Justin Ritenburgh
STATS 406
Final Project
How Quickly are Bees Reacting to Climate Change?

Introduction

Bumblebees are a vital part of many of Earth's ecosystems; they pollinate much of the plant life around the globe, enabling it to reproduce and maintain the populations of organisms that rely on those plants for food, shelter, building materials, and more. Recently, however, bumblebees have been threatened by a variety of factors, mostly those related to habitat loss. A primary driver of this habitat loss is climate change; as regions grow warmer or colder, they begin to fall outside of the habitable ranges for bumblebees, interfering with bumblebees' role in many ecosystems.

The question that I want to answer is about how quickly bumblebees are responding to the changing climate. Scientists have been observing rising average global temperatures for decades, but in recent years we have started to see more and more prominent effects from climate change, such as increases in the frequency and severity of hurricanes, droughts, and wildfires. Has the behavior of bumblebees noticeably changed in a similar way? Has the severity of their response to the changing climate scaled with the severity of the impact that we have observed?

In order to answer this question, I will analyse data collected about the location and temperature ranges of many species of bumblebees over four periods of time and compare them to see firstly, if there were changes between the time periods, and secondly, if those changes increased as the time periods approach the modern day. In theory, if the changes between time periods grow more severe as they approach the modern day, it would be confirmation of the idea that bumblebees' response to climate change is scaling with the severity of its observable impact on the Earth's climates.

Data and Methodology

The data were collected by the Global Biodiversity Information Facility, *Bumblebees of North America*, and Status and Trends of European Pollinators Collaborative Project and assembled by Kerr et. al (2015). Each observation measured the species of bee, the continent, the period in which the observation took place, distance from the equator, elevation, and minimum and maximum average annual temperature. I will be using only the data for European species because the data for North American bees has sample sizes for each of the periods that are orders of magnitude different.

The periods that the data were collected for were 1901-1974, 1975-1986, 1987-1998, and 1999-2010. I plan on comparing the distributions for the quantitative variables in the dataset across these time periods and across species to see if they differ significantly using Student's t tests for differences in means and Wilcox tests for differences in medians. Graphical evidence suggests that there are differences, but these statistical tests will more concretely confirm or reject that suggestion.

I will use the distributions of the four variables that I'll be analyzing to simulate data and perform statistical tests on that simulated data as a proof of concept before using the same methods on the actual data. I will compare the distributions of the four variables across time periods, which I have named "Historic", "Early", "Middle", and "Late" for the sake of brevity, to see if there are significant differences in the variable means across those time periods. I will then take a bootstrap sample from the differences between time periods and use a one-sided t-test and

Wilcox test to see if the difference between the later time periods is greater than the difference between the earlier time periods across all four variables.

I will be performing these calculations for all European species of bumblebee represented in the dataset, or 36 species. With an alpha level of 0.05 and 36 hypothesis tests to perform, there is an elevated risk of type I errors. This is why I'll be using a Bonferroni correction to adjust the p-values returned by these tests before comparing them to my desired alpha level.

I will then present my findings in a tabular format, making note of what proportion of species were found to have statistically significant changes in these four variables across each of the time periods, as well as the proportion of species for which the change was greater for the later time periods than for the earlier ones.

Simulations

I will be using the distributions of the variable corresponding to the distance from the equator not separating for species to test the methods before separating for species in my analysis. I took the distribution from each of the time periods and generated a normal distribution with the same mean and standard deviation. Figure 1 shows density plots of the actual distribution of this variable across time periods versus a density plot of the simulated distribution.

We can see that the original distributions are certainly not normal, but are somewhat symmetric about the mode, or the point with the highest density. There are smaller peaks on either side of the mode in various time periods, most likely corresponding to the modes of individual species of bumblebee. We can see from the x-axis that the simulated distributions are centered in different places across time periods, and it seems that the center of the distribution corresponding to the Late period is much different from the other three time periods.

