Canada-Lynx-Spatial-Analysis

Avishek Saha, Jonah Edmundson, Madison Greenough

2023-04-10

Introduction

The Canadian Lynx, known scientifically as the *Lynx canadensis Kerr*, are a marvelous species found within British Columbia. For this report, a spatial analysis of this species will be conducted. They will be referred to as their shorter name, lynx, for the duration of this report. Given that the lynx are a part of the feline family, we hypothesize that they will habitate in forested areas, in small groups, with minimal movement. In other words, we expect that they will reside in an area long term without significant travelling or relocating. We also predict that they will reside in colder locations, so their presence may not be found in the warmer areas of BC, such as the southern interior. We expect that we cannot assume homogeneity for this specific species, as we expect the lynx to be found in clusters rather than evenly distributed across BC. We expect that the lynx would be found at higher elevations due to their thicker coats and natural behaviours.

Figure 1: See above an image of two lynx found in BC.

Methods

Dataset

The dataset contains 13 variables and there are 178 rows, or entries. The variables include `countryCode` (str, which have all been filtered to CA for Canada), `stateProvince` (str, filtered for British Columbia), `higherClassification` (str, additional species information). `month` (int, month of occurrence), `locality` (str, region description), `higherGeography` (str, more broad region), `sex` (str, contains male, female, or NA), `decimalLatitude` (float), `decimalLongitude` (float), `dateIdentified` (date-time value, date of identification), `coordinateUncertaintyInMeters` (int, uncertainty value), `footprintWKT` (list of int, measurement values), and `country` (str, filtered to Canada).

Workflow

To begin, first moment descriptive statistics were visualized. Basic `spatstat` features were used to transform to a `ppp` object, as well as visualize the spatial data on the appropriate window, which in this case, is the province of BC map. The elevation was also attached to the data, and basic elevation features were visualized on the window.

After that we examined the intensity of lynx locations and assessed homogeneity through a quadrat test, which revealed that the distribution was not homogeneous. We then used inhomogeneous k-function and pair correlation function to investigate the correlation of lynx locations. Then we fitted several models and selected them based on AIC scores and likelihood ratio test. Furthermore, we performed a quadrat test to determine if the selected model accurately predicts the data or if there are deviations. Additionally, we explored the potential benefits of adding other covariates to our models. To assess the efficacy of our models, partial residual plots were used to identify whether our model capturing the data well or not.

