Phytoplankton community composition derived from optics and remote sensing: Approaches, challenges, and next steps

Ali Chase, University of Washington Applied Physics Laboratory, Seattle, WA USA
Ocean Optics Summer Course – 5 July 2023, Bowdoin College, ME

Contact: alichase@uw.edu

Slide content inspired by and with info from Jeremy Werdell (NASA), Julia Uitz (LOV), Patrick Gray (UMaine), & many colleagues and papers (see references at the end)

Week 4 roadmap

All week: continue working on projects

Monday: Monte Carlo techniques (lecture/interactive demonstration)

Mon: Arduino lab

Tuesday: Calibrating bb sensors

Tuesday late afternoon discussion: Scientific ethics and being part of a community

Wednesday: Remote sensing of zooplankton

Wednesday: Phytoplankton community and particle size inversion from remote sensing

Wednesday after dinner discussion: Career panel with instructors

Friday: Final presentations of projects

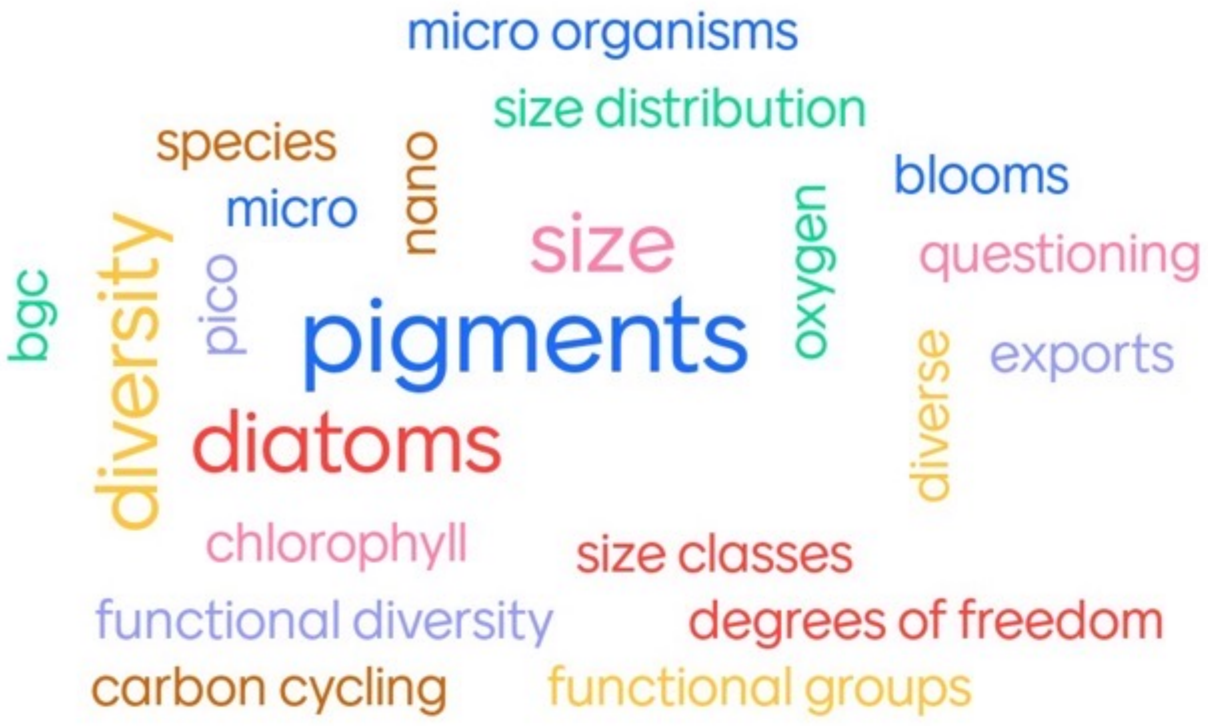
Who am I?

Overarching research goal: How are phytoplankton communities distributed in space and time? At various scales, what changes are occurring in these communities and their distributions?

→ use optical measurements to estimate parameters related to phytoplankton
→ application to remote sensing data for broad scale ocean ecosystem studies

Mentimeter

What comes to mind when you hear "Phytoplankton Community Composition"?



Responses from the 2023 Ocean Optics Class students

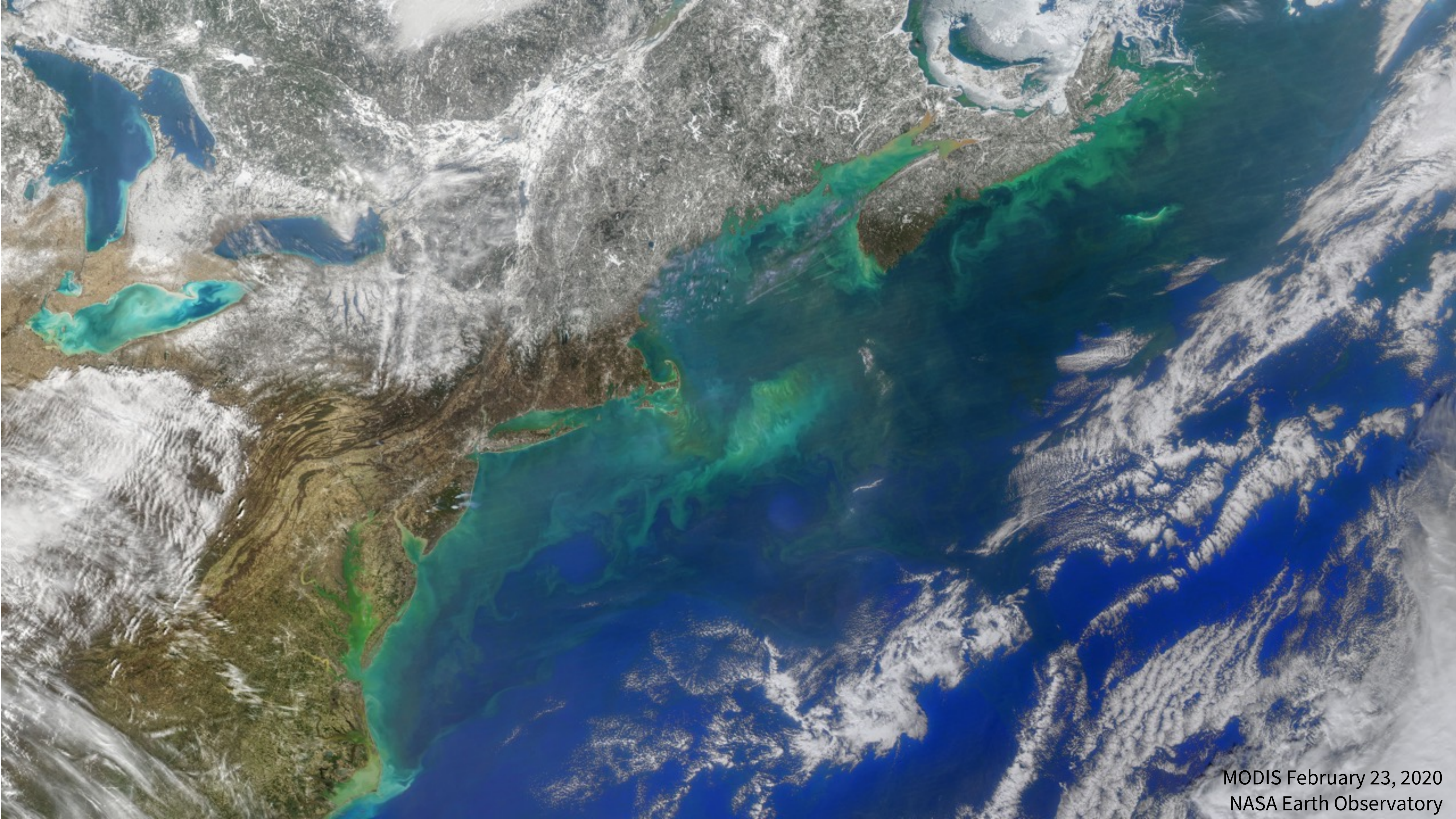
19

What comes to mind when you hear "Phytoplankton Community Composition"?

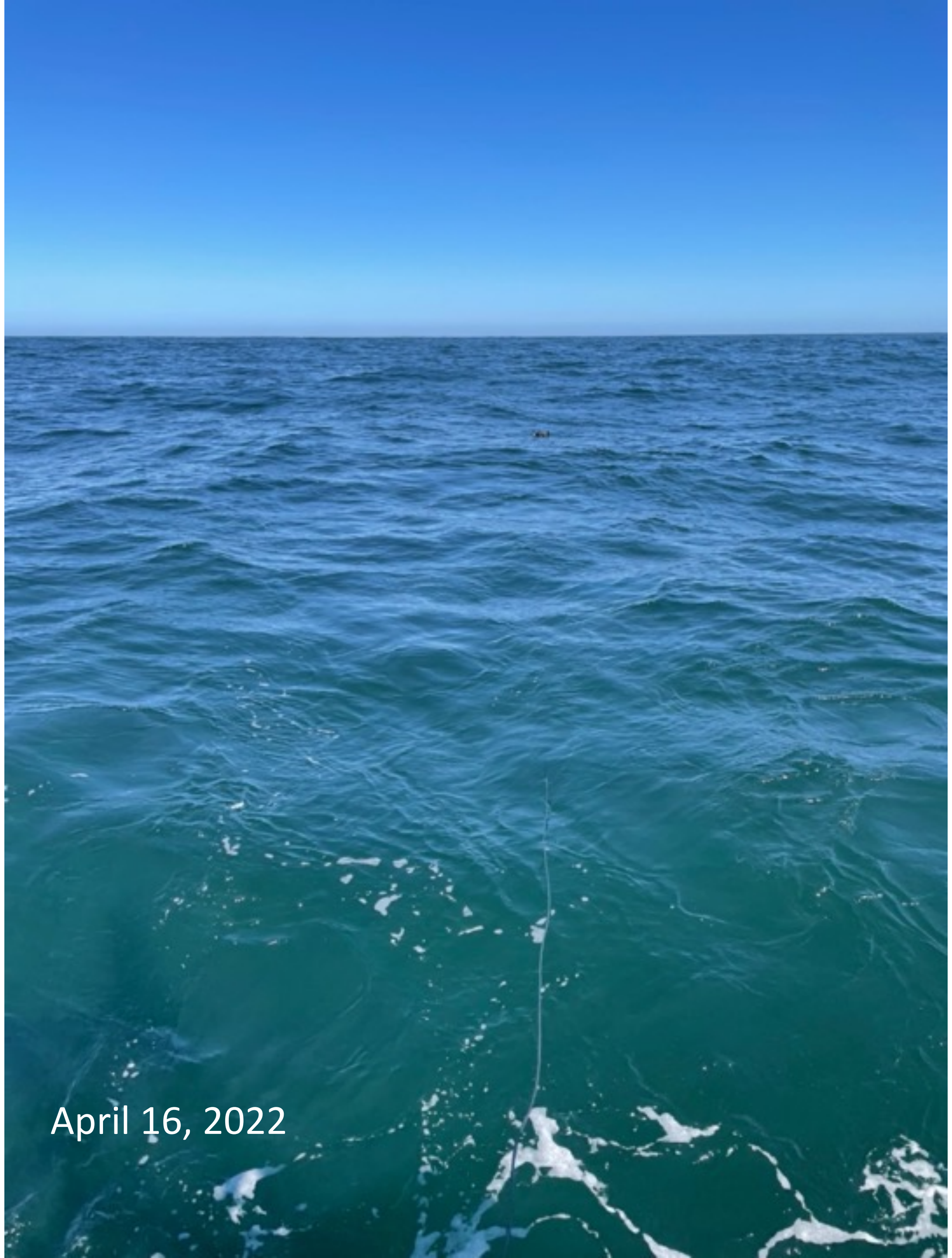
Mentimeter

19

Responses from the 2022 IOCCG Summer Lecture Series students

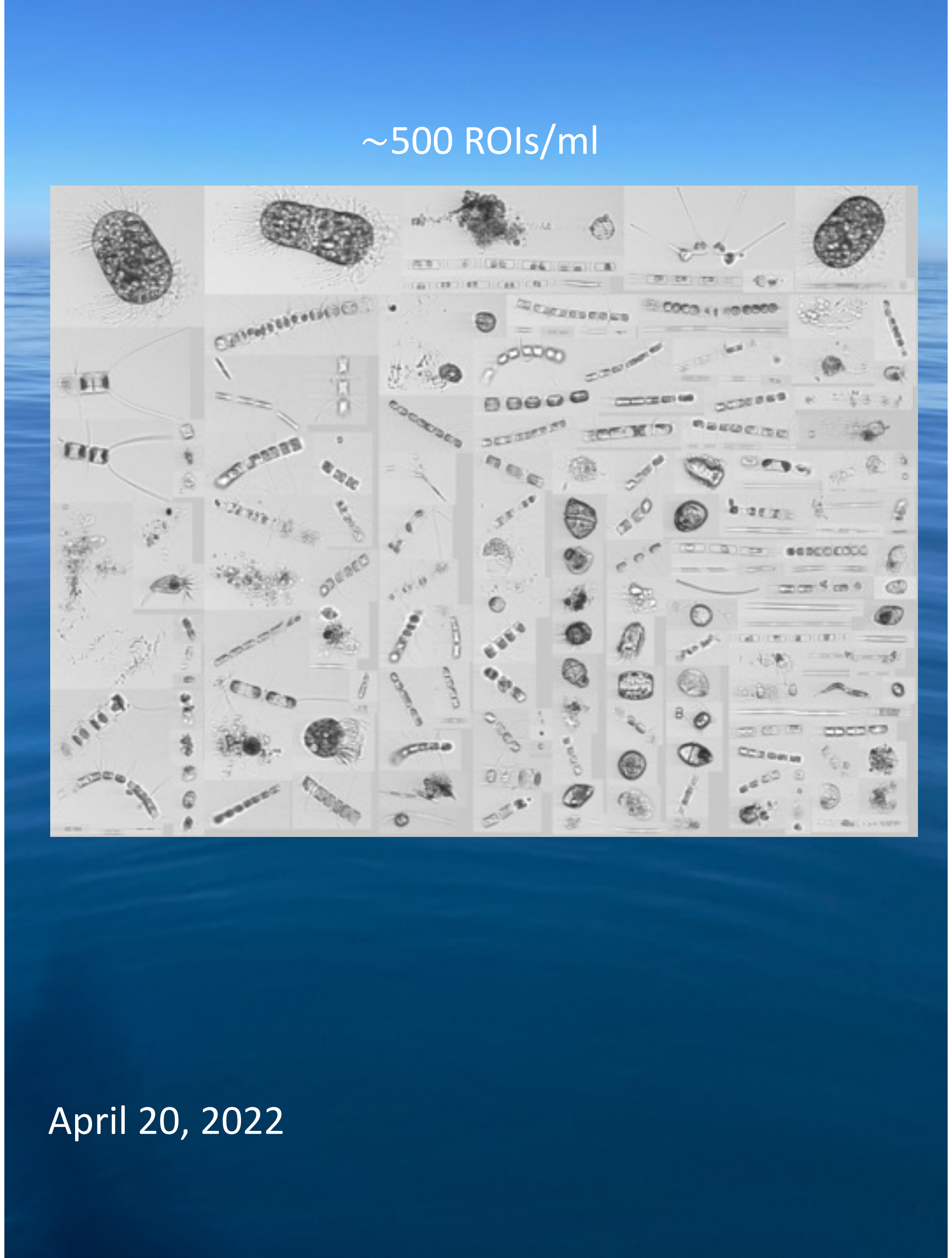


MODIS February 23, 2020
NASA Earth Observatory


April 16, 2022

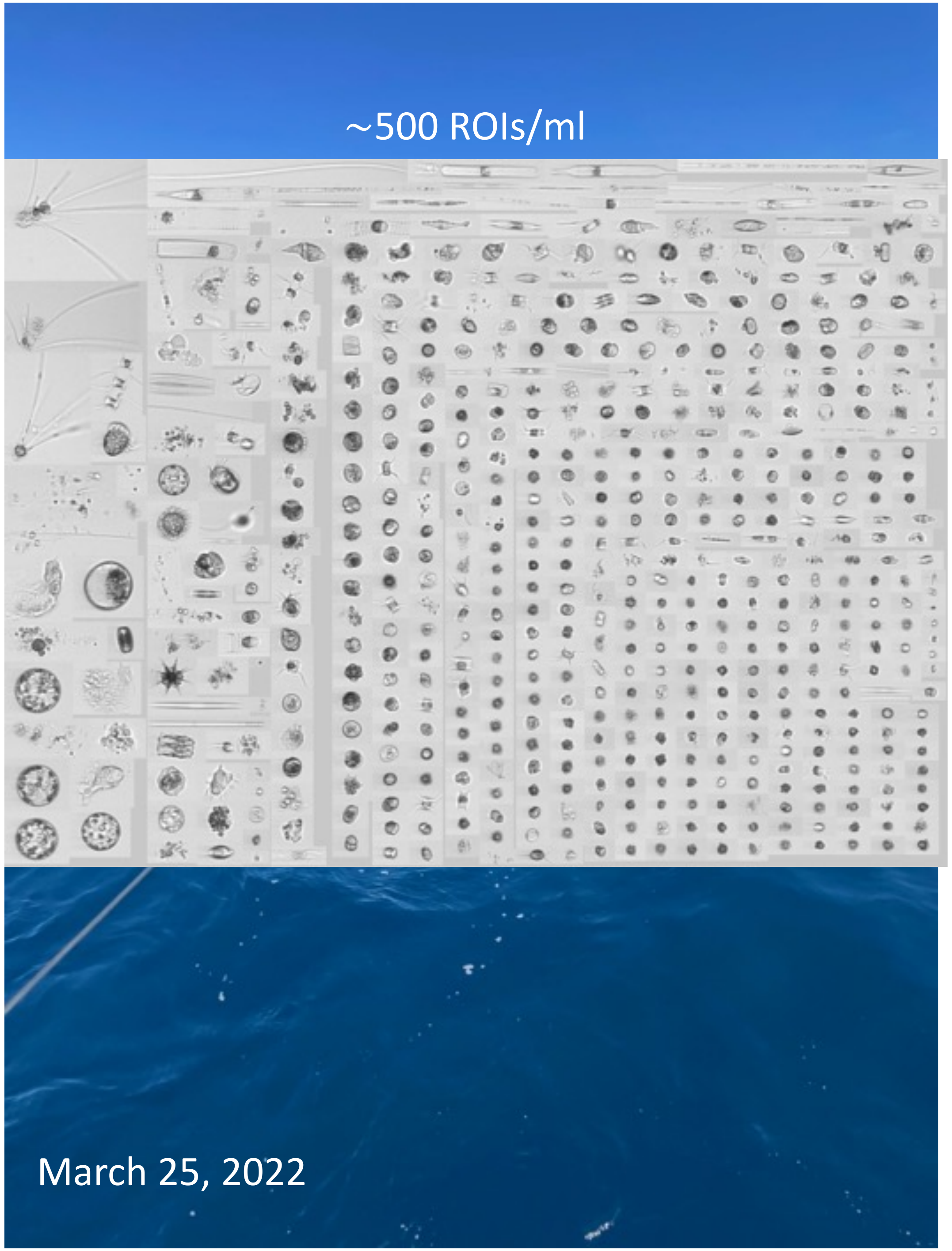

April 20, 2022

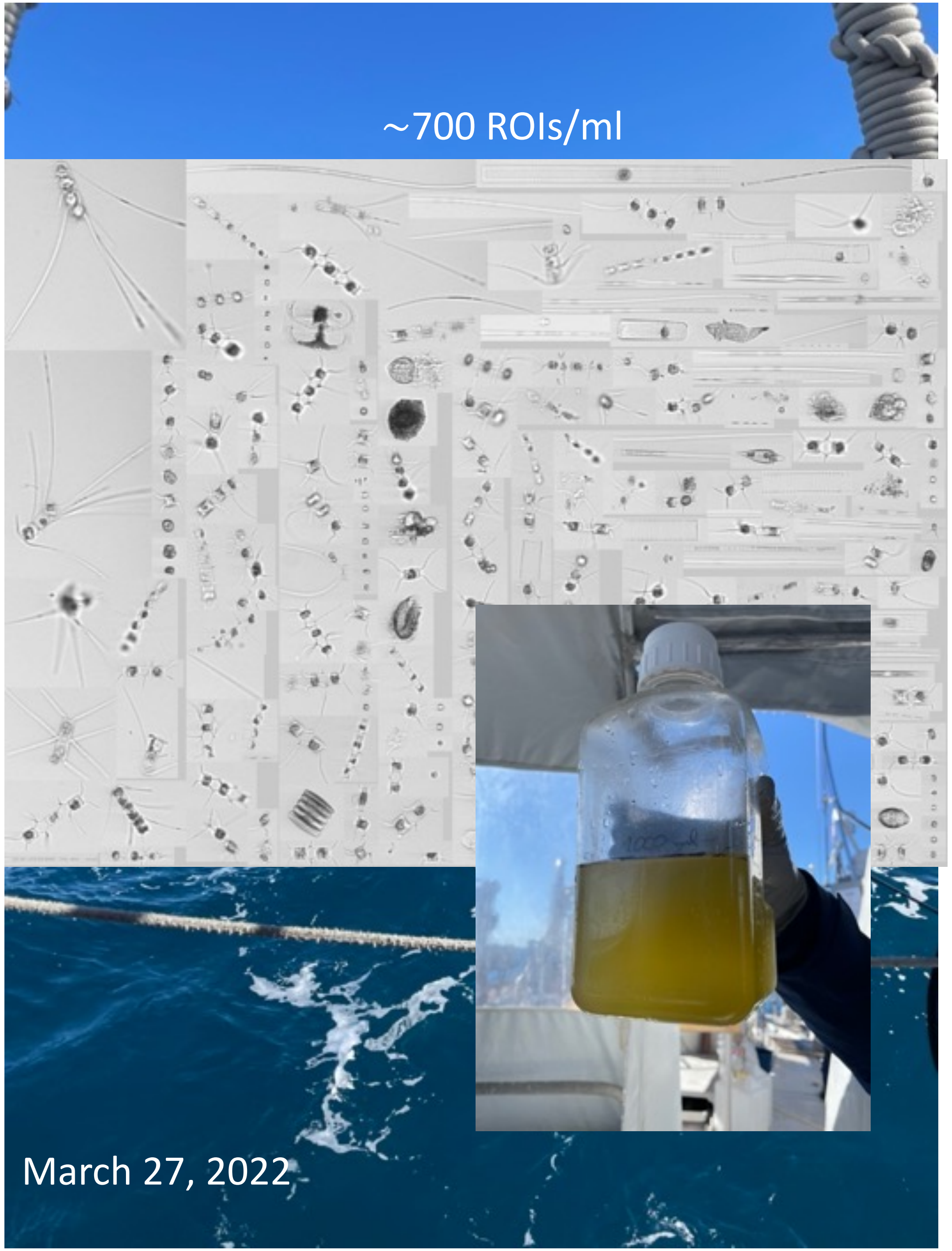

~3,000 ROIs/ml
April 16, 2022


~500 ROIs/ml
April 20, 2022

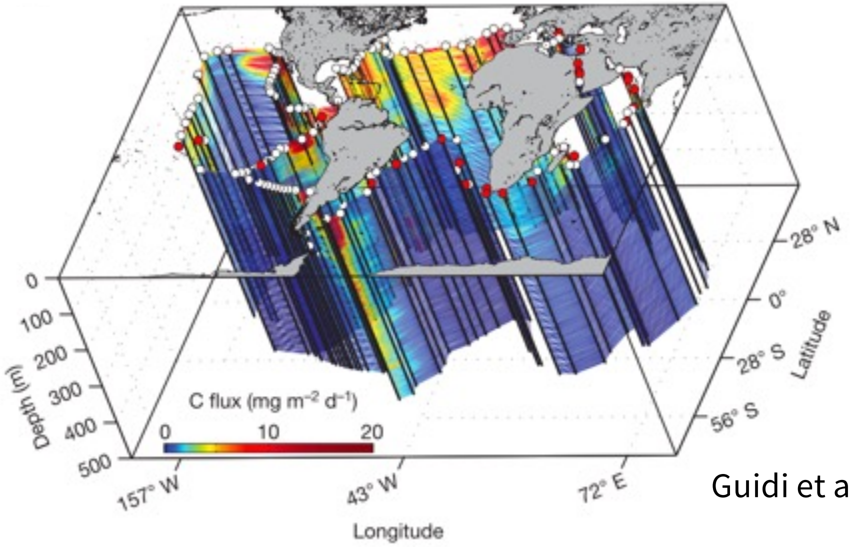

March 25, 2022


March 27, 2022


~500 ROIs/ml
March 25, 2022

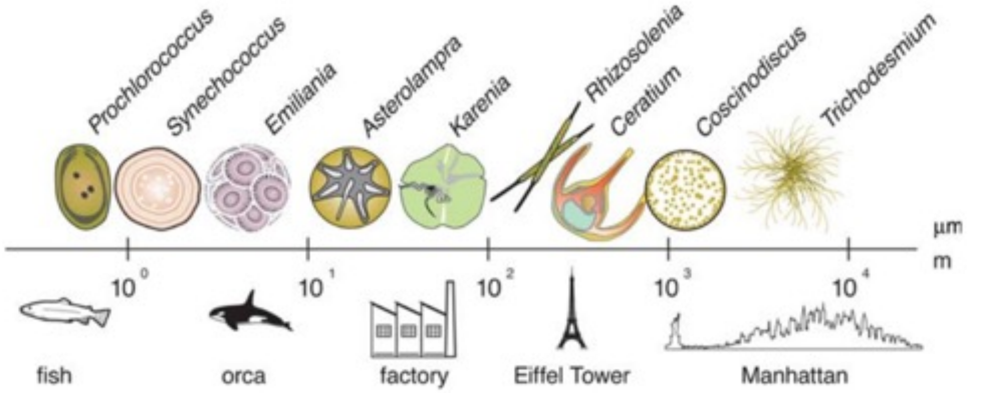

~700 ROIs/ml
March 27, 2022

What motivates our interest in phytoplankton community composition (PCC)?



