

# Climatic Trend Impact on Alaskan Stream Discharge

[https://github.com/wgrimshaw/Alaska\\_DBPs.git](https://github.com/wgrimshaw/Alaska_DBPs.git)

*Gaby Garcia, Walker Grimshaw, Tristen Townsend, Yixin Wen*

<General Guidelines> <1. Write in scientific style> <2. Global options for R chunks should be set so that only relevant output is displayed> <3. Make sure your final knitted PDF looks professional. Format tables appropriately, size figures appropriately, make sure bulleted and numbered lists appear as such, avoid awkwardly placed page breaks, etc.>

# 1 Rationale and Research Questions

<Write 1-2 paragraph(s) detailing the rationale for your study. This should include both the context of the topic as well as a rationale for your choice of dataset (reason for location, variables, etc.) A few citations should be included to give context for your topic. You may choose to configure autoreferencing for your citations or add these manually.>

## 1.1 Background

The climate is changing, in large part due to anthropogenic carbon emissions. These changes have different magnitudes around the world and local impacts of climate change vary as well. Specifically, climate change is already having greater impacts near the poles than many other parts of the globe, a process known as polar amplification or arctic amplification (Serreze et al 2009). Understanding how climate change is affecting discharge, especially in Alaska, has implications for water management, ecological processes, and the larger global system (if we consider ice-albedo feedback). Many communities rely on a given amount of water from snowmelt to arrive at certain times of the year, so a shift in the quantity or timing could drastically affect downstream users and water managers. Additionally, changing the amount of flow in rivers could affect sensitive biological communities. Furthermore, changes in temperature that result in glacial or permafrost melting could reduce the amount of reflective land cover, thus disrupting larger climate systems.

<At the end of your rationale, introduce a numbered list of your questions (or an overarching question and sub-questions). Each question should be accompanied by one or more working hypotheses, inserted beneath each question.>

## 1.2 Research Question

To what degree does climate change affect discharge in Alaskan streams and rivers?

This guiding question encompasses other questions we seek to answer through statistical analyses. If climate change is causing average temperatures to increase, and the magnitude of the increase is greater at higher latitudes, then higher latitude streams should demonstrate greater changes in timing and magnitude of peak flows. This study first seeks to examine if the magnitude of temperature changes increase with increasing latitude in Alaska. By analyzing historical streamflow records, we will investigate whether the magnitude of maximum daily temperature change causes a proportional change in the magnitude and timing of peak streamflow.

Another aspect of climate change's impacts on streamflow is the potential of melting permafrost or glaciers. This analysis uses cumulative annual streamflow and cumulative annual

precipitation to determine if interannual snowpack is melting with increasing average temperatures. If so, we expect the difference between annual precipitation and annual streamflow to increase over time.

## 2 Dataset Information

<Provide information on how the dataset for this analysis were collected, the data contained in the dataset, and any important pieces of information that are relevant to your analyses. This section should contain much of same information as the metadata file for the dataset but formatted in a way that is more narrative.>

### 2.1 Discharge

Discharge data were collected from NWIS using the Data Retrieval package. The state of Alaska was divided into 10 bins of equal latitude, and daily discharge data was downloaded for the site in each latitude bin with the greatest number of samples. This dataset includes the site location, daily discharge, and county of the site, among other variables not used in the analysis.

### 2.2 Temperature and Precipitation

Temperature and precipitation data were downloaded from the National Oceanic and Atmospheric Administration (NOAA) Climate Data Online web portal. As discharge stations do not collect data on temperature and precipitation, the climate data for each latitude bin were downloaded from a station in the same county as each selected discharge station. In each county, daily precipitation, maximum temperature, and minimum temperature, data were downloaded from one station. The criteria for station selection include data extending to the current date, beginning at the earliest date, with at least 80% data coverage. This dataset also includes site location.

### 3 Data Wrangling

<Add a table that summarizes your data structure (variables, units, ranges and/or central tendencies, data source if multiple are used, etc.). This table can be made in markdown text or inserted as a `kable` function in an R chunk. If the latter, do not include the code used to generate your table.>

Variable	Units (if known)	Type of Variable	Hypothesis
Discharge	Cubic feet per second	Response	1a, 1b, 1c
Site Number	Latitude/Longitude	Predictor	1a
Date Time	Year/Month/Day	Predictor	1c
Date of First Snowmelt	Year/Month/Day	Predictor	1a, 1b, 1c
Air Temperature	Celsius	Predictor	1b
Precipitation	Millimeters	Predictor	1a, 1c
HUC 8 Watershed Size	Square Meters	Predictor	1a, 1b
Permafrost Melt	Qualitative	Predictor	1b
Glacial Coverage/Melting	Qualitative	Predictor	1b

### 4 Data Wrangling Methods

<Describe how your team wrangled your dataset in a format similar to a methods section of a journal article.>

#### 4.0.1 Importing, Cleaning, and Addressing Date Issues

The state of Alaska was divided into 10 bins of equal latitude. Precipitation and temperature data for 10 Alaskan NOAA stations—one inside each latitude bin—was obtained from NOAA’s Climate Data Online web portal, and discharge data was retrieved from NWIS’s Data Retrieval package using one site in each latitude bin with the greatest number of samples. All discharge data resided in one CSV entitled NWIS\_Discharge. All raw CSVs were imported into our Raw Data folder within our project repository on Github, and imported into R using the `read.csv` function.