Code

```r
lc <- read.csv("../datasets/processed/lc.csv")
suppressMessages(library(spatstat))
suppressMessages(library(sf))
load("../datasets/raw/BC_Covariates.Rda")
bc_window_sf <- st_as_sf(DATA$Window)
bc_window_owin <- as.owin(bc_window_sf)
# Convert to a ppp object
lc_ppp <- ppp(x = lc$decimalLongitude, # X coordinates
              y = lc$decimalLatitude, # Y coordinates
              window = bc_window_owin, # Observation window
              )
lc_ppp <- as.ppp(lc_ppp)
# Plot the ppp object
plot(lc_ppp,
     pch = 16,
     cex = 0.5,
     cols = "#046C9A",
     main = "Marked Canada Lynx")
intensity(lc_ppp)
Q <- quadratcount(lc_ppp,
                  nx = 2,
                  ny = 4)
plot(intensity(Q, image = T),
```

```r
     main = "Lynx intensity")
plot(lc_ppp, pch = 16, cex = 0.6, cols = "white", add = T)
plot(lc_ppp, pch = 16, cex = 0.5, cols = "black", add = T)
#plotting on BC map
library(raster)
dmap1 <- density.ppp(lc_ppp, sigma = bw.ppl(lc_ppp),edge=T)
r1 <- raster(dmap1)
library(leaflet)
#make sure we have right CRS, which in this case is British National Grid
bc_crs <- "+proj=aea +lat_0=45 +lon_0=-126 +lat_1=50 +lat_2=58.5 +x_0=1000000 +y_0=0 +datum=NAD83 +units
crs(r1) <- sp::CRS(bc_crs)
# create a colour palet
pal <- colorNumeric(c("#41B6C4", "red"), values(r1),
  na.color = "transparent")
#and then make map
leaflet() %>%
  addTiles() %>%
  addRasterImage(r1, colors = pal, opacity = 0.8) %>%
  addLegend(pal = pal, values = values(r1),
    title = "Canada Lynx in BC")
quadrat.test(Q)
col_pal <- colorRampPalette(c('orange', 'green'))
# Calculate nearest neighbor distances
nn_dist <- data.frame(nndist(lc_ppp))
# Add distance to nearest neighbor mark
marks(lc_ppp) <- nn_dist
# Plot point pattern with distance to nearest neighbor mark
plot(lc_ppp, main = "Seperation Distance",
     which.marks = NULL,
     cols = col_pal(nrow(nn_dist)), #The colours of the points
     pch = 16)
# Estimate the k- function
k_lc <- Kest(lc_ppp)
# visualise the results
plot (k_lc, main = "Homogeneous K-Function", lwd = 2)
# Bootstrapped CIs
# rank = 1 means the max and min
# values will be used for CI
E_lc <- envelope(lc_ppp , Kest , rank = 1, nsim = 19, fix.n = T)
# visualise the results
plot (E_lc , main = "Homogeneous K-function")
#Estimate intensity
lambda_lc <- density(lc_ppp, bw.ppl)
Kinhom_lc <- Kinhom(lc_ppp, lambda_lc)
# Estimate a strictly positive density
lambda_lc_pos <- density(lc_ppp, sigma=bw.ppl, positive=TRUE)
# Simulation envelope (with points drawn from the estimated intensity)
E_lc_inhom <- envelope(lc_ppp,
                       Kinhom,
                       simulate = expression(rpoispp(lambda_lc_pos)),
                       correction="border",
                       rank = 1,
                       nsim = 19,
```

```r
                           fix.n = TRUE)
# visualise the results
plot(E_lc_inhom, xlim = c(0,320000), main = "Inhomogeneous K-function", lwd = 2)
#Simulation envelope (with points drawn from the estimated intensity)
pcf_lc_inhom <- envelope(lc_ppp,
                         pcfinhom,
                         simulate = expression(rpoispp(lambda_lc_pos)),
                         rank = 1,
                         nsim = 19)


# visualise the results
par(mfrow = c(1,2))
plot(pcf_lc_inhom, main = "Inhomogeneous g-function")
# Zoom in on range where significant deviations appear
plot(pcf_lc_inhom, xlim = c(0,20000), main = "", lwd = 2)
library(viridis)
library(viridisLite)
cols = terrain.colors(5)
# Create 5 elevation classes with equal width
elev_classes <- cut(DATA$Elevation,
                    breaks = 5,
                    labels = c("low", "low-medium", "medium", "high-medium", "high"))
#table(elev_classes[lc_ppp])
# Plot the elevation class image and overlay the lynx locations
# plot(elev_classes, col = cols, main = "Elevation Classes", par(bg="grey50", cex.main = 2, cex = 0.6))
# points(lc_ppp, pch = 16, cex = 0.6, col = "black")

# Create 5 forest classes with equal width
forest_classes <- cut(DATA$Forest,
                    breaks = 5,
                    labels = c("low", "low-medium", "medium", "high-medium", "high"))
#table(forest_classes[lc_ppp])
# plot(forest_classes, col = cols, main = "Forest Classes", par(bg="grey50", cex.main = 2, cex = 0.6))
# points(lc_ppp, pch = 16, cex = 0.6, col = "black")

# Create 5 human footprint index classes with equal width
hfi_classes <- cut(DATA$HFI,
                  breaks = 5,
                  labels = c("low", "low-medium", "medium", "high-medium", "high"))
#table(hfi_classes[lc_ppp])
# plot(hfi_classes, main = "Footprint Index Classes", par(bg="grey50", cex.main = 2, cex = 0.6))
# points(lc_ppp, pch = 16, cex = 0.5, col = "white")

# Create 5 dist water classes with equal width
dist_water_classes <- cut(DATA$Dist_Water,
                        breaks = 5,
                        labels = c("low", "low-medium", "medium", "high-medium", "high"))
#table(dist_water_classes[lc_ppp])
# plot(dist_water_classes, main = "Dist Water Classes", par(bg="grey50", cex.main = 2, cex = 0.6))
# points(lc_ppp, pch = 16, cex = 0.5, col = "white")

#plot all at once
par(mfrow = c(2,2))
```

```r
par(mar=c(0,0,2,1))
#1
plot(elev_classes, col = cols, main = "Elevation Classes", par(bg="grey50", cex.main = 2, cex = 0.6))
points(lc_ppp, pch = 16, cex = 0.6, col = "black")
#2
plot(forest_classes, col = cols, main = "Forest Classes", par(bg="grey50", cex.main = 2, cex = 0.6))
points(lc_ppp, pch = 16, cex = 0.6, col = "black")
#3
plot(hfi_classes, main = "Footprint Index Classes", par(bg="grey50", cex.main = 2, cex = 0.6))
points(lc_ppp, pch = 16, cex = 0.5, col = "white")
#4
plot(dist_water_classes, main = "Dist Water Classes", par(bg="grey50", cex.main = 2, cex = 0.6))
points(lc_ppp, pch = 16, cex = 0.5, col = "white")
par(mfrow = c(1,1))
#Estimate Rho
hfi = DATA$HFI
rho_hfi = rhohat(lc_ppp, hfi)
distw = DATA$Dist_Water
rho_distw = rhohat(lc_ppp, distw)
elev = DATA$Elevation
rho_elev = rhohat(lc_ppp, elev)
forest = DATA$Forest
rho_forest = rhohat(lc_ppp, forest)

par(mfrow = c(2,2))
par(mar=c(0,0,2,1))
plot(rho_elev,
     main = "",
     xlab = "Elevation (m)",
     xlim=c(0, 3600))
plot(rho_forest,
     main = "",
     xlab = "Forest Cover (%)")
plot(rho_hfi,
     main = "",
     xlab = "Human Footprint Index")
plot(rho_distw,
     main = "",
     xlab = "Distance to Water (m)")
par(mfrow = c(1,1))
#correlation
cor.im(elev, forest, hfi, distw,  use = 'complete.obs')
# Convert to a ppp object
lc_ppp <- ppp(x = lc$decimalLongitude, # X coordinates
              y = lc$decimalLatitude, # Y coordinates
              window = bc_window_owin, # Observation window
              )
lc_ppp <- as.ppp(lc_ppp)

mu <- mean(DATA$Elevation)
stdev <- sd(DATA$Elevation)
Elevation_scaled <- eval.im((elev - mu)/stdev, DATA)
mu <- mean(DATA$Dist_Water)
```

```r
stdev <- sd(DATA$Dist_Water)
Dist_Water_scaled <- eval.im((distw - mu)/stdev, DATA)
# intercept only
fit_intercept = ppm(lc_ppp ~ 1, data = DATA)
fit1 = ppm(lc_ppp ~ Elevation_scaled + I(Elevation_scaled^2) + hfi + I(hfi^2))
fit1
AIC(fit1); AIC(fit_intercept)
anova(fit_intercept, fit1, test = "LRT") #LRT
quadrat.test(fit1, nx = 2, ny = 4)
#lurking variable plot
lurking(fit1, forest, type = "raw", cumulative = F, envelope = T, xlab = "Forest Coverage (%)" )
#lurking variable plot
lurking(fit1, Dist_Water_scaled, type = "raw", cumulative = F, envelope = T, xlab = "Water Distance" )
#partial residuals
par_res_elev = parres(fit1, "Elevation_scaled")
par_res_hfi = parres(fit1, 'hfi')
par(mfrow = c(1,2))
plot(par_res_elev,
     legend = FALSE,
     lwd = 2,
     main = "",
     xlab = "Elevation Scaled")
plot(par_res_hfi,
     legend = FALSE,
     lwd = 2,
     main = "",
     xlab = "HFI")
par(mfrow = c(1,1))
fit2 = ppm(lc_ppp ~ Elevation_scaled + I(Elevation_scaled^2) + I(Elevation_scaled^3) + hfi + I(hfi^2))
fit2
AIC(fit1); AIC(fit2)
anova(fit1, fit2, test = "LRT") #LRT
#partial residuals
par_res_elev = parres(fit2, "Elevation_scaled")
par_res_hfi = parres(fit2, 'hfi')
par(mfrow = c(1,2))
plot(par_res_elev,
     legend = FALSE,
     lwd = 2,
     main = "",
     xlab = "Elevation Scaled")
plot(par_res_hfi,
     legend = FALSE,
     lwd = 2,
     main = "",
     xlab = "HFI")
par(mfrow = c(1,1))
#plotting ppm prediction
plot(fit2,
     se = FALSE,
     superimpose = FALSE,
     n = 150)
plot(lc_ppp,
```

```
      pch = 16,
      cex = 0.6,
      cols = "white",
      add = TRUE)
plot(lc_ppp,
      pch = 16,
      cex = 0.5,
      cols = "black",
      add = TRUE)
```

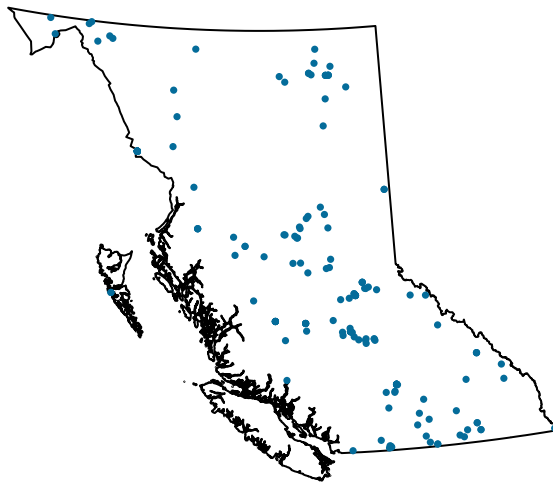# Results

First Moment Descriptives

Below we will see a plot of Canadian Lynx within BC.

