

589 Project Report

Henny Luo, Renghe Tang, Weijia Lyu

2023-04-30

Introduction

The current project attempts to explore the spatial distribution of coyotes in British Columbia (BC), Canada, and its potential interaction with environmental covariates. The scope of the study covers the entire province of BC, which is known to have a significant population of coyotes. The project explores the relationship between coyote populations and four environmental covariates within the province: elevation, forest cover, distance to water, and human footprint index (HFI). Characteristics of the spatial distribution of coyote occurrences and their potential relationships with these local environmental covariates are explored with various spatial data analysis techniques. The coyote can be identified by its rough, grey, or reddish-grey fur with a lighter buff-colored underbelly, along with long rusty or yellowish legs and a thick, bushy tail. They have notable yellow eyes and prominent ears and weigh between 20 to 40 pounds on average. Historically, coyotes were restricted to living in prairies and arid regions of the western part of the United States, but with the arrival of settlers who changed the landscape and eradicated wolves, coyotes found new opportunities to expand their range. They are now widespread throughout the Western Hemisphere, from the Pacific to the Atlantic oceans (Lariviere, 2023). According to Furbeaer Management Guidelines published by BC Government, the earliest occurrences of coyotes in BC were reported in the 1930s. In the recent decade, they rapidly expanded within the province, including in urban areas. Coyotes are generally not found around coastal slopes and islands in the province. The current project attempts to further explore the distribution of coyotes within BC and their potential relationships with the environmental covariates of interest to provide insight into the management and conservation of coyotes in BC. Various spatial data analysis techniques implemented by spatial data analysis packages in R are employed to investigate the relationships between the coyote occurrence locations and how environmental covariates influence the spatial distribution of coyotes.

Method

Data

The occurrences data of coyotes is obtained from the Global Biodiversity Information Facility. GBIF is an online data infrastructure founded with international governmental efforts to provide free data on different species on Earth to facilitate research and the conservation of biodiversity. A dataset containing 63,537 occurrences of coyote (*Canis latrans* Say, 1823) across North America is obtained from the Global Biodiversity Information Facility (GBIF.org, 2023). The coordinates are extracted for analysis of the spatial distribution of coyotes. The “BC Covariate” dataset provided for the current project contains four environmental covariates, including elevation, forest cover, Human Footprint Index, and distance to water in British Columbia, Canada.

Data processing

The coordinates of the coyote occurrences are transformed to Albers projection to match the projection method employed by the sampling window bounding the borders of BC using the R package `sp`. The resulting coordinates data is then combined with the BC covariate dataset and converted to a `ppp` (Planar Point Pattern) object using the `spatstat` package. Coyote occurrences within BC are filtered with the help of the sampling window. The final dataset for analysis contains 960 coyote occurrences.

Analysis

Visualization

The spatial distribution of coyotes in BC is visualized for preliminary visual assessment by plotting the locations of each occurrence within the sampling window outlining the border of BC. The distance between each point and its nearest neighbor is calculated and plotted with the distances reflected as the size of each point to identify the most isolated occurrences.

Test for homogeneity

The quadrat test is conducted to examine the homogeneity of the distribution of coyote occurrences. The sampling window is partitioned into 8 quadrats, and the observed number of occurrences in each quadrat is compared with the expected number of occurrences in each quadrat, assuming homogeneity. The result suggests inhomogeneity.

Intensity Analysis

The intensity of coyote occurrences across BC is estimated by kernel density estimation. Two kernel density estimations are performed using different bandwidth selection methods. The first intensity estimation employs the default Diggle's method. The second one uses the likelihood cross-validation method. The results are plotted for visual assessment. Scan LRT test is performed to identify areas of elevated intensity (hot spots). A radius of 60000 is specified for the test to aid visualization of the likelihood ratio as the estimated radius with likelihood CV bandwidth selection resulted in hot spots too small for the scale of the provided sampling window. The local p-values of the Scan test using estimated radius are visualized as it is easier to read when the radius is small. Relationship between occurrence locations Ripley's K-function is calculated to access the spatial relationship between the occurrences. The calculation is corrected for inhomogeneity.

Analysis with Covariates

Quadrat count tests are conducted for each of the four covariates to see if occurrences are evenly distributed across different levels of the covariates. The four quantiles of each covariate are calculated, and the sampling window is divided and mapped to four levels, low, medium, high, and very high, based on the value of the covariate. The number of occurrences in each level of the covariate is then counted and compared to the expected number of occurrences if the distribution was even. The intensity of the coyote occurrences as a function of each of the covariates is estimated non-parametrically. The estimated intensity, given the values of the covariate, is assessed for patterns that deviate from the constant one, which would suggest that there is a relationship between the covariate and the intensity of the coyote occurrences. Potential collinearity between the covariates is examined by computing their correlation matrix.

Poisson Point Model

The Poisson Point Process model is used to model the relationship between the intensity of the coyote occurrences and the covariates. Since the point process is inhomogeneous, a function form of $\lambda(u) = e^{\alpha + \beta_1 Z_1(u) + \beta_2 Z_2(u) + \dots + \beta_i Z_i(u)}$ is assumed. All four covariates, elevation, forest cover, HFI, and distance to water, are used to fit the model. Quadratic terms are added to model the non-linear relationships between the intensity and the covariates. A simplified model is fitted, removing the quadratic term for distance to water. AIC for the full and the reduced models are both calculated and compared. A Likelihood Ratio Test is performed to determine if the simplification is justified. The reduced model is validated by a quadrat count test. The intensity predicted by the model is visualized and mapped with the observed data to visually assess model fit. The residuals are also plotted for validation. Partial residuals of each of the covariates are plotted to examine if the model captures the relationship between the individual covariate and the intensity. A model with higher-order polynomials is fitted with the data in an attempt to better capture the complexity of the relationships between intensity and covariates. The new model includes up to the 7th power of elevation and 6th power of forest cover. The new model is evaluated by visualizing the residuals and partial residuals for each covariate.

Results

visualized the distribution of coyote across BC

preliminary visual assessment of spatial distribution

From figure 1, we know that coyotes are mainly concentrated in the central and southern parts of British Columbia, with the highest concentration being along the lower boundary of BC, while only a few scattered populations are found in the upper part. Therefore, we tentatively conclude that the distribution of coyotes in BC is inhomogeneous.

```
##
##  very low      low      medium      high very high
##      657      268       31        1         0
```

We divided the elevation into 5 levels, from ‘very low’ to ‘very high’. From figure 2, we can see that coyotes are mainly distributed in areas with ‘very low’ and ‘low’ elevations, with a small number of them found in areas with ‘medium’ elevations, while there are almost no coyotes distributed in areas with ‘high’ and above elevations. This suggests that elevation may be related to the distribution of coyotes.

```
##
##  very low      low      medium      high very high
##      468      191      111       94       75
```

We then plotted the forest cover and coyote distribution in Figure 3. From Figure 3, we observed that coyotes are mainly distributed in areas with ‘very high’ forest cover, while there are almost no coyotes in areas with ‘high’ and above forest cover. This suggests that there may be a certain relationship between the distribution of coyotes and forest cover.

```
##
##  very close      close      medium      far      very far
##      915         39         5         0         0
```