Guidi et al. 2016

Variability in carbon flux & food web dynamics



Finkel et al. 2010

★ = PCC plays a role

Dierssen et al. 2021

How is Phytoplankton Community Composition defined?

PSC = Phytoplankton Size Classes (note: PSC also = Photosynthetic Carotenoids…)

- pico, nano, and micro (what should the size cutoffs be?)

PG = Phytoplankton Groups

- a catch-all terms for species and size classes?

PFT = Phytoplankton Functional Types

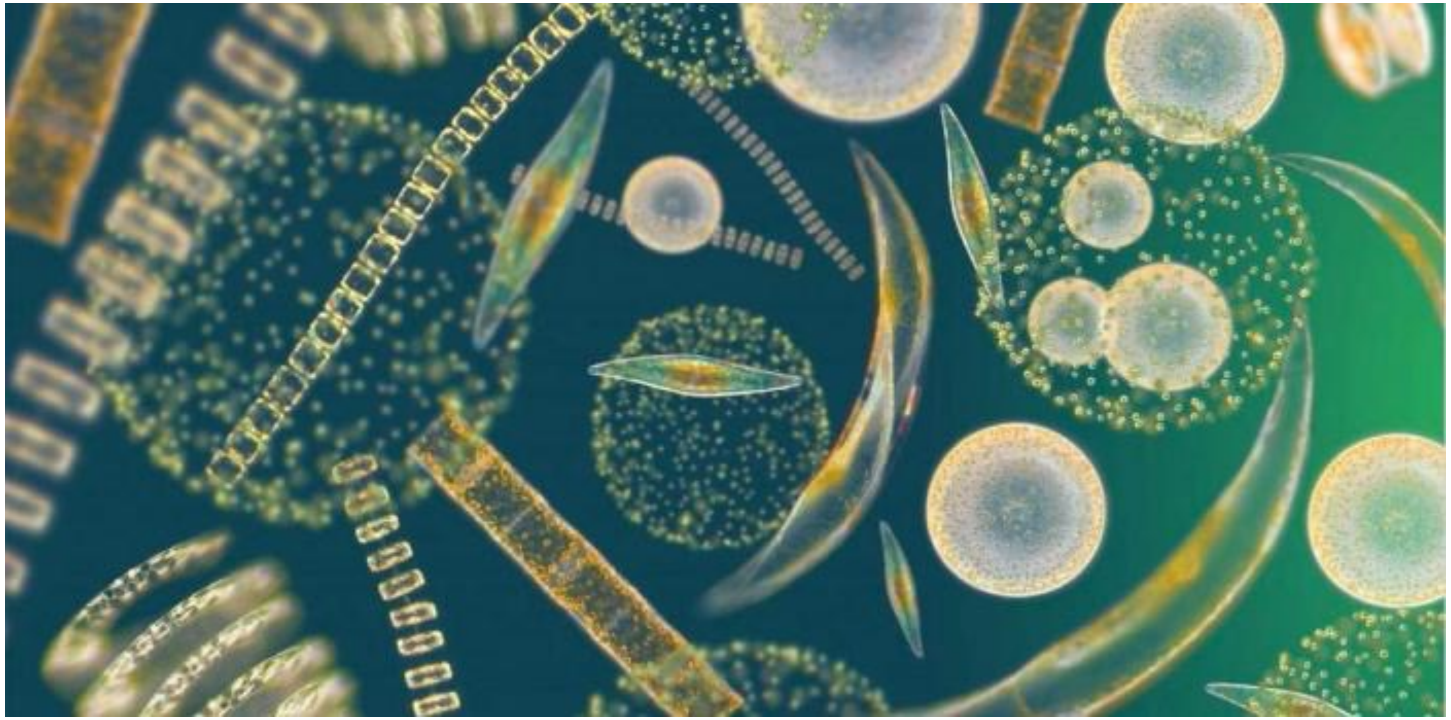
- biogeochemical function?

*Buyer Beware: The meanings of **all** these terms may change based on the user*

Bottom line: we want to define the phytoplankton present in the water by some metric that differs from/moves beyond total biomass (most commonly approximated via estimates of chlorophyll *a* concentration)

How is Phytoplankton Community Composition measured?

- Microscopy
- Pigments
- Flow cytometry
- Automated imagery
- Merged size spectra
- Optical signatures
- Molecular methods



→ see lecture by Sasha Kramer from week 1 for a very nice detailed overview:
https://misclab.umeoce.maine.edu/OceanOpticsClass2023/assets/pdfs/SJK_phytoplankton_OO23.pdf

And what about units???

Absolute

- Concentrations (cells/L)
- Biovolume (mg/m^3)
- Biomass, carbon (mg/m^3)
- Chl *a* (micrograms/L, mg/m^3)

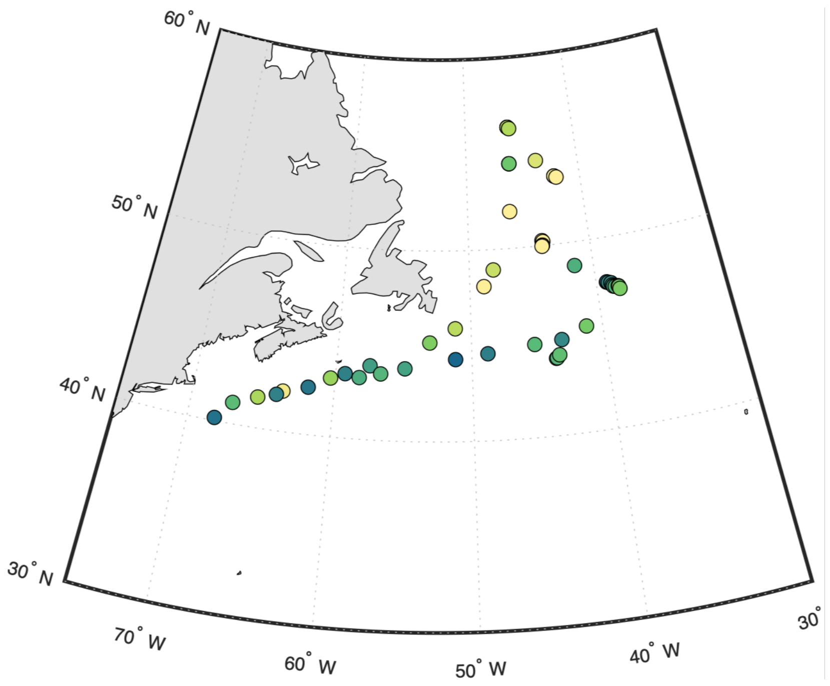
Relative

- Fraction (%) of total Chl *a*
- Fraction (%) of total biovolume
- Fraction of some subset of the total community (e.g., % of all microplankton)
- “Dominant” group (in what units?)

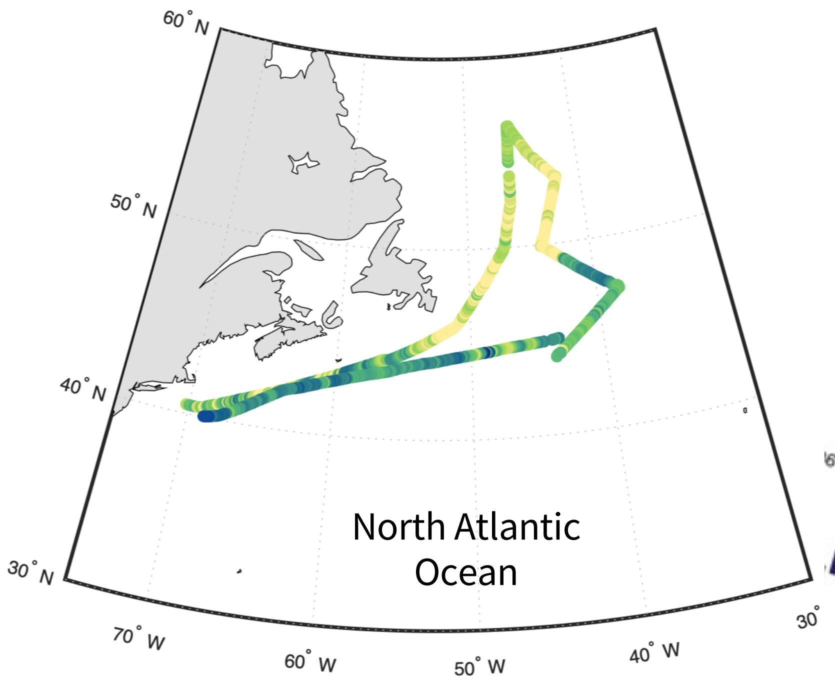
Probability of occurrence (at some threshold?)

How can we scale up to regional and global views?

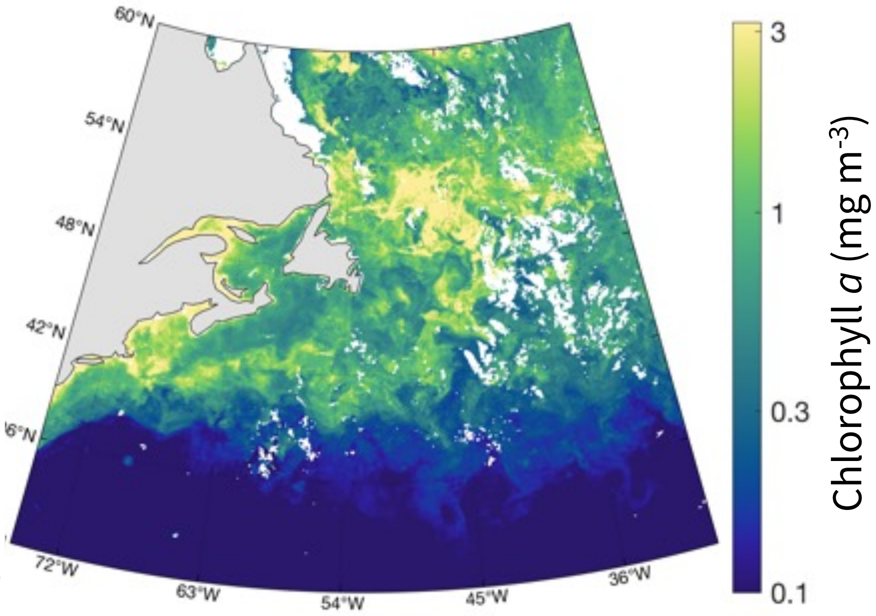
Discrete samples



Ship-board continuous measurements

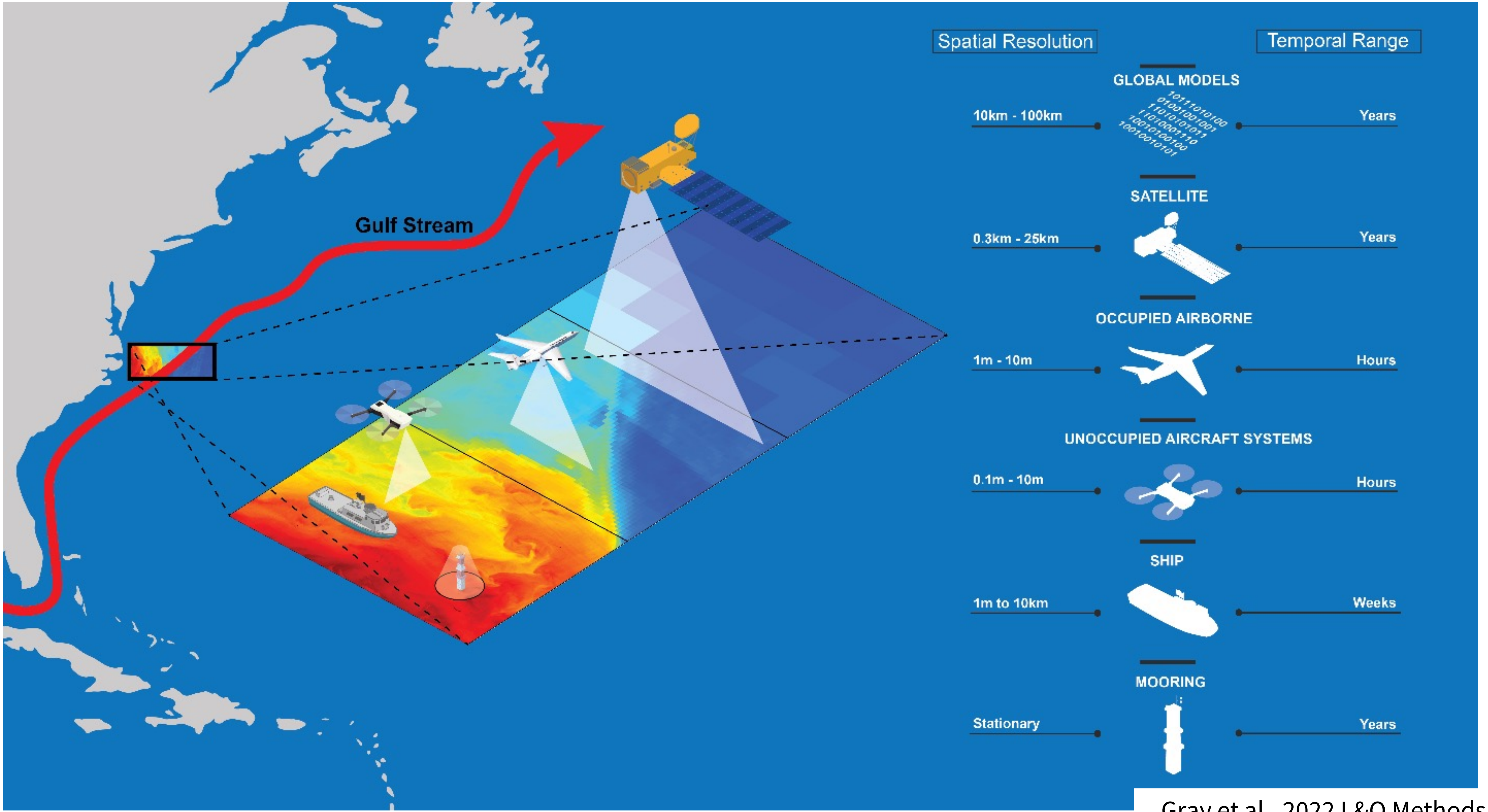


Satellite remote sensing



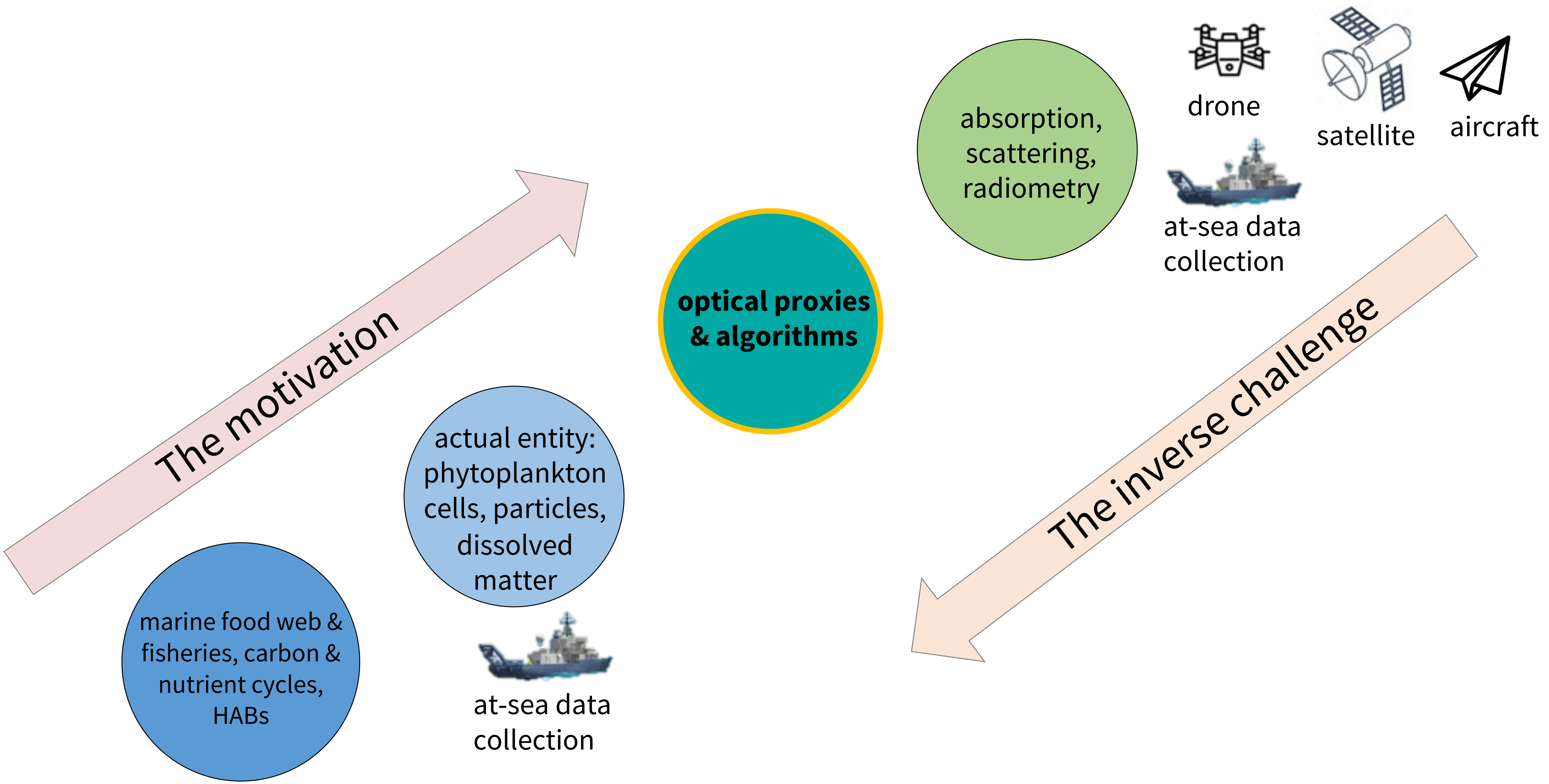
Direct measurement of cells & processes

Spatial & temporal resolution*



Gray et al., 2022 L&O Methods

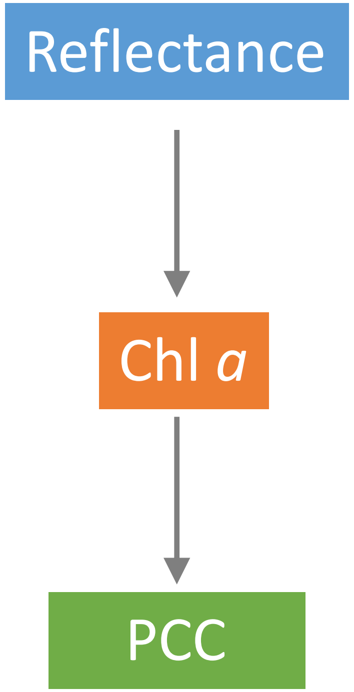
Optics: a tool to link what we can measure to what we want to know



