USGS Coding Lab

Emma Reich

11/12/2020

Just how dry was the 2020 monsoon?

Recently NOAA published a press release demonstrating that 2020 was *both* the hottest *and* driest summer on record for Arizona. In this lab we will look at USGS NWIS stream gauge and precipitation data to investigate just how anomalous 2020 data are.

```
# Load packages
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------------------------------------- tidyvers
```

```
## v ggplot2 3.3.1     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

```
## Warning: package 'tibble' was built under R version 4.0.2
```

```
## Warning: package 'tidyr' was built under R version 4.0.2
```

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

```
library(dataRetrieval)
```

```
## Warning: package 'dataRetrieval' was built under R version 4.0.2
```

```
library(plotly)
```

```
## Warning: package 'plotly' was built under R version 4.0.2
```

```
##
## Attaching package: 'plotly'
```

```
## The following object is masked from 'package:ggplot2':
##
##     last_plot
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
## The following object is masked from 'package:graphics':
##
##     layout
```

1: Use the `readNWISstat` function to retrieve the following statewide data for 2015-2020: 1. Precipitation total (inches/week) 2. Streamflow (ft^3/s)

```r
# Download precip data
# Using daily values, since the weekly 00046 parameter is returning 0 rows
precip <- readNWISdata(stateCd="AZ", service="dv",
                       parameterCd="00045",
                       startDate="2015-01-01",endDate="2020-10-01")
```

```r
head(precip)
```

```
##   agency_cd  site_no   dateTime X_00045_00006 X_00045_00006_cd tz_cd
## 1      USGS 09424380 2019-04-18          0.00                P   UTC
## 2      USGS 09424380 2019-04-19          0.00                P   UTC
## 3      USGS 09424380 2019-04-20          0.00                P   UTC
## 4      USGS 09424380 2019-04-21          0.00                P   UTC
## 5      USGS 09424380 2019-04-22          0.06                P   UTC
## 6      USGS 09424380 2019-04-23          0.00                P   UTC
```

```r
# Download stream flow data
flow <- readNWISdata(stateCd="AZ", service="dv",
                       parameterCd="00060")
```

```r
head(flow)
```

```
##   agency_cd  site_no   dateTime X_00060_00003 X_00060_00003_cd X_00060_00001
## 2      USGS 09379025 2006-10-11          6.65                A            NA
## 3      USGS 09379050 2007-01-04          2.72                A            NA
## 4      USGS 09379180 2005-12-31          8.62                A            NA
## 5      USGS 09379200 2020-11-14          0.00                P            NA
## 6      USGS 09379910 2004-08-02      14900.00                A            NA
## 7      USGS 09380000 2020-11-14      11600.00                P            NA
##   X_00060_00001_cd X_.Data.prior.to.10.1.1992._00060_00003
## 2             <NA>                                      NA
## 3             <NA>                                      NA
## 4             <NA>                                      NA
## 5             <NA>                                      NA
## 6             <NA>                                      NA
## 7             <NA>                                      NA
##   X_.Data.prior.to.10.1.1992._00060_00003_cd
## 2                                       <NA>
## 3                                       <NA>
## 4                                       <NA>
## 5                                       <NA>
## 6                                       <NA>
## 7                                       <NA>
##   X_.Data.from.10.1.1992.Forward._00060_00003
## 2                                          NA
## 3                                          NA
## 4                                          NA
## 5                                          NA
## 6                                          NA
## 7                                          NA
##   X_.Data.from.10.1.1992.Forward._00060_00003_cd X_00060_00002 X_00060_00002_cd
## 2                                           <NA>            NA             <NA>
## 3                                           <NA>            NA             <NA>
## 4                                           <NA>            NA             <NA>
## 5                                           <NA>            NA             <NA>
## 6                                           <NA>            NA              <NA>
## 7                                           <NA>            NA              <NA>
##   X_00060_32400 X_00060_32400_cd X_00060_00006 X_00060_00006_cd tz_cd
## 2            NA             <NA>            NA             <NA>   UTC
## 3            NA             <NA>            NA             <NA>   UTC
## 4            NA             <NA>            NA             <NA>   UTC
## 5            NA             <NA>            NA             <NA>   UTC
## 6            NA             <NA>            NA             <NA>   UTC
## 7            NA             <NA>            NA             <NA>   UTC
```

2: Create two timeseries plots (one for preciptitation, one for streamflow), where color is a function of year (e.g. the x axis is month, the y axis is precipitation, legend shows color by year).

