Spatial Analysis in Habitat Suitability of the Horned Grebe

Introduction

This research project focuses on the Horned Grebe (*Podiceps auritus*), an endangered species in British Columbia characterized by its unique breeding plumage and behaviors. British Columbia offers a diverse array of ecological environments, ranging from coastal wetlands to freshwater systems. This region is ideal for studying *Podiceps auritus* due to its varied geographic and climatic conditions which influence habitat availability and quality. Our study aims to explore the correlation between multiple environmental factors, which are elevation, forest cover, Human Footprint Index (HFI), distance to water bodies, and habitat suitability for *Podiceps auritus*. We use a robust dataset encompassing observations of endangered birds to analyze the specific habitat requirements of the Horned Grebe. Given BC's varied geographic and climatic conditions, the project delves into the dataset contains observations of the different endangered bird species under different environmental pressures, and specifies Horned Grebe, throughout the study, it aims to provide a better understanding of how these factors influence the Grebe's habitat preferences and survival throughout the year. By assessing the impact of both natural and anthropogenic elements on the habitats of the Horned Grebe, this study also seeks to provide insights crucial for developing effective conservation strategies and managing habitats to support the survival of this endangered species. This project aims to provide critical insights to conservationists and policymakers. By analyzing the environmental dependencies of the Horned Grebe, we seek to inform strategies to ensure the persistence and protection of this species in British Columbia.

Methods

This study uses species occurrence data obtained from the Global Biodiversity Information Facility (GBIF) for endangered bird species. The GBIF records are particularly valuable as they provide a specific focus on Podiceps auritus, offering detailed information on each sighting, including the observation location, which empowers us to perform spatial analyses to visualize and quantify distribution patterns across the varied landscapes of BC. To further enhance our understanding of the environmental determinants of habitat suitability, we have integrated several environmental covariates into our analysis. These covariates include Elevation data, which will allow us to explore the altitudinal preferences of the species; Forest cover percentage data, to investigate the influence of forest density on the species distribution; the Human Footprint Index (HFI), providing a comprehensive measure of human impact on natural habitats; and data on Water Distribution, which is particularly important for this aquatic bird species.

Our study begins with an initial phase of data visualization to understand the distribution of endangered bird species across British Columbia. We will conduct our analyses in RStudio, employing point pattern process (ppp) functions. These functions will help us visualize the locations where endangered birds have been observed, providing an overview of species distributions and identifying hotspots. After visualizing the overall distribution of endangered birds, we will refine our dataset by filtering observations to include only data collected after the year 2015. This ensures that our analysis incorporates the most recent and relevant information, reflecting current habitat conditions and species distributions. From this refined dataset, we will specifically focus on the Horned Grebe (*Podiceps auritus*), identified as the top endangered bird in this dataset.

In the detailed analysis phase, we delve deeper into the spatial analysis of the Horned Grebe's distribution. To estimate the intensity of our data, we first used quadrat counting to plot the intensity and tested for homogeneity. We employ Kernel Density Estimation to create a continuous surface that highlights areas of high density, effectively visualizing potential hotspots where conservation efforts can be concentrated. In addition, hotspot analysis pinpoints statistically significant clusters of Horned Grebe occurrences, providing a targeted approach for future detailed studies and conservation actions. As we explore the relationship between the Horned Grebe's distribution and various environmental covariates such as elevation, forest cover, the Human Footprint Index (HFI), and proximity to water bodies, we use advanced spatial statistics techniques. This includes assessing the intensity and spatial inhomogeneity of the species' occurrences to understand how population density varies with geographic and environmental heterogeneity. A non-parametric estimate of ρ was obtained via kernel estimation. Additionally, the K-function and Pair Correlation Function (PCF) are applied to model and analyze spatial point patterns at multiple scales, revealing underlying clustering or dispersion patterns relative to environmental variables.

The critical step in the methodology is constructing a predictive model to forecast habitat suitability for the Horned Grebe under various environmental scenarios. To ensure that our predictive models are based on independent factors, the collinearity analysis is required to refine the model by confirming that each variable contributes unique information. The robustness of this model can be assessed using the Akaike Information Criterion (AIC) and Analysis of Variance (ANOVA) to evaluate model fit and test the significance of model covariates. These tests are crucial for validating the predictive accuracy and reliability of our model.

The iterative process of model validation and refinement employed in our study follows a systematic approach to ensure the robustness of our statistical model. Initially, we compute the Akaike Information Criterion (AIC) for each successive model iteration. A reduction in

AIC values suggests an improved model fit despite the added complexity of additional terms. To quantitatively assess the significance of including an extra term, we conduct likelihood ratio tests comparing the fit of nested models. A small p-value from this test indicates that the additional term significantly enhances the model. Following each iteration, we generate partial residual plots, which visually represent the fit of individual covariates while controlling for others. This graphical assessment aids in detecting patterns, such as non-linearity or outliers, that might not be apparent through numerical measures alone. Each covariate's inclusion is thus substantiated by both statistical evidence and visual diagnostics, ensuring that our final model is both statistically sound and interpretable in relation to the ecological data on *Podiceps auritus*.

