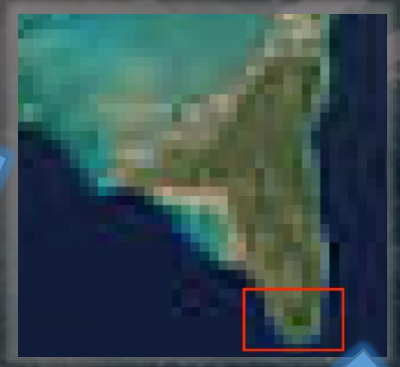
EDMW 2025

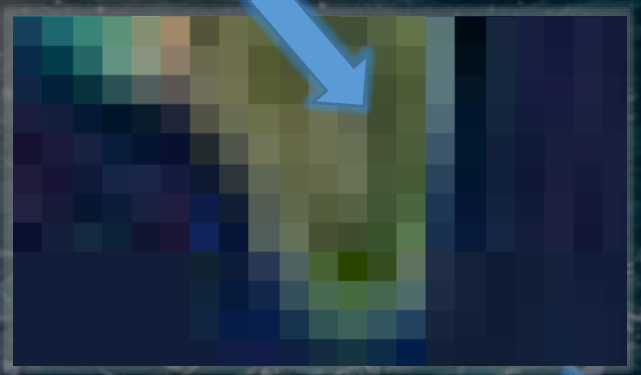
Introduction to using earth data in the cloud for scientific workflows

PACE data products

Ryan Vandermeulen | NMFS Office of Science and Technology
Eli Holmes | NMFS Northwest Fisheries Science Center

NOAA Fisheries | U.S. Department of Commerce

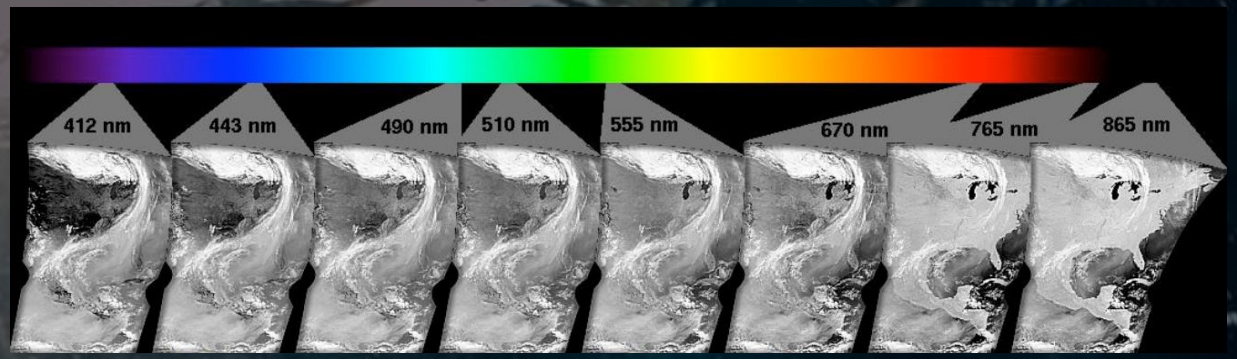




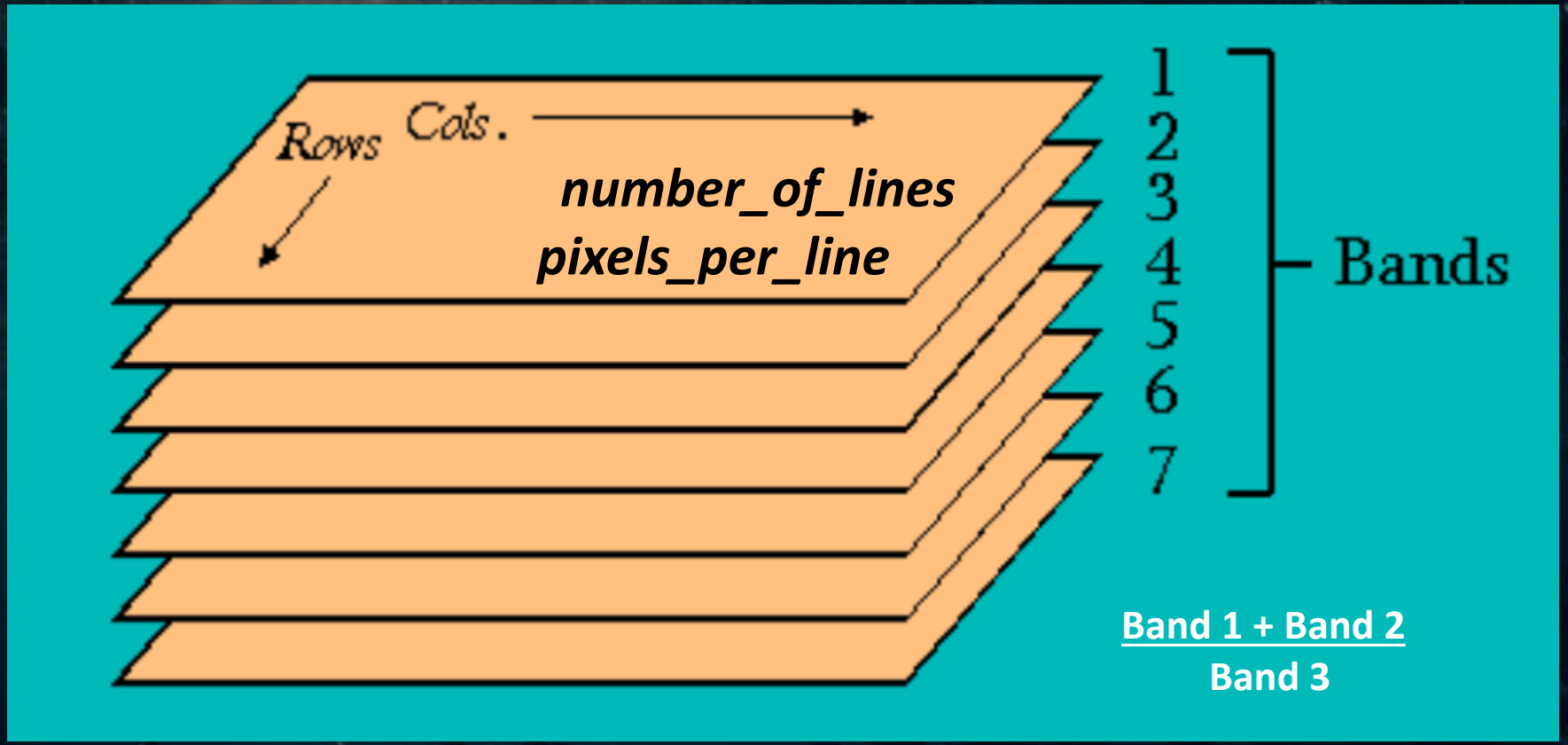
111	124	107	91	98	85	82	79	82	91	88	91	142	238	194	168	171	142	142	136	136
111	113	95	91	98	91	85	82	91	104	101	101	136	206	203	178	187	145	127	136	161
113	104	95	91	82	88	85	91	104	107	95	95	111	165	199	183	133	104	120	120	111
120	104	111	91	82	88	91	98	101	101	101	124	140	136	129	111	111	101	101	104	98
107	107	107	88	79	95	104	95	101	104	111	107	104	113	117	104	98	98	98	98	107
98	111	95	85	85	95	98	91	95	98	101	104	107	111	113	104	98	98	91	95	101
104	107	85	82	85	91	91	95	95	95	107	107	111	113	107	111	111	104	95	98	101	104
101	91	79	82	82	88	88	98	101	104	111	113	107	104	113	111	98	104	104	101	124
88	82	79	79	82	85	88	95	107	107	107	107	113	111	127	127	104	111	107	129	165
88	79	82	79	79	91	88	98	111	104	111	113	117	133	187	219	183	161	165	152	149
82	82	79	85	85	91	88	101	111	107	111	113	168	190	255	255	228	212	183	161	168
82	85	82	91	98	95	95	95	107	107	111	142	241	255	255	235	255	238	232	248	219
88	85	85	98	104	104	101	104	117	113	104	197	255	255	255	215	255	255	251	212	212
82	85	85	95	101	107	101	124	120	113	104	181	241	251	235	199	212	212	219	206	183
82	88	98	101	107	101	111	117	120	117	107	165	210	219	199	199	203	199	212	241	255
88	91	101	107	111	107	117	101	101	113	107	140	194	215	181	215	226	210	206	241	255
95	91	95	104	111	113	117	111	98	113	129	133	210	210	203	210	210	199	178	199	228
91	88	101	101	104	117	111	117	111	111	113	127	215	232	212	199	194	187	203	206	199
91	95	104	98	107	117	120	111	111	101	113	158	206	194	206	203	181	178	212	199	181
101	104	98	98	129	161	124	113	113	107	101	181	199	197	210	199	181	183	199	199	174
113	101	111	117	219	228	197	133	107	98	95	107	187	203	228	171	149	161	174	190	161