We can visualize this by first showing the observation window, which is the province of BC.

```
## Loading required package: sp
```

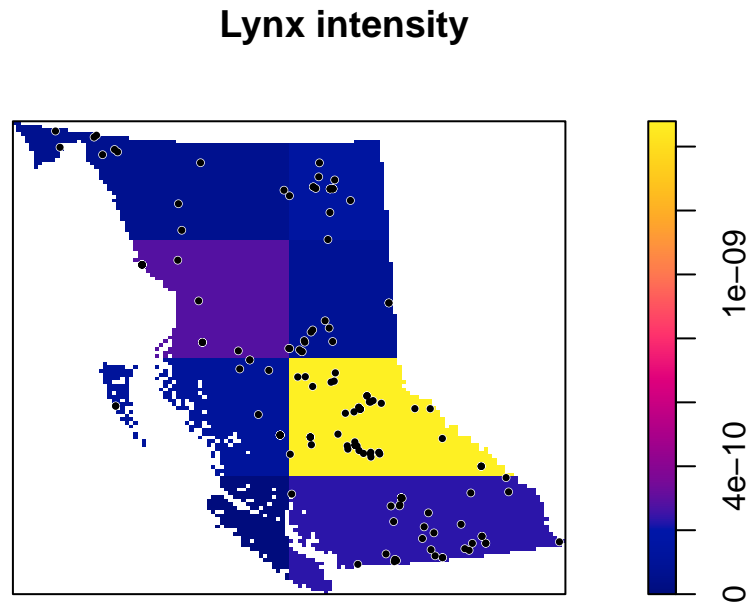Now we can overlay the two to show the observations within the window.

Marked Canada Lynx

Intensity

To determine the intensity of the Lynx across the region, we compute the value itself.

```
## [1] 1.856026e-10
```

This value means that there are 1.856026e-10 Lynx found per square meter in BC.

We can also see the intensity with observations through quadratcount.

Lynx intensity

Plot on BC Map

```
##
## Attaching package: 'raster'

## The following object is masked from 'package:nlme':
##
##     getData
```

From the plots above, it can be seen that the assumption of homogeneity is not appropriate for this dataset as the lynx tend to be clustered in certain areas of the study site, whereas others have no lynx at all.

Next, we can perform a quadrat test of homogeneity to determine if, statistically, the assumption of homogeneity is met.

```
##
##  Chi-squared test of CSR using quadrat counts
##
## data:
## X2 = 54.379, df = 7, p-value = 3.958e-09
## alternative hypothesis: two.sided
##
## Quadrats: 8 tiles (irregular windows)
```

The small p-value suggests that there is a significant deviation from homogeneity. So, the assumption of homogeneity is not met.

Nearest Neighbour Distances

Now we can compute nearest neighbor distances.
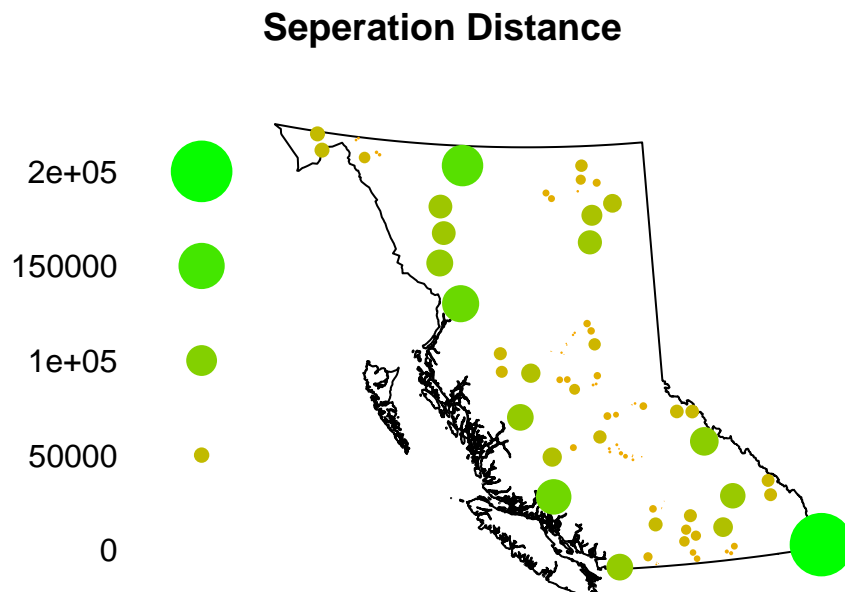
We see in the plot above that there are many regions with very large distances between the nearest neighbours. This follows our hypothesis that they are nomadic creatures in small groups or alone.

Second Moment Descriptives

Following this, second moment descriptive statistics were visualized.

K-Function

The K-Function will be the first visualization of this section.