Results

Figure 1 is a stacked bar chart that represents the distribution of observations by year and by season. Each bar corresponds to a year, ranging from 2016 to 2023, and is divided into segments that represent the count of observations made in each of the four seasons: spring, summer, autumn, and winter. From the chart, it appears that the total number of observations has generally been increasing each year. The largest proportion of observations is made in the summer and autumn seasons, which could suggest that the species being observed is more visible or active during these times, or perhaps that these seasons are more conducive to research activities. Spring also has a significant number of observations, while winter has the fewest, which might reflect the seasonal behavior of the species or could be due to the difficulty of conducting observations during the winter months. In terms of trends, we do not see a consistent increase or decrease in observations for any particular season across the years. Instead, there seems to be a general upward trend in the total count of observations.

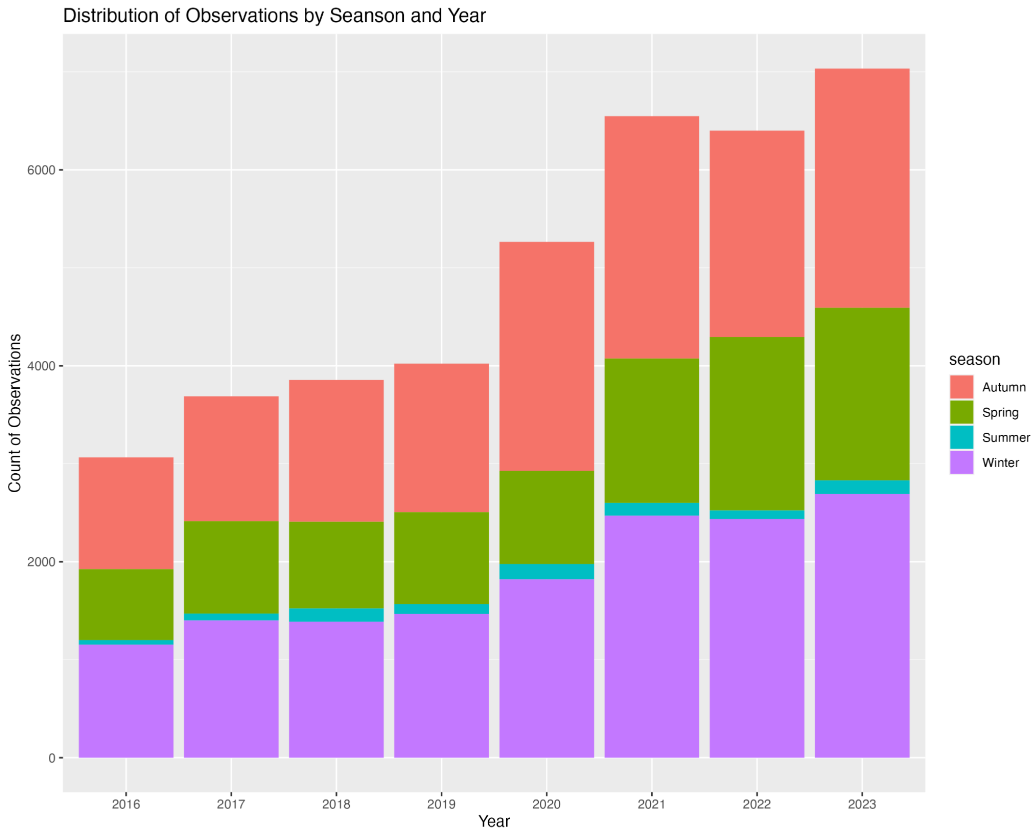

Figure 1: Distribution of Observation by Seasons and Years

The quadrat count analysis was conducted to assess the intensity and distribution of bird occurrences across British Columbia. In Figure 2, the map shows the kernel density estimate of bird observations across a region, presumably British Columbia. The intensity of observations is indicated by the color gradient on the map, with darker areas corresponding to higher observation densities and lighter areas indicating lower densities. The actual observation points are marked with dots, and their clustering is visible in certain areas. From the map, we can infer that there are particular hotspots where bird observations are significantly concentrated, as indicated by the areas in red and yellow. These hotspots may represent areas with optimal habitat conditions or may be key breeding or feeding grounds for the birds. Conversely, the regions in blue show sparse observations, suggesting these areas might be less suitable or harder to access for observation purposes.