Each of the 10 NOAA stations classified by latitude bin had a unique CSV that was cleaned prior to joining. Ancillary columns from each of the 10 dataframes were removed in order to ensure that each CSV had the same columns in the same order. In order to address date-time issues in the NOAA CSVs for Bin 1 and Bin 5, the `as.DATE` function was used to change the “DATE” column from a factor to a date, in the format month/day/year. Next, we changed the format of the “DATE” column to a two digit year, a two digit month, and the two digit day. A function was then written to create early dates that were misrepresented as years after 2019. Then, the DATE column was inserted into the function written above for each row of the DATE column. Finally, the `as.DATE` function was used to convert the DATE column into a 4 digit year, a two digit month, and 2 digit day format, our preferred format.

#### 4.0.2 Joining Data

After all 10 NOAA dataframes containing precipitation and temperature data were cleaned, they were all joined by row using the rbind function into a new CSV entitled “TempPrecip”. Next, the “site\_no” column in the NWIS\_Discharge dataframe was converted into a factor using the as.factor function. The revalue function within the “plyr” package was then used to create a new “Bin” column renaming the USGS Station Numbers in the “site\_no” column by their bin number. The “Date” column in the Discharge dataframe was renamed to “DATE” so that it matched the “DATE” column in the “TempPrecip” dataframe. Finally, the merge function was used to join the “TempPrecip” and “NWIS\_Discharge” dataframes by “DATE” and “Bin” into a new CSV entitled “AlaskaTempPrecipDischarge”, and cleaned for clarity.

## 5 Exploratory Analysis

<Insert exploratory visualizations of your dataset. This may include, but is not limited to, graphs illustrating the distributions of variables of interest and/or maps of the spatial context of your dataset. Format your R chunks so that graphs are displayed but code is not displayed. Accompany these graphs with text sections that describe the visualizations and provide context for further analyses.>

<Each figure should be accompanied by a caption, and each figure should be referenced within the text>

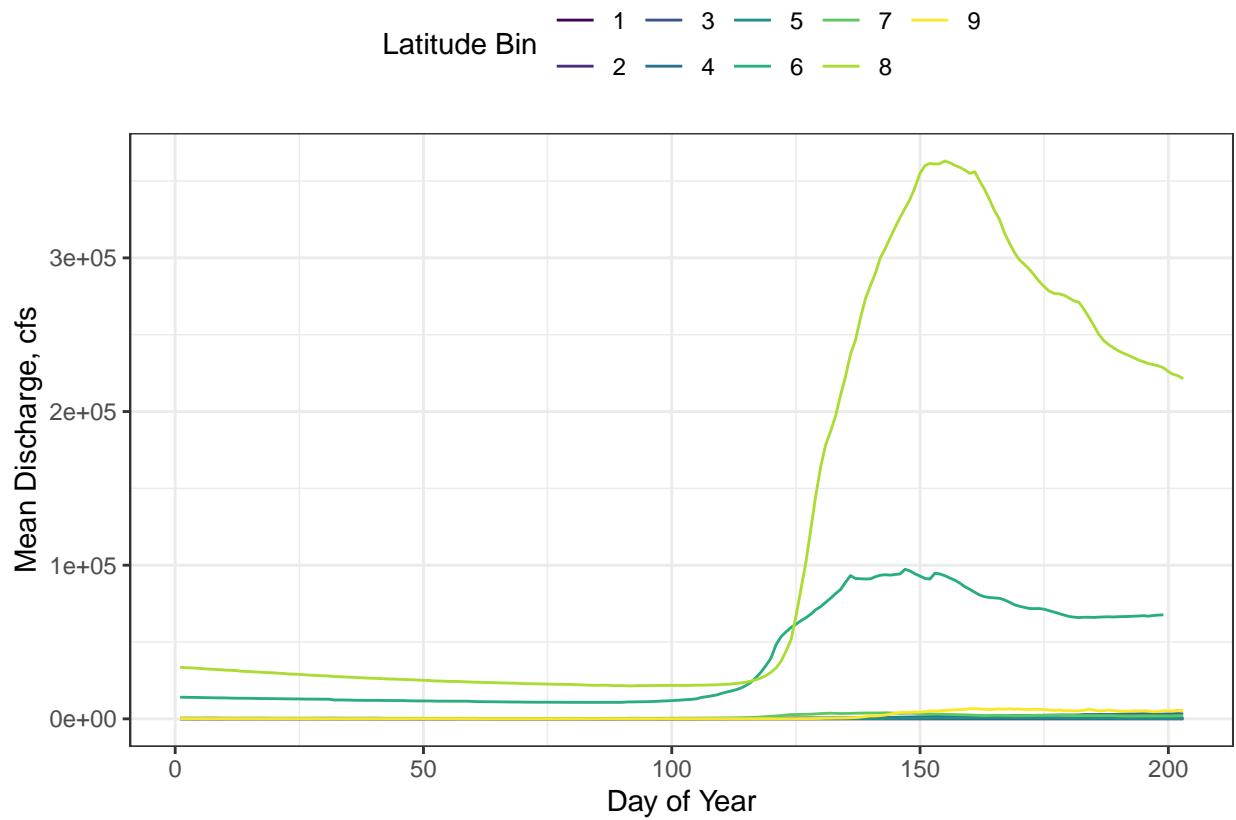


Figure 1: Day of Year vs. Mean Discharge. This figure shows mean discharge across all nine latitude bins for each day of the year. This graph served to illustrate variation across sites as to when first day of snowmelt and peak snowmelt would occur.