Homogeneous K–Function
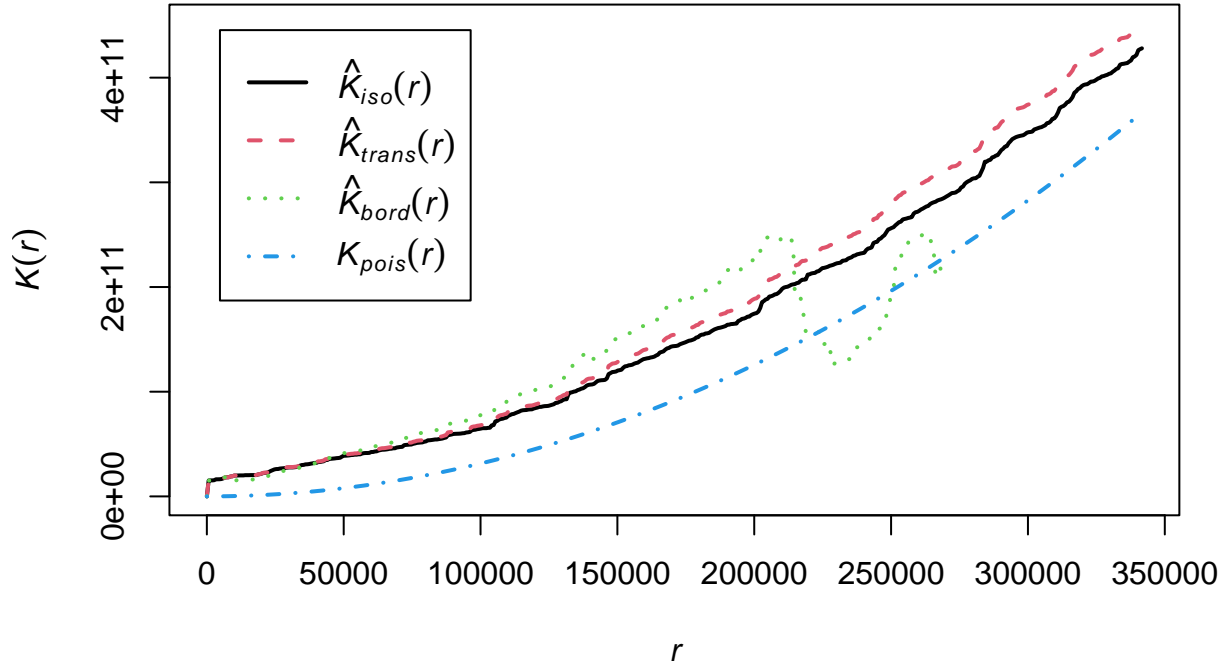
Here we can see that theoretical k-function $k_{pois}(r)$ deviates from other k-function corrections indicates clustering but these estimates assume homogeneity.

```
## Generating 19 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.
##
## Done.
```
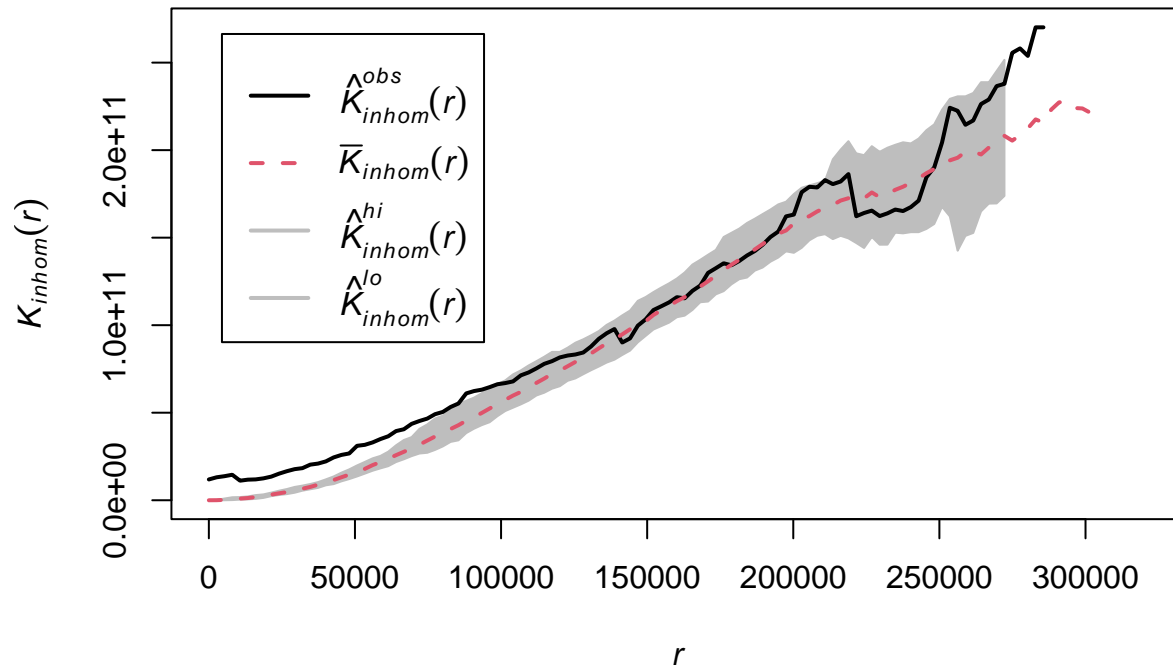
Homogeneous K–function
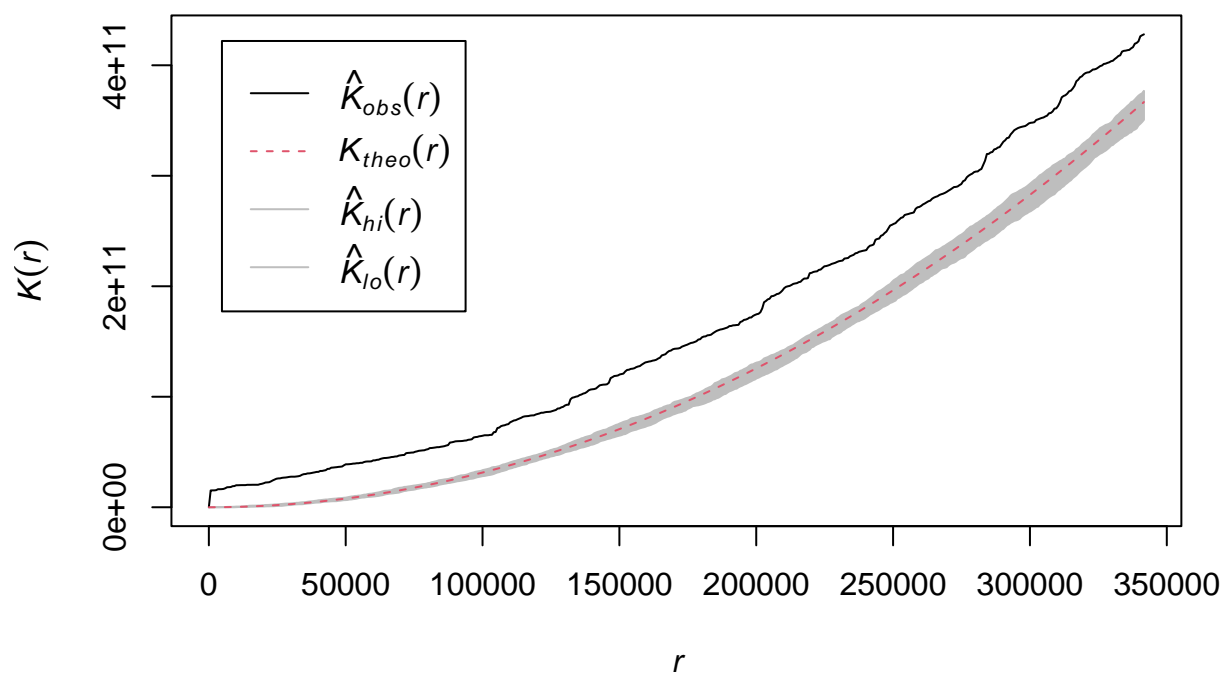
From this, we have evidence that suggests significant clustering, but these estimates assume homogeneity.
Now we can relax the homogeneity assumption.

```
## 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.
```

Inhomogeneous K−function

When correcting for inhomogeneity, the clustering is not as strong homogeneous k-function. Clustering appears to exist in and around 0 to 125000 units and 275000 to 300000 units.

Pair Correlation Function

The pair correlation function can be used next.

```
## 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.
```

Inhomogeneous g−function
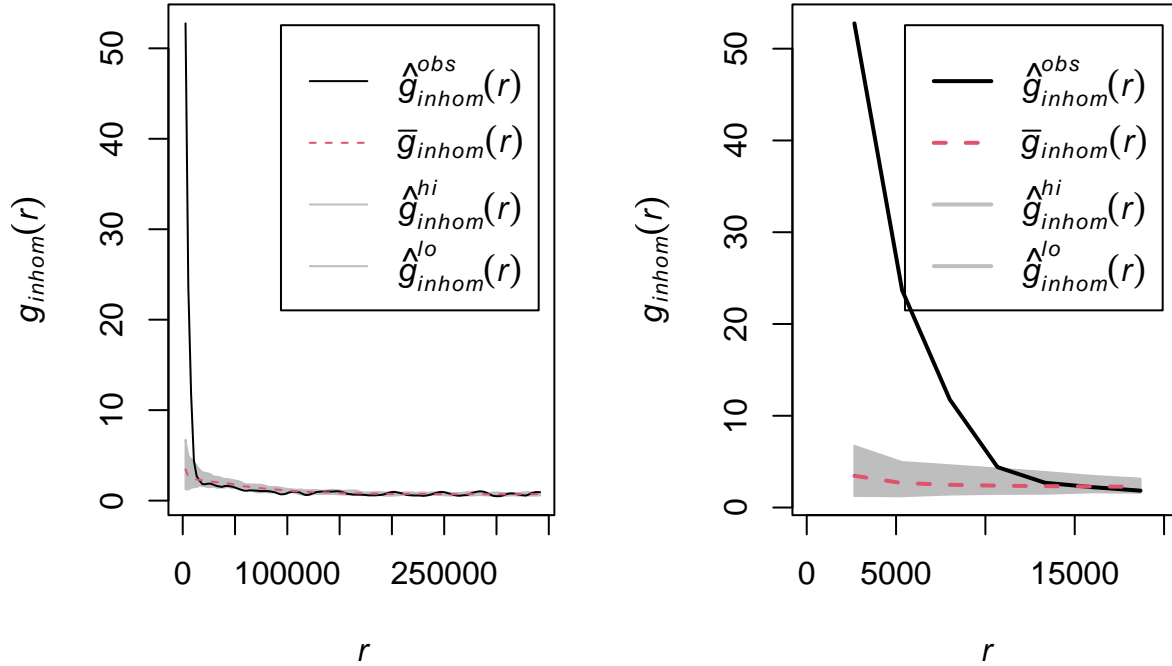
There appear to be more lynx than expected by random chance between 0 - 13500 as $g(r) > 1$. Beyond that, the locations of lynx appear not to exhibit any significant correlations.

