589Project_a

Shengyao Li

2024-04-27

GIS

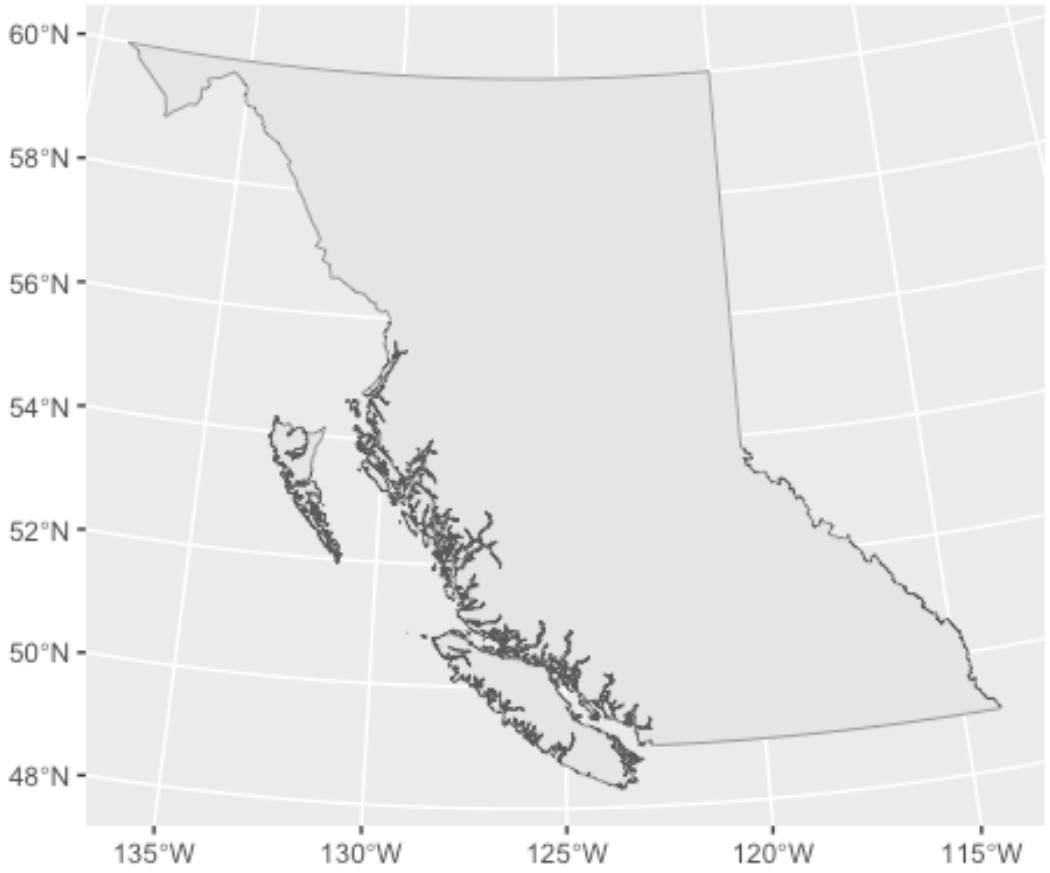
```
install.packages("data.table")
```

```
##
## The downloaded binary packages are in
##  /var/folders/3b/l64jrnxs3zb4gqrm_qnntv780000gn/T//RtmpK3zMQO/downloaded_p
ackages
```

