# **Flight in Flux:** Analyzing American Bittern (Botaurus lentiginosus) Migration Timing and Climate Change

### **Introduction:**

Climate change poses many questions relating to systems across the globe, especially ecological and evolutionary systems. One system which could potentially be heavily impacted is migration. Bird migration timing is critical for ensuring species survival and ecological balance. As anthropogenic climate change has altered seasonal patterns, there is a potential disruption to migration timing, impacting birds’ breeding success, food availability, and overall population health. The American Bittern (Botaurus lentiginosus), for example, depends on the thriving vegetation of marshlands for camouflage and nesting, as they are known to nest on the ground. Wetland birds are expected to be affected negatively by a considerable margin with the presence of a changing climate altering their regular migratory ranges and patterns (Steen et al., 2012).

### **Abstract:**

The American Bittern is a solitary, cryptic species that relies on wetlands for breeding and overwintering. Typically, these birds breed in freshwater marshes across Canada and the northern United States before migrating south to the southeastern U.S., Mexico, and Central America for the winter. However, increasing global temperatures, shifts in precipitation, and habitat loss due to climate change threaten to disrupt these migration patterns. Past studies on migratory birds indicate that climate change is leading to earlier spring arrivals, delayed fall departures, and shifts in overall migration routes (La Sorte & Thompson, 2007). For wetland-dependent birds like the American Bittern, changes in water levels due to altered precipitation cycles could further complicate their ability to find suitable stopover sites and breeding grounds. Given that wetlands are already among the most threatened ecosystems in North America, with significant losses due to urbanization and agriculture, climate-induced changes pose an additional challenge (Dahl, 2011). Studying the migration patterns of the American Bittern under a changing climate is essential for predicting future population trends and informing conservation strategies. By understanding how climate variables influence their movements, conservationists can prioritize habitat protection, restoration, and adaptive management to mitigate potential negative impacts.

### **Motivation:**

Bird migration range and timing is critical for ensuring species survival and ecological balance. However, climate change has altered seasonal patterns, potentially disrupting migration timing and impacting breeding success, food availability, and overall population health. Understanding these shifts can inform conservation strategies and mitigate risks to migratory bird species. Additionally, information gained from analyzing one species can be used to make inferences about the potential shifts in other similar species.

### **Data:**

#### Climate Data Online (NOAA Global Summary of the Month - U.S. Specific):

With NOAA’s Climate Data Online search, we focused on nine weather stations located between Arizona (two weather stations: Phoenix and Kingsman), California (three weather stations: Lake Tahoe, San Diego, and Santa Rosa), Oregon (two weather stations: Chiloquin and Hermiston), and Washington (two weather stations Shelton and Spokane), to map multiple points of migratory status for the American Bittern. The files were requested through eBird, and then downloaded into a readable CSV for Excel to initially portion the data into the portions necessary for our analysis. Each weather station collected data that contained the specific information for the weather station, temperature, precipitation, and wind-related metrics (of which are negligible for our testing metrics).

Variables within the Datasets contain:

* Metadata: Station name (NAME), location (LATITUDE, LONGITUDE, ELEVATION), and observation date (DATE)
* Climate Variables: Includes average and extreme temperatures (e.g., ADPT, ASLP, AWBT), precipitation statistics (e.g., DP01, DP10), and wind metrics (AWND)
* Attributes and Flags: Columns such as \*\_ATTRIBUTES and logical flags provide metadata on data quality and source
* Format: Character (station info and attribute flags), numeric (climate measures), and logical (indicator variables)

#### American Bittern - eBird Data:

This dataset was sourced from eBird and includes observations of American Bitterns (Botaurus lentiginosus) reported by citizen scientists. The dataset spans from 2000 to 2024 and includes over 70,000 observations.

* Geographic Coverage: Arizona, California, Oregon, and Washington within the U.S.
* Metadata: Observation Date Time Observations Started Latitude Longitude
* Bird Details: Observation Count Breeding Code Behavior Code
* Format: Primarily character and logical columns, with a few numeric fields (e.g., coordinates, observer count)

### **Methods**

To examine the effects climate change will have on the migratory timing of the American Bittern, we used a monthly summary of climate data from nine recorded NOAA weather stations within Arizona, California, Oregon, and Washington and observational data of their abundance within those states from eBird.

#### Data Acquiring:

Climate data within nine weather stations stretched between Arizona, California, Oregon, and Washington by NOAA Global Monthly Summaries, were recorded on a scale between 1947-2025. By using ‘R’ resources, we cleaned this data, removing unnecessary attributes for our study, converting units to be consistent throughout each dataset, and removing unnecessary data from past years to focus on data between 2000-2024.

#### Data Processing / Seasonal Aggregation:

Each weather station recorded the seasonal averages of temperature and precipitation, which were then manipulated into a cleaned version of that data, removing unnecessary wind records. With that, we created a list item, using ‘RStudio’, that combined the contents of the dataframes, to join them all together, ensuring columns and the contents of the dataframes are not duplicated or N/Aed. We then integrated a seasonal component into the data denoted by month: Winter = 12 (Dec.), 1 (Jan.), 2 (Feb.); Spring = 3 (Mar.), 4 (Apr.), 5 (May); Summer = 6 (Jun.), 7 (Jul.), 8 (Aug.); and Fall = 9 (Sep.), 10 (Oct.), 11 (Nov.). The seasonal component was also applied to the American Bittern migration dataframe, which was later converted to a tibble for later analysis.

#### Migratory Timing Analysis of American Bittern:

Observational data was provided to understand the timing of Bitterns within our four selected states, by grouping month and year to understand temporal patterns of their appearance based on the seasonal presence. Using a time series visualization to detect changes within the migratory patterns based on their seasonality between 2000 and 2024.

#### Comparing Seasonal Trends of Climate Variability and Bittern Observation:

To understand the correlation between seasonal changes for all the weather stations, we averaged the temperature and precipitation of each month per year. We then evaluated the long-term changes in bittern observation from plot layouts used before, by visually plotting the changes in the presence of climate seasonality.

To evaluate the influence of climate variability on American Bittern observations from 2000 to 2024, we fit a linear mixed effects model using the lmer function from the lme4 package in R. The response variable was the count of observations per year and state. Fixed effects included average temperature (avg\_temp) and average precipitation (avg\_precp), representing climatic conditions. To account for non-independence among observations across years and states, we included season year and state indicators (AZ, CA, OR, and WA) as random intercepts. This approach allowed us to control for unobserved heterogeneity due to temporal and spatial variation, isolating the effects of climate variables on bird observations. All data and observations were analyzed using RStudio version 2025.05.0 Build 496.