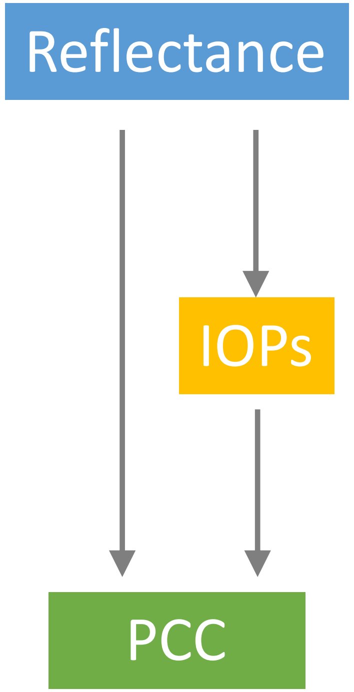
Adapted from M. J. Perry

Previously developed algorithms: two main categories

Abundance-based



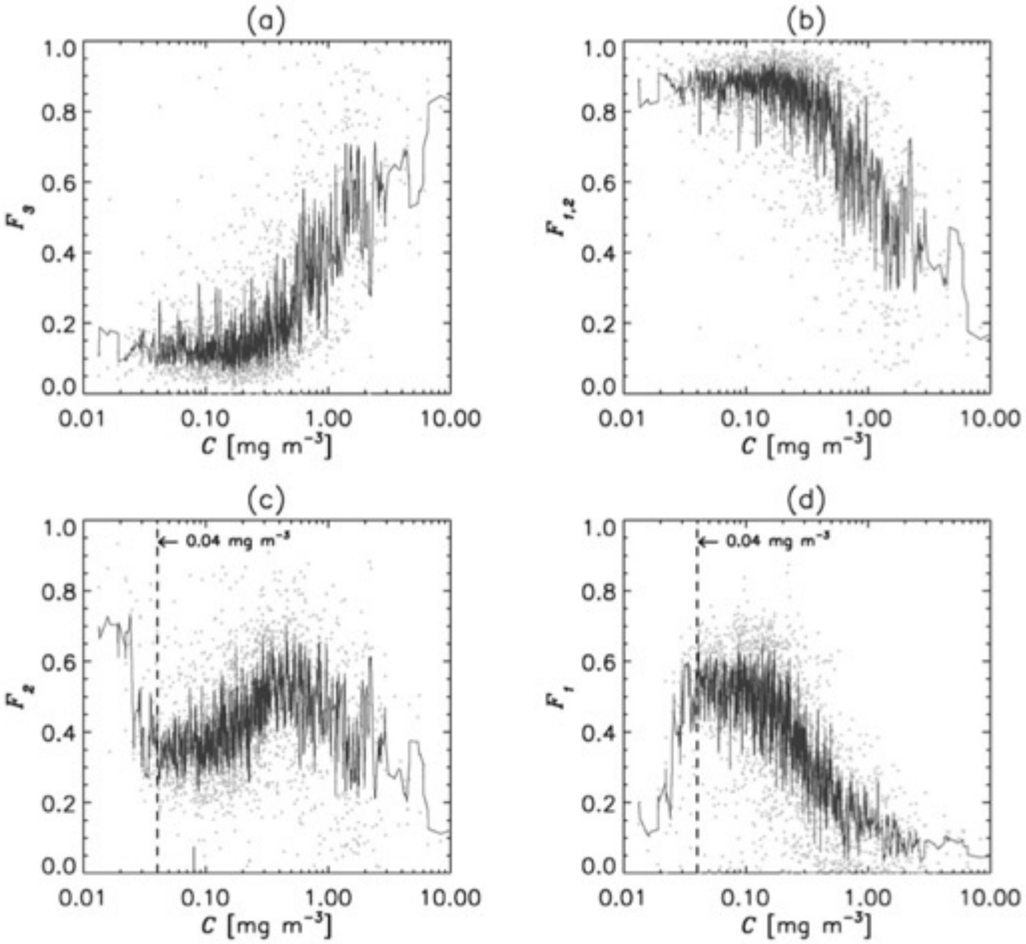
Spectral-based



Abundance-based - PSCs

Micro

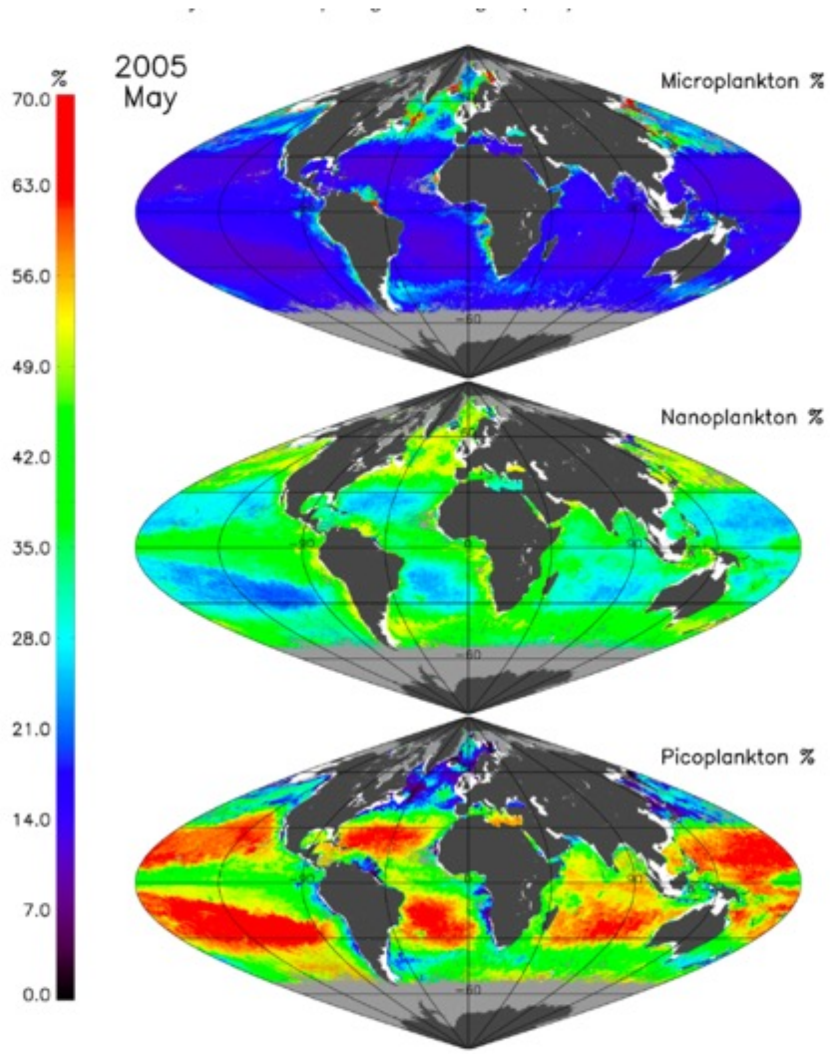
Pico+
Nano



Nano

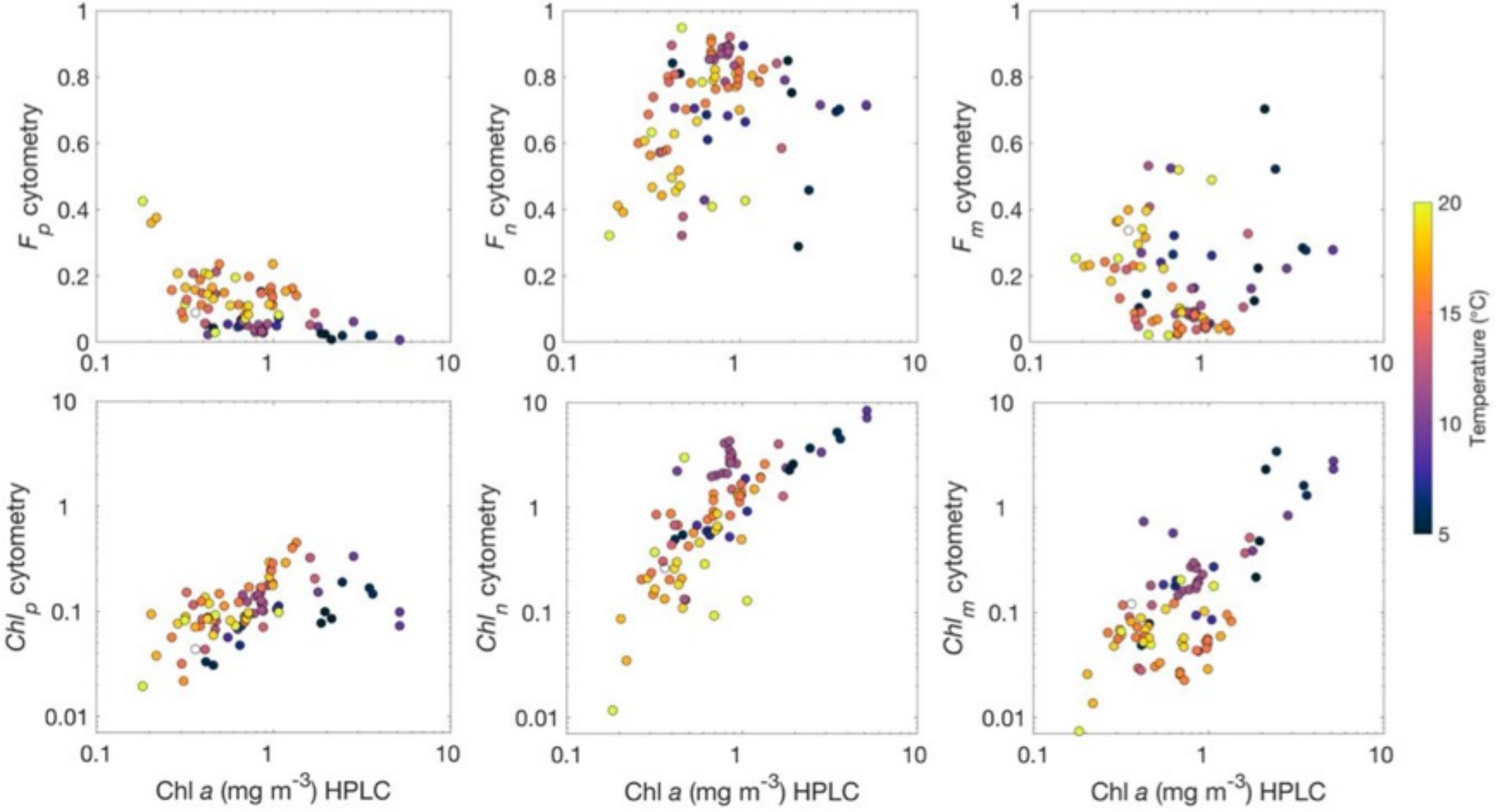
Pico

x-axes = [Chl *a*]



Chl *a* concentrations are related to fractions of pico-, nano-, and microplankton **as defined by pigments**

Brewin et al., 2010



Chl *a* concentrations vs. fraction of pico-, nano-, and microplankton from cytometry

Chase et al., 2020

Abundance-based – taxonomic groups

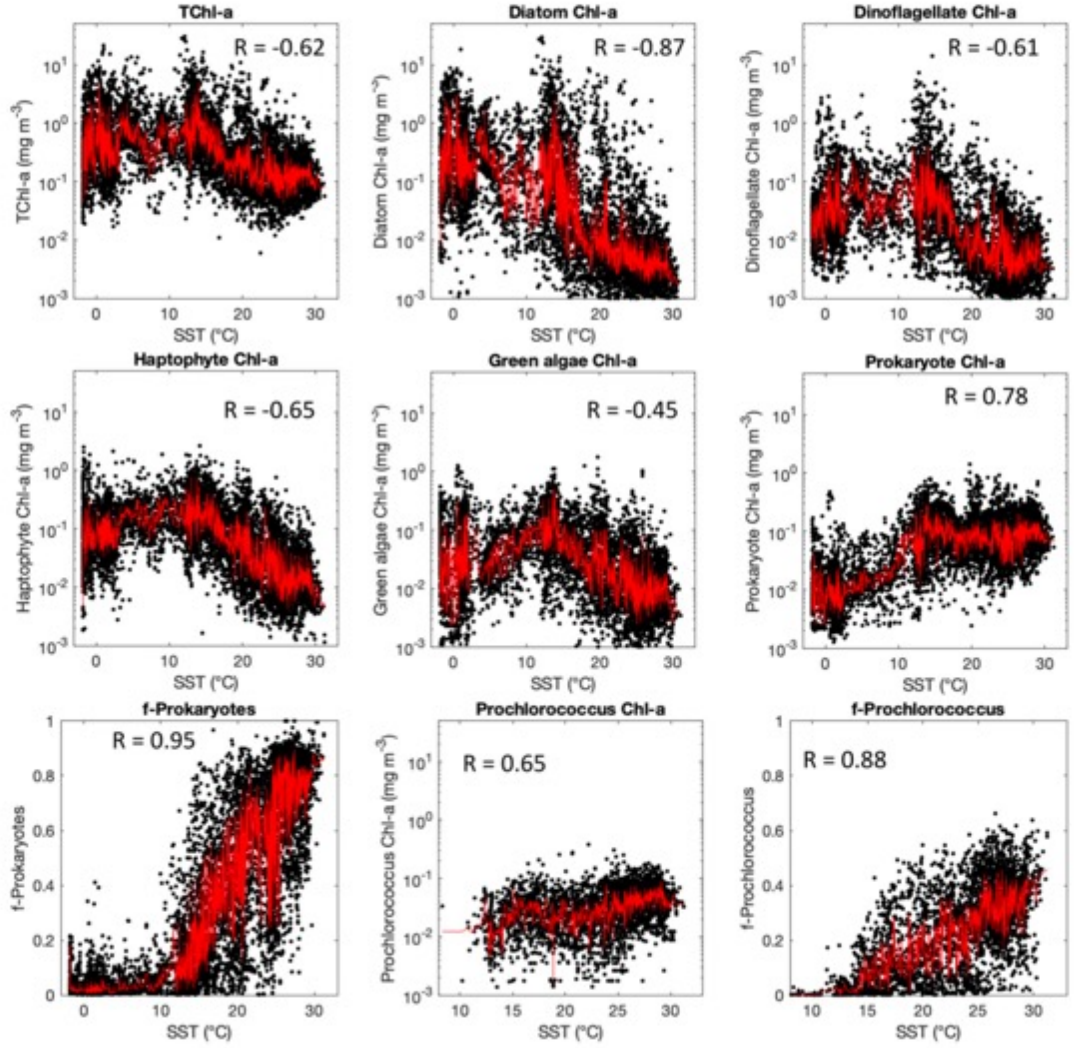
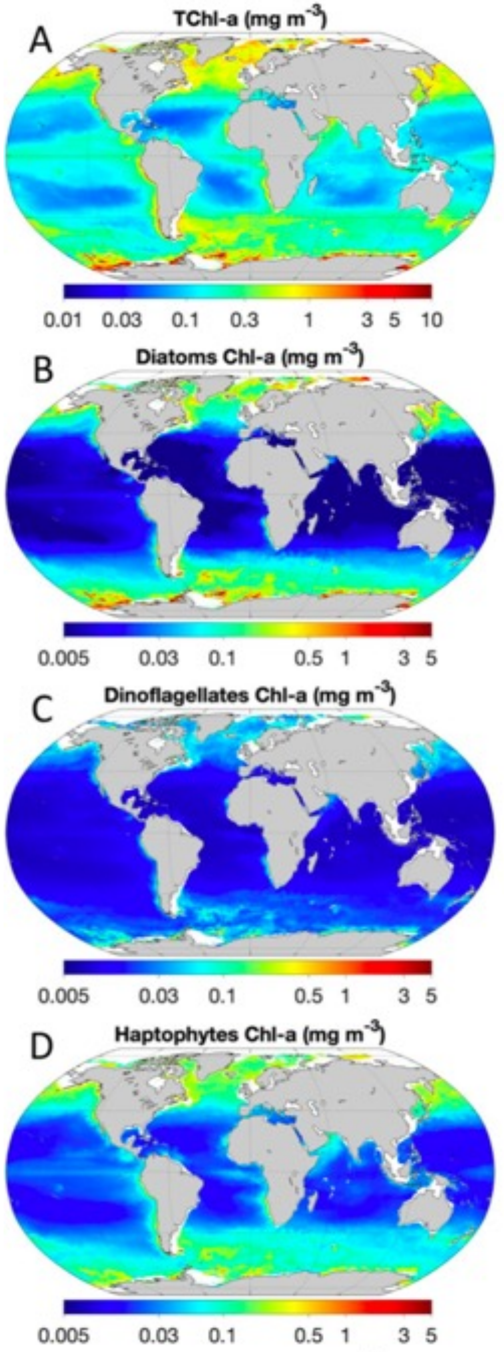


Figure 3. Scatterplots of in situ TChl-a, PFT Chl-a, and fractions of prokaryotes and *Prochlorococcus* versus collocated satellite SST data. The correlation coefficient (*R*) was calculated based on the 10-point running mean (red curve). PFT, phytoplankton functional type; SST, sea surface temperature.



Chl *a* concentrations related to phytoplankton groups **as defined by pigments**

Xi et al., 2021

Spectral-based - PSCs

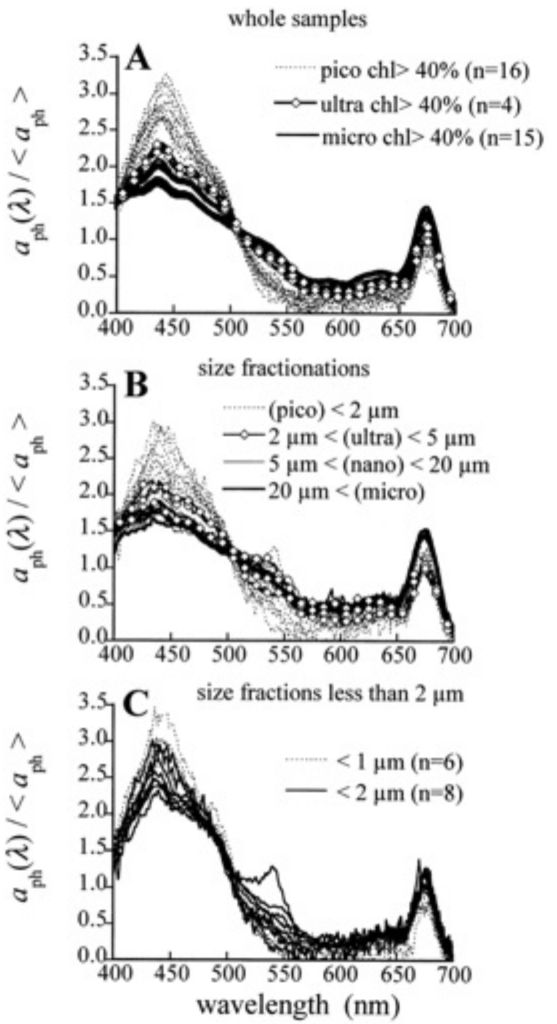


Fig. 3. (A) Phytoplankton absorption spectra measured during the Oregon cruise on whole surface samples. Spectra were normalized to the average phytoplankton absorption between 400 and 700 nm. Legend indicates the dominant Chl *a* size fractions (defined in Fig. 1). (B) Normalized phytoplankton absorption spectra measured directly for different size fractions collected on GF/F filters. (C) Comparison of absorption for size fractions from 14 stations using either 1- or 2-µm filters.

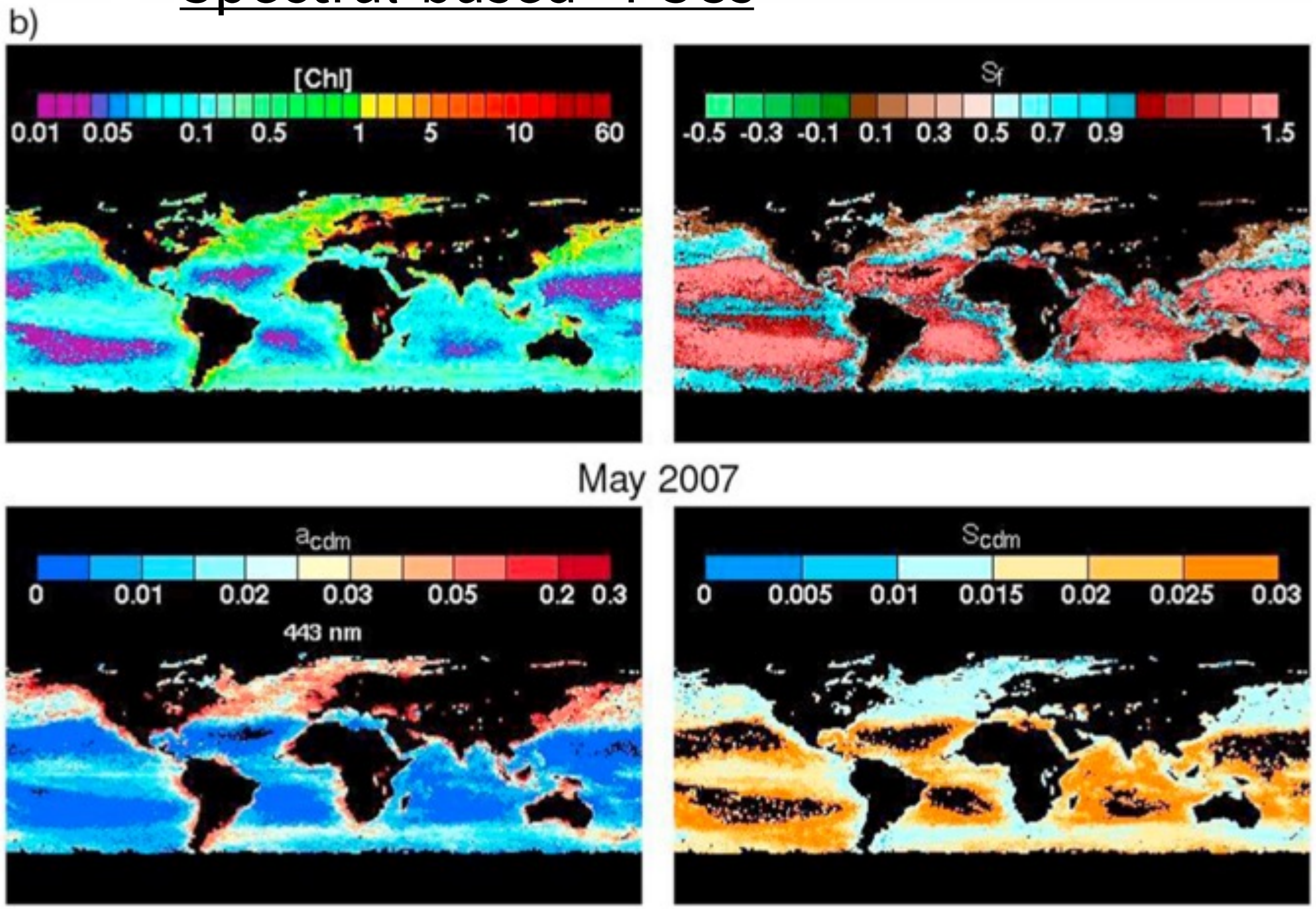


Figure 1. Monthly maps derived from the SeaWiFS monthly composite of reflectance for (a) February 2007, (b) May 2007, (c) August 2007, and (d) November 2007: (top to bottom and left to right) chlorophyll *a* concentration (in mg m^{-1}), size factor of phytoplankton S_f (dimensionless), absorption coefficient of colored detrital matter (CDM) at 443 nm (in m^{-1}), and spectral slope of CDM absorption ($S_{cdm,}$ in nm^{-1}). Color scales are given for each parameter and are common to Figures 1a–1d. Scales are logarithmic for [Chl] and $a_{cdm}(443)$, and linear for S_f and S_{cdm}.

Ciotti et al., 2002

Bricaud et al., 2012

Spectral-based – taxonomic groups

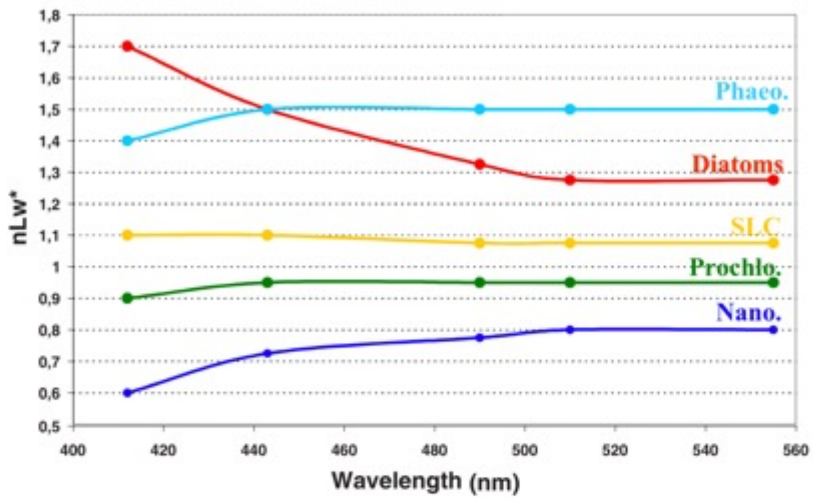


Figure 2. Mean nLw* spectra for the five PHYSAT phytoplankton groups: diatoms in red, nanoeucaryotes in blue, Synechococcus in yellow, Prochlorococcus in green, and phaeocystis-like in light blue.

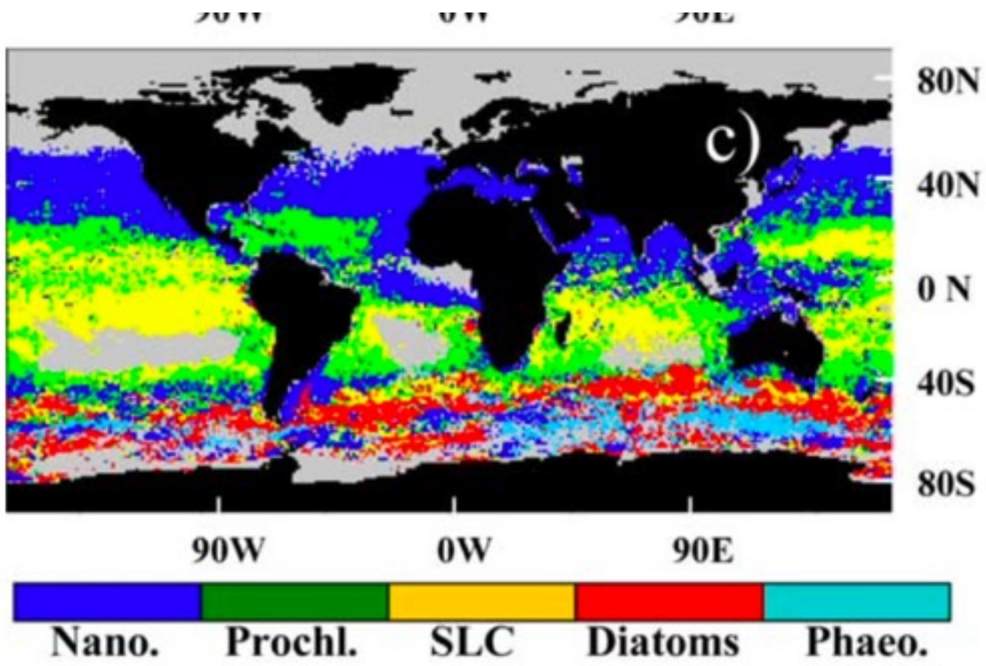
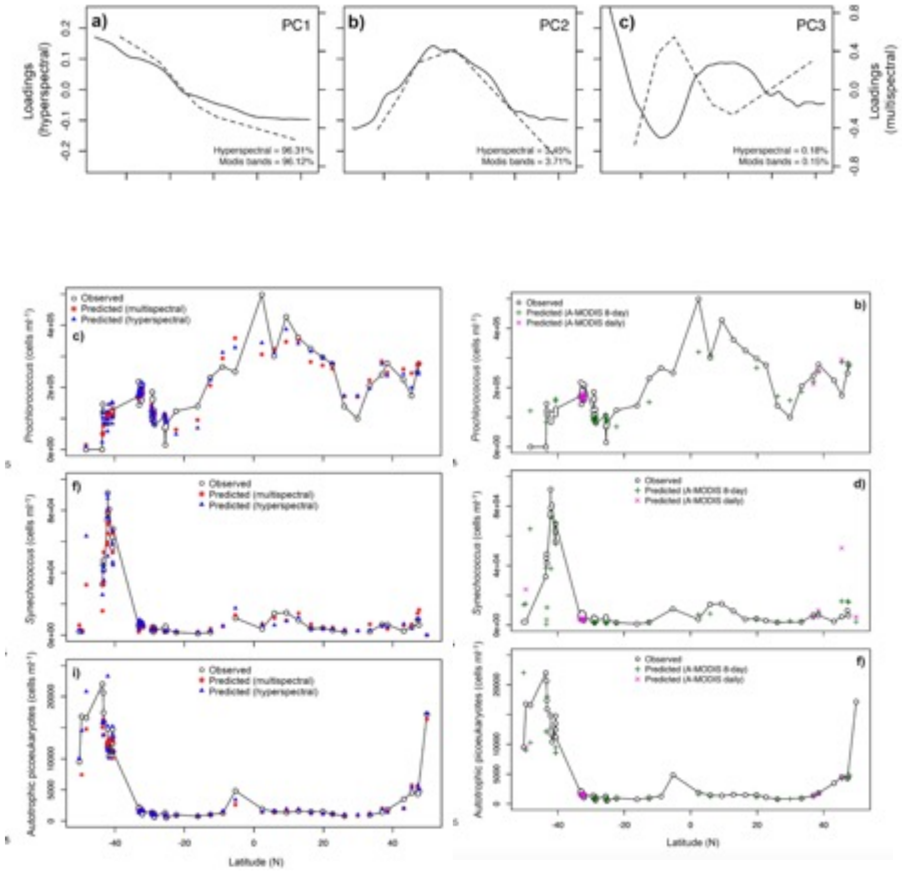


Figure 1. Maps of the dominant phytoplankton group for January 2002 obtained from (a) the standard PHYSAT method of *Alvain et al.* [2005], (b) the improved PHYSAT method used in this study, and (c) the improved PHYSAT method with the additional phaeocystis-like group.

