Tiffany Wu
UEP 254

Quant Final Project

Introduction

In the coastal plains of the Arctic, climate change impacts on ecosystems are increasingly visible, particularly on wildlife populations. The Greater White-Fronted Goose (Anser albifrons) is a circumpolar species of goose that spends its spring breeding season in tundras across North America, Greenland and Russia. Geese are an important type of waterfowl for the U.S. Geological Survey and other federal agencies to understand because of the role they play in subsistence in Alaska as well as sport hunting, which contributes to a significant portion of revenue. They are also bellwethers for shifting temperatures, and any changes in their population size are useful to know. Because the white-fronted goose migrates to wetlands along the Gulf Coast during the winter, their distribution is broad and there are various factors that go into their survival and overall productivity, some of which are little understood.

Ongoing research conducted by the USGS Alaska Science Center has focused on these migration patterns, nesting biology, nutrition, and differences in population. Alaska has been warming more than twice as fast as the rest of the United States (EDF, 2014). Wildlife habitats in the far north have been particularly impacted by a warming planet, including through receding glaciers, thawing permafrost, and the disappearance of sea ice. They can have a widespread effect on species ranging from polar bears to eelgrass.

During the spring and summer molting season, geese cannot fly until their new feathers come in – this is also when they are rearing their young, so being largely confined to the ground is not an issue. Biologists frequently utilize this time frame to round up geese to gather data about their characteristics. According to the USGS, the changing ecosystems on the Arctic Coastal Plain, which stretches from the northernmost parts of Alaska across the Yukon and the Northwest Territories, have had a predictably negative impact on many species of avifauna. However, all four goose species found in the region have actually increased in point counts, with the greater white-fronted goose having the fastest observed population growth rates.

Using one such study, I am hoping to determine whether certain trends in these greater white-fronted geese in Alaska, such as weight and wing molt, can be observed among individuals in the study. I am also interested in gauging whether there are patterns in this three year time span regarding their overall health to find any possible explanations for their thriving in unexpected circumstances.

Methodology

The research design of this study is quasi-experimental because it intends to establish causality, namely the year observations were taken and the corresponding weights of adult and juvenile geese, but is not truly randomized. For example, some geese that were corralled and captured may have been friendlier than average because they were comfortable in their environment and were less wary of humans – meaning they could have been healthier or heavier than normal – or the opposite could have been true, with geese that were weighed for the study being slightly weaker and / or lighter and not fast enough to avoid being coaxed into the temporary net area. By using weight as a proxy for health, I am also making some assumptions about the condition of the flock and understand that there are other factors that go into the equation, such as adequate nutrition and being free from injury or disease.

I am utilizing data from the USGS Alaska Science Center collected between 2012 and 2014 for the purposes of studying gosling growth rates and the weight of adult greater white-fronted geese (Anser albifrons). The study region was located in the Arctic Coastal Plain, and the specific study area was between Admiralty Bay and the eastern edge of Teshekpuk Lake. For reference, a common method of capture for these geese during the summer molting season on land is to set up a corral with nets attached to poles and herd flocks into these areas for them to be individually weighed and measured. In the water, canoes or kayaks are used to coax them onto land for the same process (MassGov, 2021).

Researchers in this USGS study gathered data from 917 greater white-fronted geese in this three-year period to better understand their population dynamics in the context of climate change. The original data included 15 variables, but I removed 7 of them (webtag, brood, nest_id, hatch_date, age_days, capture_date, and bird_id) as many contained missing values or were not filled out at all. All the birds were studied during the first week of August each year (USGS, 2021) so I dropped capture_date. The 8 variables I did keep were year, site, age, sex, culmen, tarsus, ninth_primary and mass. Culmen refers to the bill length, tarsus refers to the leg length, and ninth_primary refers to the length of each individual's ninth primary feather, which indicates new growth.

