Spatial Analysis of Gray Wolf (*Canis lupus*) Occurrences in British Columbia

Suneet D'Silva, Mohammed Bayati

2025-04-16

Contents

1 References

1. Muñoz-Fuentes, V., Darimont, C. T., Paquet, P. C., & Leonard, J. A. (2009). Ecological factors drive differentiation in wolves from British Columbia. *Journal of Biogeography*, 36(8), 1516–1531.

2. British Columbia Ministry of Environment. (2014). *Management Plan for the Grey Wolf (Canis lupus) in British Columbia*. https://www2.gov.bc.ca/assets/gov/environment/plants-animals-and-ecosystems/wildlife-wildlife-habitat/management_plan_for_the_grey_wolf_in_british_columbia.pdf

3. Darimont, C. T., Reimchen, T. E., Bryan, H. M., & Paquet, P. C. (2017). Resource selection by coastal wolves reveals seasonal importance of mature forest and open land cover. *Forest Ecology and Management*, 405, 56–65.

4. GBIF.org (2025). GBIF Occurrence Download for *Canis lupus* in British Columbia. https://www.gbif.org/occurrence/search?taxon_key=5219173&state_province=British%20Columbia&has_coordinate=true

2 Introduction

2.1 Species Background

Wolves (Canis lupus) are apex predators found throughout much of British Columbia, inhabiting ecosystems ranging from coastal rainforests to alpine regions. The province is home to both interior and coastal ecotypes, with coastal wolves—often referred to as "rainforest wolves"—exhibiting distinct traits such as smaller body size, reddish fur, and partial reliance on marine resources[1]. These wolves play a keystone role in regulating prey populations, including deer, moose, and caribou, and contribute to broader ecological processes through trophic cascades[2].

2.2 Project Motivation and

Understanding the spatial distribution of wolves in BC is important for both ecological and management purposes. Wolves are sensitive to environmental gradients such as elevation, forest cover, and human disturbance, all of which influence habitat suitability and movement patterns[3]. This project aims to explore these spatial relationships using point pattern analysis. By modeling wolf occurrence data against environmental variables, we hope to uncover meaningful spatial patterns that reflect ecological behaviour and landscape constraints.

2.3 Research Questions

1. **Do Gray Wolf observations in British Columbia exhibit spatial clustering across the landscape?**

2. **Are these spatial patterns influenced by environmental variables such as elevation and forest cover?**

We hypothesize that wolf occurrences will be spatially clustered, with higher intensity in low-elevation, densely forested areas, consistent with known habitat preferences.

3 3. Methods

3.1 Data Sources

Wolf occurrence data were obtained using the `rgbif` package in R, querying the GBIF database for *Canis lupus* observations in British Columbia. Only records with valid geographic coordinates were retained for analysis.

3.2 Preprocessing and Conversion

Occurrence records were cleaned to remove missing coordinates and converted into an `sf` spatial object using WGS84 projection. The spatial data were then transformed into a ppp point pattern object for spatial analysis. Environmental rasters were cropped and resampled to match the observation window, then converted to `im` objects for model fitting.

3.3 Spatial Analysis Workflow

Exploratory VIsualizaiton. KDE's. Rpleys test… then fitting PPP model. Answering First Research Question Spatial Clustering

