Climatic Trend Impact on Alaskan Stream Discharge

https://github.com/wgrimshaw/Alaska_DBPs.git

Gaby Garcia, Walker Grimshaw, Tristen Townsend, Yixin Wen

Abstract

Experimental overview. This section should be no longer than 250 words.

Contents

List of Tables

List of Figures

1 Research Question and Rationale

The climate is changing, in large part due to anthropogenic carbon emmissions. 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. 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.

1.1 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

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.

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

3 Exploratory Data Analysis and Wrangling

```r
# Import all Data Frames
NOAATempPrecipBin1 <- read.csv("./DATA/RAW/NOAA_Temp_Bin1.csv")
NOAATempPrecipBin2 <- read.csv("./DATA/RAW/NOAA_Temp_Bin2.csv")
NOAATempPrecipBin3 <- read.csv("./DATA/RAW/NOAA_Temp_Bin3.csv")
NOAATempPrecipBin5 <- read.csv("./DATA/RAW/Bin5TempandPrecip.csv")
NOAATempPrecip48910 <-
  read.csv("./DATA/RAW/NOAA_Precip_Temp_Bin_4_8_9_10.csv")

### Add Bin 2 Column to NOAATempPrecipBin2
NOAATempPrecipBin2$Bin <- 2

### Remove Tavg Column from Precip
NOAATempPrecip48910 <- select(NOAATempPrecip48910,-c(8))

### Add Bin Columns with values for 4, 8, and 9
NOAATempPrecip48910$Bin <- recode(NOAATempPrecip48910$STATION,
    USW00026510='4', USC00500464='8', USW00027502='9')

### Upload Bins 6 and 7
NOAATempPrecipBin67 <-
  read.csv("./DATA/PROCESSED/NOAA_Temp_Precip_Bin_6_7.csv")

### Remove Snow Column from Precip Temp 6 & 7
NOAATempPrecipBin67 <- select(NOAATempPrecipBin67,-c(8))

# Fix Date Issues in NOAA Bin 1

##Tell R to change factor to a date, in the format Month/day/year
NOAATempPrecipBin1$DATE <- as.Date(NOAATempPrecipBin1$DATE,
                                   format = "%m/%d/%y")

# Changes the format of the "datetime" column to a two digit year, a two
#digit month, and the day as a number
NOAATempPrecipBin1$DATE <- format(NOAATempPrecipBin1$DATE, "%y%m%d")


create.early.dates <- (function(d) {
       paste0(ifelse(d > 181231,"19","20"),d)
       })

### This will input the datetime column into the function we created
#above for each row of the column
```

```r
NOAATempPrecipBin1$DATE<- create.early.dates(NOAATempPrecipBin1$DATE)

### Changes the date time column into a different format to a 4 digit year,
#a two digit month, and the day as a number
NOAATempPrecipBin1$DATE <- as.Date(NOAATempPrecipBin1$DATE,
                                   format = "%Y%m%d")

# Fix Date Issues in NOAA Bin 5

## Tell R to change factor to a date, in the format Month/day/year
NOAATempPrecipBin5$DATE <- as.Date(NOAATempPrecipBin5$DATE,
                                   format = "%m/%d/%y")

## Changes the format of the "datetime" column to a two digit year,
#a two digit month, and the day as a number
NOAATempPrecipBin5$DATE<- format(NOAATempPrecipBin5$DATE, "%y%m%d")

## The 181231 is the last date of 2018. If d is over this value,
#a 19 is returned, If D<the number, a 20 is returned.
create.early.dates <- (function(d) {
       paste0(ifelse(d > 181231,"19","20"),d)
       })

### This will input the datetime column into the function we created above
#for each row of the column
NOAATempPrecipBin5$DATE<- create.early.dates(NOAATempPrecipBin5$DATE)

### Changes the date time column into a different format to a 4 digit year,
#a two digit month, and the day as
NOAATempPrecipBin5$DATE <- as.Date(NOAATempPrecipBin5$DATE,
                                   format = "%Y%m%d")

# RBind all TempPrecip Data
TempPrecip<-rbind(NOAATempPrecipBin1, NOAATempPrecipBin2,
                  NOAATempPrecipBin3, NOAATempPrecipBin5,
                  NOAATempPrecipBin67, NOAATempPrecip48910)

# Import USGS Discharge Data
NWIS_Discharge<-read.csv("./DATA/RAW/NWIS_Discharge.csv")

### Convert site_no to a factor
NWIS_Discharge$site_no<-as.factor(NWIS_Discharge$site_no)

## Add Bins to Discharge Data
```

```r
NWIS_Discharge$Bin<-recode(NWIS_Discharge$site_no,
       "15297680" = '1', "15297610" ='2', "15072000"='3',
         "15052500" ='4', "15276000" ='5', "15304000"='6',
         "15514000" ='7', "15453500" ='8', "15908000"='9',
         "15896000" ='10')

### Rename "Date" to "DATE" in NWIS_Discharge dataframe so both columns match
names(NWIS_Discharge)[3]<-"DATE"

# Join TempPrecip and NWIS_Discharge Data Frames by Date and Bin
AlaskaTempPrecipDischarge <- TempPrecip %>%
  merge(NWIS_Discharge, by = c('DATE', 'Bin')) %>%
  rename(Discharge = X_00060_00003) %>%
  select(DATE:Discharge)
```

