

INTRODUCTION

The strong impacts of climate change on biodiversity emphasize the urgent necessity to comprehend how environmental gradients shape species distributions and diversity. Bumblebees (*Bombus*), are essential pollinators that are particularly sensitive to climate variations, and their species diversity across latitudinal gradients requires increasing attention. Previous studies have found that species diversity tends to decrease with increasing latitude, likely due to harsher climates and reduced habitat availability (Kerr et al., 2015). However, in which way these patterns of diversity shift in response to the rapid and increasing effects of climate change remains unclear.

Two competing hypotheses emerge from existing literature. The first, it is expected that species diversity will be lower at higher latitudes, where environmental conditions are more harsh and resources are limited, resulting in the population to be more vulnerable to extinction (Willis & Bhagwat, 2009). Alternatively, some studies suggest that northern bumblebee populations may exhibit unique adaptations, due to evolution, that allow them to thrive despite challenging conditions, possibly sustaining diversity even as the climate changes (Sgrò et al., 2011). By examining these contrasting possibilities, this project seeks to explore the relationship between *Bombus* species diversity and latitude, with a specific focus on the potential impacts of climate change on species richness and distribution.

CODE SECTION 1: DATA EXPLORATION

```
#### DATA EXPLORATION ####  
  
## ---- Load Packages ----  
  
# Assume all libraries have been installed.  
#If not, install.packages("nameOfpackage")  
  
library(tidyverse)
```

```

library(vegan)
library(dplyr) # Used for data manipulations
library(ggplot2) # Used for visualizations
library(plotly) # Used for visualizations (specifically ggplotly())
library(sf)
library(maps)
library(mapdata)
library(cowplot) # For dual-axis plotting

## ---- Load NASA Temp Data and Clean ----

# Read and Load NASA Average Temperature Data
nasa_data <- read.csv("nasa_data.csv", header = TRUE, stringsAsFactors = FALSE)

# Convert the data frame into proper columns
nasa_data <- as.data.frame(nasa_data)

nasa_data <- nasa_data %>%
  filter(!is.na(X) & !is.na(X.1) & !is.na(X.2)) # Remove NA values

# Rename columns
names(nasa_data)[names(nasa_data) == "X"] <- "year"
names(nasa_data)[names(nasa_data) == "X.1"] <- "No_Smoothing"
names(nasa_data)[names(nasa_data) == "X.2"] <- "Lowess(5)"

nasa_data <- nasa_data[-1, ] # Remove first row

## ---- Load Data from BOLD API ----

# Extracting Bombus data from Canada + USA in TSV data format
# Extracted data in TSV is also saved in /data folder
BOLDdata <- read_tsv(file =
"http://www.boldsystems.org/index.php/API_Public/specimen?taxon=Bombus&geo=Canada,United%20States&format=tsv")

BOLDdata_filtered <- BOLDdata %>%
  filter(country == "Canada" | country == "United States")

## ---- Data Cleaning + Processing ----

# Creating data frame with necessary variables from BOLDdata
bombus <- BOLDdata_filtered %>% # Extracting BOLDdata for new bombus data frame
  select(processid, species_name, lat, lon, province_state) %>% # Selecting
  processid, species_name, lat, lon, province_state variables to add into bombus
  mutate(across(where(is.character), ~ na_if(., " "))) %>% # Changing any empty
  cells in bombus to NA
  filter(!is.na(processid) & !is.na(species_name) & !is.na(lat) & !is.na(lon) &
!is.na(province_state)) # Removing NA cells

# Summarize Key Variables and Check for Data Errors

# Summarize Numerical Variables

# Latitude mean, max, min and range
mean(bombus[["lat"]])
max(bombus[["lat"]])
min(bombus[["lat"]])

