Marcell Experimental Forest: S2 Stream Height (Water Years 2017 - 2019)

Mia Forsline

2022-06-24

Contents

Set Up

- Install/load necessary R packages
- Set working directory if necessary (or create a file path to use throughout the RMD)

Water Year 2017

- Read in the WY 2017 data from the External-MEF_DATA Box folder
- Specifically, raw data is found at the following file path: `External-MEF_DATA/Hydro/Streamflow/L0_subdaily/StripC`
- Clean and manipulate the WY 2017 data into a tidy format

    - Renaming columns to use lower_snake_case and avoid special characters
    - create a new `date` column that follows the format YYYY-MM-DD (and does not include a timestamp)

- Test that the manipulated data frame has the same number of rows as the original data frame (we do not want to accidentally lose data, so the code will throw an error is the two data frames do not match)
- Graph the WY 2017 data

```r
#read in the 2017 data
wy2017 <- read_excel(path = here(filepath, "WY2017.S2_Breakpoint.xlsx"),
                     skip =3)

#clean the 2017 data
wy2017_clean <- wy2017 %>%
  rename(datetime = "Date/time",
         stream_height_ft = "Stage.ft") %>%
  select(c("datetime", "stream_height_ft")) %>%
  clean_names() %>%
  mutate(date = datetime,
         date = format(as.POSIXct(datetime,format='%m/%d/%Y %H:%M:%S'),format='%Y-%m-%d'),
         date = as.POSIXct(date))

#test the 2017 data
if(nrow(wy2017) != nrow(wy2017_clean)) stop("Check clean dataframe dimensions")

#plot
ggplot(data = wy2017_clean) +
  geom_line(aes(x = datetime, y = stream_height_ft)) +
  theme_classic() +
  labs(x = "Time",
       y = "Stream Height (Ft)",
       title = "WY 2017 S2 Bog Stream Height")
```


WY 2017 S2 Bog Stream Height
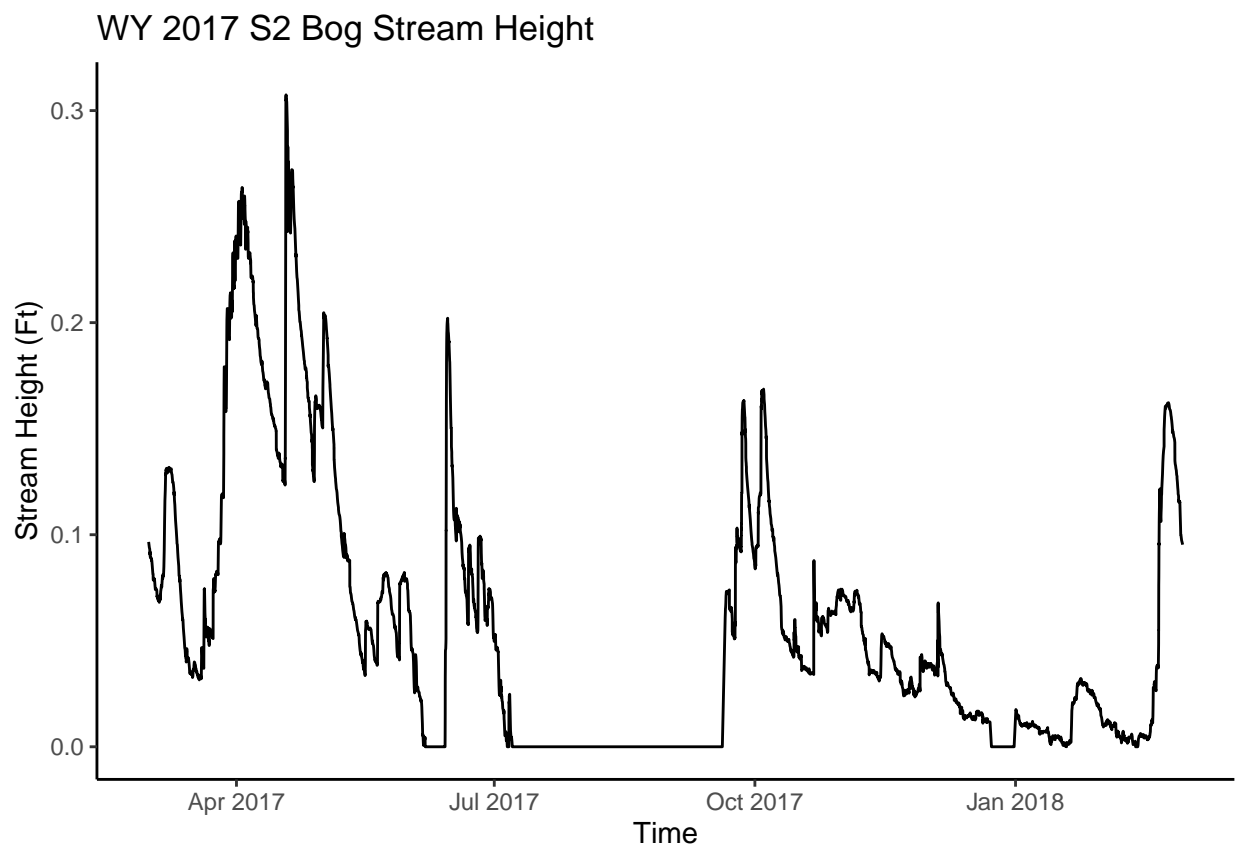
Water Year 2018