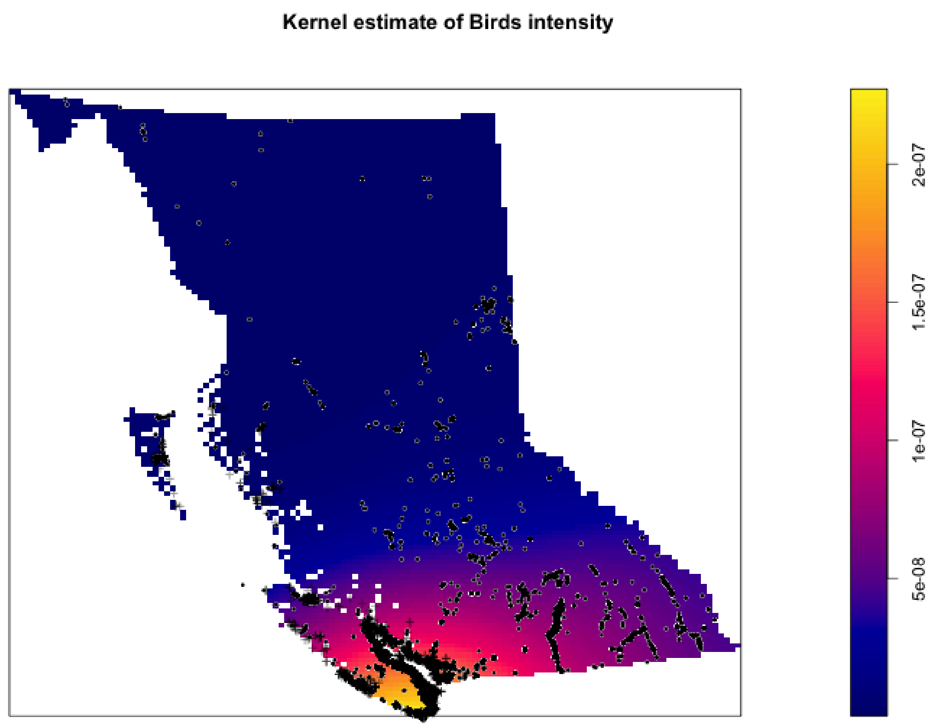

Figure 2: Kernel Estimate of Bird Intensity

Figure 3 depicts the results of a Hotspot Analysis applied to bird occurrence data. The left image shows a Likelihood Ratio (LR) statistic for each location, which assesses the intensity of occurrences. Areas with high LR values, indicated by warmer colors on the scale, suggest regions with significantly more observations than expected under a random spatial distribution, denoting potential hotspots. The right image presents the local p-values derived from the LR statistic, where lower values (closer to 0, shown in blue) correspond to areas where the observed number of occurrences is significantly higher than expected, indicating the presence of hotspots with a high degree of confidence. Conversely, areas with higher p-values (closer to 1, shown in yellow) are not significantly different from a random distribution, suggesting the absence of hotspots.

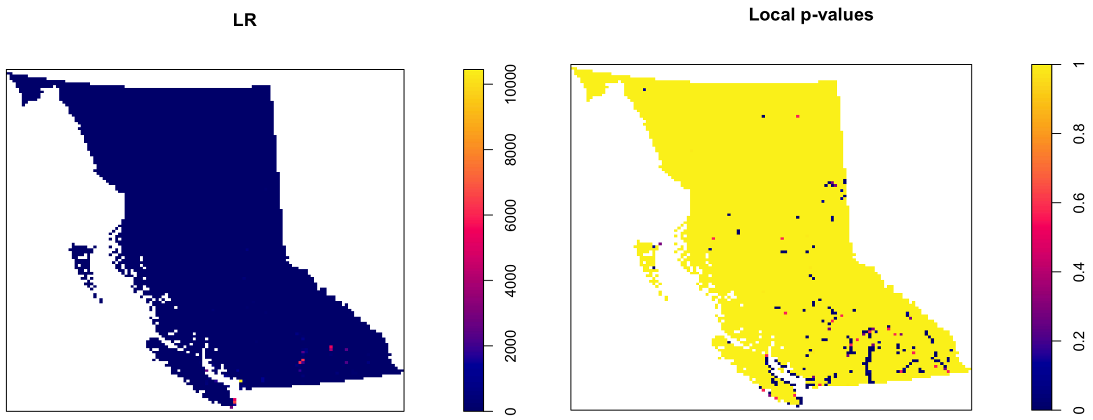

Figure 3: Hotspot Analysis

The Ripley's *K*-function analysis, as depicted in Figure 4, was utilized to quantify the spatial clustering of bird occurrences within the study area. The analysis reveals that the observed K-function (in black) lies substantially above the theoretical line (in red), suggesting a significant degree of spatial clustering at various scales of distance (r). This deviation from the CSR (Complete Spatial Randomness) indicates that bird occurrences are not randomly distributed but rather exhibit aggregation patterns, likely influenced by underlying environmental gradients or social behaviors intrinsic to the species. The inhomogeneous correction further refines these insights by considering local variations in point intensity, providing a nuanced understanding of spatial structure.

