DATA 589 Final Project

Asio Flammeus (Short-Eared Owl)

Tim Pulfer, Jacob Rosen, Dhun Sheth

Table of Contents

Introduction

The Short-eared Owl (Asio flammeus) is a medium-sized owl known for its distinctive, brown plumage and short ear tufts. This species is one of the world's most widely distributed owls. It can be spotted year round in southern parts of British Columbia (BC) and typically flies north to breed during the spring months [1]. Its preferred habitats for breeding are dry marshes or tall grass meadows since it builds nests on the ground [2]. Generally, it favours grasslands, old fields, hay meadows, pastures and marshes as its habitat. They are one of the most often spotted owl species because they are known to hunt during early morning or late evening as opposed to being purely nocturnal [3]. When hunting, they fly low over the ground in search of small rodents.

The data that is being used for the Short-eared owl location was captured 97% by human observations, 1.6% from preserved specimens, and 0.9% from machine observations. This data was collected globally with a total of 415,632 sightings, 34,291 of which are in Canada. However, it is important to note that each observed sighing may include more than one bird. There are 7787 observations in BC, 319 of which were discarded because they were outside of the window.

There are three scientific questions that the analysis attempts to answer.

1. What elevation are the Short-eared Owl typically found at?
   Hypothesis: Considering the short-eared owl is known to prefer dry marshes, grasslands, old fields, and hay meadows, a greater intensity of owl sightings is expected at lower elevations.

2. Will the intensity of Short-eared Owl sightings increase with proximity to water?
   Hypothesis: Considering the short-eared owl prefers dry marshes and grasslands, a greater intensity at dryer locations is suspected, rather than locations closer to water.

3. Is there a relationship of Short-eared Owl sightings to other covariates?
   Hypothesis: The covariate data also contains forest coverage among other covariates of which some might express some relationship with short-eared owl sightings.

Methods

Variables

Applied Variables		
	Type	**Description**
decimalLatitude	Float	Latitude of owl sighting
decimalLongitude	Float	Longitude of owl sighting

The dataset contained over 40 descriptor columns, many of which contained the same value for each sighting. For the analysis, the only required values were two descriptors, the latitude and longitude. With these two descriptors, we were able to plot the owl's locations in the BC window, this allowed us to visualize the owl's locations versus other geographical descriptors such as elevation and distance from water. Additionally, the owls location plotted allowed us to check the density of the owls in BC and check for covariate relationships.

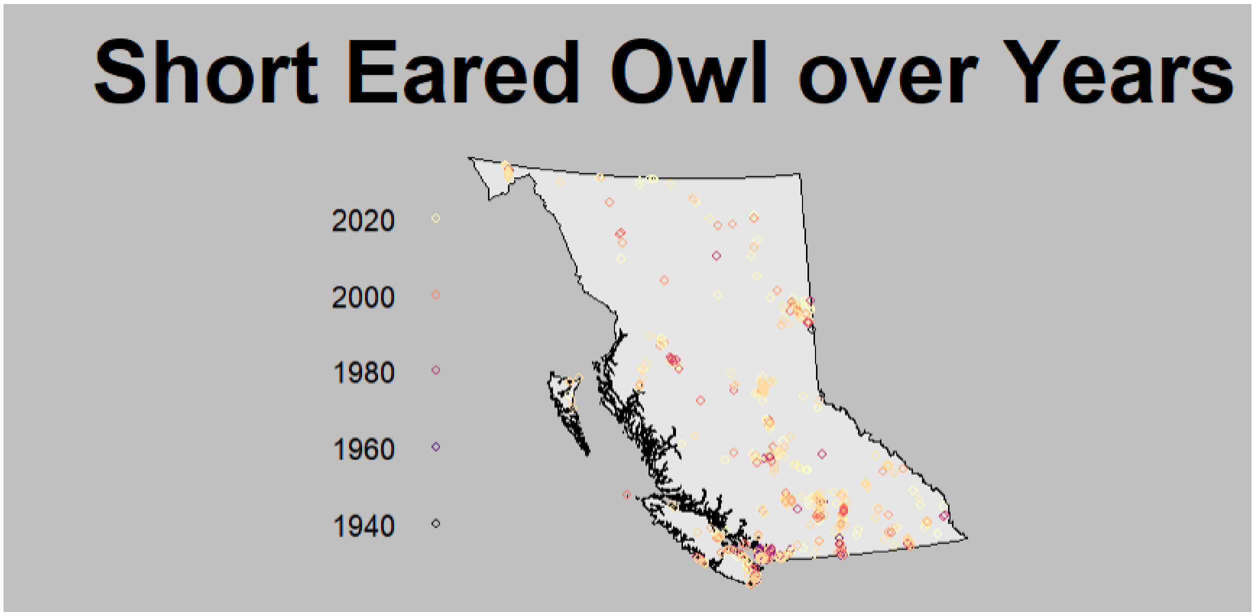
Exploratory
The data was first plotted to get a sense of the general distribution of owl sightings across BC. After, owl sightings were compared against elevation across BC and distance from bodies of water to better understand the relationship between owl sightings and covariates.