```
# Format Date

  # precip
precip$date <- as.Date(precip$dateTime,"%Y-%m-%d")
```

```
## Warning in as.POSIXlt.POSIXct(x, tz = tz): unknown timezone '%Y-%m-%d'
```

```
precip$year <- as.numeric(format(precip$date, format = "%Y"))
precip$month <- as.numeric(format(precip$date, format = "%m"))
precip$day <- as.numeric(format(precip$date, format = "%d"))

  #flow
flow$date <- as.Date(flow$dateTime,"%Y-%m-%d")
```

```
## Warning in as.POSIXlt.POSIXct(x, tz = tz): unknown timezone '%Y-%m-%d'
```

```
flow$year <- as.numeric(format(flow$date, format = "%Y"))
flow$month <- as.numeric(format(flow$date, format = "%m"))
flow$day <- as.numeric(format(flow$date, format = "%d"))
  #filter flow for 2015-2020
flow <- flow %>%
  filter(year > 2014)
```

```
# Group precip by month

precip_month <- precip %>%
  group_by(month, year) %>%
  summarise(P = mean(X_00045_00006))
```

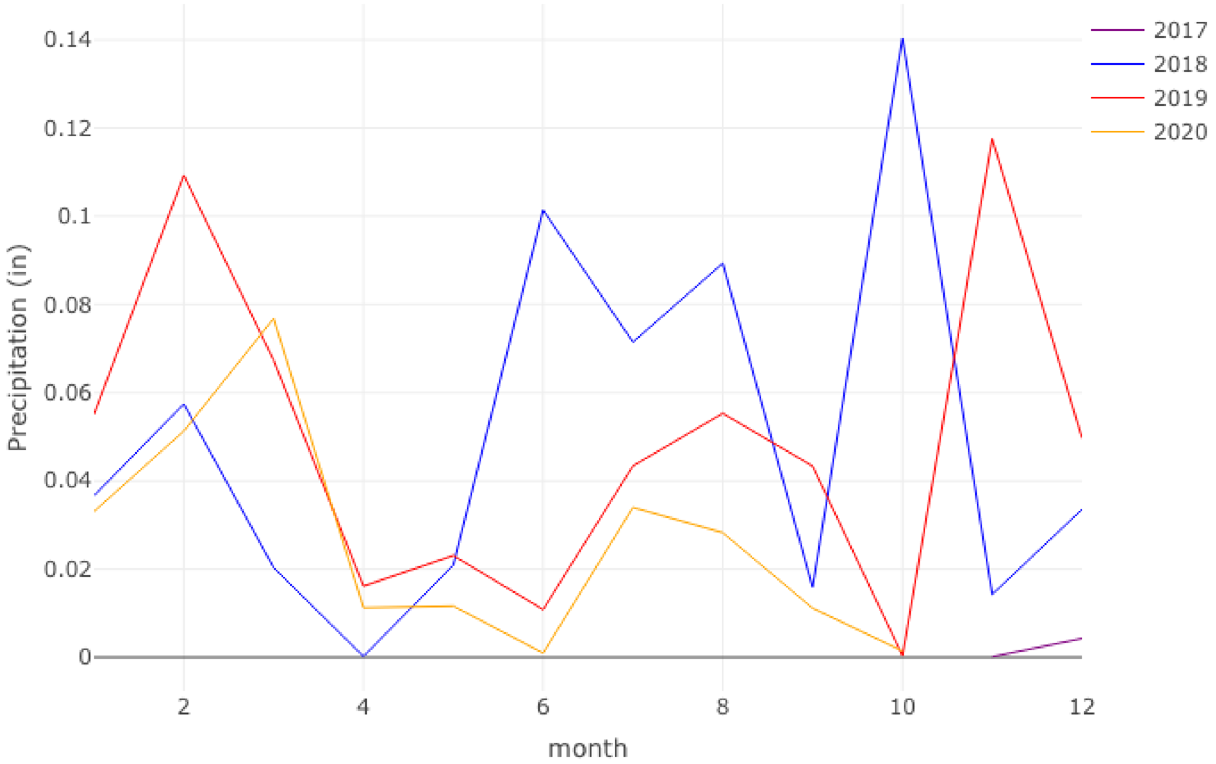
```
## `summarise()` regrouping output by 'month' (override with `.groups` argument)
```

```
precip_month$year <- as.factor(precip_month$year)
precip_month$month <- as.factor(precip_month$month)
```

```
# Group precip by year

precip_year <- precip_month %>%
  spread(year, P) %>%
  rename(P2017 = `2017`) %>%
  rename(P2018 = `2018`) %>%
  rename(P2019 = `2019`) %>%
  rename(P2020 = `2020`)
```