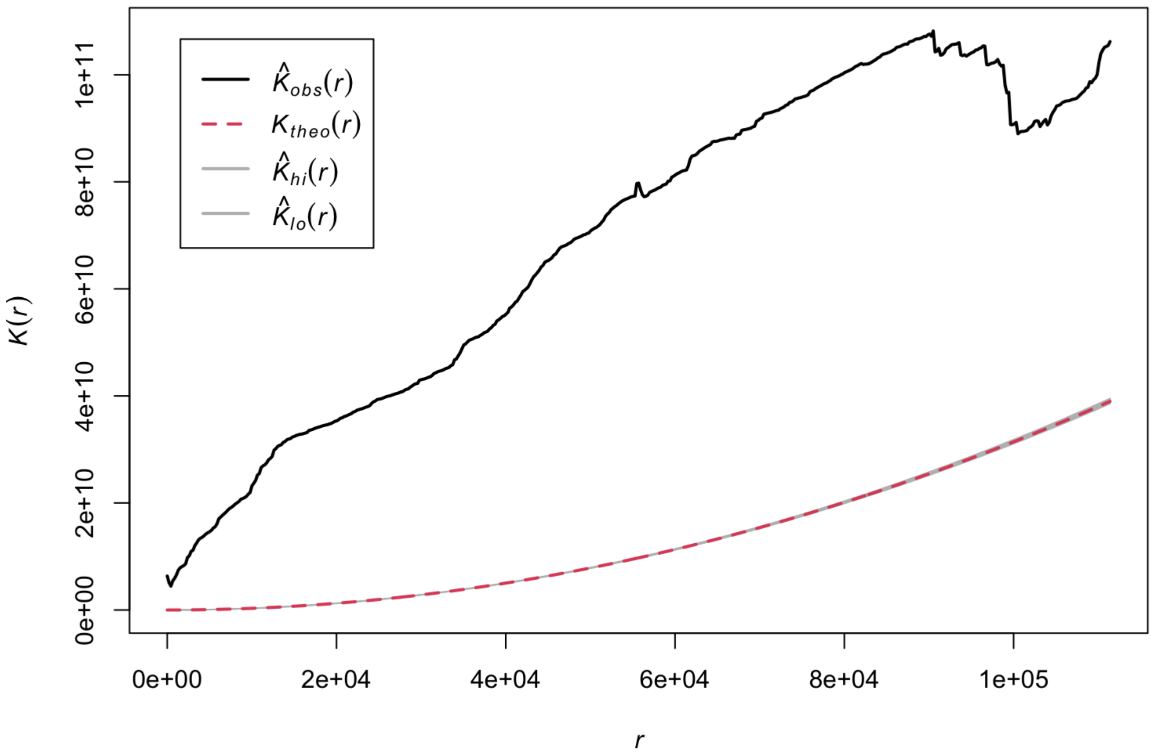

Figure 4: Homogeneous *K*-function

In Figure 5, the pair of plots illustrate the inhomogeneous K-function, which accounts for spatial inhomogeneity in the point process. The left plot provides a broad overview, while the right plot is a focused zoom-in on the data, allowing for detailed examination at smaller spatial scales. The black solid line represents the observed inhomogeneous K-function, whereas the dashed red line indicates the expected value under a model of inhomogeneity. The shaded area around the observed line represents the confidence envelope, calculated from simulations, giving a statistical context to the observed function. The fact that the observed line lies within the envelope suggests that the point pattern's spatial clustering can be explained by the model of inhomogeneity at various scales. When corrected for inhomogeneity, significant clustering only appears to exist in and around 0-20000 meters. The deviations appear much less meaningful. This would suggest that much of the correlations between points are due to relationships with covariates, rather than relationships between the points.

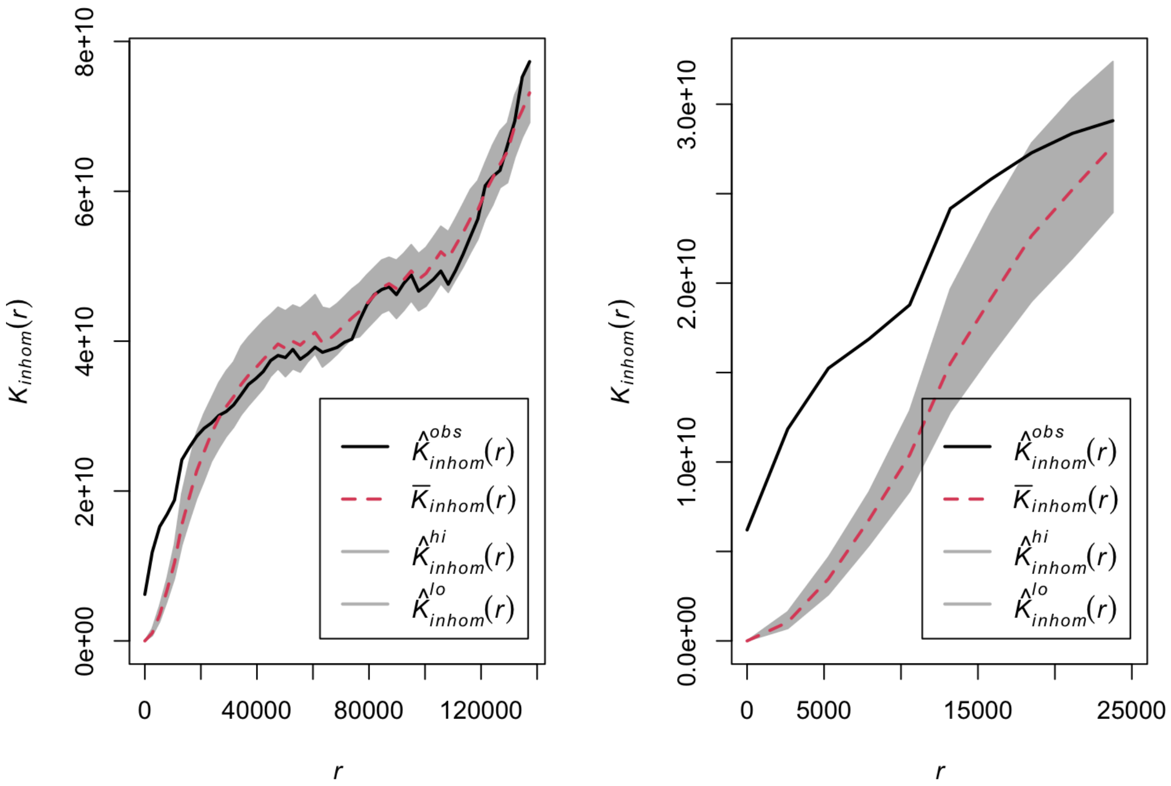

Figure 5: Inhomogeneous *K*-function

The pair of graphs in Figure 6 shows the PCFs for the point pattern, helping in determining whether a point pattern displays clustering, randomness, or regularity over a range of scales. The left graph represents a homogeneous PCF, which assumes that the point pattern has a constant intensity across the study region. The observed PCF decreases dramatically at the beginning and then settles closely to the theoretical line, suggesting that any observed clustering or repulsion only occurs at very small distances. The right graph depicts an inhomogeneous PCF, which takes into account varying intensity across the study region. The observed PCF line remains close to the expected PCF line and within the confidence envelope, indicating the point pattern is largely consistent with an inhomogeneous random process.