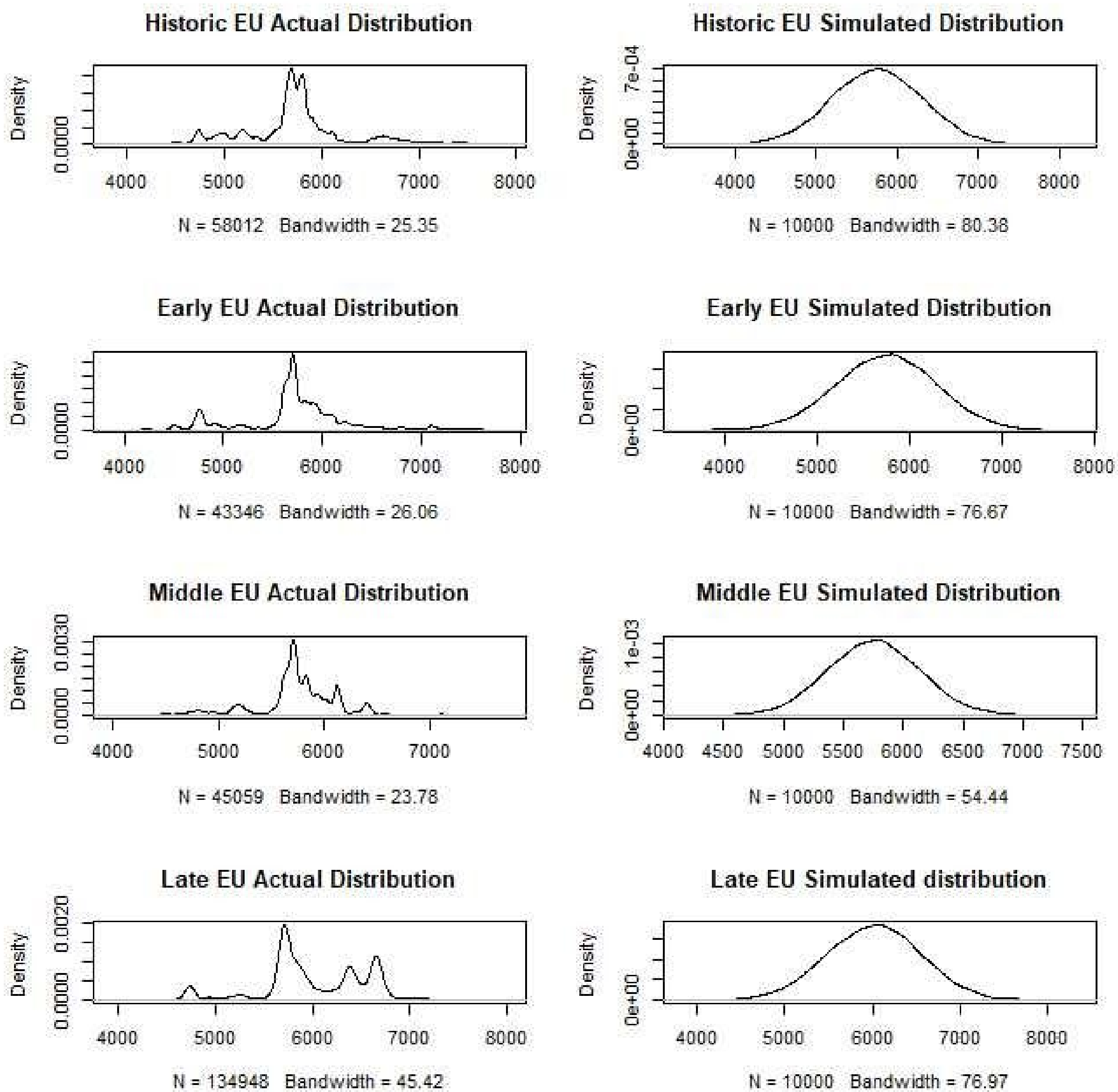

Figure 1: From top to bottom, the actual versus simulated distributions of the Distance from the Equator variable across the four time periods: Historic, Early, Middle, and Late.

The next thing I did was take the difference between the following pairs of simulated distributions: Late and Middle, Middle and Early, and Early and Historic. I plotted the density of these differences and compared them to a similar distribution with a mean of 0 and the same standard deviation. The density curves for the distributions centered at 0 are colored in red.

