NSci233 Lesson Plan Summary - Precipitation

Incorporates Lesson Plans 3 - 5 for Weeks 6 - 10

Charles Jason Tinant

5/02/2018

Overview

Precipitation is the single most important variable driving hydrologic processes and events, because it replenishes soil moisture, lakes and streams, and groundwater. Precipitation is any form of solid or liquid water the falls from the atmosphere to the earth's surface. Precipitation can be thought about in terms of events, as averages over a period of time, or as deviations from the normal or "adequate" amount. In this section we are going to discuss the role of precipition in climate, what causes precipiation, precipitation events, and rainfall frequency.

Student Learning Outcomes (SLOs) & Lesson Goals:

- *SLO1: Students will demonstrate proficiency in the components of the hydrologic cycle: precipitation, infiltration, runoff, and evapotranspiration* - Students will be able to generally describe what causes precipiation, precipitation events, and rainfall frequency.
- *SLO2: Students are able to analyze components of the hydrologic with computer software* - Students will be generally able to read data into R, and perform some analysis using packages. Students will will be able to read arguments of a package from R Documentation.

Reading Assignments

- Scan Chapter 1 of **Exploratory Data Analysis** https://leanpub.com/exdata
- Read sections 2.1 - 2.4 of the text
- Article on Cape Town, South Africa water crisis
- Data Visualization with ggplot2

Homework

You are welcome to turn in homework in Rmd, pdf, html or Word. Upload the file into a folder with your name in Google Drive.

1. Vocabulary - Summarize the terms: weather, climate, frontal precipitation, convection, orographic, stratosphere, troposphere, jet stream

2. Use the following link: Koeppen Climate to download a map of Koeppen climates for the United States. What is the climate for Western South Dakota? What is the climate for Eastern South Dakota.

3. Use the following link: Drought Monitor - What is the drought intensity for this week for the county you live in? If your county is in drought, then is the drought short-term or long-term? What is the drought summary for the High Plains? What is the the drought summary for the West?

4. Calculate the effective depth in mm and volume of water in the last snowstorm in cubic meters for two watersheds. Assume the snow has a 10% moisture content, 8 inches of snow fell over a 500 square mile watershed, and 6 inches of snow over a 300 square mile watershed. Recall: 1 inch = 22.54 mm, 1 square mile = 2.59 square kilometers. Feel free to use either R or another approach.

5. Look up the average monthly precipitation totals for the county you live in. The Drought Monitor may be a helpful resource. You can use a table to show the precipitation totals using the method discussed on the lower left part of page 2 of the R Markdown cheatsheet that you can download from https://www.rstudio.com/resources/cheatsheets/

6. Pick three short-term and three long-term intensities for Table 2.5 and look up on a map the locations they are recorded in. Based on the geographic location of these areas, what is the process causing the precipitation?

7. Take a look at the graphs for the Cottonwood station below. Identify: 1) the greatest 24-hour event, 2) the wettest year, 3) the dryest year, 4) the wettest decade, 5) the dryest decade, 6) whether the region is recieving more or less precipitation than 100 years ago.

Lecture Notes with R-code

```
# one of the biggest issues we have been having the first four weeks
# of class is setting the workspace.  I have taught the most basic
# approach to setting a workspace.  However, given the challenges many
# are experiencing, I want to try a different approach using the 'here'()'
# package.

# download 'here' package -the install function below is commented out.
# install.packages("here")
```

```
# If you need to you should 'install.packages("here")'
library("tidyverse") # load the tidyverse package
library("lubridate")
library("here")
library("janitor")
here::here() # this checks where R thinks the working directory is...
```

```
## [1] "/Users/cjtinant/Google Drive/NSci253-master/NSci253_Notes"
```

Weather, Climate, and Water Resources

Weather is the result of the day-to-day movement of moist air masses from thier sources and thier associated effects on precipitation and temperature. **Climate** is a long-term average of the weather. Intermediate to weather and climate is a moving window of recent weather history and intermediate term forecasting, and how these relate to the climate. Precipitation is an important part of both weather and climate. Too much precipitation on a particular day, *weather,* causes schools to close. Too little precipitation in a given season results in a drought. Too much water extraction for the climate generally, or during a drought results in a loss of water resources as (Cape Town is currently experiencing)[https://www.vox.com/world/2018/2/9/16964416/cape-town-water-crisis-day-zero-south-africa].