```

```

range(bombus[['lat']])

# Longitude mean, max, min and range
mean(bombus[['lon']])
max(bombus[['lon']])
min(bombus[['lon']])
range(bombus[['lon']])
# The ranges for latitude and longitude are accurate, as referenced by Google
Maps.

# Summarize species count for each specimen
species_count <- bombus %>%
  count(species_name, sort = TRUE)
summarize(species_count)

sum(species_count$n) # Check to ensure that n = 3432, for each specimen in
dataset being classified

# Location
location <- bombus %>%
  count(province_state, sort = TRUE)

sum(location$n) # Check to ensure that n = 3432, for each specimen in dataset
being classified into location

# Boxplots for outliers of numerical variables
boxplot(bombus$lat, main = "Latitude")
boxplot(bombus$lon, main = "Longitude")

# Plot Histograms
ggplot(bombus, aes(x = species_name)) +
  geom_bar(fill = "skyblue", color = "black") +
  labs(title = "Histogram of Species Types",
        x = "Species",
        y = "Count") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 60, hjust = 0.5))

ggplot(bombus, aes(x = province_state)) +
  geom_bar(fill = "skyblue", color = "black") +
  labs(title = "Histogram of Province/States",
        x = "Province/State",
        y = "Count") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 60, hjust = 0.5))

## ---- Organizing into Data Frames + Matrices ----

# Classify latitude into 3 groups (Low, Med, High)
latitude_group <- cut(bombus$lat,
                      breaks = c(25, 44, 63, 82),
                      labels = c("Low", "Mid", "High"),
                      include.lowest = TRUE)
bombus$latitude_group <- latitude_group # Add coloumn to bombus data frame

# Classify each specimen into species and by latitude groups
species_richness <- table(bombus$latitude_group, bombus$species_name)

```

```

# Extract year from last 2 digits of processid (Assuming that the year was when
the specimen was found)
year <- str_extract(bombus$processid, "\\d{2}$")
bombus$year <- year
bombus$year <- (paste0("20", bombus$year)) # Ensures that year is numeric when
adding to bombus data frame

# Compare species name with frequency count of each species with average
latitude
species_count <- bombus %>%
  group_by(species_name) %>%
  summarise(
    species_freq = n(), # Count the frequency of each species
    avg_lat = mean(lat, na.rm = TRUE) # Calculate average latitude for each
species
  )

# Find the year with the most specimen caught
year <- bombus %>%
  count(year, sort = TRUE)
year

# Counts total species records for each year
species_freq_per_year <- bombus %>%
  group_by(year) %>%
  summarise(species_count = n())

# Data frame merging average temperature data + species frequency by years in
Common
combined_df <- merge(nasa_data, species_freq_per_year, by = "year")
combined_df$No_Smoothing <- as.numeric(as.character(combined_df$No_Smoothing))
# Convert Average to Numeric

# Data cleaning for combined data frame
combined_df <- combined_df %>% # Remove Lowess(5) Column
  select(-"Lowess(5)")

```

CODE SECTION 2: ANALYSIS TO ADDRESS QUESTION

```

#### ANALYSIS TO ADDRESS QUESTION ####

## ---- Statistical Testing ----

shannon_diversity <- diversity(species_richness, index = "shannon") # Conduct
Shannon Index Test for 3 lat categories
shannon_diversity

# Spearman correlation pre-requisites
str(combined_df) # Check the structure of the combined_df dataframe

# Convert species_richness to numeric
combined_df$species_count <-
as.numeric(as.character(combined_df$species_count))

```

```

combined_df$avg_temp_change <-
as.numeric(as.character(combined_df$avg_temp_change))

# Boxplot to visualize potential outliers
boxplot(combined_df$species_count, main = "Species Richness")
boxplot(combined_df$No_Smoothing, main = "Average Temperature Change")

# Calculate Spearman correlation
spearman_corr <- cor(combined_df$species_count, combined_df$No_Smoothing,
method = "spearman")
print(spearman_corr)

## ---- Visualizations ----

### PLOT 1: Species Richness Scatter Plot Per Latitude Gradient

# Calculate species richness per latitude and species
species_richness_data <- bombus %>%
  group_by(lat, species_name) %>%
  summarise(SpeciesRichness = n(), .groups = 'drop')

# Create the scatter plot with ggplot2
scatterplot <- ggplot(data = species_richness_data, aes(x = lat, y =
SpeciesRichness, color = species_name, text = species_name)) +
  geom_point(size = 3, alpha = 0.7) +
  labs(title = "Species Richness vs Latitude",
        subtitle = "Distribution of Bombus Species Across Latitudes",
        x = "Latitude",
        y = "Species Richness") +
  theme_minimal() +
  theme(legend.position = "none") # Removes the legend

plot_1 <- ggplotly(scatterplot, tooltip = "text") # Convert the ggplot to a
plotly object
print(plot_1)

