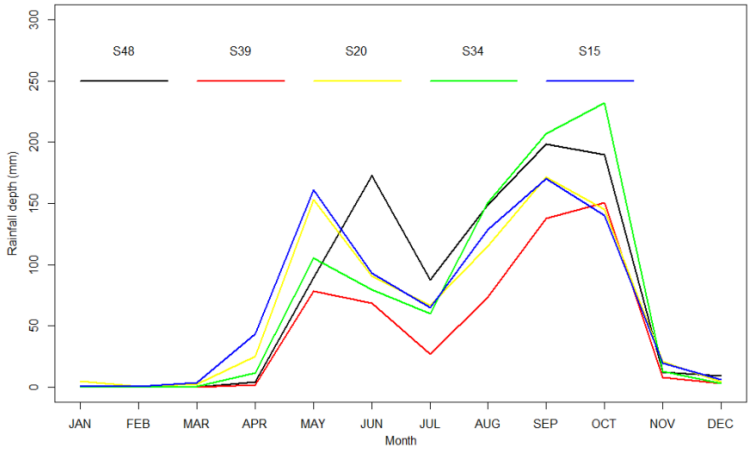
Statistics

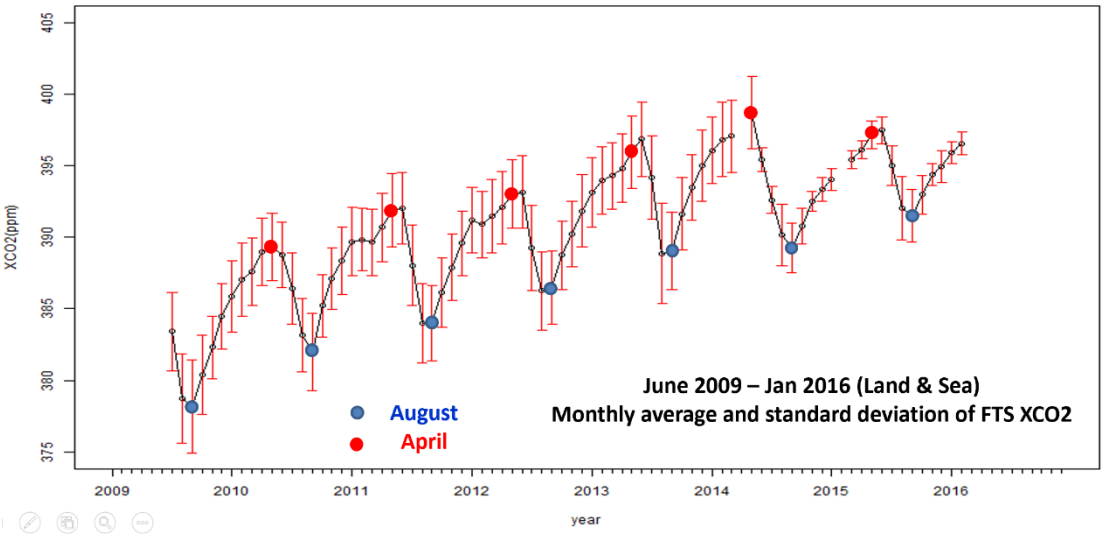
Homework 1 (Due March 19, 2021)

Exploratory data analysis using R & R graphics

1. Fifteen years of daily rainfalls observed at five weather stations in a region are given in the "raindata.csv" EXCEL file.
   (1) Read the daily rainfall data using the *read.csv* command and calculate the monthly rainfalls of individual stations.
   (2) Calculate the long-term average monthly rainfalls of individual stations.
   (3) Calculate and show the time series plot of the regional-average monthly rainfalls.
   (4) Identify the dry and wet seasons of the region.
2. The EXCEL file "Myanmar_Monthly_Rain.csv" contains the average monthly rainfalls of 54 rainfall stations in the Central Dry Zone (CDZ) of Myanmar.
   (1) In the same plot, show the average monthly rainfall time series of five randomly chosen stations. Note: You should **label the x-axis by months in 3-letter format, for example, JAN, FEB, MAR,…**



   (2) Plot a multiple-boxplot (same as shown in the class PPT file) of the average monthly rainfalls in the CDZ.
   (3) Calculate the mean and standard deviation of the average monthly rainfalls of individual months.
   (4) For July, identify the values of outliers and the upper and lower whiskers.
3. The dataset gosat_xco2_East_Asia.rds contains satellite (GOSAT) estimates of the CO_2 concentrations over the East Asia region for the period of 2009/06 to 2016/01.
   (1) Calculate the regional mean and standard deviation of the CO2 concentration for each month from 2009/06 to 2016/01.
   (2) Plot the time series of monthly mean and standard deviation of the CO2 concentration similar to the following figure.



4. The file “Wind_Speed.csv” contains 10-minute average wind speed observed at a wind farm.
   (1) Plot the boxplot of the wind speed data.
   (2) Plot the relative histogram of the wind speed.
   (3) Explain what you can observe form the above relative histogram.