Exploring Elevation and Forest coverage. Then overlaying Forest coverage and wolves. Answering Second Research Question about correlation of forest elevation and the wolves.

```r
# 4. Define the geographic filters
  # GBIF uses ISO 3166-1 alpha-2 country codes (CA for Canada)
  target_country <- "CA"
  # Use the full province name as GBIF generally expects
  target_province <- "British Columbia"

  # 5. Perform the occ_search with geographic filters
  print(paste("Searching for Canis lupus occurrences in", target_province, ",", target_country
```

```
## [1] "Searching for Canis lupus occurrences in British Columbia , CA ..."
```

```r
# Increase the limit if you expect more records, but be mindful of occ_search limitations.
  # hasCoordinate = TRUE is often useful for mapping/spatial analysis.
  # return = "data" gives you the data frame directly.
  search_result_list <- occ_search(
    taxonKey = species_key,
    country = target_country,
    stateProvince = target_province,
    limit = 100000,         # Adjust this limit based on expected results / needs
    hasCoordinate = TRUE, # Optional: Get only records with coordinates
    # You could add other filters like year, basisOfRecord etc.
    # basisOfRecord = "OBSERVATION" # Example: only observations
    # year = "2020,2023"            # Example: only from 2020-2023
  )
```

```r
# Check column names
colnames(lupus_bc_search_df)
```

```
##   [1] "key"
##   [2] "scientificName"
##   [3] "decimalLatitude"
##   [4] "decimalLongitude"
##   [5] "issues"
##   [6] "datasetKey"
##   [7] "publishingOrgKey"
##   [8] "installationKey"
##   [9] "hostingOrganizationKey"
##  [10] "publishingCountry"
##  [11] "protocol"
```

```
## [12] "lastCrawled"
## [13] "lastParsed"
## [14] "crawlId"
## [15] "basisOfRecord"
## [16] "occurrenceStatus"
## [17] "taxonKey"
## [18] "kingdomKey"
## [19] "phylumKey"
## [20] "classKey"
## [21] "orderKey"
## [22] "familyKey"
## [23] "genusKey"
## [24] "speciesKey"
## [25] "acceptedTaxonKey"
## [26] "acceptedScientificName"
## [27] "kingdom"
## [28] "phylum"
## [29] "order"
## [30] "family"
## [31] "genus"
## [32] "species"
## [33] "genericName"
## [34] "specificEpithet"
## [35] "taxonRank"
## [36] "taxonomicStatus"
## [37] "iucnRedListCategory"
## [38] "dateIdentified"
## [39] "coordinateUncertaintyInMeters"
## [40] "continent"
## [41] "stateProvince"
## [42] "year"
## [43] "month"
## [44] "day"
## [45] "eventDate"
## [46] "startDayOfYear"
## [47] "endDayOfYear"
## [48] "modified"
## [49] "lastInterpreted"
## [50] "references"
```

```
## [51] "license"
## [52] "isSequenced"
## [53] "identifier"
## [54] "facts"
## [55] "relations"
## [56] "isInCluster"
## [57] "datasetName"
## [58] "recordedBy"
## [59] "identifiedBy"
## [60] "dnaSequenceID"
## [61] "geodeticDatum"
## [62] "class"
## [63] "countryCode"
## [64] "recordedByIDs"
## [65] "identifiedByIDs"
## [66] "gbifRegion"
## [67] "country"
## [68] "publishedByGbifRegion"
## [69] "rightsHolder"
## [70] "identifier.1"
## [71] "http...unknown.org.nick"
## [72] "verbatimEventDate"
## [73] "dynamicProperties"
## [74] "verbatimLocality"
## [75] "collectionCode"
## [76] "gbifID"
## [77] "occurrenceID"
## [78] "taxonID"
## [79] "http...unknown.org.captive_cultivated"
## [80] "catalogNumber"
## [81] "institutionCode"
## [82] "eventTime"
## [83] "identificationID"
## [84] "name"
## [85] "occurrenceRemarks"
## [86] "projectId"
## [87] "vitality"
## [88] "infraspecificEpithet"
## [89] "identificationRemarks"
```

```
##  [90] "sex"
##  [91] "lifeStage"
##  [92] "informationWithheld"
##  [93] "recordedByIDs.type"
##  [94] "recordedByIDs.value"
##  [95] "identifiedByIDs.type"
##  [96] "identifiedByIDs.value"
##  [97] "gadm"
##  [98] "networkKeys"
##  [99] "institutionKey"
## [100] "collectionKey"
## [101] "preparations"
## [102] "recordNumber"
## [103] "georeferenceProtocol"
## [104] "georeferenceVerificationStatus"
## [105] "locality"
## [106] "county"
## [107] "verbatimCoordinateSystem"
## [108] "disposition"
## [109] "fieldNotes"
## [110] "verbatimElevation"
## [111] "georeferenceSources"
## [112] "higherGeography"
## [113] "georeferencedBy"
## [114] "associatedSequences"
## [115] "otherCatalogNumbers"
## [116] "samplingProtocol"
## [117] "institutionID"
## [118] "language"
## [119] "type"
## [120] "locationAccordingTo"
## [121] "georeferencedDate"
## [122] "nomenclaturalCode"
## [123] "footprintWKT"
## [124] "organismID"
## [125] "previousIdentifications"
## [126] "identificationQualifier"
## [127] "accessRights"
## [128] "higherClassification"
```

