CHLA depth profiles

Deriving Vertical Chlorophyll-a Structure and Ecological Indicators Using PACE Hyperspectral Data

Overview

We propose to develop and evaluate a machine-learning (ML) system to estimate chlorophyll-a vertical profiles (CHLA(z)) globally using PACE hyperspectral Rrs, trained on in situ vertical profile data from Bio-Argo and OOI profilers. These CHLA(z) estimates will enable a suite of new ecological and fisheries-relevant indicators that were not possible with past ocean color missions.

Key output: A validated global model delivering daily 10 m CHLA profiles from PACE.

Step 1 — Predict CHLA(z) from PACE hyperspectral Rrs

Approach

Train an ML model (BRT, CNN, or hybrid) to predict CHLA at 10 m depth bins from 0–200 m.

```
import xarray as xr
print(1)
```

1

Phase 1 Inputs:

- PACE Rrs hyperspectral radiances and solar_hour
- Depth binned (10m) CHLA averages from Bio-Argo BGC and OOI Coastal Arrays

Pilot Phase

1. Demonstrate a proof of concept using Boosted Regression Tree and CNN and Level 3 PACE data.
2. Evaluate performance using held-out profiles: RMSE by depth bin, DCM depth error, Integrated CHLA (0–100 m, 0–200 m) errors, maybe Vertical correlation metrics (profile shape similarity)

Step 2 — Derive ecological and fisheries indicators from CHLA(z)

Once CHLA(z) is predicted in 10 m bins from 0–200 m, we can derive a suite of ecosystem indicators directly from the vertical structure, and then extend that with mixed-layer depth (MLD) from existing ROMs/reanalysis products.

2A. Indicators that require only CHLA(z)

These use only the predicted CHLA profiles:

- **DCM depth**
  Depth of maximum CHLA in the 0–200 m profile. Indicates where biomass peaks and identifies deep vs shallow bloom regimes.

- **Depth-integrated CHLA (CHLA_0–100, CHLA_0–200)**
  Vertical integral of CHLA(z). Proxy for total phytoplankton biomass in the upper water column.

- **CHLA-weighted mean depth**

$$z_{\mathrm{mean}} = \frac{\sum z_i \cdot CHLA_i}{\sum CHLA_i}$$

  Captures where, on average, biomass is located vertically.

- **Productive layer thickness**
  Thickness of the layer where CHLA(z) exceeds a chosen fraction of the profile maximum (e.g., 50% of max). Quantifies how vertically concentrated vs diffuse the biomass is.

- **Bloom vertical structure class**
  Cluster/profile classification into surface-intensified, deep-submerged, double-peaked, or vertically uniform blooms. Supports regime mapping and ecosystem state characterization.

- **Deep vs shallow bloom regimes**
  Simple metrics like (DCM depth / Zeu) or (DCM depth threshold) to flag deep vs shallow biomass structure.

- **Vertical CHLA anomalies**
  Anomalies relative to a regional or climatological mean CHLA(z) profile, highlighting unusual vertical structure (e.g., unusually deep maxima, unusually surface-intensified blooms).

2B. Indicators that combine CHLA(z) with Mixed Layer Depth (MLD)

MLD is available from global reanalyses and regional ROMS products (GLORYS, HYCOM, ECCO, CCSROMS, etc.), so we can combine physical structure with PACE-derived biology:

- **Mixed Layer Chlorophyll (ML-CHL)**
  CHLA averaged from the surface down to the local MLD:

$$\text{ML--CHL} = \frac{1}{\text{MLD}} \int_0^{\text{MLD}} CHLA(z)\,dz$$

  More directly represents the biomass experienced by organisms living and feeding in the mixed layer.

- **Fraction of CHLA above vs below the MLD**

$$f_{\text{above}} = \frac{\int_0^{\text{MLD}} CHLA(z)\,dz}{\int_0^{200m} CHLA(z)\,dz}$$

  Distinguishes systems where biomass is confined to the mixed layer vs concentrated in a deep maximum.

- **DCM–MLD separation**
  Difference between DCM depth and MLD depth:

$$\Delta z = z_{\text{DCM}} - z_{\text{MLD}}$$

  Identifies "detached" deep chlorophyll maxima vs mixed-layer–attached blooms, relevant for habitat compression and nutrient/light regime classification.

- **MLD-normalized habitat metrics**
  Combine ML-CHL with MLD and CHLA-weighted depth (e.g., CHLA-weighted depth / MLD) to characterize how biomass is distributed relative to the mixed layer.

- **MLD-conditioned CHLA anomalies**
  Vertical CHLA anomalies stratified by MLD regime (e.g., shallow vs deep mixed layers), to help attribute anomalies to physical vs biological drivers.

Step 3 — Map these indicators globally using PACE

Once the CHLA(z) model is trained and MLD fields are available on the same grid:

1. Apply the CHLA(z) model to global PACE Rrs to obtain daily CHLA profiles (0–200 m).
2. Interpolate external MLD fields (e.g., GLORYS, HYCOM) to the PACE grid.
3. Compute:

   - DCM depth and magnitude

   - Integrated CHLA_0–100, CHLA_0–200

   - CHLA-weighted mean depth

   - Productive layer thickness and bloom structure class

   - ML-CHL, fraction above/below MLD, and DCM–MLD separation

4. Produce global maps and time series of these indicators, and evaluate them against Bio-Argo/OOI for validation and, in later phases, against fisheries-relevant metrics (e.g., recruitment or habitat indices).