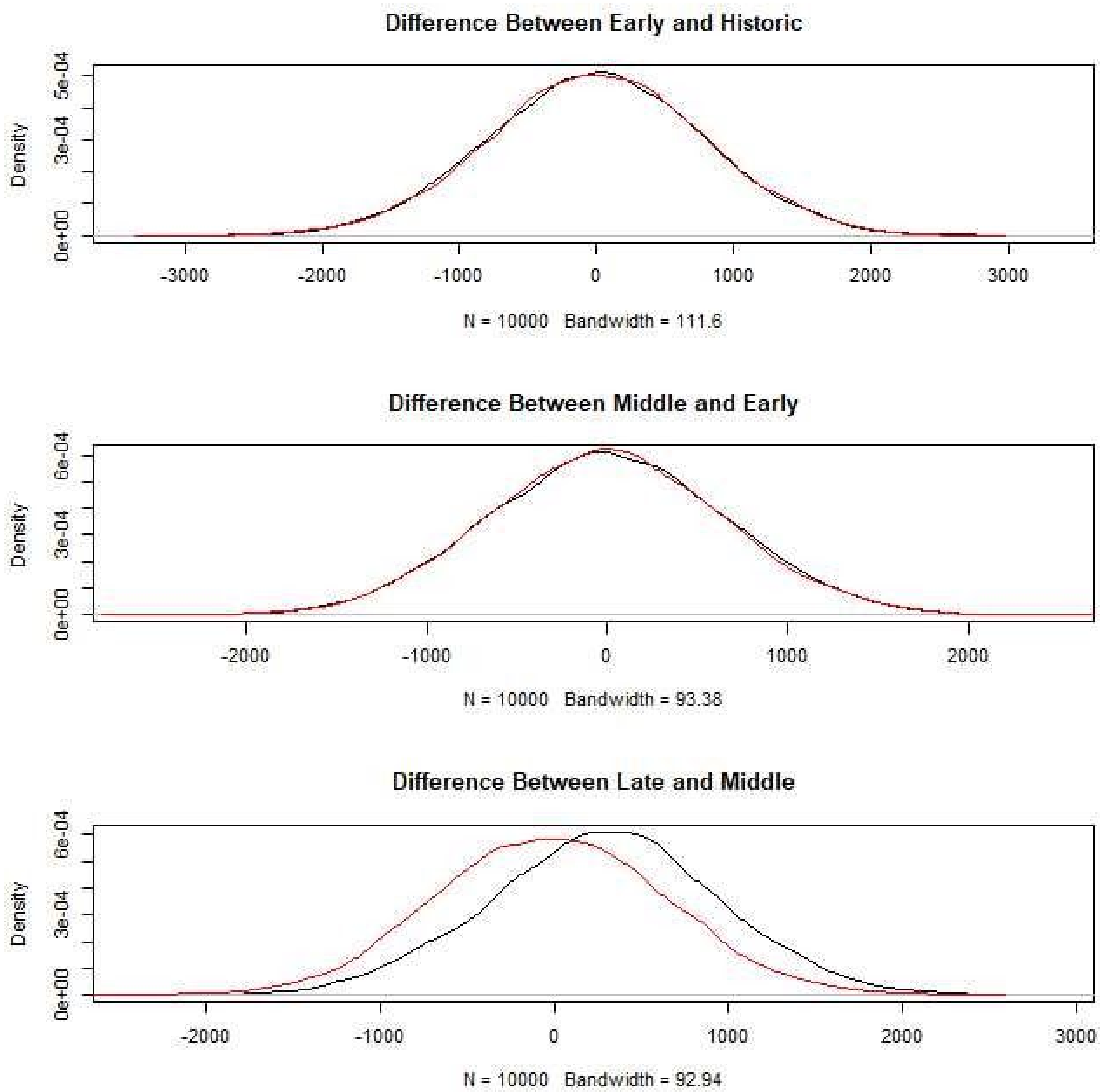

Figure 2: Density plots of the difference in Distance from the Equator between pairs of time periods (black) compared to distributions that assume no difference between the pairs of time periods (red).

The density plots in Figure 2 suggest that the difference in Distance from the Equator between the Middle and Late periods is much greater than the difference between any other pair of time periods, as there is no perceptible difference between the black and red lines in either the Historic/Early plot or the Early/Middle plot.

In order to better quantify the observations made for the plots in Figure 1 and Figure 2, I conducted Kolmogorov-Smirnov tests for goodness of fit on the original distributions, since we don't know a parameterization for the distributions. The alternative hypotheses being that the data for the pairs of time periods come from different distributions. Here are the results; R output approximated some of them to 0.

Time Period Pair	**Actual Distribution p-val**	**Sim Distribution p-val**
Historic/Early	≅ 0	0.0482
Early/Middle	≅ 0	≅ 0
Middle/Late	≅ 0	≅ 0

Figure 3: The results of Kolmogorov-Smirnov tests for pairs of time periods on the Distance from Equator variable.

Finally, I conducted 1000 bootstrap simulations to estimate the difference in the magnitude of the difference between pairs of time periods and performed one-sided t-tests for differences in means and Wilcox tests for differences in medians with the alternative hypothesis that the difference between the later pairs of distributions was greater than the difference between the earlier pairs. R again approximated some of the p-values to 0 in the output.

Time Period Pair Diff	**T-test p-val**	**Wilcox p-val**
Late-Middle vs Middle-Early	≅ 0	≅ 0
Middle-Early vs Early-Hist.	≅ 0	≅ 0

Figure 4: The results of t-tests and Wilcox tests on the difference between differences between time periods.

Having shown the results of these tests on distributions that were engineered to be different, I will move on to presenting results and analysis for these tests on the real set of data and discussing what they mean in the context of my research question.