Coyote in BC

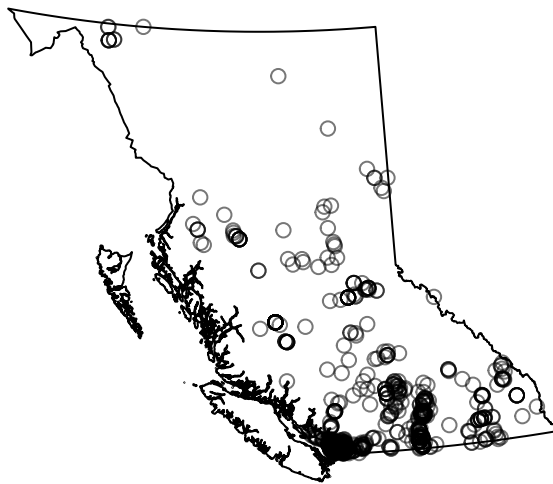


Figure 1: Distribution of Coyote in BC

Elevation classes

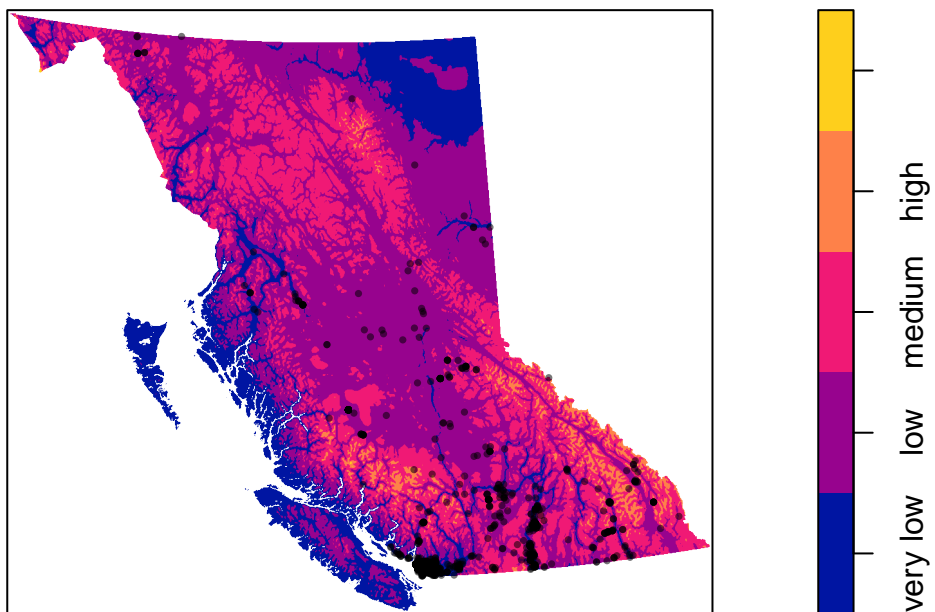


Figure 2: Elevation Classes and Coyote Overlay Plot in BC

Forest classes

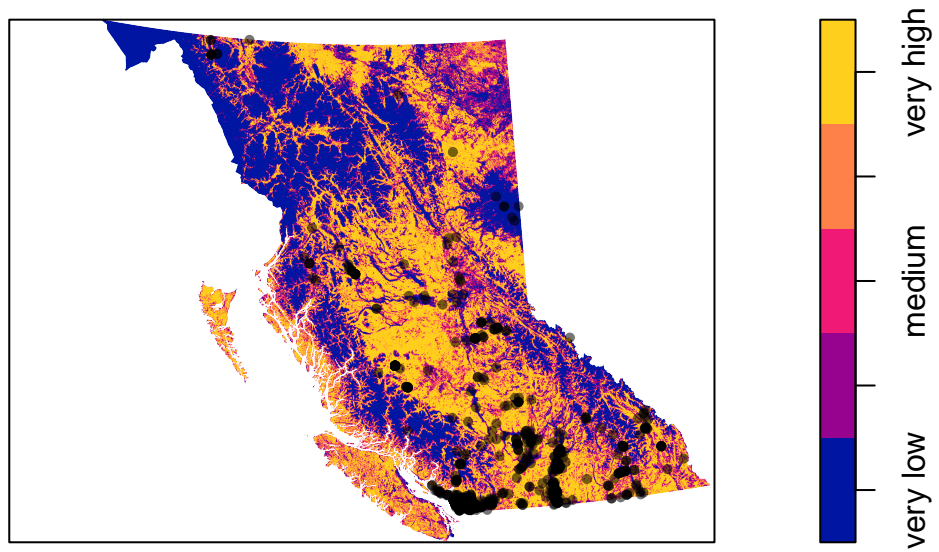


Figure 3: Forest Cover Classes and Coyote Overlay Plot in BC

Distance to Water classes

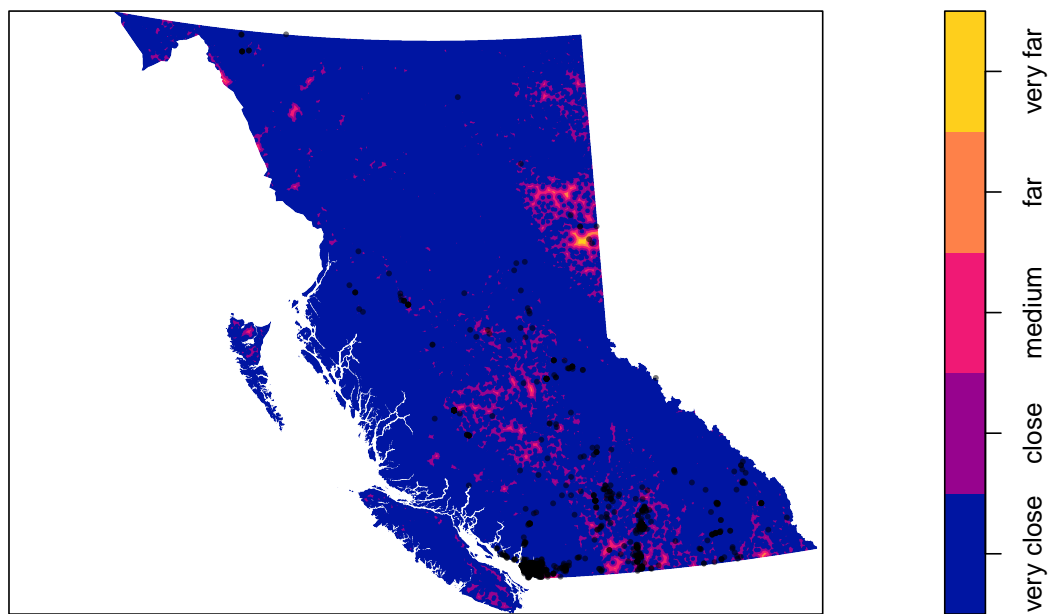


Figure 4: Distance of Water Classes and Coyote Overlay Plot in BC

We then plotted the distance to water sources and coyote distribution in Figure 4. As shown in the figure, coyotes are mainly distributed in areas close to water sources, with very few found in areas far from water sources. Therefore, we speculate that the distance to water sources may be related to the distribution of coyotes.

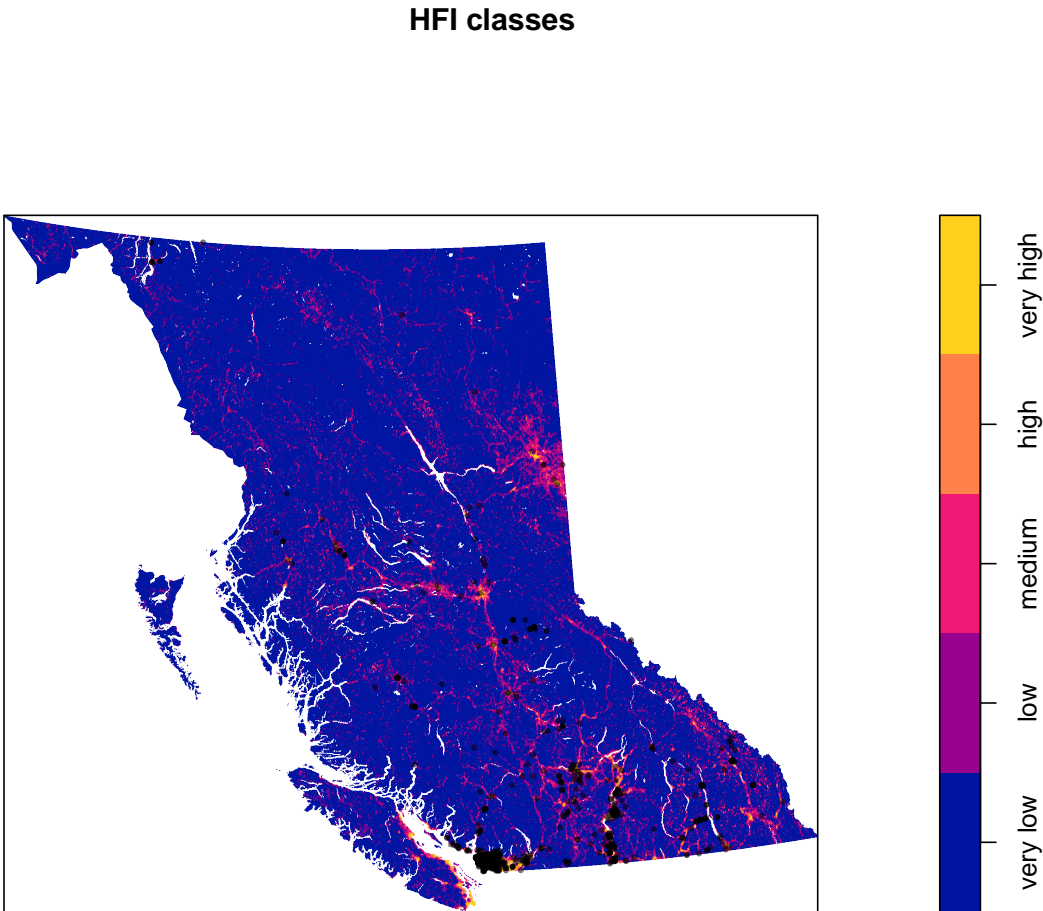


Figure 5: HFI Classes and Coyote Overlay Plot in BC

##	very low	low	medium	high	very high
##	915	39	5	0	0

We then plotted the Human Footprint Index (HFI) and coyote distribution in Figure X. From the figure 5, we can see that coyotes are mainly distributed in areas with higher HFI values, which indicates greater

human impact on those areas. This suggests that there may be a certain relationship between HFI and coyote distribution.

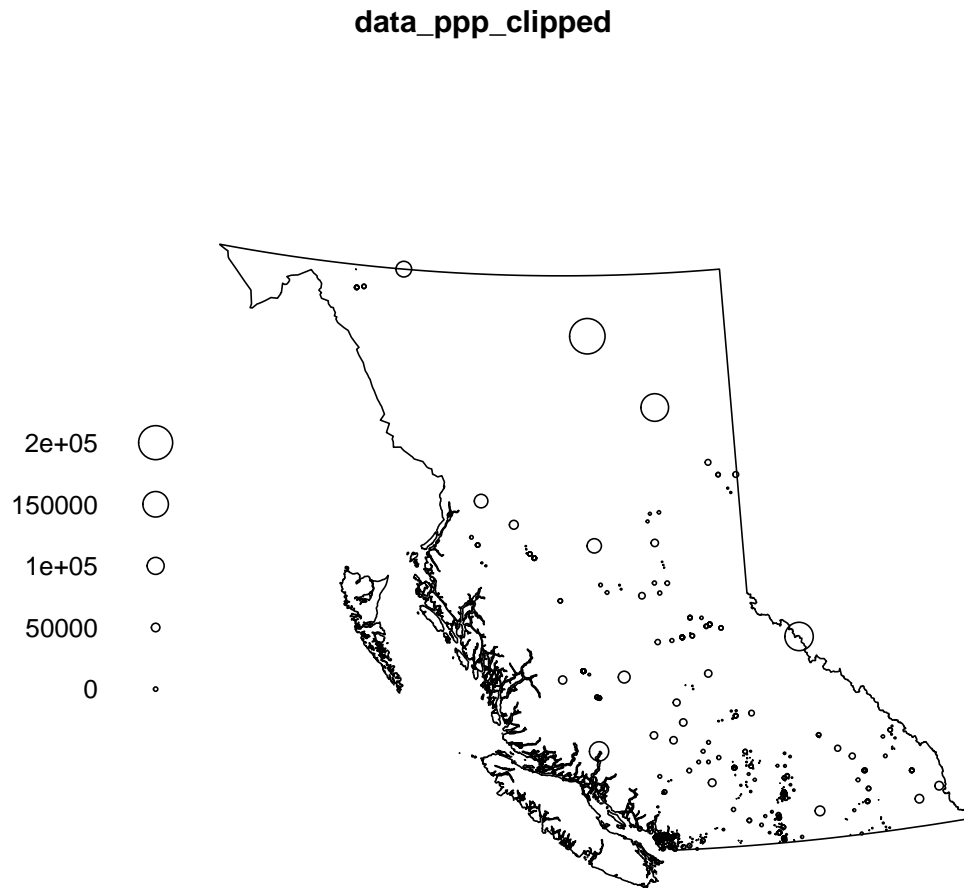


Figure 6: Nearest Coyotes Overlay Plot in BC

We then plotted the nearest neighbor distances between each coyote and its closest neighbor. We found that coyotes in the lower part of British Columbia are relatively close to each other, while those in the upper part are further apart. The results are shown in figure 6.