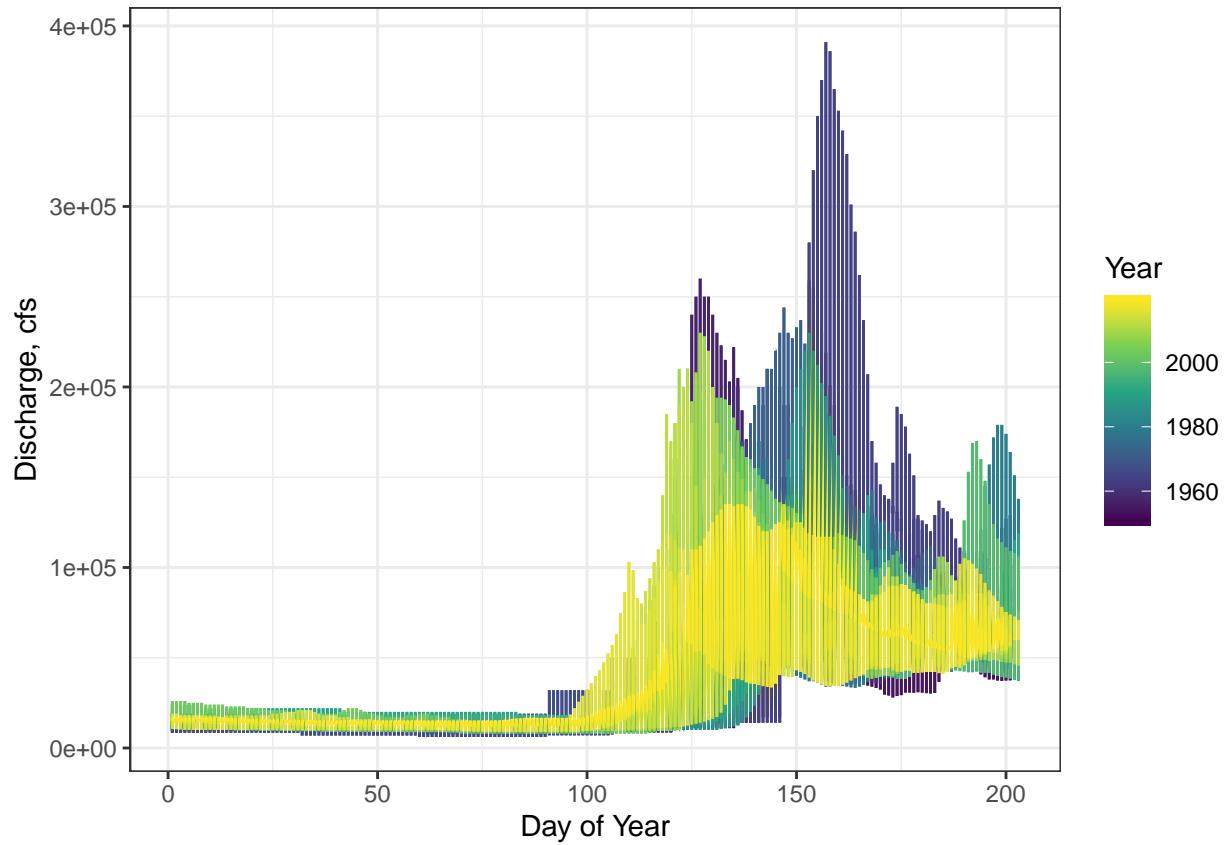


Figure 2: Day of Year vs. Discharge. This figure shows how discharge changes with day of year across the period of record for Bin 6. This provided a visual cue as to whether snowmelt was occurring earlier and during what time of the year the shift was occurring.