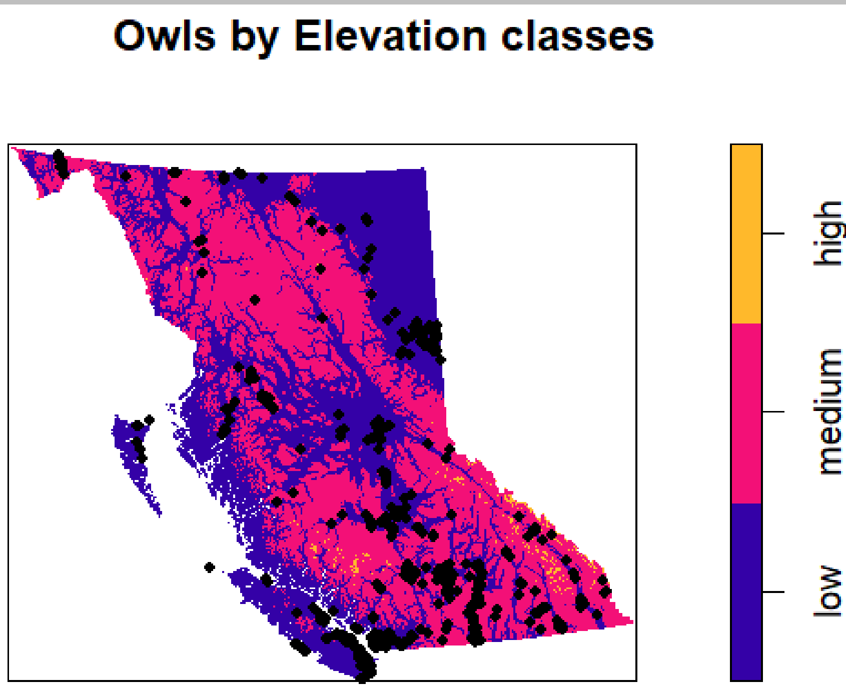
Homogeneity
The homogeneity and spatial distribution of owl sightings were analyzed through density estimation, comparing the density of BC elevation to owl elevation, and applying a quadrat test to conclude if owl intensity is homogeneous or inhomogeneous across BC.

Covariate Relationship
The relationship between owl sightings and elevation, distance from water, and forest coverage was evaluated to determine the general form of the model.

Model Building
Based on the covariate relationship, a general form was assumed. A quadrat test was used to determine if the model predictions were deviating significantly, if so, a more complex model using GAMs could be implemented. AIC difference and a likelihood ratio test would be used to compare the various models to select the best one. Once a model is selected, owl intensity can be predicted across BC.