```
## [129] "collectionID"
## [130] "ownerInstitutionCode"
## [131] "bibliographicCitation"
## [132] "individualCount"
## [133] "establishmentMeans"
## [134] "municipality"
## [135] "coordinatePrecision"
## [136] "higherGeographyID"
## [137] "locationRemarks"
## [138] "datasetID"
## [139] "fieldNumber"
## [140] "eventRemarks"
## [141] "geometry"
```

```r
# Look at the first few rows of coordinate columns (replace names if different)
head(lupus_bc_search_df[, c("decimalLongitude", "decimalLatitude")])
```

```
## Simple feature collection with 6 features and 2 fields
## Geometry type: POINT
## Dimension:     XY
## Bounding box:  xmin: -129.1466 ymin: 48.34157 xmax: -116.9289 ymax: 59.80141
## Geodetic CRS:  WGS 84
## # A tibble: 6 x 3
##   decimalLongitude decimalLatitude             geometry
##              <dbl>           <dbl>          <POINT [°]>
## 1            -126.            49.1 (-125.7105 49.08471)
## 2            -124.            48.3 (-123.5467 48.34198)
## 3            -124.            48.3 (-123.5503 48.34157)
## 4            -117.            49.7 (-116.9289 49.65639)
## 5            -129.            59.8 (-129.1413 59.80141)
## 6            -129.            59.7 (-129.1466 59.74125)
```

```r
# Check for missing coordinate data
sum(is.na(lupus_bc_search_df$decimalLongitude))
```

```
## [1] 0
```

```r
sum(is.na(lupus_bc_search_df$decimalLatitude))
```

```
## [1] 0
```

```r
# Remove rows with missing coordinates first (important!)
lupus_bc_search_df <- lupus_bc_search_df %>%
  filter(!is.na(decimalLongitude) & !is.na(decimalLatitude))

# Convert to an sf object
# Ensure coords are in the order: X (Longitude), Y (Latitude)
# GBIF data typically uses WGS84 (EPSG:4326) coordinate reference system (CRS)
lupus_bc_search_df <- st_as_sf(lupus_bc_search_df,
                       coords = c("decimalLongitude", "decimalLatitude"),
                       crs = 4326, # WGS84 CRS
                       remove = FALSE) # Keep original coordinate columns if desired

# Check the result
print(lupus_bc_search_df)
```

```
## Simple feature collection with 458 features and 140 fields
## Geometry type: POINT
## Dimension:     XY
## Bounding box:  xmin: -137.3242 ymin: 48.23257 xmax: -114.7851 ymax: 60.16684
## Geodetic CRS:  WGS 84
## # A tibble: 458 x 141
##    key        scientificName  decimalLatitude decimalLongitude issues datasetKey
##  * <chr>      <chr>                     <dbl>            <dbl> <chr>  <chr>
##  1 5063264735 Canis lupus Li~            49.1            -126. cdc,c~ 50c9509d-~
##  2 5063640978 Canis lupus Li~            48.3            -124. cdc    50c9509d-~
##  3 5063732354 Canis lupus Li~            48.3            -124. cdc,c~ 50c9509d-~
##  4 5087224829 Canis familiar~            49.7            -117. cdc,c~ 50c9509d-~
##  5 5087375311 Canis lupus Li~            59.8            -129. cdc    50c9509d-~
##  6 5087785191 Canis lupus Li~            59.7            -129. cdc,c~ 50c9509d-~
##  7 5088190704 Canis lupus cr~            49.4            -126. cdc,c~ 50c9509d-~
##  8 5087982149 Canis lupus cr~            49.4            -126. cdc,c~ 50c9509d-~
##  9 5087829289 Canis lupus cr~            49.4            -126. cdc,c~ 50c9509d-~
## 10 5088121358 Canis lupus Li~            49.4            -126. cdc,c~ 50c9509d-~
## # i 448 more rows
## # i 135 more variables: publishingOrgKey <chr>, installationKey <chr>,
```

```
## #   hostingOrganizationKey <chr>, publishingCountry <chr>, protocol <chr>,
## #   lastCrawled <chr>, lastParsed <chr>, crawlId <int>, basisOfRecord <chr>,
## #   occurrenceStatus <chr>, taxonKey <int>, kingdomKey <int>, phylumKey <int>,
## #   classKey <int>, orderKey <int>, familyKey <int>, genusKey <int>,
## #   speciesKey <int>, acceptedTaxonKey <int>, acceptedScientificName <chr>, ...
```

```
class(lupus_bc_search_df) # Should include "sf" and "data.frame"
```