We have tentatively concluded that all four variables may be related to the distribution of coyotes, but relying solely on visual inspection is not enough.

Intensity Analysis

Next, we used the `quadratcount` function to divide the study area into eight small rectangles, and plotted the intensity distribution map, followed by `quadrat.test` analysis.

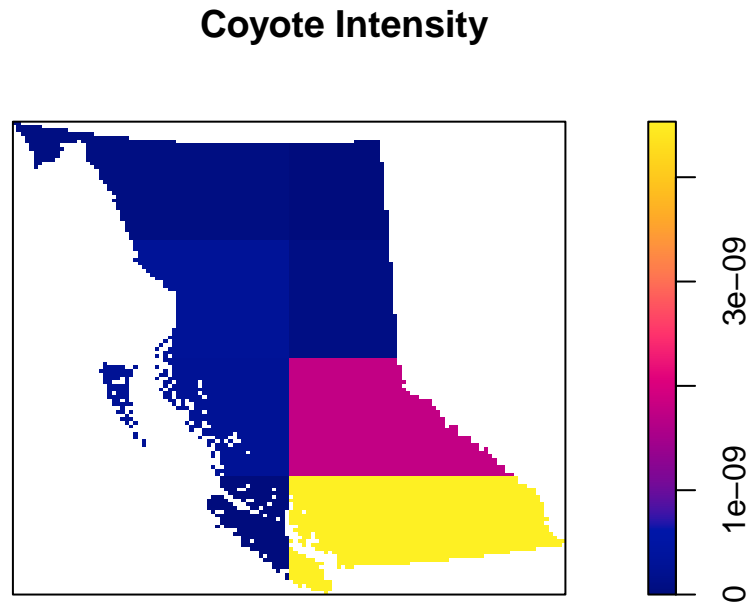


Figure 7: Intensity of Coyote

The intensity map is shown in Figure 7, and it is clear that the distribution of coyotes is inhomogeneous. We then used a χ^2 test to test for significant deviation from CSR.

```
##  
## Chi-squared test of CSR using quadrat counts  
##  
## data:  
## X2 = 2722.6, df = 7, p-value < 2.2e-16  
## alternative hypothesis: two.sided  
##  
## Quadrats: 8 tiles (irregular windows)
```

The results of the test from the `quadrat.test` function showed a p-value much smaller than 0.05, indicating that the distribution of coyotes in BC is inhomogeneous.

Then we perform density estimation of the intensity (`lambda`) with default bandwidth selection. From Figure 8, it can be observed that coyotes are mainly distributed in the lower half of BC.

We then used likelihood cross-validation bandwidth selection. As shown in Figure 9, coyotes are more concentrated in the Vancouver and Kelowna areas.

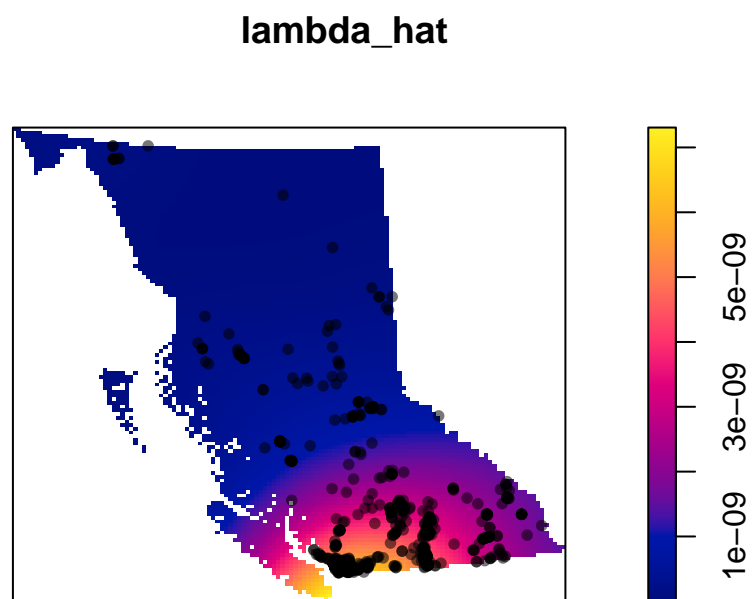


Figure 8: Kernel Density Estimate

Likelihood CV Bandwidth Selection

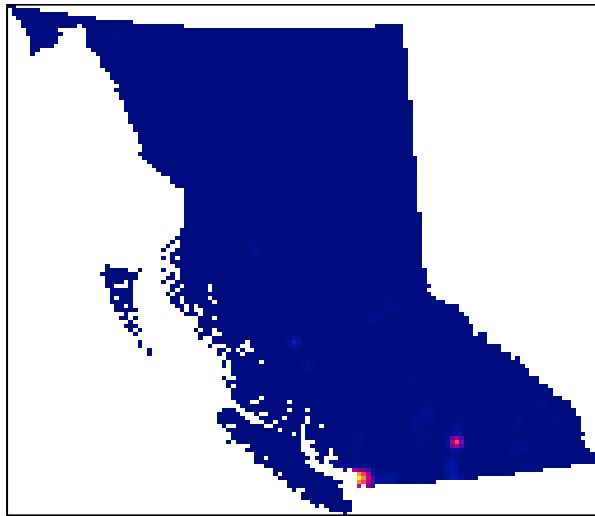


Figure 9: Kernel Density Estimate with Likelihood CV Bandwidth Selection

LR

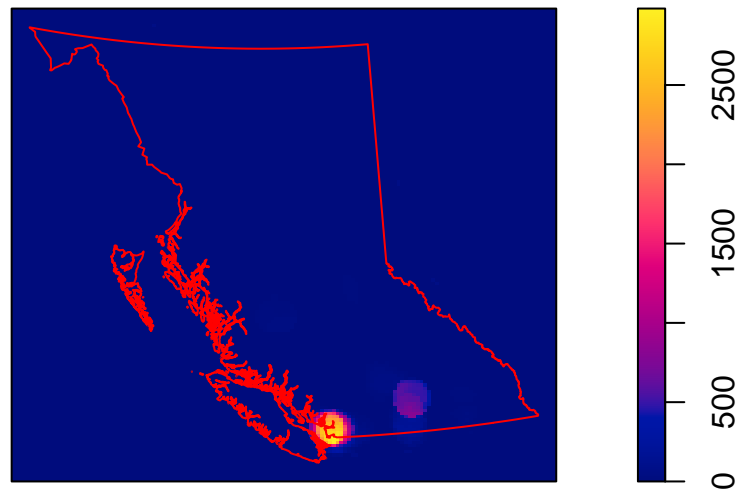


Figure 10: LR

We conducted the ScanLRT test for hot spot analysis, as shown in Figure 10, and the results were almost the same as the likelihood cross-validation bandwidth selection. Areas around Vancouver and Kelowna are two hot spots.

Covariates

Next, we visualized the density estimates of coyote sightings on the four variables to analyze the distribution density of coyotes in different variable intervals.

Elevation

```
## tile
##      (-130,761]      (761,1.1e+03]  (1.1e+03,1.46e+03]  (1.46e+03,3.56e+03]
##              729              106              101              24
```

Figure 11 displays the density estimates of coyote sightings based on elevation. We can observe from the graph that there may be a relationship between elevation and coyote distribution. Moreover, these relationships are non-linear, with higher coyote intensity observed at lower elevations (approximately 0-600).

Forest

```
## tile
##      (0,11.6] (11.6,50.2] (50.2,86.9] (86.9,100]
##          338          409          155          53
```

Next, we plotted the density estimates of coyote sightings based on forest coverage, as shown in Figure 12. We observed that there may be a non-linear relationship between forest coverage and coyote intensity, with higher coyote intensity observed at forest coverage levels between 5% and 40%.

HFI

```
## tile
##      (0,0.000337] (0.000337,0.063] (0.063,0.148] (0.148,0.924]
##              10              22              53              875
```

Then, we plotted the density estimates of coyote sightings based on the Human Footprint Index (HFI), as shown in Figure 13. We observed that there may be a non-linear relationship between HFI and coyote intensity, with higher coyote intensity observed in areas with higher human impact.

water

```
## tile
##      (0,483] (483,1.1e+03] (1.1e+03,2.18e+03] (2.18e+03,2.32e+04]
##          311          157          183          305
```

Finally, we plotted the density estimates of coyote sightings based on distance from water sources, as shown in Figure 14. We observed that there may be a non-linear relationship between distance from water sources and coyote intensity, with higher coyote intensity observed in areas close to water sources or at a moderate distance from them.