The first step of the analysis process is determining the presence and magnitude of temperature change across the latitude range of Alaska. The daily maximum temperature data show no obvious trends over time for any of the sites. Furthermore, the sites differ substantially in the period of record of their temperature data. Sites with no recent observations may need to be dropped from analysis, unless nearby streams have similar periods of record.

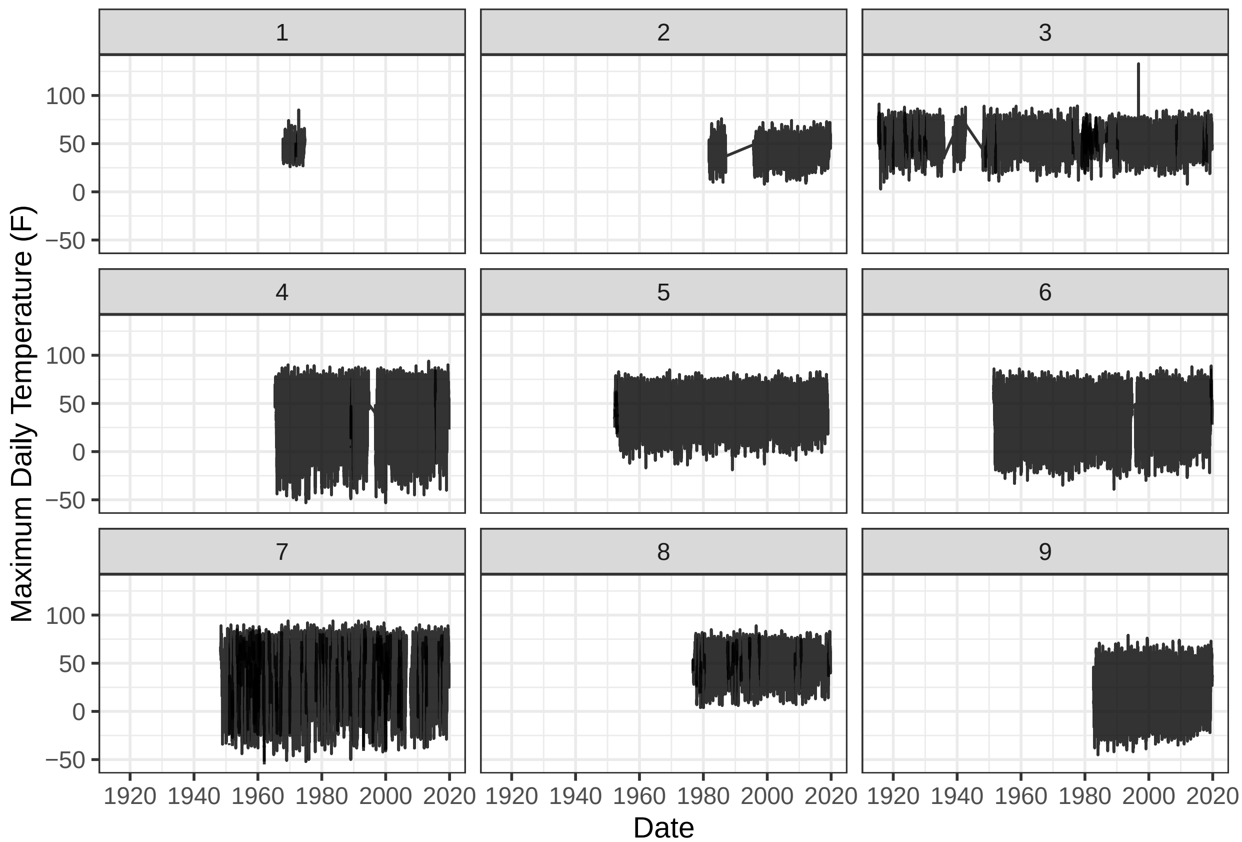
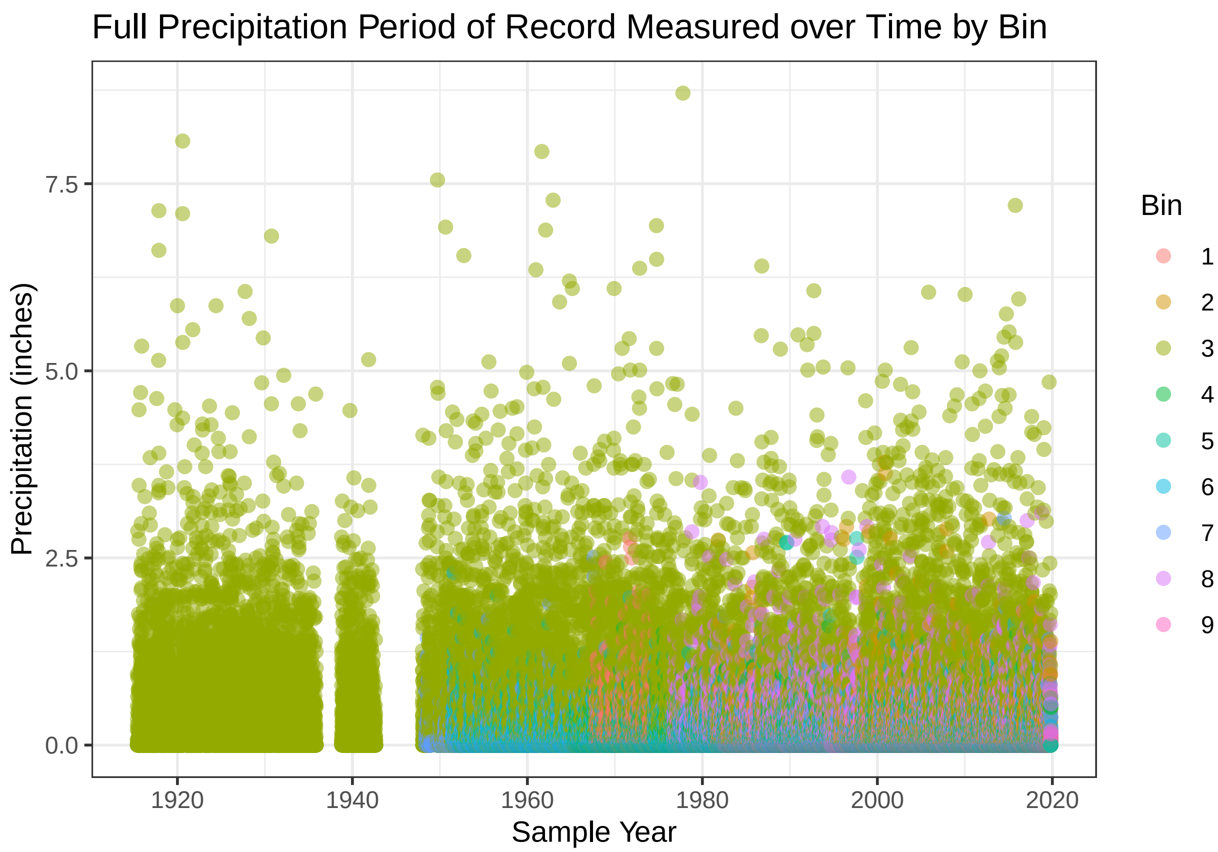