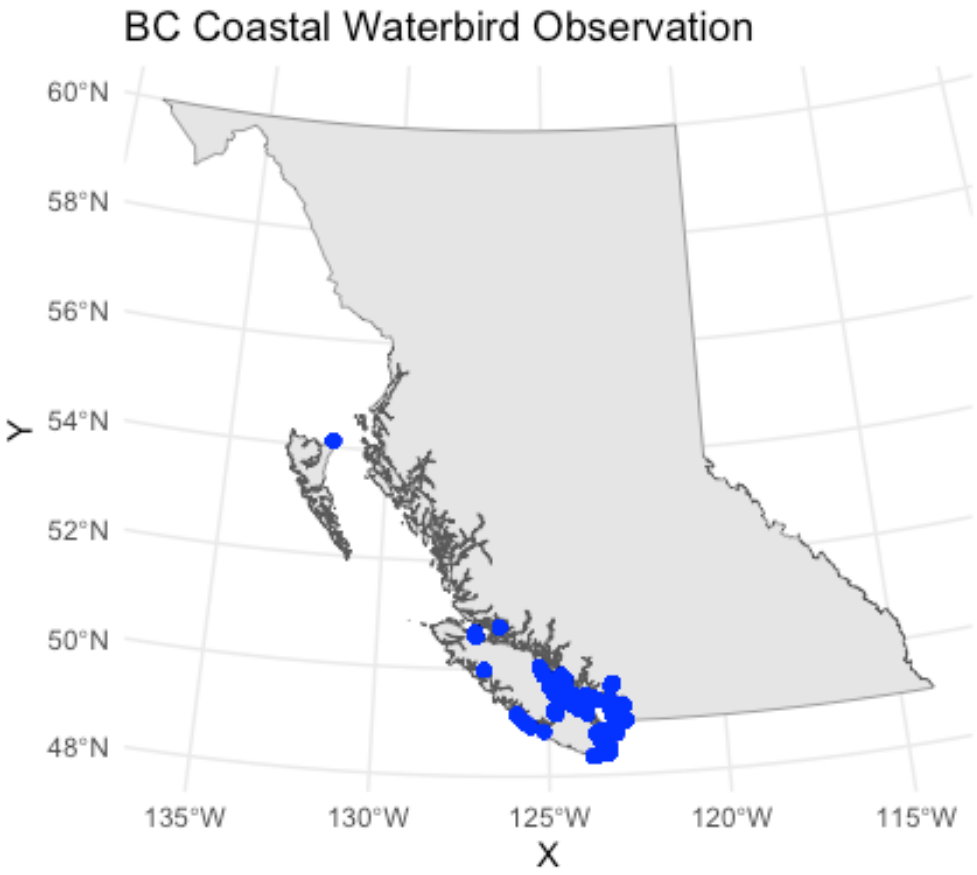
```
library(data.table)
dataset <- fread("589_RawData.csv")

library(dplyr)



selected_data <- dataset %>% select(gbifID, decimalLatitude, decimalLongitude
, eventDate, verbatimScientificName)


write.csv(selected_data, file = "selected_data.csv", row.names = FALSE)

summary(selected_data)
```

```
##      gbifID          decimalLatitude decimalLongitude   eventDate
##  Min.   : 29796652   Min.   :48.35   Min.   :-131.7   Min.   :1999
##  1st Qu.: 29827107   1st Qu.:49.02   1st Qu.:-124.7   1st Qu.:2000
##  Median : 29860059   Median :49.31   Median :-123.8   Median :2001
##  Mean   : 30511191   Mean   :49.30   Mean   :-124.0   Mean   :2001
##  3rd Qu.: 29893713   3rd Qu.:49.56   3rd Qu.:-123.2   3rd Qu.:2002
##  Max.   :865776610   Max.   :54.12   Max.   :-122.8   Max.   :2003
##                      NA's   :2186    NA's   :2186
##  verbatimScientificName
##  Length:23146
##  Class :character
##  Mode  :character
##
##
##
##
```

```r
library(sp)
library(sf)

## Linking to GEOS 3.11.0, GDAL 3.5.3, PROJ 9.1.0; sf_use_s2() is TRUE

observations <- read.csv("selected_data.csv")
observations <- na.omit(observations)

sum(is.na(observations$decimalLongitude))

## [1] 0

sum(is.na(observations$decimalLatitude))

## [1] 0

if(sum(is.na(observations$decimalLongitude)) == 0 & sum(is.na(observations$de
cimalLatitude)) == 0) {
  coordinates(observations) <- ~decimalLongitude + decimalLatitude
  proj4string(observations) <- CRS("+proj=longlat +datum=WGS84 +no_defs")
} else {
  stop("There are still missing values in the coordinates.")
}

load("BC_Parks.Rda")

library(ggplot2)

ggplot() + geom_sf(data = DATA$Window)
```




```
observations_sf <- st_as_sf(observations, coords = c("decimalLongitude", "dec
imalLatitude"), crs = 4326)
observations_sf <- st_transform(observations_sf, crs = st_crs(DATA$Window))

ggplot() +
  geom_sf(data = DATA$Window) +
  geom_point(data = as.data.frame(st_coordinates(observations_sf)), aes(x = X
, y = Y), color = "blue") +
  labs(title = "BC Coastal Waterbird Observation") +
  theme_minimal()
```



