Notes for first meeting

Mengran Li

20/09/2021

Exploratory data analysis (EDA)

- To summary for individual year and stations.

- To transform the **date** variable into `Date` format and split it into **year**, **month** and **date** to summary.

- To detect the nonlinear date/month patterns via smoothed method.

- To detect the relationship of flow against altitude.

- To explore the distribution of response variable flow and transform it appropriately.

Methodology

- The R package `mgcv` provides a general tool to select model. Maximum likelihood estimation, restricted maximum likelihood estimation etc. are used to estimate the GAMs. It will be good to use different methods to obtain robust results.

- It will be helpful to do a lot about EDA before fitting the final model, which will reduce the times of model fitting.