EGU24-18663, updated on 24 Jul 2024
https://doi.org/10.5194/egusphere-egu24-18663
EGU General Assembly 2024


# Linking Satellite and physics-informed Data with Phytoplankton communities Using Deep Learning

**Luther Ollier**[1], Roy El Hourany[2], and Marina Levy[1]
[1]Sorbonne, LOCEAN, PROTEO, France
[2]Université de Lille, LOG, France

Understanding Phytoplankton community dynamics in response to environmental shifts is crucial for assessing the impact of climate change on marine biology. To this end, satellite observations offer a dataset spanning two decades, capturing diverse sea surface parameters, including temperature, ocean color, and surface height. Notably, ocean color data is processed to derive sea surface chlorophyll-a concentration, widely acknowledged as a reliable proxy for phytoplankton biomass.

Lately, advances in ocean color observation allow us to describe the phytoplankton community structure in terms of groups (broad functional or taxonomic groups) or size classes. Although these advances provide more detailed information on phytoplankton diversity and structure, these datasets suffer from spatial and temporal coverage limitations due to strict quality control in the presence of atmospheric aerosols, clouds, sea ice, etc... As a result, studies examining phytoplankton trends over the past two decades and future projections rely on incomplete chlorophyll-a and ocean color data. Therefore this compromises the identification of consistent trends within phytoplankton datasets.

In this study, we address this issue using a deep-learning approach. Our method constructs an attention network that learns from the available satellite dataset of Chla and phytoplankton size classes images (weekly and one-degree-degraded spatial resolution) while assimilating information from gap-free sea surface physics data originating from satellite observations and assimilated numerical models). The primary objective is to estimate the phytoplankton dataset based on the knowledge of physical factors, while filling the gaps within this dataset

The trained deep-learning model allows us to discern patterns and correlations between chlorophyll concentration and the phytoplankton size classes on one hand, and the physics-based data on the other hand. From a phytoplankton weekly database spanning from 1997 to 2020, with 50% missing pixels, our approach demonstrates promising results in replicating chlorophyll concentration and accurately inferring phytoplankton size classes.

The methodology shows the potential of deep-learning for robust ecological applications but mainly lays the groundwork for future trend studies on phytoplankton communities.