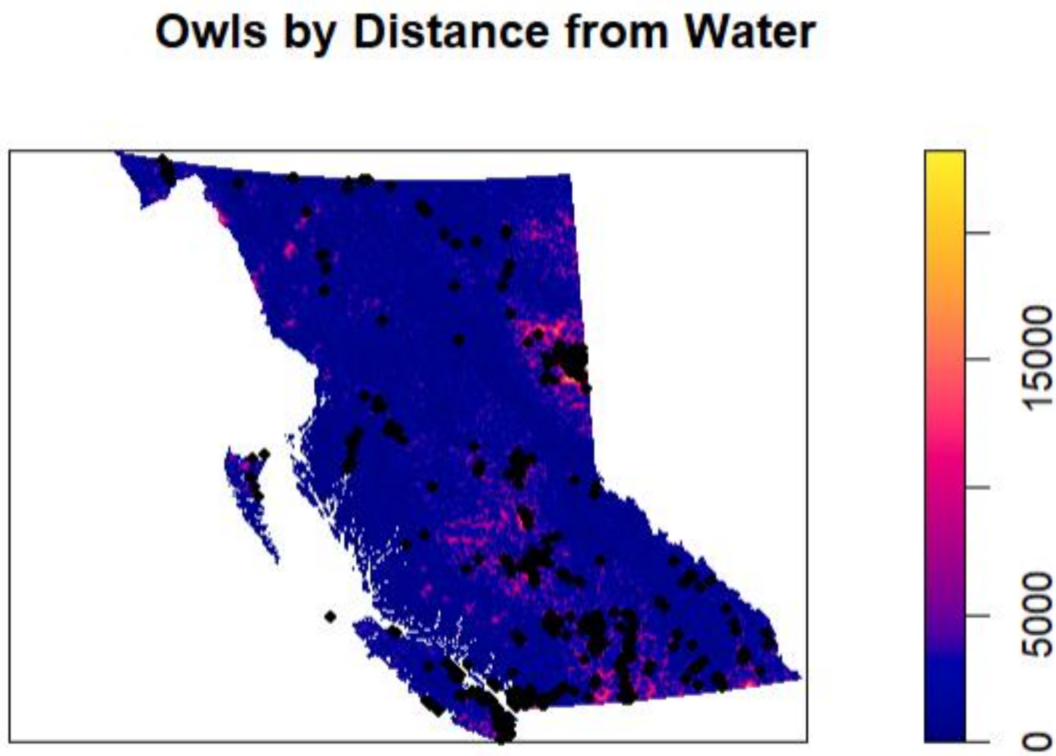
Results



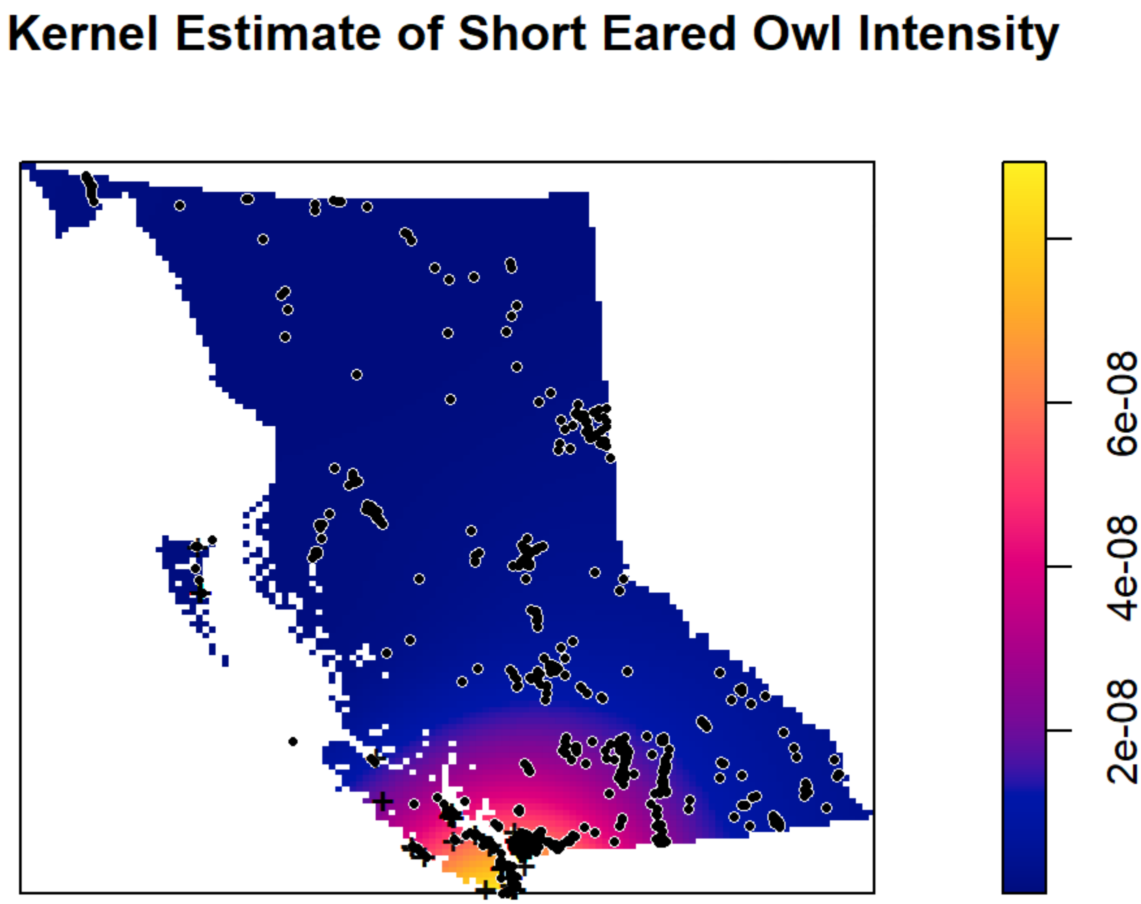


low	medium	high
7272	46	0

