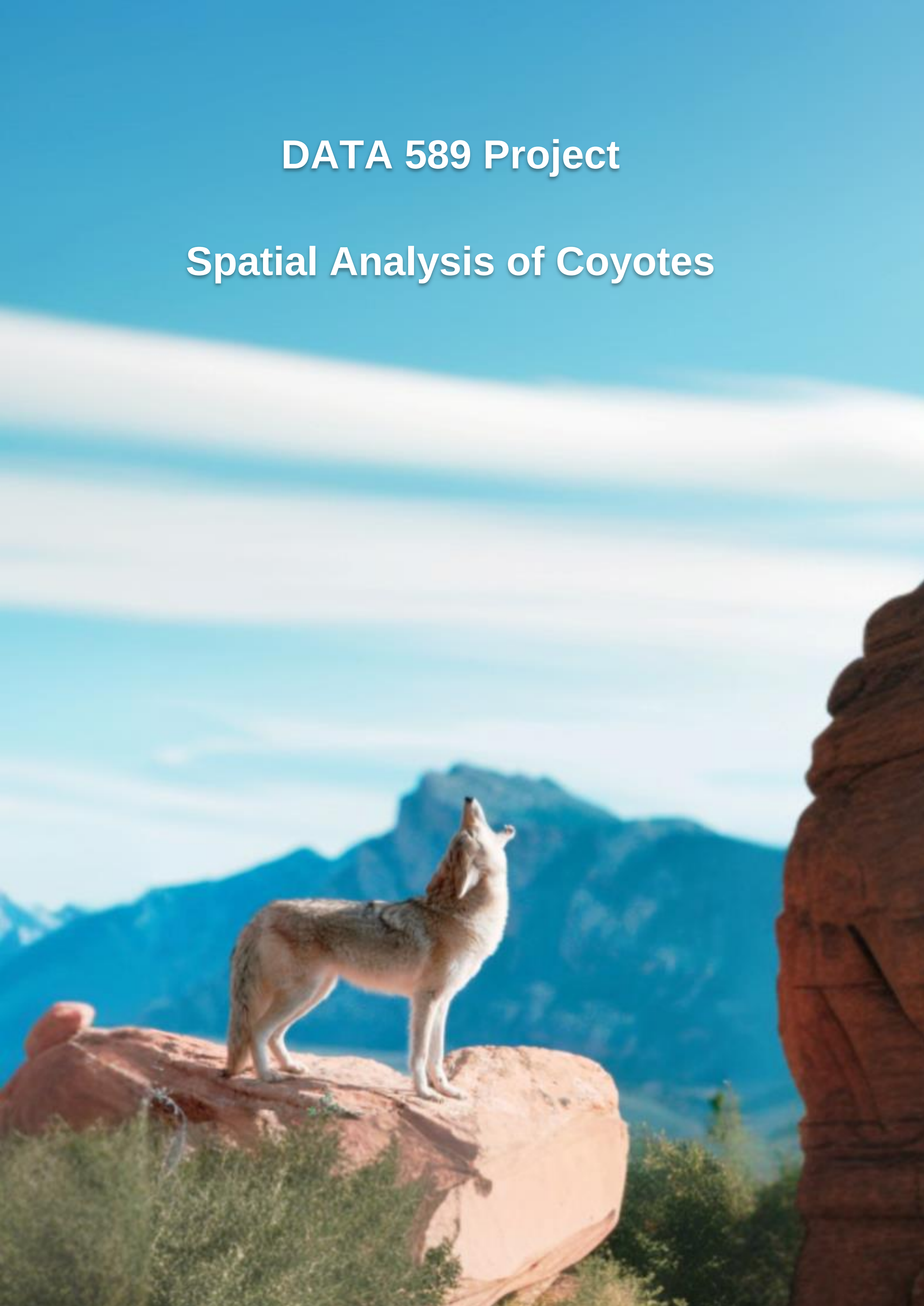
DATA 589 Project

Spatial Analysis of Coyotes



DATA 589 Project

Spatial Analysis of Coyotes

Kohsin Chen, Yahan Cong, Yihpeng Kuo, Lu Kang

1. Introduction

The coyote is a highly adaptable native carnivore found throughout North America, thriving in both urban and rural settings. Known for their intelligence, flexibility, and resilience, coyotes excel in environments heavily influenced by human activities. Their diet is diverse, encompassing small mammals, insects, fruits, and vegetables. This adaptability has sparked considerable interest in studying the spatial dynamics of their distribution.

Our project utilizes a dataset of detailed occurrence records of coyotes in British Columbia, sourced from the Global Biodiversity Information Facility (GBIF). These records, collected through various observational and tracking methods during 2023 and 2024, include precise geographic coordinates and temporal details of each sighting.

Previous studies have shown that environmental factors influence the occurrence of coyotes and conflicts in urban areas. (Poessel, Gese, & Young, 2017b). Additionally, Alexander and Quinn (2011) conducted a comprehensive analysis of coyote interactions with humans and pets as reported in Canadian print media. Their study demonstrated how emotional responses and the portrayal of coyotes often reinforce negative stereotypes, which can influence public attitudes and potentially lead to unwarranted fear and persecution of these animals.

Our research aims to use the latest data to deeply understand the lifestyles of coyotes in British Columbia, and to answer the following core research questions:

- Do coyotes exhibit clustering behaviour, and if so, under what conditions is this behaviour most pronounced?
- Are there specific areas where coyotes are more frequently observed?
- How do elevation forest cover, the Human Footprint Index (HFI) and proximity to water bodies affect the distribution of coyotes?
- Considering spatial heterogeneity, which statistical model most accurately predicts the distribution of coyotes, and how can model improvements better capture the effects of environmental variables?

By exploring these questions, we hope to provide deeper insights into how coyotes adapt to and impact their ecosystems. This is not only crucial for ecological conservation and species management but also for understanding how human activities alter the behaviour and distribution of wildlife.

2. Methold

We retrieved data on coyote sightings within British Columbia (BC) for the years 2023 and 2024 from the Global Biodiversity Information Facility (GBIF) database. To streamline our data for efficient analysis, we selectively extracted data points from 2023 to 2024, employing functions such as unique() to eliminate duplicates, resulting in a dataset comprising 347 unique observations. The dataset, named BC_covariates, encompasses four environmental attributes: elevation, forest coverage, the Human Footprint Index (HFI), and proximity to water bodies.