```
# AZ precip timeseries plotly
```

```r
p <- plot_ly()
p <- p %>%  add_trace(
    data = precip_year,
    x = ~ month,
    y = ~ P2017,
    name = '2017',
    line=list(color="purple", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%  add_trace(
    data = precip_year,
    x = ~ month,
    y = ~ P2018,
    name = '2018',
    line=list(color="blue", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%  add_trace(
    data = precip_year,
    x = ~ month,
    y = ~ P2019,
    name = '2019',
    line=list(color="red", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%  add_trace(
    data = precip_year,
    x = ~ month,
    y = ~ P2020,
    name = '2020',
    line=list(color="orange", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%
  layout(
     yaxis=list(title='Precipitation (in)')
  )
p <- p %>% layout(legend = list(x = 1, y = 1))
p
```



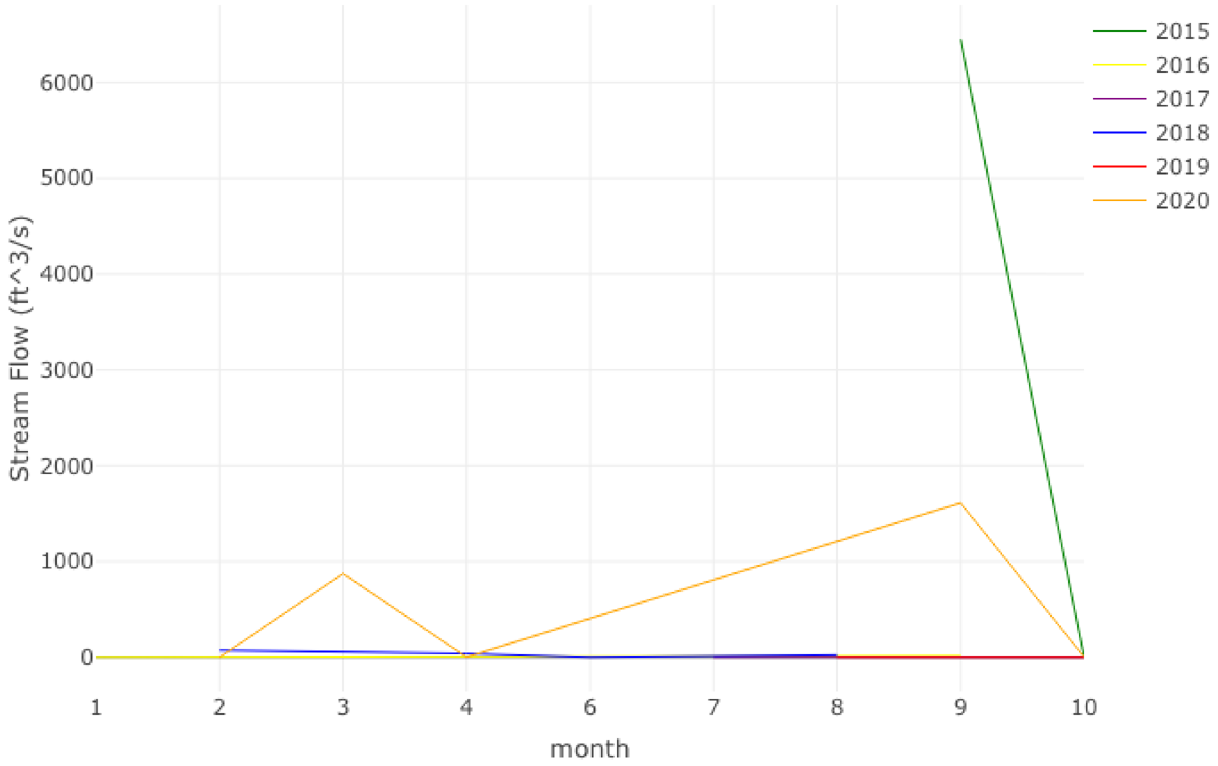
```r
# Group stream flow by month

flow_month <- flow %>%
  group_by(month, year) %>%
  summarise(stream_flow = mean(X_00060_00003, na.rm = TRUE))
```

```
## `summarise()` regrouping output by 'month' (override with `.groups` argument)
```

```r
flow_month$year <- as.factor(flow_month$year)
flow_month$month <- as.factor(flow_month$month)

flow_month <- flow_month[complete.cases(flow_month[ , 3]),]
```

```r
# Group stream flow by year

flow_year <- flow_month %>%
  spread(year, stream_flow) %>%
  rename(P2015 = `2015`) %>%
  rename(P2016 = `2016`) %>%
  rename(P2017 = `2017`) %>%
  rename(P2018 = `2018`) %>%
  rename(P2019 = `2019`) %>%
  rename(P2020 = `2020`)
```