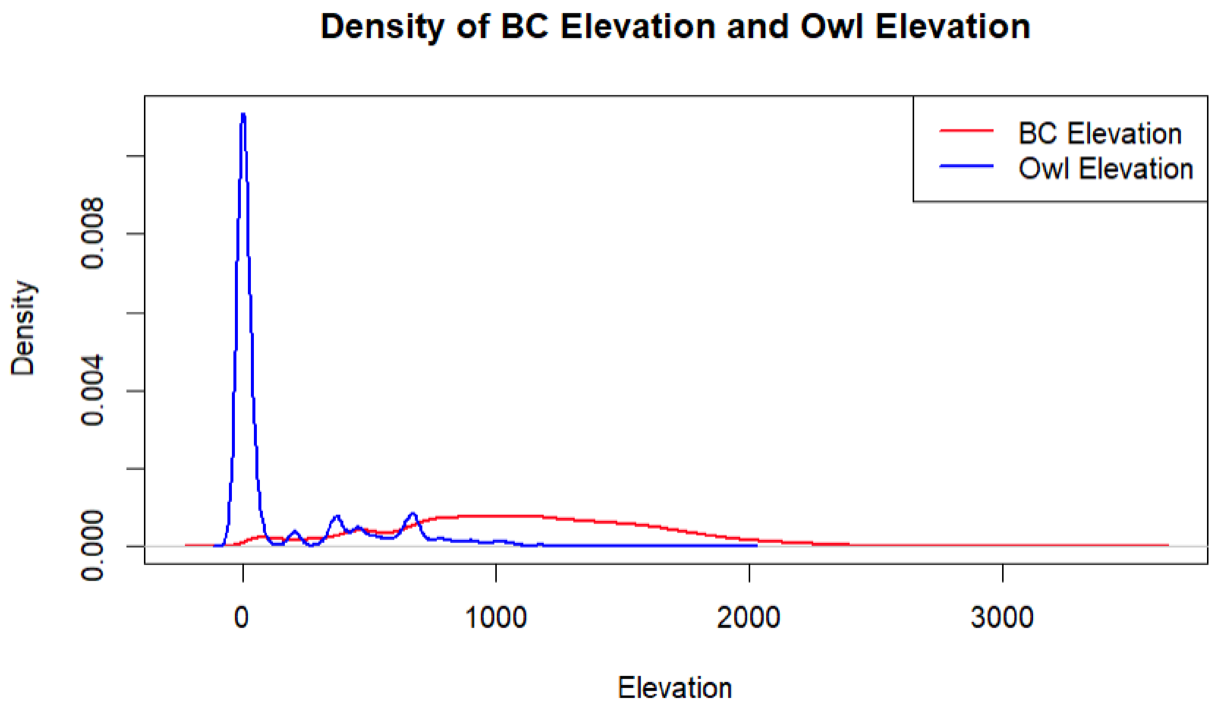
Can see from the above plot that the owls tend to stay in regions of low elevations.



