A global 3D chlorophyll-a dataset derived from multimodal deep-learning reconstruction

Ziyue Meng[1,2], Yuntao Wang[2#], Lei Han[1#]

[1] *Faculty of Information Science and Engineering, Ocean University of China, Qingdao 266100, China*

[2] *State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, China*

Correspondence to: Yuntao Wang (yuntao.wang@sio.org.cn), Lei Han (hanlei@ouc.edu.cn)

Abstract. Chlorophyll-*a* (Chl-*a*) is a key variable for characterizing marine phytoplankton biomass and upper-ocean biogeochemical variability. Existing global products are generally limited to the ocean surface and therefore cannot adequately resolve subsurface vertical structure. In this study, a multimodal Profile-Surface Transformer (PST) framework was developed by integrating Biogeochemical-Argo (BGC-Argo) Chl-*a* profiles, Core-Argo temperature and salinity profiles, and satellite-derived surface Chl-*a* observations to reconstruct the three-dimensional (3D) vertical structure of Chl-*a* from 10 to 500 m depth. A global monthly mean 3D Chl-*a* dataset at 1° spatial resolution for January 2005–June 2025 was subsequently constructed.

Evaluation demonstrates that the PST framework provides robust and stable profile reconstruction performance across different generalization scenarios. The coefficients of determination under random split, year-based split, and spatial cross-validation are 0.89, 0.87, and 0.86 ± 0.01, respectively. Furthermore, independent external validation using ship-based observations confirms that the model maintains good generalization performance across different observing systems. The resulting dataset successfully reproduces established global patterns of Chl-*a* vertical structure, including the subsurface chlorophyll maximum (SCM), as well as their characteristic seasonal variability.

Overall, the long-term 3D Chl-*a* dataset integrates satellite surface observations with vertically resolved in situ measurements to provide an observationally constrained representation of global Chl-*a* structure from the upper ocean to 500 m depth. It provides a new reference product for investigating phytoplankton vertical structure, ocean biogeochemical variability, and upper-ocean ecosystem responses to environmental and climatic variability. The dataset is publicly available from Zenodo (Meng et al., 2026a) at https://doi.org/10.5281/zenodo.22726410.