Our analysis commenced with a detailed examination of the spatial distribution of these coyote occurrences using descriptive statistics to define the scope of our study. Through the spatstat package, we calculated spatial intensity to discern patterns in the data. The homogeneity of the distribution was assessed via “quadrat count” and “quadrat test” methods, providing a baseline for understanding the data's structure.

We enhanced our statistical analysis by delving deeper into the spatial relationships of coyotes using second-moment descriptive statistics. We utilized Morisita's index (via Miplot) to analyze individual density distribution patterns. For spatial clustering, we applied Ripley's K-function (using Kest for homogeneous assumptions and Kinhom for inhomogeneous assumptions) and assessed the significance of clustering through simulation envelopes set at a 95% confidence interval (envelope() function). Additionally, we scrutinized point-pattern intricacies using both homogeneous (pcf) and inhomogeneous (pcfinhom) point correlation functions, also evaluating these with a 95% confidence interval. This comprehensive application of spatial statistical techniques enables robust inferences about the influence of environmental variables on coyote distribution.

To elucidate the association between coyote distribution and the covariates of elevation, forest cover, human footprint index, and distance to water, we categorized these four covariates into five classes based on a gradient from low to high values. Within each class, we analyzed the distribution of coyote occurrences to detect any discernible patterns. Further refining our analysis, we utilized the rhohat function from the spatstat package. This function is designed to estimate the intensity of a spatial point process, providing an initial assessment of the relationship between each covariate and the spatial distribution of coyotes. By interpreting the intensity maps generated, we were able to infer potential correlations or dependencies that coyotes may have with these environmental factors.

Before fitting the model, we examined multicollinearity among the variables. If multicollinearity is not significant, this suggests that each covariate provides unique information and can be included in the model without concern for inflating the variance of the estimated coefficients.

Finally, we employed the ppm function from the spatstat package to fit various point process models to our data, comparing models incorporating both linear and quadratic relationships. We extended our analysis using an additive modelling framework with Generalized Additive Models (GAMs) and employed the splines package to accommodate more complex relationships. Model diagnostics were conducted using Akaike's Information Criterion (AIC) and ANOVA likelihood ratio tests. We further examined model fit through partial residuals and lurking variables checks, employing the diagnosis.ppm() function to ensure robustness in our model's predictions. This comprehensive approach allowed us to identify the model that best balances fit and complexity, providing a detailed and accurate explanation of the observed distribution patterns of coyotes.

3. Analysis Results

3.1 General analysis

Figure 1 displays coyotes in British Columbia, with a notable concentration in the south and along the coast, suggesting these areas as potential hotspots for coyotes activity or habitats.

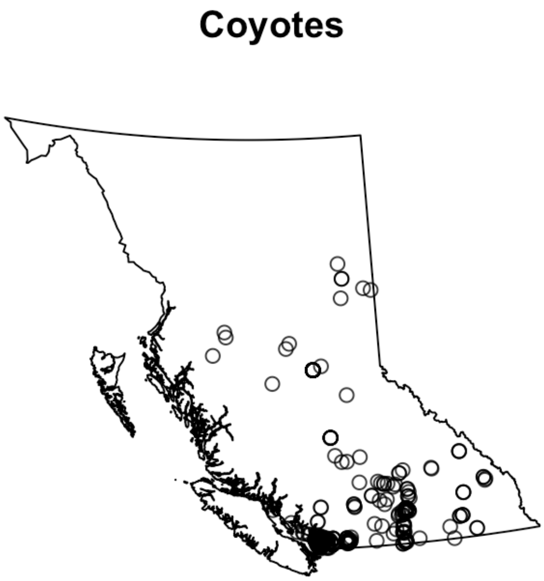


Figure 1. Coyotes in British Columbia

3.1.1 Quadrats Visualization

The quadrat visualization map for British Columbia indicates that numerous quadrats with zero, especially in the northern region which is marked by an absence of sightings. Three dense clusters are evident: one centrally located with 46 sightings, and two in the south, reporting 146 and 107 sightings respectively. Additionally, other quadrats display counts ranging from 1 to 22, indicating a varied spatial distribution of coyotes' sightings across the province.

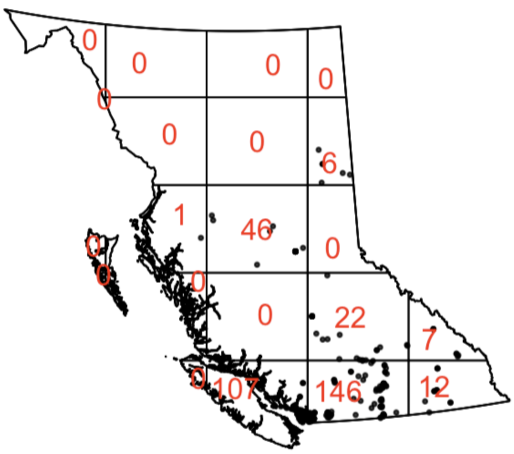


Figure 2. Quadrats Plot of Coyotes in BC

3.1.2 Intensity

Figure 3 illustrates the intensity of coyotes sightings across British Columbia. Areas shaded in green to yellow represent lower intensities, while those in orange to brown depict higher sighting intensities. The blue dots represent actual sighting locations, clustered mainly in the southern region.

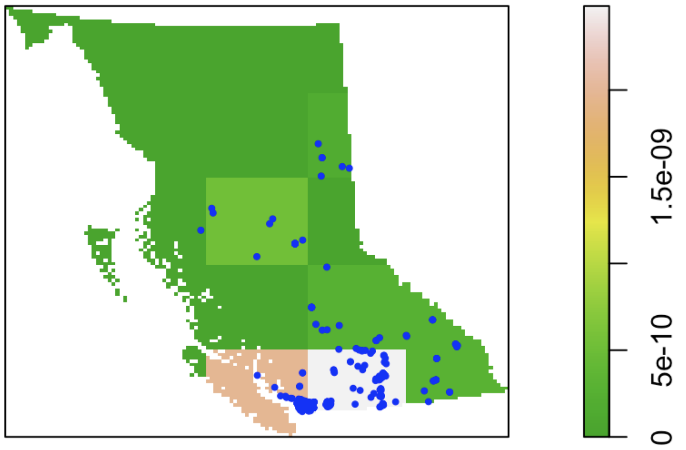


Figure 3. Intensity of Coyotes

3.1.3 Heat map

Figure 4 shows the intensity of coyote sightings across British Columbia. High-intensity areas are colored red, transitioning to yellow and green for moderate intensities, and blue for low or no sightings. Most sightings are concentrated in the southern region, with a significant hotspot evident.