Based on the above plot, can see owls in regions that are close and medium distances from water, but again these correspond to areas of lower elevation.

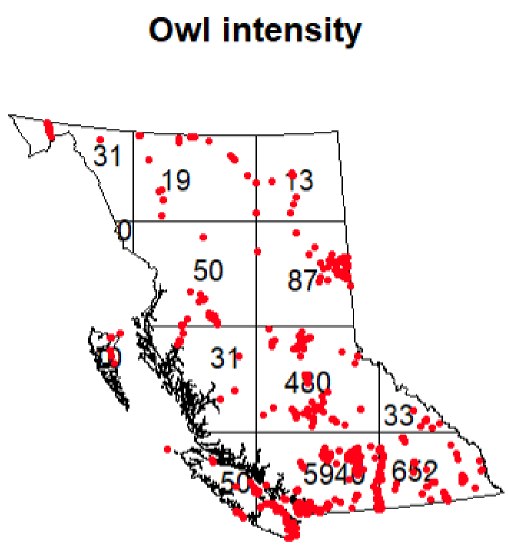


From the above plot, can see the intensity of Short-eared owls is not homogeneous across BC and seems to be more concentrated in the southern parts.



Observing the density of Owls to the density of BC elevation in general, it can be seen that the density of Owl elevations is highest near 0 whereas the density of elevation of BC is highest near 1000. Owl density falls very fast as elevation increases. This indicates the spatial distribution of owls is not random with respect to elevation, if it were random, there would be more overlap between the 2 densities.

Quadrant Plot

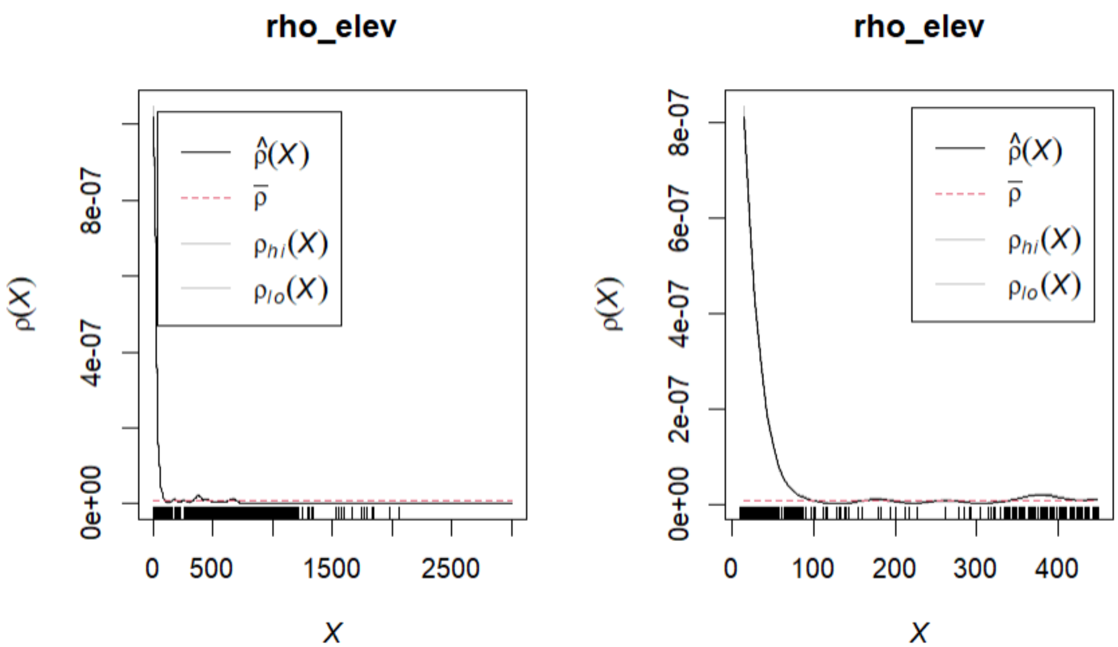


```
Q <- quadratcount(owls_ppp,
                  nx = 4,
                  ny = 4)
quadrat.test(Q)
```