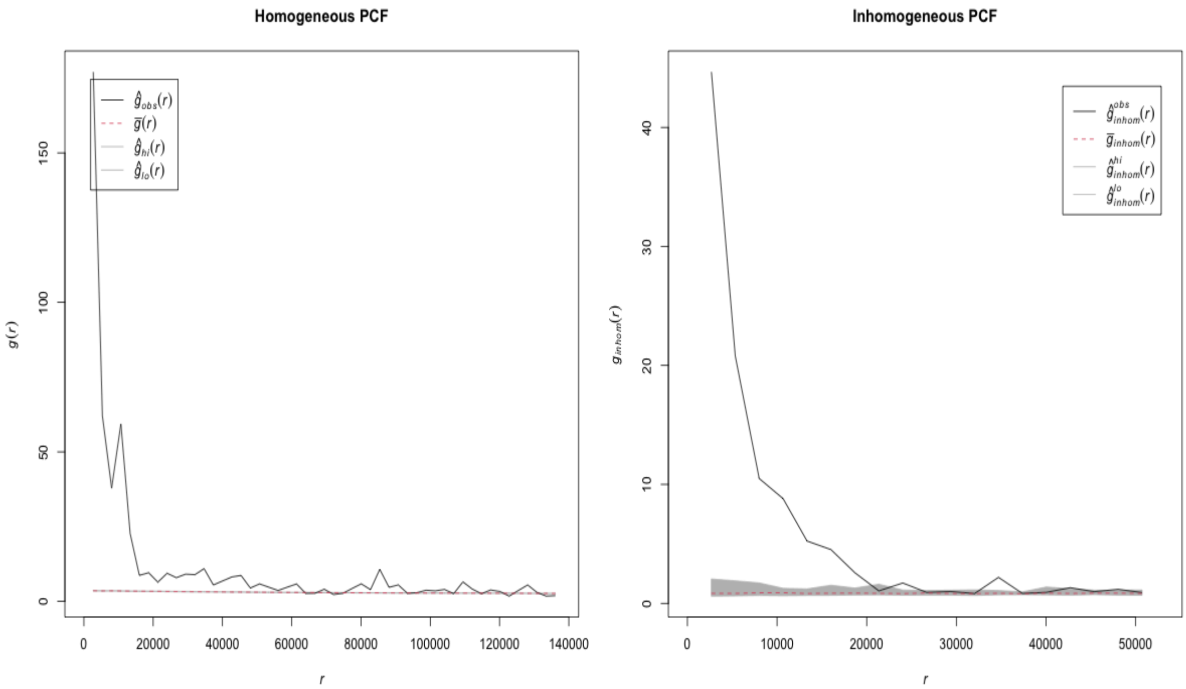

Figure 6: PCF Plot of Homogenous vs. Homogeneous

After evaluating the PCF analysis, we can probably determine that the distribution of birds is inhomogeneous, so we need to build a predictive model based on the effects of environmental factors on bird habitat. We only consider the factors of elevation, forest cover, human footprint index (HFI), and water distribution. The series plots in Figure 7 illustrate the relationship between these factors and bird distribution. The top left plot shows the probability density of observations decreases sharply as elevation increases and then tends toward stability. The top right plot shows the density of observations with respect to forest cover, there is significant variability, with peaks and troughs indicating that certain levels of forest cover are more associated with bird observations than others. The bottom left plot shows the relationship between bird observations and the Human Footprint Index (HFI); it indicates a fairly stable probability density across lower HFI values with a sharp increase at the highest values. The bottom right plot shows that there appears to be little relationship between the distance to water and the density of bird observations, it has a relatively flat line of the observed value, and some spikes randomly appear at certain distances from water bodies.