Figure 1: Daily maximum temperature for the period of record of each latitude bin. There are no obvious trends, and the period of record differs substantially for each site.

4 Temperature

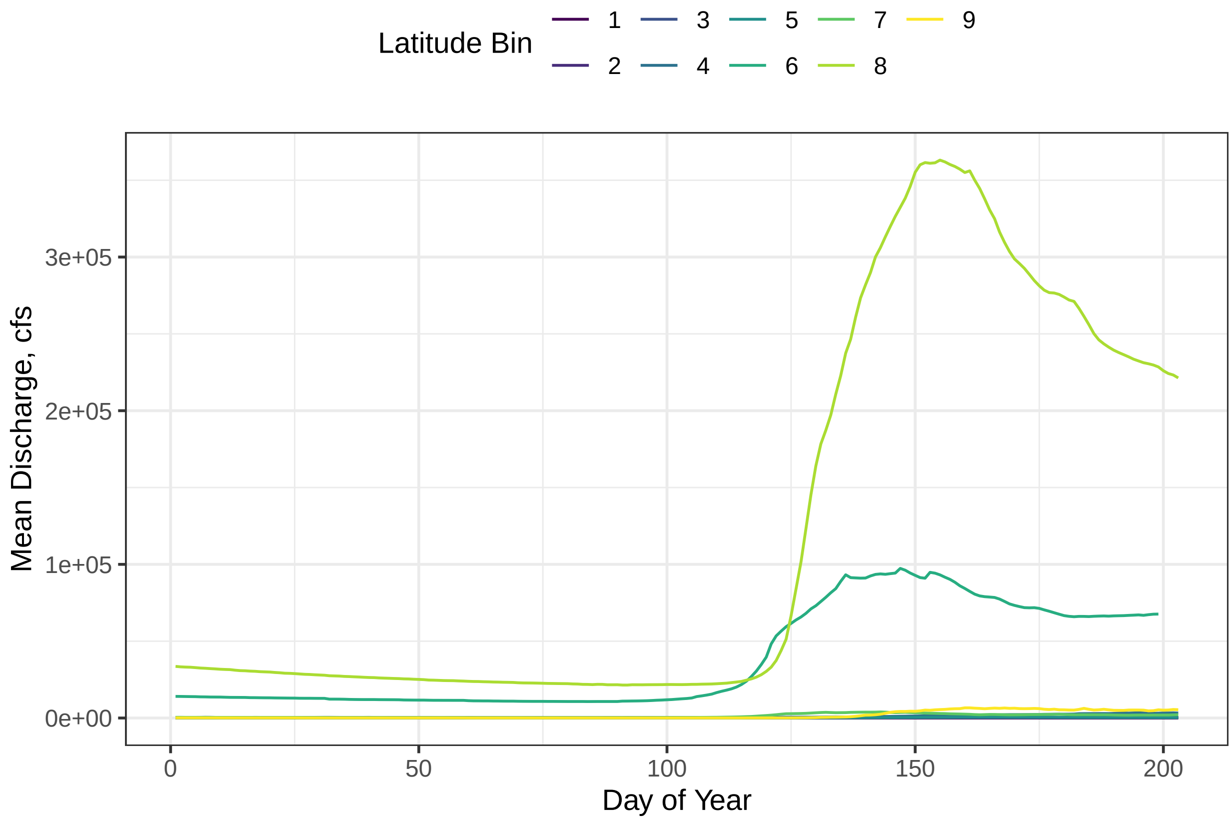


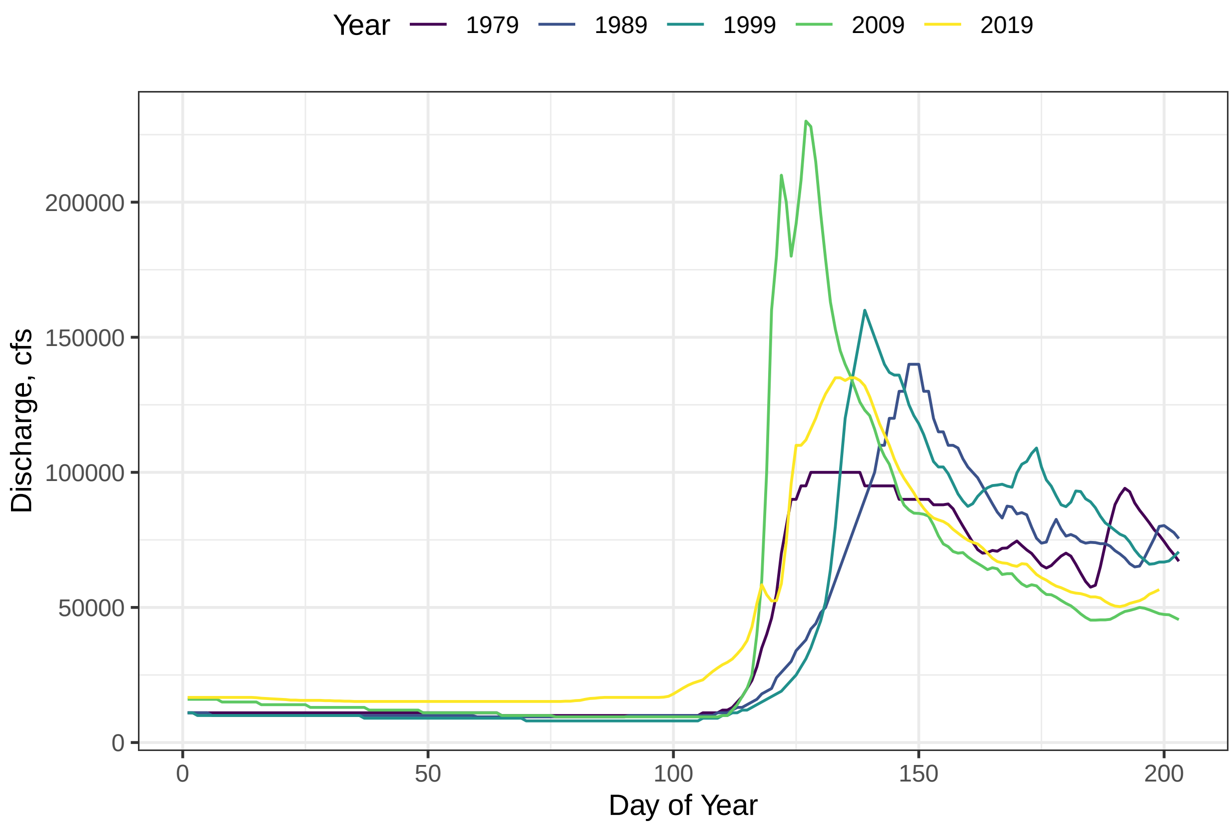
#Precipitation

Give a general image of how discharge changes over time in different latitudes.