Phytoplankton taxonomic groups as defined by pigments estimated from normalized water-leaving radiance

Alvain et al., 2008

Spectral-based – taxonomic groups



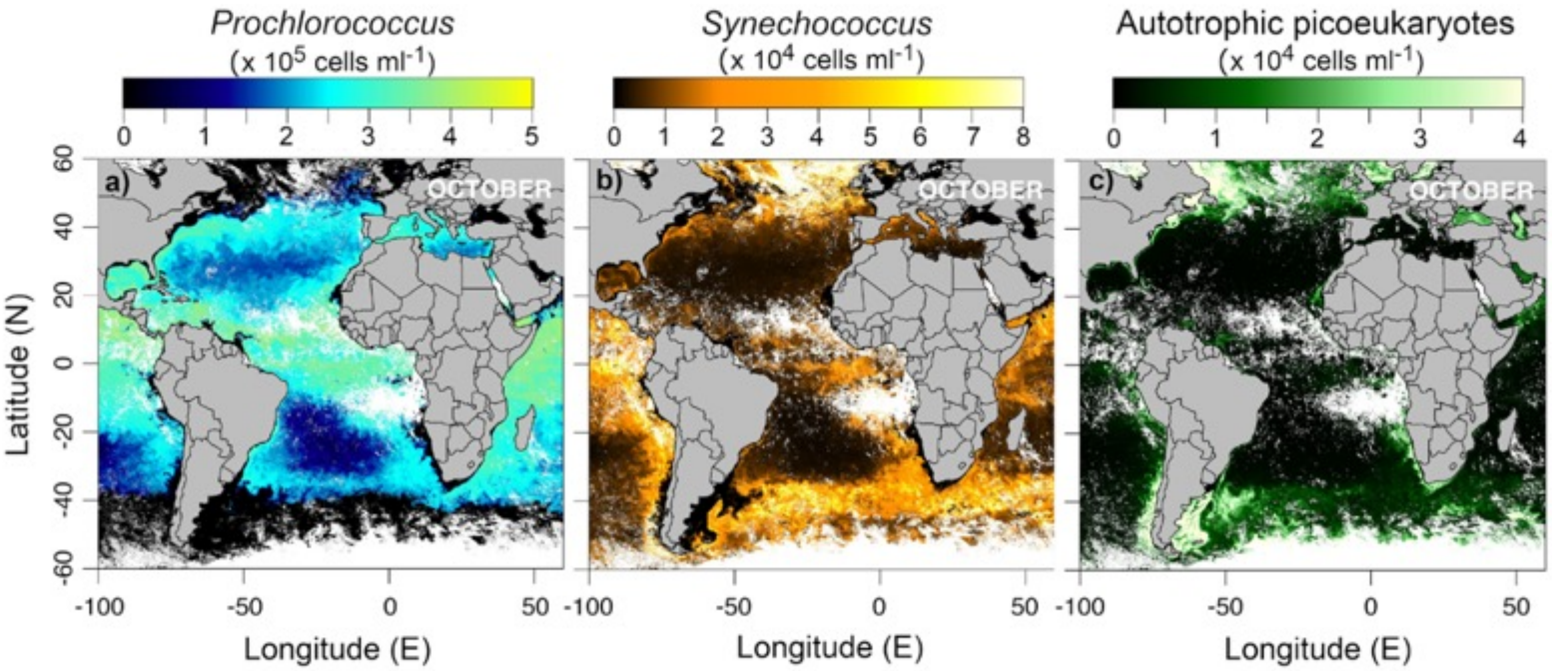


Fig. 9. Aqua-MODIS monthly composites (October 2014) showing cell abundances (cells ml^{-1}) of **a)** *Prochlorococcus*, **b)** *Synechococcus* and **c)** autotrophic picoeukaryotes at the sea surface.

Picophytoplankton estimated from R_{rs}

Lange et al., 2020

TABLE 2 | Summary of satellite inputs and outputs.

Type	Algorithm references	Algorithm abbreviation	Development inputs								Satellite inputs						Satellite Outputs							
			nLw/Rrs	*[Chl]*	a_{ph}	a_{cdm}	b_{bp}	η	*S*	HPLC pigments	*nLw/Rrs*	*[Chl]*	*at*	a_{ph}	a_{cdm}	b_{bp}	*Micro*	*Nano*	*Pico*	*Hapto (cocco)*	*Dino*	*Cyano (Pro/Syn)*	*Diatom*	*Phaeo*
Abundance	Brewin et al., 2010	BR10		x						x		x					x	x	x					
	Brewin R. J. et al., 2011	BR10		x	x					x		x					x	x	x					
	Hirata et al., 2011	OC-PFT		x						x		x					x	x	x	x	x	x	x	
	Uitz et al., 2006	UITZ06								x		x					x	x	x					
Radiance	Alvain et al., 2005, 2008	PHYSAT	x	x						x	x	x						x		x		x	x	x
	Li et al., 2013	LI13	x							x	x						x	x	x					
Absorption	Bracher et al., 2009	PhytoDOAS			x						x											x	x	
	Sadeghi et al., 2012a	PhytoDOAS	x		x						x									x	x		x	
	Ciotti and Bricaud, 2006; Bricaud et al., 2012	CB06	x	x	x	x			x			x	x				(x)		x					
	Devred et al., 2011	DSSP11	x	x	x	x	x		x	x	x						x	x	x					
	Fujiwara et al., 2011	FUJI11	x	x	x			x			x			x			x		(x)					
	Hirata et al., 2008	HIRATA08		x	x					x				x			x	x	x					
	Mouw and Yoder, 2010a	MY10	x	x	x	x	x			x	x	x			x		x		(x)					
	Roy et al., 2013	ROY13		x	x					x		x		x			x	x	x					
Scattering	Kostadinov et al., 2009, 2010	KSM09					x	x			x					x	x	x	x					

The four algorithm types are indicated by color: abundance (green), radiance (red), absorption (yellow), scattering (blue). The development inputs, satellite inputs, and satellite outputs are indicated with "x" for each algorithm. Instances where other size classes could be inferred but are not directly retrieved are indicated with "(x)". Notation for column headers can be found in ***Table 1****.*

Inputs ≠ Outputs is a fundamental algorithm limitation

Mouw et al., 2017

Table 3. Compilation of published algorithms to assess phytoplankton community composition. Algorithms are considered global if they are designed for/applied to more than one major ocean.

Application	PCC product(s)	Algorithm validation data	Remote sensing approaches	Hyperspectral (or polarization?) in situ approaches
Global	Taxonomic group(s)	Direct cell observation (cultures and/or field microscopy)	Subramaniam et al. (2001); Westberry et al. (2005) Subramaniam and Carpenter (1994)	
		Pigment concentrations	Alvain et al. (2005); Alvain et al. (2008); Ben Mustapha et al. (2014); Bracher et al. (2009); Hirata et al. (2011); Losa et al. (2017); Moore et al. (2012); Palacz et al. (2013); Sadeghi et al. (2012); Soppa et al. (2014); Xi et al. (2020)	Torrecilla et al. (2011)
		Spectral signatures	Brown and Yoder (1994)	
	Size classes, size index, or PSD	Pigment concentrations	Brewin et al. (2010); Brewin et al. (2015); Devred et al. (2006); Devred et al. (2011); Fujiwara et al. (2011); Hirata et al. (2008); Hirata et al. (2011); Kostadinov et al. (2010); Li et al. (2013); Moore and Brown (2020); Mouw and Yoder (2010); Roy et al. (2013, spectral a_ph also used in development); Uitz et al. (2006)	
		Mie modeling, Coated Spheres model	Kostadinov et al. (2009); Kostadinov et al. (2022)	
		Spectral signatures	Bricaud et al. (2012)	
	Accessory pigments	Pigment concentrations	O'Shea et al. (2021); Wang et al. (2018)	Bracher et al. (2015); Chase et al. (2013); (Chase et al. 2017); Kramer et al. (2022); Taylor et al. (2013); Uitz et al. (2015)
Regional /Local	Taxonomic group(s)	Direct cell observation (microscopy of cultures and/or field data or imaging-in-flow cytometry)	Chase et al. (2022); Raitsos et al. (2008) Rêve-Lamarche et al. (2017)	Kirkpatrick et al. (2000); Lubac et al. (2008); Millie et al. (1997); Xi et al. (2017); Xi et al. (2015)
		Pigment concentrations	Di Cicco et al. (2017); Kramer et al. (2018); Palacios et al. (2015); Sathyendranath et al. (2004); Werdell et al. (2014)	Catlett and Siegel (2018); Isada et al. (2015); Shaju et al. (2015)
		Spectral signatures		Craig et al. (2006)
	Size classes, size index, or PSD	Pigment concentrations	Gittings et al. (2019)	
		Spectral signatures	Ciotti and Bricaud (2006)	
	Accessory pigments	Pigment concentrations	Bracher et al. (2015); Pan et al. (2010); Sun et al. (2022)	Aguirre-Gómez et al. (2001); Hoepffner and Sathyendranath (1991); Hoepffner and Sathyendranath (1993); Liu et al. (2019); Lohrenz et al. (2003); Wang et al. (2016); Ye et al. (2019)

Cetinić et al., in prep

What are the advantages and limitations of defining PCC via phytoplankton pigments during algorithm development?

- Pigments are ubiquitously measured (global coverage)
- Standardized laboratory protocols (and intercalibrations possible)
- Absorption by pigments is directly related to optical properties of the water

- Pigments ≠ cellular carbon/biomass, and this relationship varies widely
- Taxonomic resolution is limited (and phycobiliproteins are excluded)
- Discrete samples have lower sampling resolution relative to continuously operated instruments

Obtaining Phytoplankton Diversity from Ocean Color: A Scientific Roadmap for Future Development

Astrid Bracher[1,2*], Heather A. Bouman[3], Robert J. W. Brewin[4,5], Annick Bricaud[6,7], Vanda Brotas[8], Aurea M. Ciotti[9], Lesley Clementson[10], Emmanuel Devred[11], Annalisa Di Cicco[12], Stephanie Dutkiewicz[13], Nick J. Hardman-Mountford[14], Anna E. Hickman[15], Martin Hieronymi[16], Takafumi Hirata[17,18], Svetlana N. Losa[1], Colleen B. Mouw[19], Emanuele Organelli[4], Dionysios E. Raitsos[4], Julia Uitz[6,7], Meike Vogt[20] and Aleksandra Wolanin[1,2,21]

https://www.frontiersin.org/articles/10.3389/fmars.2017.00055/full

Gap 1: **Information mismatch** between satellite-derived phytoplankton composition products and user group target variables

Gap 2: Lack of traceability of **uncertainties** in phytoplankton group algorithms

Gap 3: Missing **capabilities of** current ocean color **satellite measurements**

Gap 4: Lack of **regional capability** of phytoplankton group algorithms

PC: SJK lecture

Hyperspectral measurements can capture features missed by multi-spectral

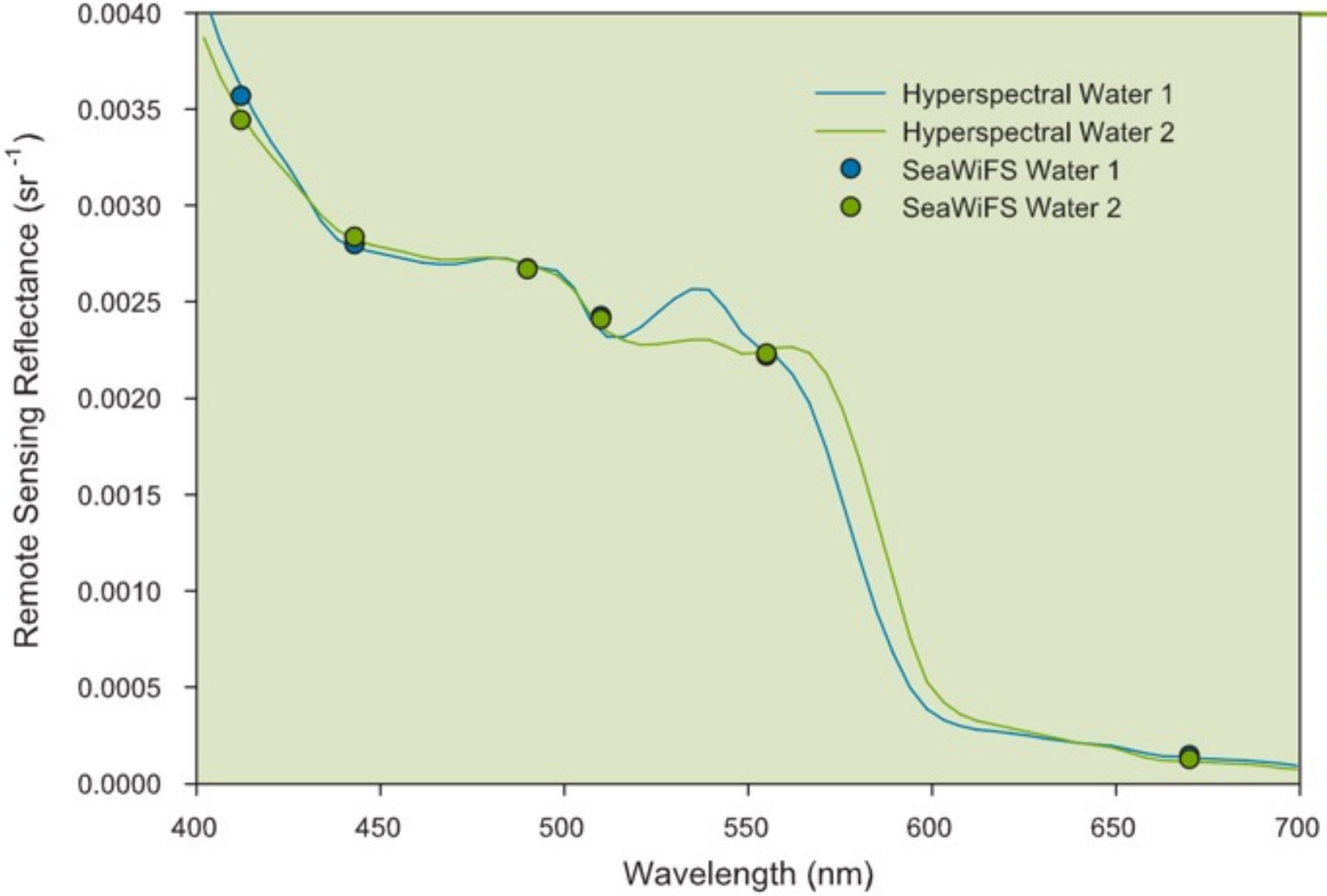
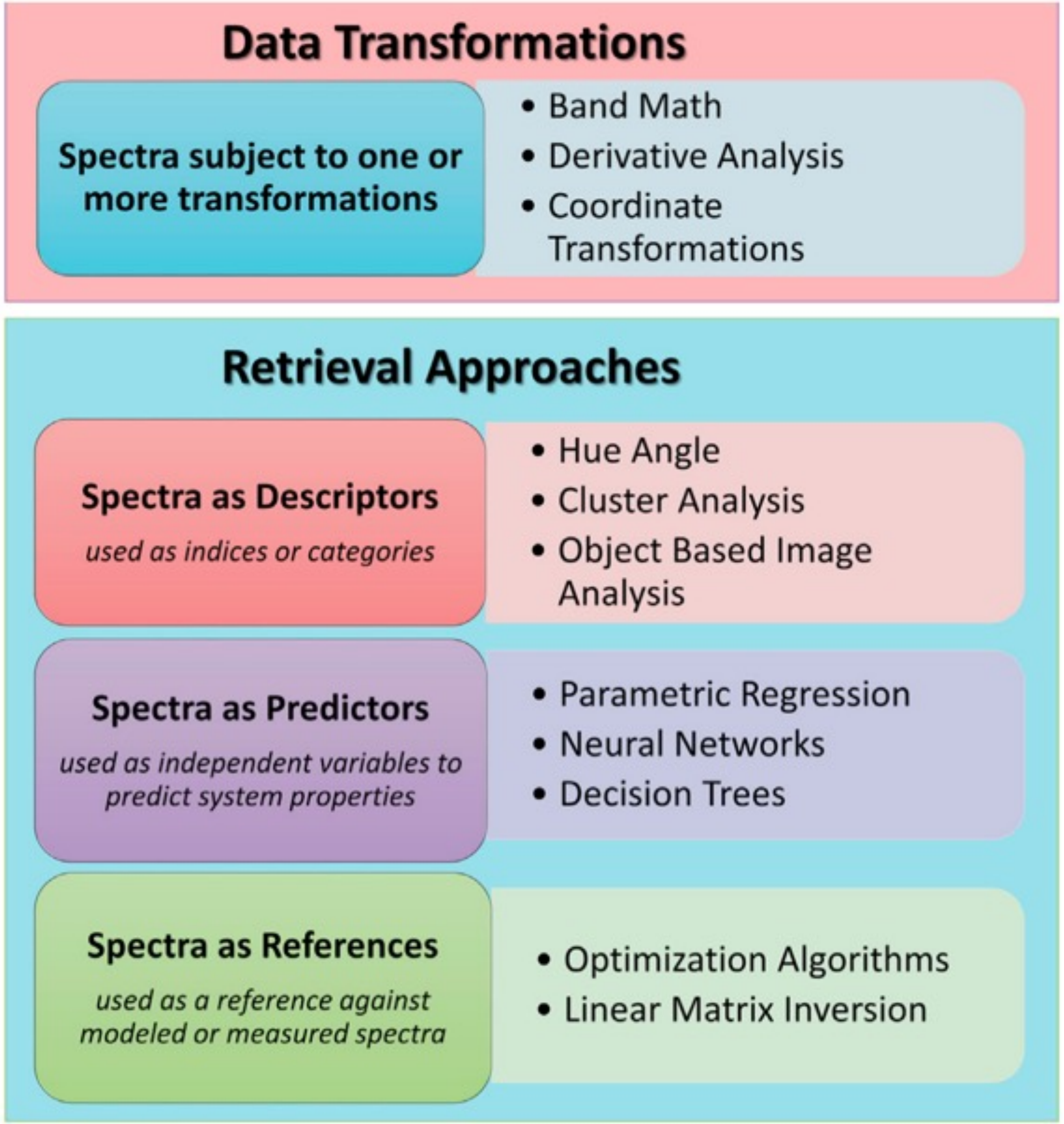


Figure 2. Bottom effects in shallow coastal waters may lead to inaccurate remote sensing retrievals of bottom depth if limited spectral bands are utilized for analysis. This figure shows modeled hyperspectral (solid lines) and multispectral (SeaWiFS wavebands; circles) spectra for two water types, generated by the Hydrolight radiative transfer model (Mobley, 1994). Water 1 (blue) is 6.5 m deep and has low chlorophyll-a and CDOM concentrations with a bottom type of a mixture of soft coral and *Sargassum*, while Water 2 (green) is 13 m deep, "pure water" with a flat green sponge bottom type. By inspection of the hyperspectral spectra, the difference between the two curves is obvious in the 500-600 nm range. However, spectra for the two water types produced using only the SeaWiFS wavebands appear almost identical. (SeaWiFS spectra in this figure were derived by applying the SeaWiFS spectral response function to the hyperspectral signatures).

"The New Age of Hyperspectral Oceanography" Chang et al., 2004 *Oceanography*

Living up to the Hype of Hyperspectral Aquatic Remote Sensing: Science, Resources and Outlook

Heidi M. Dierssen[1*], Steven G. Ackleson[2], Karen E. Joyce[3], Erin L. Hestir[4], Alexandre Castagna[5], Samantha Lavender[6] and Margaret A. McManus[7]



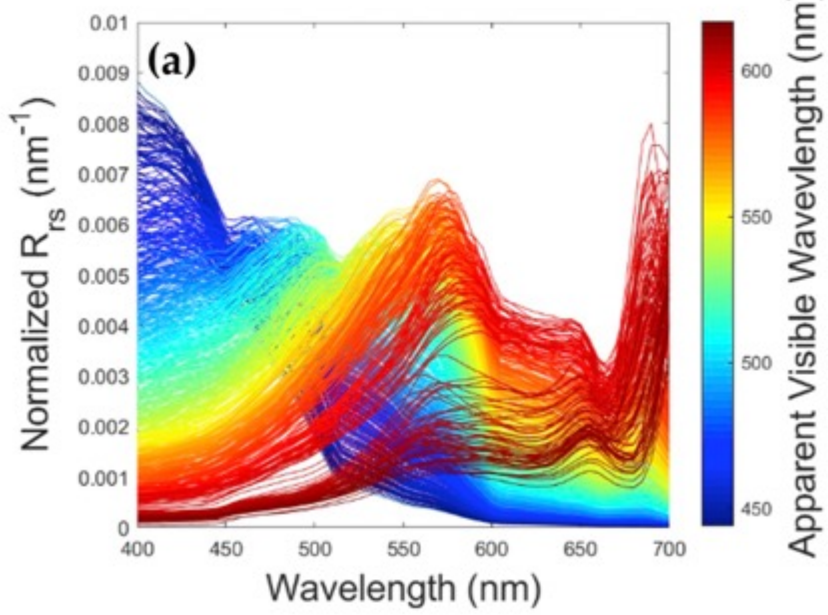
Dierssen et al., 2021

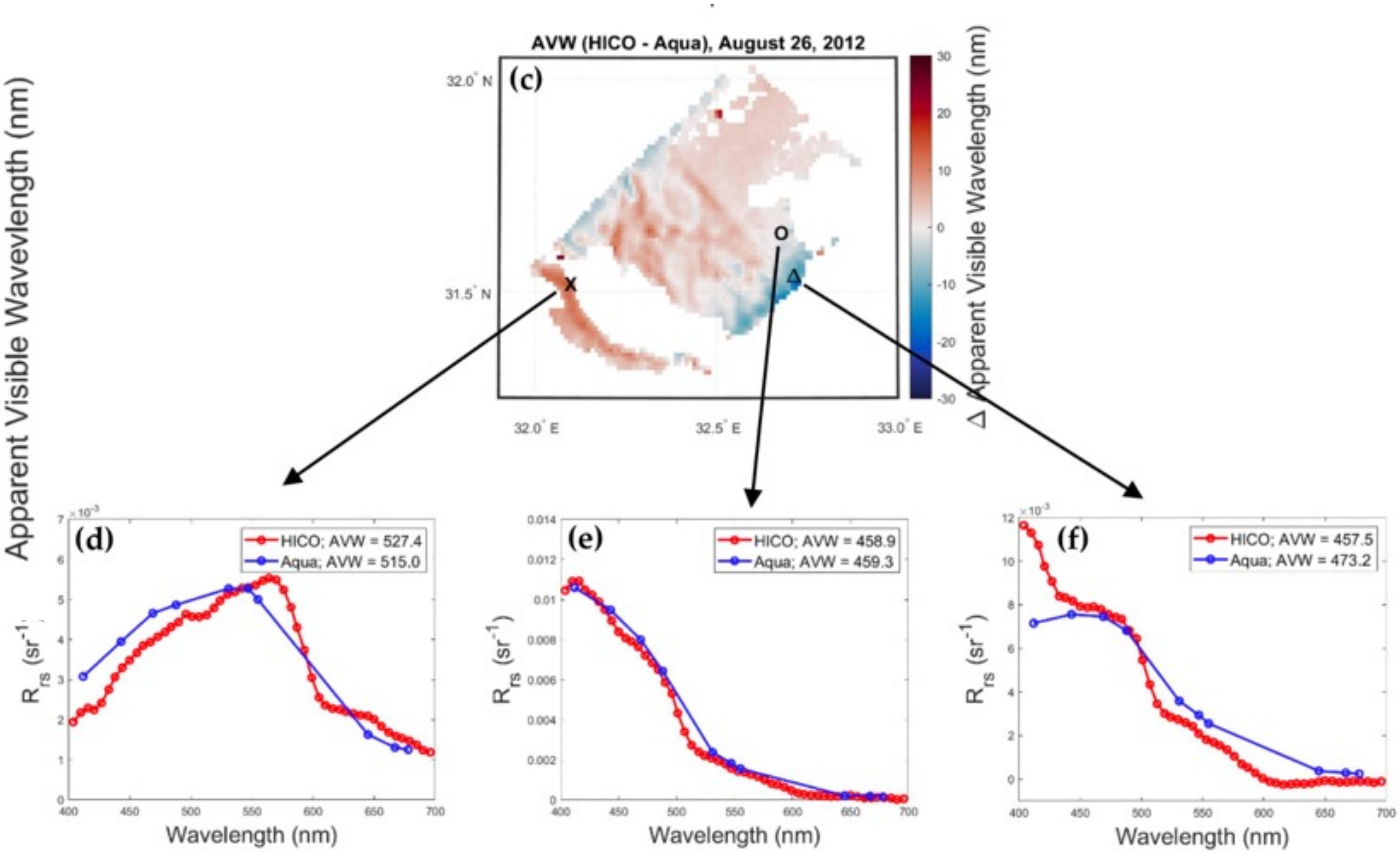
Approaches to extract information from hyperspectral data