What are we actually looking at?

What's in a NetCDF file?



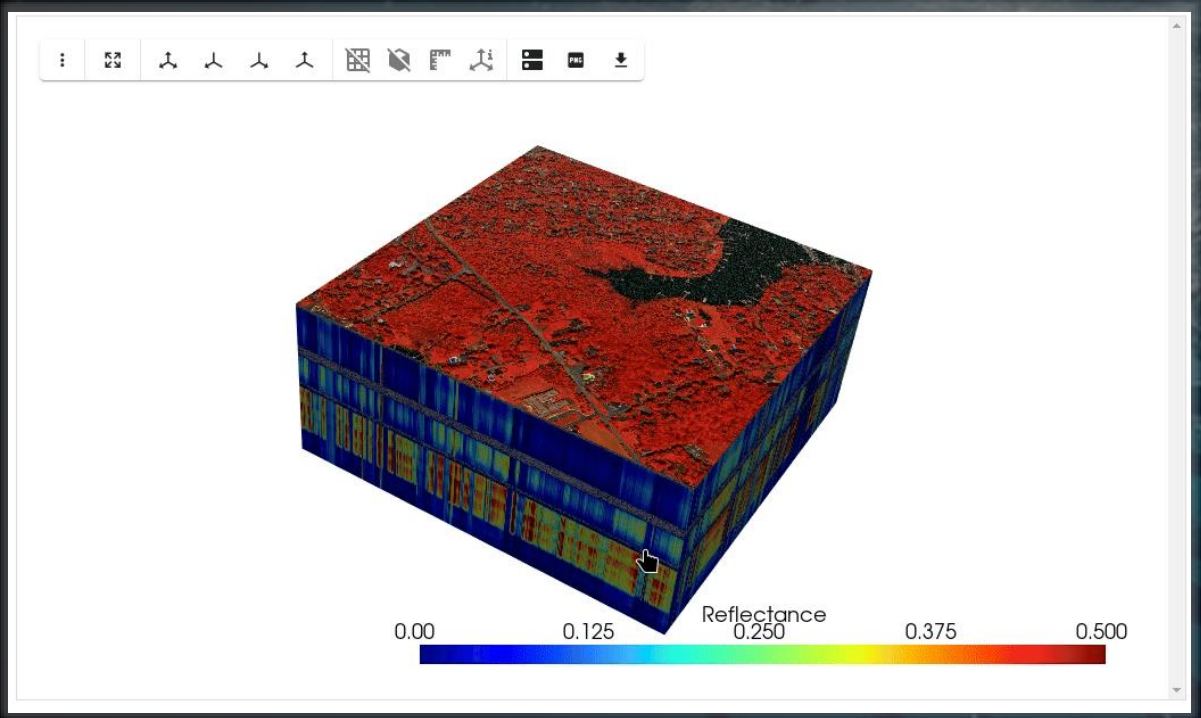
NetCDF
Network Common Data Form



What is NetCDF?

- **NetCDF** (network Common Data Form) - A platform independent format for representing multi-dimensional array-orientated scientific data.

- Self Describing - a netCDF file includes information about the data it contains.



- Direct Access - a small subset of a large dataset may be accessed efficiently, without first reading through all the preceding data.

- Sharable - one writer and multiple readers may simultaneously access the same netCDF file.

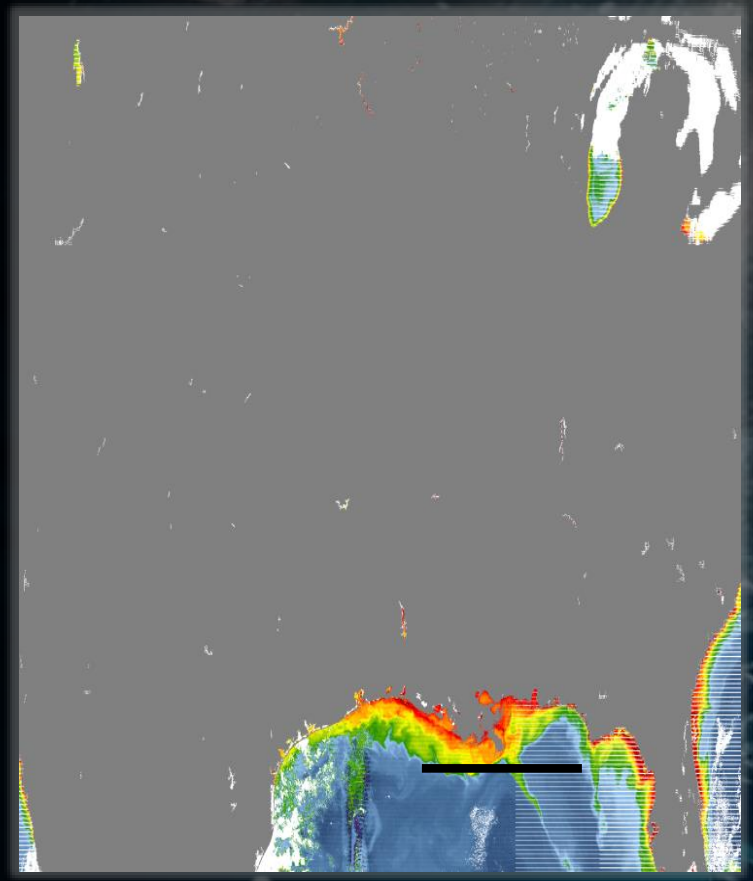
A NetCDF file consists of:

- Global Attributes: Describe the contents of the file
- Dimensions: Define the structure of the data (e.g., Time, Latitude, Longitude, Wavelengths)
- Variables: Holds the data in arrays shaped by Dimensions
- Variable Attributes: Describes the contents of each variable

Unraveling your data

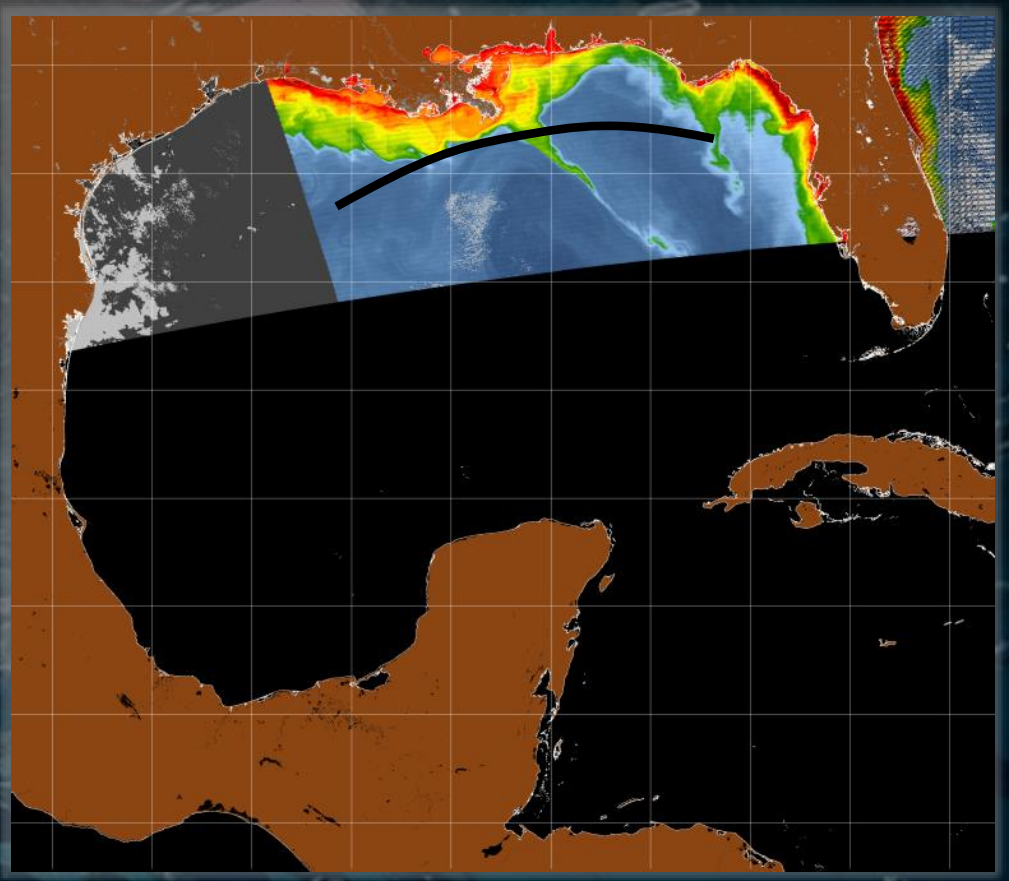
- Level 0: raw radiance counts
- Level 1a: radiance + telemetry and calibration data
- Level 1b: calibrated and geo-located
- Level 2: *derived* geo-physical values (algorithms), atmospherically corrected.
- Level 3: variables arranged in a uniform space-time grid (Earthling problems)

Level 2

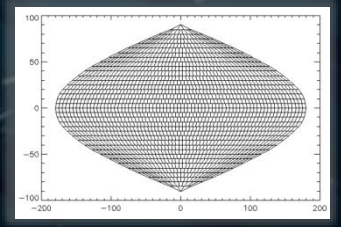


As seen from space,
Observing our marble

Level 3



Data binned to a projection
That we can use for GIS!

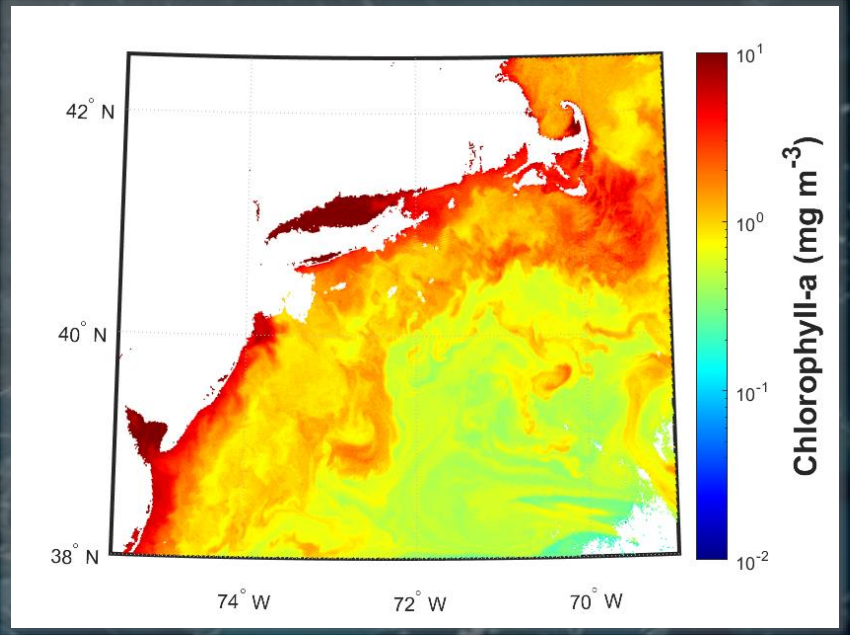


PACE “bundles”

- PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_BGC.V3_0.nc
- PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_AOP.V3_0.nc
- PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_IOP.V3_0.nc
- PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.PAR.V3_0.nc
- PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.SFREFL.V3_0.nc
- PACE_OCI.[YYYYMMDD]T[HHMMSS].L3m.MO.MOANA.V3_0.4km.nc

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_BGC.V3_0.nc

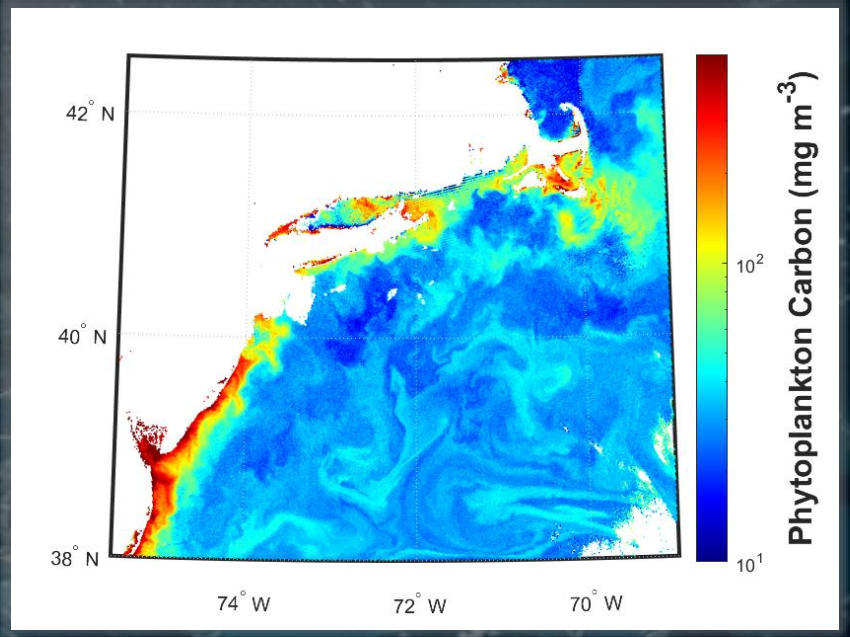
- BGC = Biogeochemistry
- **chlor_a - chlorophyll-a**
- carbon_phyto - phytoplankton carbon
- poc - particulate organic carbon