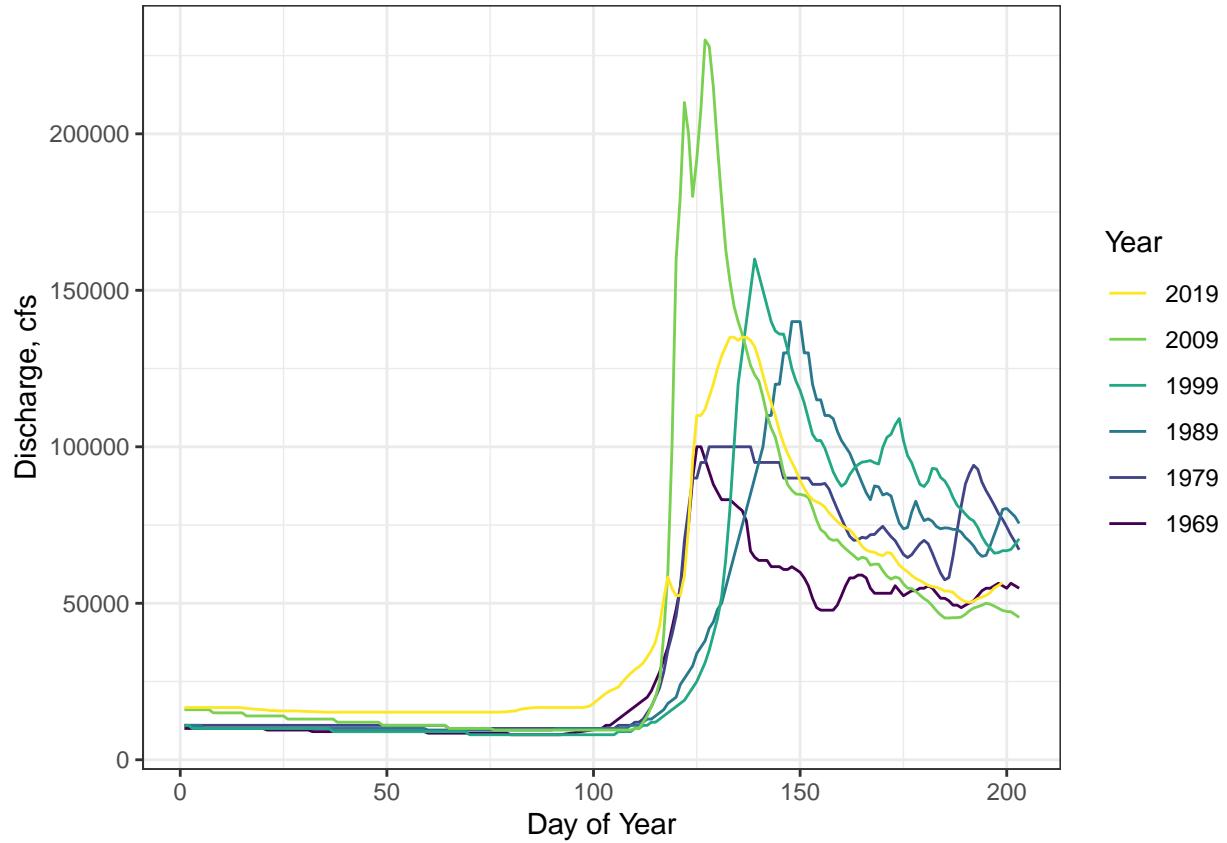


Figure 3: Day of Year vs Discharge. This figure shows how discharge changes with day of year across the every ten years across the period of record for fifty years. This provided a visual cue as to whether snowmelt was occurring earlier and during what time of the year the shift was occurring.

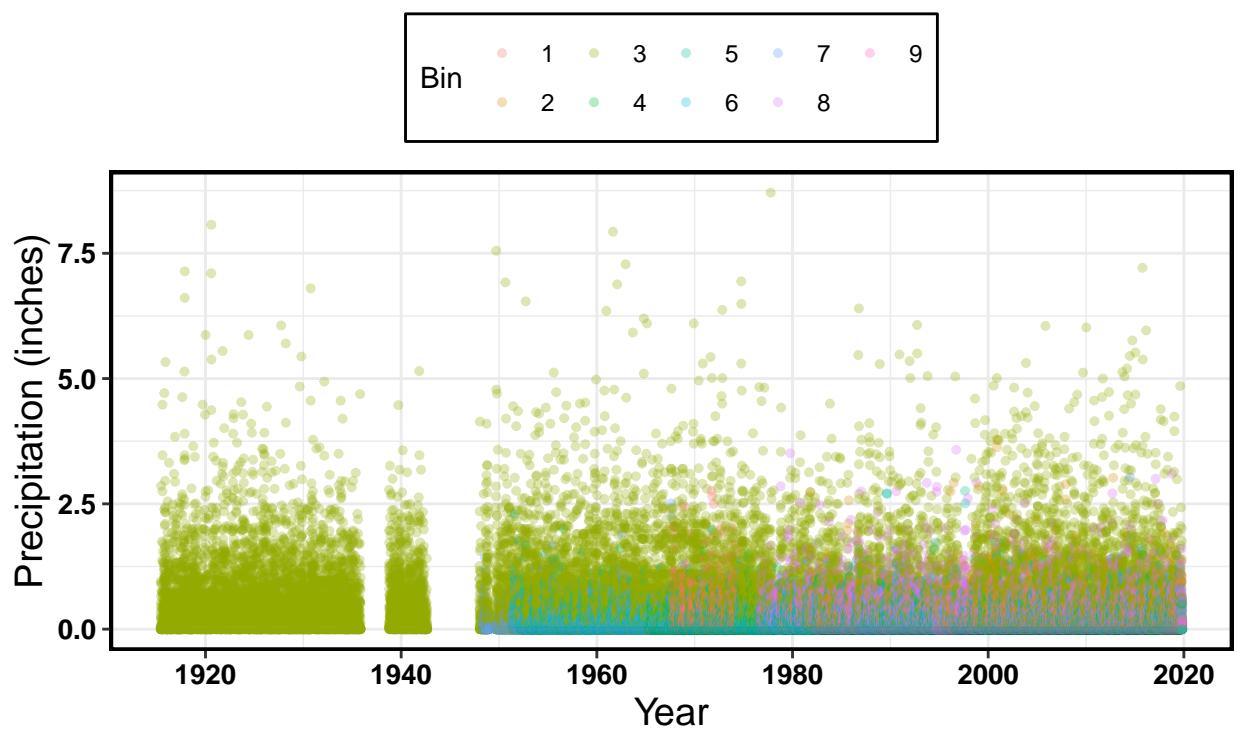


Figure 4: Precipitation Measured Over the Entire Period of Record by Bin. This figure shows the general pattern of precipitation changes over time, differentiated by the latitude bin (color). Bin 3 has the most sample points for the entire precipitation period of record. Bin 3's precipitation range has the greatest magnitude compared to the other bins.