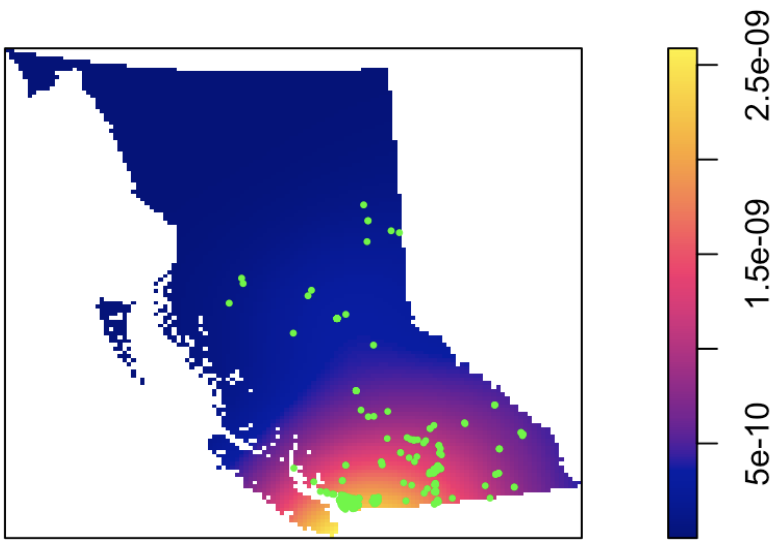


Figure 4. Heat Map

3.2 First Moment Analysis - Covariates Analysis

In order to discover the relationship between occurrences and covariates, it's crucial to map and visualize the data. These maps illustrate how covariates such as elevation, forest cover, human footprint index, and proximity to water bodies may influence where coyotes are found. By plotting occurrences against these environmental factors, we can identify trends and correlations, and inform model development.

Figure 5 and Table 1 below present the distribution of coyote sightings in British Columbia relative to various environmental covariates classified into categories ranging from very low to very high. This classification allows us to observe trends such as the majority of sightings at lower elevations and near water bodies, and varying patterns with forest cover and human footprint index. Such data is instrumental for ecologists to determine the impact of these covariates on coyote habitat selection and behavior.

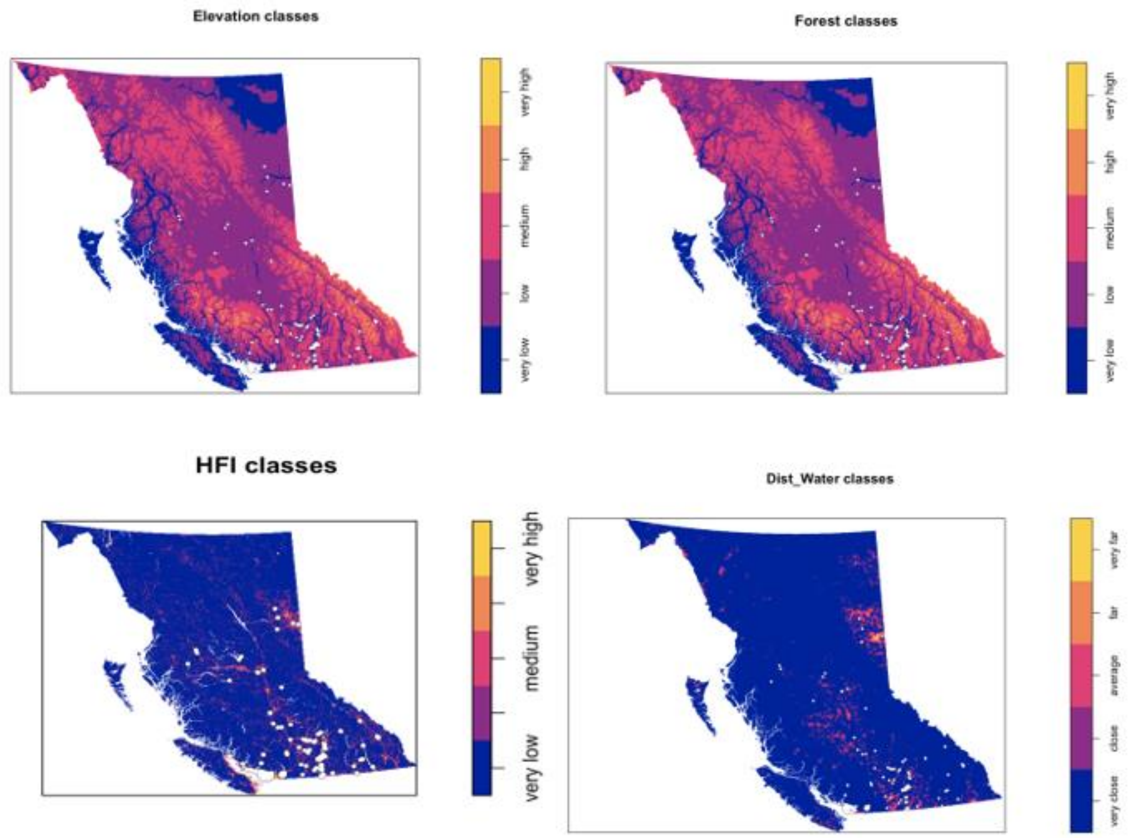


Figure 5. The distribution of Coyotes vs Different Covariates

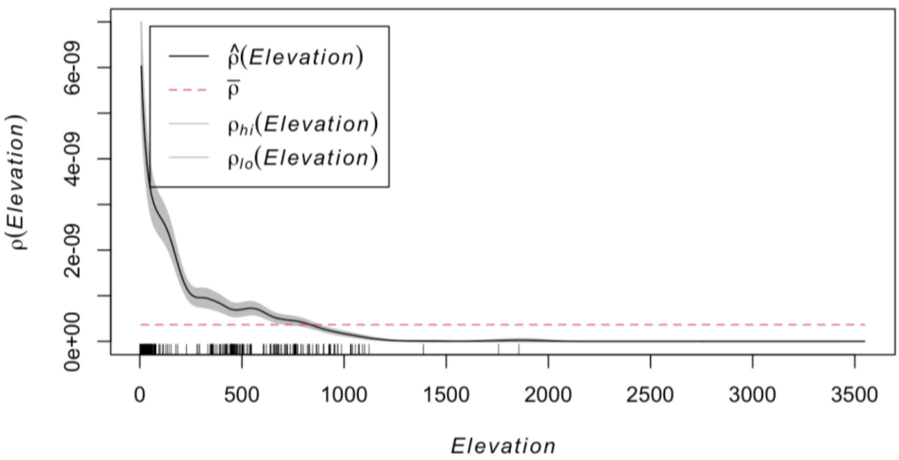
Table 1. The distribution of Coyotes vs Different Covariates

	very low (very close)	low (close)	medium (average)	high (far)	very high (very far)
Elevation	233	108	3	0	0
Forest Cover	165	83	22	52	19
HFI	19	38	73	61	127
Distance from water	332	13	2	0	0