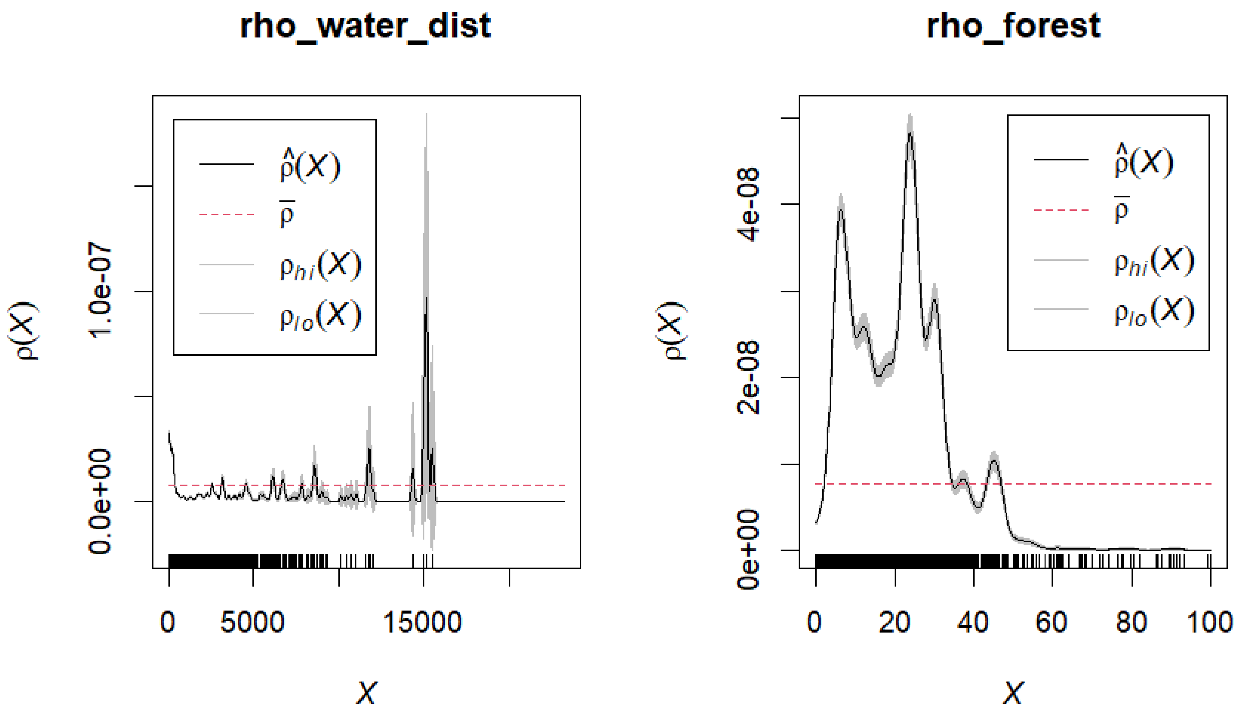
```
##
##  Chi-squared test of CSR using quadrat counts
##
## data:
## X2 = 37405, df = 12, p-value < 2.2e-16
## alternative hypothesis: two.sided
##
## Quadrats: 13 tiles (irregular windows)
```

Based on the above quadrat test, the small p-value suggests the null hypothesis should be rejected, ie. there is significant evidence against the assumption of homogeneity, therefore, the intensity, if calculated, under the assumption of homogeneity would be inaccurate. In addition, the plot of quadrat counts also shows the intensity is not homogeneous.