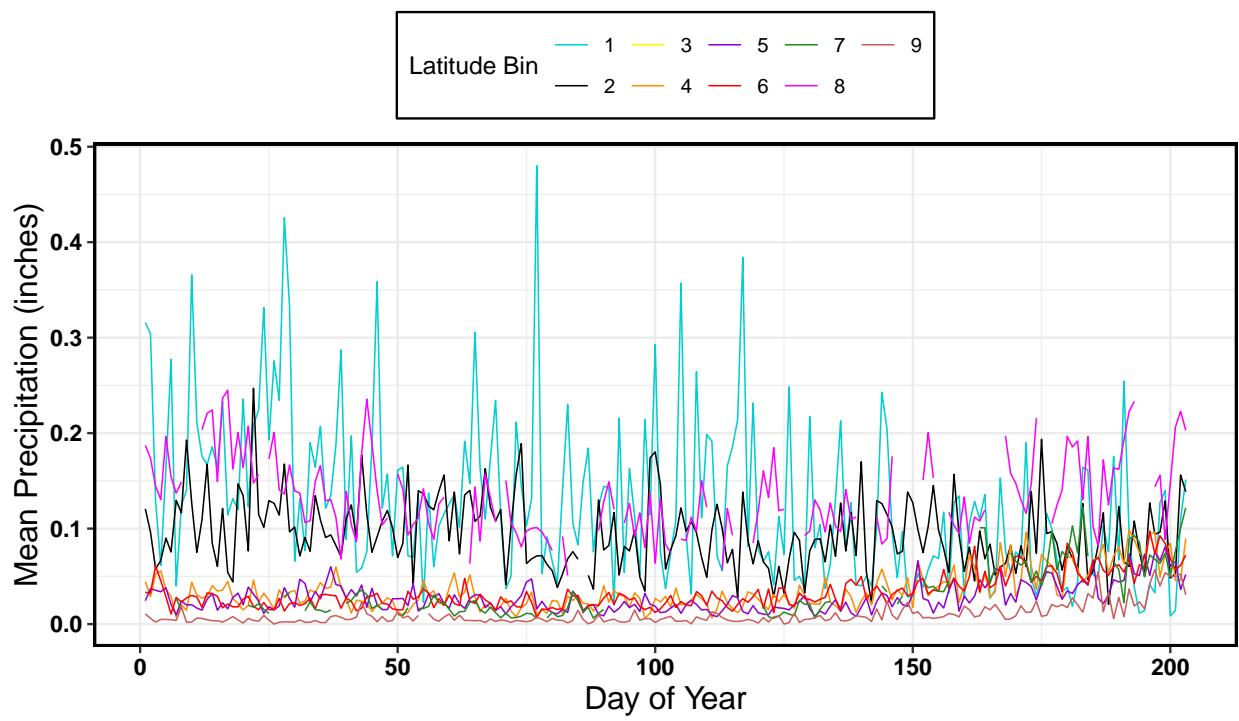


Figure 5: Day of Year vs. Mean Precipitation. This figure shows mean precipitation across all nine latitude bins for each day of the year. This graph served to illustrate precipitation variation across sites. Bin 1 has the greatest spread of mean precipitation.

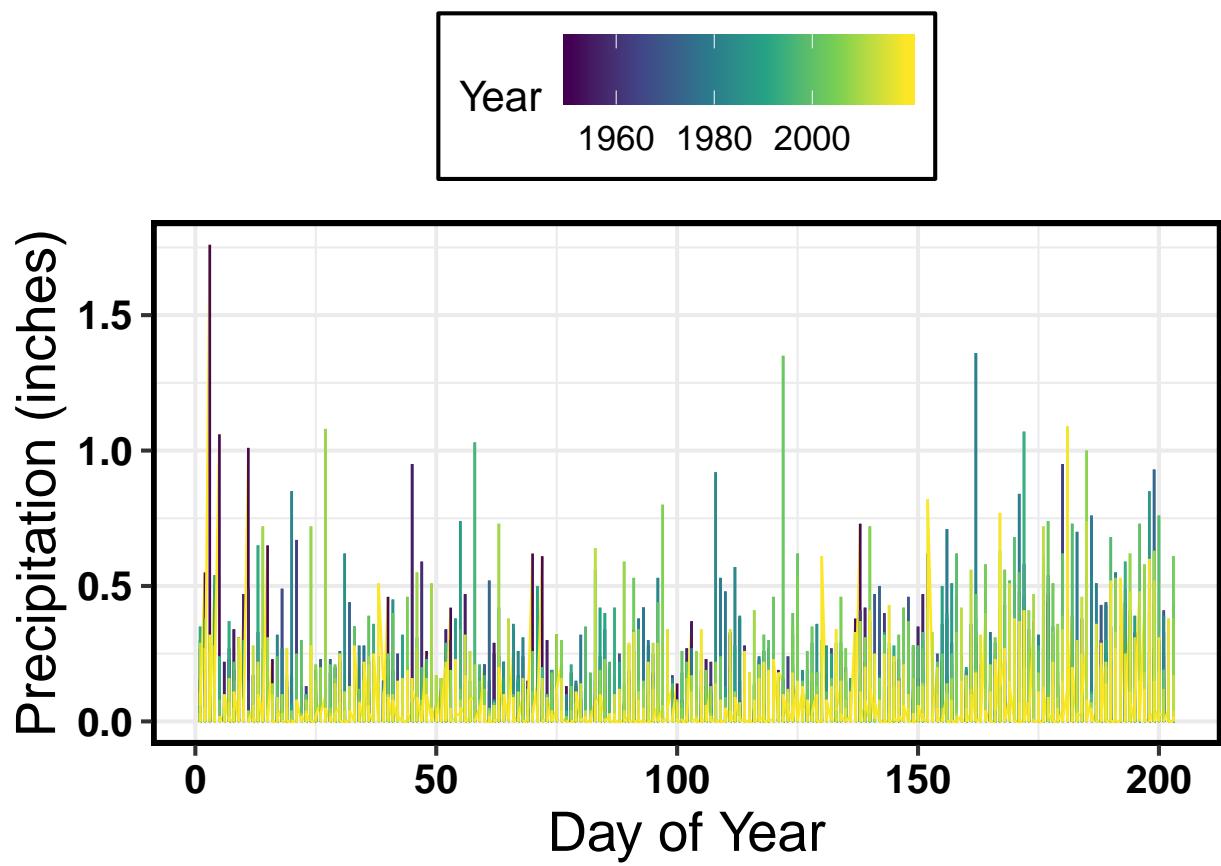


Figure 6: Day of Year vs. Precipitation. This figure shows how precipitation changes with day of year across the period of record for Bin 6.