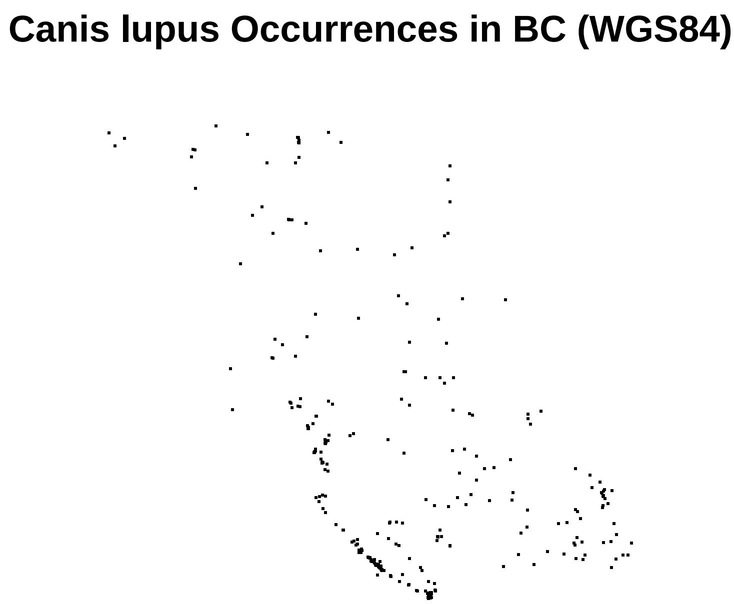
```
## [1] "sf"         "tbl_df"     "tbl"        "data.frame"
```

```
st_crs(lupus_bc_search_df) # Verify the CRS is set correctly
```

