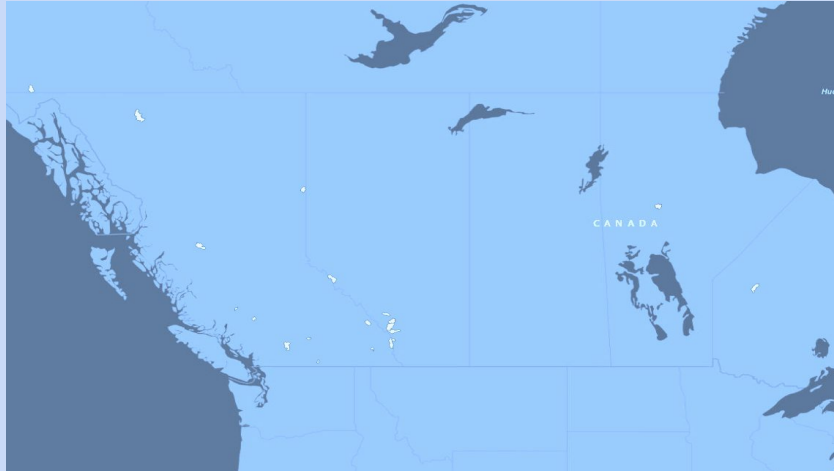# STAT450: Case Studies in Statistics
The Effects of Climate Variables on Average Stream Flow for Canadian Watersheds

Client: Asad Haris, UBC Postdoctoral Fellow

Group 2: Anjali Chauhan, Kelvin Li, Kohl Peterson, Vanessa Bayubaskoro

Introduction

- What is a watershed?
- Applications of watershed hydrology
- Want to understand impact & relationship of climate variables on stream flow

Objectives

- Find significant climate variables
- Visualize the relationships between climate variables and streamflows
- Predict the average stream flow from significant climate variables
- Detect outliers in the average streamflow values

Data

- 23 catchment areas located around Canada
- Size: 50 - 10,000 km^2
- Historical data from 1980 - 2018
- 61 unknown mean streamflow values
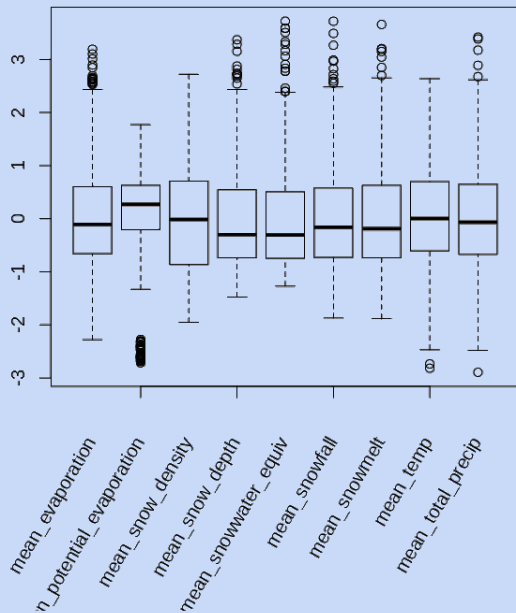- 9 climate variables of interest

Exploratory Data Analysis

Comparing different explanatory variables (before scaling)
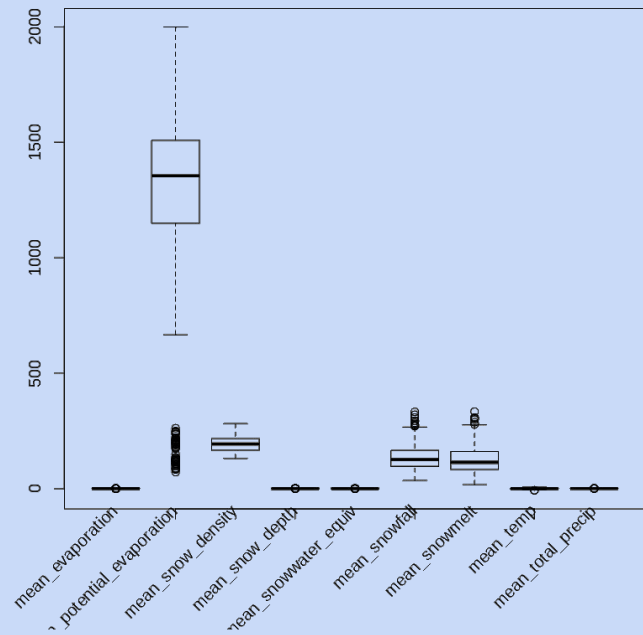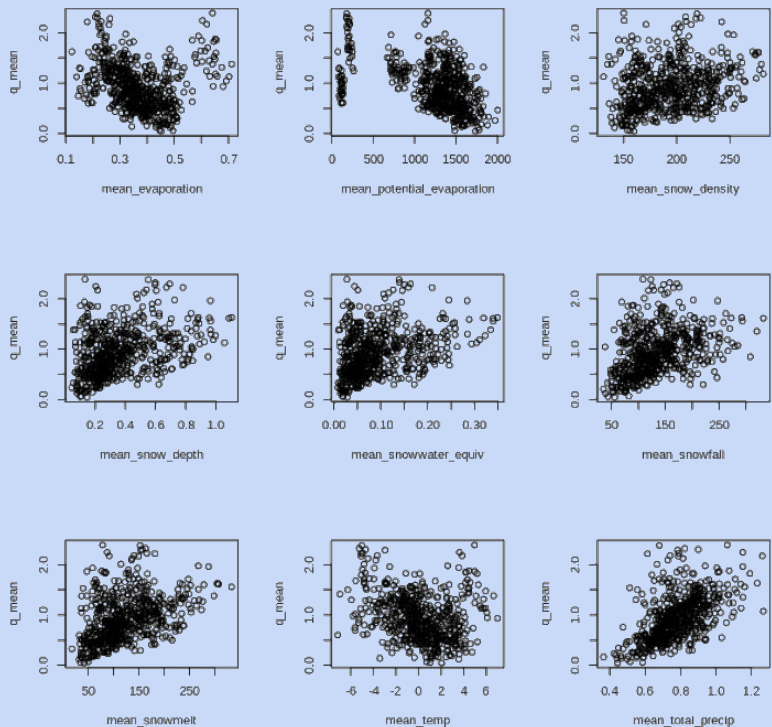