```

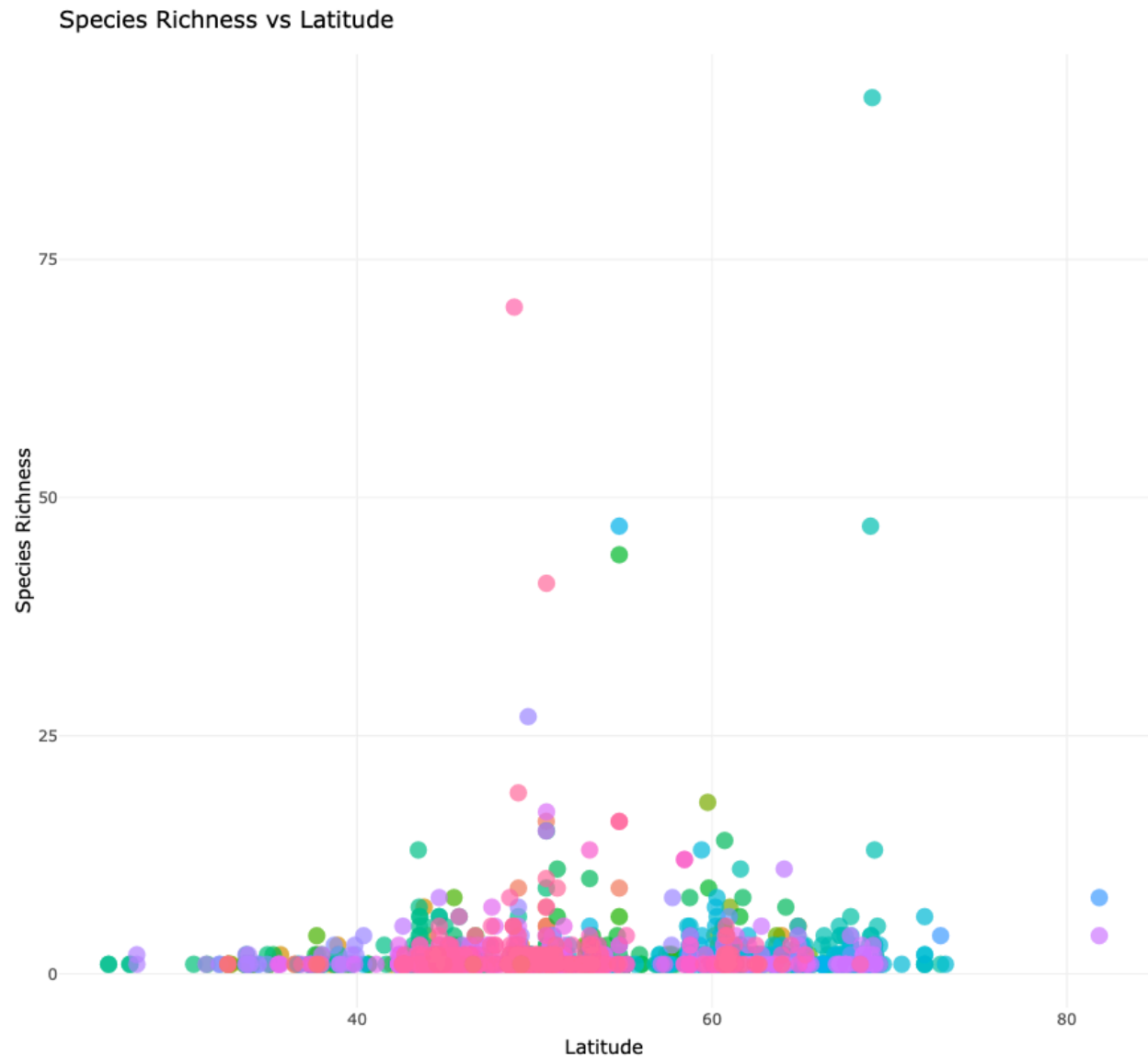


Figure 1: Scatterplot displaying species abundance and type in relation to latitude.