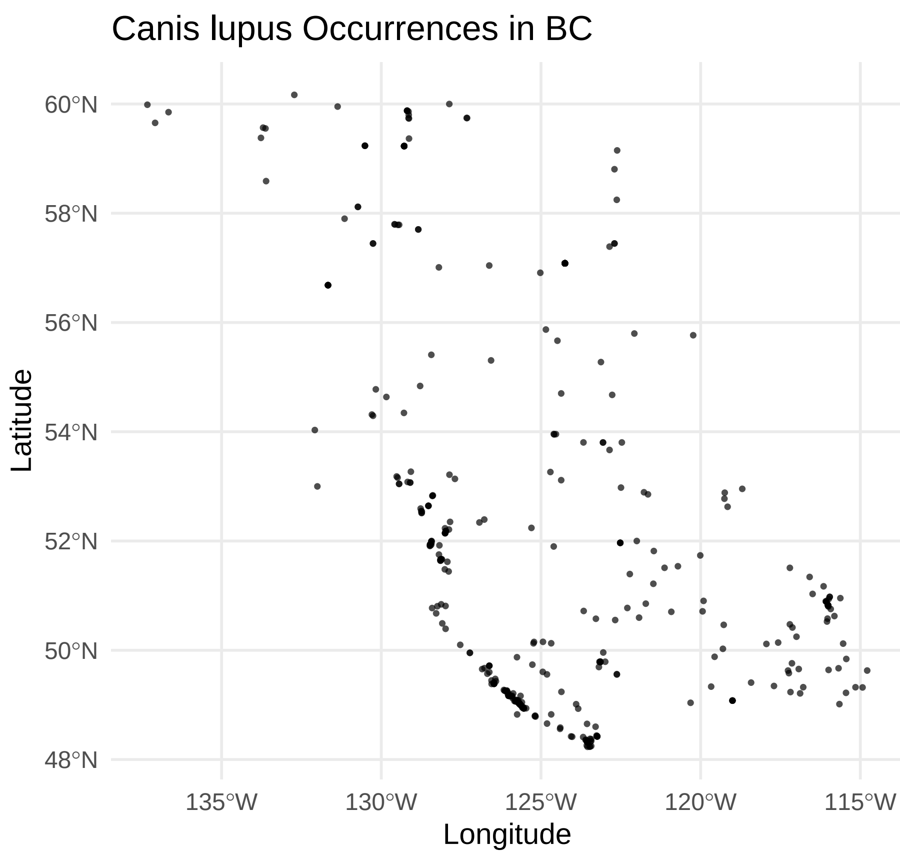
```
## Coordinate Reference System:
##   User input: EPSG:4326
##   wkt:
## GEOGCRS["WGS 84",
##     ENSEMBLE["World Geodetic System 1984 ensemble",
##         MEMBER["World Geodetic System 1984 (Transit)"],
##         MEMBER["World Geodetic System 1984 (G730)"],
##         MEMBER["World Geodetic System 1984 (G873)"],
##         MEMBER["World Geodetic System 1984 (G1150)"],
##         MEMBER["World Geodetic System 1984 (G1674)"],
##         MEMBER["World Geodetic System 1984 (G1762)"],
##         MEMBER["World Geodetic System 1984 (G2139)"],
##         MEMBER["World Geodetic System 1984 (G2296)"],
##         ELLIPSOID["WGS 84",6378137,298.257223563,
##             LENGTHUNIT["metre",1]],
##         ENSEMBLEACCURACY[2.0]],
##     PRIMEM["Greenwich",0,
##         ANGLEUNIT["degree",0.0174532925199433]],
##     CS[ellipsoidal,2],
##         AXIS["geodetic latitude (Lat)",north,
##             ORDER[1],
##             ANGLEUNIT["degree",0.0174532925199433]],
##         AXIS["geodetic longitude (Lon)",east,
##             ORDER[2],
##             ANGLEUNIT["degree",0.0174532925199433]],
##     USAGE[
##         SCOPE["Horizontal component of 3D system."],
```

```
##         AREA["World."],
##         BBOX[-90,-180,90,180]],
##     ID["EPSG",4326]]
```

```r
 # Simple plot using sf's default plot method
plot(st_geometry(lupus_bc_search_df), pch = ".", main = "Canis lupus Occurrences in BC (WGS84)"
```

