Assignment 6: Time Series Analysis

Yutao Gong

OVERVIEW

This exercise accompanies the lessons in Hydrologic Data Analysis on time series analysis

Directions

1. Change "Student Name" on line 3 (above) with your name.
2. Work through the steps, **creating code and output** that fulfill each instruction.
3. Be sure to **answer the questions** in this assignment document.
4. When you have completed the assignment, **Knit** the text and code into a single pdf file.
5. After Knitting, submit the completed exercise (pdf file) to the dropbox in Sakai. Add your last name into the file name (e.g., "A06_Salk.html") prior to submission.

The completed exercise is due on 11 October 2019 at 9:00 am.

Setup

1. Verify your working directory is set to the R project file,
2. Load the tidyverse, lubridate, trend, and dataRetrieval packages.
3. Set your ggplot theme (can be theme_classic or something else)
4. Load the ClearCreekDischarge.Monthly.csv file from the processed data folder. Call this data frame ClearCreekDischarge.Monthly.

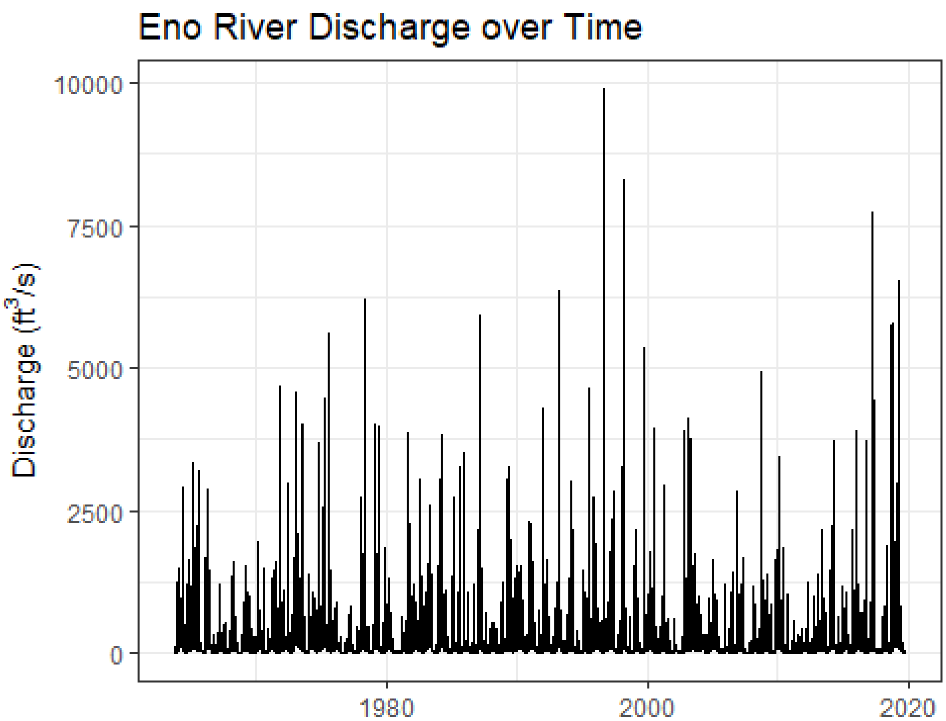
```
getwd()

## [1] "C:/Users/gongy/Documents/Hydrologic_Data_Analysis/Assignments"

library(tidyverse)

## -- Attaching packages -------------------------------------- tidyverse
1.2.1 --

## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   0.8.3     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0

## -- Conflicts ----------------------------------------
tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(lubridate)
```

```
##
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':
##
##     date

library(trend)
library(dataRetrieval)

theme_set(theme_bw())

ClearCreekDischarge.Monthly =
read.csv("../Data/Processed/ClearCreekDischarge.Monthly.csv")
```