### **Results:**

##### Coding and Figures:

# Beginning Mumbo Jumbo Libraries:  
  
library(flextable)

Warning: package 'flextable' was built under R version 4.4.3

library(readr)

Warning: package 'readr' was built under R version 4.4.3

library(tidyverse)

Warning: package 'tidyverse' was built under R version 4.4.3

Warning: package 'ggplot2' was built under R version 4.4.3

Warning: package 'tidyr' was built under R version 4.4.3

Warning: package 'purrr' was built under R version 4.4.3

Warning: package 'dplyr' was built under R version 4.4.3

Warning: package 'stringr' was built under R version 4.4.3

Warning: package 'lubridate' was built under R version 4.4.3

── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
✔ dplyr 1.1.4 ✔ purrr 1.0.4  
✔ forcats 1.0.0 ✔ stringr 1.5.1  
✔ ggplot2 3.5.2 ✔ tibble 3.2.1  
✔ lubridate 1.9.4 ✔ tidyr 1.3.1  
── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
✖ purrr::compose() masks flextable::compose()  
✖ dplyr::filter() masks stats::filter()  
✖ dplyr::lag() masks stats::lag()  
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(tidymodels)

Warning: package 'tidymodels' was built under R version 4.4.3

── Attaching packages ────────────────────────────────────── tidymodels 1.3.0 ──  
✔ broom 1.0.8 ✔ rsample 1.3.0  
✔ dials 1.4.0 ✔ tune 1.3.0  
✔ infer 1.0.8 ✔ workflows 1.2.0  
✔ modeldata 1.4.0 ✔ workflowsets 1.1.0  
✔ parsnip 1.3.1 ✔ yardstick 1.3.2  
✔ recipes 1.3.0

Warning: package 'broom' was built under R version 4.4.3

Warning: package 'dials' was built under R version 4.4.3

Warning: package 'infer' was built under R version 4.4.3

Warning: package 'parsnip' was built under R version 4.4.3

Warning: package 'recipes' was built under R version 4.4.3

Warning: package 'rsample' was built under R version 4.4.3

Warning: package 'tune' was built under R version 4.4.3

Warning: package 'workflows' was built under R version 4.4.3

Warning: package 'yardstick' was built under R version 4.4.3

── Conflicts ───────────────────────────────────────── tidymodels\_conflicts() ──  
✖ purrr::compose() masks flextable::compose()  
✖ scales::discard() masks purrr::discard()  
✖ dplyr::filter() masks stats::filter()  
✖ recipes::fixed() masks stringr::fixed()  
✖ dplyr::lag() masks stats::lag()  
✖ yardstick::spec() masks readr::spec()  
✖ recipes::step() masks stats::step()

library(ggplot2)   
library(purrr)   
library(dplyr)   
library(lubridate)   
library(tidyr)   
library(gridExtra)

Warning: package 'gridExtra' was built under R version 4.4.3

Attaching package: 'gridExtra'  
  
The following object is masked from 'package:dplyr':  
  
 combine

library(stringr)  
library(cowplot)

Warning: package 'cowplot' was built under R version 4.4.3

Attaching package: 'cowplot'  
  
The following object is masked from 'package:lubridate':  
  
 stamp

library(gridExtra)

1. Create time period for seasons

# Read in all climate data  
AZ\_kingman <- read\_csv("data/Climate Data/AZ-kingman-climate.csv")

Rows: 345 Columns: 90  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (42): STATION, DATE, NAME, ADPT\_ATTRIBUTES, ASLP\_ATTRIBUTES, ASTP\_ATTRIB...  
dbl (45): LATITUDE, LONGITUDE, ELEVATION, ADPT, ASLP, ASTP, AWBT, AWND, CDSD...  
lgl (3): DYFG\_ATTRIBUTES, DYHF\_ATTRIBUTES, DYTS\_ATTRIBUTES  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

AZ\_phoenix <- read\_csv("data/Climate Data/AZ-phoenix-climate.csv")

Rows: 319 Columns: 82  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (38): STATION, DATE, NAME, AWND\_ATTRIBUTES, CDSD\_ATTRIBUTES, CLDD\_ATTRIB...  
dbl (41): LATITUDE, LONGITUDE, ELEVATION, AWND, CDSD, CLDD, DP01, DP10, DP1X...  
lgl (3): DYFG\_ATTRIBUTES, DYHF\_ATTRIBUTES, DYTS\_ATTRIBUTES  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

CA\_lake\_tahoe <- read\_csv("data/Climate Data/CA-lake-tahoe-climate.csv")

Rows: 657 Columns: 96  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (45): STATION, DATE, NAME, ADPT\_ATTRIBUTES, ASLP\_ATTRIBUTES, ASTP\_ATTRIB...  
dbl (48): LATITUDE, LONGITUDE, ELEVATION, ADPT, ASLP, ASTP, AWBT, AWND, CDSD...  
lgl (3): DYFG\_ATTRIBUTES, DYHF\_ATTRIBUTES, DYTS\_ATTRIBUTES  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

CA\_sandiego <- read\_csv("data/Climate Data/CA-sandiego-climate.csv")

Rows: 849 Columns: 84  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (39): STATION, DATE, NAME, AWND\_ATTRIBUTES, CDSD\_ATTRIBUTES, CLDD\_ATTRIB...  
dbl (42): LATITUDE, LONGITUDE, ELEVATION, AWND, CDSD, CLDD, DP01, DP10, DP1X...  
lgl (3): DYFG\_ATTRIBUTES, DYHF\_ATTRIBUTES, DYTS\_ATTRIBUTES  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

CA\_santarosa <- read\_csv("data/Climate Data/CA-santarosa\_climate.csv")

Rows: 322 Columns: 96  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (45): STATION, DATE, NAME, ADPT\_ATTRIBUTES, ASLP\_ATTRIBUTES, ASTP\_ATTRIB...  
dbl (48): LATITUDE, LONGITUDE, ELEVATION, ADPT, ASLP, ASTP, AWBT, AWND, CDSD...  
lgl (3): DYFG\_ATTRIBUTES, DYHF\_ATTRIBUTES, DYTS\_ATTRIBUTES  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

OR\_chiloquin <- read\_csv("data/Climate Data/OR-chiloquin-climate.csv")

Rows: 511 Columns: 68  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (30): STATION, DATE, NAME, CLDD\_ATTRIBUTES, DP01\_ATTRIBUTES, DP10\_ATTRIB...  
dbl (36): LATITUDE, LONGITUDE, ELEVATION, CDSD, CDSD\_ATTRIBUTES, CLDD, DP01,...  
lgl (2): DYFG\_ATTRIBUTES, DYTS\_ATTRIBUTES  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

OR\_hermiston <- read\_csv("data/Climate Data/OR-hermiston-climate.csv")

Rows: 324 Columns: 82  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (38): STATION, DATE, NAME, AWND\_ATTRIBUTES, CDSD\_ATTRIBUTES, CLDD\_ATTRIB...  
dbl (41): LATITUDE, LONGITUDE, ELEVATION, AWND, CDSD, CLDD, DP01, DP10, DP1X...  
lgl (3): DYFG\_ATTRIBUTES, DYHF\_ATTRIBUTES, DYTS\_ATTRIBUTES  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

