Occurrences mapping

Norah Saarman

2024-11-10

Point data by time frame

```r
# Load necessary libraries
library(ggplot2)
library(sf)
```

```
## Linking to GEOS 3.10.2, GDAL 3.4.1, PROJ 8.2.1; sf_use_s2() is TRUE
```

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(maps)

# Load the data from your CSV files, filter out NAs, and ensure `year` column is numeric
pipiens <- read.csv("coords_pip_v4.csv") %>%
  filter(!is.na(year) & latitude > 20) %>%
  mutate(year = as.numeric(year))
```

```
## Warning: There was 1 warning in `mutate()`.
## i In argument: `year = as.numeric(year)`.
## Caused by warning:
## ! NAs introduced by coercion
```

```r
quinquefasciatus <- read.csv("coords_qui_v4.csv") %>%
  filter(!is.na(year) & latitude > 20) %>%
  mutate(year = as.numeric(year))

hybrids <- read.csv("coords_hyb_v4.csv") %>%
  filter(!is.na(year) & latitude > 20) %>%
  mutate(year = as.numeric(year))

# Add time period labels to each dataset and make Time_Period a factor in chronological order
pipiens <- pipiens %>% mutate(Time_Period = factor(case_when(
  year < 1970 ~ "Before 1970",
  year >= 1970 & year <= 2013 ~ "1970-2013",
```

```r
    year >= 2014 & year <= 2024 ~ "2014-2024",
    TRUE ~ "Unknown"  # Ensures no NA values in Time_Period
), levels = c("Before 1970", "1970-2013", "2014-2024")))

quinquefasciatus <- quinquefasciatus %>% mutate(Time_Period = factor(case_when(
  year < 1970 ~ "Before 1970",
  year >= 1970 & year <= 2013 ~ "1970-2013",
  year >= 2014 & year <= 2024 ~ "2014-2024",
  TRUE ~ "Unknown"
), levels = c("Before 1970", "1970-2013", "2014-2024")))

hybrids <- hybrids %>% mutate(Time_Period = factor(case_when(
  year < 1970 ~ "Before 1970",
  year >= 1970 & year <= 2013 ~ "1970-2013",
  year >= 2014 & year <= 2024 ~ "2014-2024",
  TRUE ~ "Unknown"
), levels = c("Before 1970", "1970-2013", "2014-2024")))

# Remove any entries labeled "Unknown" in Time_Period
pipiens <- pipiens %>% filter(Time_Period != "Unknown")
quinquefasciatus <- quinquefasciatus %>% filter(Time_Period != "Unknown")
hybrids <- hybrids %>% filter(Time_Period != "Unknown")

# Combine all datasets
all_data <- bind_rows(
  pipiens %>% mutate(Species = "pipiens"),
  quinquefasciatus %>% mutate(Species = "quinquefasciatus"),
  hybrids %>% mutate(Species = "hybrid")
)

# Load USA state outlines from the maps package
us_states <- map_data("state")

# Plot with facets for each time period, arranged in chronological order
ggplot() +
  # Add USA basemap with state outlines
  geom_polygon(data = us_states, aes(x = long, y = lat, group = group),
               fill = "white", color = "black", alpha = 0.3) +

  # Plot individual points with consistent colors and shapes
  geom_point(data = all_data, aes(x = longitude, y = latitude, color = Species, shape = Species), size =

  # Set color and shape scales
  scale_color_manual(values = c("pipiens" = "blue", "quinquefasciatus" = "red", "hybrid" = "purple")) +
  scale_shape_manual(values = c("pipiens" = 16, "quinquefasciatus" = 17, "hybrid" = 18)) +

  # Facet the plot into three panels by time period in a single row
  facet_wrap(~Time_Period, nrow = 1) +

  # Add title and labels
  labs(title = "Occurrence of Culex Mosquitoes in the United States by Time Period",
       x = "Longitude", y = "Latitude", color = "Species", shape = "Species") +
  coord_fixed(ratio = 1.3) +  # Fixed aspect ratio
```
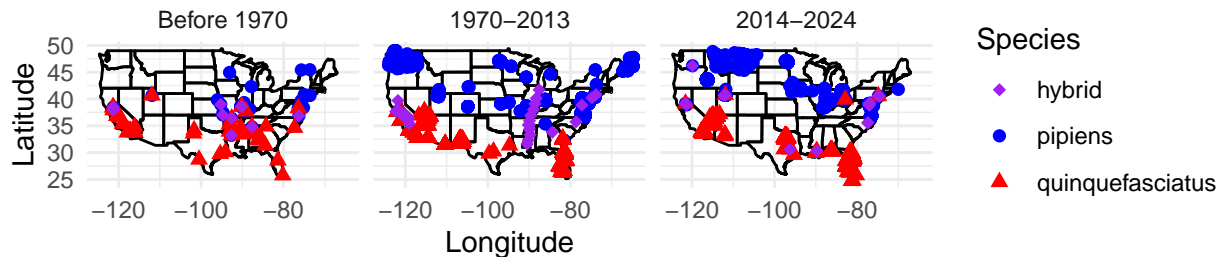
```
# Apply minimal theme
theme_minimal()
```

Occurrence of Culex Mosquitoes in the United States by Time Period

Density plot of hybrids by time period