A lot of systems have been used to describe climate. The **Koeppen Climate Classification System** is the most widely adopted approach, Köppen's classification is based on a subdivision of terrestrial climates into five major types, which are represented by the capital letters A, B, C, D, and E. Except for type B, each of these climate types is defined by temperature criteria. **Type B designates climates in which the controlling factor on vegetation is dryness** (rather than coldness). **Aridity** is not a matter of precipitation alone but is defined by the relationship between the precipitation input to the soil in which the plants grow and the evaporative losses. Since evaporation is difficult to evaluate and is not a conventional measurement at meteorological stations, Köppen was forced to substitute a formula that identifies aridity in terms of a temperature-precipitation index (that is, evaporation is assumed to be controlled by temperature). Dry climates are divided into arid (BW) and semiarid (BS) subtypes, and each may be differentiated further by adding a third code, h for warm and k for cold.

Temperature defines the other four major climate types. These are subdivided, with additional letters again used to designate the various subtypes. Type A climates (the warmest) are differentiated on the basis of the seasonality of precipitation: Af (no dry season), Am (short dry season), or Aw (winter dry season). Type E climates (the coldest) are conventionally separated into tundra (ET) and snow/ice climates (EF). The mid-latitude C and D climates are given a second letter, f (no dry season), w (winter dry), or s (summer dry), and a third symbol (a, b, c, or d [the last subclass exists only for D climates]), indicating the warmth of the summer or the coldness of the winter. Although Köppen's classification did not consider the uniqueness of highland climate regions, the highland climate category, or H climate, is sometimes added to climate classification systems to account for elevations above 1,500 metres (about 4,900 feet; source: https://www.britannica.com/science/Koppen-climate-classification). You can check out the changing climate at http://koeppen-geiger.vu-wien.ac.at/.

For land and water resources management, one important tool is the **Drought Monitor** (http://droughtmonitor.unl.edu/). The Drought Monitor describes weather trends, and relates them to long-range weather forecasts and the climate.

Causes of Precipitation

The precipitation process begins when water droplets condense in the atmosphere around dust, smoke, or ice particles. As the droplet size increases it reaches a size at which gravity begins to pull them down. Condensensation is caused by cooling, which is caused by the lifting of air masses. There are three processes that lift air masses: * Frontal air masses - frontal precipitation, * Convection - thunderstorms, * Mountains - orographic precipitation.

** Frontal precipitation ** occurs when warm moist air masses meet colder, heavier air masses. It is the dominant precipitation type in the northern United States, and other mid-latitude areas, and is caused by differences in the temperature and pressure of air packets. Weather fronts, the places where air packets meet, are associated with mid-latitude cyclones (low pressure) moving from west to east. Pressure refers to **barometric pressure**, the force exerted by the weight of air above a surface. At sea level, that average pressure exerted by air is 1 atm = 29.01.325 kPa. This is also equal to 29.92 inches of mercury (Hg) or 406.6 inches of liquid water. In the Northern Hemisphere, air circulates counterclockwise around low-pressure areas, so warm air is most often found in the south and east of low-pressure areas. The **troposphere**, the 4 to 11 mile thick layer of atmosphere where packets of air that weather occur, is steered by the **jet stream** in the **stratosphere**.

** Convection** occurs as moist air near the surface heated by solar radiation becomes lighter and rises up. The rising air and latent heat released by water changing phase causes atmospheric instability, and possibly late afternoon thundershowers. Convective precipitation also can result as air flows into the center of a low pressure system and chills until the moisture in the air reaches saturation. Hurricanes are a large-scale convective storm cell that forms up in the tropics and generally move northerly.

Orographic effects are caused by moist air rising over mountains. These effects cause the high amounts of snow in mountains, as well as the rain shadow east of the mountains.