Note: A ubiquitous indicator of phytoplankton biomass. In some cases, increases/decreases in concentration do not explicitly reflect biomass change.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_BGC.V3_0.nc

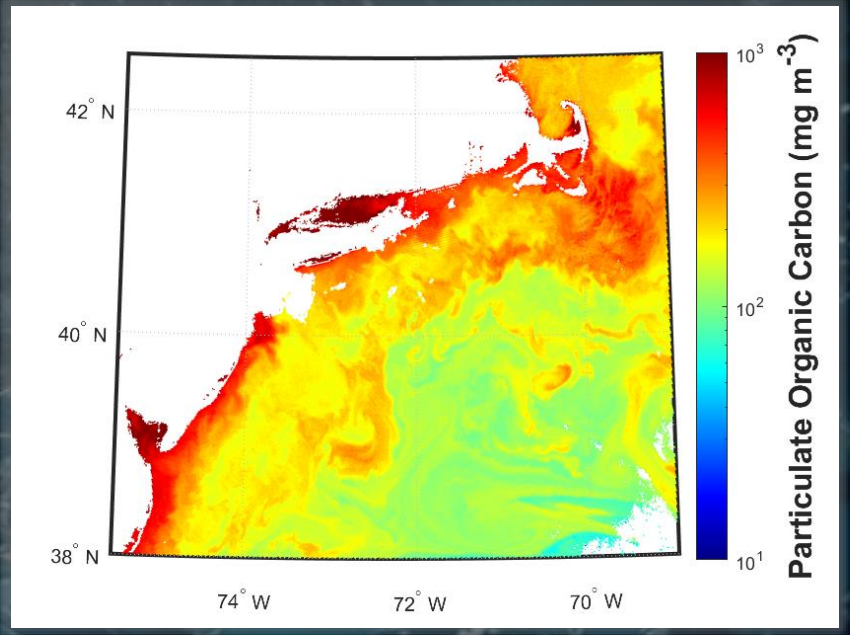
- BGC = Biogeochemistry
- chlor_a - chlorophyll-a
- **carbon_phyto - phytoplankton carbon**
- poc - particulate organic carbon



Note: This product is related to particle backscattering. Arguably, a more direct indicator of phytoplankton biomass, but it hasn't been as well vetted.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_BGC.V3_0.nc

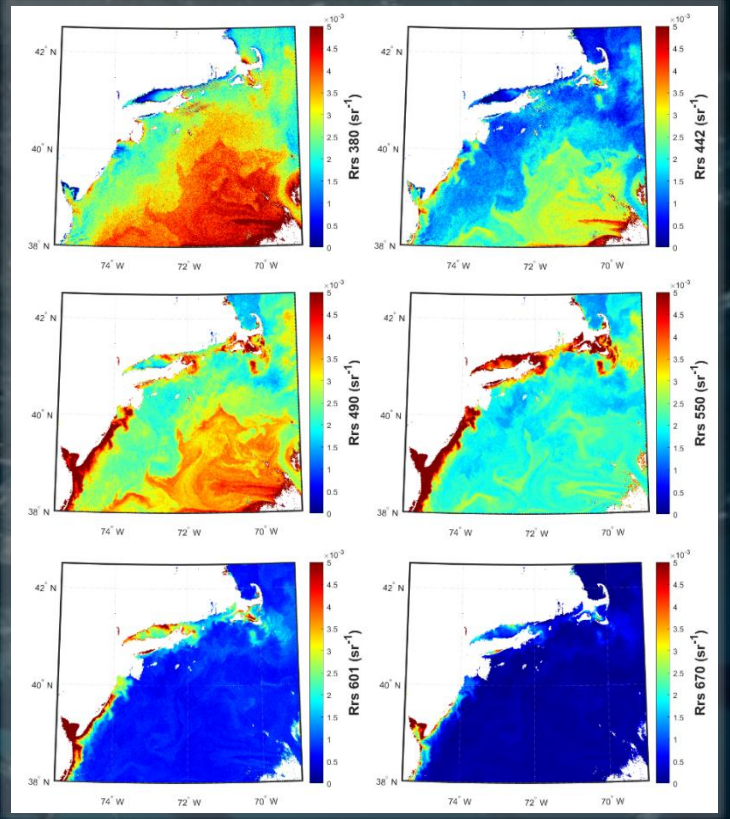
- BGC = Biogeochemistry
- chlor_a - chlorophyll-a
- carbon_phyto - phytoplankton carbon
- **poc - particulate organic carbon**



Note: All carbon-containing particles suspended in seawater (phytoplankton, detritus, bacteria)- particularly relevant for carbon export processes.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_AOP.V3_0.nc

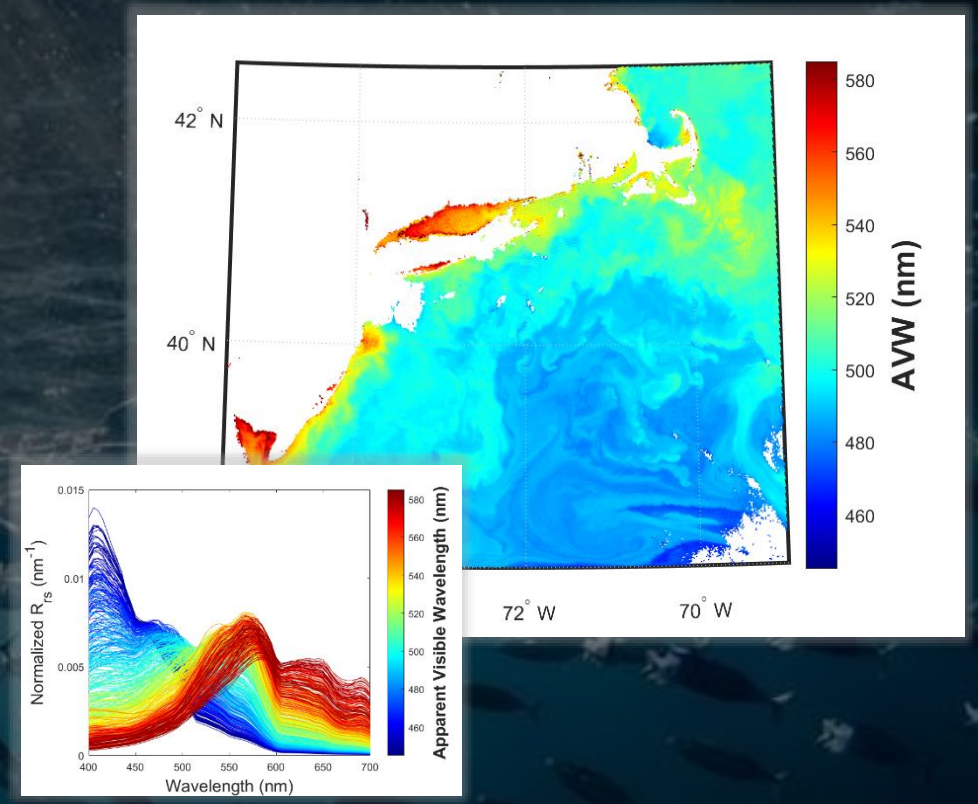
- AOP = Apparent Optical Properties
- **Rrs - remote sensing reflectance at 184 wavelengths (339–719 nm)**
- avw - apparent visible wavelength
- nflh – normalized fluorescence line height



