Assignment 6: Time Series Analysis

Walker Grimshaw

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.

```
## prevent warnings in knitted pdf
knitr::opts_chunk$set(warning = FALSE)

getwd()
```

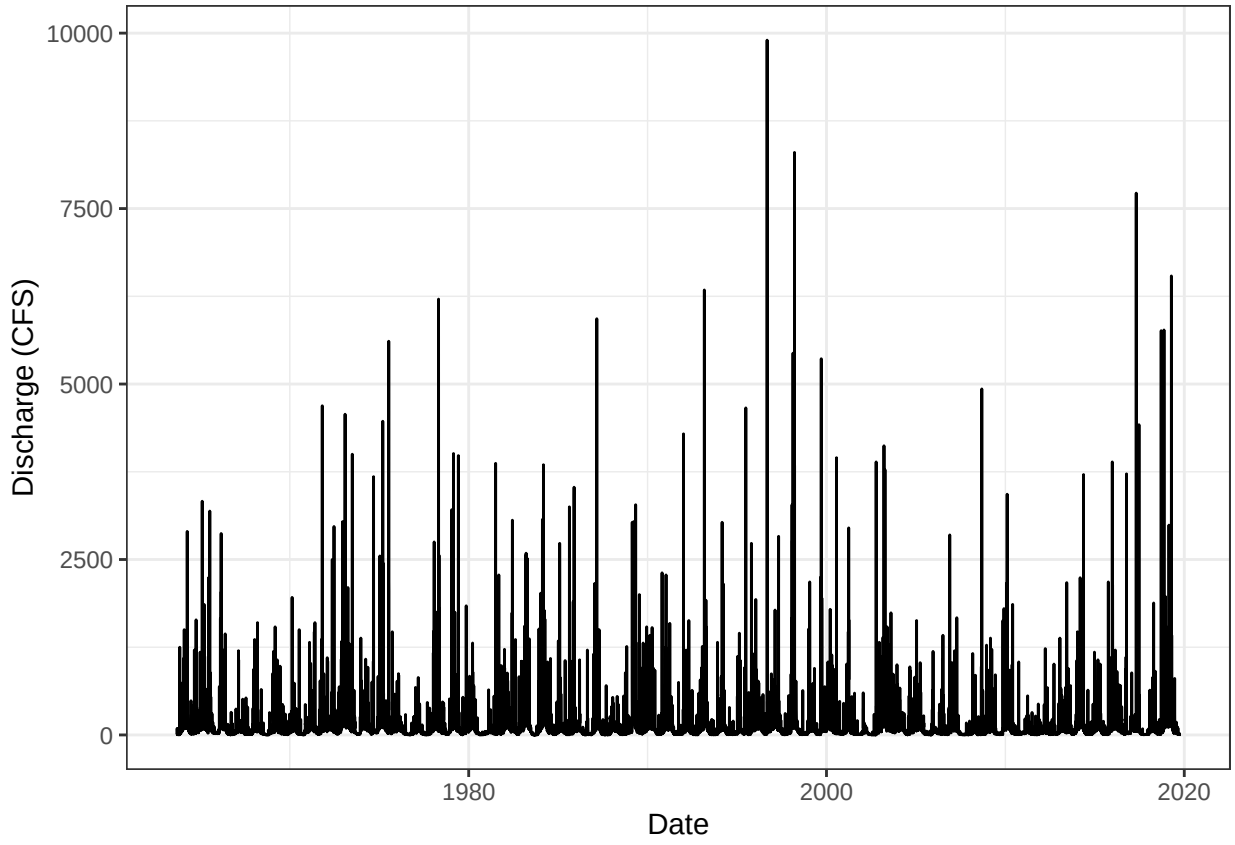
```
## [1] "C:/Users/walke/OneDrive/Documents/Duke/Courses/Fall_2019/Hydrologic_Data_Analysis/Assignments"
```

```
packages <- c("tidyverse",
              "lubridate",
              "trend",
              "dataRetrieval")
invisible(lapply(packages, library, character.only = TRUE))

theme_set(theme_bw())

## Load Data
ClearCreekDischarge.Monthly <-
  read.csv("../Data/Processed/ClearCreekDischarge.Monthly.csv", header = T)
```

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.

```r
## 5
EnoRiverDischarge <- readNWISdv(siteNumbers = "02085070",
                     parameterCd = "00060", # discharge (ft3/s)
                     startDate = "",
                     endDate = "")
names(EnoRiverDischarge)[4:5] <- c("Discharge", "Approval.Code")

## 6
ggplot(EnoRiverDischarge, aes(x = Date, y = Discharge)) +
  geom_line() +
  labs(y = "Discharge (CFS)")
```