5 Discharge

Discharge is to be used as a proxy for snowmelt. The data was loaded and wrangled to include the year, month, day, day of year, and week in order to better analyze if and how the day of first snowmelt is changing. The data was filtered for weeks 1-30 of each year (mid-winter to mid-summer), to minimize the amount of data being plotted in exploratory graphs. The mean discharge for each day of year through time by latitude bin was plotted. This served to provide insight into roughly when the day of first snowmelt was for each site, and to see how different sites varied. The data was then filtered to look at a single latitude bin, in this case bin 6, and see how discharge changed by the day of year across the period of record. In order to better visual any trend, bin 6 was filtered for data on a ten year frequency for the past fifty years (1979, 1989, 1999, 2009, 2019). After looking at the various exploratory graphs, it was decided to proceed with using peak discharge as a proxy for peak snowmelt, rather than try to determine when the first snowmelt of each year was occuring.


Latitude Bin
1
2
3
4
5
6
7
8
9
Mean Discharge, cfs
0e+00
1e+05
2e+05
3e+05
0
50
100
150
200
Day of Year



Year
1979
1989
1999
2009
2019
Discharge, cfs
200000
150000
100000
50000
0
0
50
100
150
200
Day of Year

6 Analysis

6.1 Long-term Temperature and Discharge Trends

A seasonal Mann-Kendall test was used to determine if there was a long-term trend in maximum temperature and stream discharge for each latitude.

A seasonal Sen's Slope quantifies the magnitude of the temperature and discharge trends.

From general discharge image of 10 bins, Bin 6 seems to have seasonal change. Thus apply smk.test on bin 6 discharge data to see if there's a seasonal trend. adf.test is also applied to see if the data is stationary. If it is, use auto.arima function to find the best fit model to predict the future discharge and plot the graph.