- **Elevation**: The table below shows that coyote observations are significantly higher in areas with very low to low elevation, with 233 sightings in very low and 108 in low elevation areas, and no sightings in areas classified as medium to very high elevation.
- **Forest coverage**: This map portrays forest coverage with the same classification range. Here, coyote sightings are also distributed mostly in the lower classes (very low and low), with a gradual decrease in sightings as forest coverage increases, suggesting a preference for greater ease of detection in areas with less dense forestation.
- **Human footprint index**: The data shows a broader spread of sightings across HFI classes, with the highest number of sightings in the very high class, which may indicate coyote adaptability to human-modified environments or could reflect biased sampling near human habitats.
- **Distance from water**: As the distance from water increases, coyote sightings become increasingly sparse, which could suggest that water availability may be a key factor in coyote habitat selection or that these areas are more frequently surveyed.

Because we are interested in modeling λ(u) as a function of covariates Z(u), we use the rhohat() function to help us identify a reasonable starting model.

Figure 6. ρ vs Elevation by Kernel Estimation



From graph 6, we can see that lower elevations have a higher than average density of coyote occurrences. This suggests that coyotes are far more prevalent at lower elevations, with their presence dropping off significantly at higher altitudes. The relationship looks almost exponential, with a rapid decrease in sightings as elevation rises from the base level.

Figure 7. ρ vs Forest by Kernel Estimation

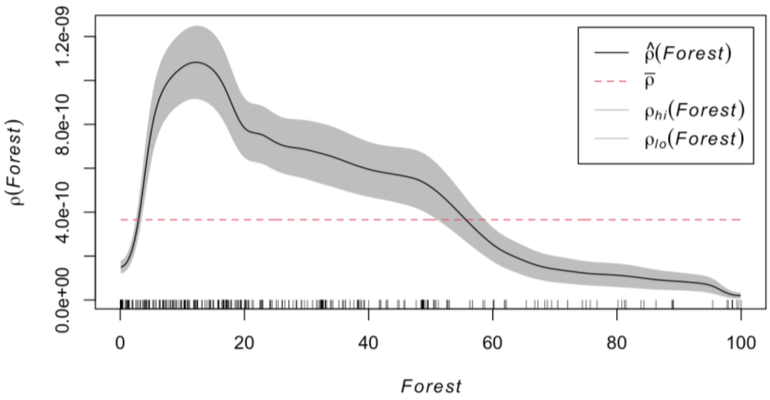


Figure 7 suggests coyotes are predominantly observed in regions with less dense forest cover, especially noticeable in areas where the forest coverage is below 50%. The relationship between coyote sightings and forest coverage appears to be quadratic, with sightings peaking at lower forest densities and then diminishing, suggesting a preference or easier detection in less dense forests.

Figure 8. ρ vs HFI by Kernel Estimation

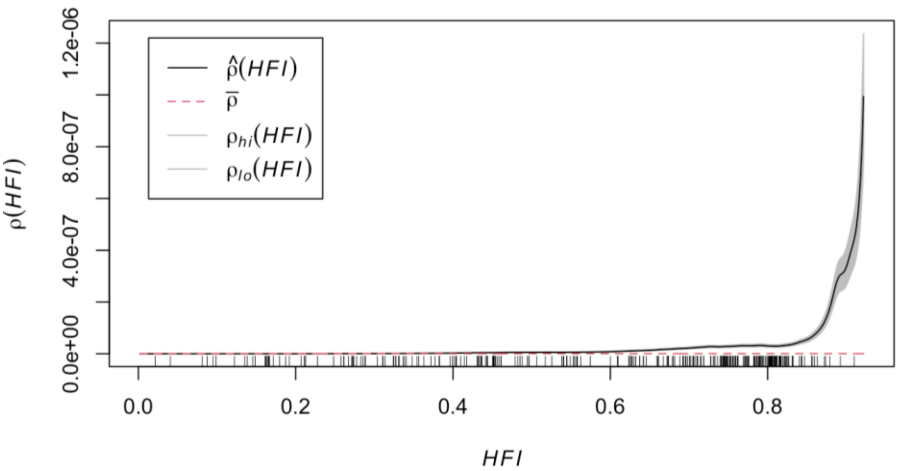


Figure 8 showing coyote occurrences versus the Human Footprint Index (HFI) suggests a complex relationship. The intensity of sightings is initially low and increases sharply as the HFI approaches 1. This might indicate that coyote sightings are more common in areas with higher human activity or impact. The shape of the curve seems to suggest a sigmoidal or logistic relationship, where the probability of coyote sightings is low in both very low and medium HFI areas but increases significantly in areas of high human footprint. The confidence intervals (shown by the shaded area around the estimated curve) widen as HFI increases, indicating more uncertainty in areas with a higher human footprint.

Figure 9. ρ vs Distance from Water by Kernel Estimation

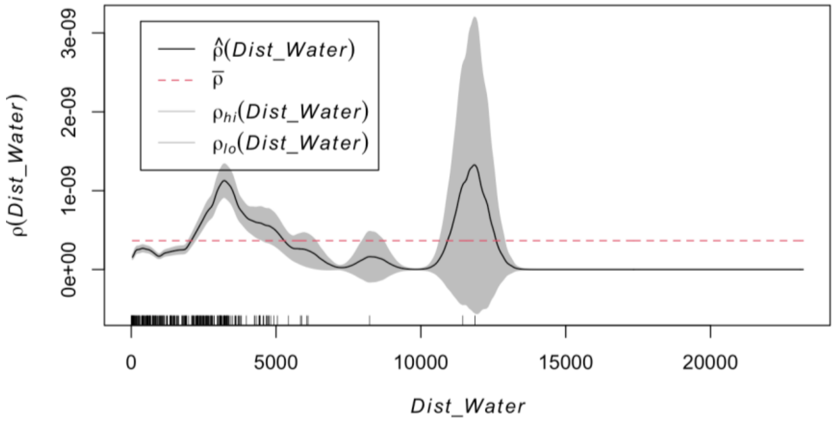


Figure 9 regarding the distance of coyote occurrences from water sources appears to indicate a significant peak in intensity at a certain distance, after which the sightings decrease. This suggests that there is a specific range from water sources where coyotes are most frequently observed. The relationship does not seem to be linear or simply exponential. The significant width of the confidence interval in the latter part of the plot for the distance from water indicates there is considerable uncertainty in the intensity estimates at greater distances.