Precipitation Events

Storms can be described by **type**, **amount**, and **duration**. * type - hail, rain, snow, sleet; * amount - depth of precipitation (as rain or as snow); * duration - length of time the storm is occurring.

** intensity** is the *average rate* of precipitation found by $intensity(mm/hr) = \frac{depth(mm)}{duration(hr)}$. Another useful thing to consider is an event's **return period**, the average number of years between storms of a particular magnitude. *We will discuss return period in more detail next week.*

We usually think about precipitation as a **depth** rather than a volume. We can use the equation $Vol = depth * area$. To find out the average volume of water from a **depth** of water that falls on each country each year, we need to multiply the land area by the precipitation depth.

```
# We are going to calculate a volume from a given depth of water
# -> we need to pay close attention to units
# ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
# Review: we can quickly add data together in R rather than read it in
# from a .csv file using the 'concatinate' function, c().  Than using the
# 'tidyverse', we can easily bind rows and columns together.

# make a vector of strings with continent names and a vector of area values
continent <- c("Africa", "North America", "South America", "Asia",
               "Europe", "Oceania")
depth_in <- c(27, 26.4, 64.9, 28.6, 28.9, 29.0)
area_thos_km2 <- c(3022, 2471, 1784, 4458, 1018, 853)

# turn the vectors into data.frames & remove the old data
continent <- as.data.frame(continent)
depth_in <- as.data.frame(depth_in)
area_thos_km2 <- as.data.frame(area_thos_km2)

# bind the two vectors together and then make it into a data-frame
precip <- bind_cols(continent, depth_in)
precip <- bind_cols(precip, area_thos_km2)
rm(continent, depth_in, area_thos_km2)
```

```
# 1 inch equals 22.54 mm
# 1,000,000 mm (10^6 mm) = 1 km
```

```r
# Use the 'mutate' function
# You can use the '%>%' called a pipe to connect parts together.

# Use the 'mutate' function
# You can use the '%>%' called a pipe to connect parts together.
# In class answer
precip <- precip %>%
  # 1 inch = 22.54 mm
  mutate(depth_mm = depth_in * 22.54) %>%
  # 1000 m = 1 km
  # 1000 m = 1 km
  # 1,000,000 m2 = 1 km2
  mutate(area_thos_m2 = area_thos_km2 * 10^3 * 10^3) %>%
  # 1000 mm = 1 m
  mutate(depth_m = depth_mm * 10^-3) %>%
  # 1000 m = thos_m
  # 1000 m = thos_m
  # 1000 * 1000 m2 =  thos_m2
  mutate(area_m2 = area_thos_m2 * 10^6) %>%
  mutate(vol_m3 = area_m2 * depth_m)

# here is the answer
precip <- precip %>%
  mutate(depth_mm = 22.54 * depth_in) %>%
  mutate(precip_km3 = area_thos_km2 * 10^3 * depth_mm *10^-6) %>%
  mutate(precip_m3 = precip_km3 * 10^9)
print(precip)
```

```
##       continent depth_in area_thos_km2 depth_mm area_thos_m2  depth_m
## 1        Africa     27.0          3022  608.580    3.022e+09 0.608580
## 2 North America     26.4          2471  595.056    2.471e+09 0.595056
## 3 South America     64.9          1784 1462.846    1.784e+09 1.462846
## 4          Asia     28.6          4458  644.644    4.458e+09 0.644644
## 5        Europe     28.9          1018  651.406    1.018e+09 0.651406
## 6       Oceania     29.0           853  653.660    8.530e+08 0.653660
##     area_m2       vol_m3 precip_km3    precip_m3
## 1 3.022e+15 1.839129e+15  1839.1288 1.839129e+12
## 2 2.471e+15 1.470383e+15  1470.3834 1.470383e+12
## 3 1.784e+15 2.609717e+15  2609.7173 2.609717e+12
## 4 4.458e+15 2.873823e+15  2873.8230 2.873823e+12
## 5 1.018e+15 6.631313e+14   663.1313 6.631313e+11
## 6 8.530e+14 5.575720e+14   557.5720 5.575720e+11
```

```
rm(precip)
```