## 6 Snowmelt Data Wrangling

## 7 Snowmelt Graph 1

## 8 Snowmelt Graph 2

## 9 Snowmelt Graph 3

## 10 Precipitation Data Wrangling

### 10.1 Scatterplot of Full Precipitation Period of Record Colored by Bin

10.1.1 Group Precip Data by Bin and DOY, create Mean Precip column, and filter weeks to only include the spring

## 11 GRAPH 1

## 12 GRAPH 2

## 13 Exploratory Precipitation Mapping

```
## Reading layer `cb_2018_us_state_20m' from data source `/Users/gabrielagarcia/Desktop/
## Simple feature collection with 52 features and 9 fields
## geometry type:  MULTIPOLYGON
## dimension:      XY
## bbox:            xmin: -179.1743 ymin: 17.91377 xmax: 179.7739 ymax: 71.35256
## epsg (SRID):    4269
## proj4string:    +proj=longlat +ellps=GRS80 +towgs84=0,0,0,0,0,0,0 +no_defs
```

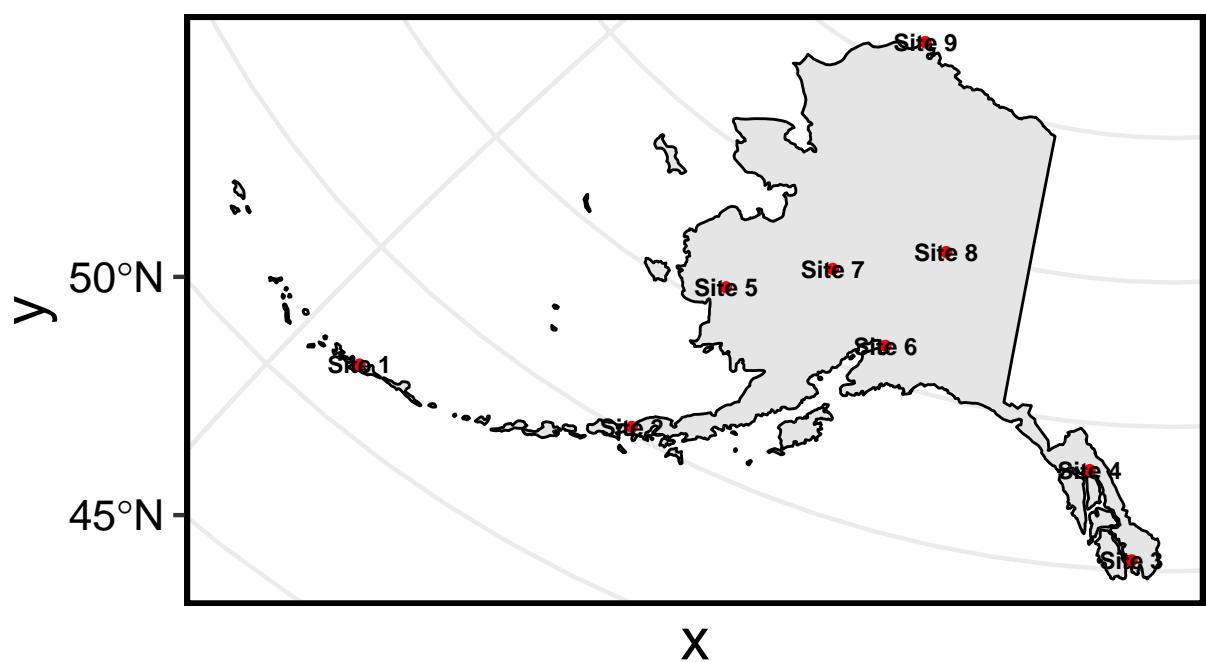
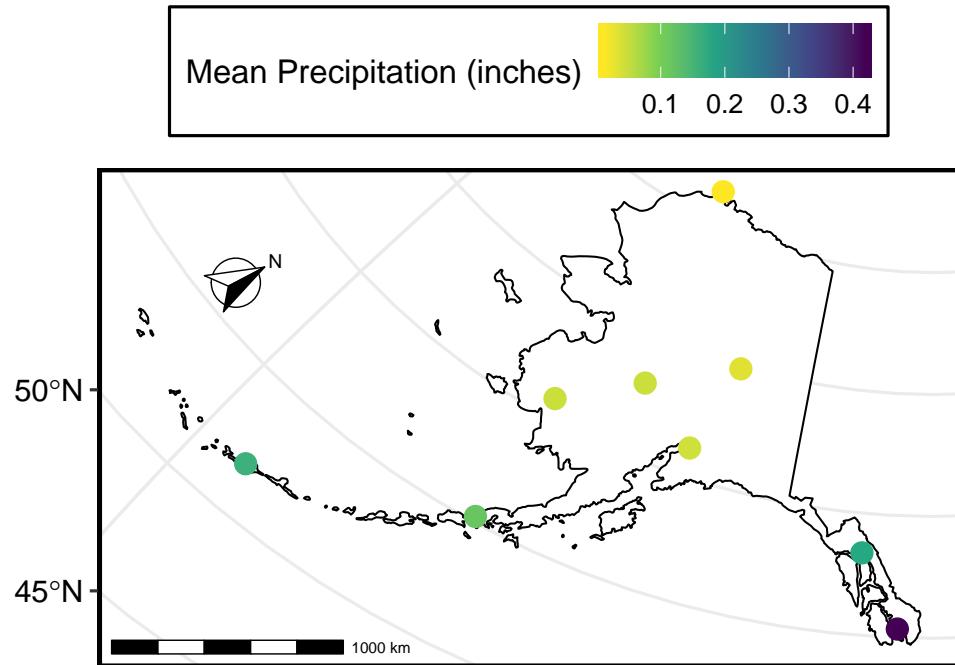
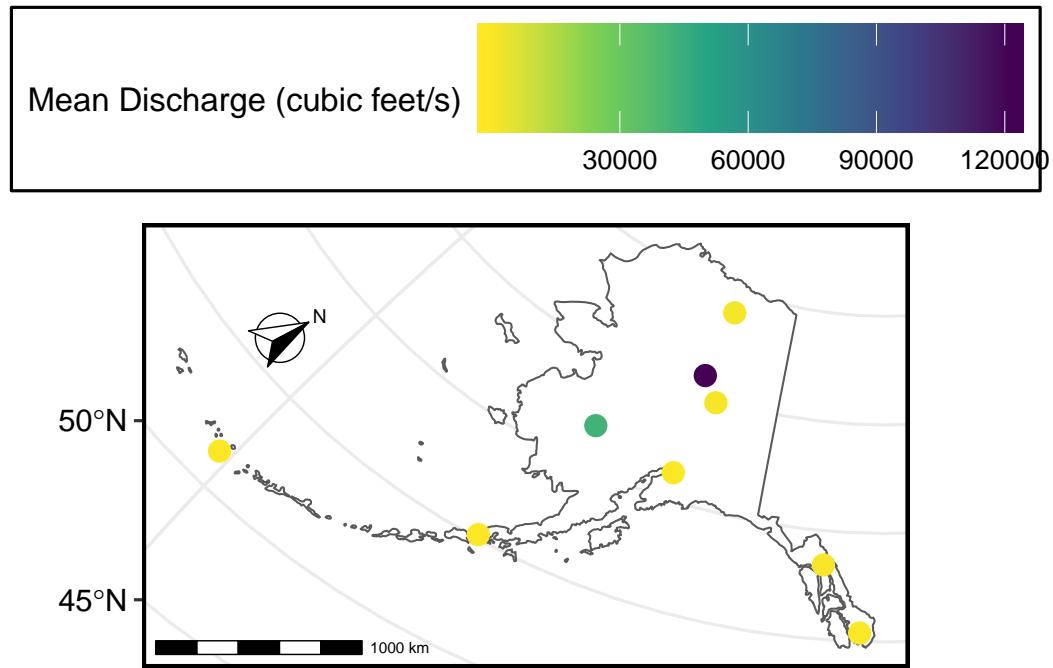