The fact that this data collection took place over only a 2-3 year period may limit the reliability of this study as it is a relatively short amount of time. An anomalous year in breeding or nesting habits could have influenced the results. Additionally, given that the study took place in a region where the climate is rapidly changing, it could have had a greater than expected influence on the participants (in this case, the geese) with regards to which locations they chose to migrate to and whether they would have shown up in this study at all, which would diminish internal validity. Since the geese data was specifically from ones that were caught, it is difficult to say whether the data was fully randomly sampled. However, the process by which individuals ended up being corralled would have been random given their large flock sizes, which can be close to 1000 (Gupte, et al. 2019). This research can be applied to other species or situations, so the external validity is high. The study is also capable of being replicated and hopefully will be extended further at some point.

My plan for data analysis is to use a hypothesis test to see whether gosling weights and adult weights changed from 2012 to 2014. If observational studies bear that greater white-fronted geese appear to be breeding in greater numbers, then I would imagine that they are in good health and have been able to receive enough nutrition, resulting in increasing body mass per year. This would be the case for both goslings and adults. I also believe that other physical characteristics, such as the length of their ninth primary flight feathers, may be stronger to indicate greater population health. In the data, age is given as AHY ("After Hatching Year") or L ("Local"). The former indicates the individual is an adult and the latter means the individual was recently hatched and is considered a nestling or young bird.

Because the sex of each individual (M, F, Unknown) is given as a string, I first converted it to an integer using the encode function under the new variable name "gender". I then dropped the original column.

```
. encode sex, generate(gender)

. drop sex
```

Using the Stata codebook command, I checked to see my results had been converted to integers as intended.

```
-------------------------------------------------------------------------------
gender                                                                      Sex
-------------------------------------------------------------------------------

                  Type: Numeric (long)
                 Label: gender

                 Range: [1,3]                    Units: 1
         Unique values: 3                    Missing .: 0/921

           Tabulation: Freq.   Numeric  Label
                         513         1  F
                         406         2  M
                           2         3  U

. drop if gender == 3
(2 observations deleted)
```

There were two individuals of unknown gender that I dropped from the data, leaving 919 observations.

I repeated the process with age (AHY, L) and also dropped the original column. Note that AHY (adults) were labeled with the numeric 1 and L (juveniles) were labeled with the numeric 2.

```
. encode age, generate (age2)
```

```
age2                                                                    Age

              Type: Numeric (long)
             Label: age2

             Range: [1,2]                    Units: 1
     Unique values: 2                    Missing .: 0/921

        Tabulation: Freq.   Numeric  Label
                      322         1  AHY
                      599         2  L
```

With these values converted into numerics, I was able to proceed with running a one-way ANOVA on the data with a significance level of 5%. I first wanted to see the mean weights of the geese in 2012, 2013, and 2014. The next thing I wanted to do was to analyze the mean weights of adults and juveniles in this time frame. The null hypothesis is that there is no statistically significant difference between the mean weight of greater white-fronted geese during the three year study period. The alternative hypothesis is that there is, which could potentially reflect their increasing robustness. This data was independent and there was no pairing that occurred.

H0: $\mu_1 = \mu_2 = \mu_3$
HA: at least one mean is different than the others

Results

Adult Geese

I tabulated the data for mass and year for adults and juveniles, starting with adult geese.

```
. oneway mass year if age == 1, tabulate
```

Year	Summary of Mass Mean	Std. dev.	Freq.
2012	2202.6124	248.33751	129
2013	2284.8046	205.71875	87
2014	2297.2453	254.01627	106
Total	2255.972	242.83328	322

At a glance, it appears that the mean weight of adult greater white-fronted geese is relatively similar in each observation year. But does a closer look warrant the same interpretation?

```
. anova mass year if age == 1
```