Geographical and seasonal variations

Precipitation varies by season and by location. Average annual precipitation in the United States is 30 inches (75 cm), but there is great spatial variation across the country by location and by season (Figures 2.4 - 2.6 in text).

```
# I downloaded station data for Cottonwood station, SD from
# NOAA's National Centers for Environmental Information (NCEI)
# https://www.ncdc.noaa.gov/cdo-web/
# the column descriptions are below.  The depths are in metric (mm)

prcp_raw <- read_csv("1228041_cottonwood_sta.csv") %>%
  # drop the columns we dont need
  select(-(4:12)) %>%
  select(-(6))

precip <- prcp_raw %>%
  # clean the names
  clean_names() %>%
  # make a column of years, months, and days
  mutate(year = year(date)) %>%
  mutate(month = month(date)) %>%
  mutate(day = day(date)) %>%
  # change the mm to inches
  mutate(prcp_in = prcp/22.54) %>%
  # drop 2018 from the data
  filter(year != "2018")

precip_event <- precip %>%
  filter(prcp != 0)

# this is to check on NA values
prcp_na <- precip %>%
  filter(is.na(prcp))

# this is to check on the dryest year
prcp_test <- precip %>%
  filter(year == "1961")

# make an annual precipitation
prcp_ann <- precip %>%
# make a group of the years so that we can see how much prcp per year
```

```r
  group_by(year) %>%
  summarize(prcp = sum(prcp, na.rm = TRUE),
            prcp_in = sum(prcp_in, na.rm = TRUE)) %>%
  arrange(prcp) %>%
  # add back in the month and days (helps with the stuff below)
  mutate(month = 1) %>%
  mutate(day = 1) %>%
  mutate(date = ymd(paste(year, month, day, sep="-"))) %>%
  select(date, year, month, day, prcp, prcp_in)

# Here is what the columns from the raw data mean:
#EVAP - Evaporation of water from evaporation pan
#DAPR - Number of days included in the multiday precipitation total (MDPR)
#DAEV - Number of days included in the multiday evaporation total (MDEV)
#SNOW - Snowfall
#PRCP - Precipitation
#MNPN - Daily minimum temperature of water in an evaporation pan
#MDEV - Multiday evaporation total (use with DAEV)
#MXPN - Daily maximum temperature of water in an evaporation pan
#MDPR - Multiday precipitation total (use with DAPR and DWPR, if available)
#MDSF - Multiday snowfall total
#SNWD - Snow depth
#DASF - Number of days included in the multiday snow fall total (MDSF)
```