- Read in the WY 2018 data from the External-MEF_DATA Box folder
- Clean and manipulate the WY 2018 data into a tidy format
  - Renaming columns to use lower_snake_case and avoid special characters
  - create a new `date` column that follows the format YYYY-MM-DD (and does not include a timestamp)
- Change values of 0 to missing NA values during the following time periods:
  - December 1-3 clock stopped, no record
  - December 8-10 pin reached end of chart, no record
  - Dec 11-end, clock is barely running. No way to tell which days are what on the chart.
  - November 10-14 clock stopped, no record
  - November 25- December 4 clock stopped, no record Possibly the clock was moving super slowly for some of it?
  - For Oct 9-10 there was no rain for flow to increase
  - Oct 11-15 clock was stuck and the pin did not move forward in time
- Test that the manipulated data frame has the same number of rows as the original data frame
- Graph the WY 2018 data

```r
#read in the 2018 data
wy2018 <- read_excel(path = here(filepath, "WY2018.S2_Breakpoint.xlsx"),
                     skip =3)

#clean the 2018 data
wy2018_clean <- wy2018 %>%
  rename(datetime = "Date/time",
         stream_height_ft = "Stage.ft") %>%
  select(c("datetime", "stream_height_ft")) %>%
  clean_names() %>%
  mutate(date = format(as.POSIXct(datetime,format='%m/%d/%Y %H:%M:%S'),format='%Y-%m-%d'),