```r
# AZ stream flow timeseries plotly

p <- plot_ly()
p <- p %>%  add_trace(
    data = flow_year,
    x = ~ month,
    y = ~ P2015,
    name = '2015',
    line=list(color="green", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%  add_trace(
    data = flow_year,
    x = ~ month,
    y = ~ P2016,
    name = '2016',
    line=list(color="yellow", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%  add_trace(
    data = flow_year,
    x = ~ month,
    y = ~ P2017,
    name = '2017',
    line=list(color="purple", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%  add_trace(
    data = flow_year,
    x = ~ month,
    y = ~ P2018,
    name = '2018',
    line=list(color="blue", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%  add_trace(
    data = flow_year,
    x = ~ month,
    y = ~ P2019,
    name = '2019',
    line=list(color="red", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%  add_trace(
    data = flow_year,
    x = ~ month,
    y = ~ P2020,
    name = '2020',
    line=list(color="orange", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%
  layout(
     yaxis=list(title='Stream Flow (ft^3/s)')
  )
p <- p %>% layout(legend = list(x = 1, y = 1))
p
```



3. Calculate the monthly mean precipitation and streamflow from 2015-2019, and use that mean to calculate a 2020 anomaly timeseries. Create two new plots (like #2 above) with the 2015-2019 mean as a thick black line, and 2020 anomaly as a thin red line.

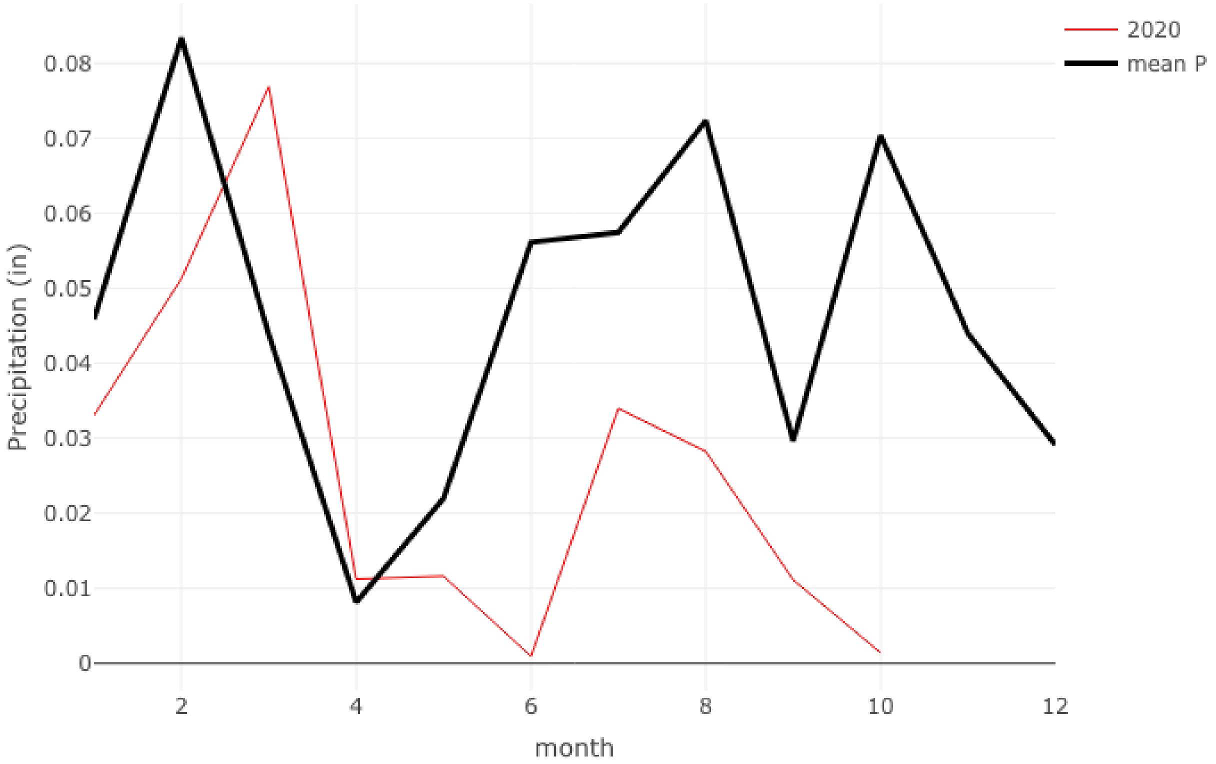
```
# precip monthly mean before 2020
precip_below20 <- precip_month %>%
  filter(year != 2020) %>%
  group_by(month) %>%
  summarise(meanP = mean(P))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