Figure 7: Map of NOAA Site Locations in Alaska

- 14 Create Alaska Temp Precip Discharge Summary using Summarize function
- 15 NOAA Site Locations Map
- 16 Mean Precipitation Map



## 17 Mean Discharge Map



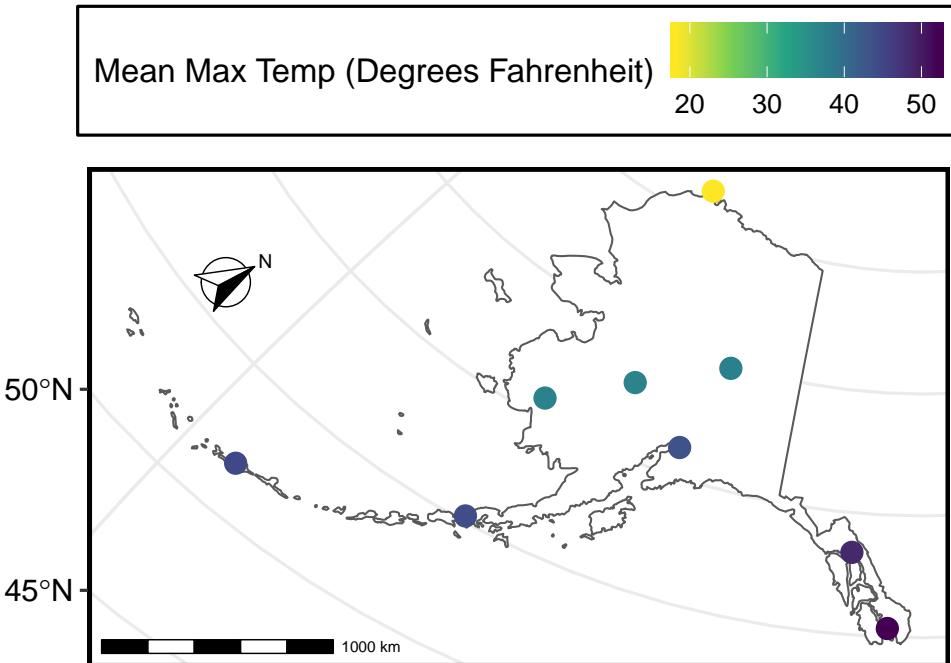


Figure 8: Map of Mean Maximum Temperature across Alaska by Bin

## 18 Comparison of Mean Precipitation, Mean Discharge, and Mean Maximum Temperature by Bin

## 19 Analysis

<Insert visualizations and text describing your main analyses. Format your R chunks so that graphs are displayed but code and other output is not displayed. Instead, describe the results of any statistical tests in the main text (e.g., “Variable x was significantly different among y groups (ANOVA; df = 300, F = 5.55, p < 0.0001)”). Each paragraph, accompanied by one or more visualizations, should describe the major findings and how they relate to the question and hypotheses. Divide this section into subsections, one for each research question.>

<Each figure should be accompanied by a caption, and each figure should be referenced within the text>

```
##  
## Call:  
## lm(formula = YEAR ~ DOY, data = Snowmelt.Discharge.Peaks5)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max  
## -36.067 -16.464    3.743  11.960  35.689
```