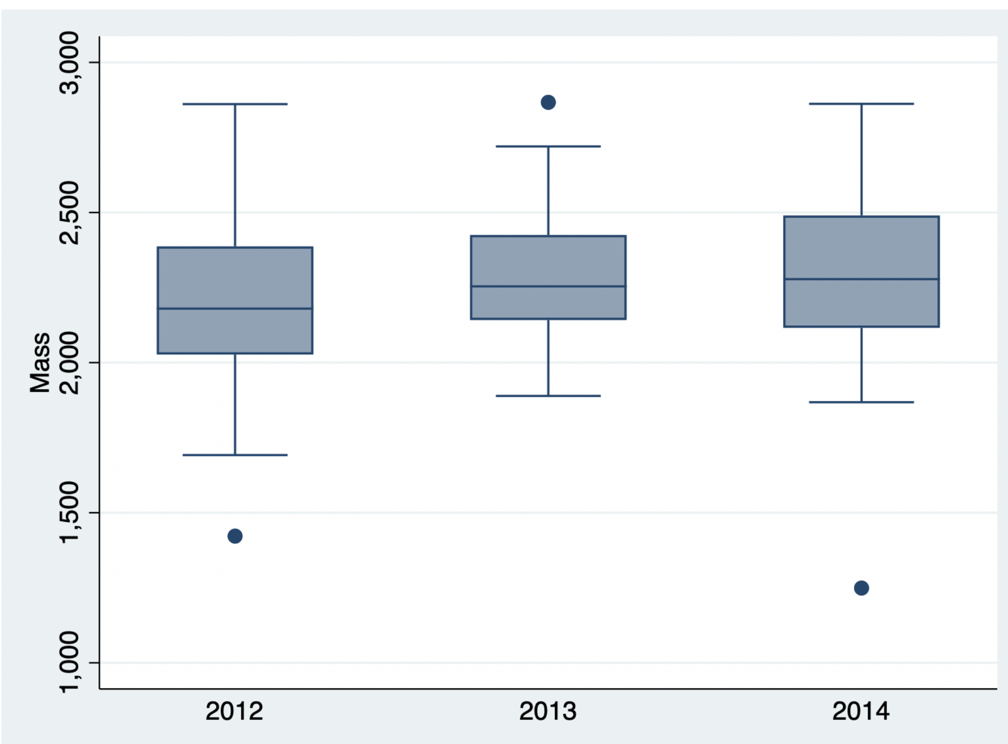
	Number of obs =	322	R-squared	=	0.0328	
	Root MSE =	239.569	Adj R-squared =	0.0267		

Source	Partial SS	df	MS	F	Prob>F
Model	620188.83	2	310094.41	5.40	0.0049
year	620188.83	2	310094.41	5.40	0.0049
Residual	18308540	319	57393.542		
Total	18928729	321	58968.002		

If the p-value is > 0.05, the null hypothesis is true. Here, Prob>F, or the p-value for the entire model test is 0.005, so we will reject the null hypothesis. The difference in mean weight over year for adult greater white-fronted geese is statistically significant between at least two of the years.

To determine what this data looked like more visually, I made a boxplot.

```
. graph box mass if age == 1, over(year)
```



It appears the median weight for adult greater white-fronted geese was highest in 2013, and the whiskers (1.5 * IQR of the first and third quartiles) were shorter. The outlier was also a value that was higher, compared to 2012 and 2014, where the outlier was lower than the rest of the dataset. The interquartile ranges of 2012 and 2014 appeared to be similar, but the bottom whisker in 2012 was noticeably lower than in 2014. The tops of the whiskers appeared to be very similar and was around the same as the outlying 2013 value.

Juvenile Geese

I then tabulated the data for mean weight and year for juveniles.

```
. oneway mass year if age == 2, tabulate
```

	Summary of Mass		
Year	Mean	Std. dev.	Freq.
2012	1471.8333	360.61341	18
2013	1320.2825	234.09917	177
2014	1514.4801	242.56529	402
Total	1455.6181	259.35318	597

It appears there is more variance in this category over the observation years, but with only 18 observations of juveniles in the first year, the results may be influenced due to the lower sample size.

```
. anova mass year if age == 2
```