3.3 Second-moment Description - Relationships between Points

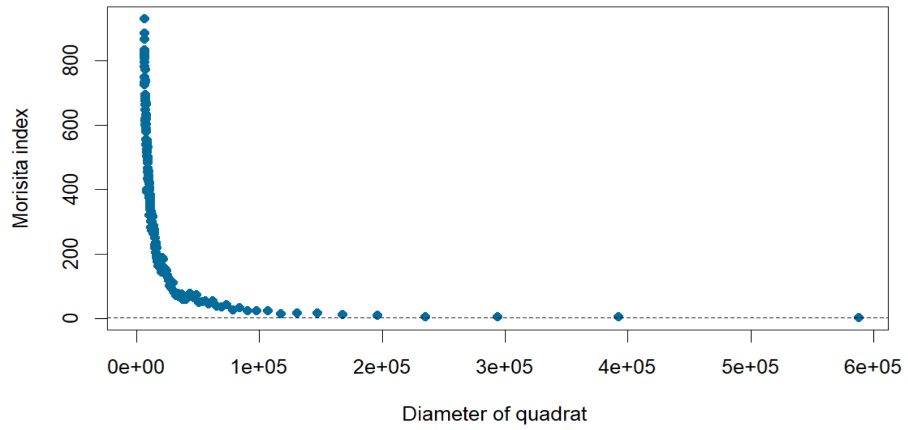
The spatial intensity of a point process, such as the locations where coyotes are observed, informs us about the expected number of occurrences at any given spot but doesn't explain the relationships between these occurrences. Coyotes' observed points might cluster, avoid each other, or be placed independently, indicating patterns that could stem from external factors affecting the distribution of coyotes, or from correlations among the points themselves. Next, we will analyze the relationships between points under both homogeneous and inhomogeneous assumptions.

3.3.1 Under Assumption of Homogeneity

(1) Morisita's index

Under the assumption of homogeneity, we can observe clear clustering phenomena, especially within a radius of 100,000 meters.

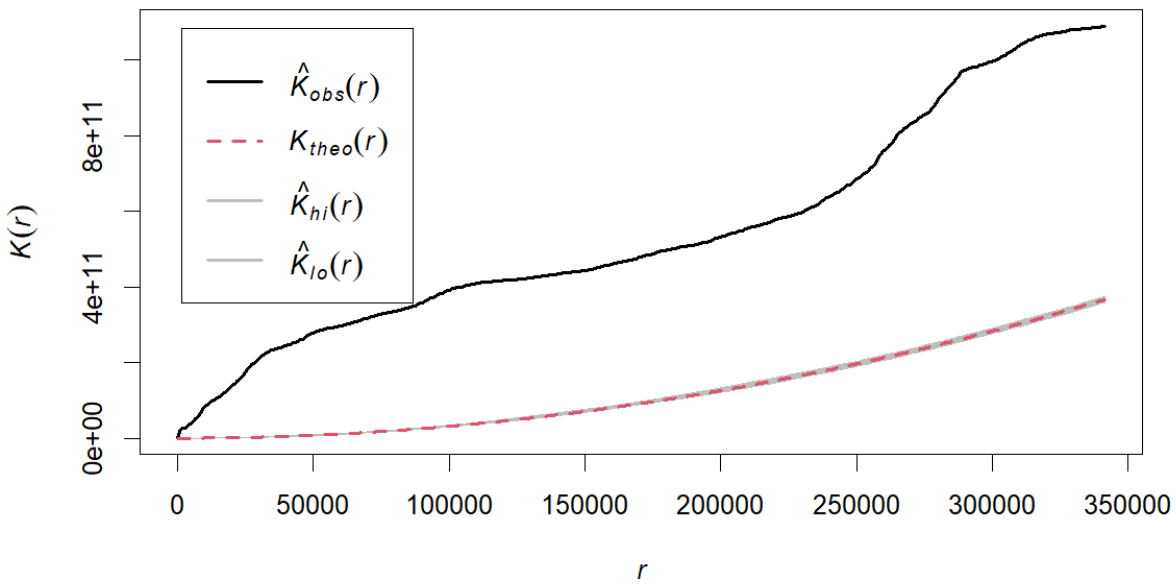
Figure 10 Moristia's Index



(2) Ripley's K-function

According to the results from Ripley's K-function, the distribution of coyotes also shows significant clustering phenomena.

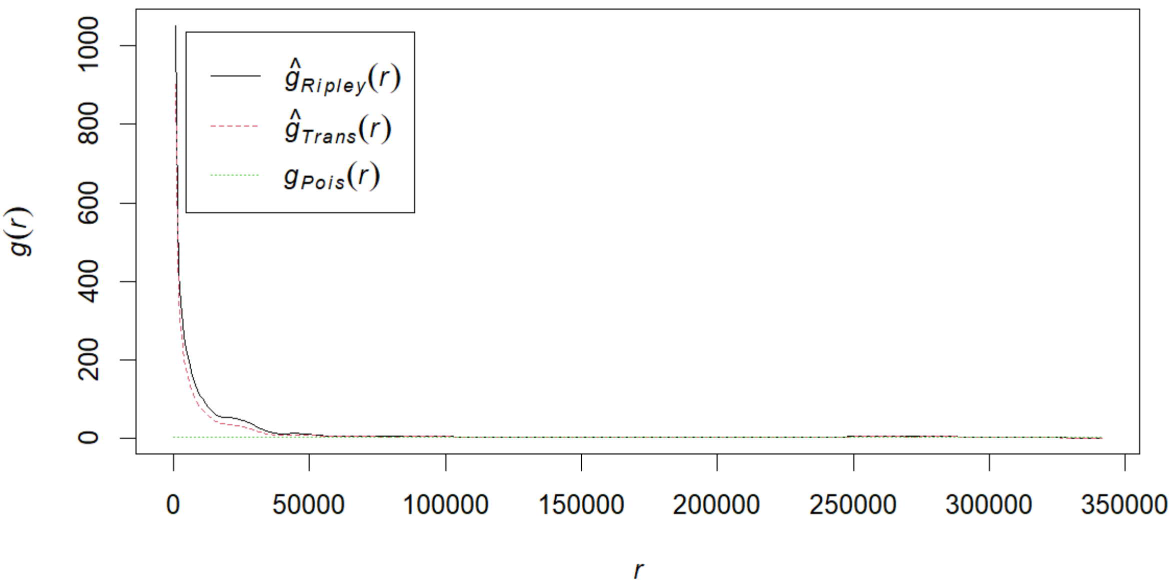
Figure 11. Ripley's K-function (Homogeneity)



(3) Pair Correlation Function

The Pair Correlation Function also shows clustering effects, which are mainly concentrated within 50,000 meters.