WA\_shelton <- read\_csv("data/Climate Data/WA-shelton-climate.csv")

Rows: 321 Columns: 96  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (45): STATION, DATE, NAME, ADPT\_ATTRIBUTES, ASLP\_ATTRIBUTES, ASTP\_ATTRIB...  
dbl (48): LATITUDE, LONGITUDE, ELEVATION, ADPT, ASLP, ASTP, AWBT, AWND, CDSD...  
lgl (3): DYFG\_ATTRIBUTES, DYHF\_ATTRIBUTES, DYTS\_ATTRIBUTES  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

WA\_spokane <- read\_csv("data/Climate Data/WA-spokane-climate.csv")

Rows: 344 Columns: 88  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (41): STATION, DATE, NAME, CLDD\_ATTRIBUTES, DP01\_ATTRIBUTES, DP10\_ATTRIB...  
dbl (45): LATITUDE, LONGITUDE, ELEVATION, CDSD, CDSD\_ATTRIBUTES, CLDD, DP01,...  
lgl (2): DYFG\_ATTRIBUTES, DYTS\_ATTRIBUTES  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

# Clarify joining values:  
  
AZ\_kingman$LATITUDE <- as.numeric(AZ\_kingman$LATITUDE)  
AZ\_kingman$LONGITUDE <- as.numeric(AZ\_kingman$LONGITUDE)  
  
AZ\_phoenix$LATITUDE <- as.numeric(AZ\_phoenix$LATITUDE)  
AZ\_phoenix$LONGITUDE <- as.numeric(AZ\_phoenix$LONGITUDE)  
  
CA\_lake\_tahoe$LATITUDE <- as.numeric(CA\_lake\_tahoe$LATITUDE)  
CA\_lake\_tahoe$LONGITUDE <- as.numeric(CA\_lake\_tahoe$LONGITUDE)  
  
CA\_sandiego$LATITUDE <- as.numeric(CA\_sandiego$LATITUDE)  
CA\_sandiego$LONGITUDE <- as.numeric(CA\_sandiego$LONGITUDE)  
  
CA\_santarosa$LATITUDE <- as.numeric(CA\_santarosa$LATITUDE)  
CA\_santarosa$LONGITUDE <- as.numeric(CA\_santarosa$LONGITUDE)  
  
OR\_chiloquin$LATITUDE <- as.numeric(OR\_chiloquin$LATITUDE)  
OR\_chiloquin$LONGITUDE <- as.numeric(OR\_chiloquin$LONGITUDE)  
  
OR\_hermiston$LATITUDE <- as.numeric(OR\_hermiston$LATITUDE)  
OR\_hermiston$LONGITUDE <- as.numeric(OR\_hermiston$LONGITUDE)  
  
WA\_shelton$LATITUDE <- as.numeric(WA\_shelton$LATITUDE)  
WA\_shelton$LONGITUDE <- as.numeric(WA\_shelton$LONGITUDE)  
  
WA\_spokane$LATITUDE <- as.numeric(WA\_spokane$LATITUDE)  
WA\_spokane$LONGITUDE <- as.numeric(WA\_spokane$LONGITUDE)  
  
  
#Cleaning Data - the SAGA:  
  
AZ\_kingman <- AZ\_kingman %>%  
 select(-contains("ATTRIBUTES"))  
  
AZ\_phoenix <- AZ\_phoenix %>%  
 select(-contains("ATTRIBUTES"))  
  
CA\_lake\_tahoe <- CA\_lake\_tahoe %>%  
 select(-contains("ATTRIBUTES"))  
  
CA\_sandiego <- CA\_sandiego %>%  
 select(-contains("ATTRIBUTES"))  
  
CA\_santarosa <- CA\_santarosa %>%  
 select(-contains("ATTRIBUTES"))  
  
OR\_chiloquin <- OR\_chiloquin %>%  
 select(-contains("ATTRIBUTES"))  
  
OR\_hermiston <- OR\_hermiston %>%  
 select(-contains("ATTRIBUTES"))  
  
WA\_shelton <- WA\_shelton %>%  
 select(-contains("ATTRIBUTES"))  
  
WA\_spokane <- WA\_spokane %>%  
 select(-contains("ATTRIBUTES"))  
  
#Adding to a collective list to make join easier!  
station\_list <- list(AZ\_kingman, AZ\_phoenix, CA\_lake\_tahoe, CA\_sandiego, CA\_santarosa, OR\_chiloquin, OR\_hermiston, WA\_shelton, WA\_spokane)  
  
  
cleaned\_list <- map(station\_list, ~ select(.x, -contains("ATTRIBUTES")))  
  
  
#Binding the cleaned data!  
climate\_data <- bind\_rows(cleaned\_list)

1. Building Seasons:

# Changing into month and year  
climate\_data <- climate\_data %>%  
 mutate(  
 DATE = paste0(DATE, "-01"), # Adding a day component since it was missing!  
 DATE = as.Date(DATE, format = "%Y-%m-%d"),  
 year = year(DATE),  
 month = month(DATE)  
 )  
  
  
# Creating seasons:  
climate\_data <- climate\_data %>%  
 mutate(  
 season = case\_when(  
 month %in% c(12, 1, 2) ~ "Winter",  
 month %in% c(3, 4, 5) ~ "Spring",  
 month %in% c(6, 7, 8) ~ "Summer",  
 month %in% c(9, 10, 11) ~ "Fall"  
 ))  
  
  
# Creating seasons with the year to compare for climate  
climate\_data <- climate\_data %>%  
 mutate(  
 season\_year = if\_else(month == 12, year + 1, year)  
 )  
  
  
#Separating temp and precp to clean further!  
avg\_temp <- climate\_data %>%  
 mutate(TAVG\_F = (TAVG \* 9/5) + 32) %>%  
 group\_by(STATION, NAME, month, season, season\_year) %>%  
 summarize(avg\_temp = mean(TAVG\_F, na.rm = TRUE), .groups = "drop") %>%  
 filter(season\_year >= 2000, season\_year <= 2024)  
  
avg\_prec <- climate\_data %>%  
 group\_by(STATION, NAME, month, season, season\_year)%>%  
 summarize(avg\_prec = mean(PRCP, na.rm = TRUE), .groups = "drop") %>%  
 filter(season\_year >= 2000, season\_year <= 2024)

1. Understanding the changes of average temperature and precipitation in a U.S. map!
2. Temperature Difference between 2000 - 2024 based on seasonality

The seasonal average temperature varies respectively within each recorded season showing similarities between Fall and Winter temperatures and Spring and Summer temperatures for all nine weather stations (Fig. 1).

#TEMPERATURE!!!  
  
