 0))
title <- "Red-headed Woodpecker encounter rate"
image.plot(zlim = range(brks), legend.only = TRUE,
           col = cols, breaks = brks,
           smallplot = c(0.25, 0.75, 0.06, 0.09),
           horizontal = TRUE,
           axis.args = list(at = lbl_brks, labels = lbl_brks,
                            fg = "black", col.axis = "black",
                            cex.axis = 0.75, lwd.ticks = 0.5,
                            padj = -1.5),
           legend.args = list(text = title,
                              side = 3, col = "black",
                              cex = 1, line = 0))
```