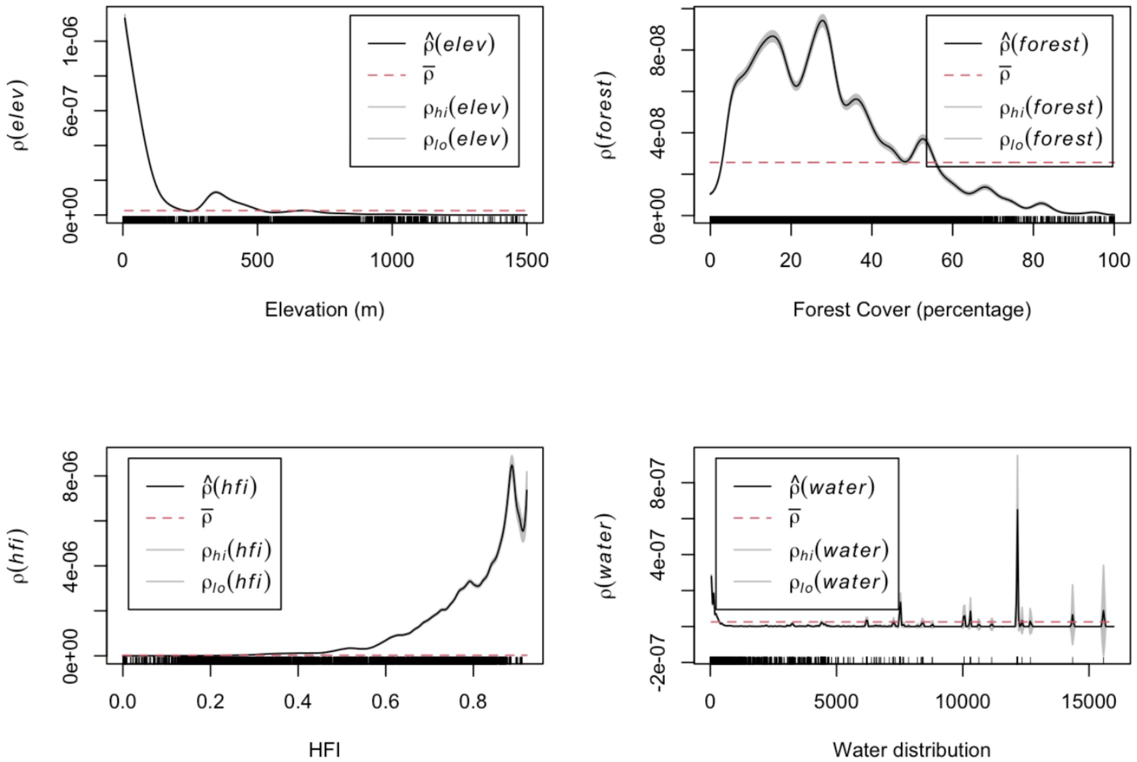

Figure 7: Environmental Factors Plot

The provided model plot in Figure 8 represents the fitted trend of bird observations across British Columbia based on a point process model. The model has been fitted using a combination of environmental variables of forest cover with a linear and a quadratic term, elevation, and the Human Footprint Index (HFI) based on the above collinearity analysis of environmental variables, suggesting that the relationship between bird observations and forest cover is not merely linear, and the variable of elevation and HFI indicate that they are considered as potential factors. In addition, the AIC of the model is 719736, which provides a starting point for model selection.

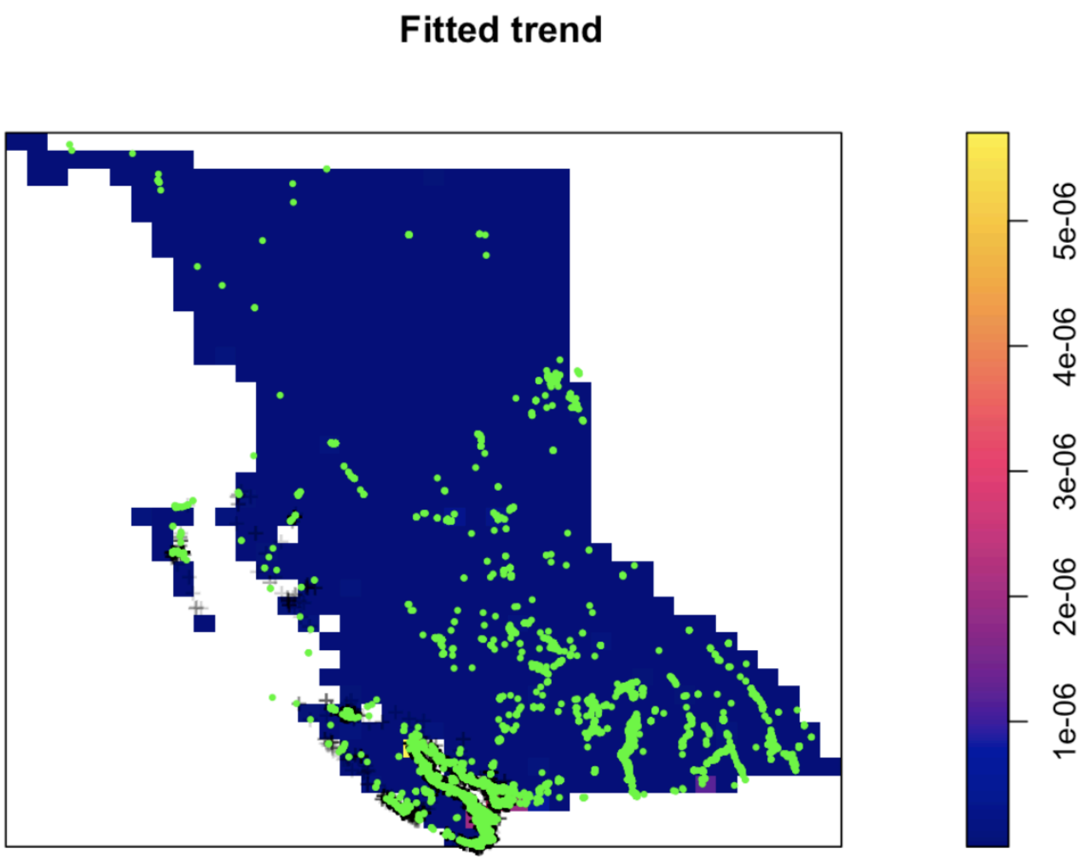

Figure 8: Fitted Predictive Model

Figure 9 shows the partial residual plots for the model in Figure 8. Specifically, the form of the model is:

$$Intensity(\lambda) = exp(\beta_0 + \beta_1 forest + \beta_2 forest^2 + \beta_3 elev + \beta_4 hfi)$$

The series of partial residual plots illustrate the fitted effects of environmental covariates on bird occurrence intensity while holding other variables constant. For forest cover percentage, the plot shows a nonlinear relationship with the intensity of bird occurrences; initially increasing, then plateauing as forest cover approaches 100%. Elevation appears to have a mostly stable effect with a slight peak at higher elevations, indicating an optimal elevation range for these occurrences. The Human Footprint Index (HFI) shows an increasing trend, suggesting that bird occurrence intensity rises with increasing human impact up to a certain point. Each plot includes a solid line for observed effects against the covariate's value, with a dashed line indicating the fitted model effect and a grey shaded area representing the confidence interval around the fitted line. These plots are critical for understanding the individual impact of each environmental factor on bird occurrence patterns within the study area.