#creating a separte coordination dataframe  
station\_coords <- climate\_data %>%  
 select(STATION, LATITUDE, LONGITUDE) %>%  
 distinct()  
  
#joining the two to relay to the map  
avg\_temp\_map <- avg\_temp %>%  
 left\_join(station\_coords, by = "STATION")  
  
#loading the u.s. map raster file!  
us\_map <- map\_data("state")  
  
#Creating map of avg temperature for Selected stations based on seasonality!  
  
ggplot(avg\_temp\_map %>% filter(season\_year >= 2000 & season\_year <= 2024)) +  
 geom\_polygon(data = us\_map, aes(x = long, y = lat, group = group),  
 fill = "gray90", color = "white") +  
 geom\_point(aes(x = LONGITUDE, y = LATITUDE, color = avg\_temp), size = 2) +  
 scale\_color\_viridis\_c(option = "plasma", name = "Avg Temp (°F)") +  
 coord\_fixed(1.3) +  
 facet\_wrap(~season) +  
 labs(title = "Seasonal Average Temperatures by Station (2000–2024)",  
 x = "", y = "") +  
 theme\_minimal()

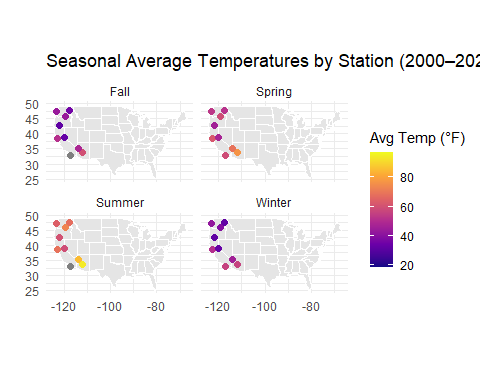


Figure 1: Average seasonal temperatures (°F) at nine selected climate stations across Arizona, California, Oregon, and Washington from 2000–2024, faceted by season.

All seasons but Spring showed several degree increase in average temperature from 2000 to 2024 (Fig. 2).

# Let's compare first just what 2000 vs 2024 looks like!  
  
# separating 2000 and 2024  
  
avg\_temp\_2000 <- avg\_temp\_map %>% filter(season\_year == 2000)  
avg\_temp\_2024 <- avg\_temp\_map %>% filter(season\_year == 2024)  
  
# Create the ggplot for 2000  
temp\_2000 <- ggplot(avg\_temp\_2000) +  
 geom\_polygon(data = us\_map, aes(x = long, y = lat, group = group),  
 fill = "gray90", color = "white") +  
 geom\_point(aes(x = LONGITUDE, y = LATITUDE, color = avg\_temp), size = 2) +  
 scale\_color\_viridis\_c(option = "plasma", name = "Avg. Temp. (°F)") +  
 coord\_fixed(1.3) +  
 facet\_wrap(~season) +  
 labs(title = "Seasonal Average Temperatures (2000)",  
 x = "",  
 y = "")+  
 theme\_minimal()

# Create the ggplot for 2024  
temp\_2024 <- ggplot(avg\_temp\_2024) +  
 geom\_polygon(data = us\_map, aes(x = long, y = lat, group = group),  
 fill = "gray90", color = "white") +  
 geom\_point(aes(x = LONGITUDE, y = LATITUDE, color = avg\_temp), size = 2) +  
 scale\_color\_viridis\_c(option = "plasma", name = "Avg. Temp. (°F)") +  
 coord\_fixed(1.3) +  
 facet\_wrap(~season) +  
 labs(title = "Seasonal Average Temperatures (2024)",  
 x = "", y = "") +  
 theme\_minimal()

# Arrange them to be side by side!  
  
plot\_grid(temp\_2000, temp\_2024, labels=c("", ""), ncol = 1, nrow = 2, widths = c(1), align = "v")

Warning in as\_grob.default(plot): Cannot convert object of class numeric into a  
grob.

Warning: Graphs cannot be vertically aligned unless the axis parameter is set.  
Placing graphs unaligned.

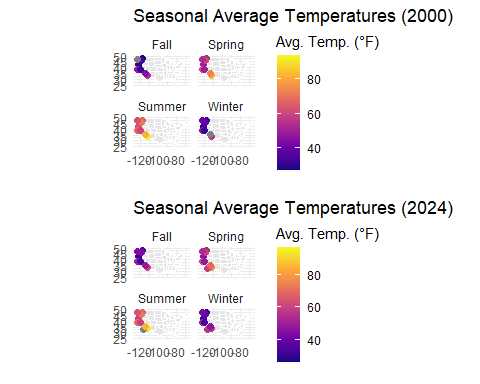


Figure 2: Average seasonal temperatures (°F) at nine selected weather stations across Arizona, California, Oregon, and Washington combined to show difference side by side. Faceted by season, featuring 2000 and 2024 temperature data.

1. Precipitation Difference between 2000 - 2024 based on seasonality

The precipitation average that is experienced between 2000-2024 for all nine weather stations have no stark difference in patterns across all four seasons (Fig. 3), although it is slightly wetter across the western coast weather stations.

#PRECIPITATION!!!  
  
#joining the two to relay to the map  
avg\_prec\_map <- avg\_prec %>%  
 left\_join(station\_coords, by = "STATION")  
  
#Creating map of avg precipitation for Selected stations based on seasonality!  
  
ggplot(avg\_prec\_map %>% filter(season\_year >= 2000 & season\_year <= 2024)) +  
 geom\_polygon(data = us\_map, aes(x = long, y = lat, group = group),  
 fill = "gray90", color = "white") +  
 geom\_point(data = avg\_prec\_map %>%  
 filter(season\_year >= 2000 & season\_year <= 2024),  
 aes(x = LONGITUDE, y = LATITUDE, color = avg\_prec), size = 2) +  
 scale\_color\_viridis\_c(option = "plasma", name = "Avg. Prec. (100 μm)") +  
 coord\_fixed(1.3) +  
 facet\_wrap(~season) +  
 labs(title = "Seasonal Average Precipitation by Station (2000–2024)",  
 x = "", y = "") +  
 theme\_minimal()

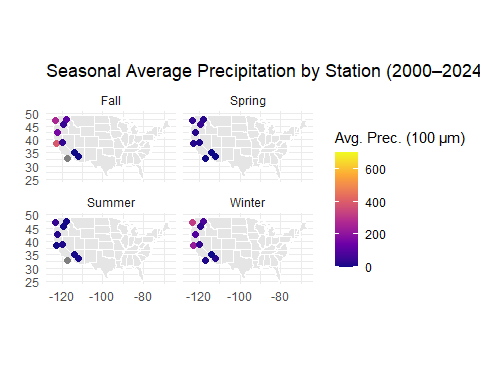


Figure 3: Average seasonal precipitation (100μm) at nine selected weather stations across Arizona, California, Oregon, and Washington from 2000–2024, faceted by season.

Precipitation in the all four seasons, Winter, Spring, Summer, and Fall, has dramatically increased from 2000 to 2024, in the nine respective weather stations (Fig. 4).