Before scaling
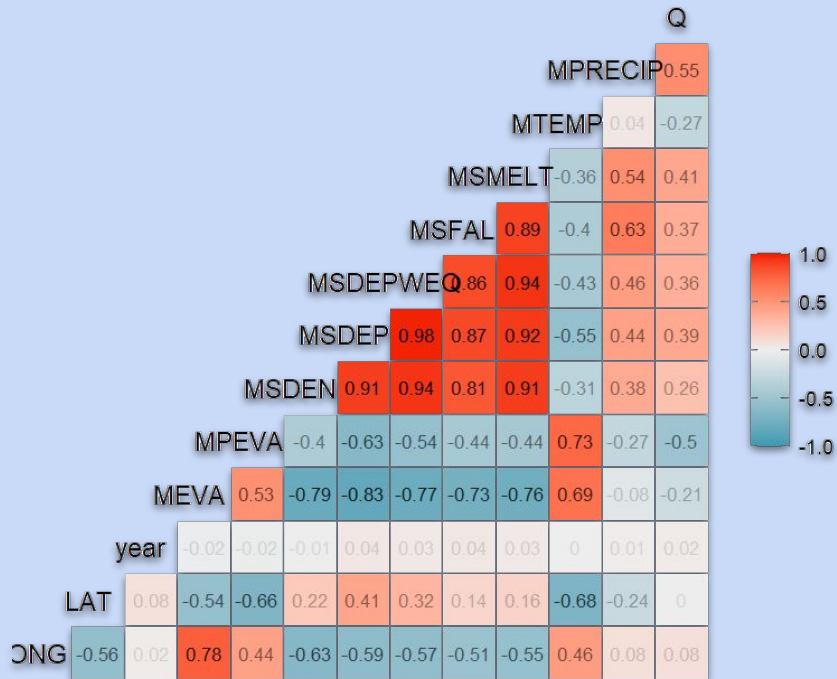
Comparing different explanatory variables (after scaling)

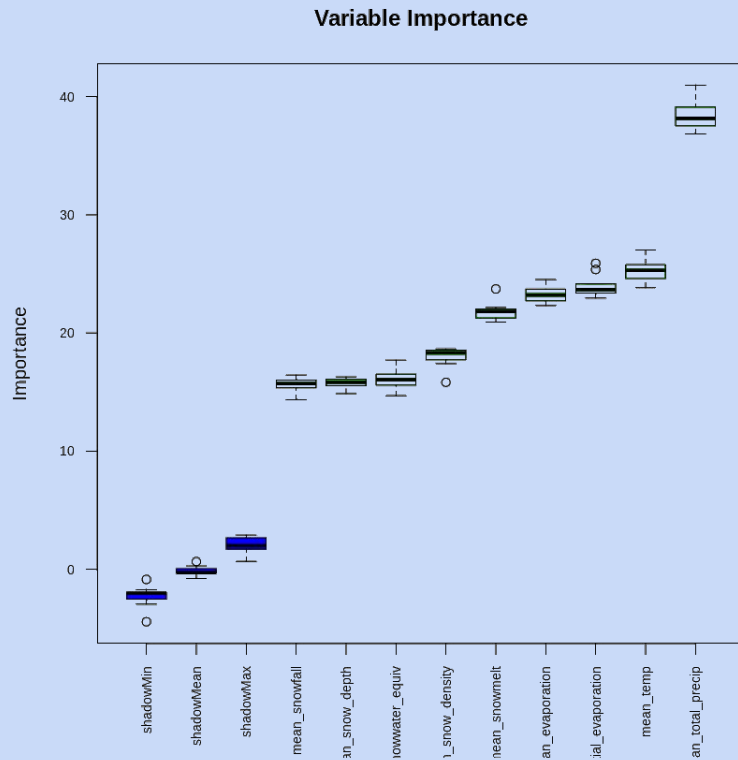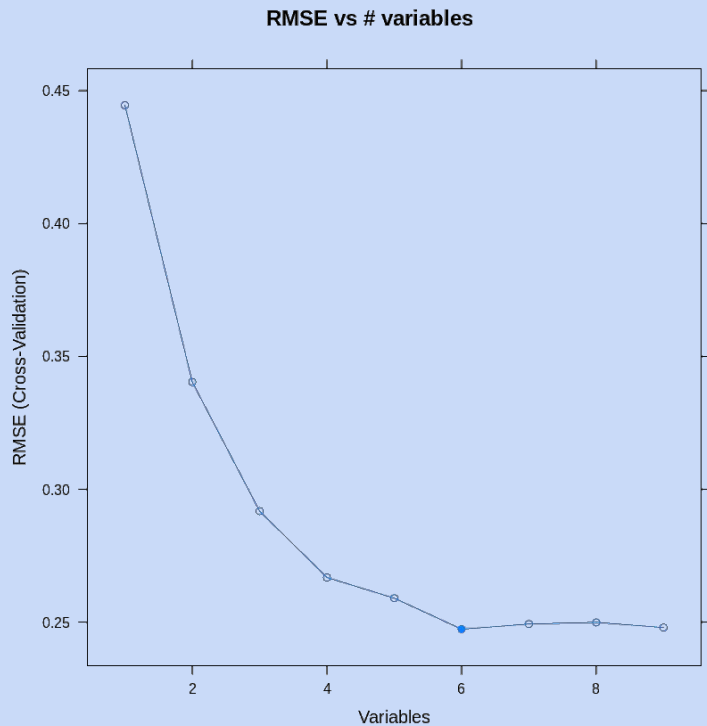*After scaling*