Relationships with Covariates

We are usually interested in determining whether the intensity depends on a covariate(s). One simple approach to check for a relationship between inhomogeneous $\lambda(u)$ and a spatial covariate $Z(u)$ is via quadrat counting. We can create five elevation classes with equal width below.

Now we can look at the distribution of the elevation classes in the window, the province of BC.
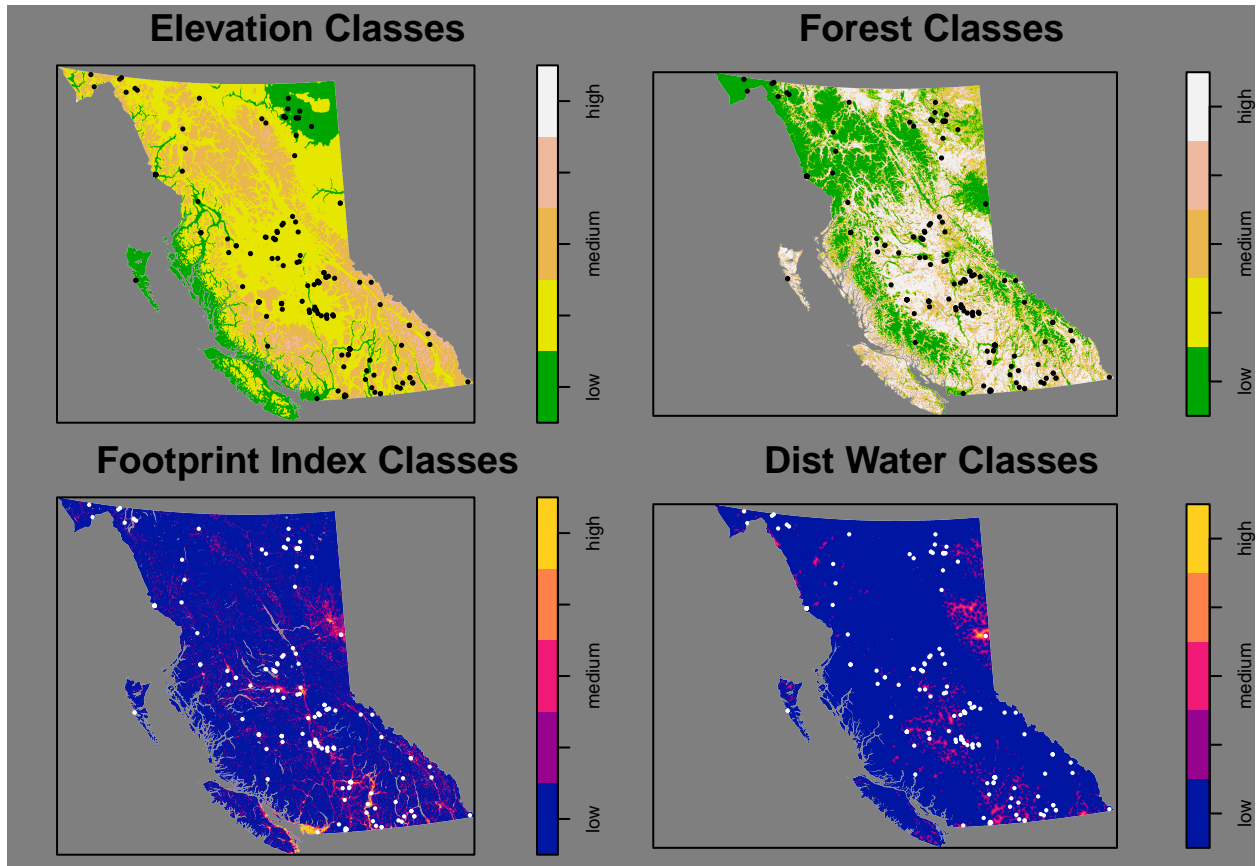
Based on a visual inspection, it appears that the heterogeneity observed in the data may be associated with a preference for certain elevations and distances from water.

More formally, in testing for relationships with covariates we are assuming that λ is a function of Z, such that

$$\lambda(u) = \rho(Z(u))$$

A non-parametric estimate of ρ can be obtained via kernel estimation, available via the rhohat() function.

We can see the plot of Rho with elevation.
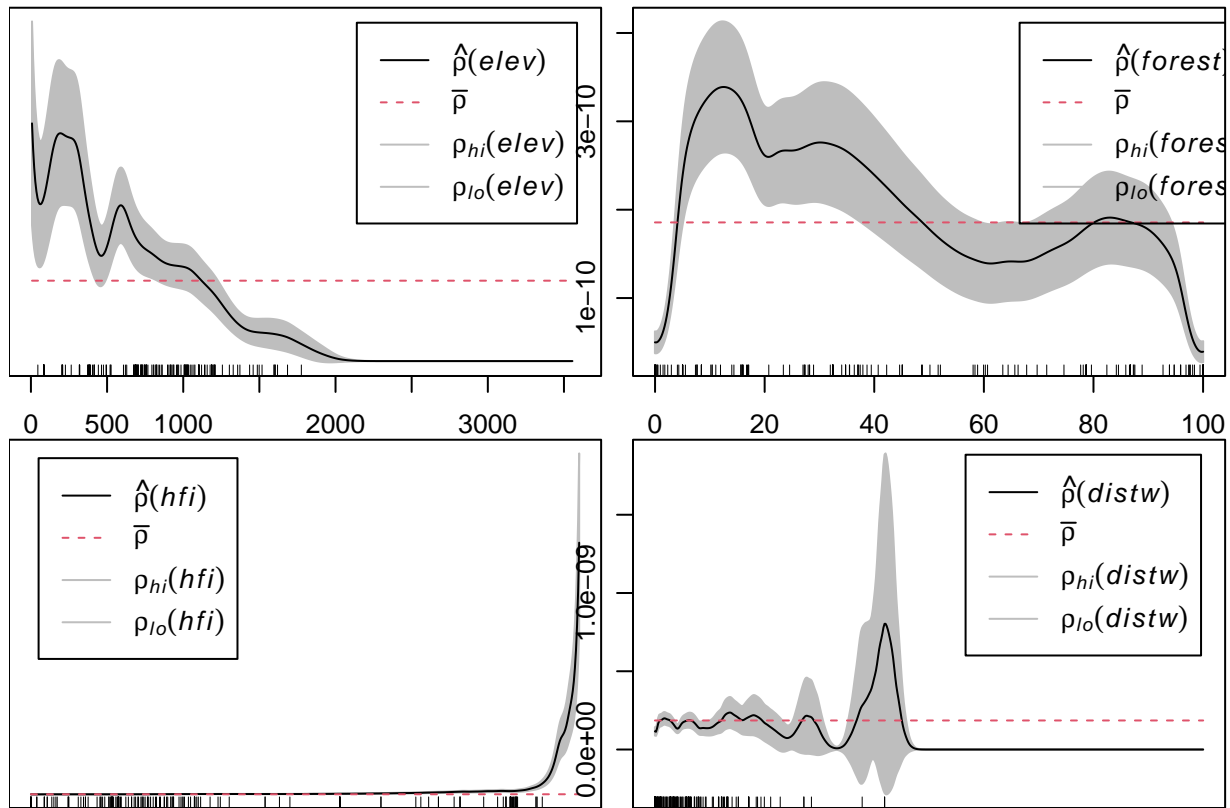
There is a non-linear relationship between elevation and lynx intensity. Majority of the lynx are at low elevations than would be expected by chance, fewer lynx at intermediate elevations and no lynx at high elevations.

```
##              ..1         ..2         ..3         ..4
## ..1  1.00000000 -0.26225376 -0.26625626 -0.03493453
## ..2 -0.26225376  1.00000000  0.06618592  0.04818598
## ..3 -0.26625626  0.06618592  1.00000000  0.13246899
## ..4 -0.03493453  0.04818598  0.13246899  1.00000000
```