*Please execute the code for the figure, as it is interactive; selecting a point will display the species name.

```
### PLOT 2: Species Density Heat Map

world <- st_as_sf(map("world", plot = FALSE, fill = TRUE))

species_location_data <- bombus %>%
  filter(!is.na(lat) & !is.na(lon) & !is.na(province_state)) # Filter out
missing values

# Plot the map with a heat map
plot_2 <- ggplot() +
  # Add the map of Canada and the USA
  geom_sf(data = world, fill = "grey90", color = "white") +

  # Add heat map layer based on species density
```

```

stat_density_2d(
  data = species_location_data,
  aes(x = lon, y = lat, fill = ..level..), # 'fill' for heat map
  geom = "polygon", # Use polygons for filled density
  alpha = 0.5 # Adjust transparency
) +

# Add contour lines
geom_density_2d(
  data = species_location_data,
  aes(x = lon, y = lat),
  color = "red"
) +

# Add points for species locations
geom_point(data = species_location_data, aes(x = lon, y = lat), size = 3,
alpha = 0.7) +

# Customize the labels and theme
labs(
  title = "Species Locations with Density Heat Map in Canada and USA",
  # subtitle = "Visualizing Species Concentrations with Contour Lines",
  x = "Longitude (°W)",
  y = "Latitude (°N)",
  fill = "Density",
  color = "Province/State"
) +

# Focus on Canada and USA region
coord_sf(xlim = c(-150, -50), ylim = c(20, 85)) +

scale_fill_viridis_c(option = "plasma", name = "Density") +
scale_color_viridis_d(name = "Province/State", option = "plasma") +

# Improving overall aesthetics of map
theme_minimal() +
theme(
  plot.title = element_text(size = 16, face = "bold"),
  plot.subtitle = element_text(size = 12, face = "italic"),
  axis.title.x = element_text(size = 10),
  axis.title.y = element_text(size = 10),
  legend.position = "bottom",
  legend.key.width = unit(2, "cm")
)

print(plot_2)