# Let's compare first just what 2000 vs 2024 looks like!  
  
# separating 2000 and 2024  
  
avg\_prec\_2000 <- avg\_prec\_map %>% filter(season\_year == 2000)  
avg\_prec\_2024 <- avg\_prec\_map %>% filter(season\_year == 2024)  
  
# Create the ggplot for 2000  
prec\_2000 <- ggplot(avg\_prec\_2000) +  
 geom\_polygon(data = us\_map, aes(x = long, y = lat, group = group),  
 fill = "gray90", color = "white") +  
 geom\_point(aes(x = LONGITUDE, y = LATITUDE, color = avg\_prec), size = 2) +  
 scale\_color\_viridis\_c(option = "plasma", name = "Avg. Prec. (100 μm)") +  
 coord\_fixed(1.3) +  
 facet\_wrap(~season) +  
 labs(title = "Seasonal Average Precipitation (2000)",  
 x = "", y = "") +  
 theme\_minimal()

# Create the ggplot for 2024  
prec\_2024 <- ggplot(avg\_prec\_2024) +  
 geom\_polygon(data = us\_map, aes(x = long, y = lat, group = group),  
 fill = "gray90", color = "white") +  
 geom\_point(aes(x = LONGITUDE, y = LATITUDE, color = avg\_prec), size = 2) +  
 scale\_color\_viridis\_c(option = "plasma", name = "Avg. Prec. (100 μm)") +  
 coord\_fixed(1.3) +  
 facet\_wrap(~season) +  
 labs(title = "Seasonal Average Precipitation (2024)",  
 x = "", y = "") +  
 theme\_minimal()

# Arrange them to be side by side!  
  
plot\_grid(prec\_2000, prec\_2024, labels=c("", ""), ncol = 1, nrow = 2, widths = c(1), align = "v")

Warning in as\_grob.default(plot): Cannot convert object of class numeric into a  
grob.

Warning: Graphs cannot be vertically aligned unless the axis parameter is set.  
Placing graphs unaligned.

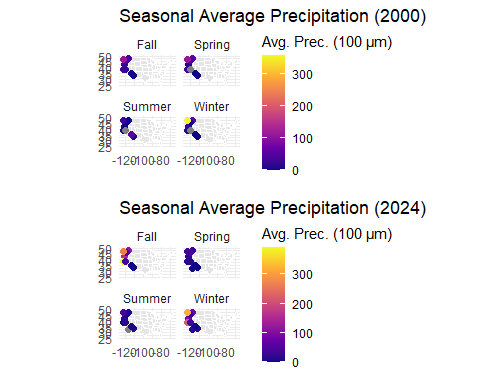


Figure 4: Average seasonal precipitation (100μm) at nine selected weather stations across Arizona, California, Oregon, and Washington combined to show difference side by side. Faceted by season, featuring 2000 and 2024 precipitation data.

1. Bird migration pattern time!

Based on eBird data, we are focusing in the Western Pacific region of the U.S. (Arizona, California, Oregon, and Washington). Based on the eBird data set variables for understanding migratory codes:

# read in bird data:  
  
amebit\_data <- read\_csv("data/AB Data/amebit\_data.csv")

New names:  
• `` -> `...22`  
• `` -> `...23`  
• `` -> `...24`  
• `` -> `...25`  
• `` -> `...26`  
• `` -> `...27`  
• `` -> `...28`  
• `` -> `...29`  
• `` -> `...30`  
• `` -> `...31`  
• `` -> `...32`  
• `` -> `...33`  
• `` -> `...34`  
• `` -> `...35`  
• `` -> `...36`  
• `` -> `...37`  
• `` -> `...38`  
• `` -> `...39`  
• `` -> `...40`  
• `` -> `...41`  
• `` -> `...42`  
• `` -> `...43`  
• `` -> `...44`  
• `` -> `...45`  
• `` -> `...46`  
• `` -> `...47`  
• `` -> `...48`  
• `` -> `...49`  
• `` -> `...50`  
• `` -> `...51`  
• `` -> `...52`  
• `` -> `...53`  
• `` -> `...54`  
• `` -> `...55`

Warning: One or more parsing issues, call `problems()` on your data frame for details,  
e.g.:  
 dat <- vroom(...)  
 problems(dat)

Rows: 70649 Columns: 55  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (20): OBSERVATION DATE, OBSERVATION COUNT, BREEDING CODE, BREEDING CATE...  
dbl (4): BCR CODE, LATITUDE, LONGITUDE, NUMBER OBSERVERS  
lgl (30): AGE/SEX, ...22, ...28, ...29, ...30, ...31, ...32, ...33, ...34, ...  
time (1): TIME OBSERVATIONS STARTED  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

View(amebit\_data)  
#I already precleaned it because of how much data there was for the American Bittern, so I don't have to clean it that much!  
  
  
# Cleaning the Data to focus on what we want:  
  
amebit\_monthly <- amebit\_data %>%  
 mutate(ob\_date = mdy(`OBSERVATION DATE`),  
 year = year(ob\_date),  
 month = month(ob\_date, label = TRUE),  
 state\_abb = state.abb[match(STATE, state.name)]  
 ) %>%  
 group\_by(year, month, state\_abb) %>%  
 summarize(observations = n(), .groups = "drop") %>%  
 filter(year >= 2000, year <= 2024)

ggplot(amebit\_monthly, aes(x = state\_abb, y = observations, fill = state\_abb)) +  
 geom\_boxplot() +  
 scale\_fill\_brewer(palette = "Set3",  
 name = "State Name",  
 labels = c("AZ" = "Arizona",  
 "CA" = "California",  
 "OR" = "Oregon",  
 "WA" = "Washington")) +  
 labs(title = "Distribution of eBird Bittern Observations by State from 2000 - 2024",  
 x = "State",  
 y = "Observation Count") +  
 theme\_minimal()

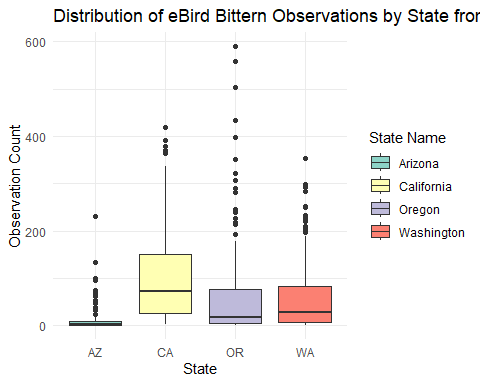


Figure 5: Distribution of American Bittern eBird observations within Arizona, California, Oregon, and Washington. Recorded over 2000-2024.