         date = as.POSIXct(date))

#create a function to specify custom time intervals
`%between%` <- function(x, interval) x >= interval[1] & x <= interval[2]

#assign October NA values
wy2018_clean$stream_height_ft[wy2018_clean$date %between% as.Date(c('2018-10-09', '2018-10-10'))] <- NA
wy2018_clean$stream_height_ft[wy2018_clean$date %between% as.Date(c('2018-10-11', '2018-10-15'))] <- NA

#assign November NA values
wy2018_clean$stream_height_ft[wy2018_clean$date %between% as.Date(c('2018-11-10', '2018-11-14'))] <- NA
wy2018_clean$stream_height_ft[wy2018_clean$date %between% as.Date(c('2018-11-25', '2018-12-04'))] <- NA

#assign December NA values
wy2018_clean$stream_height_ft[wy2018_clean$date %between% as.Date(c('2018-12-01', '2018-12-03'))] <- NA
wy2018_clean$stream_height_ft[wy2018_clean$date %between% as.Date(c('2018-12-08', '2018-12-10'))] <- NA
wy2018_clean$stream_height_ft[wy2018_clean$date %between% as.Date(c('2018-12-11', '2018-12-31'))] <- NA

#test the 2018 data
if(nrow(wy2018) != nrow(wy2018_clean)) stop("Check clean dataframe dimensions")
```
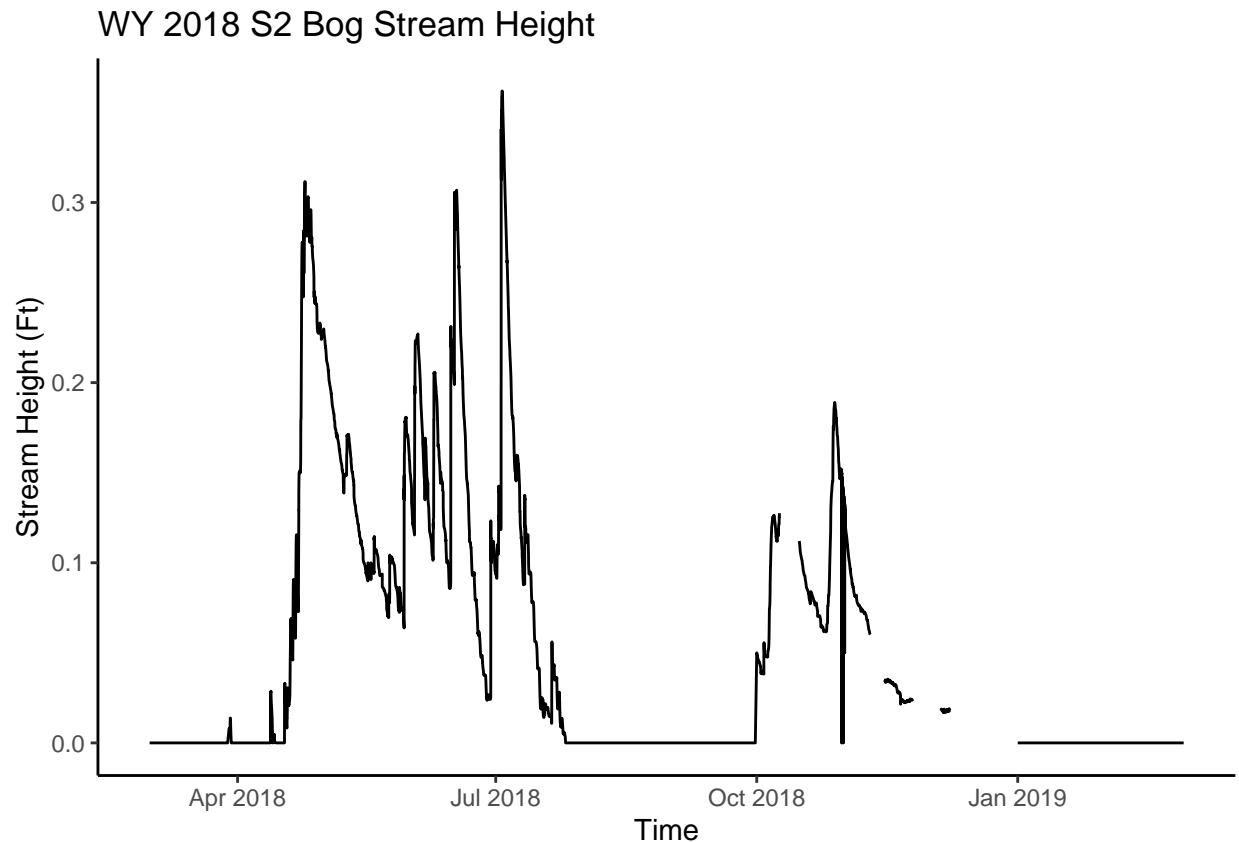
```
ggplot(data = wy2018_clean) +
  geom_line(aes(x = datetime, y = stream_height_ft)) +
  theme_classic() +
  labs(x = "Time",
       y = "Stream Height (Ft)",
       title = "WY 2018 S2 Bog Stream Height")
```