```

Species Locations with Density Heat Map in Canada and USA

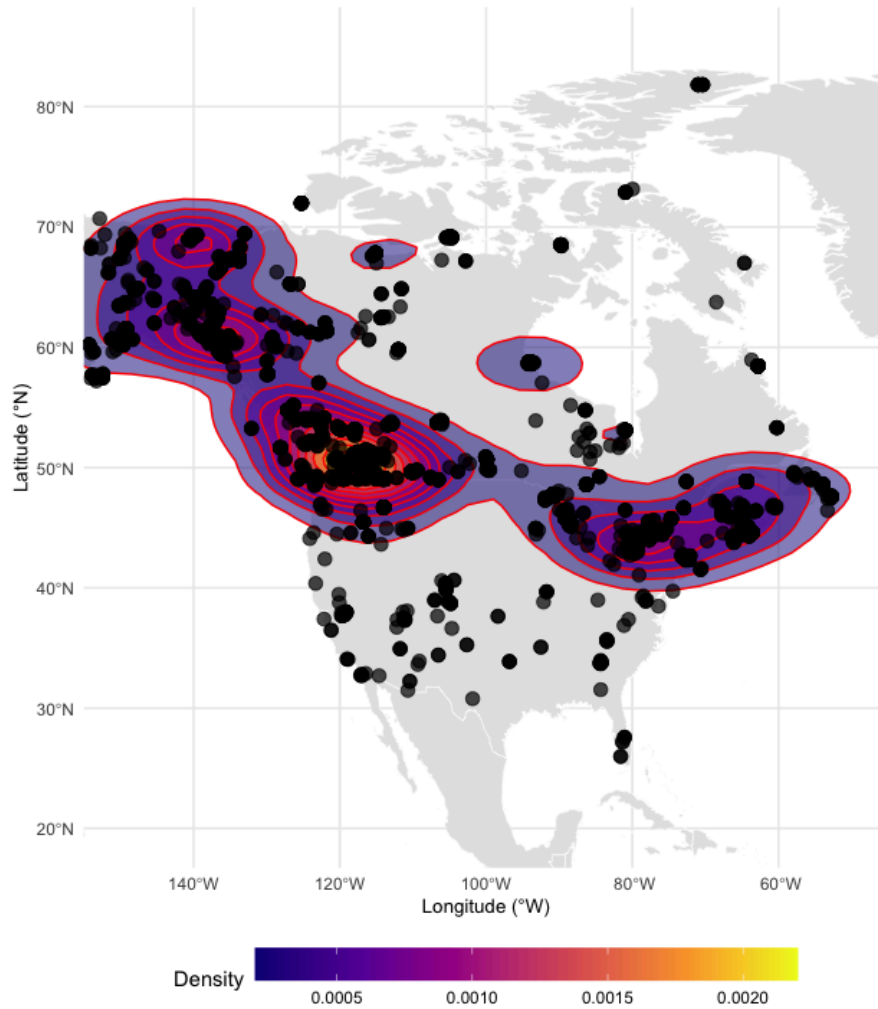


Figure 2: Species density heat map across Canada and USA.

```
### PLOT 3: Time-Series Species Richness vs. Avg. Temperature

# Species Richness over time
species_richness_plot <- ggplot(combined_df, aes(x = year, y = species_count))
+
  geom_line(color = "blue", size = 1) +
  geom_point(color = "blue", size = 2) +
  scale_y_continuous(name = "Species Richness") + # Y-axis for Species
Richness
  theme_minimal() +
  theme(axis.title.y = element_text(color = "blue", size = 12))

# Second plot: Temperature Change over time (on the same x-axis)
temperature_change_plot <- ggplot(combined_df, aes(x = year, y = No_Smoothing))
+
  geom_line(color = "red", size = 1) +
  geom_point(color = "red", size = 2) +
  scale_y_continuous(name = "Temperature Change (°C)",
```



```

sec.axis = sec_axis(~ ., name = "Temperature Change
(°C)") + # Secondary axis for temperature
  theme_minimal() +
  theme(axis.title.y.right = element_text(color = "red", size = 12))

# Combine the plots
combined_plot <- plot_grid(
  species_richness_plot + theme(axis.title.x = element_blank()), # Remove
x-axis label from first plot
  temperature_change_plot + theme(axis.title.x = element_blank()), # Remove
x-axis label from second plot
  ncol = 1, align = "v", axis = "l" # Align both plots vertically
)

# Add title
title <- ggdraw() + draw_label("Species Richness and Temperature Change Over
Time",
                                fontface = 'bold', size = 14)

# Add shared x-axis
x_axis <- ggdraw() + draw_label("Year", size = 12)

# Final plot with title and shared x-axis
plot_3 <- plot_grid(title, combined_plot, x_axis, ncol = 1, rel_heights =
c(0.1, 1, 0.05))

print(plot_3)