Quadrat Count

```
library(spatstat)

window_data <- as.owin(DATA$Window)


coords <- st_coordinates(observations_sf)
observations_ppp <- ppp(x = coords[,1], y = coords[,2], window = window_data)

## Warning: 18374 points were rejected as lying outside the specified window

## Warning: data contain duplicated points

Qua <- quadratcount(observations_ppp,
                    nx = 10,
```

```
                      ny = 10)
quad_test <- quadrat.test(Qua)

## Warning: Some expected counts are small; chi^2 approximation may be inaccu
rate

quad_test

##
##  Chi-squared test of CSR using quadrat counts
##
## data:
## X2 = 68990, df = 63, p-value < 2.2e-16
## alternative hypothesis: two.sided
##
## Quadrats: 64 tiles (irregular windows)

plot(Qua, main="Quadrat count of bird observations")
points(observations_ppp)
```

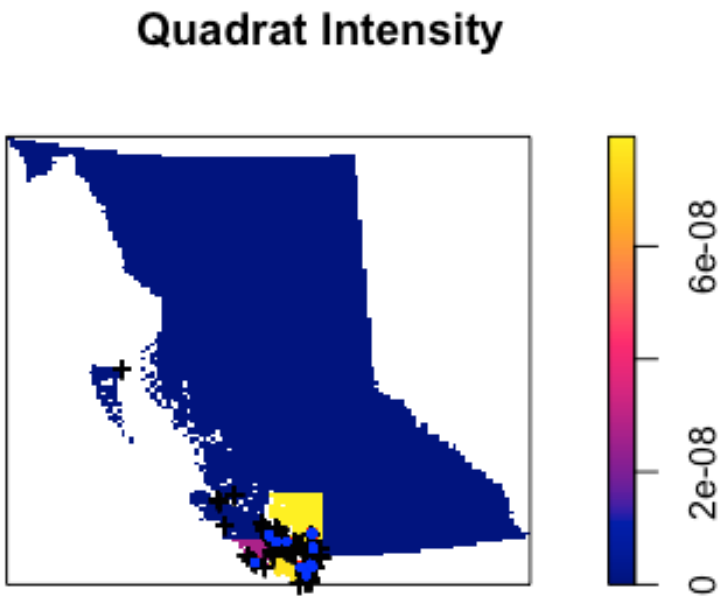
Quadrat Intensity

```
plot(intensity(Qua, image = T),
     main = "Quadrat Intensity")

plot(observations_ppp,
     pch = 16,
     cex = 0.6,
     cols = "red",
     add = T)
```

```
## Warning in plot.ppp(observations_ppp, pch = 16, cex = 0.6, cols = "red", :
## 18374 illegal points also plotted
```

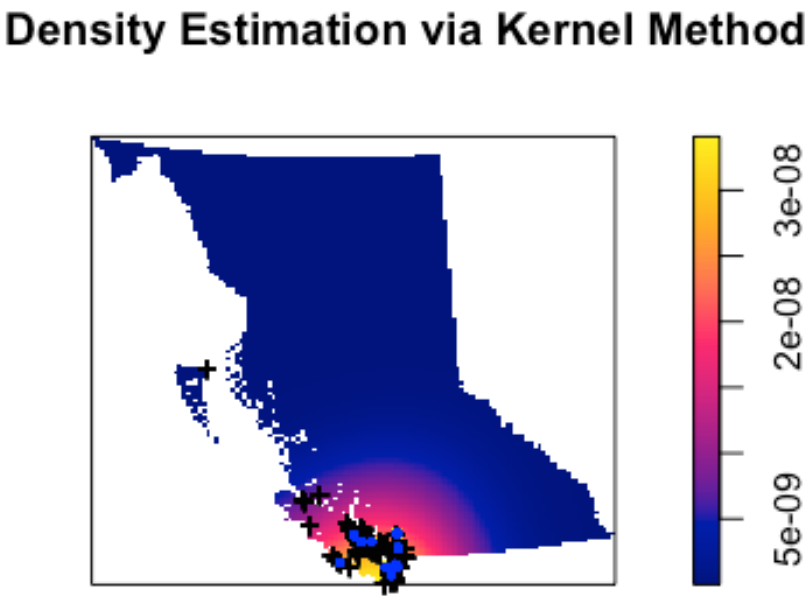
```
plot(observations_ppp,
     pch = 16,
     cex = 0.5,
     cols = "blue",
     add = T)
```

```
## Warning in plot.ppp(observations_ppp, pch = 16, cex = 0.5, cols = "blue",
:
## 18374 illegal points also plotted
```