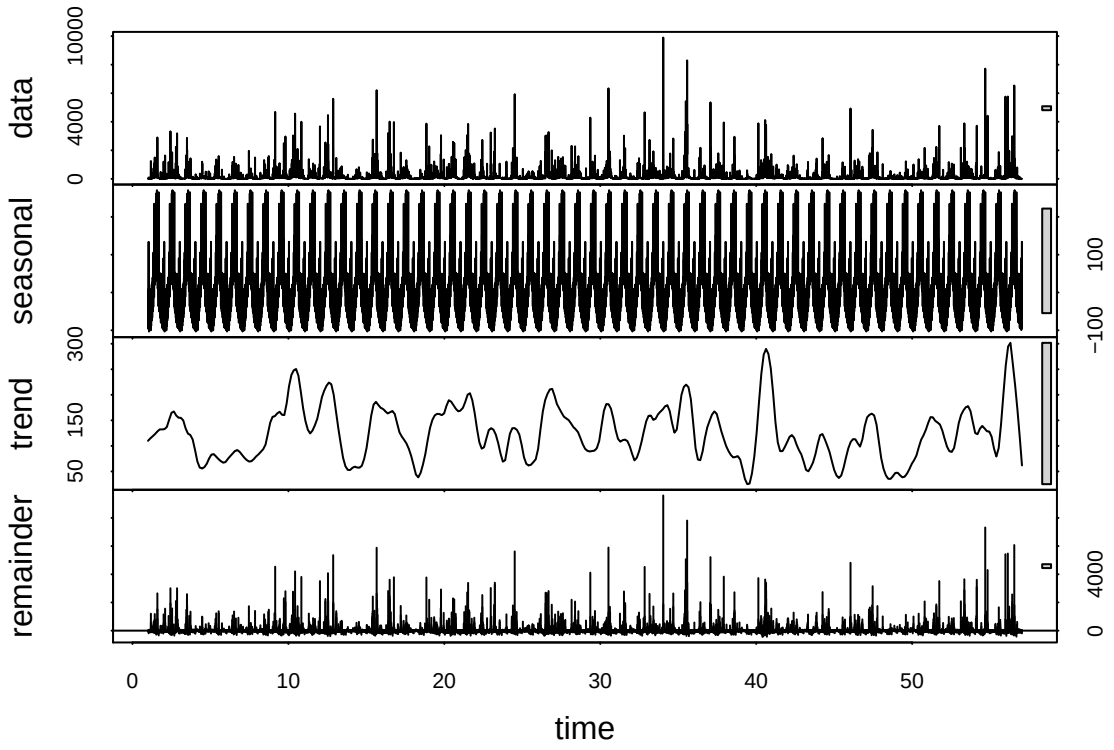



```r
## 7
## check spacing of data
table(diff(EnoRiverDischarge$Date))
```

```
##
##     1    39
## 20454     1
```

```r
## there is only one gap in the data of 39 days
Eno_ts <- ts(EnoRiverDischarge[[4]], frequency = 365)

## 8
Eno_Decomposed <- stl(Eno_ts, s.window = "periodic")

## 9
```

```
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 the Eno Discharge is smaller about half the magnitude of the seasonal component of the Clear Creek Discharge, but it also inlcudes larger negative values. However, the total discharge of the Eno River is usually greater than the total discharge of the Clear Creek, so the seasonal discharge of the Eno is even less comparatively.

    Trend: As with the Clear Creek Discharge, there is little trend in the Eno River Discharge. The trends of the two discharge sets have similarly small magnitudes.

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.

```
## Monthly time series
ClearCreek_ts <- ts(ClearCreekDischarge.Monthly[[3]], frequency = 12)

## Seasonal Mann Kendall
ClearCreekTrend <- smk.test(ClearCreek_ts)
```

```
ClearCreekTrend
```

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

```
summary(ClearCreekTrend)
```

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

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

    There is not an overall monotonic trend in the discharge of Clear Creek 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?

    The only statistically significant monotonic trends in Clear Creek discharge occur in June and July. In both months, discharge has increased over the period of record.

Reflection

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

    Long-term trends are difficult to detect in high frequency data over long time periods. Stream discharge is especially difficult to predict.

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

The majority of the discharge data does not fall into the trend or seasonality for either of the rivers, even though they are two very different rivers.

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

    The seasonality visualization for the Eno River actually decreased my understanding slightly because there was such little seasonality to the data. Overall, it was good to actually visualize the trends though.

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

    The difference in seasonality of discharge between a Western, snow fed river and an Eastern river was very much what I expected. Trends are however more difficult to detect and similarly difficult to recognize visually in plots.