Figure 12. Pair Correlation Function (Homogeneity)

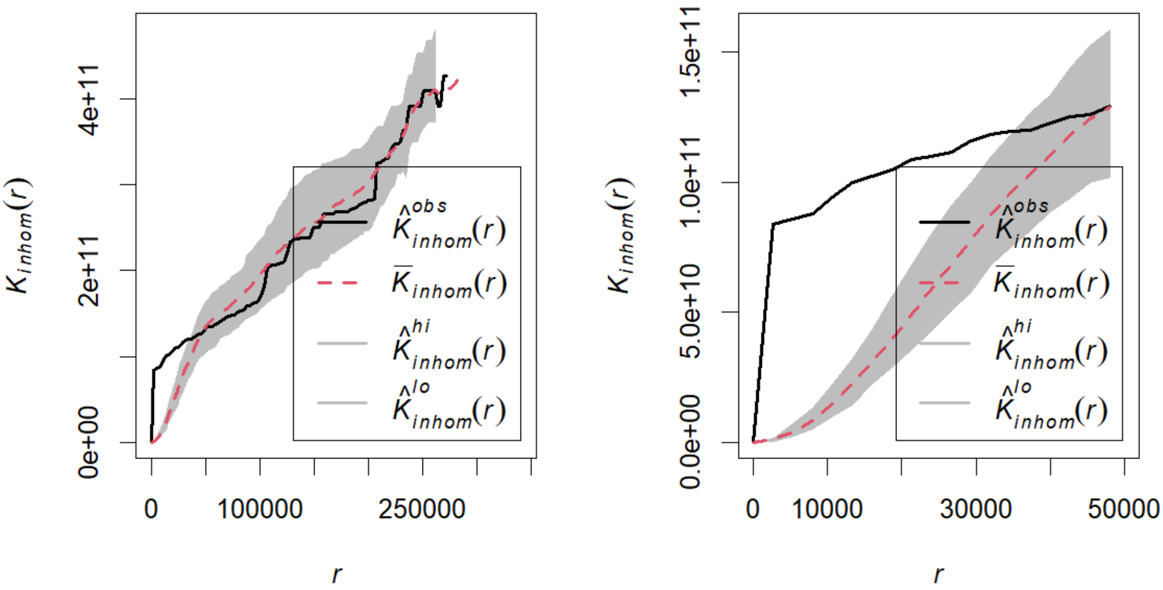


3.3.2 Under Assumption of Inhomogeneity

(1) Ripley's K-function

Ripley's K-function indicates that, after incorporating the assumption of inhomogeneity, significant clustering only appears to exist within around 30000 meters. This can be more clearly seen in the zoomed-in graph on the right side.

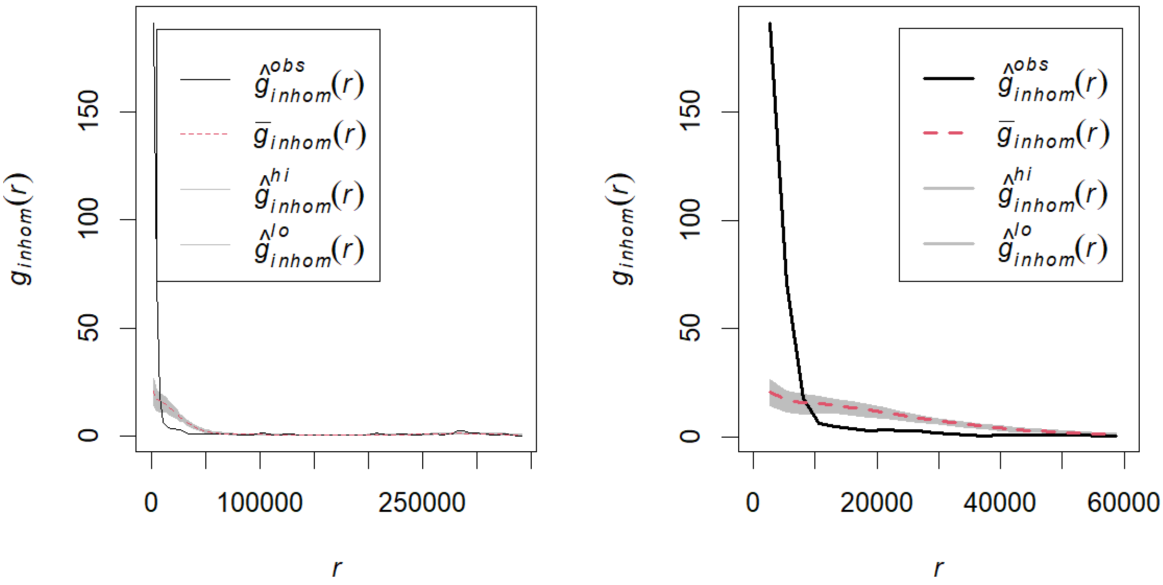
Figure 13. Ripley's K-function (Imhomogeneity)



(2) Pair Correlation Function

From the pair correlation function, it is evident that clustering is more significant within 10,000 meters. Between 10,000 and 60,000 meters, however, there is an indication of avoidance.

Figure 14. Pair Correlation Function (Imhomogeneity)



Given the complex and varied landscapes in British Columbia, along with the adaptive behaviors of coyotes that respond to these environmental variables, the inhomogeneous assumption likely provides more actionable and reliable insights for managing and conserving coyote populations effectively. It accommodates the dynamic interactions between the coyotes and their diverse habitats, thereby offering a more nuanced and applicable understanding of their spatial patterns.

