The growing use of pesticides in agricultural activities has led to higher water pollution especially within the context of small streams. The extent of streams summing into thousands of kilometres makes it unfeasible to sample locations across the stream network, hence the need for prediction modelling. There are different statistical methods for modelling of stream, they include linear modelling method, general addictive modelling method, statistical stream network modelling method, and machine learning methods. Each of the above methods have their respective strengths and weakness. This research compares the prediction accuracy of linear model, general addictive model and statistical stream network model with regards to water quality in small streams.

We took a total of 101, 752, 868 samples from Baden-Wurttemberg, Nordrhein-Westfalen and Sachen federal states of Germany respectively. We then fitted linear models, general addictive models, and statistical stream network models. Prediction analysis was done using the leave one out cross validation. The whole analysis was done with two response variables with normal and count distribution respective. The result shows that statistical stream network models outperformed other methods, while the general addictive model performed better than linear models in both cases of the response variable. This further confirms the strong spatial autocorrelation between sites in streams and the non-linear nature of stream water quality and various land uses.