

DS421 SUMMER RESEARCH EXPERIENCE SUMMARY

SARA STOUDT

This summer I worked in collaboration with David Senn and Erica Spotswood at the San Francisco Estuary Institute and Perry de Valpine in the Environmental Science and Policy Management department to better understand what variables help explain the variation of chlorophyll over time and at each station in the Bay Delta. We want to understand if these relationships differ over time and/or over space. I built various Generalized Additive Models (GAMs) to explore these relationships. For each model, I fit it in a nested way so that we can see the contribution of each variable to the overall predicted value. As I built different models, I displayed their fitted values, the components of each predictor variable on the overall fitted values, and other relevant plots in a Shiny application to allow for easy exploration and comparison of models.

As a part of this project, I also worked on a comparison of GAMs and Weighted Regression on Time, Discharge, and Season (WRTDS) in collaboration with Marcus Beck from the Environmental Protection Agency. He is also collaborating with SFEI and building WRTDS models for many of the same stations that I was looking at for chlorophyll but for different nutrients. I built GAM models using his setup and the comparison framework that he used in another GAM/WRTDS comparison in a different location to provide another case study for comparison. For this portion of the project, I also created a Shiny application for easier exploration and comparisons.

Throughout work on this project, I became more familiar with GAMs including strategies for reducing the wait time for more computationally complex ones, learned about WRTDS and its approach to bringing flow into models for nutrients, familiarized myself with the context of the ecology of the Bay Delta as well as improved my Shiny application and GitHub skills.

Towards the end of the project, I have been trying to incorporate information about the flow structure into the models, forcing the flow-connectedness property to be respected. This kind of spatial statistics on flow networks is still a fairly open problem that I hope to continue to delve into deeper in the near future. Even if this topic doesn't turn into dissertation work, data from SFEI could prove to be a perfect test site to work on other spatial methods and environmental applications.

Track my progress and find out more about my work here https://github.com/sastoudt/DS421_summerProject. The links to the online hosted Shiny apps can be found here as well.