Exploratory Data Analysis

Relationship b/w explanatory variables and response
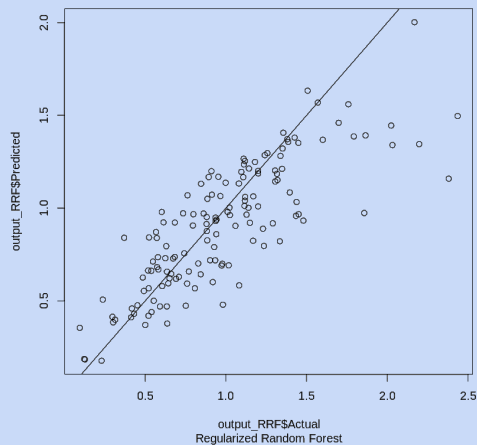
Heatmap: Correlation of explanatory variables
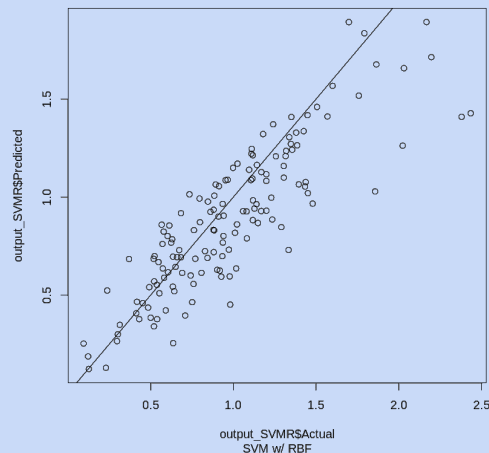
Statistical Analysis: Feature Selection
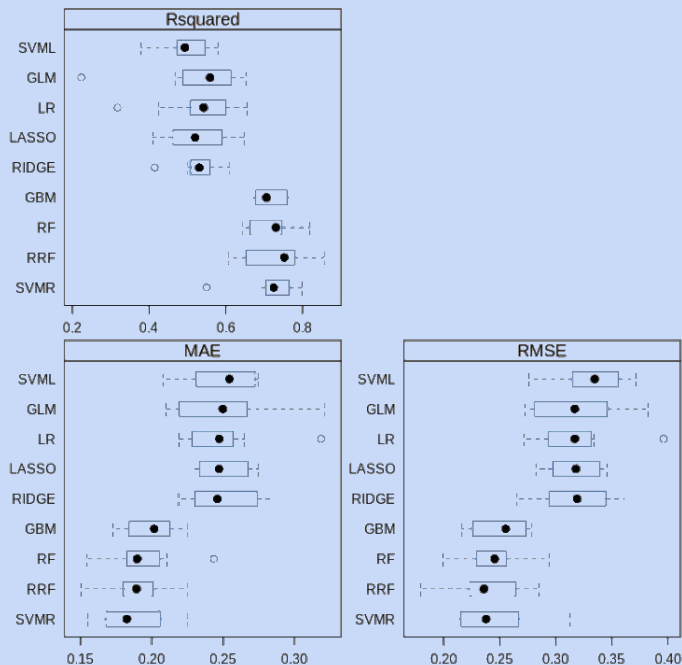
Statistical Analysis: Results

Future step(s)

- Revise the existing feature selection pipeline and incorporate feature engineering

- Perform hyperparameter tuning to improve existing model performance

- Outlier Detection using tree-based models