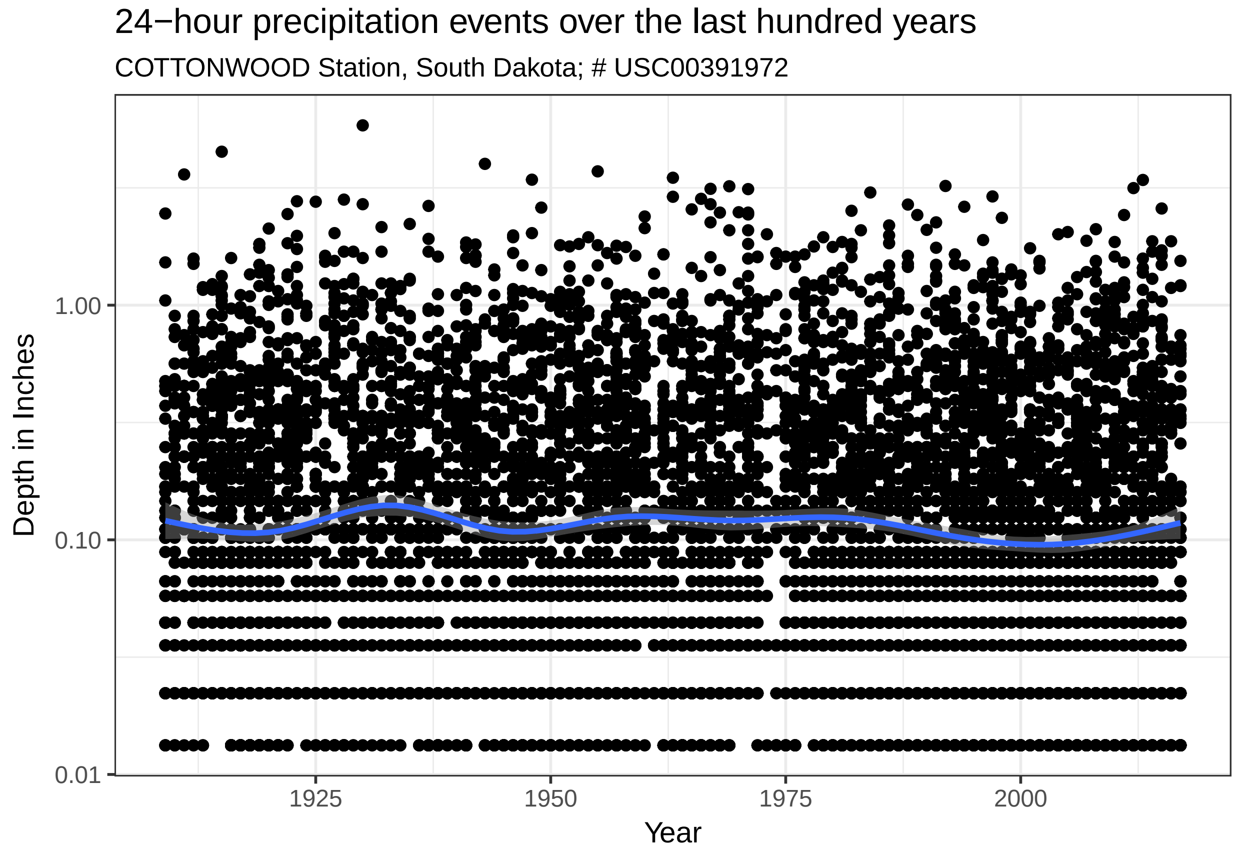
Averaging precipitation over space and time

Precipitation is measured at points often using tipping rainfall buckets, or by pillows to measure the snowfall mass. These point measurements can be averaged over an area such as a watershed using different approaches such as Theissen polygons and contouring that can be automated in ArcGIS or R.

The averaging period for rainfall depth and intensity can be important, depending on the question of interest. Shorter averaging periods (15 minute data or shorter) can show differences in intensity over a storm. Oftentimes, the magnitude of 24-hour rainfall events is used for hydrologic studies. However, daily, monthly, or annual data is useful for most hydrologic budget studies depending on what you are interested in

```r
# make a plot of the daily precip values
ggplot(precip_event, aes(x = year, y = prcp_in)) +
  geom_point() +
  geom_smooth() +
  theme_bw() +
  scale_y_log10() +
  labs(title = "24-hour precipitation events over the last hundred years", subtitle = "C
```

```
       x = "Year", y = "Depth in Inches")
```



Historic time trends

Climate tends to have distinct wetter and drier periods (Figure 2.7), with periods of 3-7, 15-20, 100 or so recurrence intervals having been defined. Different approaches to showing variability in the region influenced by the Great Lakes using trend lines, running averages to smooth the data, and standard deviations.