Approach	Input measurements	Result/product	Target/validation data	Reference
Direct use of optical measurements: Similarity Index, EOF, and/or clustering analysis	$a_\phi(\lambda)$ & 4th derivative of spectra	% contribution of *G. breve*	*G. breve* field and culture data	Millie et al. 1997
	2nd derivative of $a_\phi(\lambda)$	Diatom contribution to Chl *a*	CHEMTAX diatom Chl *a*	Isada et al. 2015
	$a_p(\lambda)$	Cell counts and Chl *a* fraction of *G. breve*	*G. breve* field and culture data	Kirkpatrick et al. 2000
	2nd derivative of $R_{rs}(\lambda)$	Detection of *Phaeocystis* blooms	Microscopic identification of phytoplankton	Lubac et al. 2008
	4th derivative of $a_\phi(\lambda)$ and $R_{rs}(\lambda)$	Differentiation of phytoplankton groups; cyanobacteria dominance in inland waters	Cultures, Hydrolight simulations, field $R_{rs}(\lambda)$ measurements	Xi et al. 2015; 2017
	Derivatives of $a_p(\lambda)$ or $a_\phi(\lambda)$	Pigment assemblages or concentrations	HPLC pigments or Chl *a* concentration from fluorescence	Catlett and Siegel 2018; Shaju et al. 2015; Torrecilla et al. 2011
	$R_{rs}(\lambda)$	Pigment concentrations	HPLC pigments	Bracher et al. 2015; Kramer et al. 2022
	$a_\phi(\lambda)$ and $R_{rs}(\lambda)$, and derivatives	Bio-optical water categories	HPLC pigments	Uitz et al. 2015
	$L_u(\lambda)$	Relative phycoerythrin concentrations	PE concentration	Taylor et al. 2013
	$a_\phi(\lambda)$ and $R_{rs}(\lambda)$, and $a_\phi(\lambda)$ derivatives	*K. brevis* relative bloom strength	*K. brevis* absorption spectrum	Craig et al. 2006
	$R_{rs}(\lambda)$	Apparent Visible Wavelength		Vandermuelen et al. 2020; Dierssen et al. 2022
Methods of spectral inversion: Spectral inversion and Gaussian decomposition	$a_p(\lambda)$ or $a_\phi(\lambda)$	Pigment concentrations or absorption	HPLC pigments	Aguirre-Gomez et al. 2001; Chase et al. 2013; Hoepffner and Sathyendranath 1991, 1993; Liu et al. 2019; Lohrenz et al. 2003; Ye et al. 2019
	$R_{rs}(\lambda)$	Contribution of phytoplankton groups to absorption	Microscopic cell counts	Roesler et al. 2004
	$R_{rs}(\lambda)$	Pigment concentrations	HPLC pigments	Chase et al. 2017; Wang et al. 2016
	$R_{rs}(\lambda)$	$a_\phi(\lambda)$ and Chl *a* concentrations	In situ $R_{rs}(\lambda)$	Pahlevan et al., 2020; Pahlevan et al., 2021

Retrieval Algorithms

- Spectra as descriptors: optical indices, cluster analyses



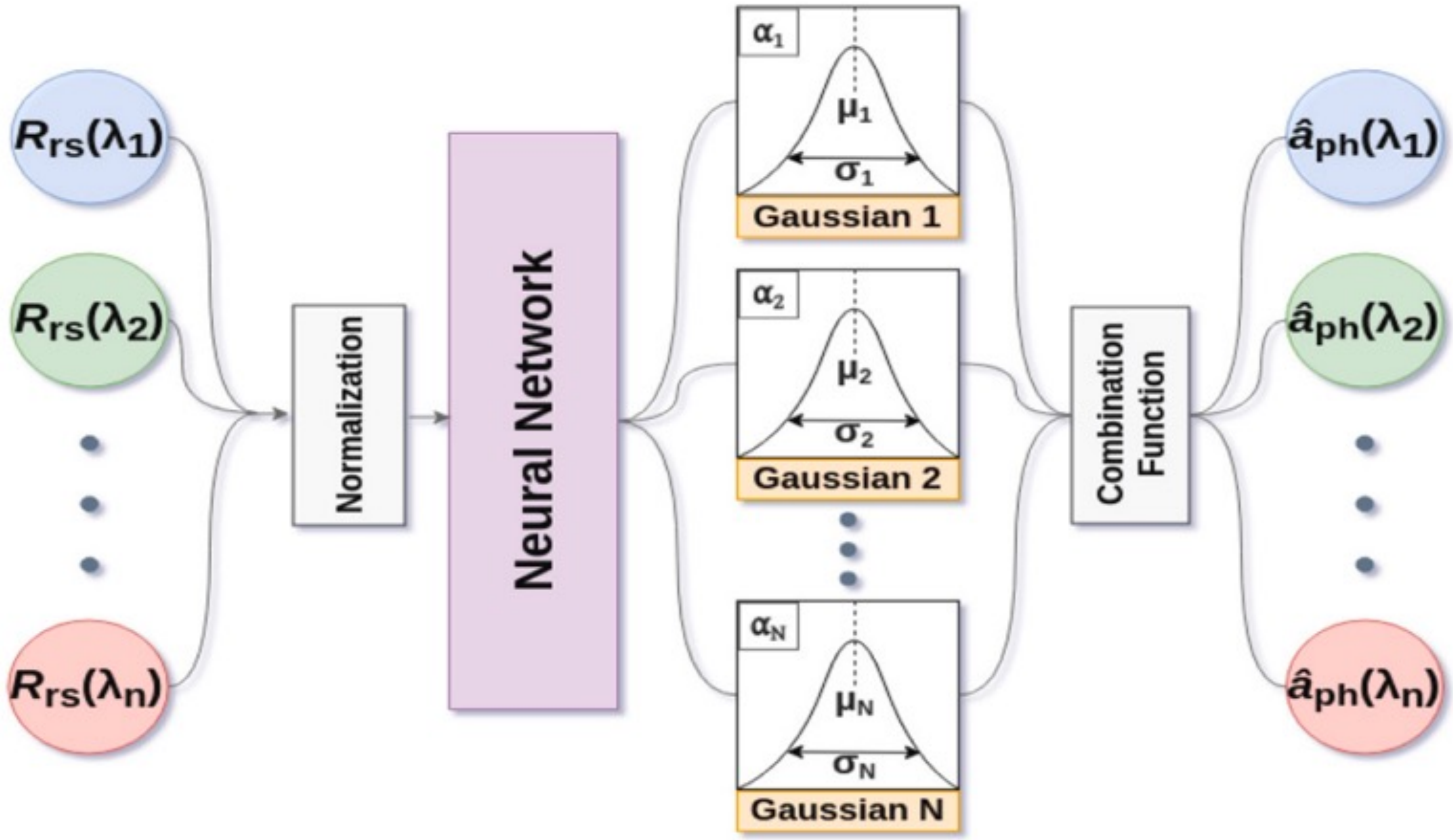


QWIP: A Quantitative Metric for Quality Control of Aquatic Reflectance Spectral Shape Using the Apparent Visible Wavelength

Heidi M. Dierssen[1*], Ryan A. Vandermeulen[2,3], Brian B. Barnes[4], Alexandre Castagna[5], Els Knaeps[6] and Quinten Vanhellemont[7]

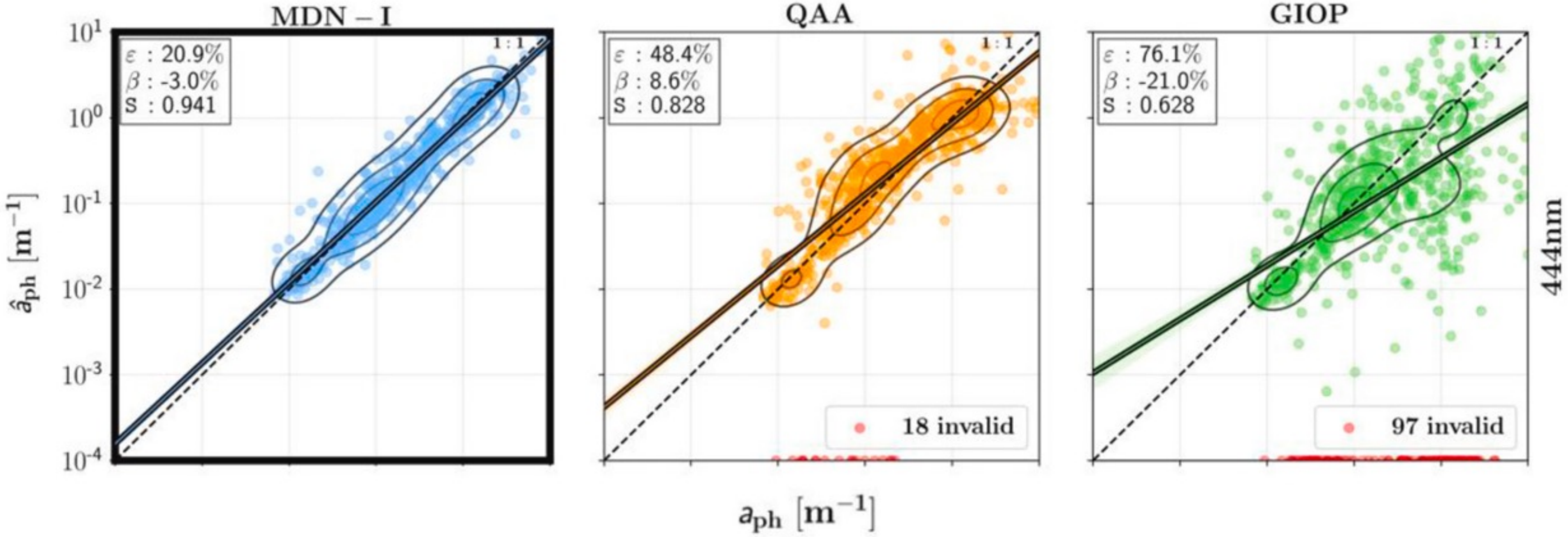
Vandermuelen et al., 2020

Applying Mixture Density Networks (MDN) to hyperspectral R_{rs}



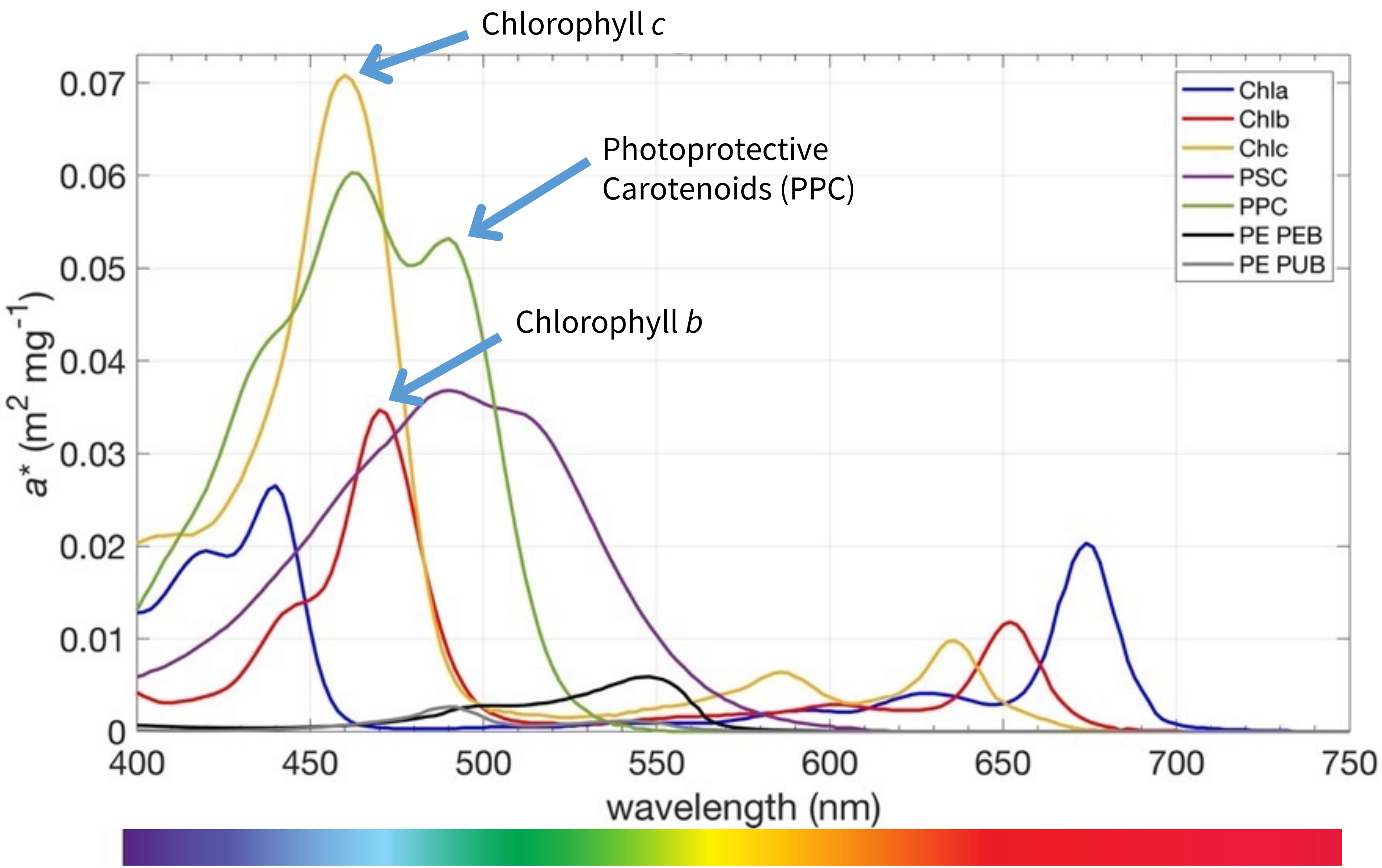
Pahlevan et al., 2021

Applying Mixture Density Networks (MDN) to hyperspectral R_{rs}



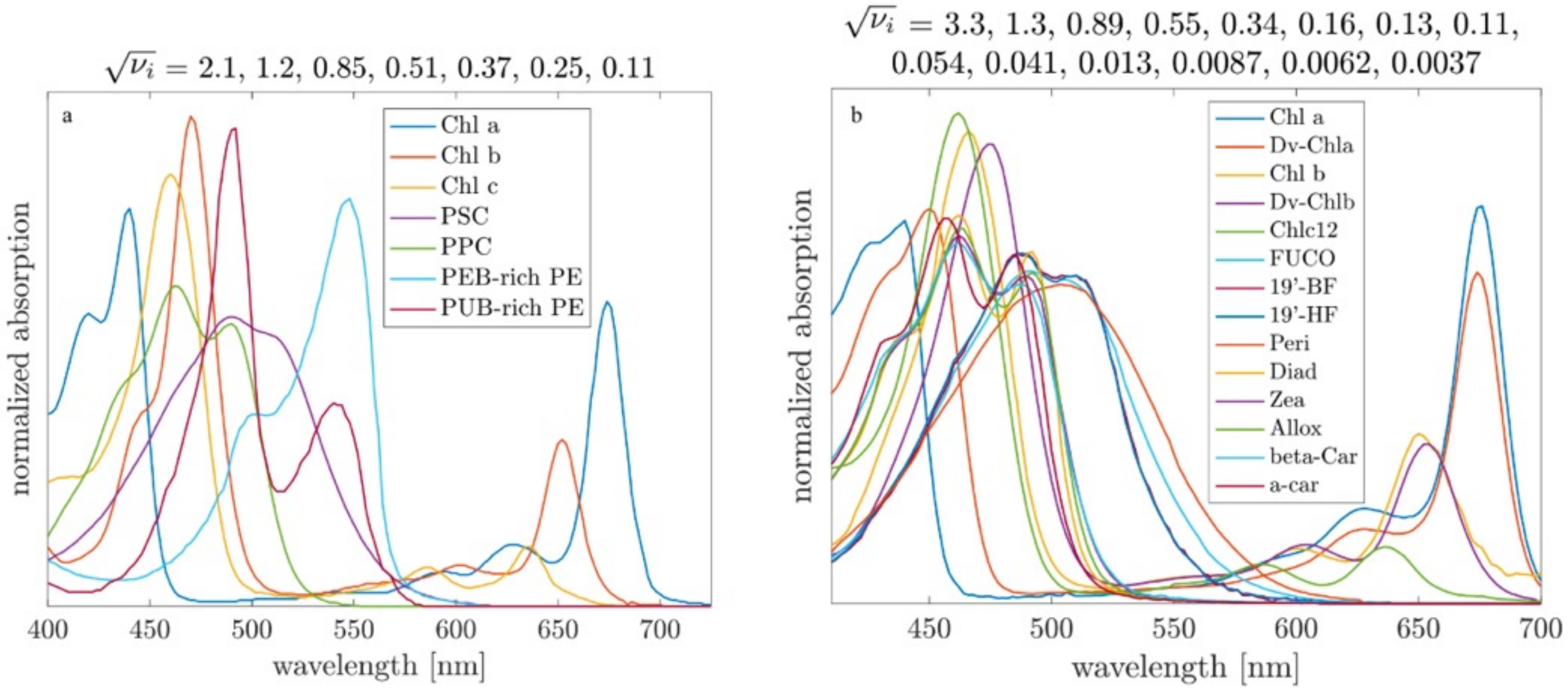
Pahlevan et al., 2020

Phytoplankton pigments drive spectral absorption features



data from Bidigare et al. 1990

But does the inversion problem become ill-posed?



Cael et al. 2020

Phytoplankton pigments estimated using Gaussian decomposition

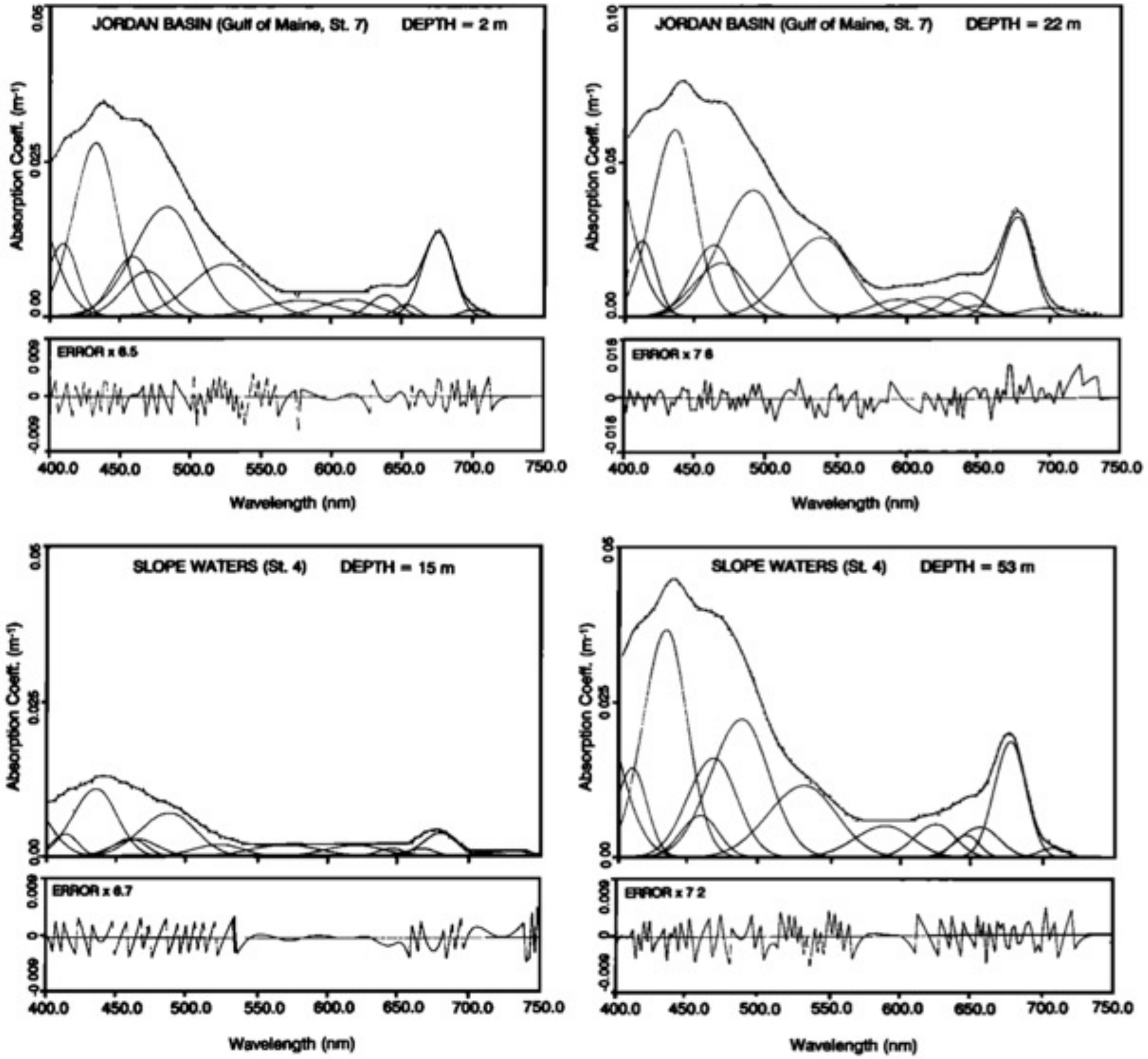
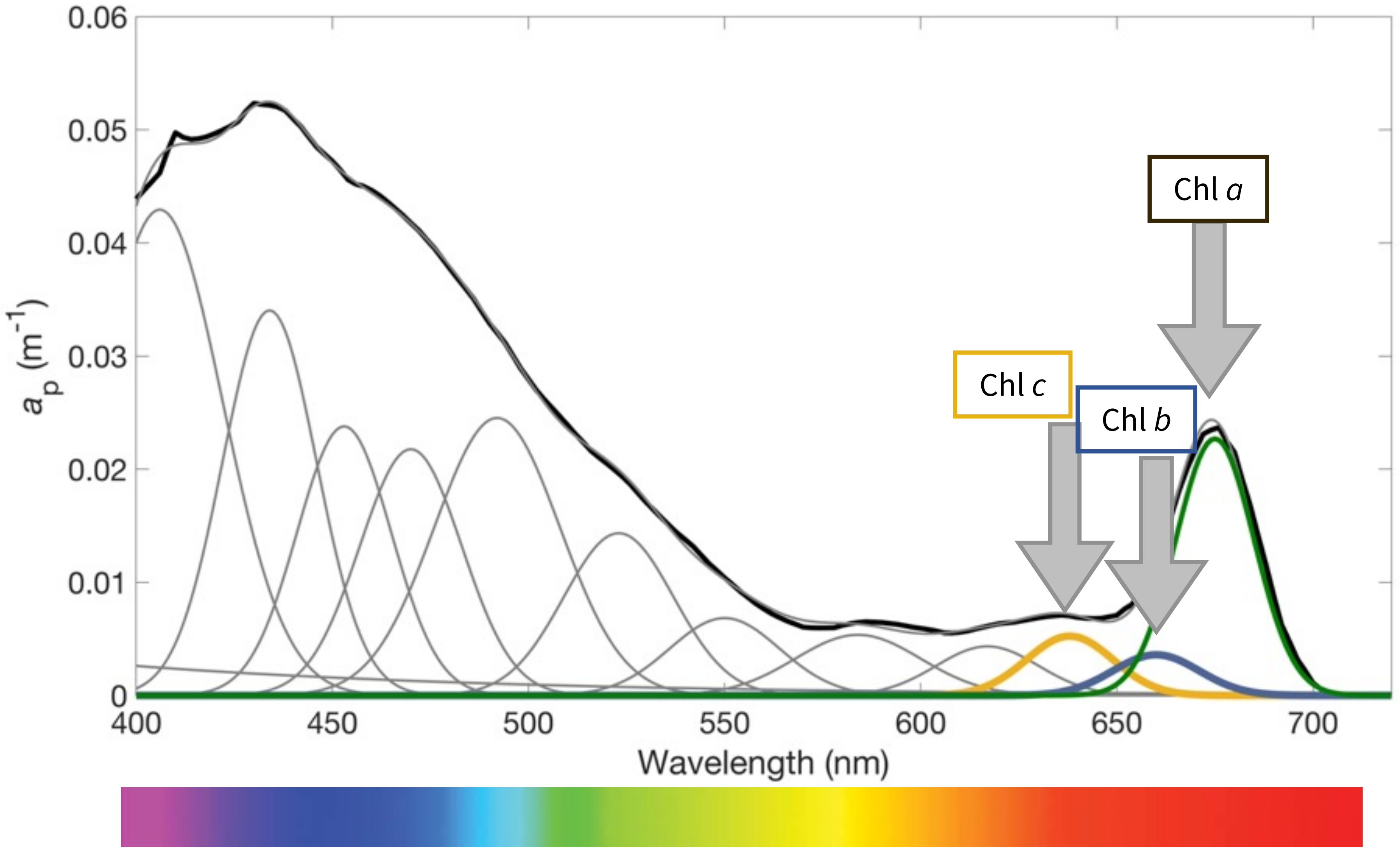


Fig. 7. Decomposition of absorption spectra of natural phytoplankton communities into 13 Gaussian bands, reflecting the absorption characteristics of chlorophylls *a*, *b*, and *c* and of carotenoids at two locations in the western North Atlantic. Variations in the residual error are shown in the lower panel for each spectrum.