WY 2018 S2 Bog Stream Height

Water Year 2019

- Read in the WY 2019 data from the External-MEF_DATA Box folder
- Clean and manipulate the WY 2019 data into a tidy format

    - Rename columns to use lower_snake_case and avoid special characters
    - Create a new `date` column that follows the format YYYY-MM-DD (and does not include a timestamp)

- Change values of 0 to missing NA values during the following time periods:

    - April 1-2, Line and time is off

    - April 8-17. Line doesn't fluctuate though there should've been rain or if not rain, it should've fallen from how high the flow is

    - Proven in that the adjustment for 4/15 flow is great and S6 Weir chart response. Therefore, no record

    - April 17-22 discounted because pen level is off with no way to account for it

4

- – February: Clock errors, record compromised. Flow is below zero for month Clock restarted 2/21/2019. Measurements still below zero.
    - – January: Clock errors meant that time could not be used to record accurate measurements. Measurements tend to be around 0.001 and chart line is there or below zero. Physical measurements indicate flow until at least 1/15/2019
    - – March 14, Pen reverses and goes off the chart
    - – March 18th, Pen placed back on chart, still reversed (fixed 4/23)
    - – March 30-31 Pen stays in place even though there should be a response. Likely because it reached the bottom of the chart (pen was reversed)
    - – September: Chart did not run from 9/3-9/10
- Test that the manipulated data frame has the same number of rows as the original data frame
- Graph the WY 2019 data

```
#read in the 2019 data
wy2019 <- read_excel(path = here(filepath, "WY2019.S2_Breakpoint.xlsx"),
                     skip =3)

#clean the 2019 data
wy2019_clean <- wy2019 %>%
  rename(datetime = "Date/time",
         stream_height_ft = "Stage.ft") %>%