```
# Plot separate density plots by species
ggplot() +
  # Add USA basemap with state outlines
  geom_polygon(data = us_states, aes(x = long, y = lat, group = group),
               fill = "white", color = "black", alpha = 0.3) +

  # Add density layer for hybrids in purple
  stat_density_2d(data = all_data %>% filter(Species == "hybrid"),
                  aes(x = longitude, y = latitude, fill = after_stat(level)),
                  geom = "polygon", alpha = 0.4) +
  scale_fill_gradient(low = "lavender", high = "purple", name = "Hybrid Density") +

  # Facet the plot into three panels by time period in a single row
  facet_wrap(~Time_Period, nrow = 1) +

  # Add title and labels
  labs(title = "Density of Culex Mosquito Occurrences by Time Period",
       x = "Longitude", y = "Latitude") +
  coord_fixed(ratio = 1.3) +   # Fixed aspect ratio
  # Apply minimal theme
  theme_minimal()
```
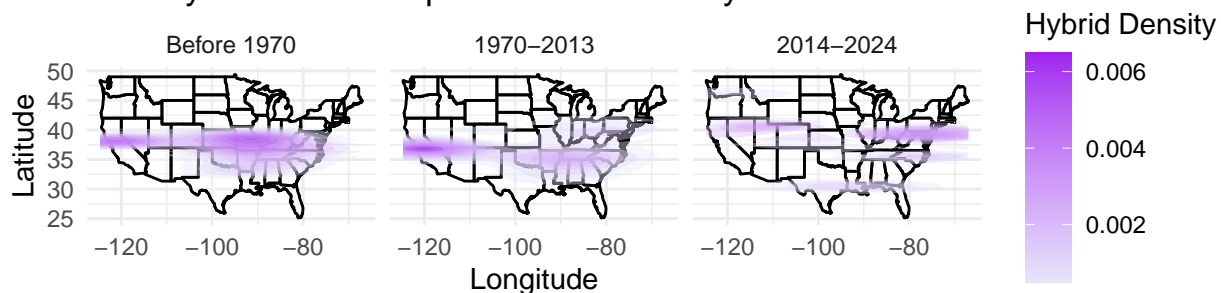

For each species, aggregating all years

```
library(gridExtra)   # for arranging multiple plots in a grid
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##      combine
```

```r
# Combine all datasets, adding species labels
all_data <- bind_rows(
  pipiens %>% mutate(Species = "pipiens"),
  quinquefasciatus %>% mutate(Species = "quinquefasciatus"),
  hybrids %>% mutate(Species = "hybrid")
)

# Remove any cases where data is insufficient for density calculation
all_data <- all_data %>% filter(!is.na(longitude), !is.na(latitude))

# Load USA state outlines from the maps package
us_states <- map_data("state")

# Create density plot for each species individually with separate color scales and fixed aspect ratio
pipiens_plot <- ggplot() +
  geom_polygon(data = us_states, aes(x = long, y = lat, group = group),
               fill = "white", color = "black", alpha = 0.3) +
  stat_density_2d(data = all_data %>% filter(Species == "pipiens"),
                  aes(x = longitude, y = latitude, fill = after_stat(level)),
                  geom = "polygon", alpha = 0.5) +
  scale_fill_gradient(low = "lightblue", high = "blue", name = "pip") +
  labs(x = NULL, y = NULL) +
  coord_fixed(ratio = 1.3) +   # Fixed aspect ratio
  theme_minimal()

quinquefasciatus_plot <- ggplot() +
  geom_polygon(data = us_states, aes(x = long, y = lat, group = group),
               fill = "white", color = "black", alpha = 0.3) +
  stat_density_2d(data = all_data %>% filter(Species == "quinquefasciatus"),
                  aes(x = longitude, y = latitude, fill = after_stat(level)),
                  geom = "polygon", alpha = 0.5) +
  scale_fill_gradient(low = "pink", high = "red", name = "quinq") +
  labs(x = NULL, y = NULL) +
  coord_fixed(ratio = 1.3) +   # Fixed aspect ratio
  theme_minimal()

hybrid_plot <- ggplot() +
  geom_polygon(data = us_states, aes(x = long, y = lat, group = group),
               fill = "white", color = "black", alpha = 0.3) +
  stat_density_2d(data = all_data %>% filter(Species == "hybrid"),
                  aes(x = longitude, y = latitude, fill = after_stat(level)),
                  geom = "polygon", alpha = 0.5) +
  scale_fill_gradient(low = "lavender", high = "purple", name = "hybrid") +
  labs(x = NULL, y = NULL) +
  coord_fixed(ratio = 1.3) +   # Fixed aspect ratio
  theme_minimal()

# Arrange the plots in a row
grid.arrange(pipiens_plot, quinquefasciatus_plot, hybrid_plot, nrow = 3)
```