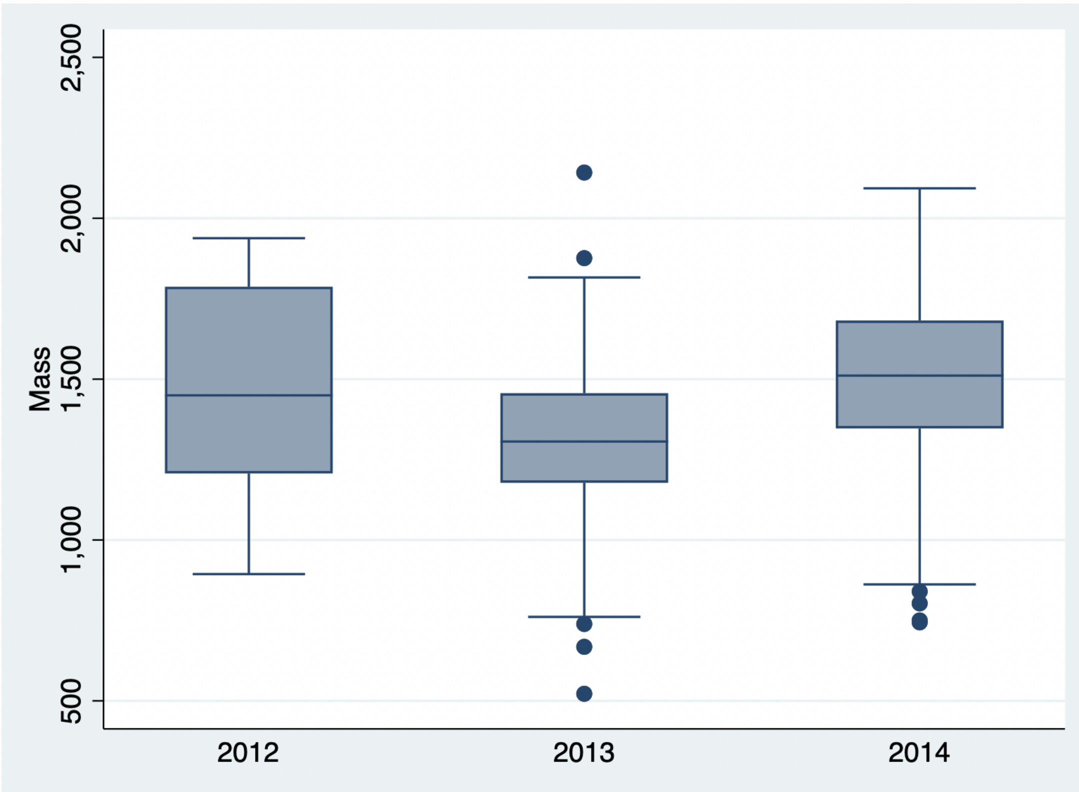
Number of obs = 597 R-squared = 0.1157
Root MSE = 244.295 Adj R-squared = 0.1128

Source	Partial SS	df	MS	F	Prob>F
Model	4639440.2	2	2319720.1	38.87	0.0000
year	4639440.2	2	2319720.1	38.87	0.0000
Residual	35449947	594	59680.045		
Total	40089387	596	67264.072		

With Prob>F less than 0.05, this indicates we should reject the null hypothesis. The difference in mean weight over year for juvenile greater white-fronted geese is also statistically significant between at least two of the years.

Below is a boxplot for the weights of juveniles over the three years.

```
. graph box mass if age == 2, over(year)
```



It appears there was quite a bit of variation in weights of juvenile greater-fronted white geese. The interquartile range for 2013 was relatively narrow and indicated the lowest median weight out of the three years. The top and bottom whiskers (1.5 * IQR of the first and third quartiles) were quite long, and there were both high and low outliers. In 2014, the whiskers on both ends were similarly lengthy, indicating the collected data for the latter two years covered a greater range. Since there were less than 20 juveniles weighed for 2012, this may have contributed to the results.

