Wunderpus Octopus: Modeling the relationship between biogeochemical layers and chlorophyll density

Francesca Balestrieri, Deniz Olgu Devecioglu, Nadir Hajouji, Saswat Mishra, Kshitiz Parihar, Ingrida Semenec

Chlorophyll concentration in the ocean reflects the interaction of factors like nutrient availability and sunlight, influencing oceanic biomass productivity. High chlorophyll levels suggest nutrient-rich waters where phytoplankton, the ocean's primary producers, thrive, which in turn impacts fishery population dynamics and ecosystem health. Additionally, from chlorophyll patterns, we can gain insight into climate change, ocean currents, and mixing processes. A predictive and/or explanatory analysis using machine and deep learning models will help address these environmental challenges.

Stakeholders: NOAA, Fish and Wildlife Department, local fisheries, and coastal communities. **KPI**: Predicting/Forecasting chlorophyll levels accurately.

Datasets: We used various datasets from Copernicus Marine (Global Ocean Color, Global Ocean Biochemistry Hindcast, Global Ocean OSTIA, Multi Observation Global Ocean, and Global Ocean Surface Carbon). We focused our analysis on the North Sea region, looking at monthly data from Sept. 1997 to Dec. 2021.

Methods and Architectures: We did our initial modeling using XGBoost, and then designed two different types of neural networks.

- A "pure" CNN, inspired by NetResDeep, which predicts chlorophyll levels.
- A CNN+LSTM tries to predict the next month of data (chlorophyll and features).

Results

- The CNN models demonstrated promising results by identifying spatial features and patterns, although there is still room for improvement, especially in areas with higher chlorophyll concentrations near the shore.
- We found the NetResDeep model to perform the best among the three models.

Conclusion

Iron concentration and pH levels were the major contributors to chlorophyll levels. Geoengineering by iron fertilization can lead to a better fishery ecosystem.