EGU24-16166, updated on 07 Dec 2024
https://doi.org/10.5194/egusphere-egu24-16166
EGU General Assembly 2024

A global daily gap-filled chlorophyll-*a* dataset in open oceans during 2001–2021 from multisource information using convolutional neural networks

Zhongkun Hong[1], Di Long[1], Xingdong Li[1], Yiming Wang[1], Jianmin Zhang[1], Mohamed Hamouda[2], and Mohamed Mohamed[2]

[1]State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing, China

[2]Department of Civil and Environmental Engineering, United Arab Emirates University, Al Ain, United Arab Emirates

Ocean color data are essential for developing our understanding of biological and ecological phenomena and processes and also of important sources of input for physical and biogeochemical ocean models. Chlorophyll-*a* (Chl-*a*) is a critical variable of ocean color in the marine environment. Quantitative retrieval from satellite remote sensing is a main way to obtain large-scale oceanic Chl-*a*. However, missing data are a major limitation in satellite remote-sensing-based Chl-*a* products due mostly to the influence of cloud, sun glint contamination, and high satellite viewing angles. The common methods to reconstruct (gap fill) missing data often consider spatiotemporal information of initial images alone, such as Data Interpolating Empirical Orthogonal Functions, optimal interpolation, Kriging interpolation, and the extended Kalman filter. However, these methods do not perform well in the presence of large-scale missing values in the image and overlook the valuable information available from other datasets for data reconstruction. Here, we developed a convolutional neural network (CNN) named Ocean Chlorophyll-*a* concentration reconstruction by convolutional neural NETwork (OCNET) for Chl-*a* concentration data reconstruction in open-ocean areas, considering environmental variables that are associated with ocean phytoplankton growth and distribution. Sea surface temperature (SST), salinity (SAL), photosynthetically active radiation (PAR), and sea surface pressure (SSP) from reanalysis data and satellite observations were selected as the input of OCNET to correlate with the environment and phytoplankton biomass. The developed OCNET model achieves good performance in the reconstruction of global open ocean Chl-*a* concentration data and captures spatiotemporal variations of these features. The reconstructed Chl-*a* data are available online at https://doi.org/10.5281/zenodo.10011908. This study also shows the potential of machine learning in large-scale ocean color data reconstruction and offers the possibility of predicting Chl-*a* concentration trends in a changing environment.