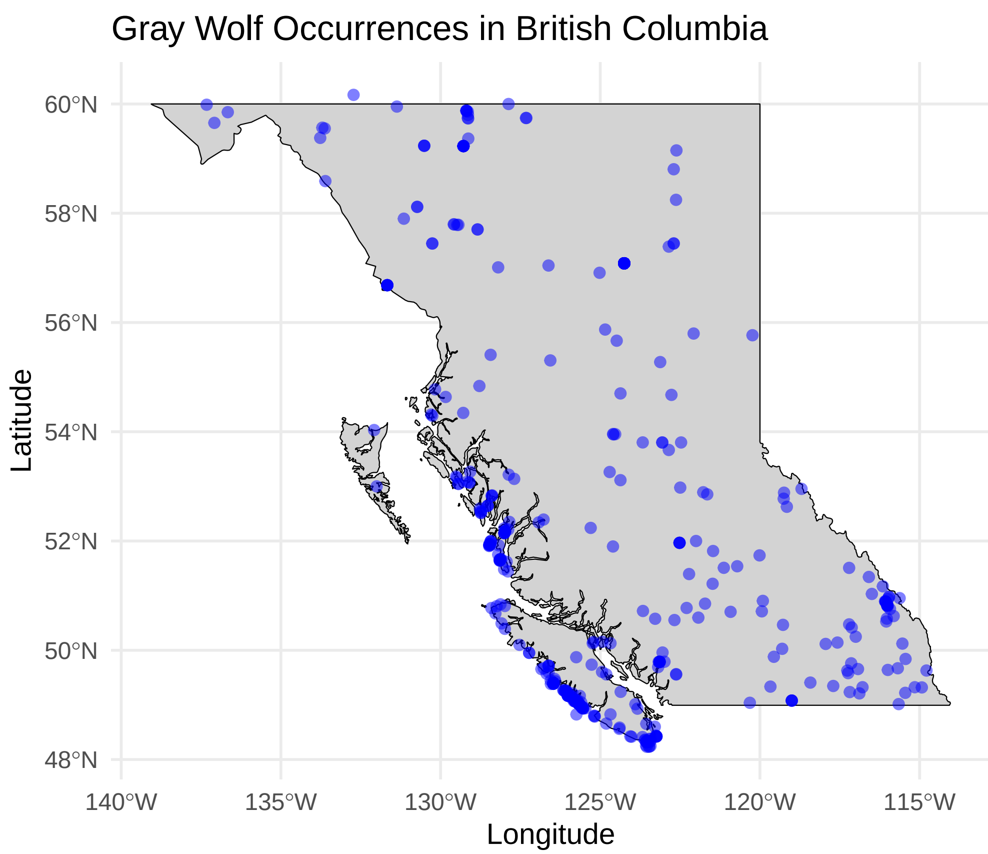



```r
# Nicer plot using ggplot2
ggplot() +
  geom_sf(data = lupus_bc_search_df, size = 0.5, alpha = 0.7) + # Plot the points
  labs(title = "Canis lupus Occurrences in BC",
       x = "Longitude", y = "Latitude") +
  theme_minimal() # Use a minimal theme
```




```r
# Load British Columbia shapefile from rnaturalearth
bc_shapefile <- ne_states(country = "Canada", returnclass = "sf") %>%
  filter(name == "British Columbia")

# Plot Gray Wolf occurrences with BC map overlay
ggplot() +
  # Plot the BC boundary (light gray fill)
  geom_sf(data = bc_shapefile, fill = "lightgray", color = "black") +
  # Plot Gray Wolf occurrence points
  geom_point(data = lupus_bc_search_df, aes(x = decimalLongitude, y = decimalLatitude),
             alpha = 0.5, color = "blue") +
  labs(title = "Gray Wolf Occurrences in British Columbia",
       x = "Longitude", y = "Latitude") +
  theme_minimal() +
  theme(legend.position = "none")  # Remove legend if not needed
```




```r
library(mapview)

mapview(lupus_bc_search_df)
```