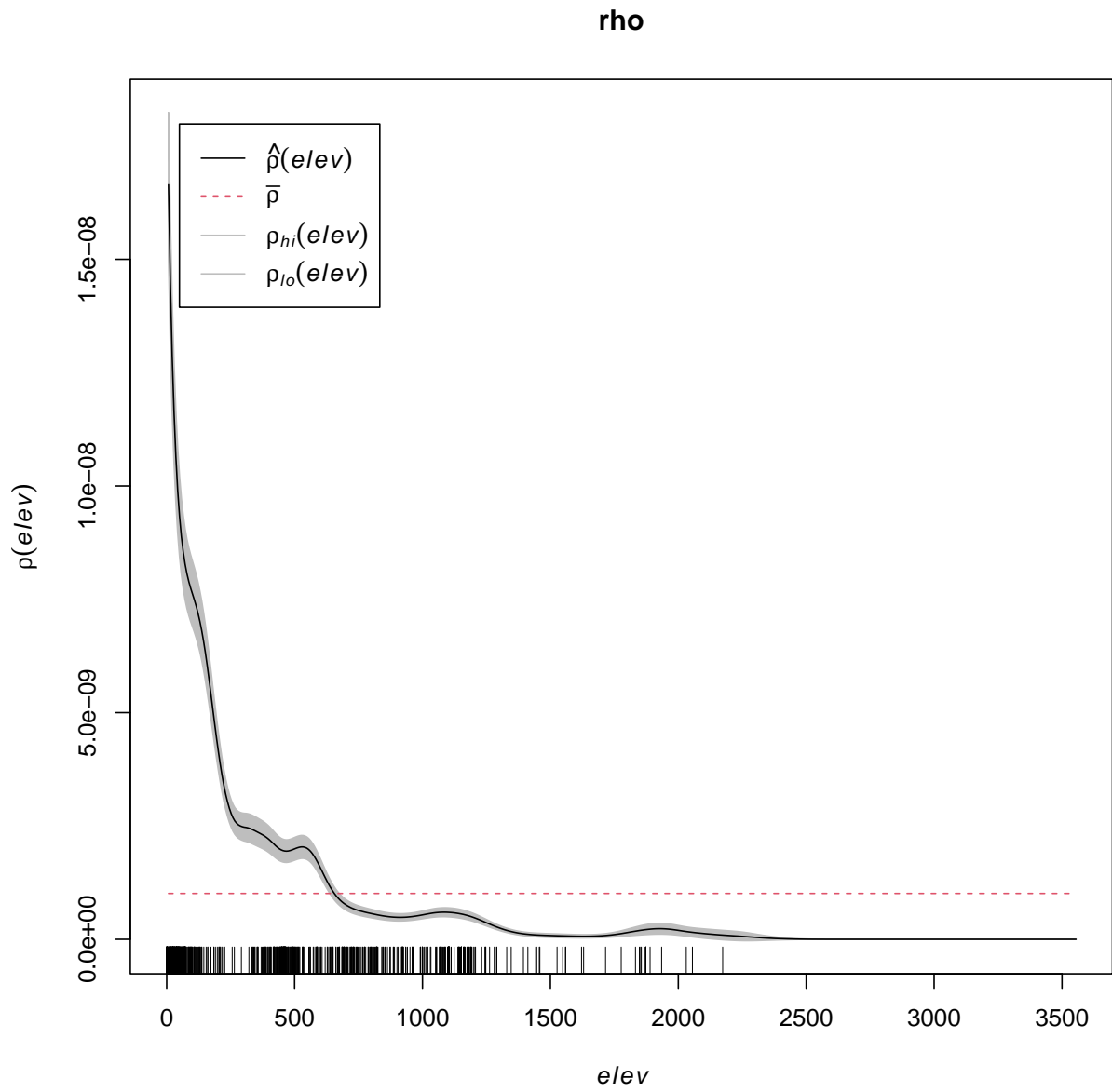


Figure 11: Density Estimates of Coyote on Elevation

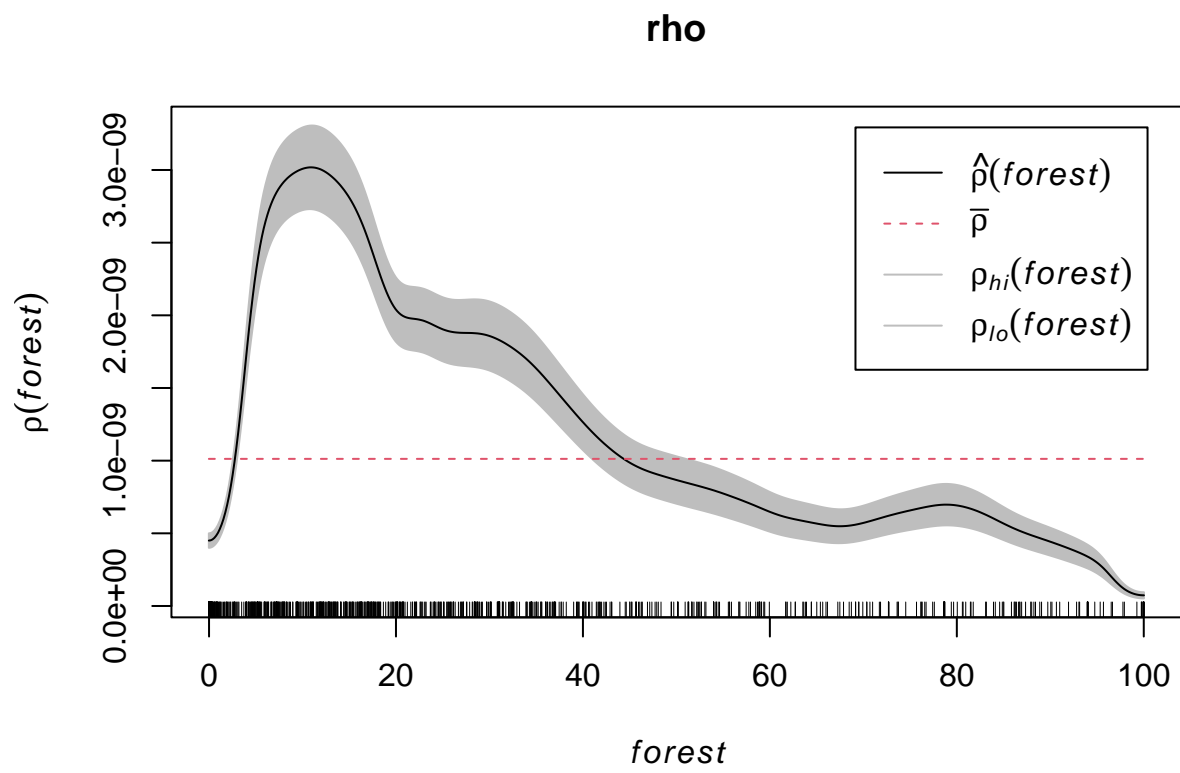


Figure 12: Density Estimates of Coyote on Forest

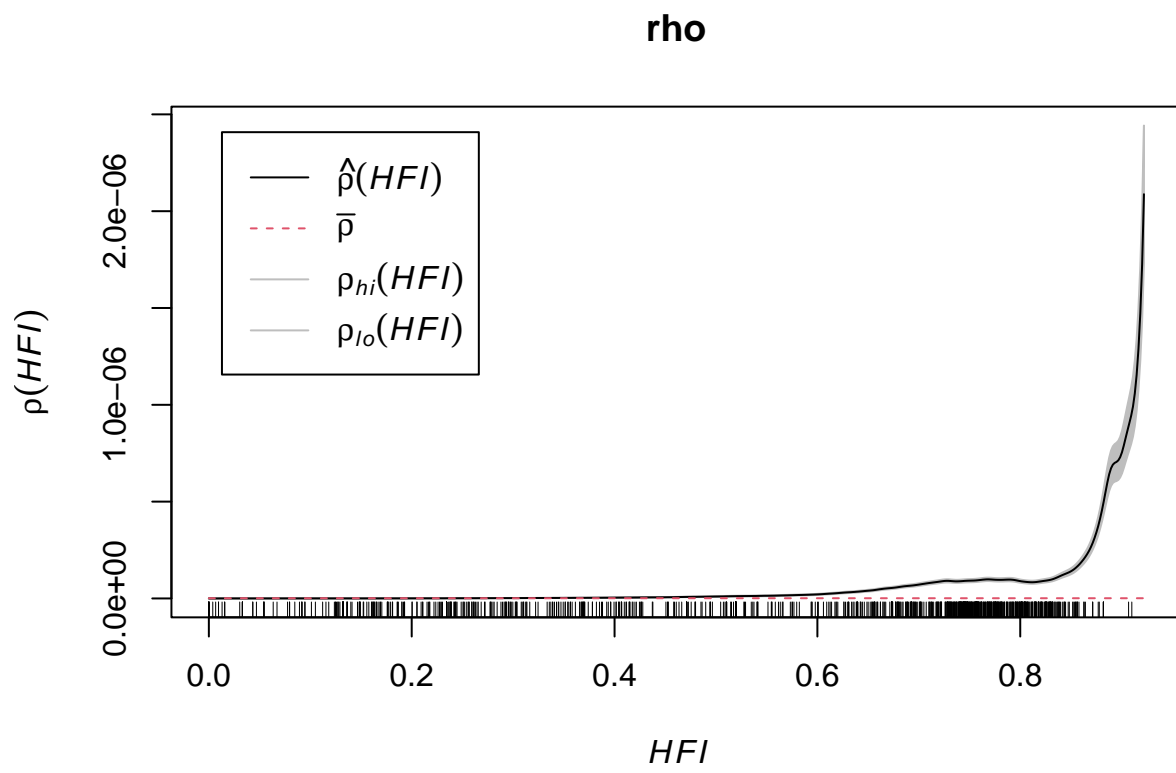


Figure 13: Density Estimates of Coyote on HFI

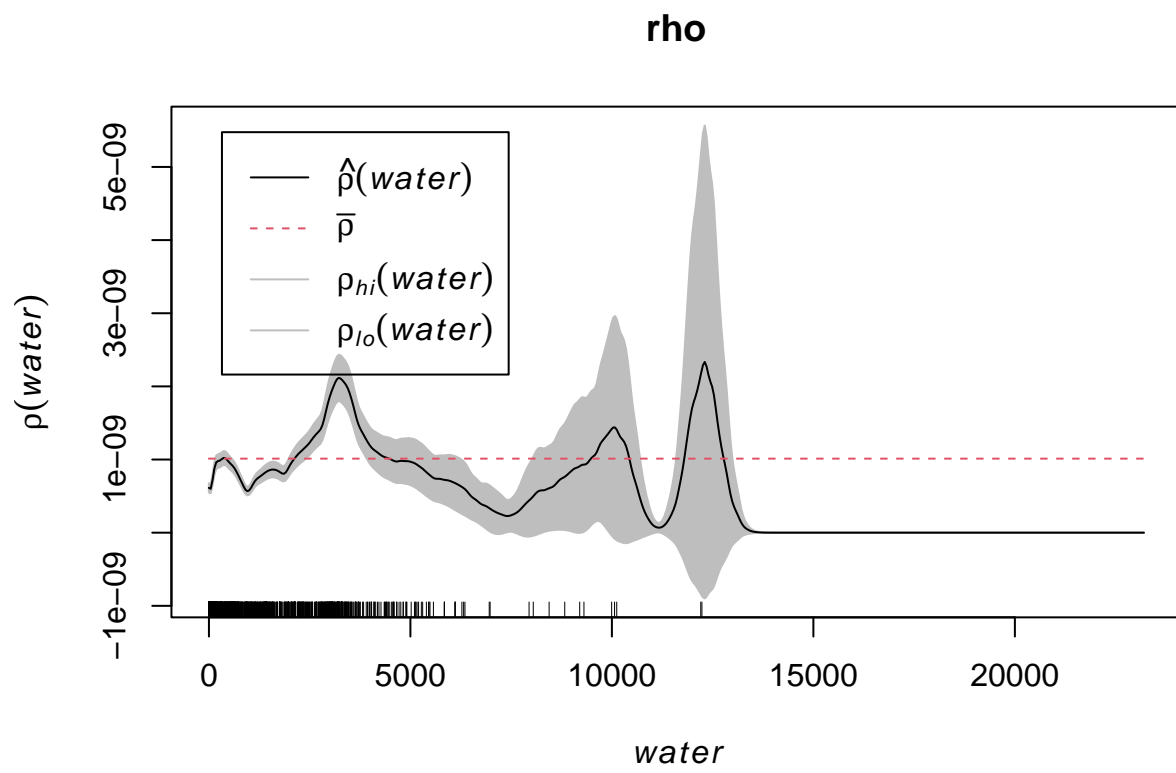


Figure 14: Density Estimates of Coyote on Distance to Water

Furthermore, to gain more insights into the relationship among coyotes, we conducted Second Moment Descriptives analysis. By examining the distribution of the distances between pairs of coyotes, we can determine whether coyotes tend to cluster or disperse spatially. This analysis can also provide information on the spatial scale at which coyotes exhibit strong or weak clustering. This approach not only complements the previous density estimation analysis but also sheds light on the social behavior of coyotes and their habitat preferences.

Second Moment Descriptives

Firstly, we conducted Ripley's K function analysis to examine the spatial clustering of coyotes.

```
## Generating 99 simulations of CSR with fixed number of points ...
## 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99.
##
## Done.
```

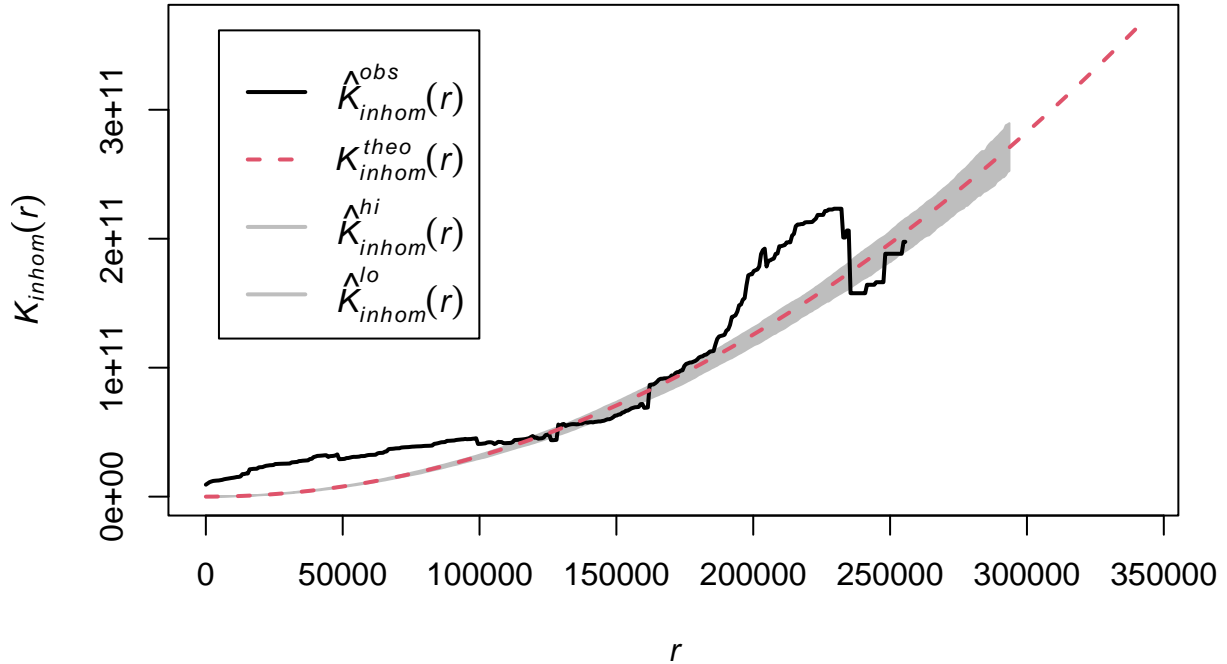


Figure 15: K function

From Figure 15, it is evident that significant clustering of coyote sightings occurs in the distance ranges of 1-100,000 and 180,000-240,000 after adjusting for inhomogeneity. Conversely, significant avoidance occurs in the distance range of 240,000-251,000.

Next, we conducted pair correlation analysis to further investigate the spatial pattern of coyotes.

```
## Generating 19 simulations by evaluating expression ...
## 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19.
##
## Done.
```

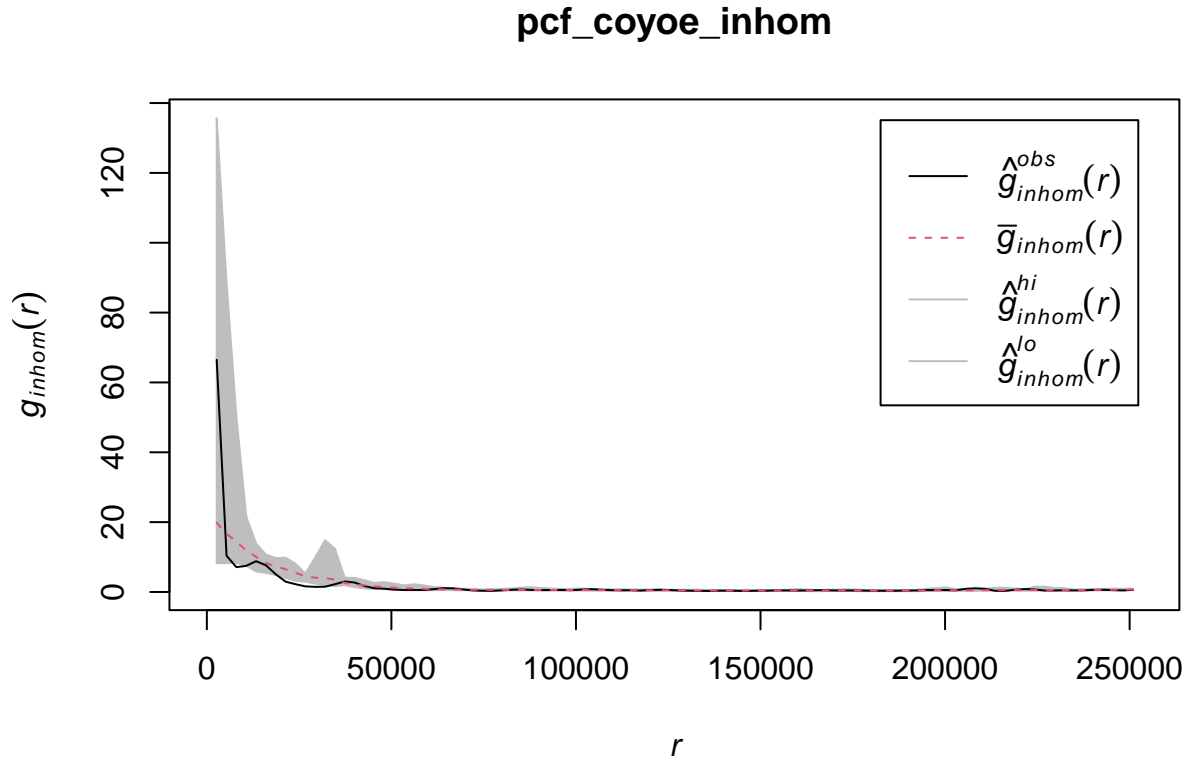


Figure 16: Paired Correlation Function

From Figure 16, it can be observed that after correcting for homogeneity, the number of coyote sightings appears to be significantly lower than expected by random chance in a range of approximately 20,000 to 40,000 meters. Outside of this range, the distribution of coyotes does not appear to exhibit any significant correlation.

Poisson Point Process Model