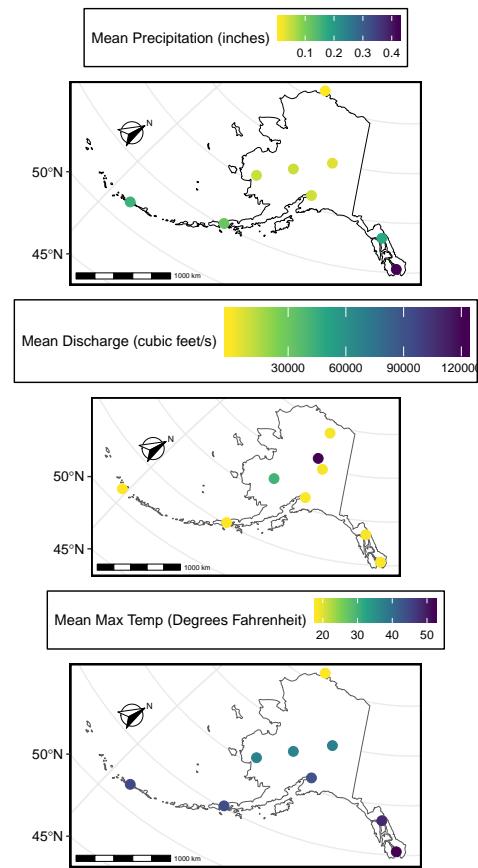


Figure 9: Comparison of Mean Precipitation, Mean Discharge, and Mean Maximum Temperature across Alaska by Bin

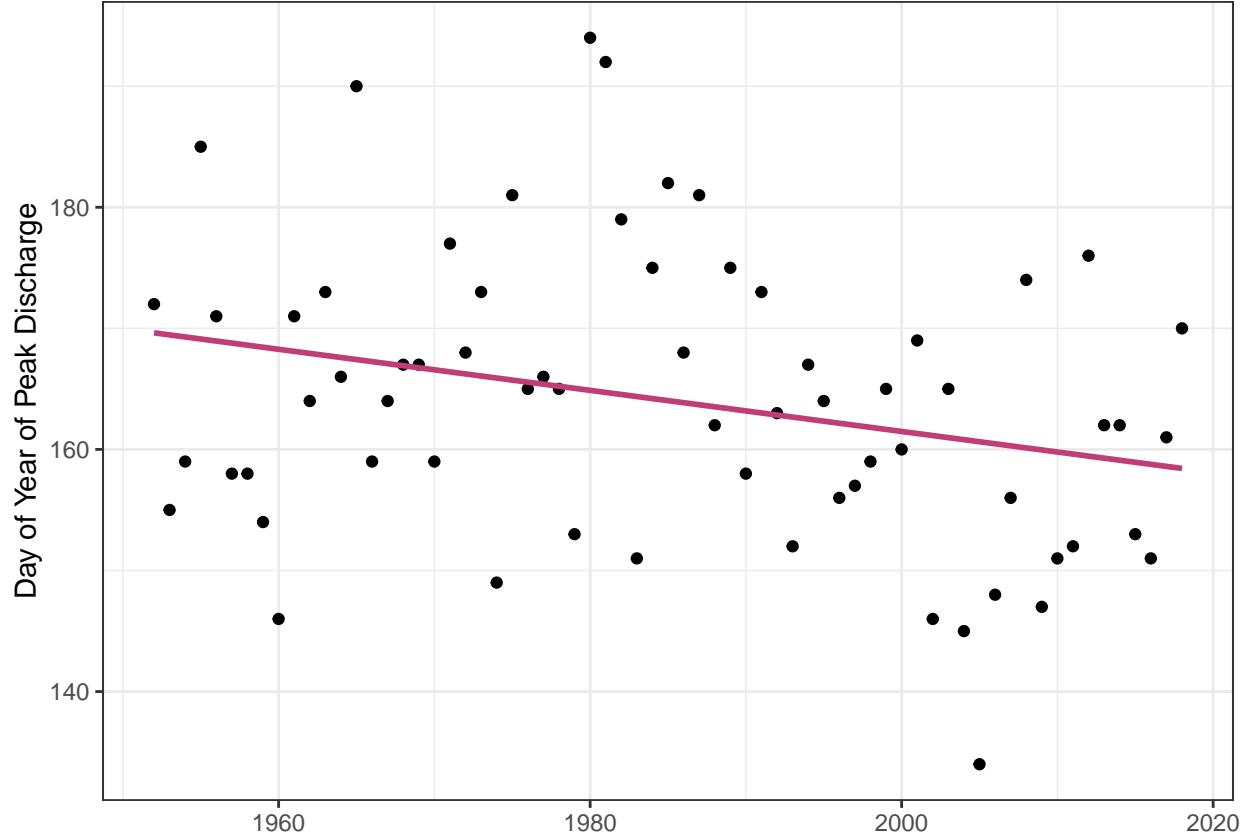


Figure 10: Year vs. Day of Year of Peak Discharge. Bin 5 is the only Latitude Bin with a significant change in the day of year of peak discharge ( $p = 0.02353$ ). There is a decreasing trend in the data, indicating the day of peak snowmelt is happening sooner across 1952-2018.

```
##  
## Coefficients:  
##             Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 2058.8834    31.9392 64.463 <2e-16 ***  
## DOY         -0.4504     0.1942 -2.319  0.0235 *  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 18.87 on 65 degrees of freedom  
## Multiple R-squared:  0.07643,   Adjusted R-squared:  0.06222  
## F-statistic: 5.379 on 1 and 65 DF,  p-value: 0.02353
```

## 19.1 Question 1:

## 19.2 Question 2:

The Climate Science Special Report estimates that maximum temperature increases in Alaska since between the first half of the 20th century and the past 30 years have been 1.43 degrees F (<https://science2017.globalchange.gov/chapter/6/>).

## **20 Summary and Conclusions**

<Summarize your major findings from your analyses in a few paragraphs. What conclusions do you draw from your findings? Relate your findings back to the original research questions and rationale.>

## **21 References**

Serreze, M. C., Barrett, A. P., Stroeve, J. C., Kindig, D. N., and Holland, M. M.: The emergence of surface-based Arctic amplification, *The Cryosphere*, 3, 11–19, <https://doi.org/10.5194/tc-3-11-2009>, 2009.