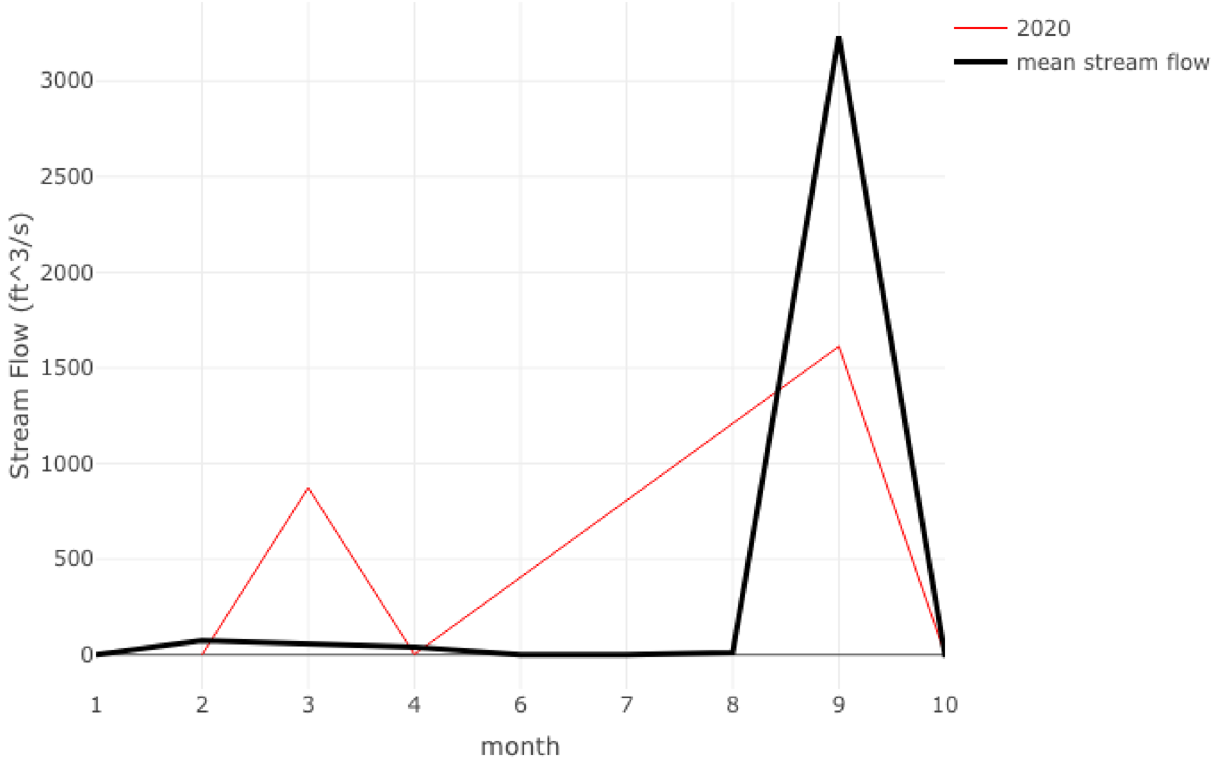
```
# streamflow monthly mean before 2020
flow_below20 <- flow_month %>%
  filter(year != 2020) %>%
  group_by(month) %>%
  summarise(meanflow = mean(stream_flow))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
# AZ precip timeseries plotly

p <- plot_ly()
p <- p %>%  add_trace(
    data = precip_year,
    x = ~ month,
    y = ~ P2020,
    name = '2020',
    line=list(color="red", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%  add_trace(
    data = precip_below20,
    x = ~ month,
    y = ~ meanP,
    name = 'mean P',
    line=list(color="black", width=3),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')

p <- p %>%
  layout(
     yaxis=list(title='Precipitation (in)')
  )
p <- p %>% layout(legend = list(x = 1, y = 1))
p
```



```
# AZ precip timeseries plotly

p <- plot_ly()
p <- p %>%  add_trace(
    data = flow_year,
    x = ~ month,
    y = ~ P2020,
    name = '2020',
    line=list(color="red", width=1),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')
p <- p %>%  add_trace(
    data = flow_below20,
    x = ~ month,
    y = ~ meanflow,
    name = 'mean stream flow',
    line=list(color="black", width=3),
    showlegend = TRUE,
    type = 'scatter',
    mode = 'line')

p <- p %>%
  layout(
     yaxis=list(title='Stream Flow (ft^3/s)')
  )
p <- p %>% layout(legend = list(x = 1, y = 1))
p
```


mean stream flow
Stream Flow (ft^3/s)
month