

Name: Cristian Swift

Mentor and Department: Dr. François Riblet and Dr. E. Virginia Armbrust, School of Oceanography

Project Title: Random Forest Regression Models for Predicting Picophytoplankton Biomass Partitioning in the Eastern Pacific.

The partitioning of phytoplankton biomass is crucial for understanding how climate change impacts the carbon cycle as it sheds light on the distribution and dynamics of different phytoplankton groups. However, there is a significant knowledge gap regarding the environmental factors that drive the partitioning of phytoplankton biomass, particularly in warm, nutrient-poor regions that are predicted to expand in future ocean conditions. This study focus on the dominant phytoplankton groups in these regions: *Prochlorococcus*, *Synechococcus*, and small eukaryotic phytoplankton (less than 2 μm in diameter). Using random forest regression models, we assessed the predictability of phytoplankton biomass based on salinity, temperature, light intensity, and dissolved inorganic nutrient concentrations (nitrate, phosphorus, and iron). Model performance was evaluated using 2,800 observations per population were obtained through high-frequency flow-cytometry in surface water of the North Subtropical and Subpolar gyres. Mean Absolute Error Percentage varied among picoplankton population's Random Forest models, *Prochlorococcus*: 7.71%, *Synechococcus*: 2.30%, Nanoeukaryotes: 3.34%, and Picoeukaryotes: 6.57%. Results highlight the importance of nitrate, salinity, and to a lesser extent, iron, as predictors of *Prochlorococcus* biomass. Conversely, phosphate and nitrate emerge as the primary drivers of *Synechococcus* and picoeukaryote biomass, respectively. The upper and lower limits of the biomass and environmental data in the dataset define the boundaries within which a Random Forest can generate its predictions, making these models specific to the location and timing of this study. Further research is needed to explore additional factors and refine prediction models by considering factors like nutrient uptake rates, grazing pressure, and microbial interactions.