Table 1: Correlation Between Covariates

	..1	..2	..3	..4
..1	1.0000000	-0.2622538	-0.2662563	-0.0349345
..2	-0.2622538	1.0000000	0.0661859	0.0481860
..3	-0.2662563	0.0661859	1.0000000	0.1324690
..4	-0.0349345	0.0481860	0.1324690	1.0000000

The table 1 shows that there is no obvious collinearity between the four variables.

```

## Nonstationary Poisson process
## Fitted to point pattern dataset 'data_ppp_clipped'
##
## Log intensity: ~Elevation_scaled + I(Elevation_scaled^2) + Forest +
## I(Forest^2) + HFI + I(HFI^2) + Dist_Water_scaled + I(Dist_Water_scaled^2)
##
## Fitted trend coefficients:
##           (Intercept)      Elevation_scaled  I(Elevation_scaled^2)
##      -2.370165e+01      -2.983170e-01      2.169039e-01
##           Forest           I(Forest^2)           HFI
##      -9.761252e-03      8.749969e-05      1.207885e+01
##           I(HFI^2)      Dist_Water_scaled  I(Dist_Water_scaled^2)
##      -5.786288e+00      -2.683963e-01      1.268933e-02
##
##           Estimate      S.E.      CI95.lo      CI95.hi
## (Intercept)      -2.370165e+01  0.1516761572  -2.399893e+01  -2.340437e+01
## Elevation_scaled      -2.983170e-01  0.0592147839  -4.143758e-01  -1.822581e-01
## I(Elevation_scaled^2)      2.169039e-01  0.0292257746  1.596224e-01  2.741854e-01
## Forest      -9.761252e-03  0.0039058495  -1.741658e-02  -2.105928e-03
## I(Forest^2)      8.749969e-05  0.0000415225  6.117086e-06  1.688823e-04
## HFI      1.207885e+01  0.6125856238  1.087820e+01  1.327949e+01
## I(HFI^2)      -5.786288e+00  0.6231222672  -7.007585e+00  -4.564991e+00
## Dist_Water_scaled      -2.683963e-01  0.0497755698  -3.659547e-01  -1.708380e-01
## I(Dist_Water_scaled^2)      1.268933e-02  0.0183822635  -2.333925e-02  4.871790e-02
##           Ztest      Zval
## (Intercept)      *** -156.2648638
## Elevation_scaled      ***  -5.0378800
## I(Elevation_scaled^2)      ***   7.4216644
## Forest      *  -2.4991368
## I(Forest^2)      *   2.1072838
## HFI      ***  19.7178131
## I(HFI^2)      ***  -9.2859592
## Dist_Water_scaled      ***  -5.3921300
## I(Dist_Water_scaled^2)      0.6903027
## Problem:
## Values of the covariate 'HFI' were NA or undefined at 0.56% (18 out of 3219)
## of the quadrature points