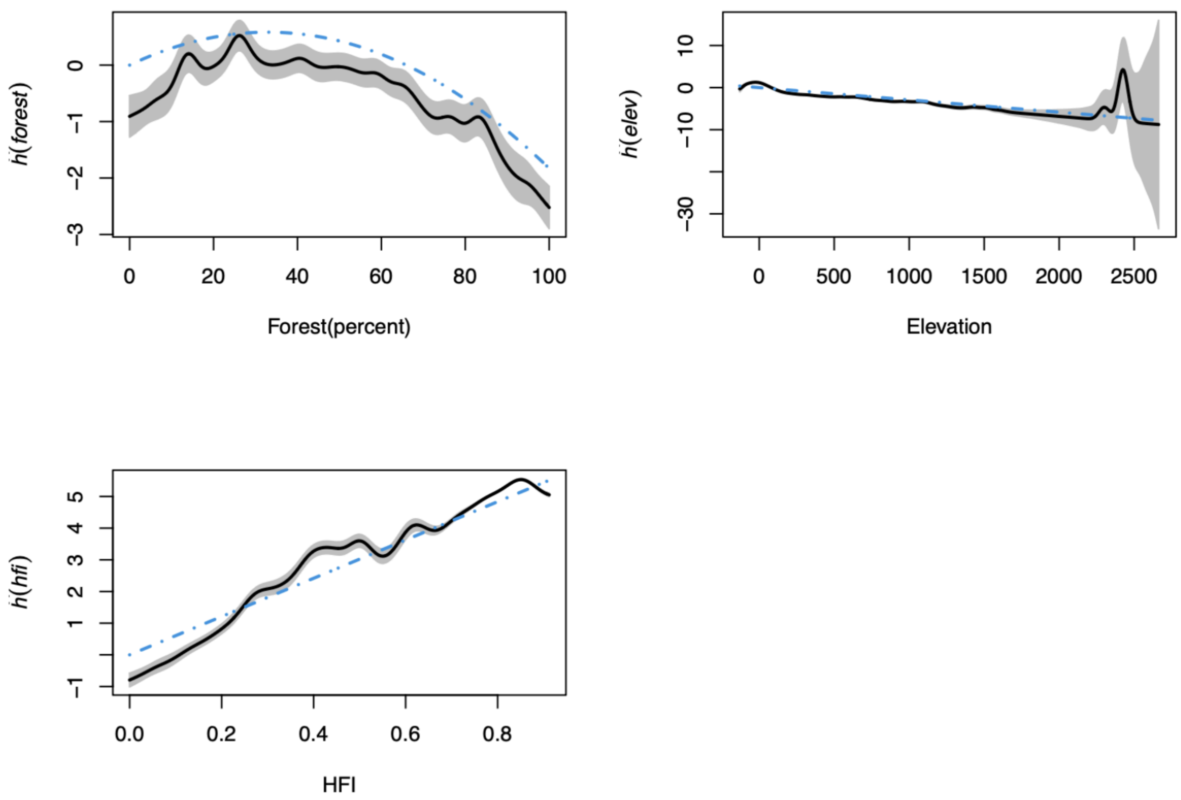

Figure 9: Partial Residual Plots

Figure 10 is a lurking variable plot for the model, which determines whether it’s worth considering water as an additional covariate. In the context of horned grebe, which is waterbird, proximity to water could be a crucial determinant of habitat suitability. Lurking variable plot is created by summing up the residuals for each possible covariate value. It should be around 0 if the covariate is unrelated to trends in the residuals. From the plot, it looks like water is an important variable.

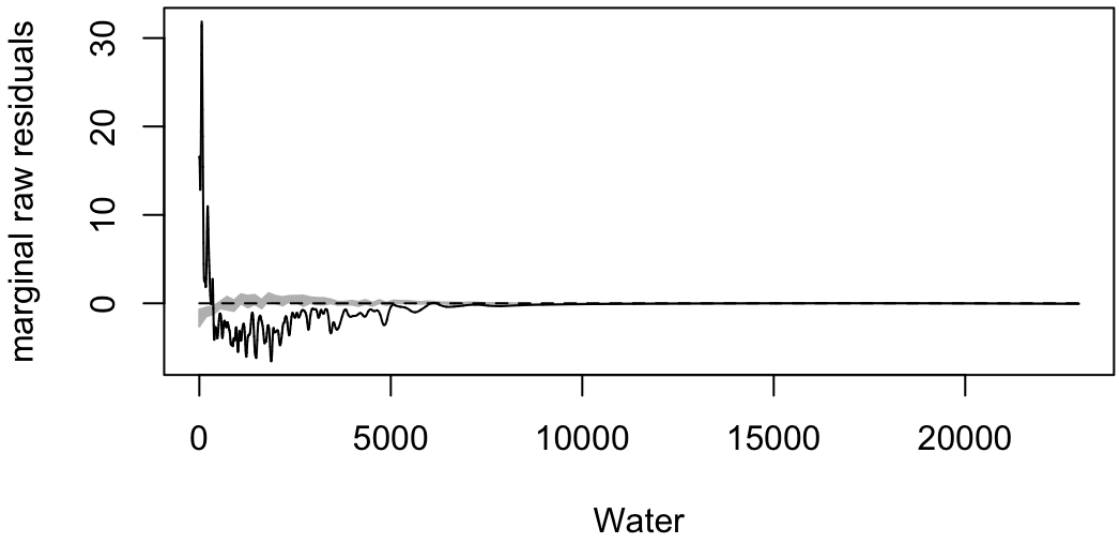

Figure 10: Lurking Variable Plot

By going through the iterative process of model validation and refinement, we obtained the partial residual plot as in Figure 11. We fit the model via `gam()` from the `spline` package, where the AIC of the model was reduced to 683118. However, from these plots we can see that the higher order terms are not capturing the patterns in our data well.