Hoepffner and Sathyendranath, 1993

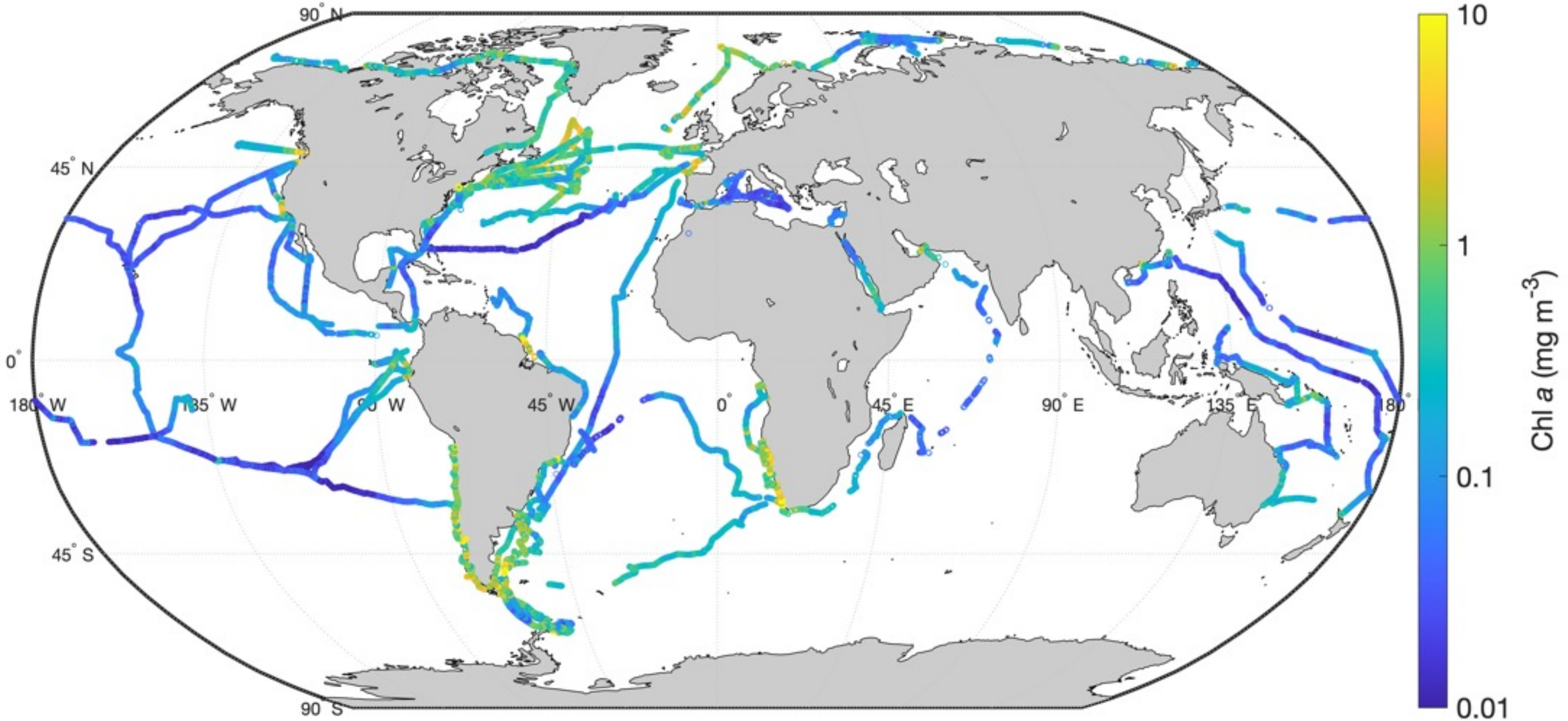
Phytoplankton pigments estimated from ac-s absorption spectra



ac-s

Chase et al., 2013

Chlorophyll *a* estimated from hyperspectral a_p measurements

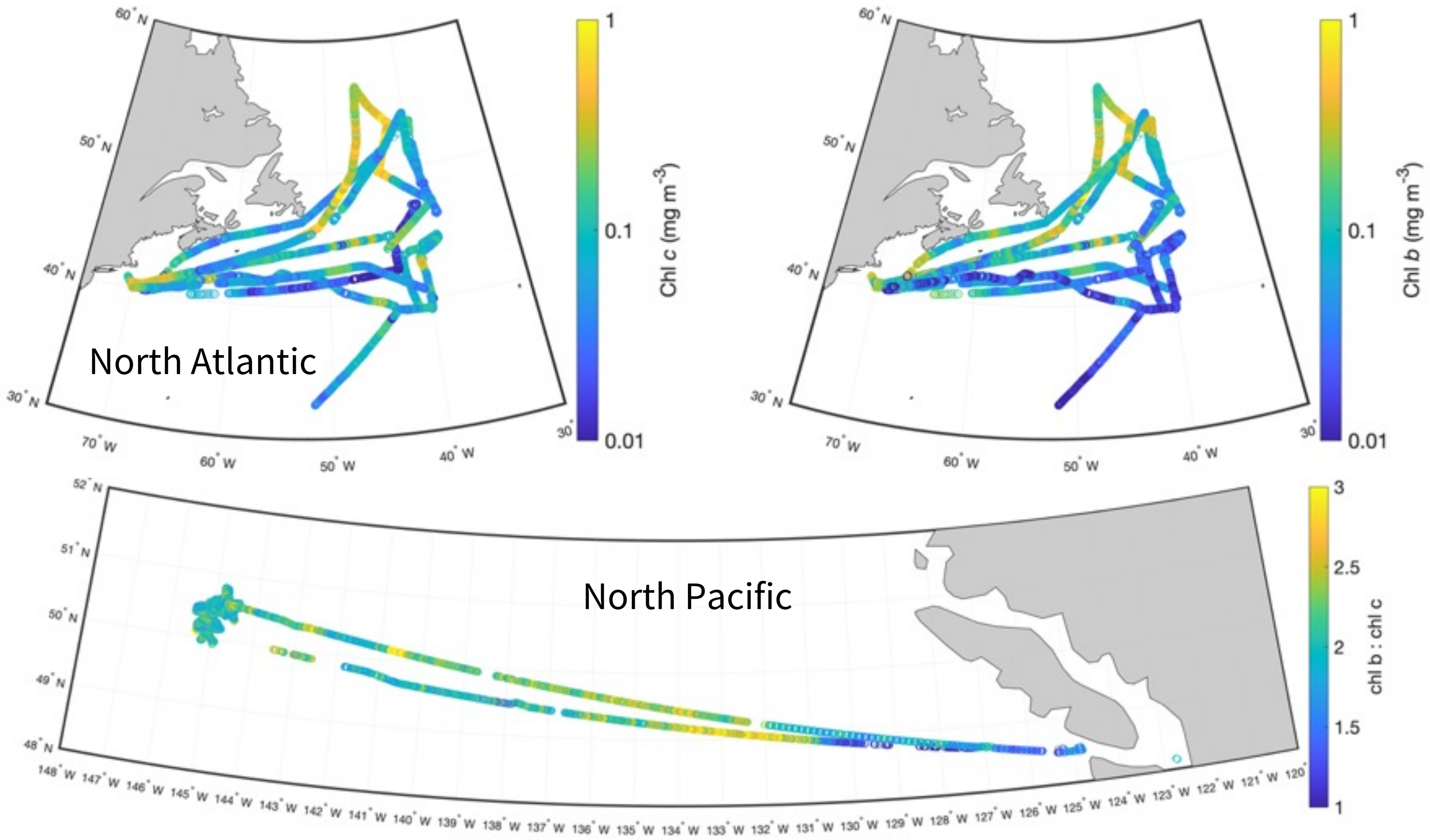


ac-s

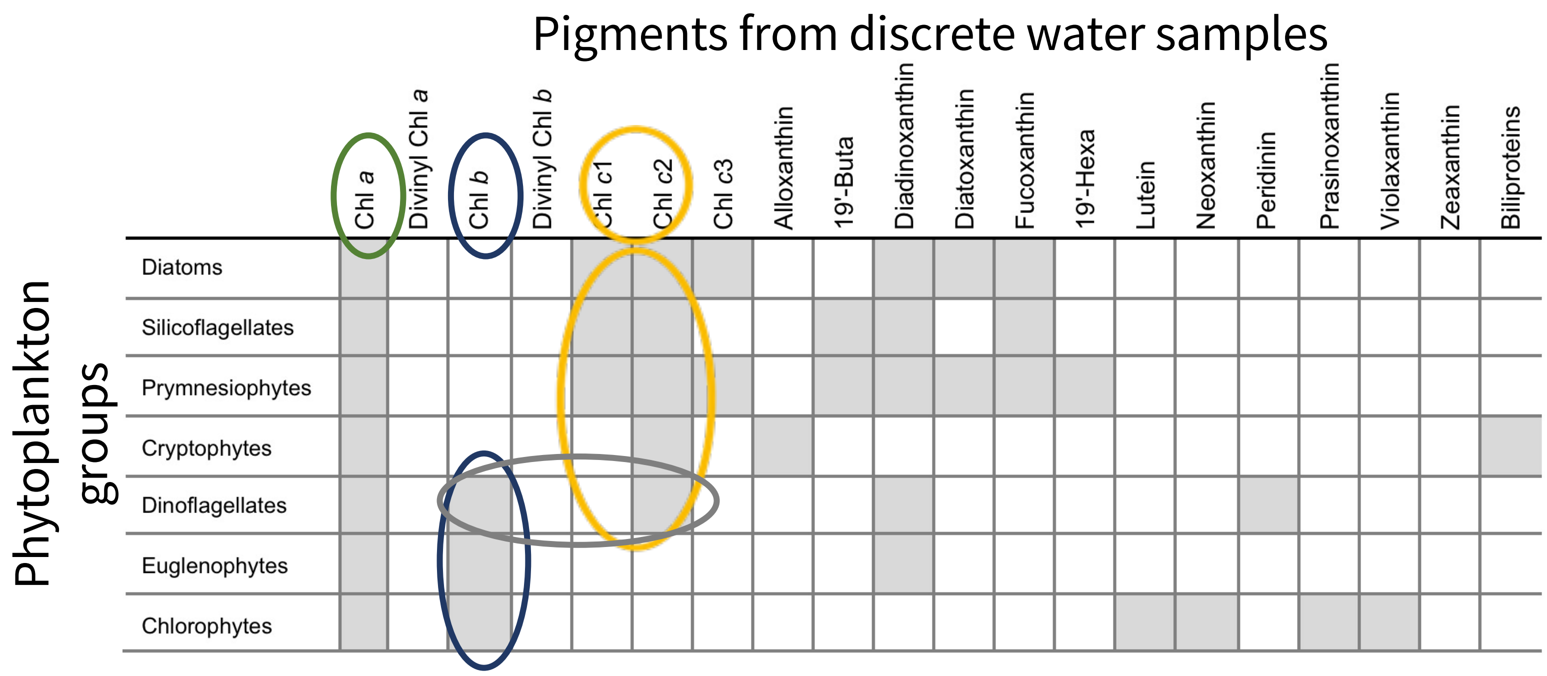
Phytoplankton accessory pigments estimated from hyperspectral a_p

Accessory pigment absolute concentrations:

Accessory pigment ratios:

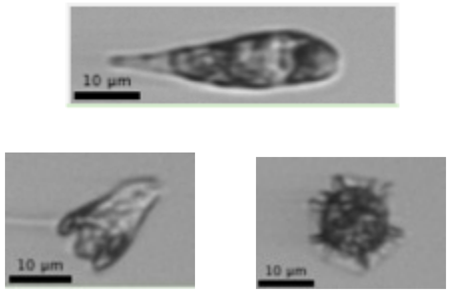


Phytoplankton pigments attributed to different groups



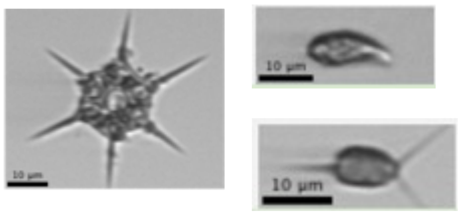
Chlorophyll *b*

Chlorophytes, Euglenoids

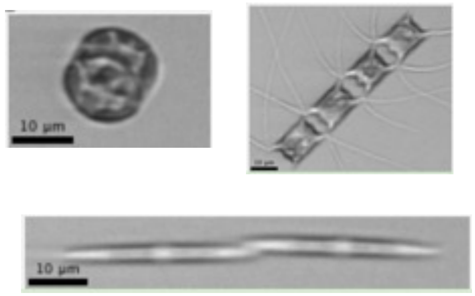


Chlorophyll *c*

Silicoflagellates
Prymnesiophytes
Cryptophytes

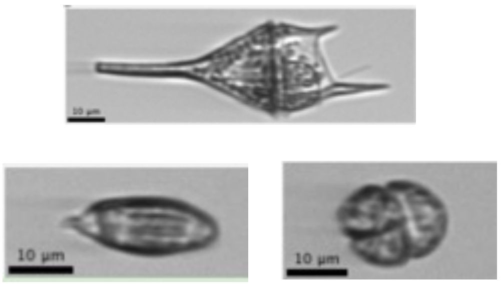


Diatoms



Chlorophylls *b* & *c*

Dinoflagellates



Incorporating Gaussian functions into $R_{rs}(\lambda)$ inversion

$$r_{rs}(\lambda)=\frac{R_{rs}(\lambda)}{0.52+1.7R_{rs}(\lambda)}$$

Lee et al. 2002

$$u(\lambda)\equiv\frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)},$$

$$r_{rs}(\lambda)=g_1u(\lambda)+g_2u(\lambda)^2,$$

$u = u_{meas}$
$g1 = 0.0949$ and $g2 = 0.0794$
(Gordon et al. 1988)

$$u_{mod}(\lambda)=\frac{b_{bp}(\lambda)+b_{bw}(\lambda)}{a_\varphi(\lambda)+a_{CDOM}(\lambda)+a_{NAP}(\lambda)+a_w(\lambda)+b_{bp}(\lambda)+b_{bw}(\lambda)},$$

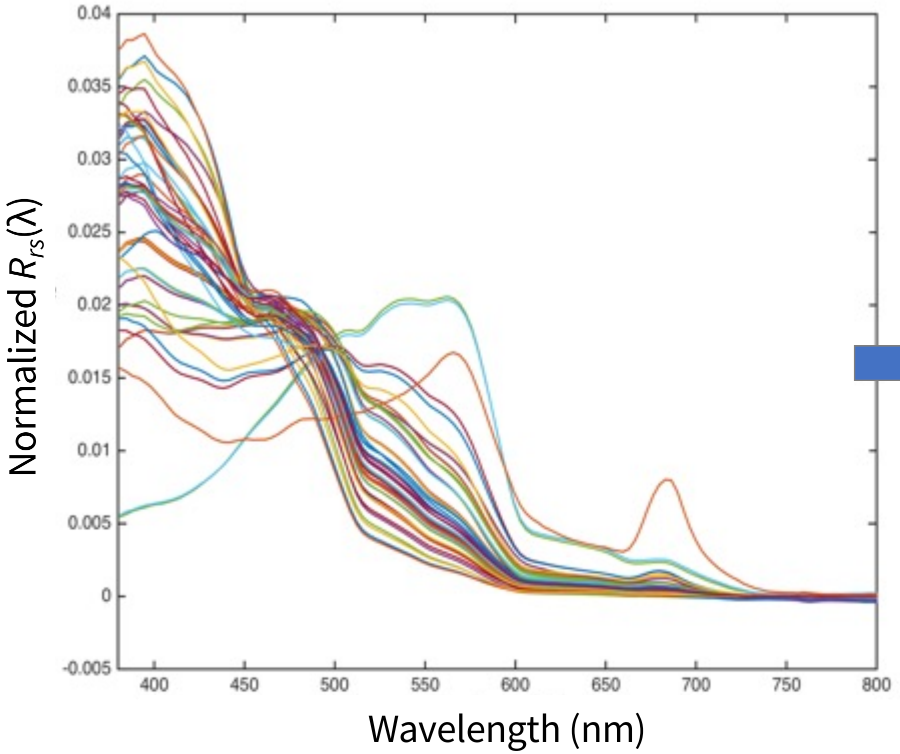
$$a_\varphi(\lambda)=\sum_{i=1}^{8}a_{gaus}(peak_i,\lambda)e^{\left(-0.5\left(\frac{\lambda-peak_i}{\sigma_i}\right)\right)^2},$$

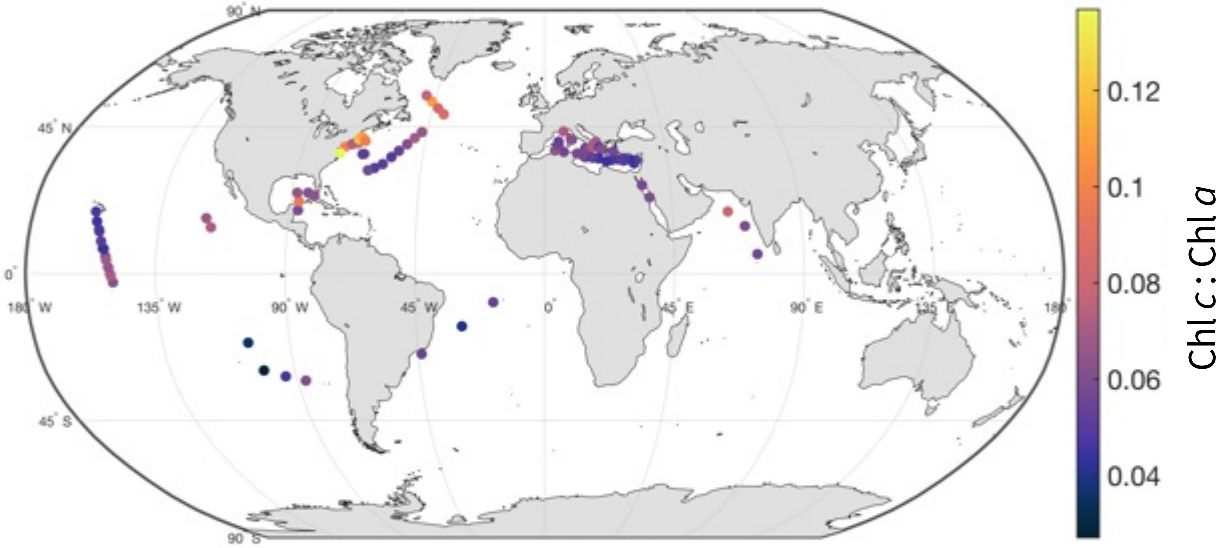
$$\chi^2=\sum_{i=1}^{60}\left(\frac{u_{meas}(\lambda_i)-u_{mod}(\lambda_i)}{u_{std}(\lambda_i)}\right)^2,$$