3.4 Poisson Point Process model

3.4.1 Collinearity Check

Before fitting the model, we checked for multicollinearity to ensure that our covariates were not exhibiting strong correlations, which could distort model results. The analysis of the data did not reveal significant multicollinearity. This allows us to proceed with modeling, confident that each covariate can provide distinct and valuable information about the factors influencing the distribution of coyotes. Subsequently, we can fit our models without the risk of inflating the variance in our estimations due to inter-correlated predictors.

	Elevation	Forest	HFI	Dist_Water
Elevation	1.00000	-0.26225376	-0.26625626	-0.03493453
Forest	-0.26225376	1.00000	0.06618592	0.04818598
HFI	0.26625626	0.06618592	1.00000	0.13246899
Dist_Water	-0.03493453	0.04818598	0.13246899	1.00000

Table 2. Correlation Matrix

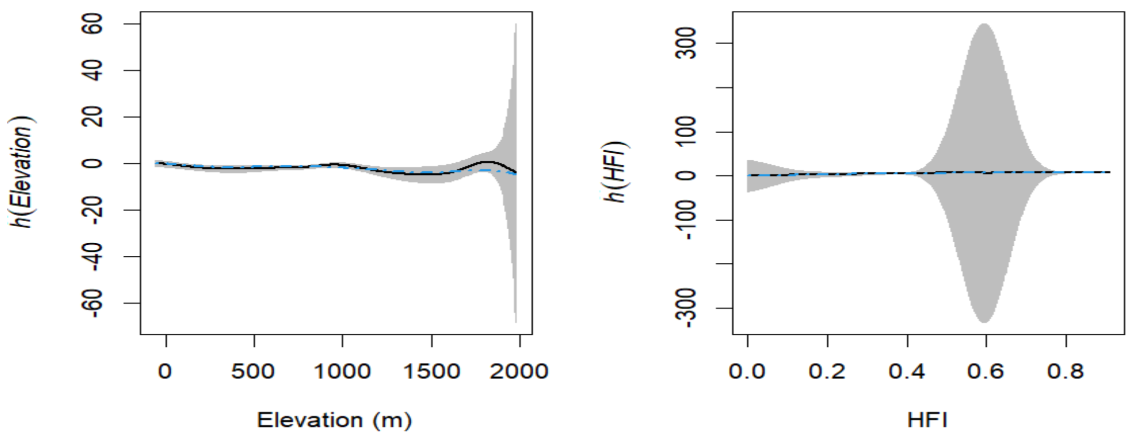
3.4.1 Model fitting

After discussing the distribution of coyotes and the influence of various environmental factors, we sought to develop a predictive model for their distribution. We initially created a Poisson Point Process model for coyote distribution based on environmental variables in BC.

From the model's Z-test value, we found that the variables 'Forest' and 'Distance from Water' lacked sufficient explanatory power. And the effects of 'HFI' and 'elevation' could be significantly enhanced by adding their higher-level polynomial terms into our model. We removed 'Distance from Water', employed a Generalized Additive Model (GAM) to capture the nuanced effects of 'elevation' and 'HFI' on coyote distribution more effectively, and tuned hyperparameters utilizing AIC and partial residuals.

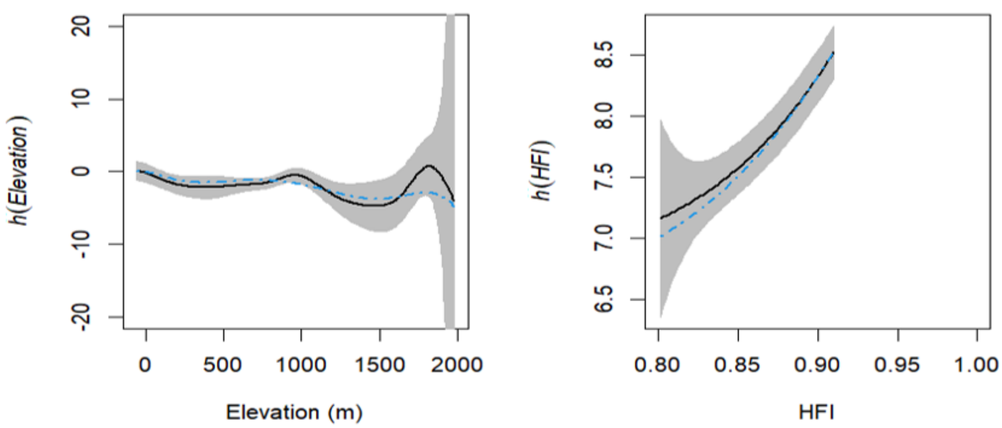
In the final model, we used bs(HFI, degree = 4, df = 4) and bs(Elevation, knots = c(300, 1400), degree = 4, df = 3) to fit the impact of HFI and elevation on the distribution of coyotes. Since the distribution changes of coyotes are concentrated in areas where HFI > 0.8, we introduced a dummy variable, I(HFI > 0.8), to enhance the influence of this range, aligning with the actual distribution patterns of coyotes. The model performed well in fitting the original data according to the partial residual checks.

Figure 15. Patial Residual Plot of Elevation(Left) and HFI(Right)



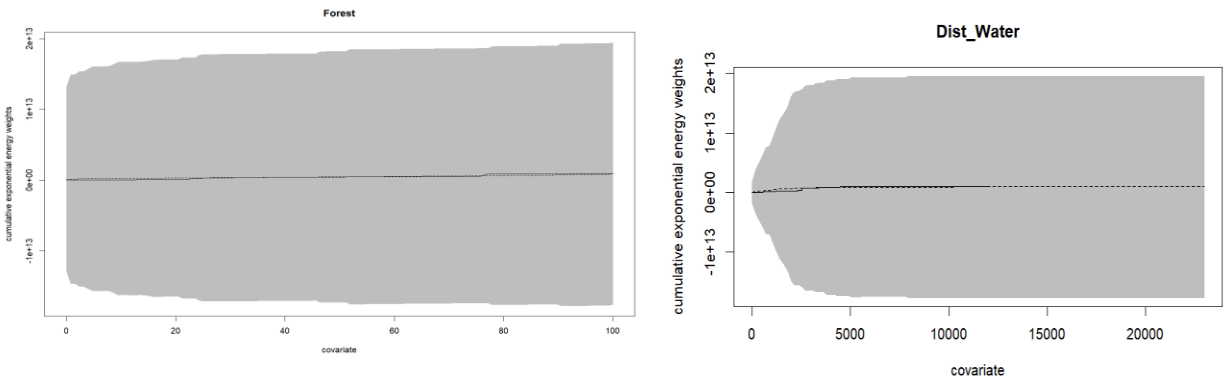
Since there are almost no data samples when HFI < 0.8, we focused on amplifying the area where HFI > 0.8 to examine the fit.