```

since the process is inhomogeneous, fit the model with the function form $\lambda = e^{-2.37016510 \times 10 - 2.983170 \times 10^{-01} elev + 2.169039 \times 10^{-01} elev^2}$

since the p-value of $water^2$ is smaller than 0.05, which means it's not important, so we delete it. Then we fit a new model.

```
## [1] 36670.39
```

```
## [1] 36668.83
```

```
## [1] -1.555642
```

```
## Analysis of Deviance Table
```

```
##
```

```
## Model 1: ~Elevation_scaled + I(Elevation_scaled^2) + Forest + I(Forest^2) + HFI + I(HFI^2) + Dist_Wa
```

```
## Model 2: ~Elevation_scaled + I(Elevation_scaled^2) + Forest + I(Forest^2) + HFI + I(HFI^2) + Dist_Wa
```

```
##      Npar Df Deviance Pr(>Chi)
## 1      8
## 2      9  1  0.44436    0.505
```

The reduced model has better AIC. Other than that, p-value of all the parameters are greater than 0.05, so we retain the null hypothesis that the reduced model can sufficiently explain the data.

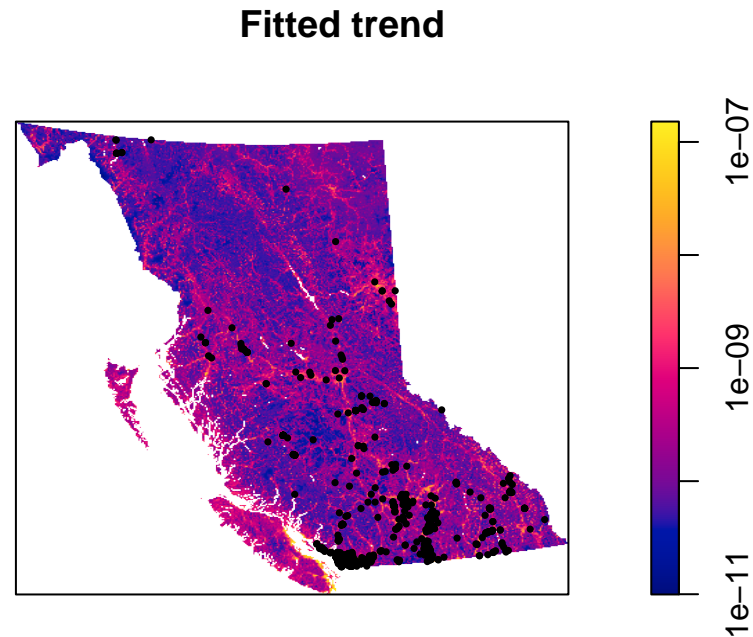


Figure 17: Predicted Intensity

We then plot the model prediction plot and the distribution points of the coyotes, as shown in Figure 17. The plot shows that the model prediction for the coyote distribution is quite accurate, but there is over-prediction in some areas.

```
##
## Chi-squared test of fitted Poisson model 'fit_reduced' using quadrat
## counts
##
## data: data from fit_reduced
## X2 = 264.72, df = 0, p-value < 2.2e-16
## alternative hypothesis: two.sided
##
## Quadrats: 8 tiles (irregular windows)
```

We conducted a `quadrat.test` with `nx = 2`, `ny = 4` to split the area into 8 small rectangles. The result showed that the p-value was less than 0.05, indicating that the fitted model was not good enough. This is consistent with the result of over-prediction in some areas that we found in the previous analysis.

Then we validated the model with residual plots.

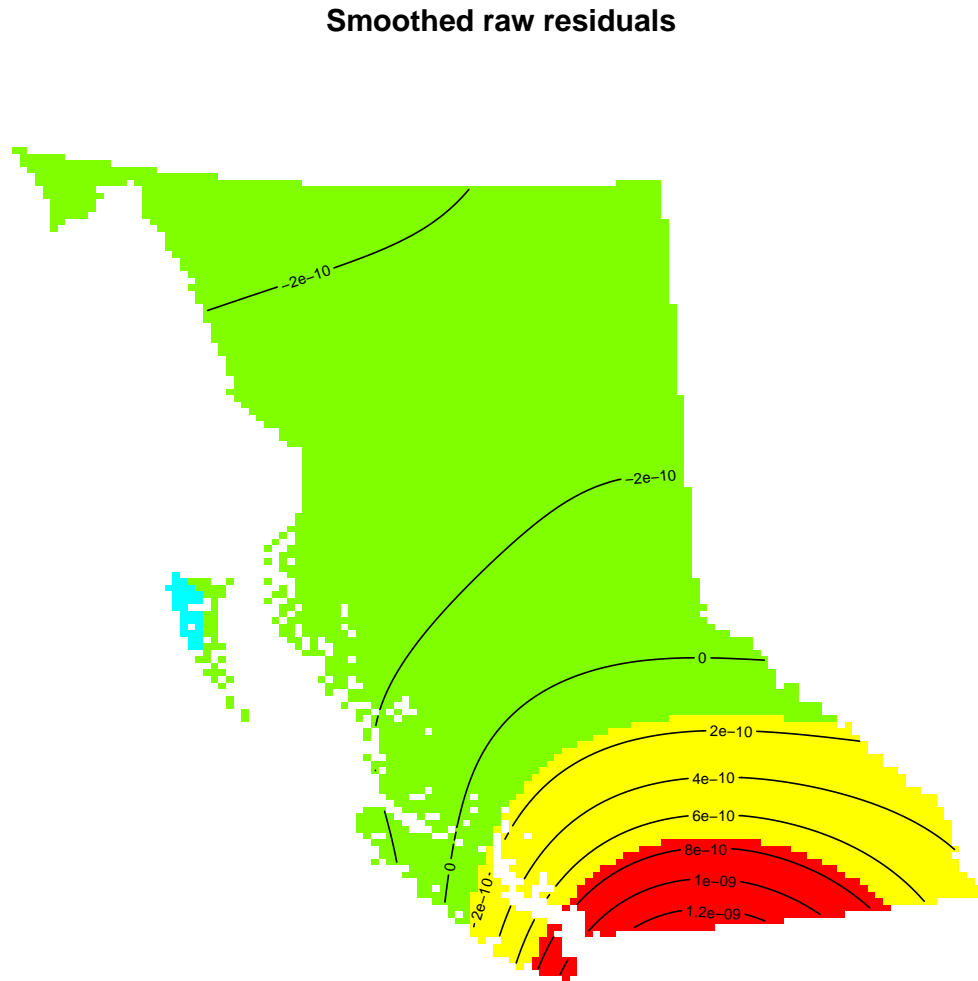


Figure 18: Residual Plot

The residual plot shown in Figure 18 indicates that the residuals are mostly centered around 0, suggesting a good fit of the model. However, the largest residuals are found in the southern part of BC, indicating over-prediction in this area.

Then we accessed model fit with partial residual plots of each of the covariates.

Firstly, elevation.

As shown in Figure 19, the quadratic term of elevation did not capture the pattern well in the data, so we added higher-order polynomial terms to elevation in later analysis to improve the fitting accuracy of the model.