Plot of Owl Intensity w.r.t Elevation



Owl Intensity w.r.t Distance from Water and Forest Coverage

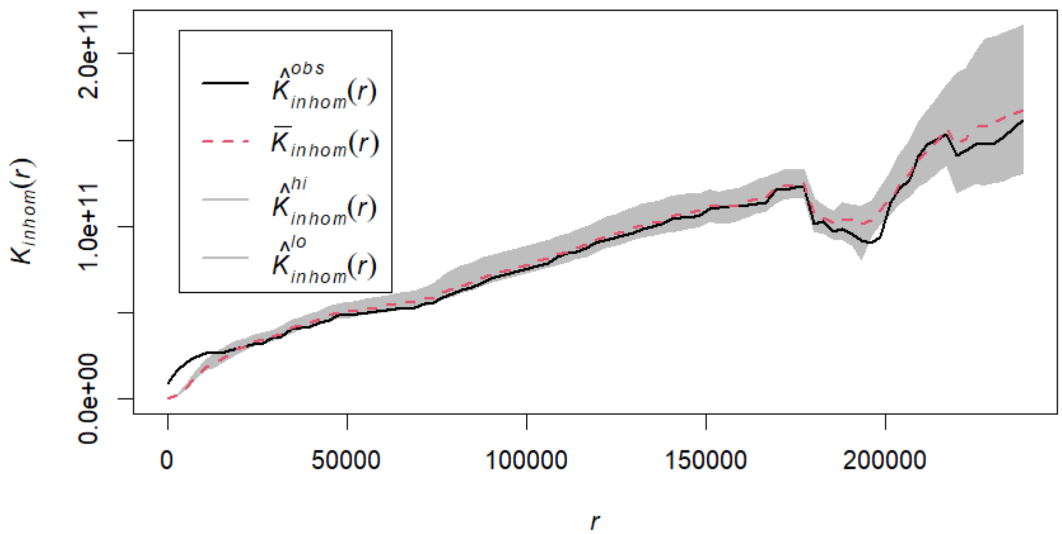


Based on the above plots, there seems to be some sort of relationship between owl intensity and elevation up to 100m and then hovers around the mean. This falls in line with prior expectations because owls prefer low-elevation regions.

Similarly, based on the distance from water, there seems no relationship between owl intensity and distance from water.

Finally, there does seem to be a non-linear relationship between owl intensity and the percent of forest coverage, where owl intensity is high for forest coverage between 0 to 40%.

Correlation between Owl Sightings



The above plot suggests there is some correlation between owl sightings for a radius of less than 5000 meters, after which there is no significant clustering results.

Fitting a Model for Owl Intensity

$$\lambda(u) = e^{intercept + Elevation(u) + Elevation(u)^2 + Forest(u) + Forest(u)^2}$$

$$AIC = 237555.9$$

```
        Chi-squared test of fitted Poisson model 'fit' using quadrat counts

data:  data from fit
X2 = 11583, df = 2, p-value < 2.2e-16
alternative hypothesis: two.sided

Quadrats: 7 tiles (irregular windows)
```

The quadrat test results in a very small p-value, which indicates a significant deviation from the model's predictions. As an improvement, rather than using a quadratic model, a GAM could be used to fit instead, to better handle a more non-linear relationship.

Fitting using a GAM:

```
fit_smooth <- ppm(owls_ppp ~ bs(Elevation,8) + bs(Forest, 8), data = DATA, use.gam = TRUE)
```