# Visualize time:  
  
  
amebit\_monthly %>%  
 filter(state\_abb %in% c("CA", "AZ", "WA", "OR")) %>%  
 ggplot(aes(x = month, y = observations, group = year, color = as.factor(year))) +  
 geom\_line(alpha = 0.5) +  
 facet\_wrap(~ state\_abb, ncol = 2) +  
 labs(  
 title = "Seasonal Observation Trends of American Bittern (2000–2024)",  
 x = "Month",  
 y = "Number of Observations",  
 color = "Year"  
 ) +  
 theme\_minimal() +  
 theme(  
 axis.text.x = element\_text(angle = 45, hjust = 1),  
 strip.text = element\_text(face = "bold")  
 )

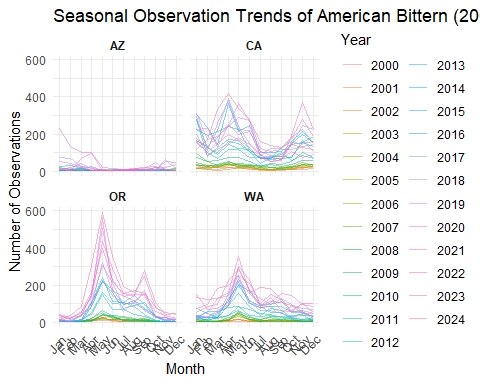


Figure 6: Observation trends of American Bitterns measured monthly from 2000 to 2024 within the four selected states: Arizona, California, Oregon, and Washington. Measured by eBird data analysts.

1. Looking at seasonal changes over time

annual\_summary <- climate\_data %>%  
 mutate(  
 TAVG\_F = (TAVG \* 9/5) + 32,  
 PRCP\_mm = PRCP / 10,  
 state\_abb = str\_extract(NAME, "[A-Z]{2}(?=\\sUS)")  
 ) %>%  
 filter(state\_abb %in% c("AZ", "OR", "WA", "CA")) %>%  
 group\_by(season\_year, season, state\_abb) %>%  
 summarize(  
 avg\_temp = mean(TAVG\_F, na.rm = TRUE),  
 avg\_precp = mean(PRCP\_mm, na.rm = TRUE),  
 .groups = "drop"  
 ) %>%  
 filter(season\_year >= 2000, season\_year <= 2024)  
  
# Reshaping to make plotting easier to understand shifts in Temp and Precip!  
  
long\_annual <- annual\_summary %>%  
 pivot\_longer(cols = c(avg\_temp, avg\_precp),  
 names\_to = "variable", values\_to = "value")  
  
  
# Plotting:  
  
ggplot(long\_annual, aes(x = season\_year, y = value, color = season)) +  
 geom\_line(linewidth = 1) +  
 facet\_grid(variable ~ state\_abb, scales = "free\_y",  
 labeller = labeller(variable = c(  
 avg\_temp = "Avg Temp (°F)",  
 avg\_precp = "Avg Precip (mm)"  
 ))) +  
 theme(panel.spacing = unit(1.2, "lines")) +  
 scale\_x\_continuous(breaks = seq(2000, 2024, by = 4)) +  
 theme\_minimal() +  
 theme(  
 axis.text.x = element\_text(angle = 45, hjust = 1)  
 ) +  
 labs(  
 title = "Seasonal Climate Variation by State (2000–2024)",  
 x = "Year", y = NULL, color = "Season"  
 ) +  
 scale\_color\_brewer(palette = "Set1")

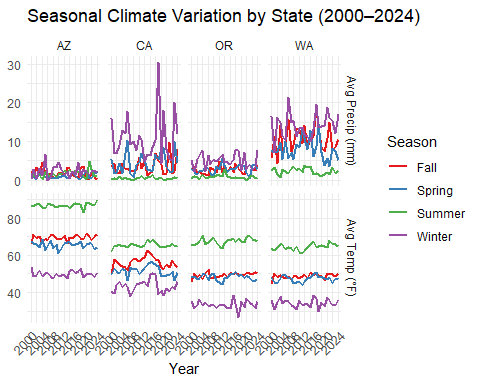


Figure 7: Seasonal trends of average temperature ((°F)) and precipitation (mm) as recorded by the nine weather stations in Arizona, California, Oregon, and Washington. Recorded between 2000 and 2024.

1. Let’s combine them to see if there is a difference!

The correlation graph displays the relationship between average seasonal temperature (°F) and the number of Bittern observations, with separate linear trends plotted for each season. The visualized data reveals distinct seasonal patterns in both temperature and observation count (Fig. 7). Each point representing the yearly average temperature and respective count of American Bitterns.

# Relationship between climate and bittern evolving together!  
amebit\_seasonal <- amebit\_data %>%  
 mutate(  
 ob\_date = mdy(`OBSERVATION DATE`),  
 year = year(ob\_date),  
 month = month(ob\_date),  
 season = case\_when(  
 month %in% c(12, 1, 2) ~ "Winter",  
 month %in% c(3, 4, 5) ~ "Spring",  
 month %in% c(6, 7, 8) ~ "Summer",  
 month %in% c(9, 10, 11) ~ "Fall"  
 ),  
 season\_year = if\_else(month == 12, year + 1, year),  
 state\_abb = state.abb[match(STATE, state.name)]  
 ) %>%  
 filter(state\_abb %in% c("CA", "AZ", "WA", "OR")) %>%  
 group\_by(season\_year, season, state\_abb) %>%  
 summarize(observations = n(), .groups = "drop") %>%  
 filter(season\_year >= 2000, season\_year <= 2024)  
  
# Combining all data sets made so far:  
combined\_data <- left\_join(amebit\_seasonal, annual\_summary,  
 by = c("season\_year", "season", "state\_abb"))  
  
combined\_long <- combined\_data %>%  
 pivot\_longer(cols = c(avg\_temp, avg\_precp),  
 names\_to = "variable", values\_to = "value")

ggplot(combined\_long %>% filter(variable == "avg\_temp"),   
 aes(x = value, y = observations, color = season)) +  
 geom\_point(alpha = 0.7) +  
 geom\_text(data = combined\_long %>%  
 filter(variable == "avg\_temp", season\_year %in% c(2000, 2024)),  
 aes(label = season\_year),  
 vjust = -0.5, size = 3, show.legend = FALSE) +  
 facet\_wrap(~ state\_abb, scales = "free\_x") +  
 labs(  
 title = "Seasonal Average Temperature and American Bittern Observations",  
 x = "Average Temperature (°F)",  
 y = "Number of Observations",  
 color = "Season"  
 ) +  
 theme\_minimal() +  
 theme(  
 axis.text.x = element\_text(angle = 0, hjust = 1),  
 strip.text = element\_text(size = 12),  
 legend.position = "bottom"  
 ) +  
 scale\_color\_brewer(palette = "Set1")

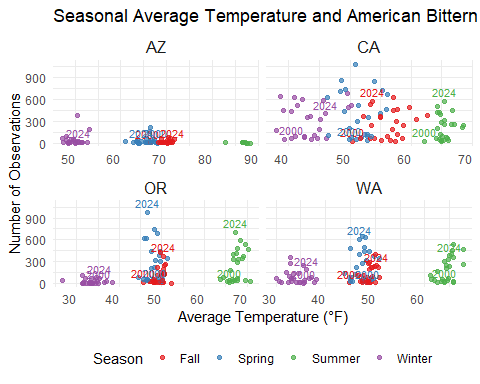


Figure 8: Average temperature of year (2000-2024) in relation towards American Bittern observation count faceted by seasonality. Each point represents a year ranging from 2000-2024, with it’s average temperature measured, by season, and number of recorded American Bittern observations. Recorded within Arizona, California, Oregon, and Washington.