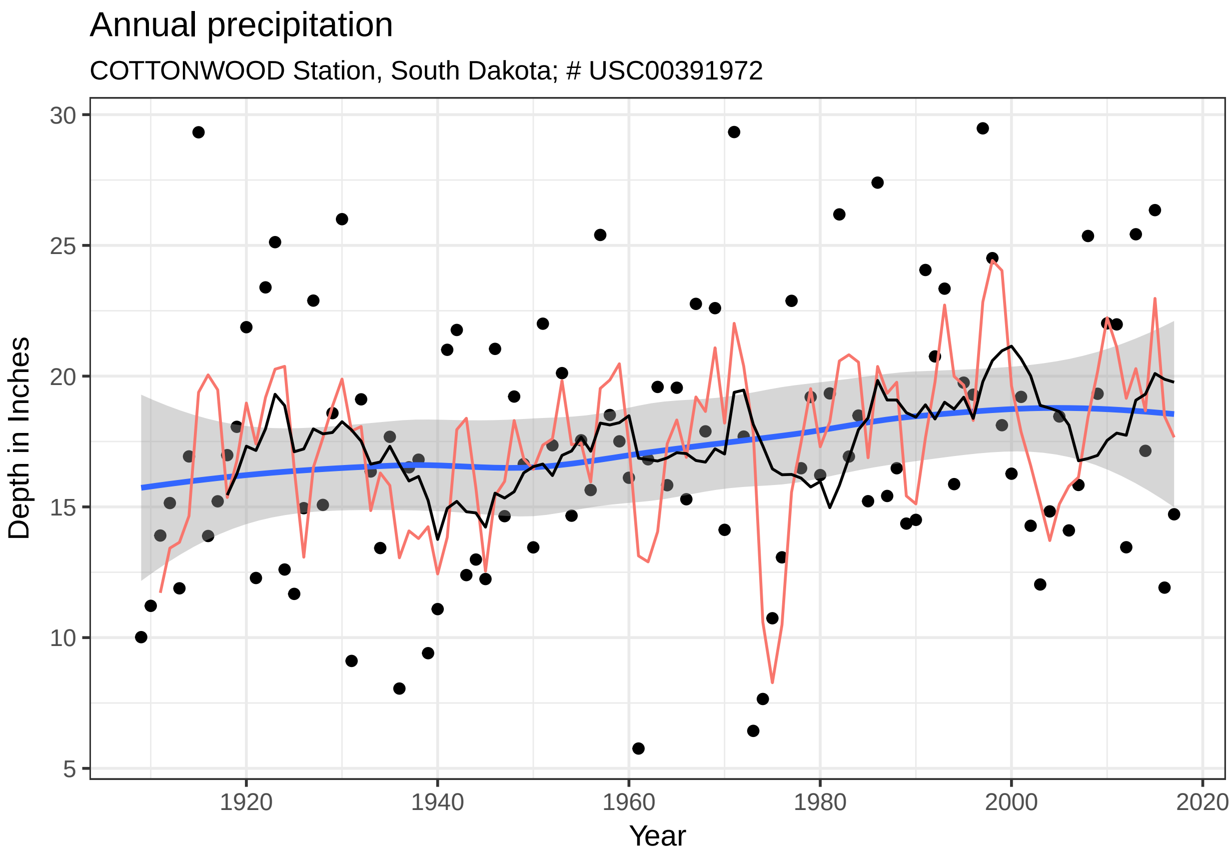
```
# ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
# add the tidyquant package to create a running mean of the years
# The package 'tidyquant' is for integrating
# quantitative financial analysis tools with the tidyverse.
# Specifically, we want to create a rolling mean like in figure 2.7.
# This is the article I used to create the rolling mean.
# http://www.business-science.io/timeseries-analysis/2017/07/23/tidy-timeseries-analys
# ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
#install.packages("tidyquant")
library(tidyquant)

# We want to calculate a rolling mean to look at the three year and
```

```
# 10 year wetter and drier cycles.
# The code was adapted from the article above...  Googling the thing
# you want to do and seeing how others have done it is a great way
# to learn new things.

prcp_ann <- prcp_ann %>%
    tq_mutate(
        # tq_mutate arguments to select the data and apply a rolling mean
        select     = prcp_in,
        mutate_fun = rollapply,
        # rollapply arguements to make a 3-yr window for the mean
        width      = 3,
        align      = "right",
        FUN        = mean,
        # arguments about calculating the mean;
        # 'na.rm = TRUE' means remove the NA values
        na.rm      = TRUE
    ) %>%
    tq_mutate(
        # tq_mutate args
        select     = prcp_in,
        mutate_fun = rollapply,
        # rollapply args (10-year rolling average)
        width      = 10,
        align      = "right",
        FUN        = mean,
        # mean args
        na.rm      = TRUE
    ) %>%
  # rename the columns made above
  rename(mean_3yr = prcp_in.1) %>%
  rename(mean_10yr = prcp_in.2)

# make a plot of the annual precip values
ggplot(prcp_ann, aes(x = date, y = prcp_in)) +
  geom_point() +
  geom_smooth() +
  geom_line(aes(y = mean_3yr, color = "red")) +
  geom_line(aes(y = mean_10yr)) +
  theme_bw() +
  labs(title = "Annual precipitation",
         subtitle = "COTTONWOOD Station, South Dakota; # USC00391972",
       x = "Year", y = "Depth in Inches") +
  theme(legend.position = "none")
```



