BHAG: Identify a quantitative predictive model of land use activities that lead to water quality issues in order to inform and support policies preventing these problems. First step: find best method.

**Study area:**
- 4 GTM NERR SWMP stations + 1000m buffer

**Data:**
- 18 water quality/nutrient parameters
  - monthly averages 2013-2017
- 26 land use variables for each station.

**Dependent Variables:**
- NO23, PO4, FecCol (CFU)

**Methods:**
- MLR, PLS, PCR, Lasso, RF
  - all with 5-fold cross validation

**Lasso method showed dramatically smaller test error, but was not a valid result**

- All other models returned error rates similar to one another
- All variables were used to test ML’s ability to perform variable selection. Prior selection is necessary
- Machine Learning has not been widely applied to environmental/ecological data, more studies would improve applications of ML to environmental/ecological data
- Next study includes more sites and less parameters