```

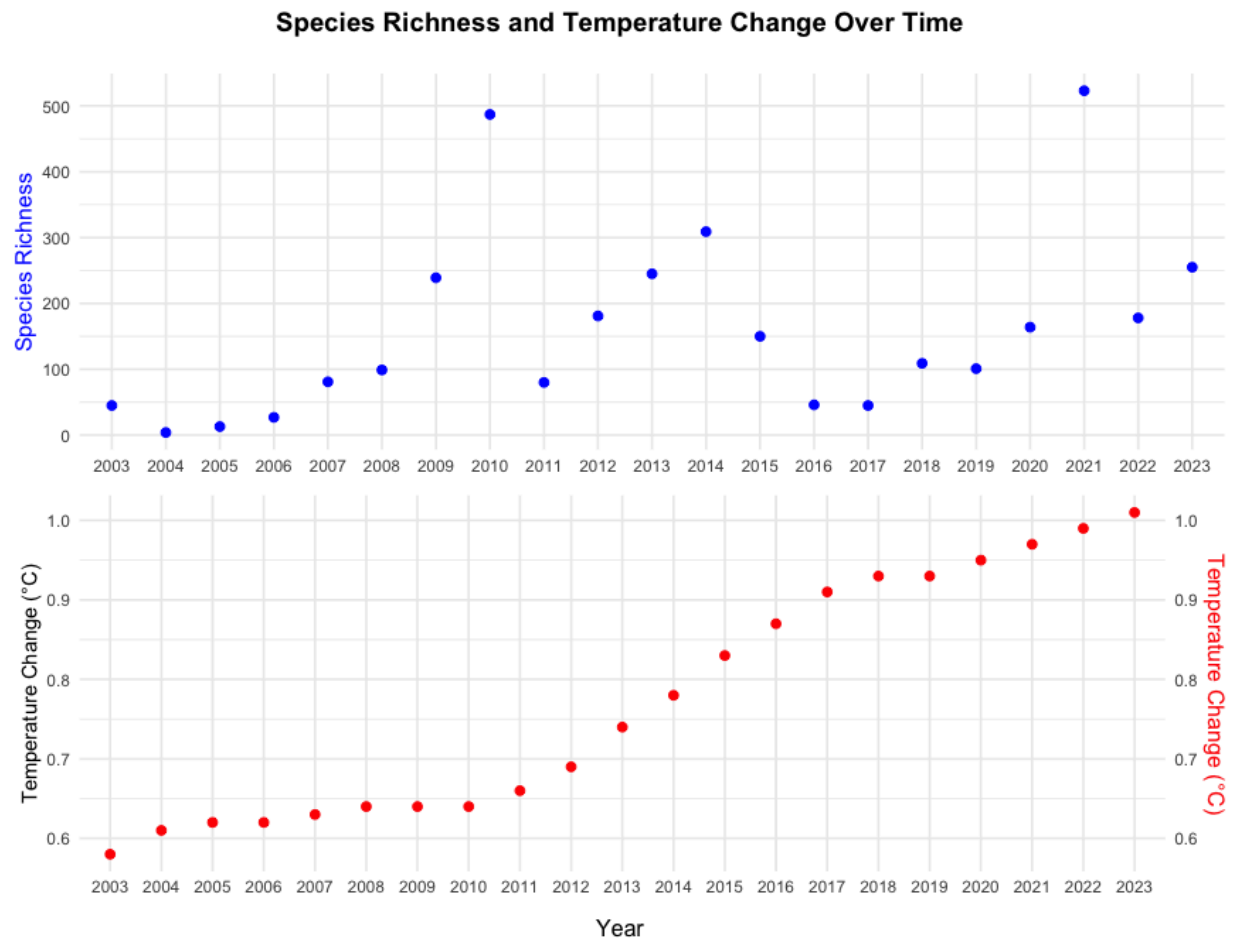


Figure 3: Scatterplot displaying species richness and average temperature change over time

DISCUSSION AND CONCLUSION

The data visualized across the three graphs provides compelling yet nuanced insights into the hypothesis that bumblebee populations at higher latitudes exhibit lower genetic diversity and are more susceptible to extinction due to climate change.

Figure 1 directly addresses the research hypothesis. The scatter plot of species richness versus latitude demonstrates a clear negative trend, with species richness declining as latitude increases. This reinforces the idea that higher latitudes, where climate conditions are more extreme, can

relate with lower species diversity. The trend line reveals variability, especially at mid-latitudes, indicating that factors beyond latitude and climate may influence species richness. This variation highlights a limitation, while latitude can indicate environmental conditions, it doesn't fully represent the complex ecological interactions that influence species diversity.

Figure 2 provides a density heat map confirming the distributional patterns of bumblebee species across Canada and USA, with higher densities in mid-latitudes, similar to Figure 1. The clustering observed may indicate suitable habitat availability at these latitudes, which tend to have a more stable climate. This spatial pattern supports Goulson's (2010) hypothesis that bumblebees have adapted to particular ecological niches, with populations narrowing at both high and low latitudes. Lower population density in higher latitudes suggests an environment less hospitable for diverse populations, possibly due to harsher climatic conditions. It should also be noted that these patterns may also be influenced by sampling bias, as some northern regions might have fewer observational records due to accessibility challenges.