  select(c("datetime", "stream_height_ft")) %>%
  clean_names() %>%
  mutate(date = format(as.POSIXct(datetime,format='%m/%d/%Y %H:%M:%S'),format='%Y-%m-%d'),
         date = as.POSIXct(date))

#February NA values
wy2019_clean$stream_height_ft[wy2019_clean$date %between% as.Date(c('2019-04-01', '2019-04-02'))] <- NA

#March NA values
wy2019_clean$stream_height_ft[wy2019_clean$date == as.Date('2019-03-14')] <- NA
wy2019_clean$stream_height_ft[wy2019_clean$date == as.Date('2019-03-18')] <- NA
wy2019_clean$stream_height_ft[wy2019_clean$date %between% as.Date(c('2019-03-30', '2019-03-31'))] <- NA


#April NA values
wy2019_clean$stream_height_ft[wy2019_clean$date %between% as.Date(c('2019-04-01', '2019-04-02'))] <- NA
wy2019_clean$stream_height_ft[wy2019_clean$date %between% as.Date(c('2019-04-08', '2019-04-17'))] <- NA
wy2019_clean$stream_height_ft[wy2019_clean$date %between% as.Date(c('2019-04-17', '2019-04-22'))] <- NA

#September NA values
wy2019_clean$stream_height_ft[wy2019_clean$date %between% as.Date(c('2019-09-03', '2019-09-10'))] <- NA

#test the 2019 data
if(nrow(wy2019) != nrow(wy2019_clean)) stop("Check clean dataframe dimensions")

#plot
ggplot(data = wy2019_clean) +
  geom_line(aes(x = datetime, y = stream_height_ft)) +
  theme_classic() +
  labs(x = "Time",
```
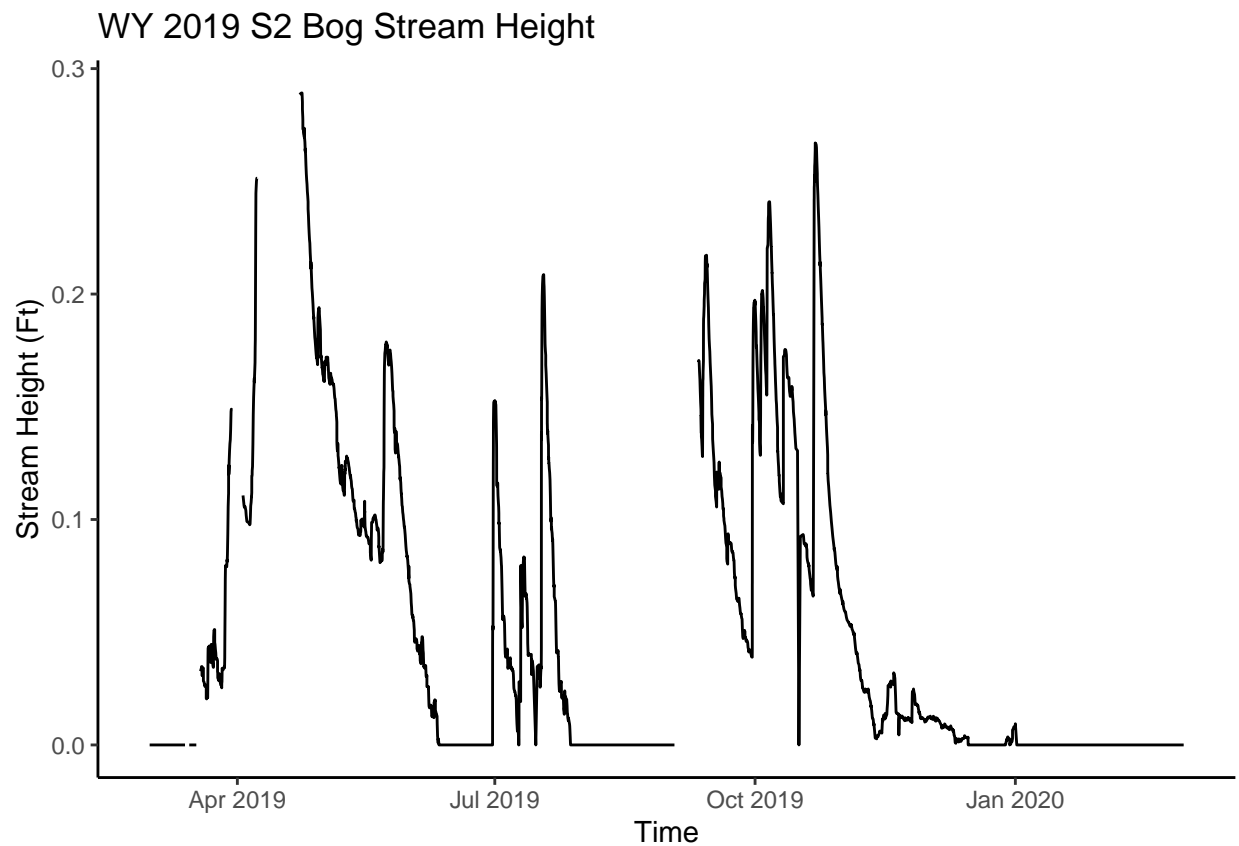
```
      y = "Stream Height (Ft)",
      title = "WY 2019 S2 Bog Stream Height")
```

WY 2019 S2 Bog Stream Height

All Water Years Combined

- Combine all 3 water years using `rbind()`

- Plot

```
all_streamflow <- rbind(wy2017_clean, wy2018_clean, wy2019_clean)

ggplot(data = all_streamflow) +
  geom_line(aes(x = datetime, y = stream_height_ft)) +
  theme_classic() +
  labs(x = "Time",
      y = "Stream Height (Ft)",
      title = "WY 2017 - 2019 S2 Bog Stream Height")
```

WY 2017 – 2019 S2 Bog Stream Height