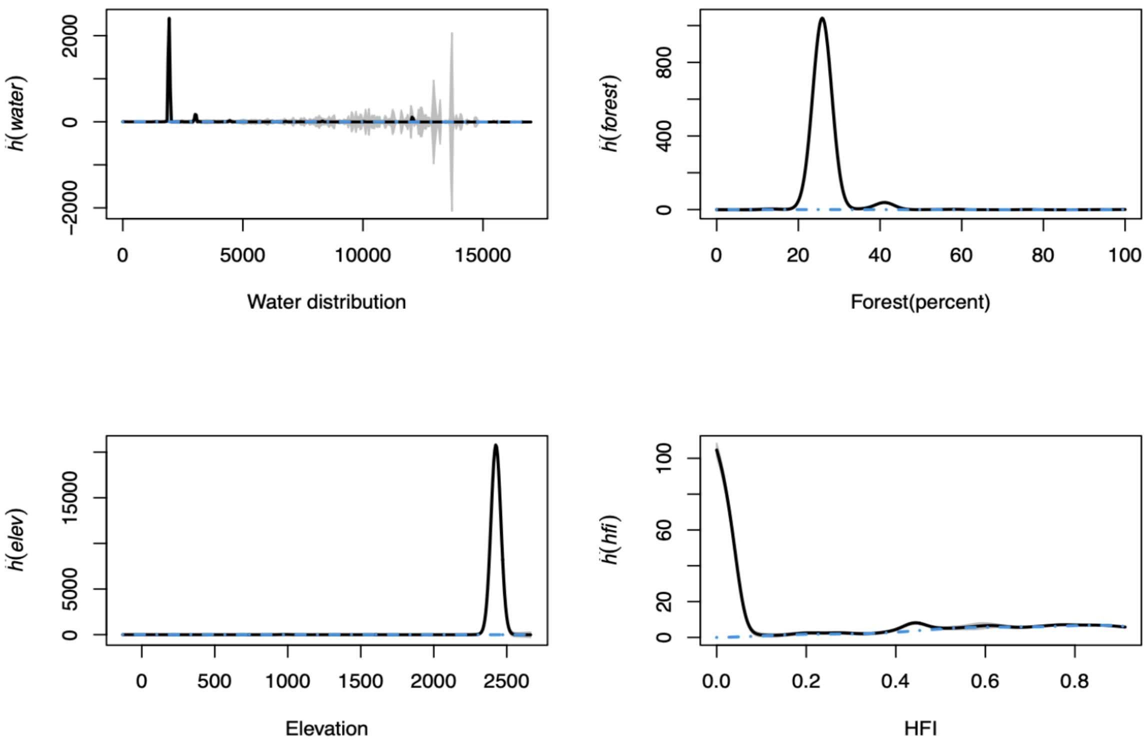

Figure 11: Partial Residual Plots for Smooth Model

Discussion

The analysis of environmental factors affecting the distribution of the Horned Grebe in British Columbia reveals several key insights. Our findings suggest that this species has distinct habitat preferences influenced by elevation, forest cover, human activity, and proximity to water bodies. The decline in observation density with elevation may indicate a preference for lower elevations or reflect the accessibility challenges of conducting surveys in higher-elevation areas. Lower elevations often correspond to wetland habitats, which would be crucial for the breeding habit of Horned Grebe. The observed stability of low observed bird data at higher elevations may also point to the Horned Grebe being less likely to be found in a certain high elevation value, possibly due to lower availability of food resources. However, the variability in the density of observations with forest cover highlights a complex relationship. Peaks in observations several times at certain percentages of forest cover could indicate optimal habitat conditions. The Human Footprint Index (HFI) plot indicates a generally stable observation density across lower HFI values, suggesting that the Horned Grebe may be somewhat tolerant of low-level human activities, a trait that might be beneficial for the species in landscapes altered by human presence. However, the sharp increase in observation density at the highest HFI values is unexpected; it is possible that the preference of birds to live in areas where humans are present may stem from more favorable climatic conditions.

The fitted point process model with a high AIC relative value indicates that while our model has predictive power, it may benefit from further refinement. Moreover, it hints at other unmeasured environmental variables or complex interactions not captured in the current model that could be influencing the distribution of the Horned Grebe. Thus, future research should consider additional environmental variables that might interact with those already studied.

In summary, our findings confirm that the Horned Grebe exhibits distinct preferences for low elevation wetlands, with an optimal range of forest coverage suggesting an intricate balance between open water and vegetative shelter. To enhance habitat conservation for the Horned Grebe, we recommend focusing efforts on regions that exhibit the identified environmental characteristics correlated with higher observation densities and consider more environmental variables to improve the predictive model.

References

GBIF.org (20 April 2024) GBIF Occurrence Download https://doi.org/10.15468/dl.rxnc5q