$$AIC = 229532$$

$$AIC\ Difference = 8023.84$$

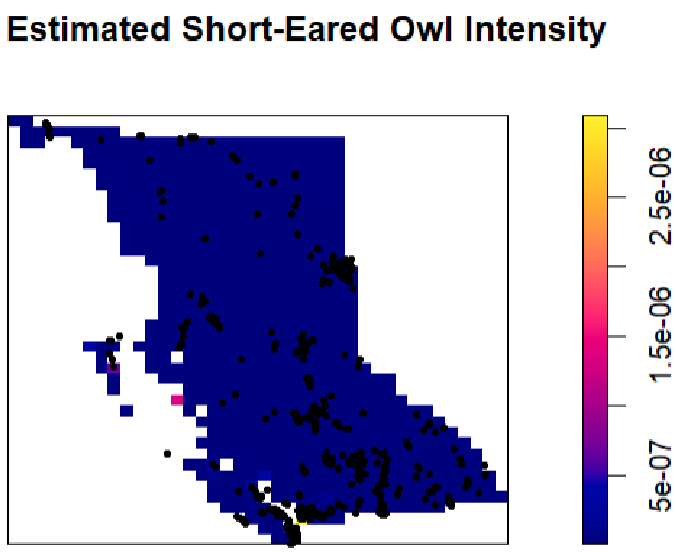
Likelihood Ratio Test

```
Analysis of Deviance Table

Model 1: ~Elevation + I(Elevation^2) + Forest + I(Forest^2)      Poisson
Model 2: ~bs(Elevation, 8) + bs(Forest, 8)       Poisson
  Npar Df Deviance  Pr(>Chi)
1    5
2   17 12   8047.8 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Based on the difference in AIC and the likelihood ratio test, the more complex model, that was fit using GAMs, was chosen over the quadratic model.

Visualizing Model Predictions vs. Observations



Discussion

Initially, when the owl sightings are plotted on the BC window, the spatial density across BC seems somewhat random. However, upon further investigation, after the owl sightings are plotted against the elevation across BC, a clear pattern emerges, in which a large majority of owl sightings are found in low-elevation regions. This helps answer the first scientific question that was proposed and falls in line with the initial hypothesis.

A similar plot of owl sightings vs. distance from bodies of water was made. Although some clusters were observed where the distance from water was medium, many sightings were close to bodies of water. Due to the range of distances from water primarily being small, sometimes it was visually difficult to identify the medium to high areas but a lot of owl sightings overlap with the pink/yellow regions, corresponding to medium or a far distance from water, which again was expected based on the species preferences.

In addition, observing the density of BC elevation compared to the elevation at which owls were seen, comparing the densities, there was clear evidence that the spatial distribution of short-eared owls with respect to elevation is not random.

Although from the initial plot, homogeneity seems reasonable to assume, doing a quadrat count test yielded a very small p-value, offering significant evidence against the assumption of homogeneity, therefore, the intensity, if calculated, under the assumption of homogeneity would be very inaccurate.

Plotting owl intensity against elevation, distance from water, and forest coverage, there was a clear relationship between elevation and forest coverage, whereas the distance from bodies of water seemed more random. This helped answer the third scientific question, indicating the short-eared owl sightings may also depend on forest coverage, in addition to elevation.

Before fitting a model to predict owl intensity, the correlation between owl sightings themselves was tested. The results suggested there was some correlation between owl sightings within a radius of 5000 meters and could suggest clustering.

When fitting a model, initially a quadratic model using elevation and forest coverage was implemented. However, distance from bodies of water was dropped due to the randomness observed in the owl intensity vs. distance from water plot as mentioned before. The quadrat test indicated significant deviations from the model's predictions suggesting the model could be improved further.

A more complex model using GAMs was fit to better handle the non-linear relationship between owl sighting intensity and elevation/forest coverage. Comparing the AIC values and performing a likelihood ratio test, both favored the more complex model over the initial quadratic form.

Lastly, visualizing the model predictions versus the observations showed an almost identical prediction of owl intensity across BC compared to the observations. This shows the model is predicting some average intensity across all of BC. One reason for this could be missing covariates that would better explain owl locations rather than elevation and forest coverage, such as temperature. Another reason could be the owl observations span several years which could lead to autocorrelation, which was not accounted for in the spatial analysis done above. Although future improvements could be made, the current state offers a reasonable model to predict short-eared owl sighting intensity based on elevation and forest coverage.

References

1. British Columbia Ministry of Environment. (15 April 2024). Short-eared owl. Retrieved from https://www.env.gov.bc.ca/okanagan/esd/atlas/species/short_eared.html

2. Cornell Lab of Ornithology. (13 April 2024). Short-eared owl overview. All About Birds. Retrieved from https://www.allaboutbirds.org/guide/Short-eared_Owl/overview

3. National Audubon Society. (13 April 2024). Short-eared owl. Audubon Field Guide. Retrieved from https://www.audubon.org/field-guide/bird/short-eared-owl

Dataset: GBIF.org (13 April 2024) GBIF Occurrence Download https://doi.org/10.15468/dl.44x7ax

GitHub repository for the project: https://github.com/Dhunsheth/short-eared-owl