The correlation graph displays the relationship between average seasonal precipitation (mm) and the number of Bittern observations, with separate linear trends plotted for each season. The visualized data reveals distinct seasonal patterns in both precipitation and observation count (Fig. 8). Each point representing the yearly average precipitation and respective count of American Bitterns.

# Visualize precipitation seasonality and bittern data:  
  
ggplot(combined\_long %>% filter(variable == "avg\_precp"),   
 aes(x = value, y = observations, color = season)) +  
 geom\_point(alpha = 0.7) +  
 geom\_text(data = combined\_long %>%  
 filter(variable == "avg\_precp", season\_year %in% c(2000, 2024)),  
 aes(label = season\_year),  
 vjust = -0.5, size = 3, show.legend = FALSE) +  
 facet\_wrap(~ state\_abb, scales = "free\_x") +  
 labs(  
 title = "Seasonal Average Precipitation and American Bittern Observations",  
 x = "Average Precipitation (mm)",  
 y = "Number of Observations",  
 color = "Season"  
 ) +  
 theme\_minimal() +  
 theme(  
 axis.text.x = element\_text(angle = 0, hjust = 1),  
 strip.text = element\_text(size = 12),  
 legend.position = "bottom"  
 ) +  
 scale\_color\_brewer(palette = "Set1")

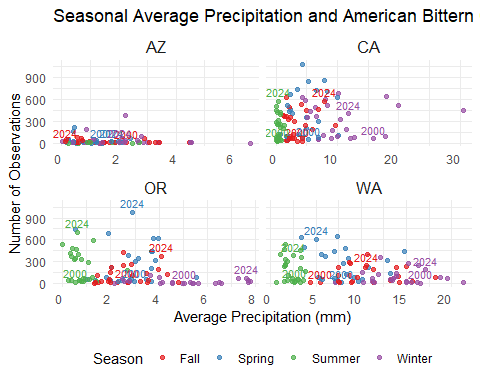


Figure 9: Average precipitation (mm) of year (2000-2024) in relation towards American Bittern observation count faceted by seasonality. Each point represents a year ranging from 2000-2024, with it’s average precipitation measured, by season, and number of recorded American Bittern observations. Recorded within Arizona, California, Oregon, and Washington.

During Winter, average temperatures ranged from approximately 37°F to 47°F, with Bittern observations generally below 1,000 (Fig. 7). A slight positive trend was noted, suggesting a weak relationship between increasing temperature and higher observations (Fig. 9). In Spring, temperatures were around 50°F to 58°F (Fig. 7). Although this season had the highest overall number of Bittern observations—with some counts exceeding 2,000—there was a slight negative trend in the linear fit, indicating that the number of Bittern observations tended to decrease slightly with increasing temperature (Fig. 9). Fall showed a moderate positive trend, with observations increasing alongside temperatures ranging from 52°F to 59°F (Fig. 7). However, overall observations remained lower compared to Spring and Summer (Fig. 7). Summer displayed the strongest positive trend. As temperatures increased from 67°F to 74°F, Bittern observations increased significantly (Fig. 9). This season also showed more consistent and clustered high observations, suggesting a stronger relationship between higher temperatures and increased presence of Bitterns (Fig. 9).

##### **ANOVA Tests:**

# 1. creating a linear model test  
combined\_data$season\_year <- as.factor(combined\_data$season\_year)  
combined\_data$state\_abb <- as.factor(combined\_data$state\_abb)  
  
  
lm\_mod <- lm(observations ~ avg\_temp \* season\_year + avg\_precp \* season\_year + state\_abb, data = combined\_data)  
  
library(car)

Warning: package 'car' was built under R version 4.4.3

Loading required package: carData

Warning: package 'carData' was built under R version 4.4.3

Attaching package: 'car'

The following object is masked from 'package:dplyr':  
  
 recode

The following object is masked from 'package:purrr':  
  
 some

print(Anova(lm\_mod, type = 2))

Anova Table (Type II tests)  
  
Response: observations  
 Sum Sq Df F value Pr(>F)   
avg\_temp 45597 1 2.3702 0.1247   
season\_year 6633544 24 14.3678 <2e-16 \*\*\*  
avg\_precp 2225 1 0.1157 0.7340   
state\_abb 3408960 3 59.0684 <2e-16 \*\*\*  
avg\_temp:season\_year 219475 24 0.4754 0.9840   
season\_year:avg\_precp 371196 24 0.8040 0.7315   
Residuals 5732734 298   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

library(lme4)

Warning: package 'lme4' was built under R version 4.4.3

Loading required package: Matrix

Attaching package: 'Matrix'

The following objects are masked from 'package:tidyr':  
  
 expand, pack, unpack

lmer\_mod <- lmer(observations ~ avg\_temp + avg\_precp + (1 | season\_year) + (1 | state\_abb), data = combined\_data)  
  
print(summary(lmer\_mod))

Linear mixed model fit by REML ['lmerMod']  
Formula: observations ~ avg\_temp + avg\_precp + (1 | season\_year) + (1 |   
 state\_abb)  
 Data: combined\_data  
  
REML criterion at convergence: 4815.1  
  
Scaled residuals:   
 Min 1Q Median 3Q Max   
-2.2167 -0.6062 -0.1211 0.4991 4.4588   
  
Random effects:  
 Groups Name Variance Std.Dev.  
 season\_year (Intercept) 17451 132.1   
 state\_abb (Intercept) 14941 122.2   
 Residual 17908 133.8   
Number of obs: 376, groups: season\_year, 25; state\_abb, 4  
  
Fixed effects:  
 Estimate Std. Error t value  
(Intercept) 81.1051 85.1900 0.952  
avg\_temp 1.4967 0.8374 1.787  
avg\_precp -0.7658 2.3499 -0.326  
  
Correlation of Fixed Effects:  
 (Intr) avg\_tm  
avg\_temp -0.611   
avg\_precp -0.462 0.635

ggplot(combined\_data, aes(x = avg\_temp, y = observations)) +  
 geom\_point(alpha = 0.4) +  
 geom\_smooth(method = "lm", se = TRUE, color = "steelblue1") +  
 labs(  
 title = "Effect of Average Temperature on Bittern Observations",  
 x = "Average Temperature (°F)",  
 y = "American Bittern Observations"  
 ) +  
 theme\_minimal()

`geom\_smooth()` using formula = 'y ~ x'