Note: The mathematical description of color in the ocean. This product is the baseline for the construction of most algorithms.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_AOP.V3_0.nc

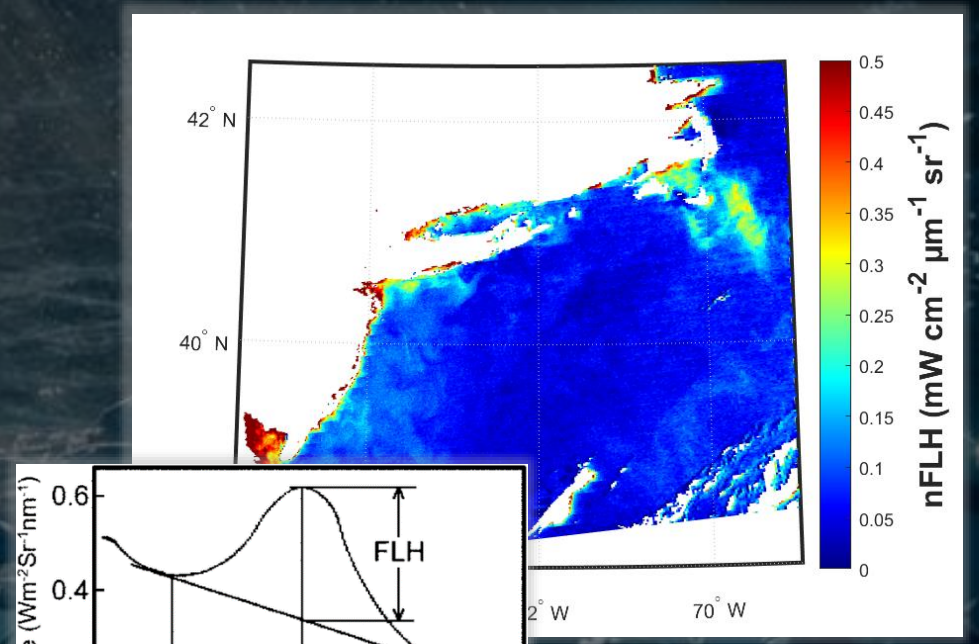
- AOP = Apparent Optical Properties
- Rrs - remote sensing reflectance at 184 wavelengths (339–719 nm)
- **avw - apparent visible wavelength**
- nflh – normalized fluorescence line height



Note: Objective descriptor of Rrs spectral shape. Higher value = red-shifted, lower value = blue-shifted.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_AOP.V3_0.nc

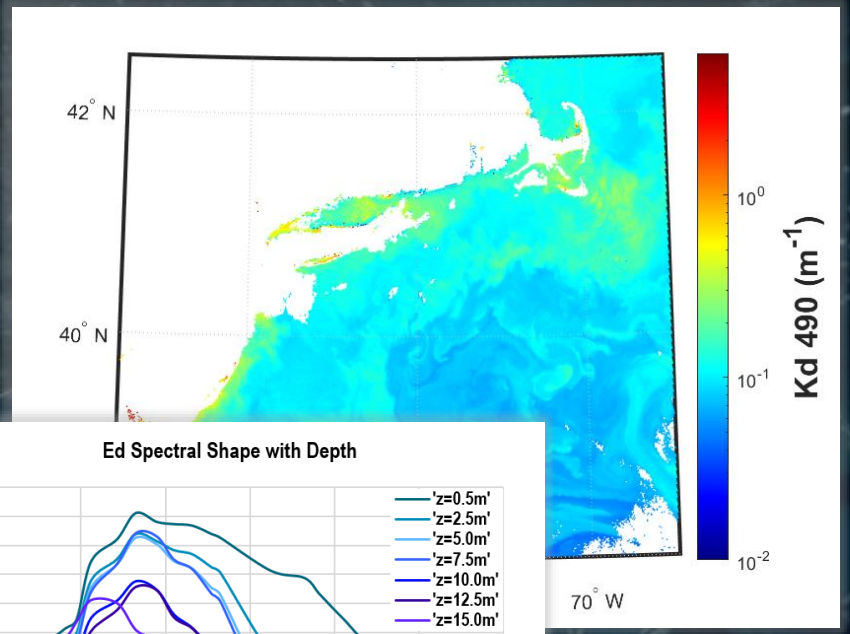
- AOP = Apparent Optical Properties
- Rrs - remote sensing reflectance at 184 wavelengths (339–719 nm)
- avw - apparent visible wavelength
- **nflh – normalized fluorescence line height**

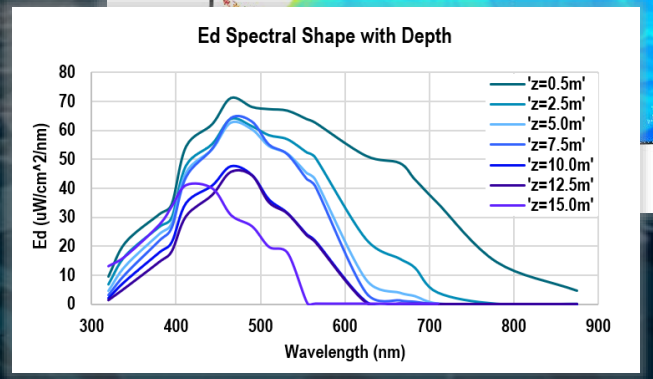


Note: Light leaving the ocean surface due to sun-induced chlorophyll fluorescence. Provides an indicator of phytoplankton physiology.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_IOP.V3_0.nc

- AOP = Inherent Optical Properties
- **Kd - diffuse attenuation coefficients at 19 wavelengths (351–711 nm)**
- a - total absorption coefficients at 19 wavelengths (351–711 nm)
- aph - phytoplankton absorption coefficients at 19 wavelengths (351–711 nm)
- adg - detrital and gelbstoff absorption coefficient at 442 nm
- adg_s - detrital and gelbstoff absorption spectral slope parameter
- bb - total backscatter coefficients at 19 wavelengths (351–711 nm)
- bbp - particle backscatter coefficient at 442 nm
- bbp_s - particle backscatter spectral slope parameter