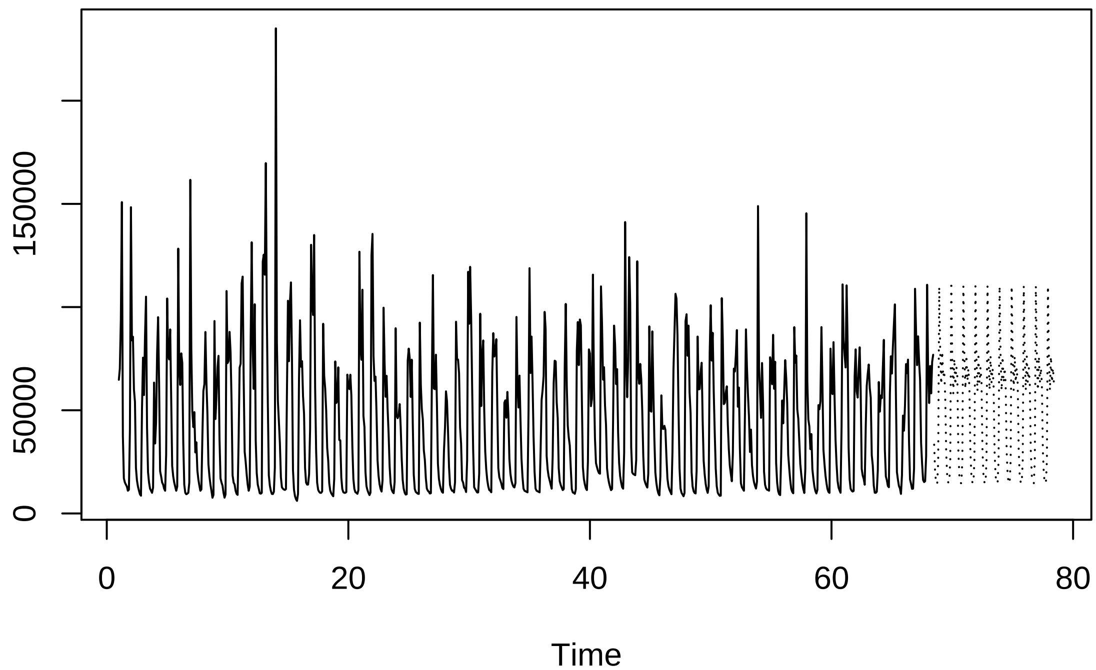
```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  AlaskaDischarge_6_ts.Monthly
## z = 3.778, p-value = 0.0001581
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##        S     varS
##   2446.0 418818.7
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: AlaskaDischarge_6_ts.Monthly
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                        S    varS    tau      z   Pr(>|z|)
## Season 1:   S = 0  -332 35688.7 -0.146 -1.752   0.079754   .
## Season 2:   S = 0  -220 35688.7 -0.097 -1.159   0.246352
## Season 3:   S = 0  -228 35688.7 -0.100 -1.202   0.229518
## Season 4:   S = 0  -174 35688.7 -0.076 -0.916   0.359794
## Season 5:   S = 0   372 35688.7  0.163  1.964   0.049547   *
## Season 6:   S = 0   796 35682.7  0.350  4.209 2.5695e-05 ***
## Season 7:   S = 0   438 34123.3  0.199  2.366   0.017997   *
## Season 8:   S = 0   270 34099.3  0.123  1.457   0.145191
## Season 9:   S = 0   291 34078.3  0.133  1.571   0.116197
## Season 10:   S = 0  400 34099.3  0.182  2.161   0.030716   *
## Season 11:   S = 0  724 34144.7  0.328  3.913 9.1269e-05 ***
## Season 12:   S = 0  109 34147.7  0.049  0.584   0.558921
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##  Augmented Dickey-Fuller Test
##
## data:  AlaskaDischarge_6_ts.Monthly
## Dickey-Fuller = -11.464, Lag order = 9, p-value = 0.01
## alternative hypothesis: stationary
##
##  Fitting models using approximations to speed things up...
##
##  ARIMA(2,0,2)(1,1,1)[12] with drift         : Inf
##  ARIMA(0,0,0)(0,1,0)[12] with drift         : 18130.39
##  ARIMA(1,0,0)(1,1,0)[12] with drift         : 17827.96
##  ARIMA(0,0,1)(0,1,1)[12] with drift         : 17689.86
##  ARIMA(0,0,0)(0,1,0)[12]                    : 18128.39
##  ARIMA(0,0,1)(0,1,0)[12] with drift         : 18068.86
##  ARIMA(0,0,1)(1,1,1)[12] with drift         : 17681.43
##  ARIMA(0,0,1)(1,1,0)[12] with drift         : 17835.47
##  ARIMA(0,0,1)(2,1,1)[12] with drift         : Inf
##  ARIMA(0,0,1)(1,1,2)[12] with drift         : 17645.1
##  ARIMA(0,0,1)(0,1,2)[12] with drift         : 17691.51
##  ARIMA(0,0,1)(2,1,2)[12] with drift         : Inf
##  ARIMA(0,0,0)(1,1,2)[12] with drift         : 17704.11
##  ARIMA(1,0,1)(1,1,2)[12] with drift         : Inf
##  ARIMA(0,0,2)(1,1,2)[12] with drift         : 17643.61
##  ARIMA(0,0,2)(0,1,2)[12] with drift         : 17689.73
##  ARIMA(0,0,2)(1,1,1)[12] with drift         : 17679.29
##  ARIMA(0,0,2)(2,1,2)[12] with drift         : Inf
##  ARIMA(0,0,2)(0,1,1)[12] with drift         : 17688
##  ARIMA(0,0,2)(2,1,1)[12] with drift         : Inf
##  ARIMA(1,0,2)(1,1,2)[12] with drift         : Inf
##  ARIMA(0,0,3)(1,1,2)[12] with drift         : 17643.1
##  ARIMA(0,0,3)(0,1,2)[12] with drift         : 17691.24
##  ARIMA(0,0,3)(1,1,1)[12] with drift         : 17680.44
##  ARIMA(0,0,3)(2,1,2)[12] with drift         : Inf
##  ARIMA(0,0,3)(0,1,1)[12] with drift         : 17689.47
##  ARIMA(0,0,3)(2,1,1)[12] with drift         : Inf
##  ARIMA(1,0,3)(1,1,2)[12] with drift         : Inf
##  ARIMA(0,0,4)(1,1,2)[12] with drift         : 17645.15
##  ARIMA(1,0,4)(1,1,2)[12] with drift         : Inf
##  ARIMA(0,0,3)(1,1,2)[12]                    : 17641.07
##  ARIMA(0,0,3)(0,1,2)[12]                    : 17689.22
##  ARIMA(0,0,3)(1,1,1)[12]                    : 17678.4
```