KDE

```
lambda_u_hat_value <- density(observations_ppp)
plot(lambda_u_hat_value, main = "Density Estimation via Kernel Method")
plot(observations_ppp,
     pch = 16,
     cex = 0.6,
     col = "yellow",
     add = TRUE)

## Warning in plot.ppp(observations_ppp, pch = 16, cex = 0.6, col = "yellow",
 :
## 18374 illegal points also plotted

plot(observations_ppp,
     pch = 16,
     cex = 0.5,
     col = "blue",
     add = TRUE)
```



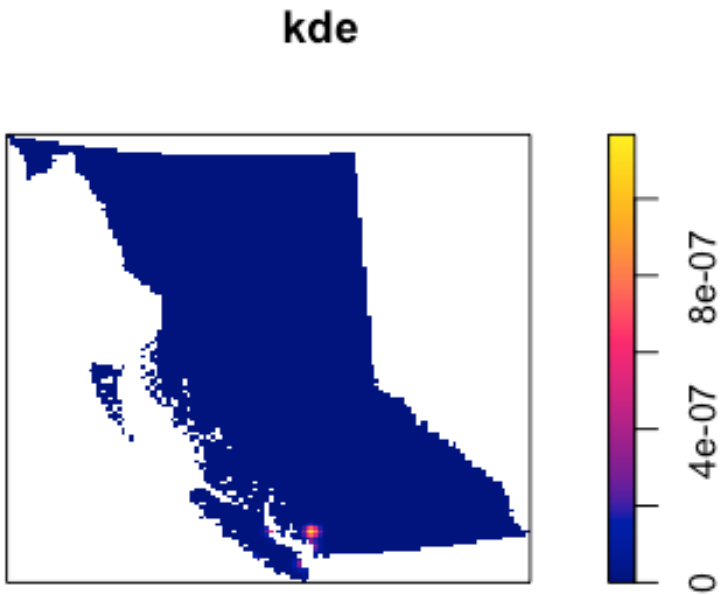

```
observations_ppp$n
```

```
## [1] 2586
```

```
sum(is.na(observations_ppp$x), is.na(observations_ppp$y), is.infinite(observa
tions_ppp$x), is.infinite(observations_ppp$y))
```

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

```
sigma_val <- bw.scott(observations_ppp)
kde <- density(observations_ppp, sigma = sigma_val)
plot(kde)
```



Birds Count Distribution

```
observations <- read.csv("selected_data.csv")
observations <- na.omit(observations)


sum(is.na(observations$decimalLongitude))

## [1] 0

sum(is.na(observations$decimalLatitude))

## [1] 0
```