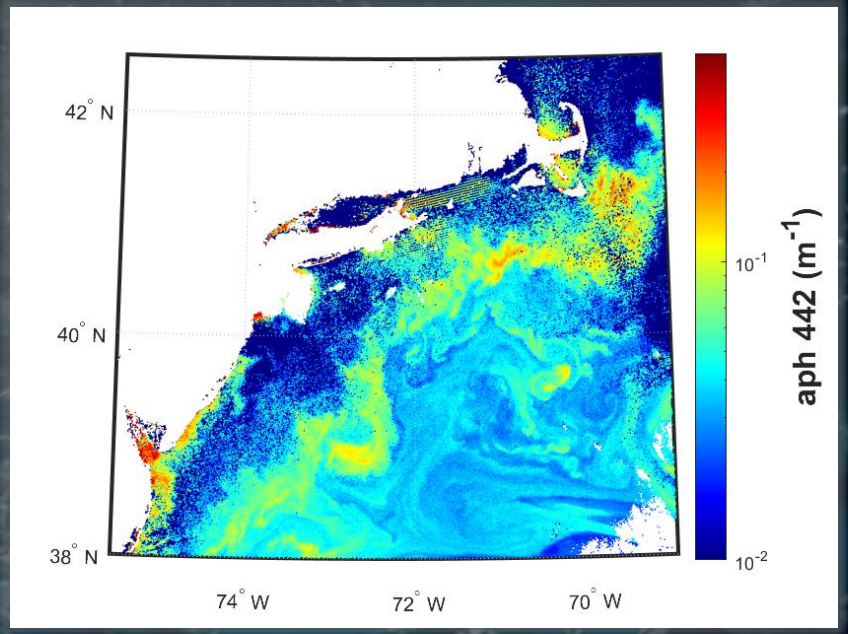




Note: A direct indicator of how deep light can penetrate into the water. These coefficients can be used to directly estimate light intensity at any depth.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_IOP.V3_0.nc

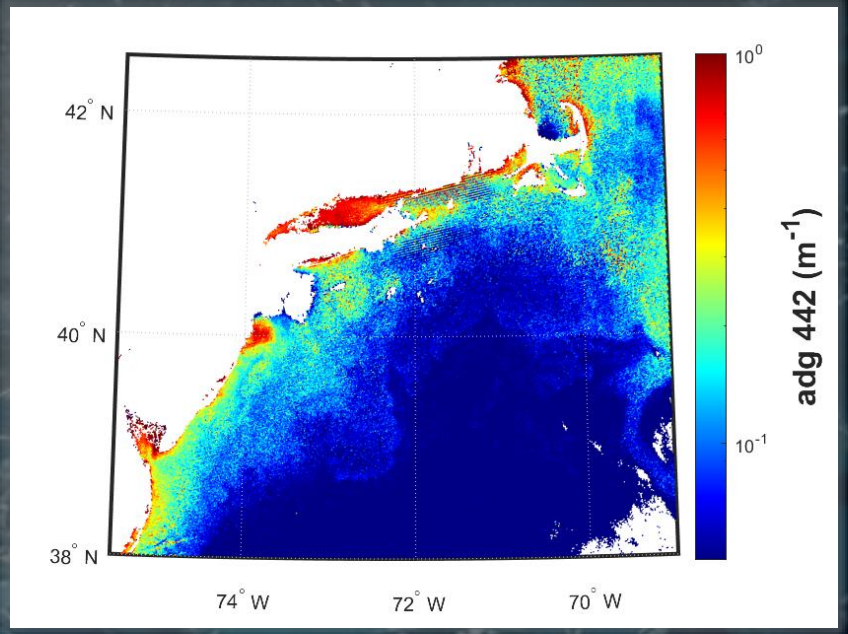
- AOP = Inherent Optical Properties
- Kd - diffuse attenuation coefficients at 19 wavelengths (351–711 nm)
- a - total absorption coefficients at 19 wavelengths (351–711 nm)
- **aph - phytoplankton absorption coefficients at 19 wavelengths (351–711 nm)**
- adg - detrital and gelbstoff absorption coefficient at 442 nm
- adg_s - detrital and gelbstoff absorption spectral slope parameter
- bb - total backscatter coefficients at 19 wavelengths (351–711 nm)
- bbp - particle backscatter coefficient at 442 nm
- bbp_s - particle backscatter spectral slope parameter



Note: The absorption of light by phytoplankton can vary by a factor of 4 or more at a constant chlorophyll-a value, so this parameter more accurately describes how much light has been utilized by living phytoplankton cells.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_IOP.V3_0.nc

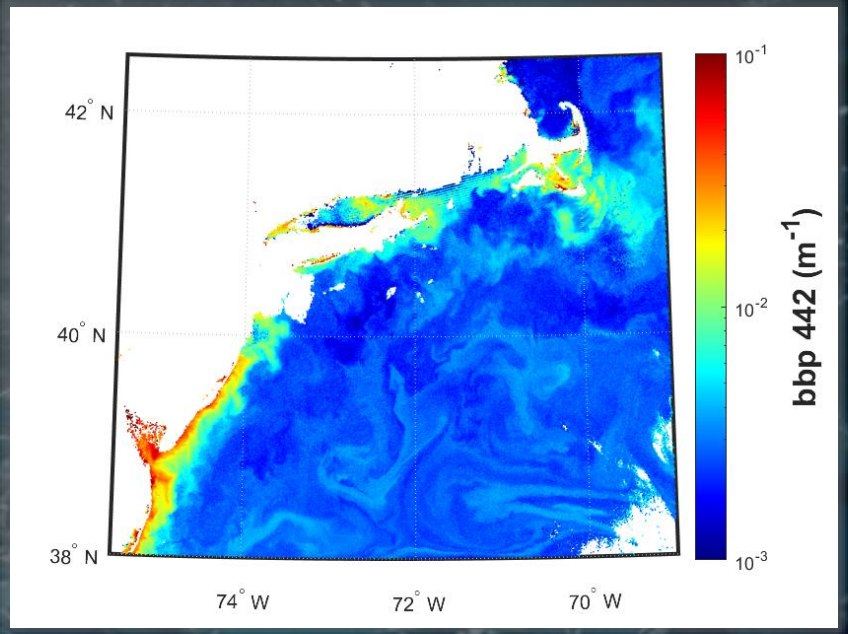
- AOP = Inherent Optical Properties
- Kd - diffuse attenuation coefficients at 19 wavelengths (351–711 nm)
- a - total absorption coefficients at 19 wavelengths (351–711 nm)
- aph - phytoplankton absorption coefficients at 19 wavelengths (351–711 nm)
- **adg - detrital and gelbstoff absorption coefficient at 442 nm**
- adg_s - detrital and gelbstoff absorption spectral slope parameter
- bb - total backscatter coefficients at 19 wavelengths (351–711 nm)
- bbp - particle backscatter coefficient at 442 nm
- bbp_s - particle backscatter spectral slope parameter



Note: Absorption of light by detritus + CDOM can indicate the presence of a declining phytoplankton bloom, or river runoff.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.OC_IOP.V3_0.nc