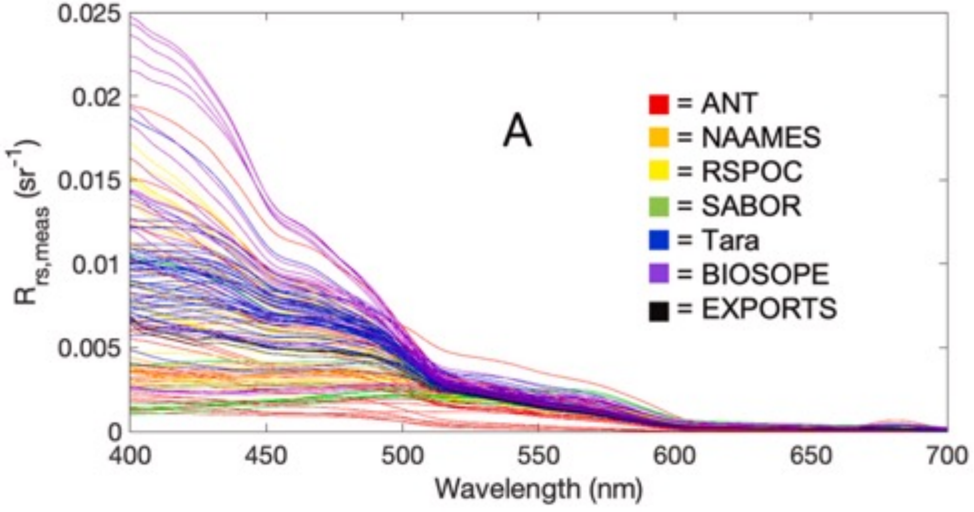
60 wavelengths
between 400-600 nm

Chase et al. 2017

Pigments estimated from $R_{rs}(\lambda)$ spectra measured *in situ*







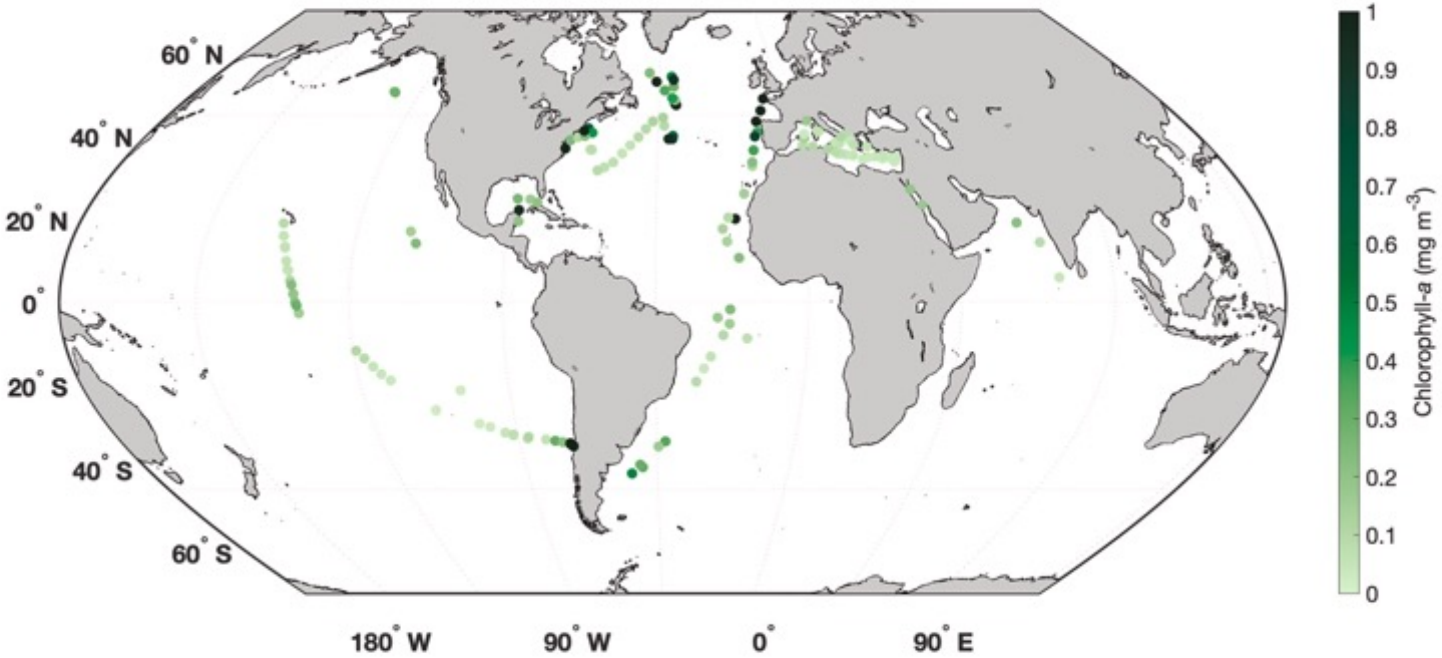
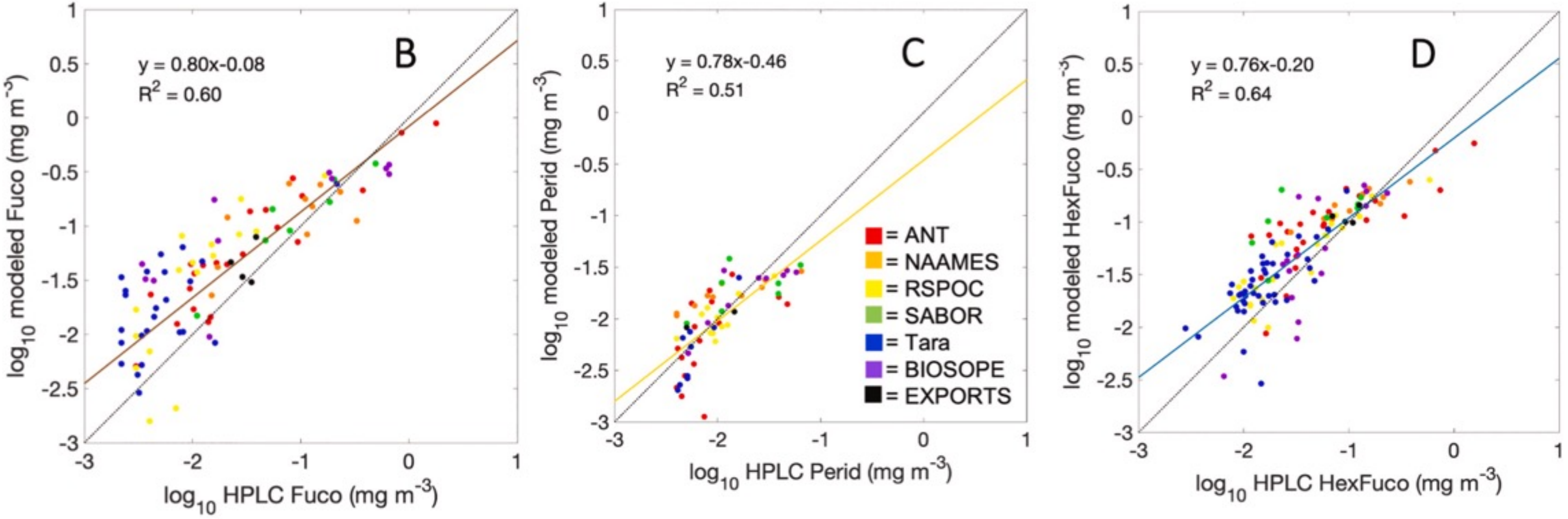
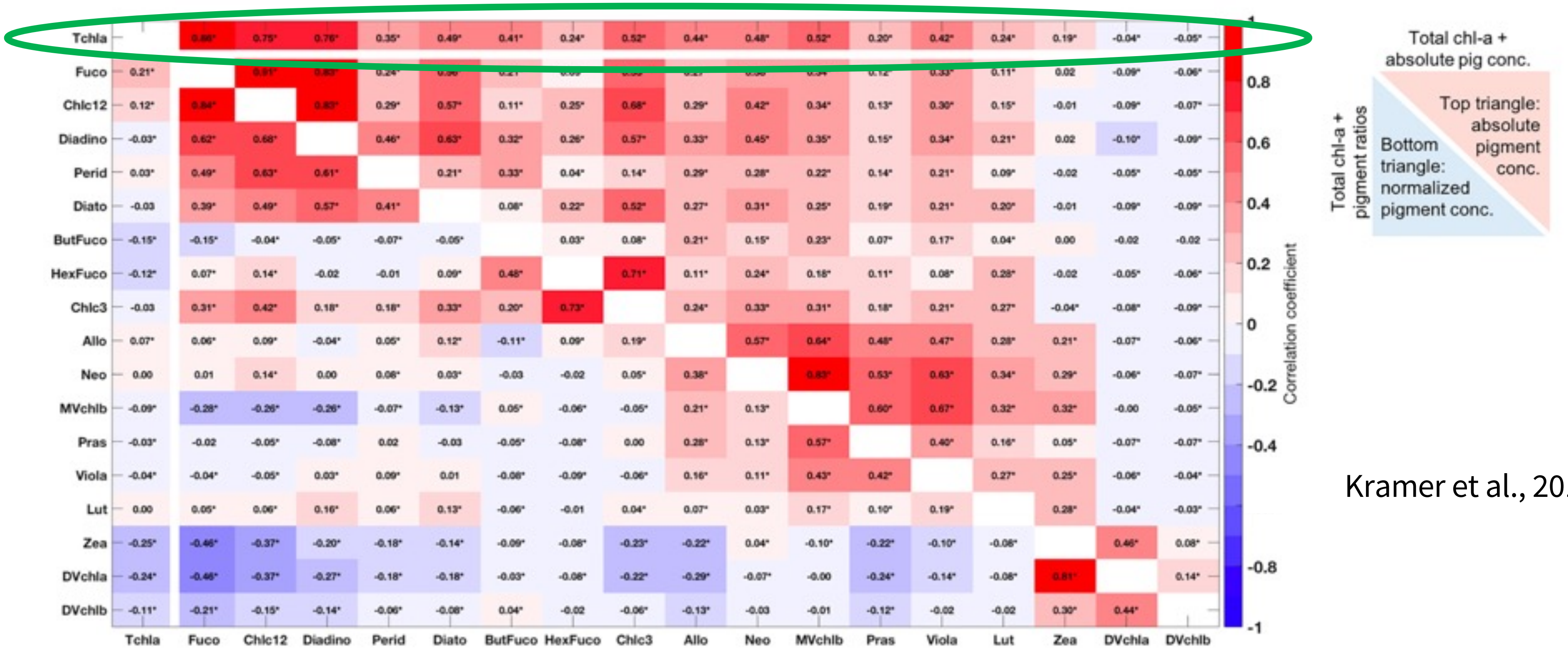


Fig. 1. Global distribution of 145 matched HPLC and hyperspectral $R_{rs}(\lambda)$ samples, colored by chlorophyll-*a* concentration (Tchla).

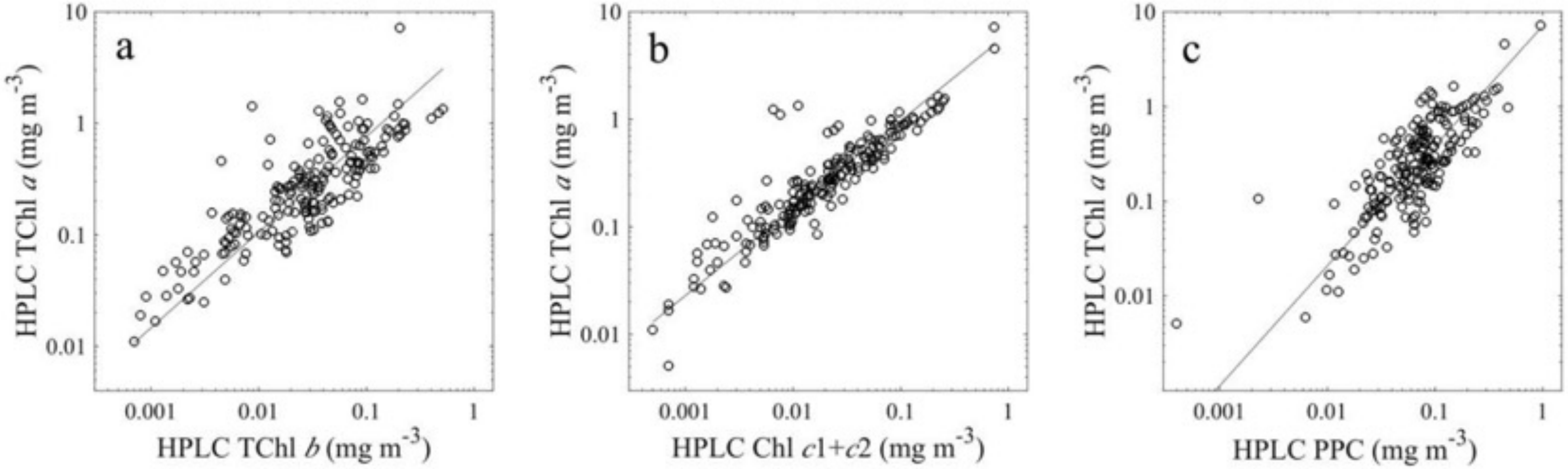


Pigment	Mean R^2	SD R^2	Mean normalized MAD	SD normalized MAD
Allo	0.40	0.19	1.221	0.400
But	0.62	0.16	0.588	0.185
Chlc3	0.68	0.13	0.639	0.212
Chlc12	0.70	0.13	0.703	0.235
DVchla	0.55	0.12	0.594	0.103
Fuco	0.65	0.15	0.844	0.274
Hex	0.54	0.16	0.692	0.201
MVchlb	0.42	0.19	0.975	0.295
Neo	0.42	0.21	1.127	0.354
Perid	0.49	0.13	0.783	0.166
Tchla	0.72	0.15	0.498	0.127
Viola	0.38	0.18	1.101	0.370
Zea	0.37	0.10	0.472	0.071

But most pigments are correlated with Chl *a*…



Kramer et al., 2019



Chase et al., 2017

Considerations of error, & “going beyond Chl *a*”

From Cael et al. (2020):

- Error is the difference between having four to five DoF rather than >60, and the difference between being able to meaningfully invert for four spectra versus 44

- Some errors such as random electronic noise can be reduced by averaging many measurements in time or space. Others, such a bias in calibration, cannot.

- While all optical variation in the water cannot be said to fall along a single axis, it does appear that much of the variation in the surface covaries with [Chl]. Thus, the interest in going “beyond chlorophyll” can be considered an interest in deviations from this axis.

- Polarization will help better separate oceanic and atmospheric contributions to the total signal, and UV will help better separate CDOM, NAP, and phytoplankton contributions to the oceanic signal. These deviations are by definition second order—though we note emphatically that this does not make them unimportant or uninteresting!

- Take home: Judicious use of available DoF; use basis vectors that are specific to your needs in the case of a regional or tuned algorithm

Take-home messages re: hyperspectral measurements:

The question is not as simple as *“how much information can we extract from hyperspectral measurements?”*, **but rather,**

“which approaches and methods that take advantage of the added information in hyperspectral measurements are relevant to my research question(s)?”

With limited degrees of freedom in hyperspectral measurements alone, consider the incorporation of other types of optical and/or environmental data during algorithm development and application, as well as spatial and temporal resolution requirements.

Recent work, new approaches, and expanding data types & tools

phytoclass: A pigment-based chemotaxonomic method to determine the biomass of phytoplankton classes

Alexander Hayward [1,2]*, Matthew H. Pinkerton [1], Andres Gutierrez-Rodriguez [1,3]
[1]National Institute of Water and Atmospheric Research, Wellington, New Zealand
[2]University Of Otago, Dunedin, New Zealand
[3]Instituto Español de Oceanografía, Centro Oceanográfico de Gijón, Gijón, Spain

Limnol. Oceanogr.: Methods 2023

doi: 10.1002/lom3.10541

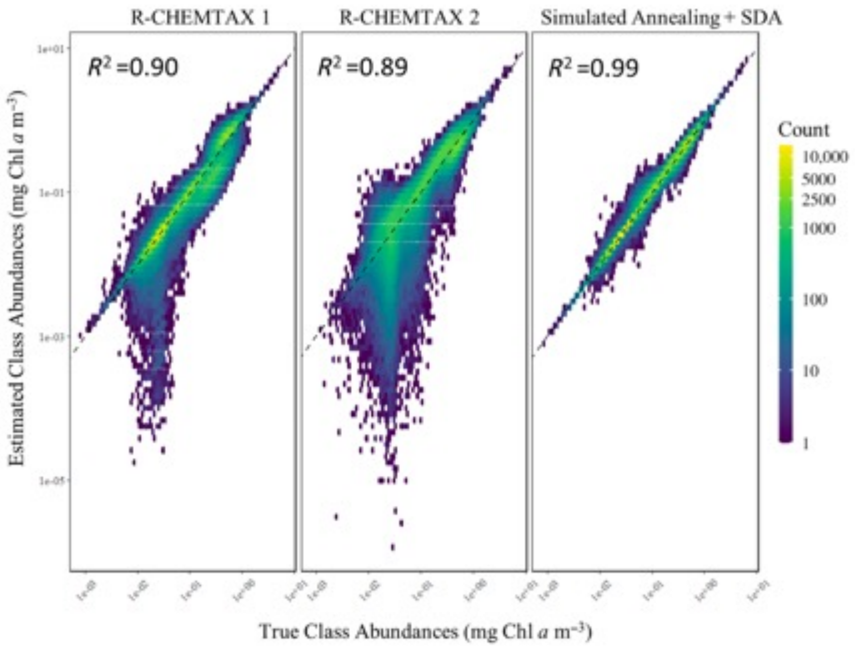


Fig. 7. Density plot of true class abundances and predicted class abundances for R-CHEMTAX-1, R-CHEMTAX-2, and simulated annealing + SDA approaches in synthetic dataset-1.

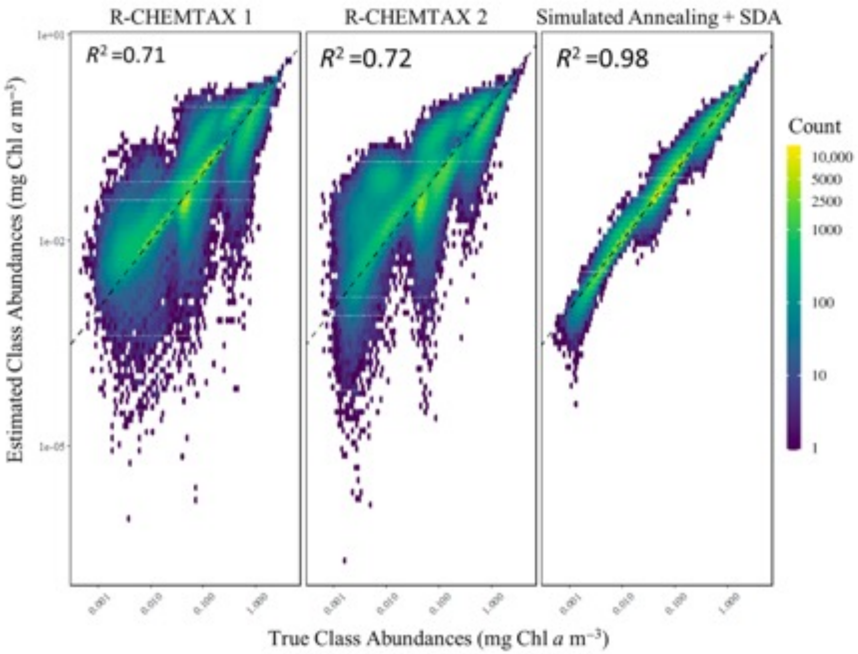
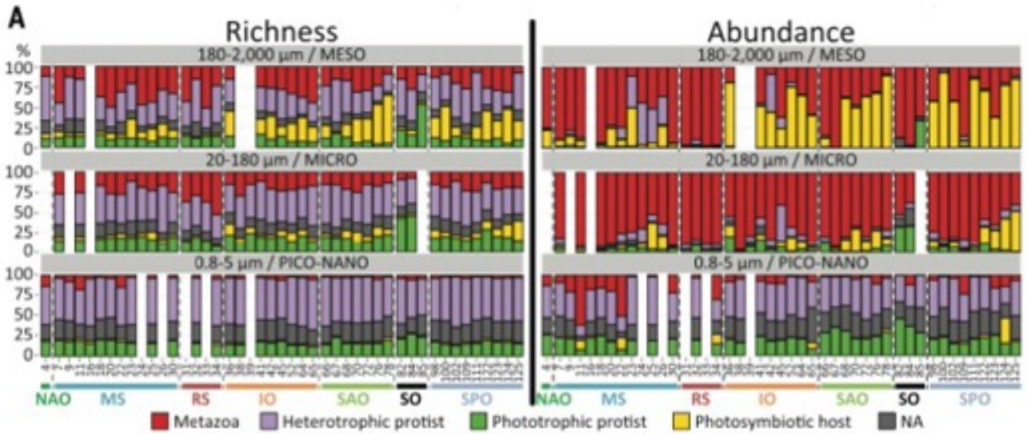


Fig. 13. Density plots of true class abundances and predicted class abundances for R-CHEMTAX-1, R-CHEMTAX-2, and simulated annealing + SDA approaches using synthetic datasets in synthetic dataset-2.

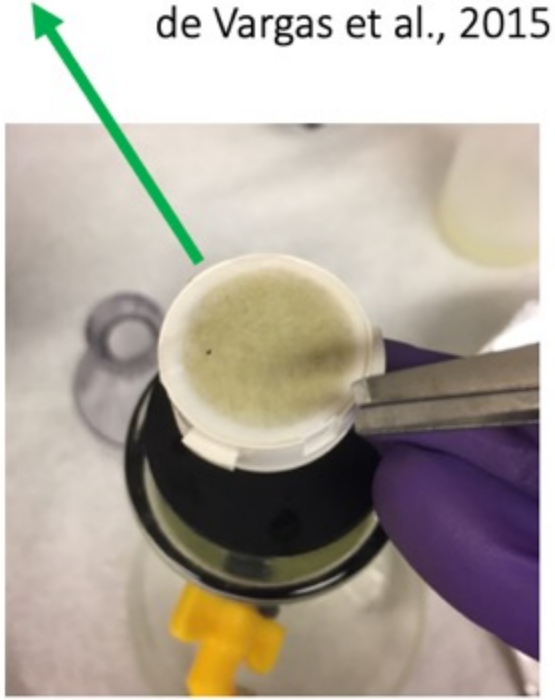
→ New method to estimate phytoplankton groups from pigments

Molecular methods

Quick method summary: Large sample (2-10L) collected, filtered, DNA is extracted, and “barcode” genes are targeted for amplification and sequencing



de Vargas et al., 2015



Strengths: High taxonomic resolution (often to genus or species level), can compare well to other methods (microscopy, pigments)

Weaknesses: Primers may be limited for some groups, gene copies =\= abundance

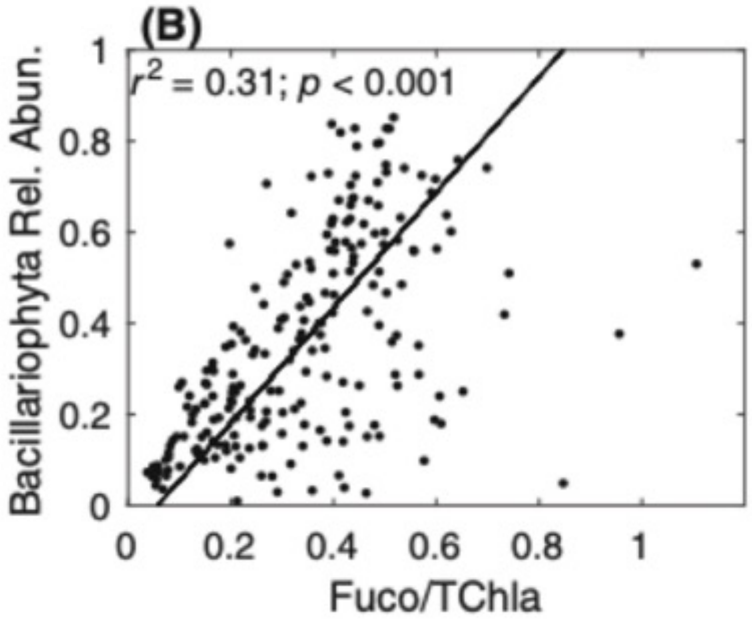
Output: *Relative* sequence abundances

Slide credit: S. Kramer

Molecular methods

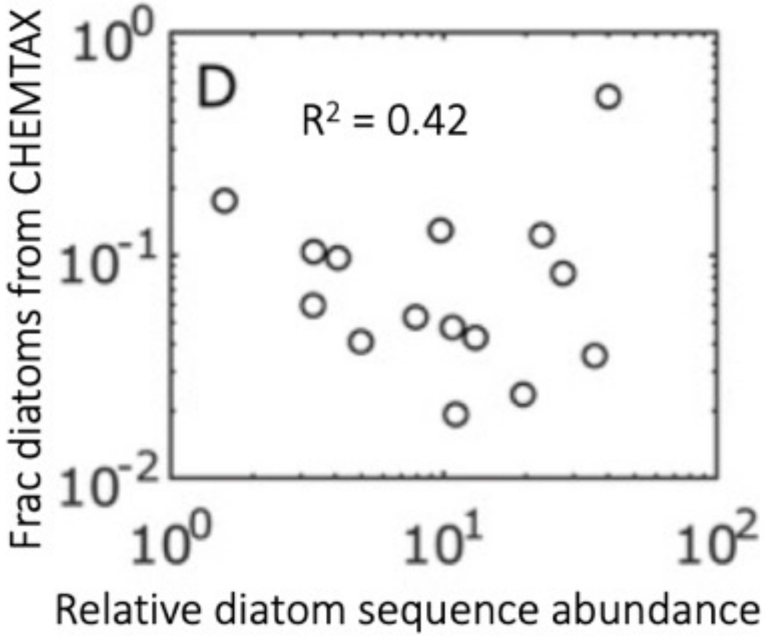
Comparing DNA metabarcoding to other methods: results can really depend on ecosystem dynamics

Santa Barbara Channel



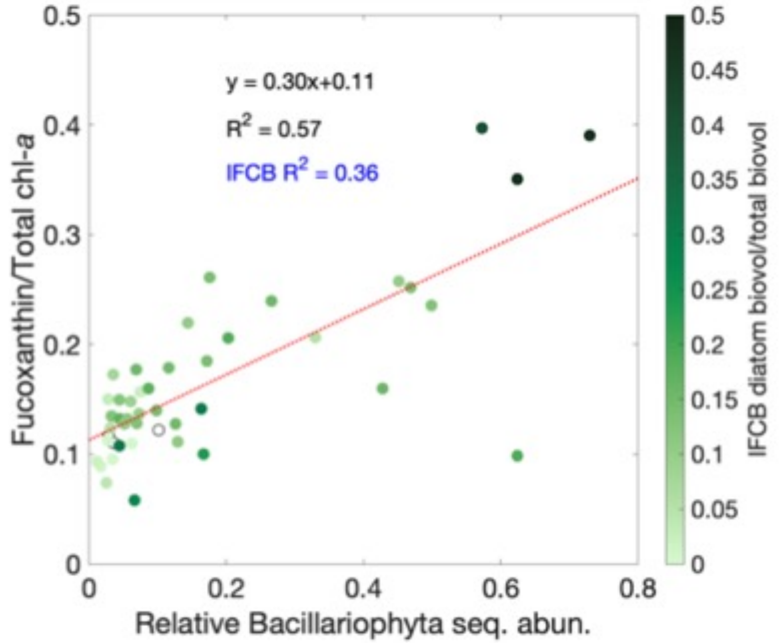
Catlett et al., 2022

West Antarctic Peninsula



Lin et al., 2019

Open ocean (N. Atl + N. Pac)