Results and Analysis

I conducted tests with the real data across all four variables and time periods to see what proportion of species had statistically significant differences according to Kolmogorov-Smirnov tests with Bonferroni corrections. Here are my results:

Time Periods	Elevation	Distance	Min Temp	Max Temp
Historic/Early	0.9167	1.0000	0.7778	0.7500
Early/Middle	0.6111	0.9167	0.6944	0.6389
Middle/Late	0.9722	1.0000	0.9722	1.0000

Figure 5: Proportions of species determined to have statistically significant differences in given between given time periods according to Kolmogorov-Smirnov tests with Bonferroni corrections.

We can see that a majority of species have statistically significant differences for every variable across all time periods, but the Middle/Late differences have consistently higher proportions of species with significant differences in variables than the other two pairs of time periods. This is especially pronounced in the two variables relating to species' temperature extremes. This suggests that the answer to our first question, whether or not bumblebees are changing environments with changes in climate, is yes.

To answer the second question, whether or not these changes are increasingly drastic with the severity of climate change, I performed the same bootstrap process from the Simulations section on all four variables and measured the proportion of species that had statistically significant differences in the difference between pairs of time periods.

Distance from the Equator:

Time Periods	**Prop. of significant results with a t-test**	**Proportion of significant results with a Wilcox test**
Middle-Early vs Early-Hist.	0.3889	0.4167
Late-Middle vs Middle-Early	0.7500	0.7500

Figure 6: The proportion of species found to have increased change in their Distance from the Equator over time.

The proportion of species that have had significant changes in their distance from the equator seems to have increased over time according to this table.

Elevation:

Time Periods	**Prop. of significant results with a t-test**	**Proportion of significant results with a Wilcox test**
Middle-Early vs Early-Hist.	0.4722	0.4444
Late-Middle vs Middle-Early	0.6111	0.6111

Figure 7: The proportion of species found to have increased change in their Elevation over time.

The results are similar to those for the change in distance from the equator, though not quite as drastic, with the earlier time periods seeing a slightly greater change and the later ones seeing a slightly lesser change.

Minimum Average Annual Temperature:

Time Periods	**Prop. of significant results with a t-test**	**Proportion of significant results with a Wilcox test**
Middle-Early vs Early-Hist.	0.4722	0.4722
Late-Middle vs Middle-Early	0.7222	0.7222

Figure 8: The proportion of species found to have increased change in their Minimum Average Annual Temperature over time.

These results are similar to the first two with the proportion of species seeing significant changes being higher in later time periods than in earlier ones.

Maximum Average Annual Temperature:

Time Periods	**Prop. of significant results with a t-test**	**Proportion of significant results with a Wilcox test**
Middle-Early vs Early-Hist.	0.6111	0.6111
Late-Middle vs Middle-Early	0.3611	0.3611

Figure 9: The proportion of species found to have increased change in their Maximum Average Annual Temperature over time.

These results are the most surprising. One would expect that the maximum temperature to which species are exposed to increase on average as climate change becomes more severe, but the proportion of species with significant changes to this condition are significantly lower in the later time periods than in the earlier ones.

Discussion

In this paper, I have presented methods for answering the question of whether bumblebees are reacting to climate change and whether the severity of their reaction is scaling with the severity of climate change. I demonstrated methods that would identify statistically significant differences in distributions across variables and significant changes in how fast those variables change across time periods.

I demonstrated that a majority of European bumblebee species, in many cases an overwhelming majority, are changing environments in statistically significant ways. I also demonstrated that in every variable but one, those changes are more significant in later time periods than in earlier ones. The exception being the maximum average annual temperature that bumblebee species are observed in.

I took precautions against the possibility of type I errors by using the Bonferroni correction and addressed the problem of these data being difficult to fit to a common distribution by using the non-parametric Kolmogorov-Smirnov test for goodness of fit. I also took care to use the appropriate one-sided tests to come to my conclusions.

One drawback of my methods is the fact that the lack of a common parameterization for the distributions I was working with made it difficult to perform simulations using Monte Carlo techniques, limiting the sophistication and size of the simulation data that I was able to work with. This could be a useful method to explore in future study of this topic. Another drawback is the fact that I only measured proportions of species that met a given significance threshold in my tests rather than measuring the relative significance of the test on the species. Finally, these results are exclusive to European bumblebees, which could bias the results to European climate.

Works Cited

Kerr, J. T., et al. “Climate Change Impacts on Bumblebees Converge across Continents.” *Science*, vol. 349, no. 6244, 2015, pp. 177–180., doi:10.1126/science.aaa7031.

McRae, Mike. “Bumblebees Are in Trouble, as Rising Temperatures Drive Their Decline.” *ScienceAlert*

Frank J. Massey Jr. (1951) The Kolmogorov-Smirnov Test for Goodness of Fit, Journal of the American Statistical Association, 46:253, 68-78, DOI: 10.1080/01621459.1951.10500769

Bland JM, Altman DG: Multiple significance tests: The Bonferroni method. BMJ 1995;310(6973):170.