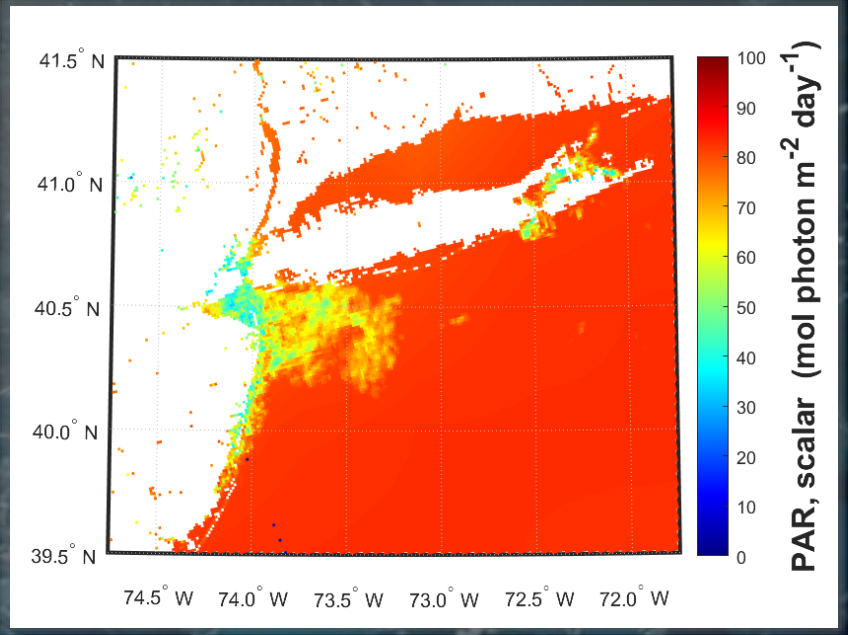
- AOP = Inherent Optical Properties
- Kd - diffuse attenuation coefficients at 19 wavelengths (351–711 nm)
- a - total absorption coefficients at 19 wavelengths (351–711 nm)
- aph - phytoplankton absorption coefficients at 19 wavelengths (351–711 nm)
- adg - detrital and gelbstoff absorption coefficient at 442 nm
- adg_s - detrital and gelbstoff absorption spectral slope parameter
- bb - total backscatter coefficients at 19 wavelengths (351–711 nm)
- **bbp - particle backscatter coefficient at 442 nm**
- bbp_s - particle backscatter spectral slope parameter



Note: Defines how light is scattered in the backwards direction by particles in the water. This product provides an indicator of particle density.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.PAR.V3_0.nc

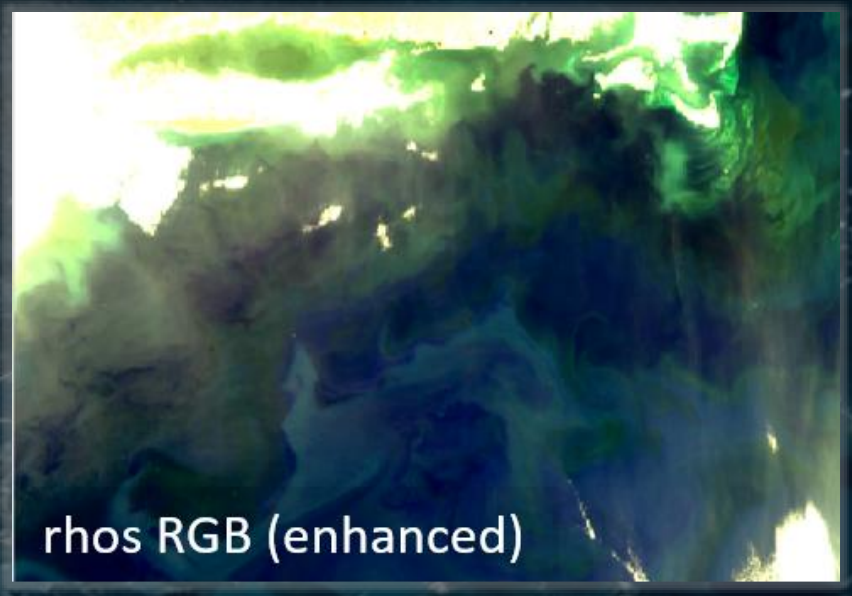
- PAR = Photosynthetically Available Radiation
  - ipar_planar_below - instantaneous photosynthetically available radiation (below water surface)
  - ipar_planar_above - instantaneous photosynthetically available radiation (above water surface)
  - **par_day_scalar_below - daily scalar photosynthetically available radiation (below water surface)**
  - par_day_planar_above - daily planar photosynthetically available radiation (above water surface)
  - par_day_planar_below - daily planar photosynthetically available radiation (below water surface)



Note: Quantum energy flux from the Sun in the 400 - 700 nm range. Impacts photosynthetic activity (how much light is actually converted to carbon)

PACE_OCI.[YYYYMMDD]T[HHMMSS].L2.SFREFL.V3_0.nc

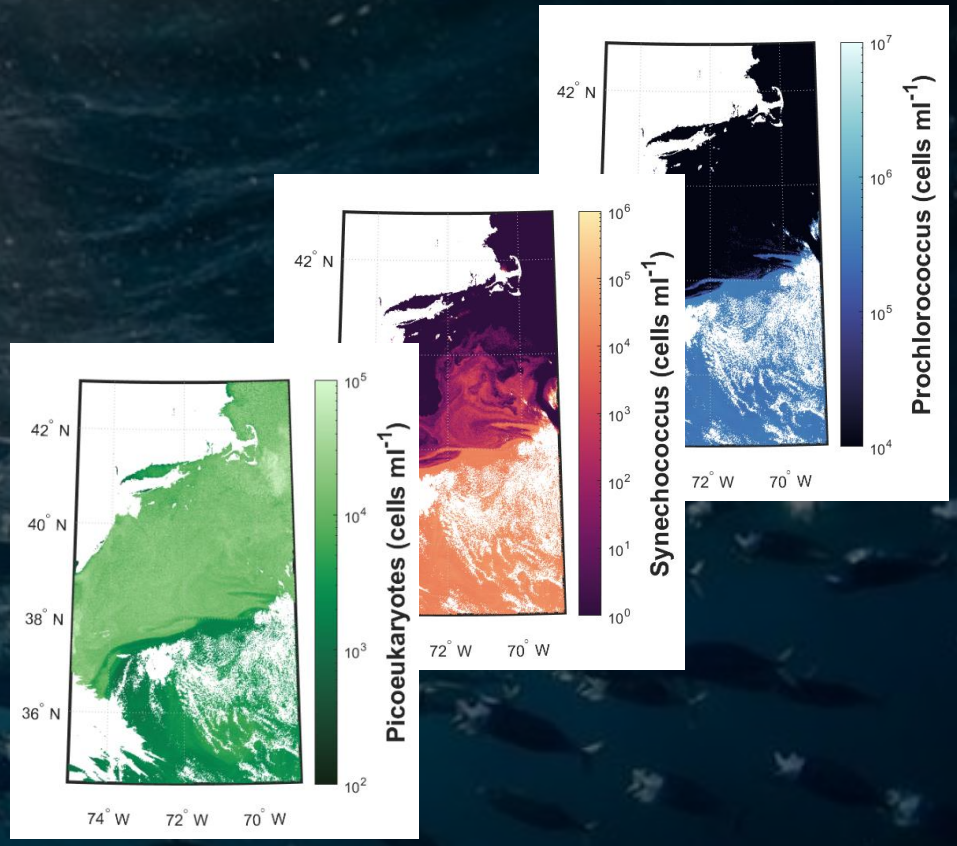
- SFREFL = Surface Reflectance
- **rhos – surface reflectance at 52 wavelengths (339–2258 nm), corrected for Rayleigh scattering**



Note: The basis for “true color” imagery. The is radiance corrected for the angular effects of Rayleigh scattering, as 90% of signal is atmosphere.