Figure 16. Patial Residual Plot (HFI>0.8) of Elevation(Left) and HFI(Right)



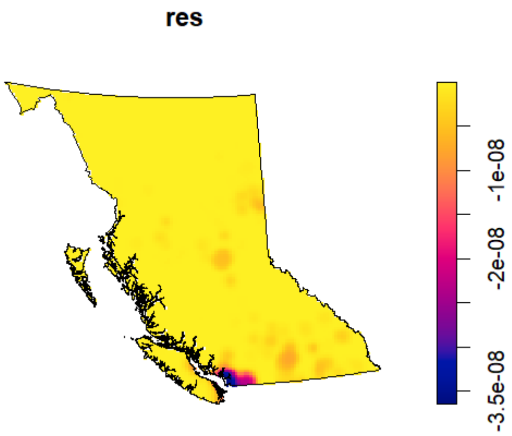
From figure 16 shown above, our model fits the data well. Additionally, our examination of lurking variables following the removal of 'Forest' and 'Distance from Water' confirmed that the model was not affected by lurking variables. The variables we selected, 'Elevation' and 'HFI', are persuasive independent variables for predicting the distribution of coyotes.

Figure 17 Examination of Lurking Variables



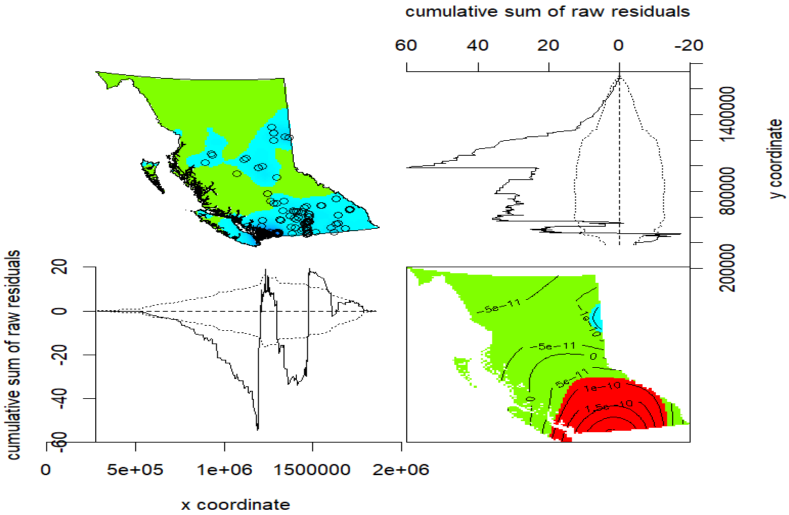
Our residual plot (figure 17) shows that the model accurately predicts the overall distribution of coyotes in British Columbia. However, there is a relatively larger estimation error along the southwestern coast of the province.

Figure 18. Residual Plot



From the intensity plot of coyotes, we got that coyotes are relatively more prevalent in the southwestern region of British Columbia. We utilized diagnostic tools to correlate the significant errors with geographical coordinates to explore why areas with higher coyote data density exhibited larger estimation errors. We found that the coordinates with larger errors correspond to major southern cities in BC, such as Vancouver and Kelowna. Our model's larger residual in the south coast area might stem from data sampling bias. Most coyote data rely on human sighting reports, leading to more recorded sightings in high Human Footprint Index (HFI) areas, while less populated areas have fewer records, thus revealing a sampling bias. Additionally, the impact of geographical settings, especially city distribution, results in an uneven distribution of coyotes. Therefore, incorporating more scientifically sampled coyote data and introducing new variables in future studies might further enhance our model's performance.

Figure 19. Cumulative Residuals



In summary, our model effectively predicts the overall trend of coyote distribution in British Columbia, which shows a strong correlation between elevation and the Human Footprint Index (HFI). However, the influences of forest coverage and proximity to water sources are not significantly represented. The distribution of coyotes in BC is likely affected by urban development and other anthropogenic factors. For future research, incorporating more diverse data sources, especially those not limited to sighting reports, and adding more environmental factors might enhance our model's performance.

4. Discussion

Our research into the distribution of coyotes in British Columbia has provided multiple insights:

Clustering Patterns: We found that coyotes exhibit significant clustering, especially within areas characterized by lower elevations and high human activity. This behaviour aligns with the animals' adaptability to urban environments.

Influence of Environmental Factors: Preliminary analyses suggested that all four environmental factors—elevation, forest cover, HFI, and distance to water—might be related to coyote distribution. However, after more rigorous modeling, it became apparent that elevation and the Human Footprint Index (HFI) were the more critical factors, showing a stronger influence on coyote presence compared to forest cover and proximity to water bodies.

Model Performance: Our models revealed specific trends in coyote distribution across British Columbia. Predominantly, coyotes are found more frequently in areas with lower elevations and high human activity—traits characteristic of urbanized environments. This suggests that coyotes are not only adapting to but possibly thriving in such areas due to the availability of resources and lesser predation threats compared to more remote, higher-elevation locales. The Generalized Additive Models (GAMs) employed allowed for a detailed understanding of these trends, showing how elevation and human presence play crucial roles in shaping coyote habitats.

Data Biases and Limitations: Our findings suggest a possible sampling bias, as data collection concentrated in urban areas with high human presence might have overshadowed other regions. This skew could impact the accuracy of predictions about rural coyote populations.

Future Directions: To enhance our models' accuracy and comprehensiveness, incorporating a broader range of environmental variables and a more diverse dataset—especially from less populated areas—will be crucial. This approach will help mitigate biases and provide a more balanced view of coyote distribution across different landscapes.

These insights underscore the complex interactions between coyotes and their environments in British Columbia, highlighting the need for continued research to understand the implications of urban expansion and ecological changes on wildlife distribution.

References:

1. https://www.gbif.org/species/5219153

2. Alexander, S.M. and M.S. Quinn. 2011." Coyote (Canis latrans) interactions with humans and pets reported in the Canadian print media (1995–2010)." Human Dimensions of Wildlife 16:345-359.

3. Spatstat, https://spatstat.org/.

4. Poessel, S. A., Gese, E. M., & Young, J. K. 2017b. Environmental factors influencing the occurrence of coyotes and conflicts in urban areas. Landscape and Urban Planning, 157, 259–269.