The correlation coefficients are relatively weak. We can proceed without too much worry.

Model Fitting

```
## Nonstationary Poisson process
## Fitted to point pattern dataset 'lc_ppp'
##
## Log intensity:  ~Elevation_scaled + I(Elevation_scaled^2) + hfi + I(hfi^2)
##
## Fitted trend coefficients:
##           (Intercept)      Elevation_scaled I(Elevation_scaled^2)
##           -23.5142168            -0.5496413           -0.1806727
##                   hfi              I(hfi^2)
##             7.9345208            -5.1926304
##
```

15

```
##                          Estimate        S.E.      CI95.lo      CI95.hi Ztest
## (Intercept)            -23.5142168 0.17827617 -23.8636317 -23.1648020   ***
## Elevation_scaled        -0.5496413 0.13384529  -0.8119732  -0.2873093   ***
## I(Elevation_scaled^2)   -0.1806727 0.08343033  -0.3441932  -0.0171523     *
## hfi                      7.9345208 1.15645659   5.6679075  10.2011340   ***
## I(hfi^2)                -5.1926304 1.33813978  -7.8153361  -2.5699246   ***
##                              Zval
## (Intercept)            -131.897701
## Elevation_scaled         -4.106542
## I(Elevation_scaled^2)    -2.165552
## hfi                       6.861062
## I(hfi^2)                 -3.880484
## Problem:
##   Values of the covariate 'hfi' were NA or undefined at 0.44% (3 out of 689) of
## the quadrature points
##
## *** Fitting algorithm for 'glm' did not converge ***
```

In the fit1 model we can see that the coefficients are significant. We can also check the AIC score and likelihood ratio test:

```
## [1] 7965.547
```

```
## [1] 8241.409
```

```
## Analysis of Deviance Table
##
## Model 1: ~1    Poisson
## Model 2: ~Elevation_scaled + I(Elevation_scaled^2) + hfi + I(hfi^2)   Poisson
##   Npar Df Deviance  Pr(>Chi)
## 1    4
## 2    8  4   237.45 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

So based on AIC score and likelihood ratio test fit1 model is better.

```
##
##   Chi-squared test of fitted Poisson model 'fit1' using quadrat counts
##
## data:  data from fit1
## X2 = 56.478, df = 3, p-value = 6.641e-12
## alternative hypothesis: two.sided
##
## Quadrats: 8 tiles (irregular windows)
```

The small p value from the quadrat test tells us that the data deviate significantly from the model's predictions.

Let's check whether adding other covariates going to improve the model fit or not.

```
## Generating 39 simulated patterns ...1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19
## Processing.. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24,
## Done.
```
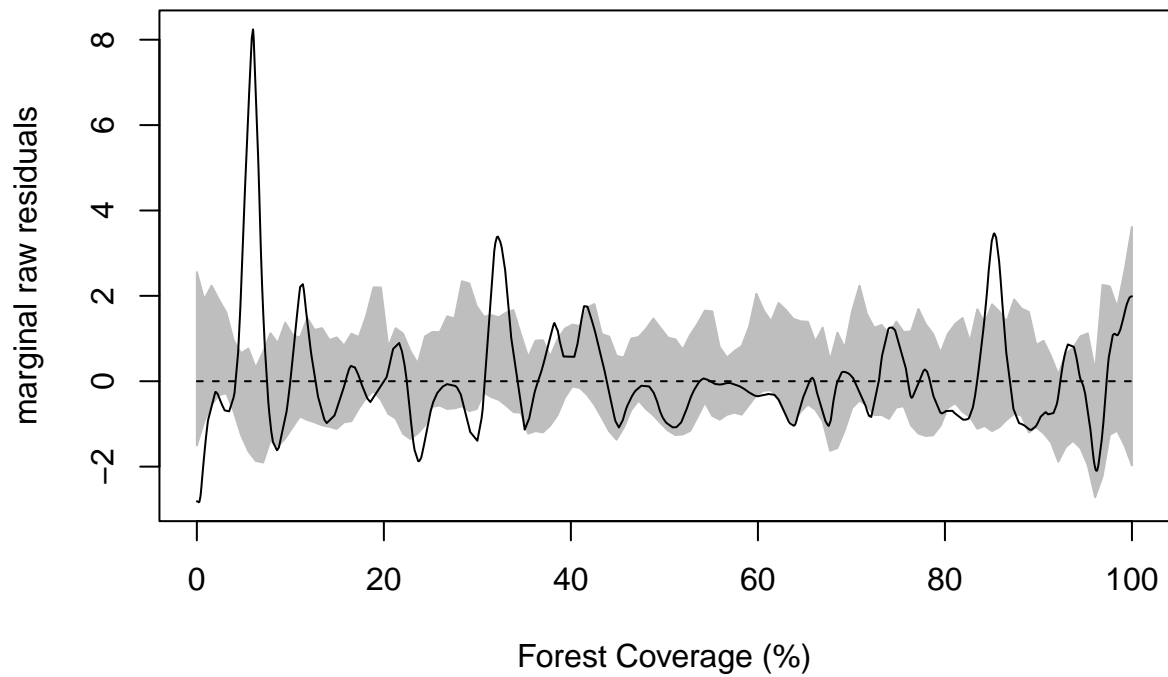
It appears that the forest coverage variable may not be significant in the model, hence we could investigate the impact of water distance instead.