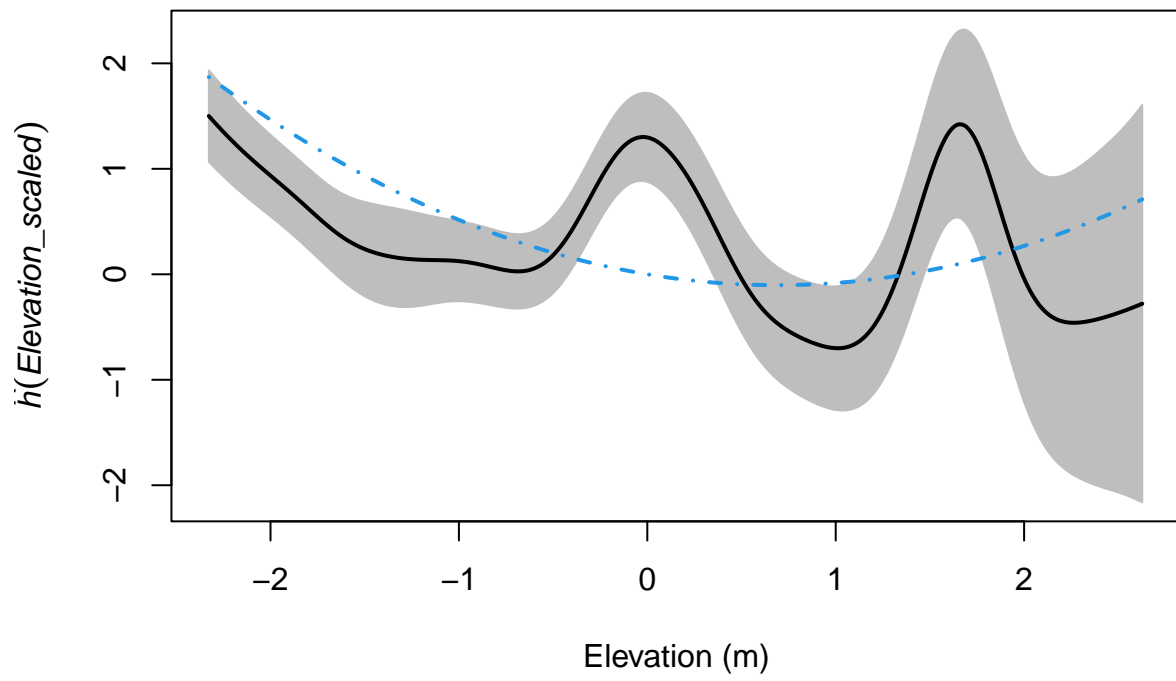


Figure 19: Relative Intensity as a Function of Elevation

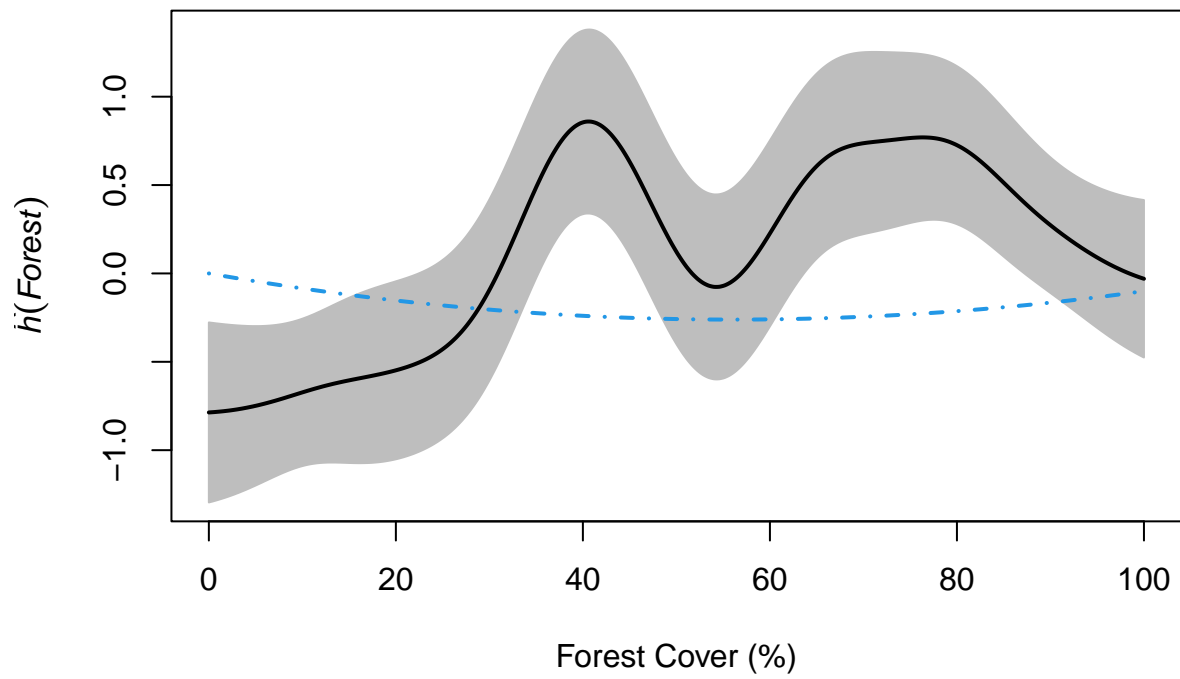


Figure 20: Relative Intensity as a Function of Forest

As shown in Figure 20, the quadratic term of forest cover did not capture the pattern well in the data, so we added higher-order polynomial terms to forest cover in later analysis.

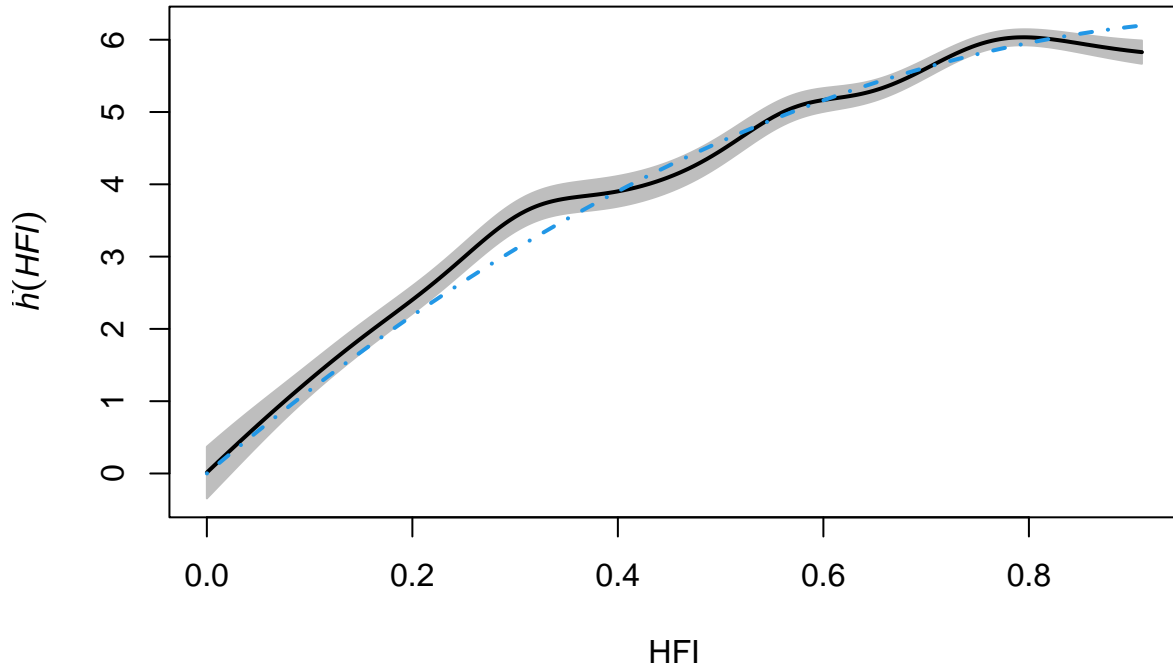


Figure 21: Relative Intensity as a Function of HFI

As shown in Figure 21, the quadratic term of HFI captures the pattern in the data well, so we did not make any changes to HFI.

As shown in Figure 22, the linear term of the variable “water” captures the pattern in the data well, so we do not make any modifications to this variable.

Next, based on the analysis of the partial residual plots as shown above, we improved our model.

```
## Nonstationary Poisson process
## Fitted to point pattern dataset 'data_ppp_clipped'
##
## Log intensity: ~bs(Elevation_scaled, 7) + bs(Forest, 6) + HFI + I(HFI^2) +
## Dist_Water_scaled
##
## Fitted trend coefficients:
##           (Intercept) bs(Elevation_scaled, 7)1 bs(Elevation_scaled, 7)2
##           -20.17505836          -4.16821031          -1.88055388
## bs(Elevation_scaled, 7)3 bs(Elevation_scaled, 7)4 bs(Elevation_scaled, 7)5
##           -4.48334726          -2.02942424          -7.53347469
## bs(Elevation_scaled, 7)6 bs(Elevation_scaled, 7)7          bs(Forest, 6)1
##           1.65313127          -13.31767272           0.50035274
##           bs(Forest, 6)2          bs(Forest, 6)3          bs(Forest, 6)4
```

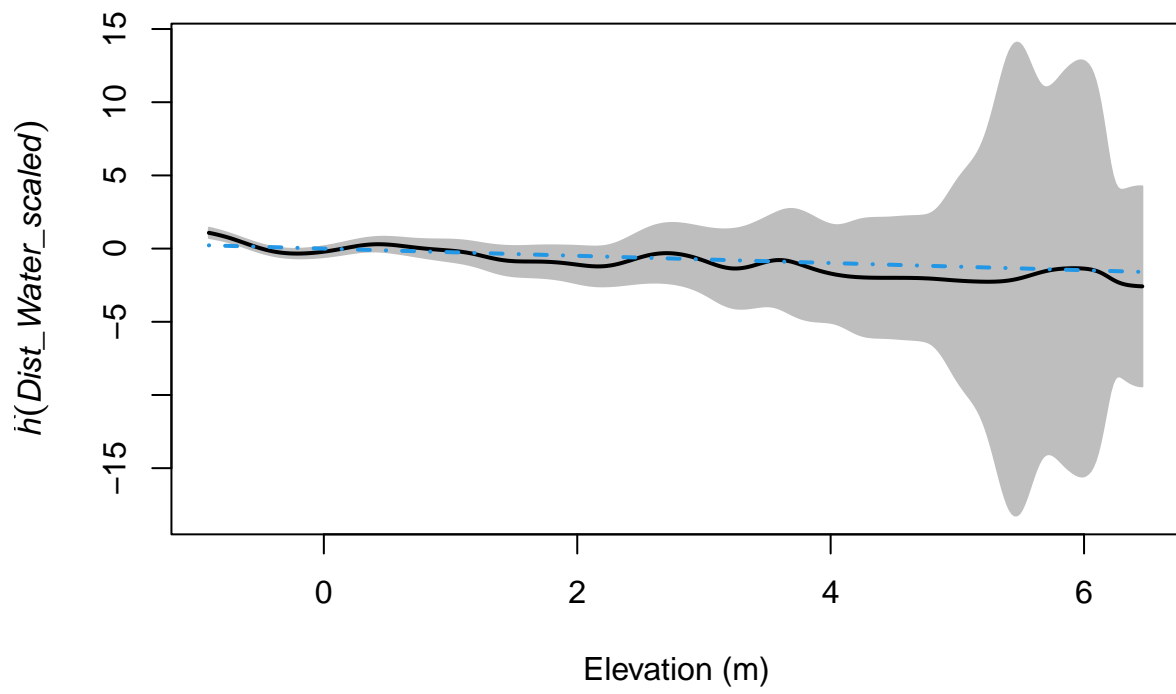


Figure 22: Relative Intensity as a Function of Distance to Water

```
##          0.07092284          -0.60415436          0.84362384
##      bs(Forest, 6)5      bs(Forest, 6)6          HFI
##      -0.08145046      -0.54986923      12.09916981
##          I(HFI^2)      Dist_Water_scaled
##      -5.71865475      -0.15858633
##
## For standard errors, type coef(summary(x))
## Problem:
## Values of the covariate 'HFI' were NA or undefined at 0.56% (18 out of 3219)
## of the quadrature points
```

The new model includes the 7th power of elevation and the 6th power of forest cover, while the other parameters remain the same.