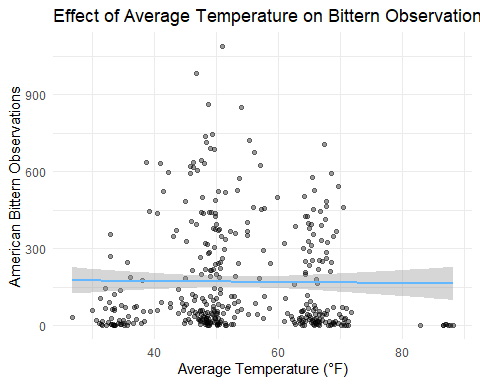


Figure 10: Average temperature (mm) of year (2000-2024) effect on American Bittern observation count taken during that time period, within Arizona, California, Oregon, and Washington. Points represent individual observations, and the blue line shows a linear trend with 95% confidence interval.

ggplot(combined\_data, aes(x = factor(season\_year), y = observations)) +  
 geom\_boxplot(fill = "dodgerblue") +  
 labs(title = "Distribution of Observations by Year",  
 x = "Year",  
 y = "Observations") +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 35, hjust = 1))

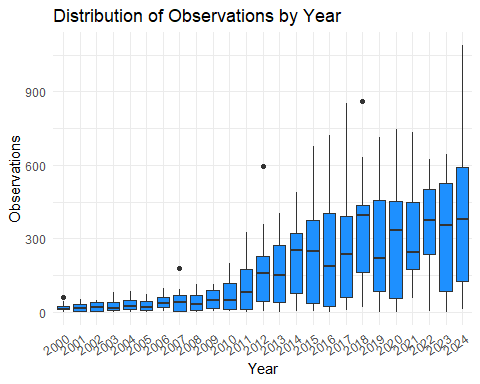


Figure 11: Interannual variation in American Bittern observations from 2000 to 2024. Spread of observations across years, highlighting changes over time.

library(ggeffects)

Warning: package 'ggeffects' was built under R version 4.4.3

Attaching package: 'ggeffects'

The following object is masked from 'package:cowplot':  
  
 get\_title

# Predict marginal effects of temperature  
pred\_temp <- ggpredict(lmer\_mod, terms = "avg\_temp")  
  
plot(pred\_temp) +  
 labs(  
 title = "Predicted Observations vs. Avg Temp (Mixed Model)",  
 x = "Average Temperature (°F)",  
 y = "Predicted Observations"  
 )

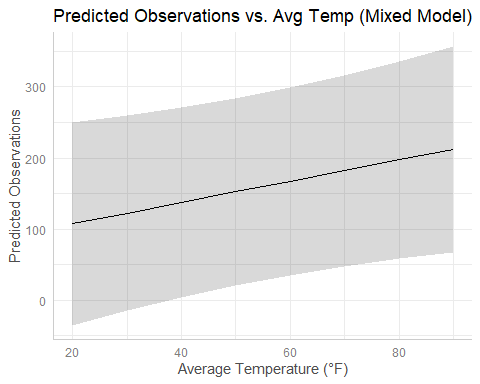


Figure 12: Predicted effect of average temperature on American Bittern observations from the mixed effects model. Marginal predictions are based on fixed effects, holding precipitation, year, and state constant.

library(sjPlot)

Warning: package 'sjPlot' was built under R version 4.4.3

Install package "strengejacke" from GitHub (`devtools::install\_github("strengejacke/strengejacke")`) to load all sj-packages at once!

Attaching package: 'sjPlot'

The following objects are masked from 'package:cowplot':  
  
 plot\_grid, save\_plot

# Plot random intercepts for year and state  
plot\_model(lmer\_mod, type = "re", sort.est = TRUE, title = "Random Effects: Year and State on Bittern Observations")

Sorting each group of random effects ('sort.all') is not possible when 'facets = TRUE'.

Sorting each group of random effects ('sort.all') is not possible when 'facets = TRUE'.

[[1]]

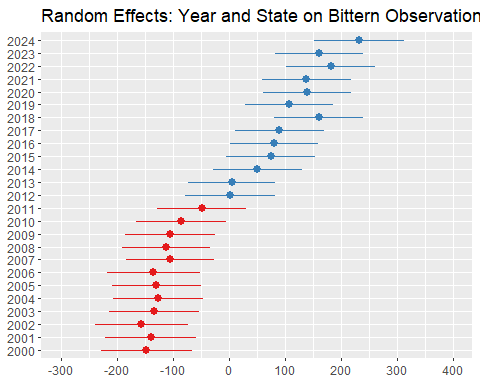


Figure 13: Random intercept effects of the year [[1]] and state [[2]] on American Bittern observations from the mixed model. Positive values (blue) indicate above-average observation rates relative to the grand mean. Negative values (red) indicative a below-average observation rates relative to the grand mean.

[[2]]

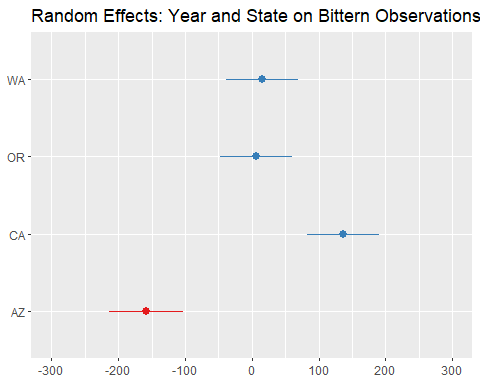


Figure 13: Random intercept effects of the year [[1]] and state [[2]] on American Bittern observations from the mixed model. Positive values (blue) indicate above-average observation rates relative to the grand mean. Negative values (red) indicative a below-average observation rates relative to the grand mean.

### **Discussion:**

These results suggest that seasonal temperature plays a variable role in Bittern observations, likely reflecting differences in migratory behavior, breeding activity, and habitat use throughout the year.

The strong positive trend in Summer may indicate that Bitterns are more active or easier to observe during warmer months, potentially due to breeding activity or greater availability of wetland habitats. Similarly, the positive trend in Fall could reflect pre-migration behavior or favorable foraging conditions as temperatures cool.

The Spring season, despite having the highest individual observation counts, showed a slight negative trend. This may reflect a peak in migration or nesting activity occurring at lower spring temperatures, after which observations decline as temperatures rise.

Winter observations were consistently low, which aligns with the species’ known migratory behavior—many populations leave colder regions during winter months. The weak positive correlation may suggest that in relatively milder winters, some individuals remain present or return earlier.

Overall, this analysis underscores how temperature and seasonality can influence bird detection and presence. Further investigation into habitat variables, precipitation, and food availability could clarify the mechanisms behind these seasonal patterns. Understanding these relationships is especially important in the context of climate change, which may shift seasonal temperature regimes and, in turn, influence Bittern migration and habitat use.

### **References:**

Status and trends of wetlands in the conterminous United States 2004 to 2009. US Department of the Interior, US Fish and Wildlife Service, Fisheries and Habitat Conservation.

eBird. (2025). eBird: An online database of bird distribution and abundance [web application]. eBird, Cornell Lab of Ornithology, Ithaca, New York. Available: http://www.ebird.org. (Accessed: April 16, 2025).

La Sorte, F. A., & Thompson III, F. R. (2007). Poleward shifts in winter ranges of North American birds. Ecology, 88(7), 1803-1812.