4 4. Exploratory Data Analysis

4.1 Spatial Distribution of Observations

Figure 1 shows the spatial distribution of Gray Wolf observations across British Columbia. The data appear to be unevenly distributed, with notable concentrations in the southern and central interior regions, as well as along portions of the coast. Large gaps in the northern interior and far southeastern areas may reflect either true absence, lower survey effort, or limited accessibility. These patterns suggest potential clustering that warrants further investigation through formal spatial analysis.

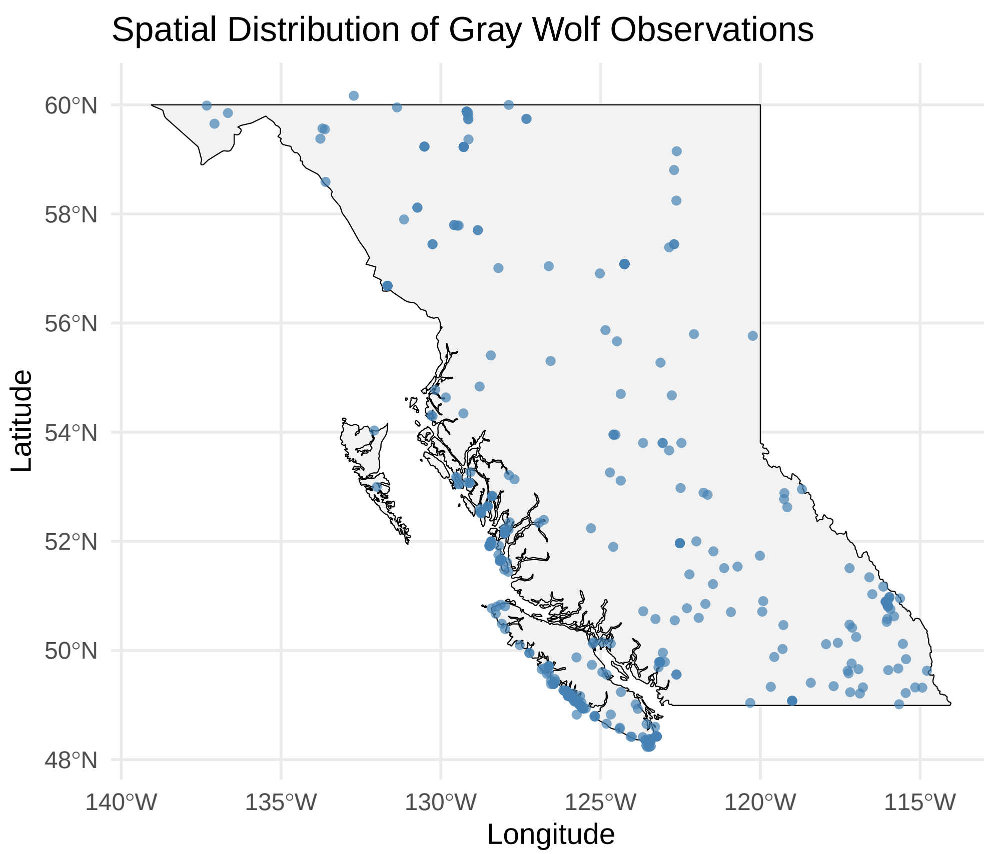


Figure 1: **Figure 1.** Spatial distribution of Gray Wolf (*Canis lupus*) observations in British Columbia. Points represent GBIF occurrence records with valid coordinates.

4.2 Kernel Density Estimation (KDE)

4.3 Quadrat Count or Ripley's K-Test

4.4 Elevation / Forest Cover Overlay

5 5. Results

6 5. Results

7 6. Discussion

7.1 6.1 Summary of Key Findings

- Were wolves spatially clustered? (based on KDE + Quadrat test)
- Were those patterns influenced by elevation or forest cover?
- Did your results support your hypothesis?

7.2 6.2 Interpretation

- Why might wolves cluster in certain areas?
- Are the patterns ecological, observational, or both?
- Could gaps be due to accessibility, observer effort, or habitat?

7.3 6.3 Limitations

- GBIF data may be biased toward accessible areas
- Resolution of environmental rasters
- Modeling assumptions (e.g., CSR, IPPM)