```
## Generating 39 simulated patterns ...1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19
## Processing.. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 2
## Done.
```
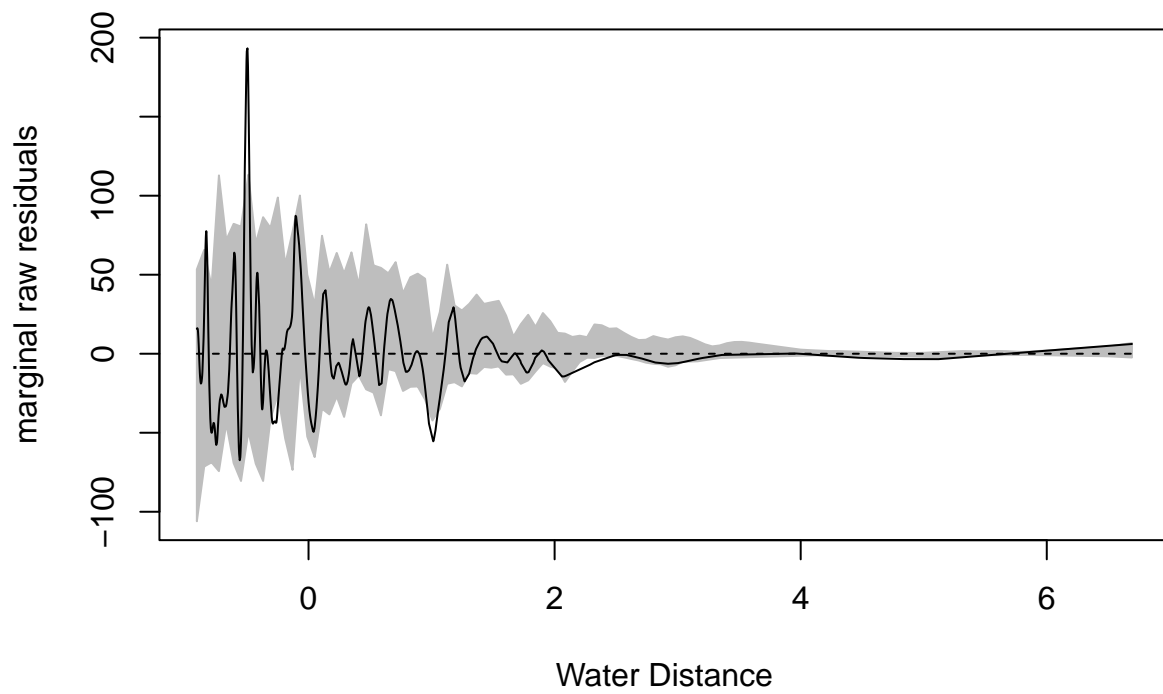
So, water distance is unrelated to trends in the residuals. We can diagnose our model further by partial residual plots to check how we can improve.
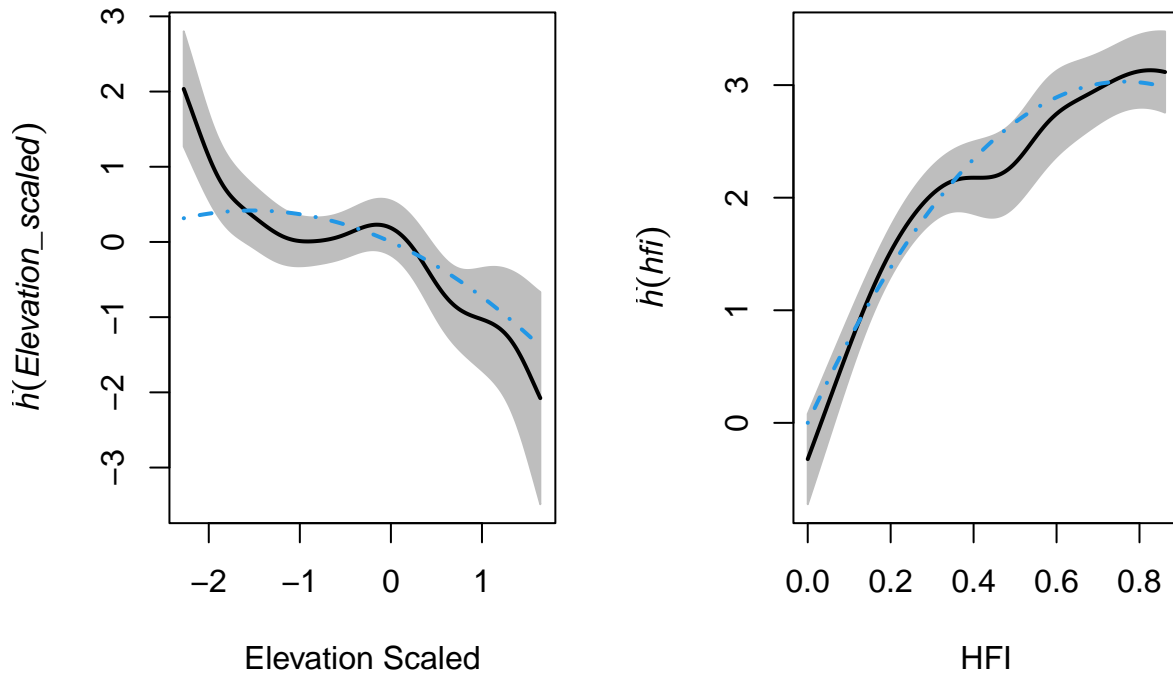
From these figures we can see that the quadratic term of HFI is capturing the patterns in our data particularly well but we can improve the elevation by adding higher order plynomials.

```
## Nonstationary Poisson process
## Fitted to point pattern dataset 'lc_ppp'
##
## Log intensity:  ~Elevation_scaled + I(Elevation_scaled^2) +
## I(Elevation_scaled^3) + hfi + I(hfi^2)
##
## Fitted trend coefficients:
##         (Intercept)      Elevation_scaled I(Elevation_scaled^2)
##         -23.3153723           -0.3313493           -0.5725852
## I(Elevation_scaled^3)               hfi            I(hfi^2)
##          -0.2390356            7.8927824           -5.1130689
##
##                        Estimate       S.E.      CI95.lo       CI95.hi Ztest
## (Intercept)          -23.3153723 0.19003181 -23.6878278 -22.942916820   ***
## Elevation_scaled      -0.3313493 0.17000882  -0.6645604   0.001861899
## I(Elevation_scaled^2) -0.5725852 0.20504110  -0.9744583  -0.170711992    **
## I(Elevation_scaled^3) -0.2390356 0.09724718  -0.4296365  -0.048434603     *
## hfi                    7.8927824 1.14898528   5.6408127  10.144752191   ***
## I(hfi^2)              -5.1130689 1.33437971  -7.7284051  -2.497732762   ***
##                          Zval
## (Intercept)          -122.691945
## Elevation_scaled       -1.949012
## I(Elevation_scaled^2)  -2.792538
```

19

```
## I(Elevation_scaled^3)   -2.458021
## hfi                      6.869350
## I(hfi^2)                -3.831795
## Problem:
##  Values of the covariate 'hfi' were NA or undefined at 0.44% (3 out of 689) of
## the quadrature points
##
## *** Fitting algorithm for 'glm' did not converge ***
```

These values indicate that:

- Lynx are less often observed at higher altitudes
- Lynx are more often observed in areas of low human development (ie. rural areas/parks, etc.) compared to no development (ie. "middle of nowhere"), but less often in areas that are highly developed (ie. cities)
- Lynx distribution is not related to forest coverage or distance to water

AIC values for squared model and cubed model, respectively:

```
## [1] 7965.547
```

```
## [1] 7960.116
```

Likelihood ratio test for squared and cubed models:

```
## Analysis of Deviance Table
##
## Model 1: ~Elevation_scaled + I(Elevation_scaled^2) + hfi + I(hfi^2)   Poisson
## Model 2: ~Elevation_scaled + I(Elevation_scaled^2) + I(Elevation_scaled^3) + hfi + I(hfi^2)   Poisso
##   Npar Df Deviance Pr(>Chi)
## 1    5
## 2    6  1   7.4317 0.006409 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The small p-value indicates that the additional complexity supported by data.
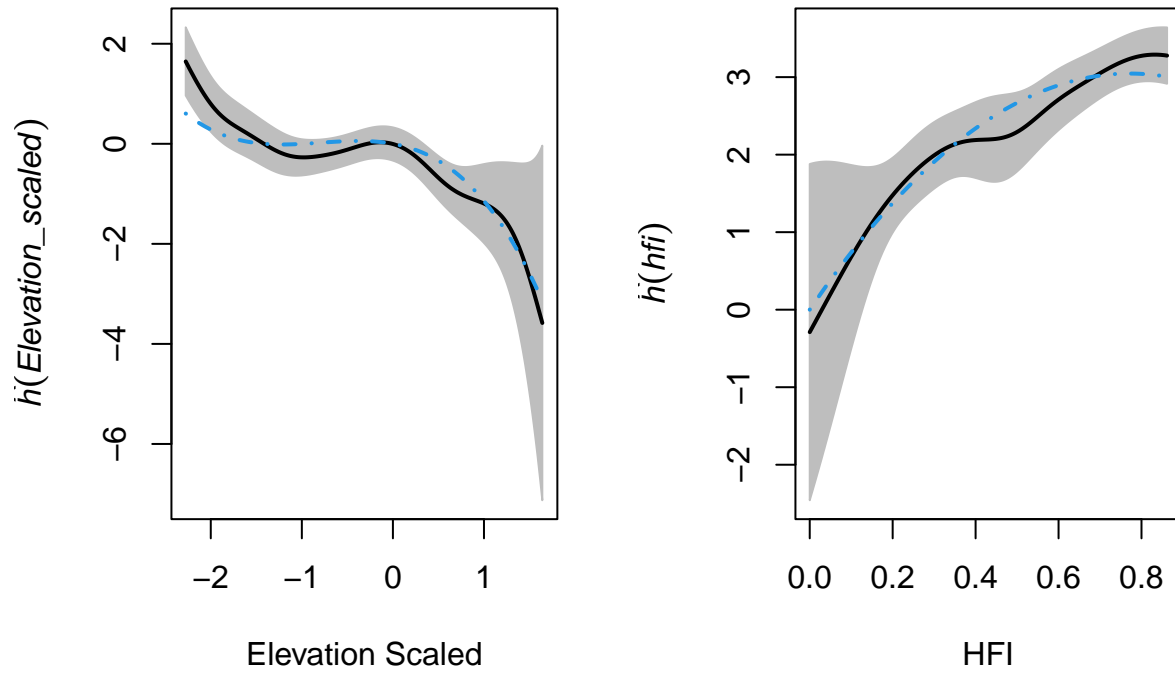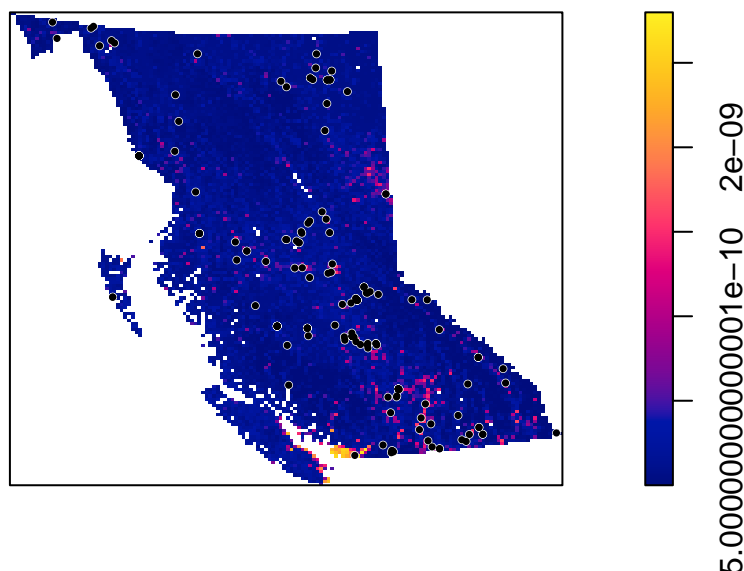
In addition, the inclusion of extra elevation polynomial term slightly enhances the accuracy of the model predictions.

Since the model is already performing well, we are not considering to add any additional complexity to it.

Plotting `ppm` predictions:

Fitted trend

There are still locations where the model seems to not be a particularly good match to the data, but it is actually performing reasonably well considering that we are predicting the locations of lynx in BC based on only elevation and human footprint index, and have no information on all of the many other factors that might influence lynx locations.

Discussion

Many of our hypotheses were found to be true. We found in the plots that the assumption of homogeneity is not appropriate for this Lynx dataset because the Lynx were clustered in certain areas of the observation window, whereas other locations had no Lynx at all. We also found that Lynx are animals that travel in either small groups or nomadically, as the nearest neighbour distances were quite large, showing that they typically stick to their territory and are not located close by each other.

We expected that lynx would be found at higher elevations, but looking at the plots, this was not the case. However, we believe the data is slightly biased because it only includes lynx sightings by humans. As lynx typically avoid humans, and areas where humans reside (as showing in the HFI plot), we believe that the majority of locations where the lynx are found have not been observed in this data. Humans are typically not in higher elevation areas, so they would not observe these areas as frequently as lower elevation areas, so the observations of the lynx are not truly representative.

As mentioned above when summarizing the model, forest coverage was not predictive in the model. The rho plot appeared as though there may be a relationship, but it was being influenced by other factors in the model that were correlated with this covariate. Similarly, distance to water was not influential in the lynx distribution of this dataset. Although this may seem strange, we can again question these results using the same logic as mentioned above regarding the elevation. Naturally, lynx need water to survive, so you would expect that they do reside within a certain distance of a water source, so this leads us to question our results

slightly. Additionally, as eluded to above, lynx are more often observed in areas of low human development, such as rural areas or parks, rather than highly developed areas such as cities. However, this does not include areas with no development, such as forests, where we would expect to find lynx. This means that the true areas where lynx are found are not properly represented in this dataset, as areas with no HFI are difficult to access and are not regularly tracked, so this is important to consider when interpreting results.

Overall, we can make a few inferences as mentioned above regarding the lynx species in BC, but given that the data had fewer than 200 observation points, despite the lynx population in BC being estimated at much larger amounts, this needs to be taken into account. In the future, it would be beneficial to complete an analysis with a larger dataset to either confirm or reassess our results.

References

GBIF.org (01 April 2023) GBIF Occurrence Download https://doi.org/10.15468/dl.zeymyk