```
##  ARIMA(0,0,3)(2,1,2)[12]                    : Inf
##  ARIMA(0,0,3)(0,1,1)[12]                    : 17687.45
##  ARIMA(0,0,3)(2,1,1)[12]                    : Inf
##  ARIMA(0,0,2)(1,1,2)[12]                    : 17641.58
##  ARIMA(1,0,3)(1,1,2)[12]                    : Inf
##  ARIMA(0,0,4)(1,1,2)[12]                    : 17643.1
##  ARIMA(1,0,2)(1,1,2)[12]                    : Inf
##  ARIMA(1,0,4)(1,1,2)[12]                    : Inf
##
##  Now re-fitting the best model(s) without approximations...
##
##  ARIMA(0,0,3)(1,1,2)[12]                    : Inf
##  ARIMA(0,0,2)(1,1,2)[12]                    : Inf
##  ARIMA(0,0,3)(1,1,2)[12] with drift         : Inf
##  ARIMA(0,0,4)(1,1,2)[12]                    : Inf
##  ARIMA(0,0,2)(1,1,2)[12] with drift         : Inf
##  ARIMA(0,0,1)(1,1,2)[12] with drift         : Inf
##  ARIMA(0,0,4)(1,1,2)[12] with drift         : Inf
##  ARIMA(0,0,3)(1,1,1)[12]                    : Inf
##  ARIMA(0,0,2)(1,1,1)[12] with drift         : Inf
##  ARIMA(0,0,3)(1,1,1)[12] with drift         : Inf
##  ARIMA(0,0,1)(1,1,1)[12] with drift         : Inf
##  ARIMA(0,0,3)(0,1,1)[12]                    : Inf
##  ARIMA(0,0,2)(0,1,1)[12] with drift         : Inf
##  ARIMA(0,0,3)(0,1,2)[12]                    : Inf
##  ARIMA(0,0,3)(0,1,1)[12] with drift         : Inf
##  ARIMA(0,0,2)(0,1,2)[12] with drift         : Inf
##  ARIMA(0,0,1)(0,1,1)[12] with drift         : Inf
##  ARIMA(0,0,3)(0,1,2)[12] with drift         : Inf
##  ARIMA(0,0,1)(0,1,2)[12] with drift         : Inf
##  ARIMA(0,0,0)(1,1,2)[12] with drift         : Inf
##  ARIMA(1,0,0)(1,1,0)[12] with drift         : 18096.57
##
##  Best model: ARIMA(1,0,0)(1,1,0)[12] with drift
```

```
## Series: AlaskaDischarge_6_ts.Monthly
## ARIMA(1,0,0)(1,1,0)[12] with drift
##
## Coefficients:
##          ar1     sar1    drift
##       0.3063  -0.4923  10.0841
## s.e.  0.0339   0.0313  57.7744
##
## sigma^2 estimated as 408872984:  log likelihood=-9044.26
## AIC=18096.52   AICc=18096.57   BIC=18115.25
```



6.2 Predictability of snowmelt

Is there a change over time of the period when snow melts? Peak discharge will be used as a proxy for peak snowmelt. At each latitude bin, the day of year of peak snowmelt will be found for each year in the period of record. If thare are any gaps in the data, it will be addressed through linear interpolation. Once the day of year of peak snowmelt is found, a dataframe of 'Year' and 'Day of Year of Peak Snowmelt' will be created, and a time series will be run to see if there is any significant change at each site in the day of peak snowmelt.

7 Summary and Conclusions

<Summarize your major findings from your analyses. What conclusions do you draw from your findings? Make sure to apply this to a broader application for the research question you have answered.>