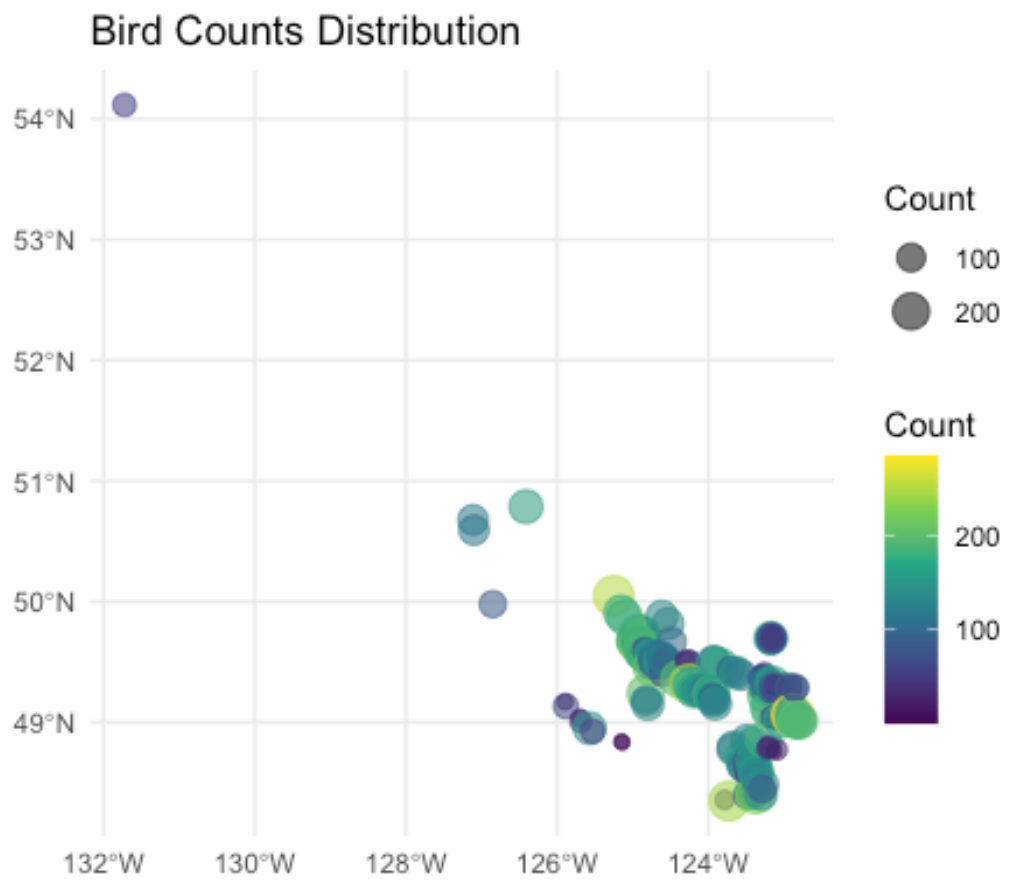
```r
observations_sf <- st_as_sf(observations, coords = c("decimalLongitude", "dec
imalLatitude"), crs = 4326)
install.packages("dplyr")
```

```
##
## The downloaded binary packages are in
##  /var/folders/3b/l64jrnxs3zb4gqrm_qnntv780000gn/T//RtmpK3zMQO/downloaded_p
ackages
```

```r
library(dplyr)
coords <- st_coordinates(observations_sf)
coords_df <- as.data.frame(coords)
bird_counts <- coords_df %>%
  group_by(X, Y) %>%
  summarise(birdCount = n())
```

```
## `summarise()` has grouped output by 'X'. You can override using the `.grou
ps`
## argument.
```

```r
bird_counts_sf <- st_as_sf(bird_counts, coords = c("X", "Y"), crs = 4326)
library(ggplot2)
library(sf)

ggplot(data = bird_counts_sf) +
  geom_sf(aes(size = birdCount, color = birdCount), alpha = 0.6) +
  scale_color_viridis_c() +
  labs(title = "Bird Counts Distribution",
       color = "Count",
       size = "Count") +
  theme_minimal()
```



Moran's I

```
install.packages("spdep")

library(sf)
library(spdep)

k <- 5

neighbors <- knn2nb(knearneigh(st_coordinates(bird_counts_sf), k=k), sym=TRUE
)

weights <- nb2listw(neighbors, style="W")

moran <- moran.test(bird_counts_sf$birdCount, listw = weights)
```

```r
print(moran)
```

```
##
##  Moran I test under randomisation
##
## data:  bird_counts_sf$birdCount
## weights: weights
##
## Moran I statistic standard deviate = 7.4456, p-value = 4.824e-14
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##       0.285359004      -0.004975124       0.001520519
```