Rainfall freqency distributions

Hydrologists use estimates of the probability of rainfall events occurring to design hydrologic structures. The main parameters used in estimation are *duration*, *intensity*, and *frequency*. **Frequency** means the probability of an event of a particular magnitude or greater occurring, usually in a given year. **Return interval**, also called **return period**, is the average period of time in years that an event of some magnitude or greater to occur. Some examples for design are for agricultural drainage (10-year return interval), culverts (20 year recurrence interval), placement of houses near a stream (100-year return interval). Return interval and frequency are two ways of communicating how often should we expect a flood (or drought) of a particular magnitude to return. They two parameters are related by the formula

$$Frequency, F = \frac{100}{return - interval, T}$$

.

```
# make a vector of return intervals and turn it into a dataframe
T_return_yr <- c(2, 5, 10, 20, 50, 100, 500)
T_return_yr <- as.data.frame(T_return_yr)

freq_return_table <- T_return_yr %>%
  mutate(F_prob_percent = 100/T_return_yr)
```

```r
rm(T_return_yr)
print(freq_return_table)
```

```
##   T_return_yr F_prob_percent
## 1           2           50.0
## 2           5           20.0
## 3          10           10.0
## 4          20            5.0
## 5          50            2.0
## 6         100            1.0
## 7         500            0.2
```

The approach above works great if you have all of the information in a population. However, we almost never have this, so we use a statistical method to plot the observations of rainfall or floods or drought and estimate how rare or common is an event of a given magnitude. Most of the methods mentioned in the text don't work great for hydrologic data (see Cunnane 1977 for a discussion of this), but the Hazen Method works pretty good.

The Hazen method uses the following plotting position to find the percent probability, F_a of some event n occurring in a given year is

$$F_a = \frac{100(2n-1)}{2y}$$

where n is the event from littlest to biggest, y is the total number of events.

The 24-hour depth - duration - frequency (DDF) plot for Cottonwood Station is shown below. If we were investigating this more in detail, I would take additional time to investigate the low precipitation years, to ensure there wasn't an error with the data. I marked the 50% probability of occurrance on the graph. The graph reads that we should expect 16 or more inches of rain half of the time (odds = 1:1), so the recurrence interval is 2-years. Note, that (if the data are correct) there seems to be a climate (something) that is leading to an upper-interval for precipitation in very wet years.