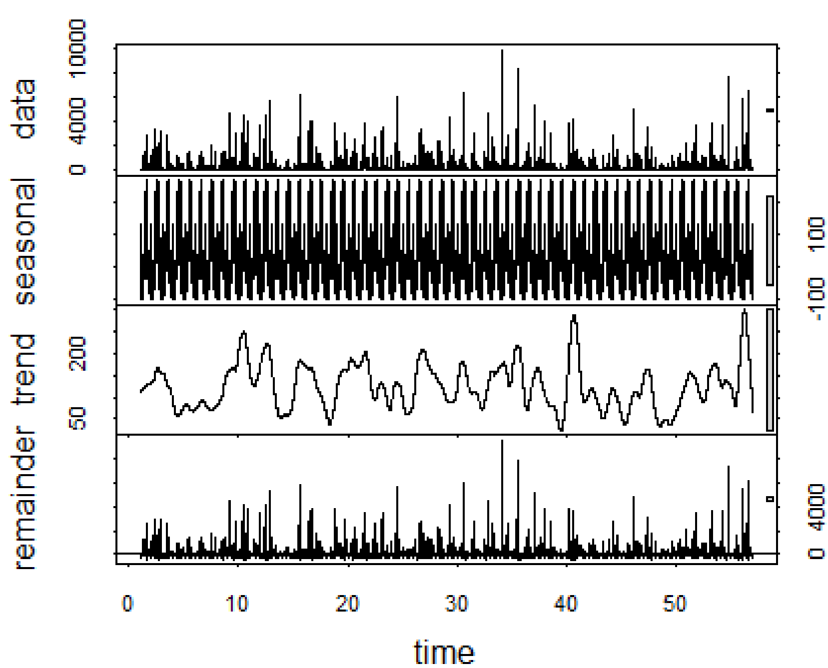
Time Series Decomposition

5. Create a new data frame that includes daily mean discharge at the Eno River for all available dates (`siteNumbers = "02085070"`). Rename the columns accordingly.
6. Plot discharge over time with geom_line. Make sure axis labels are formatted appropriately.
7. Create a time series of discharge
8. Decompose the time series using the `stl` function.
9. Visualize the decomposed time series.

```
Enodata <- readNWISdv(siteNumbers = "02085070",
                      parameterCd = "00060",
                      startDate = "",
                      endDate = "")
names(Enodata)[4:5] <- c("Discharge", "Approval.Code")
EnoPlot <-
  ggplot(Enodata, aes(x = Date, y = Discharge)) +
         geom_line() + labs(x = "", y = expression("Discharge (ft"^3*"/s)"),
title = "Eno River Discharge over Time")
print(EnoPlot)
```




```
Eno_ts <- ts(Enodata[[4]], frequency = 365)
Eno_Decomposed <- stl(Eno_ts, s.window = "periodic")
plot(Eno_Decomposed)
```



10. How do the seasonal and trend components of the decomposition compare to the Clear Creek discharge dataset? Are they similar in magnitude?

Seasonal: The seasonal component of Eno River discharge dataset looks "denser", which implies that its discharge seasonality is not as strong as the Clear Creek. The magnitudes are different: Eno River's seasonality ranges from around -100 to 100, while Clear Creek's seasonality ranges from around 0 to 400.

Trend: The trend of Eno River discharge dataset looks comparatively more similar to that of Clear Creek, except that Eno River has more sudden peaks in later times. The magnitudes are similar.

Trend Analysis

Research question: Has there been a monotonic trend in discharge in Clear Creek over the period of study?

11. Generate a time series of monthly discharge in Clear Creek from the ClearCreekDischarge.Monthly data frame. This time series should include just one column (discharge).
12. Run a Seasonal Mann-Kendall test on the monthly discharge data. Inspect the overall trend and the monthly trends.

```
ClearCreek.Monthly_ts <- ts(ClearCreekDischarge.Monthly$Discharge, frequency
= 12, start = c(1974, 10), end = c(2019,10))
ClearCreektrend <- smk.test(ClearCreek.Monthly_ts)
summary(ClearCreektrend)
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data: ClearCreek.Monthly_ts
## alternative hypothesis: two.sided
##
## Statistics for individual seasons
##
## H0
##                      S  varS    tau      z Pr(>|z|)
## Season 1:   S = 0   35 10449  0.035  0.333 0.739425
## Season 2:   S = 0    4 10450  0.004  0.029 0.976588
## Season 3:   S = 0  204 10450  0.206  1.986 0.047054 *
## Season 4:   S = 0  230 10450  0.232  2.240 0.025081 *
## Season 5:   S = 0  148 10450  0.149  1.438 0.150434
## Season 6:   S = 0   94 10450  0.095  0.910 0.362951
## Season 7:   S = 0  -54 10450 -0.055 -0.518 0.604135
## Season 8:   S = 0  -99 10449 -0.100 -0.959 0.337703
## Season 9:   S = 0  -90 10450 -0.091 -0.871 0.383958
## Season 10:   S = 0  64 11154  0.062  0.597 0.550828
## Season 11:   S = 0  24 10450  0.024  0.225 0.821984
## Season 12:   S = 0  30 10450  0.030  0.284 0.776650
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
ClearCreektrend
```

```
##
##  Seasonal Mann-Kendall trend test (Hirsch-Slack test)
##
## data:  ClearCreek.Monthly_ts
## z = 1.6586, p-value = 0.09719
## alternative hypothesis: true S is not equal to 0
## sample estimates:
##      S   varS
##    590 126102
```

13. Is there an overall monotonic trend in discharge over time? If so, is it positive or negative?

p-value is over 0.05 so evidence is not significant enough for me to reject the null hepothesis, so there is no overall monotonic trend in discharge over time.

14. Are there any monthly monotonic trends in discharge over time? If so, during which months do they occur and are they positive or negative?

Yes, they occur in March and April and they are positive

Reflection

15. What are 2-3 conclusions or summary points about time series you learned through your analysis?

1. Time series data is very useful when it comes to analyzing seasonality
2. There are multiple effects layered together on a time series data so it is important to decompose it

16. What data, visualizations, and/or models supported your conclusions from 12?

The plot of decomposed time series data

17. Did hands-on data analysis impact your learning about time series relative to a theory-based lesson? If so, how?

Yes, it made time series data more intuitive (by visualizing)

18. How did the real-world data compare with your expectations from theory?

It needs more significant evidence to determine seasonality trends in a river than I thought it would.