

GRADUATE SEMINAR

Sichen Liu

Multiple Linear Model Selection: A Data Study on Methylmercury Production

MSc Student co-supervised by Dr. Andrei Volodin and Dr. Britt Hall

March 11, 2020

3:30P.M.

Mathematics and Statistics Lounge, CW307.20

Abstract:

Data on methylmercury (MeHg) concentrations and other environment parameters was collected between 2006 to 2012 from wetlands in the prairie pothole region in Saskatchewan by members of the Department of Biology, University of Regina. The goal was to use the backward elimination model selection to develop the best fitted linear regression model to test the impact of seven environment variables on methylmercury concentration.

I analyzed the whole data set and removed all missing values and transformed some variables such as methylmercury, sulfate, and conductivity. I compared linear regression model with two-way interactions and no interaction linear model. After applying the linear model selection, I used several methods to check the model including Residual vs. Fitted Values, Normal Q-Q plots, compare AIC; after I checked for outliers in the combined data set with final model.