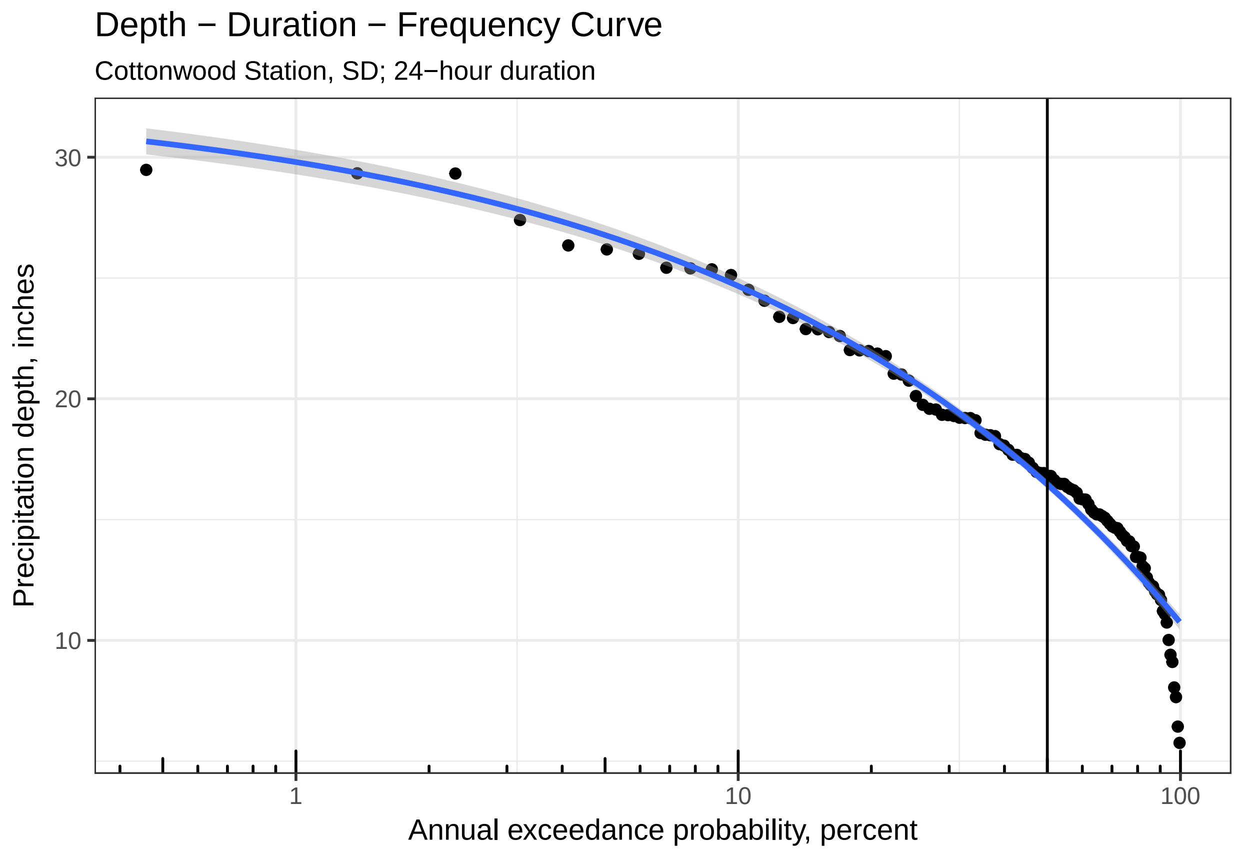
```r
# answers to the question above are: a) 14-inches;
#   b) 20% or 5-year recurrence (odds = 1:4)

# we can create the variables we need to calculate the probability
# of wet years in Cottonwood station

test <- prcp_ann %>%
  # arrange prcp_in from biggest to littlest
  arrange(desc(prcp_in)) %>%
  # there are many ways to do things, but this is one way to get
  # observations from 1 to 109
  rowid_to_column(var = "n") %>%
  # this makes the max observation
```

```r
  mutate(y = n()) %>%
  # this calculates the probability of the event or greater occurring
  mutate(Fa_prob = (100 * ((2 * n) - 1))/(2 * y)) %>%
  mutate(T_return_yr = 100 / Fa_prob)

# note that we are using a bit of trick here to get the number of
# observations, because day is the same all the way down.
ggplot(test, aes(x = Fa_prob, y = prcp_in)) +
  # plot the x, y points
  geom_point() +
  # make a smooth line through the points using a linear model
  geom_smooth(method = "lm", formula = y ~ exp(x))+
  # make the x-axis a log 10 axis & make log ticks along the bottom
  scale_x_log10() +
 # scale_y_log10() +
  annotation_logticks(base = 10, sides = "b") +
  # black and white theme
  theme_bw() +
  # labels for the graph
  labs(title = "Depth - Duration - Frequency Curve",
       subtitle = "Cottonwood Station, SD; 24-hour duration",
       x = "Annual exceedance probability, percent",
       y = "Precipitation depth, inches") +
  # make a vertical line at the x = 50 intercept
  geom_vline(xintercept = 50)
```


Depth − Duration − Frequency Curve
Cottonwood Station, SD; 24−hour duration
Precipitation depth, inches
30
20
10
1
10
100
Annual exceedance probability, percent