PACE_OCI.[YYYYMMDD]T[HHMMSS].L3m.MO.MOANA.V3_0.4km.nc

- MOANA = Multiple Ordination ANAlysis
- **prococcus_moana – concentration of prochlorococcus**
- **syncoccus_moana – concentration of synechococcus**
- **picoeuk_moana – concentration of picoeukaryotes**



Note: The first phytoplankton community composition algorithm for PACE. These products are intended for open ocean applications, not nearshore.

What else can we expect?

Coming soon

- **Pigments** (Chase): Phytoplankton pigments chlorophyll-a, -b, and -c along with photoprotective and photosynthetic carotenoids.
- **Diatom carbon** (Chase): Satellite-based diatom carbon estimates.
- **Taxonomic groups** (Kramer): Diatoms, dinoflagellates, nanoplankton, haptophytes, picoplankton, based on phytoplankton pigment estimates.
- **Carbon, Absorption, and Fluorescence Euphotic-resolving (CAFÉ)** model (Silsbe et al. 2016): Absorption-based estimates of net primary productivity.

Available, but not operational

- **Particle size class** (Kostadinov): Size partitioning of oceanic particles, particle size distribution.
- **Phytoplankton size class** (Turner): Chlorophyll-a based partitioning of phytoplankton size classes (pico-, nano-, micro-plankton).

Applications of hyperspectral ocean color data for aquaculture and fisheries management

Ryan Vandermeulen, Veronica Lance, Clete Otoshi, Ken Riley, Michelle Tomlinson, and Rebecca Trinh

NOAA Technical Memorandum NMFS-F/SPO-###
July 2024

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Oceanic and Atmospheric Administration
Richard W. Spinrad, NOAA Administrator

National Marine Fisheries Service
Janet Coit, Assistant Administrator for Fisheries

Table of Contents

Product 6: Spectral particle backscattering coefficients (b_{bp})

What is it?

- These backscatter coefficients specifically define how light is scattered in the backwards direction by particles in the water. This product provides an indicator of the concentration of particles in the ocean and a proxy indicator of particulate carbon concentrations.

How does it impact Aquaculture/Fisheries?

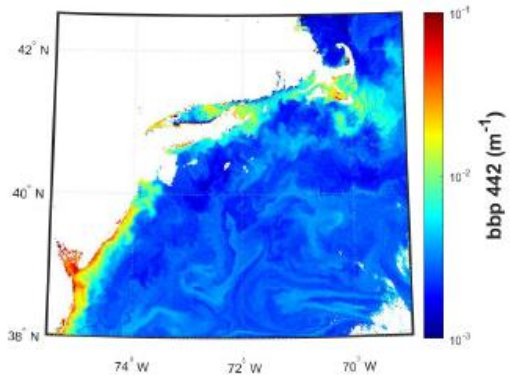
- Many phytoplankton exhibit unique backscattering characteristics[65], primarily as a function of cell size[66], and sometimes composition (e.g. Coccolithophore blooms[67]). Backscatter is also used to derive Particulate Inorganic Carbon (PIC)[68] estimates from satellites. Particle backscatter is a particularly useful tool to determine high sediment loads in nearshore environments, which tends to heavily scatter light. High sediment loads can cause gill saturation[69] in certain oyster species, and some fish species exhibit hypersensitivity to suspended sediment[70]. While not a direct measurement of suspended particulate matter (SPM), it can be used to develop those products.

What are the limitations/caveats?

- The backscatter product is one of the most robust products offered in the "inherent optical property (IOP)" suite of ocean color products. The only caveat is that the IOP algorithms can sometimes fail to arrive at a solution (i.e. no data) in waters with extreme scattering or CDOM concentrations. Nearshore environments present challenges in disentangling phytoplankton backscatter from other optical constituents (re-suspended sediment, SPM, etc.).

Does HYPERSPECTRAL directly improve/enable this product?

- Operational improvements to IOP products using hyperspectral data are anticipated, but still in development (at the time of this publication). PACE Science and Applications Team members are actively working to improve this product using new approaches and techniques.



Product(s) 2: Harmful Algal Blooms (HABs)

What is it?

- Harmful algal blooms, or HABs, occur when colonies of phytoplankton produce toxic or harmful effects on people, fish, shellfish, marine mammals and birds. Various spectral techniques have been used to remotely identify different HABs from ocean color[100]. While not all-inclusive, some examples of HAB detection are listed below.

  - ***Microcystis aeruginosa*** (CyAN[101]): Freshwater algae that can produce a toxin known as microcystin, which causes fish kills and contamination of drinking water.
  - ***Karenia brevis*** (Craig[102], Soto[103]): Ubiquitous red tide species occurring on the Florida coast, causing fish kills and human respiratory issues.
  - ***Pseudo-nitchzia*** (Anderson[104], Smith[105]): Diatom that produces domoic acid, accumulates in shellfish, invertebrates, and sometimes fish, leading to mammal illness and death.
  - ***Alexandrium cantenella*** (Bucci[106]): Dinoflagellate that produces a saxitoxin and causes Paralytic Shellfish Poisoning (PSP).
  - ***Cochlodinium polykrikoides*** (Ahn[107], Kim[108]): Dinoflagellate causing "rust tides" that are toxic to finfish and shellfish.
  - ***Noctiluca scintillans*** (Qi[109]): Large dinoflagellate that can cause disruptions to trophic energy dynamics, even impacting fish yield[110].
  - **Floating algae index** (Hu[111], Sargassum Watch[112]): Used to detect surface slicks, including nuisance algae such as *Sargassum*.
  - **Red-band difference, RBD** (Amin[113]): A generalized indicator frequently used to detect a variety of HABs.
  - **Maximum chlorophyll index, MCI** (Gower[114]): A generalized indicator frequently used to detect high biomass blooms.
  - **Regional Forecast systems:** NCOOS and external partners supply operational forecast systems for various regions of the U.S., including Gulf of Mexico & Florida[115], Gulf of Maine[116], Lake Erie[117], Pacific Northwest[118], and California[119].

How does it impact Aquaculture/Fisheries?

- The impact of HABs can be economically and ecologically disruptive, owing to direct mortality of fish and marine mammals, seafood contamination and crop loss, fisheries and aquaculture closures, trophic-food web disruptions, drinking water contamination, human health impacts, clean-up costs, tourism losses, and even long-term losses in property values.

What are the limitations/caveats?

- With some exceptions, these classes of algorithms and products are not typically supplied by agencies on an operational basis, leaving the burden of implementation on the user. It should be considered that some toxic species are not always producing toxins, and that some species may become toxic at concentrations below detection limits. Tracking of HABs from satellites is often most effective when paired with in situ ground verification and monitoring.

Does HYPERSPECTRAL directly improve/enable this product?

- Hyperspectral data enables the detection of subtle pigment signatures associated with specific phytoplankton, and can thus help determine the likelihood of toxicity.

NOAA Fisheries | U.S. Department of Commerce

Thank you!

Contact me at:
Ryan.Vandermeulen@noaa.gov