In Figure 3, the comparison of species richness and temperature change over time reveals a subtle but critical point of intersection. There is a visible decrease in species richness coinciding with periods of rising temperature, lending proof to the argument that climate change plays a central role in biodiversity loss. The decline in species richness, most seen during steep temperature increases, aligns with literature suggesting bumblebee populations are sensitive to even slight climatic shifts (Sørensen et al., 2021). However, the visual overlap of the curves alone is insufficient to confirm causation, other environmental factors need to be taken into consideration, as well.

Additionally, Shannon index analysis yielded the following results: low (2.703), medium (3.362), and high (2.388) for the varying latitude ranges. These results indicate that mid-latitude regions exhibit the highest diversity, suggesting that these areas may provide a balance of environmental conditions conducive to sustaining diverse *Bombus* populations. The low diversity at higher latitudes supports the hypothesis that climatic extremities may hinder species adaptability.

Furthermore, a Spearman correlation analysis yielded a coefficient of 0.5349828 between temperature change and species richness, indicating a moderate positive relationship. This correlation suggests that as temperature increases, species richness tends to decrease, reinforcing the idea that climate change negatively impacts bumblebee biodiversity.

While these findings do suggest that latitude and climate change influence bumblebee species diversity, the project is limited by its reliance on observational data, which may be subject to geographic and temporal biases. A logical next step would involve conducting genomic analysis of bumblebee populations across various latitudes to test whether populations in higher latitudes indeed possess lower genetic diversity. Furthermore, longitudinal studies that monitor population dynamics for shifting climate regimes would provide a more thorough understanding of the species richness changes in response to climate factors.

Overall, these findings contribute valuable data to ongoing discussions regarding biodiversity loss in relation to climate change. While temperature appears to be significant in shaping bumblebee diversity, further research is necessary to understand the complexities of environmental influence.

ACKNOWLEDGMENTS

I would like to extend my appreciation to Brittany MacIntyre for her insightful suggestions and guidance in creating unique graphs and reviewing the prerequisites for statistical tests.

REFERENCES

- Bold Systems v4*. (n.d.). Retrieved October 4, 2024, from <https://www.boldsystems.org/>
- Change, N. G. C. (n.d.). *Global surface temperature*. Climate Change: Vital Signs of the Planet. Retrieved October 4, 2024, from <https://climate.nasa.gov/vital-signs/global-temperature/?intent=121>
- Goulson, D. (2010, January 1). *Bumblebees: Behaviour, ecology, and conservation*. Oxford University Press. https://www.researchgate.net/publication/305346861_Bumblebees_behaviour_ecology_and_conservation
- Kerr, J. T., Pindar, A., Galpern, P., Packer, L., Potts, S. G., Roberts, S. M., Rasmont, P., Schweiger, O., Colla, S. R., Richardson, L. L., Wagner, D. L., Gall, L. F., Sikes, D. S., & Pantoja, A. (2015). CLIMATE CHANGE. Climate change impacts on bumblebees converge across continents. *Science (New York, N.Y.)*, 349(6244), 177–180. <https://doi.org/10.1126/science.aaa7031>
- Rahimi, E., & Jung, C. (2024). Global trends in climate suitability of bees: Ups and downs in a warming world. *Insects*, 15(2), 127. <https://doi.org/10.3390/insects15020127>
- Sgrò, C. M., Lowe, A. J., & Hoffmann, A. A. (2011). Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications*, 4(2), 326–337. <https://doi.org/10.1111/j.1752-4571.2010.00157.x>
- Sørensen, H. H., MacIvor, J. S., & Liow, L. H. (2021). Climate-driven declines in bumblebee abundance and species richness: A global synthesis. *Science Advances*, 7(22), eabc3765.
- Willis, K. J., & Bhagwat, S. A. (2009, November 1). *Biodiversity and climate change*. American Association for the Advancement of Science (AAAS). https://www.researchgate.net/publication/38070460_Biodiversity_and_Climate_Change