Conclusion

The primary findings of this analysis indicate that the change in mean weights of adult greater white-fronted geese were statistically significant from 2012 to 2014 and increased with each year. (The change in mean weights of juvenile greater white-fronted geese was also statistically significant during this time period). Assuming that the weights of adults indicate their overall health and ability to successfully rear young, this would suggest that the adult goose population is doing well, confirming the USGS Alaska Science Center's observation that this species of waterfowl is growing in population despite the impacts of a changing climate. This could mean that they are outcompeting year-round avian species in the Arctic Coastal Plain that do not migrate, or that slightly warmer global temperatures are better for their breeding outcomes – although harmful for the planet.

Regarding the significance in the weights of juvenile geese between 2012-2014, young geese grow very rapidly and even if they all hatched around the same time, their sizes could have varied significantly in the early months. The original dataset, which is provided with the project, indicates that they were studied between August 5th and August 9th of each year. This is a narrow and consistent timeframe, which is optimal for the study. There would need to be more information about how the goslings were selected in order to draw a conclusion about their overall health each year using their weight, and indeed it may not be necessary, given they reach maturity by the time they are a year old. It may be more helpful to see this as a snapshot of how much juveniles have weighed in late summer and focus on results for the adults.

Some additional information that would be helpful to know include exactly how the geese were captured since only the dataset was provided. I included the most common method of capturing geese for study purposes in the continental United States, but realize that it may be different in Alaska, where geography may require a different approach to handling them. Having this information would allow a stronger conclusion to be drawn about the randomness of the sample and increase the robustness of the study. Other pieces of information that would be useful include documentation of interactions with other types of arctic geese due to competition and whether the availability of food in a certain year can explain their changes in weight. Most importantly, there should be additional factors beyond weight that is studied in the context of an individual’s health and ability to reproduce (or some confirmation that body mass is indeed the most important contributor to these outcomes). A future study should incorporate the bill length, leg length, and primary feather growth of each individual to gauge whether or how much these physical indicators play a role.

Being able to analyze spatial characteristics such as the topography of the Arctic Coastal Plain would be useful to determine the accuracy of the sample and whether it is truly representative of the tens of thousands of greater white-fronted geese that migrate to the region each summer. Understanding their historical population patterns – including whether there is a natural ebb and flow or if these resilient geese are nearing carrying capacity – would be useful in determining patterns that may occur in the future.

Works Cited

Environmental Defense Fund. (2014). “How will climate change affect Alaska?” https://www.edf.org/sites/default/files/content/regional_releases_alaska_web_version.pdf

Government of Massachusetts. (2021). “When geese can’t fly.” https://www.mass.gov/news/when-geese-cant-fly

Gupte, P. et al. (2018). “Family size dynamics in wintering geese.” *Journal of Ornithology.* https://link.springer.com/article/10.1007/s10336-018-1613-5

Larsson, K. (1996). “Genetic and environmental effects on the timing of wing moult in the barnacle goose”. *Heredity.* https://www.nature.com/articles/hdy199613.pdf?origin=ppub

Meixell, B.W. and P. Flint. (2021). “Capture and Measurement Data of Greater White-Fronted Geese (Anser albifrons) from the Arctic Coastal Plain of Alaska, 2012-2014”. *U.S. Geological Survey, Alaska.* https://www.usgs.gov/data/capture-and-measurement-data-greater-white-fronted-geese-anser-albifrons-arctic-coastal-plain.

U.S. Geological Survey. n.d. “Age (year class) codes.” https://www.pwrc.usgs.gov/bbl/manual/age.cfm

U.S. Geological Survey. (2017). “Goose research.” https://www.usgs.gov/centers/alaska-science-center/science/goose-research

U.S. Geological Survey. (2016). “USGS Changing Arctic Ecosystems.” https://alaska.usgs.gov/portal/project.php?project_id=287

Warren, K. (2006). Identification Field Guide to the Geese of the Willamette Valley. *Oregon Department of Fish and Wildlife.* https://www.dfw.state.or.us/resources/hunting/waterfowl/goose-permits/docs/Goosefieldguide_2ndEdition_final.pdf