Then we analyzed the partial residual plots of the new model for elevation and forest cover, as shown in the following figures.

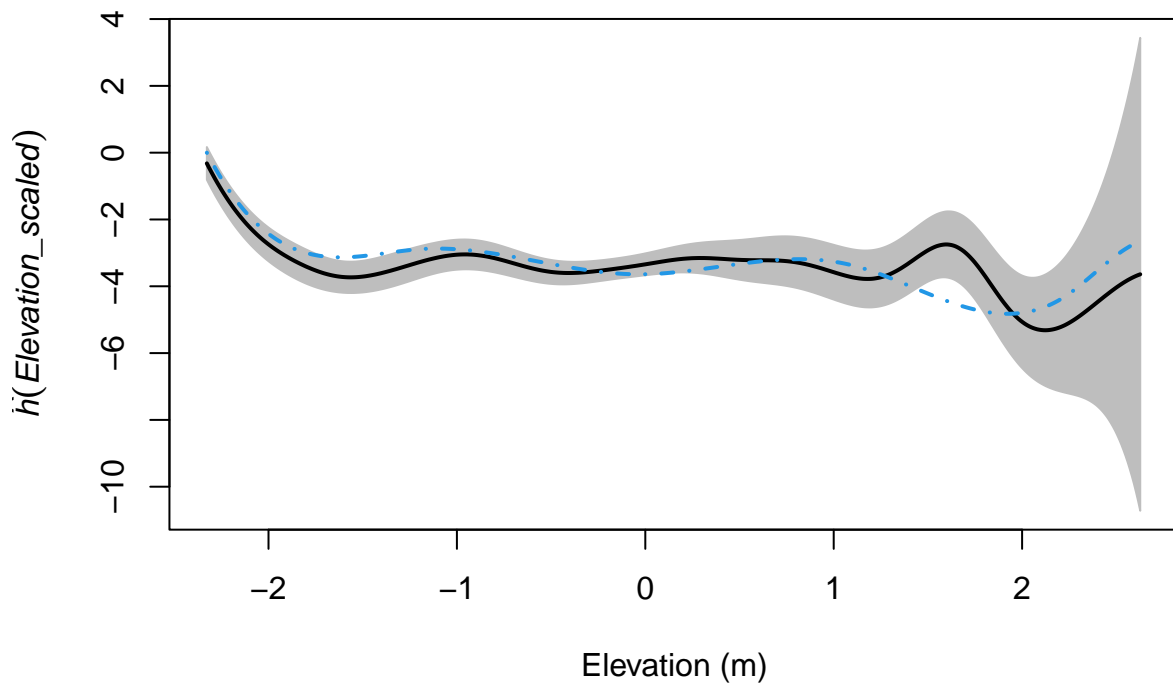


Figure 23: Relative Intensity as a Function of Elevation

From figure 23 and figure 24, the models look like a much better fit to the data, but we have added a lot of complexity. To ensure we are not overfitting, we again use our model selection techniques.

```
## [1] 36668.83
```

```
## [1] 36486.55
```

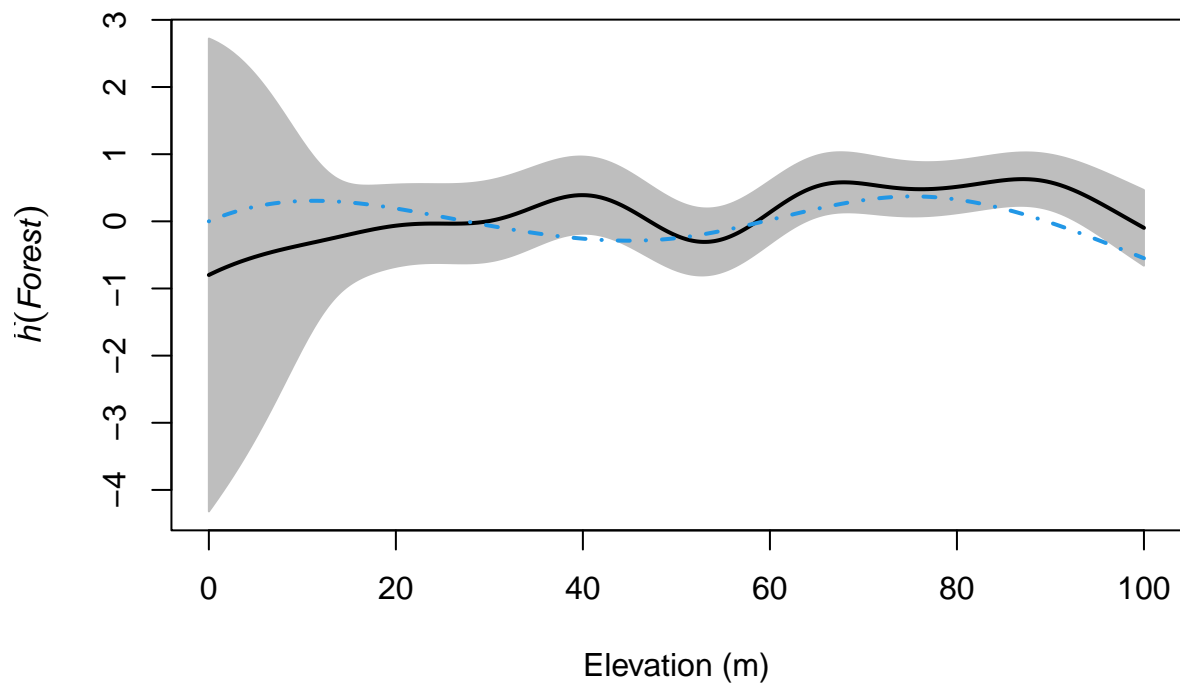


Figure 24: Relative Intensity as a Function of Forest

```
## [1] -182.2779

##
## Chi-squared test of fitted Poisson model 'fit_smooth' using quadrat
## counts
##
## data: data from fit_smooth
## X2 = 1008.6, df = 47, p-value < 2.2e-16
## alternative hypothesis: two.sided
##
## Quadrats: 64 tiles (irregular windows)

## Analysis of Deviance Table
##
## Model 1: ~bs(Elevation_scaled, 7) + bs(Forest, 6) + HFI + I(HFI^2) + Dist_Water_scaled Poisson
## Model 2: ~Elevation_scaled + I(Elevation_scaled^2) + Forest + I(Forest^2) + HFI + I(HFI^2) + Dist_Wa
## Npar Df Deviance Pr(>Chi)
## 1 17
## 2 9 -8 -199.83 < 2.2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The AIC value of the fit_reduced model is 36668.83, while the AIC value of the fit_smooth model is 36486.55. A lower AIC value indicates a better fit of the model while considering the complexity of the model. In this example, the fit_smooth model has a lower AIC value, indicating that it may have a better fit. Additionally, the LRT test results show a p-value less than $2.2e-16$, rejecting the null hypothesis that the two models have the same fit, indicating that the fit_smooth model has a significantly better fit than the fit_reduced model. The chi-squared test result also shows a p-value less than $2.2e-16$, indicating that the explanatory variables in the fit_smooth model are significant in explaining the observed data. Overall, the output indicates that the fit_smooth model has a better fit than the fit_reduced model.

The final model is too long to write down, but we can visualise the residual plot just as before.

```
## Model diagnostics (raw residuals)
## Diagnostics available:
## four-panel plot
## mark plot
## smoothed residual field
## x cumulative residuals
## y cumulative residuals
## sum of all residuals
## sum of raw residuals in entire window = -2.787e-11
## area of entire window = 9.483e+11
## quadrature area = 9.411e+11
## range of smoothed field = [-3.37e-10, 1.173e-09]
```

As shown in Figure 25, the residuals of the modified model in the southern part of BC are closer to 0, indicating a better fit to the reality and thus a better performance of the model.

Smoothed raw residuals

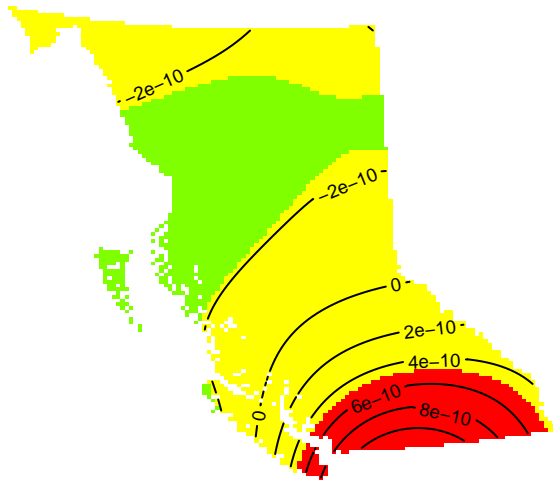


Figure 25: Residual Plot

Discussion

The spatial distribution of coyotes is inhomogeneous, and it has been observed that in the southern part of British Columbia, coyotes tend to cluster together. Several factors seem to influence the clustering of coyotes, including elevation, forest cover, the Human Footprint Index (HFI), and the distance from water sources. More specifically, coyotes tend to congregate in areas with low elevation (below 550 meters), low forest cover (less than approximately 43%), significant human impact, and close to moderate distances from water sources.

Our final model incorporates the 7th power of elevation and the 6th power of forest cover, as well as the quadratic terms for distance to water sources and the HFI. This model has been assessed for its goodness of fit using the AIC and residual plots, which indicate that the model is performing well.

In summary, the clustering of coyotes in southern British Columbia appears to be influenced by a combination of environmental and human factors. Areas with lower elevation, less forest cover, greater human impact, and not far away from water sources are more likely to have higher concentrations of coyotes. These environmental characteristics are in accord with the environmental features of metropolitan areas. In fact, hot spot analysis shows that coyotes are significantly concentrated in areas around Vancouver and Kelowna, the two biggest metropolitan areas of BC. The understanding of these factors and their influence on coyote distribution can inform wildlife management and conservation efforts. Future city development and expansions could take the insights gained into account to minimize potential conflict between human settlements and coyotes.