Title: Exploring the relationship between streamflow statistics and drainage area
Group Members (Group 2): Meelisha Maharjan; Xinchen He

Hydrologistis generally use Q5 and Q95 as metrics to indicate the magnitudes of the streamflows. Q95 is defined as the flow equalled or exceeded 95% of the time, and it is a significant low flow metric. Similarly, Q5 is defined as the flow equalled or exceeded 5% of the time, which is a significant high flow metric. There is a general consensus in catchment hydrology that Q5 and Q95 are highly correlated with the drainage area, which is the land area where precipitation falls off into water bodies. Therefore, our objective is to study the relationship in the Massachusetts region. We obtained the streamflow and the drainage area of 149 streamflow gauges from 2000-01-01 to 2020-12-31. The Q5 and the Q95 of the streamflow were computed, then we fitted the regression models between the drainage area and the flow statistics.

The daily streamflow data was scraped from the website using regular expressions and were converted to relevant data types and each basin data was saved as a csv. The area was also obtained from the same website and stored as a csv. The data was validated by using the official R package: ‘dataRetrival’ released by the USGS to retrieve streamflow data published in their website, and the difference of the values were plotted. Q5 and Q95 were computed and plotted against the drainage area. We then used the psi regression function as a metric to fit a log linear regression between Q5 / Q95 and drainage area. We further used the bootstrap sampling method to generate a distribution for each coefficient and test the robustness of the log linear regression.

Objective: Fit a regression between Q5, Q95 and drainage area of a basin. Use bootstrap sampling method to create a distribution of the coefficients of these regression and generate a confidence interval.

Data source: National Water Information System (https://waterdata.usgs.gov/nwis/dv?referred_module=sw&site_no=01197000)

Responsibility description: Meelisha: streamflow data scraping, bootstrap sampling
Xinchen He: drainage area data scraping, dataframe operation, visualization

Conclusion: Using visualization and bootstrapping technique, we find that a log linear regression fits well between the streamflow statistics and drainage area.