Kramer et al., *in prep*

Slide credit: S. Kramer

Globally Consistent Quantitative Observations of Planktonic Ecosystems

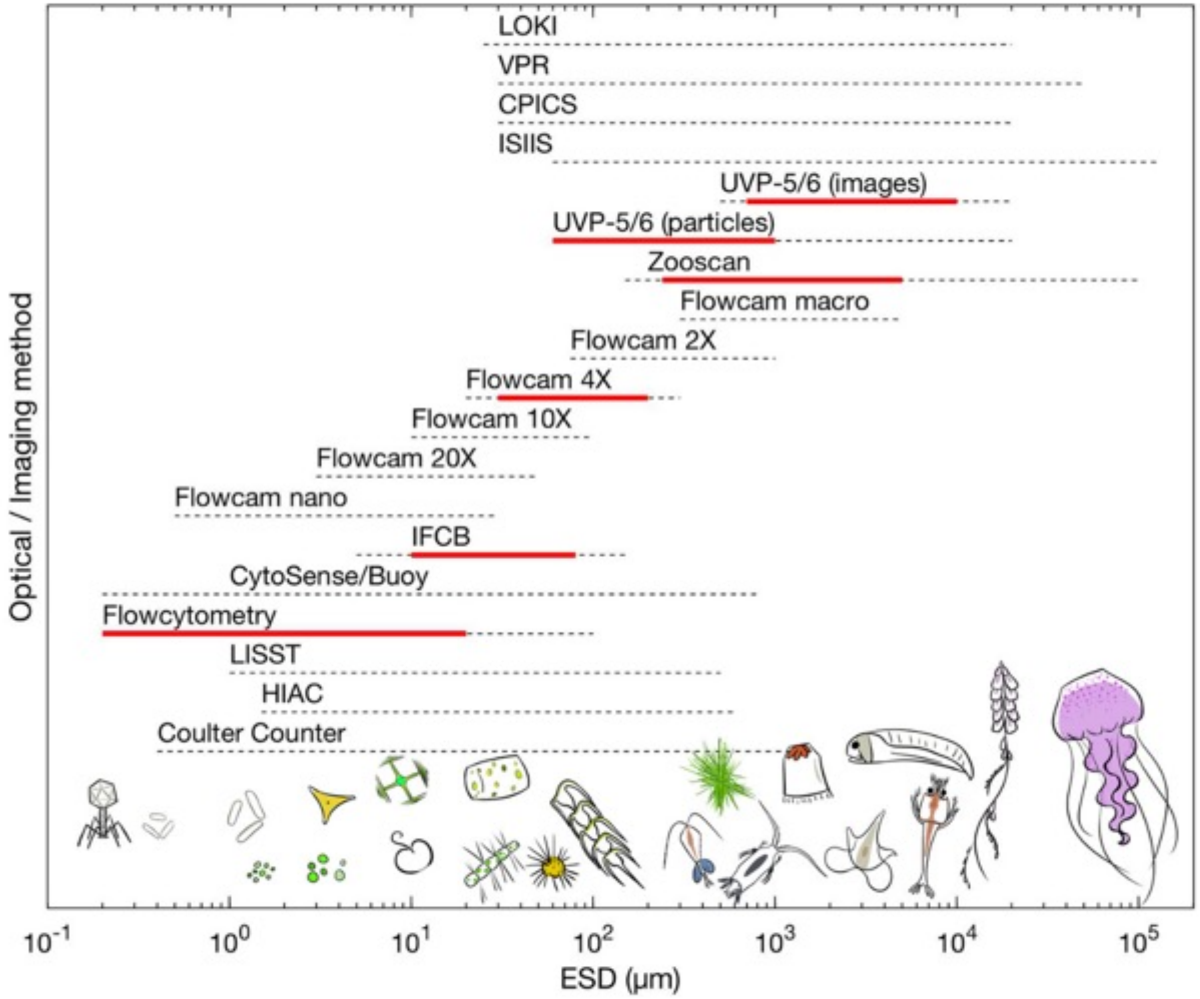
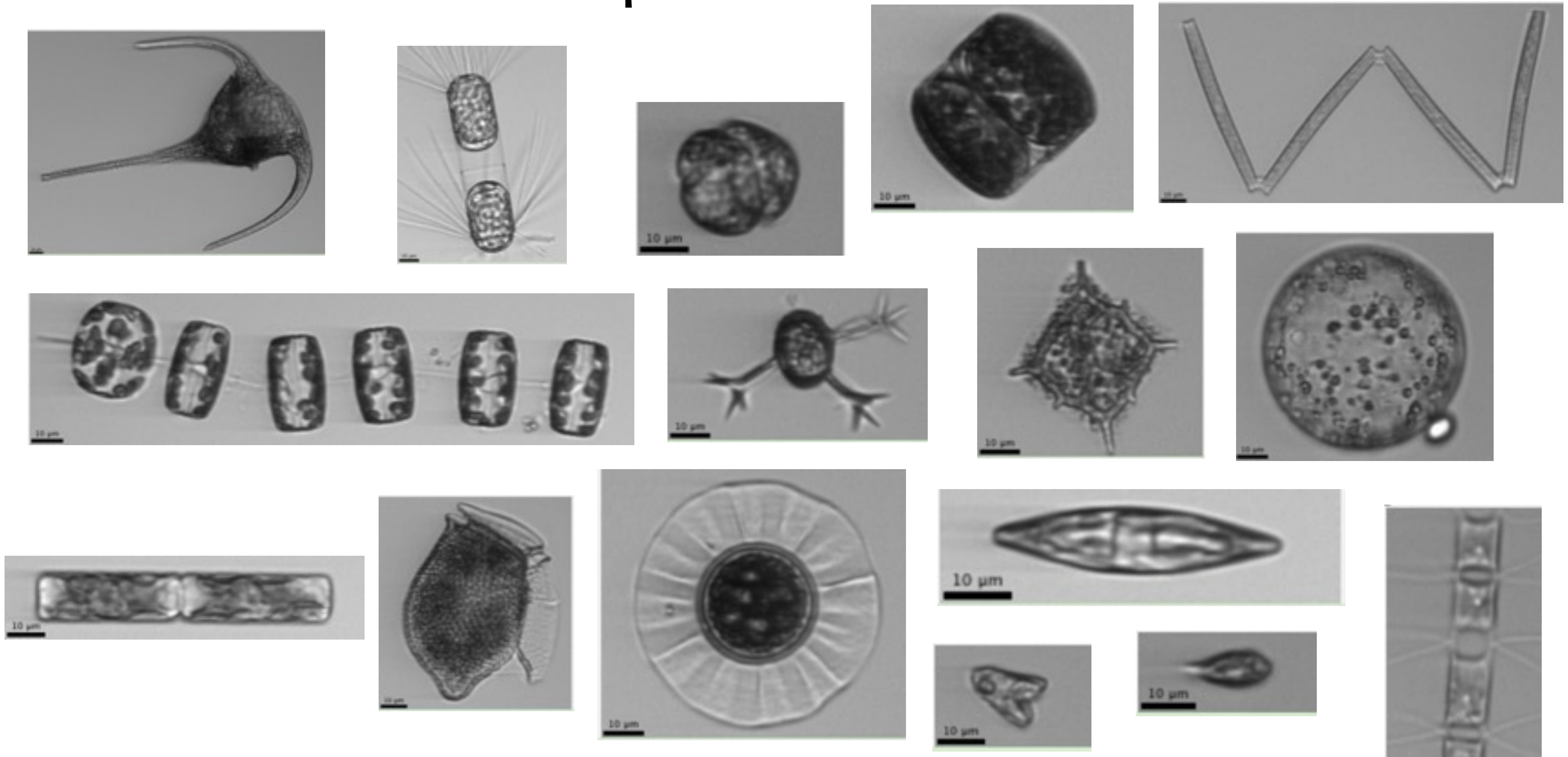


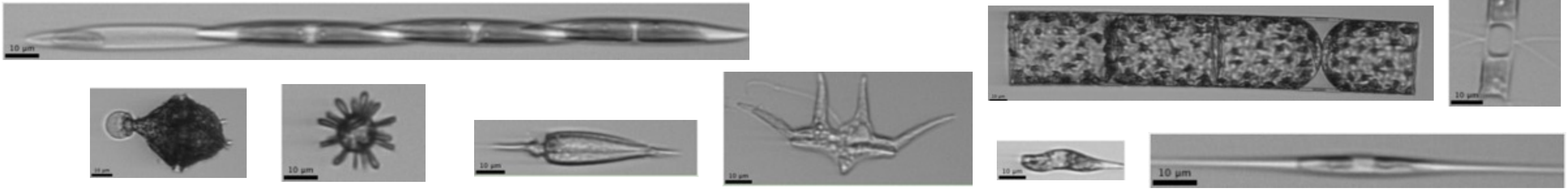
FIGURE 1 | Comparison of the total size range of plankton (in equivalent spherical diameter; ESD) that available optical and imaging methods can sample. Dashed lines represent the total operational size range from commercial information while the red line represent the practical size range which is efficient to obtain quantitative information, for an example see **Figure 2**. Drawings by Justine Courboules.

Lombard et al., 2019

Plankton imagery used to determine community composition of cells ~8-150 µm

Imaging FlowCytobot (IFCB)

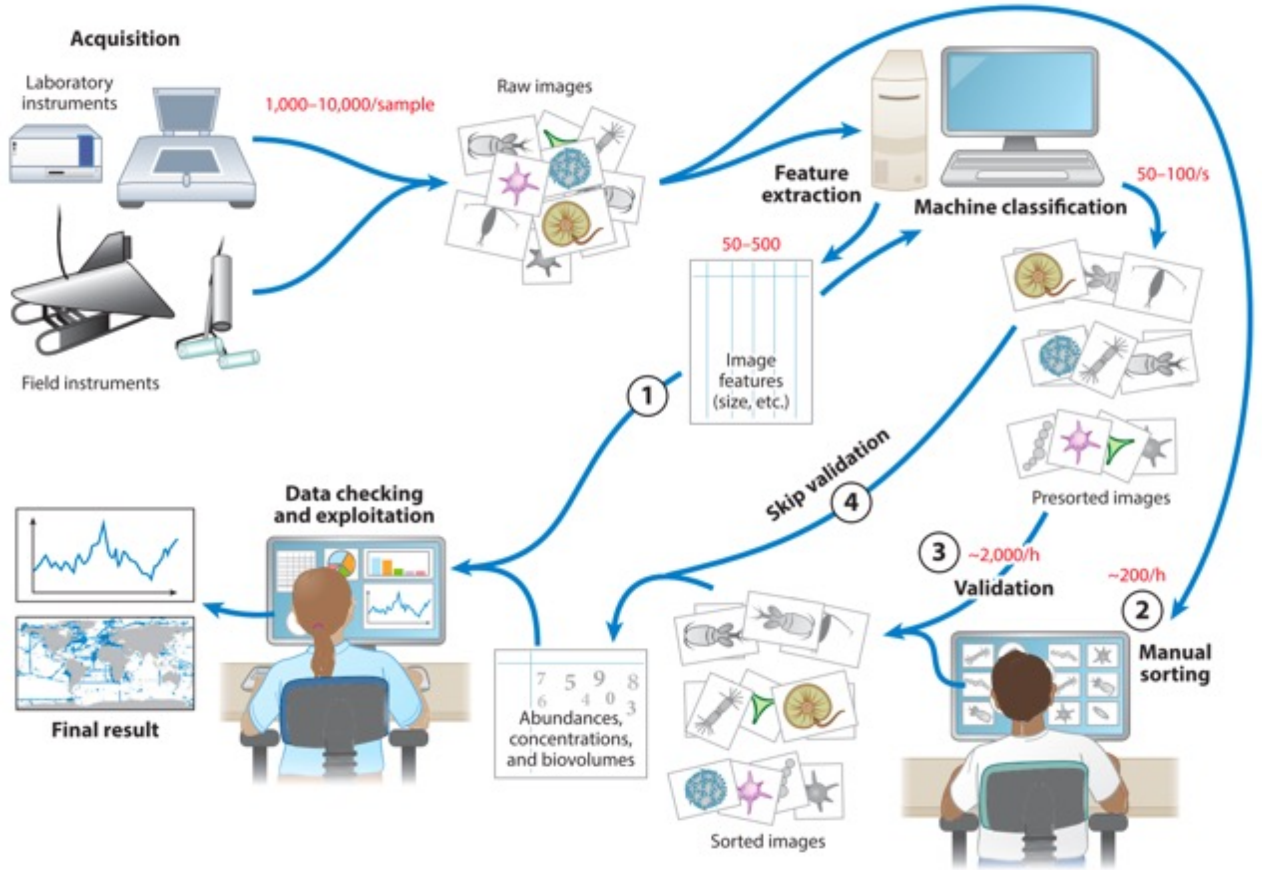




Annual Review of Marine Science

Machine Learning for the Study of Plankton and Marine Snow from Images

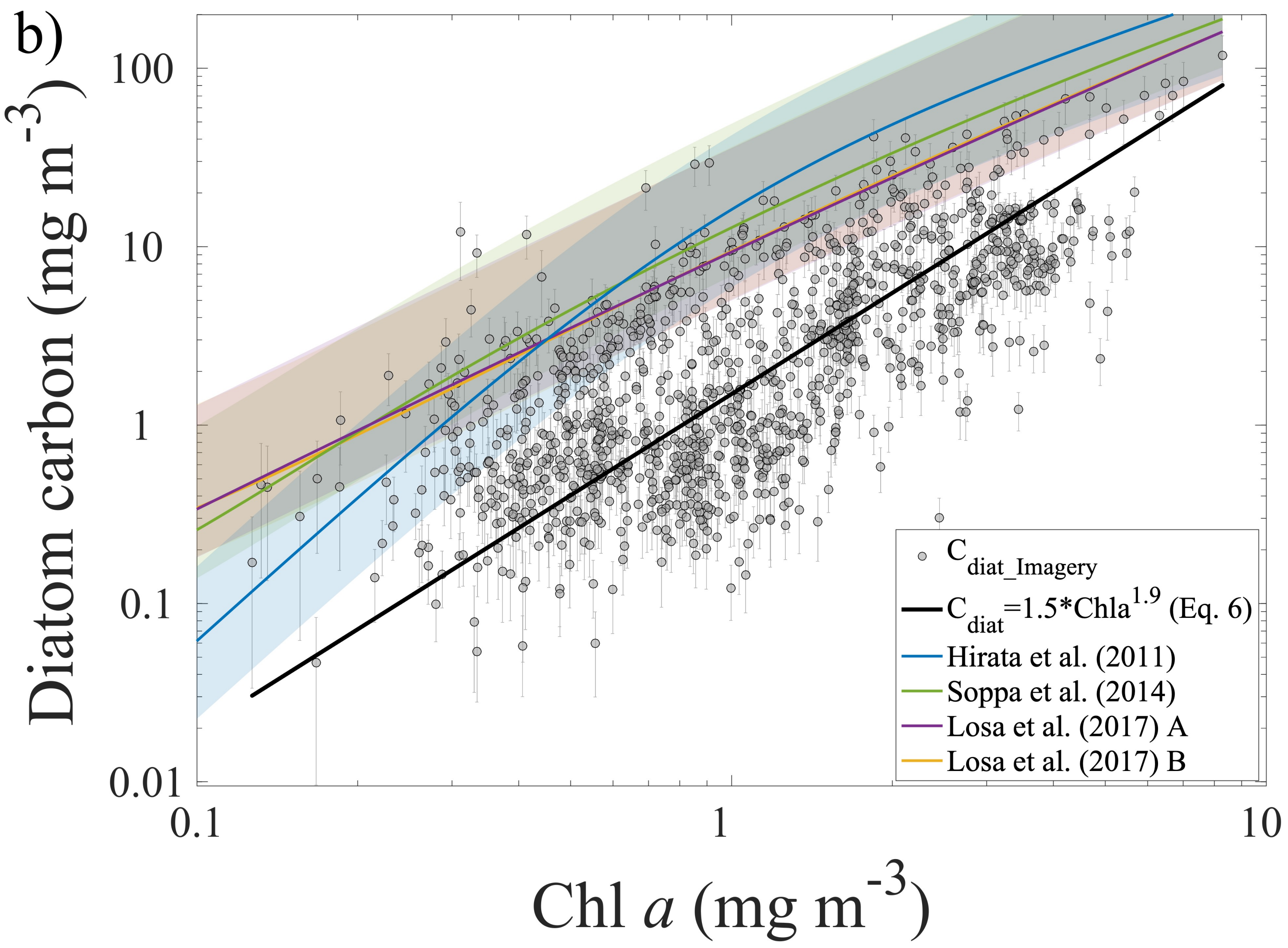
Jean-Olivier Irisson,[1] Sakina-Dorothée Ayata,[1] Dhugal J. Lindsay,[2] Lee Karp-Boss,[3] and Lars Stemmann[1]



→ Note that deep learning networks do not necessarily require a separate feature extraction step

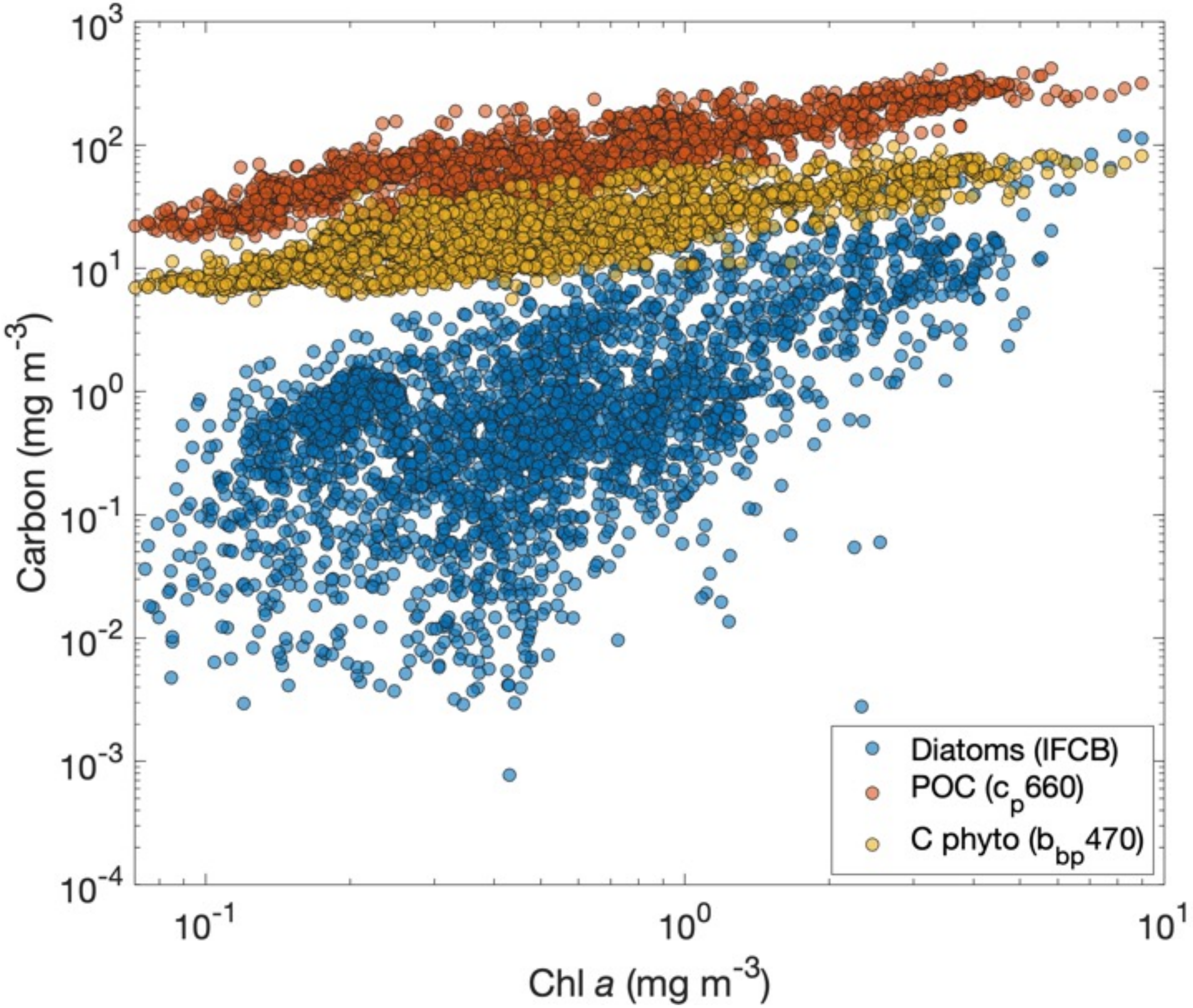
https://github.com/ifcb-utopia - Tools for processing IFCB images including a CNN demo w/a “toy” dataset

Variability in diatom carbon across chlorophyll *a* concentrations

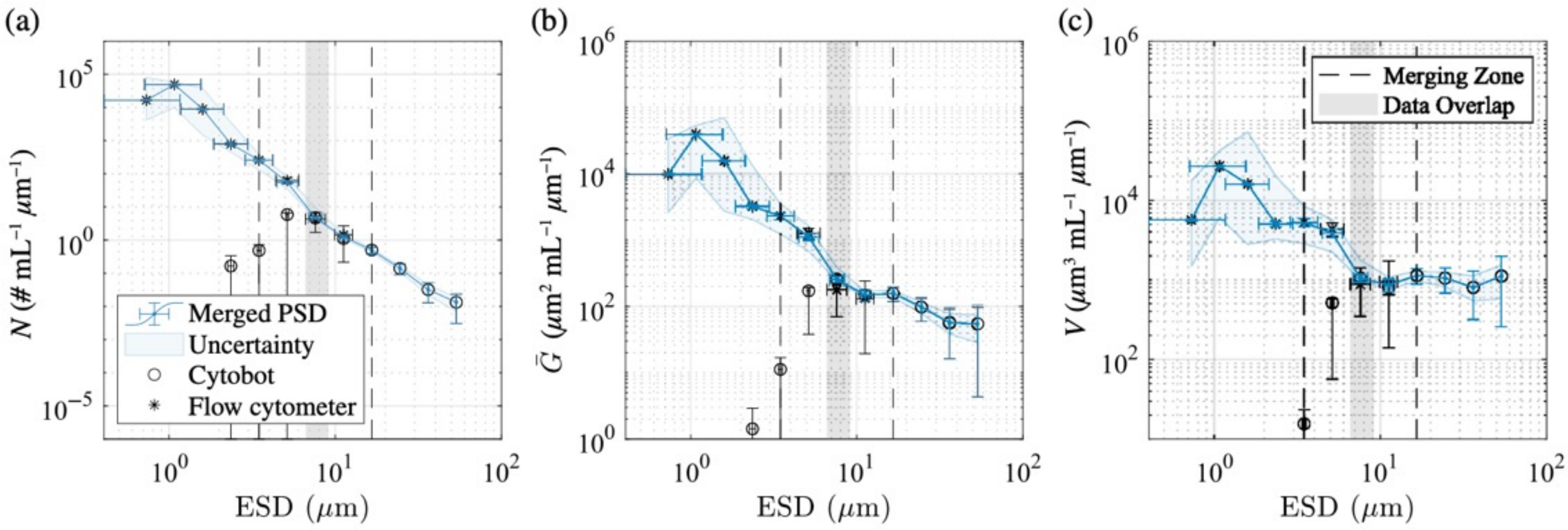


Chase et al., 2022

Variability in diatom carbon, phytoplankton carbon, and POC across chl *a*

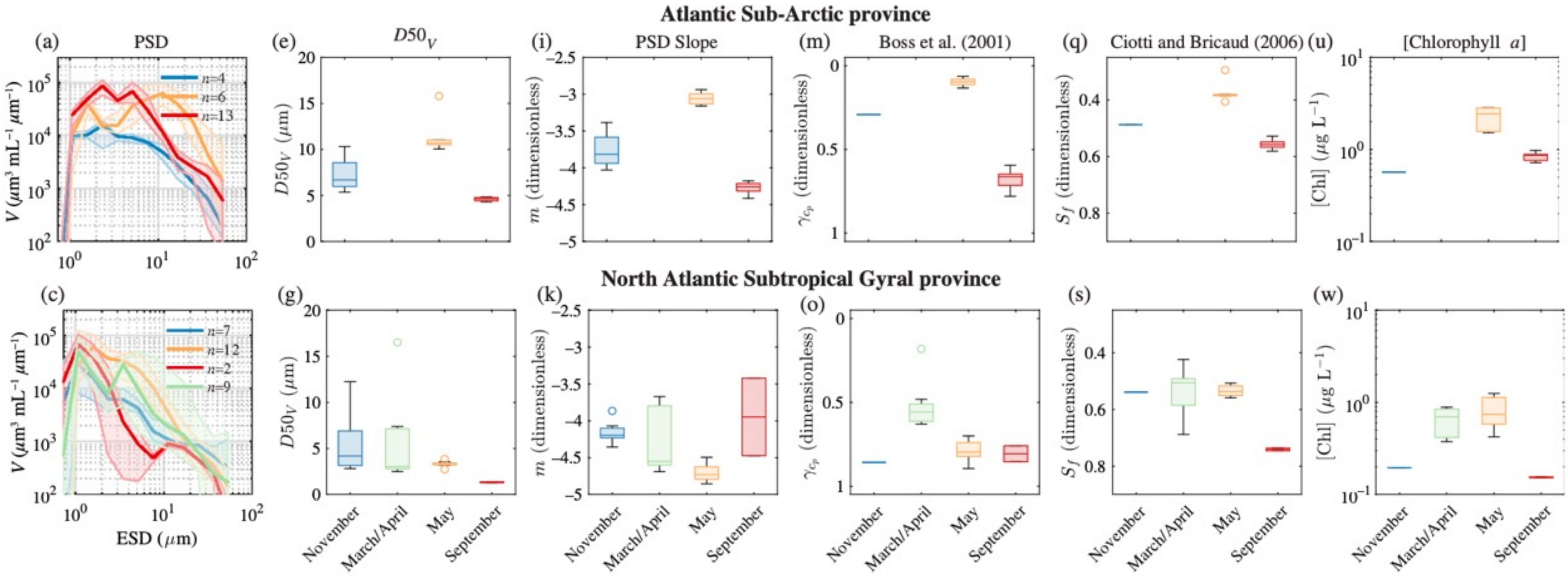


Merged cytometry-based phytoplankton size distributions



Haëntjens et al., 2022

PSDs and optical size proxies



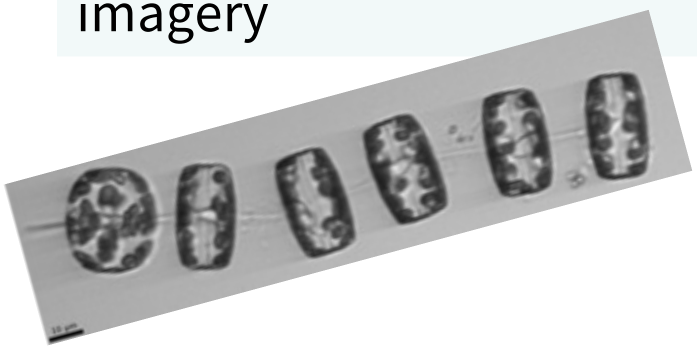
Haëntjens et al., 2022

Complementary machine learning workflows

Prepare phytoplankton imagery data and train CNNs

Apply trained network to unlabeled phytoplankton images

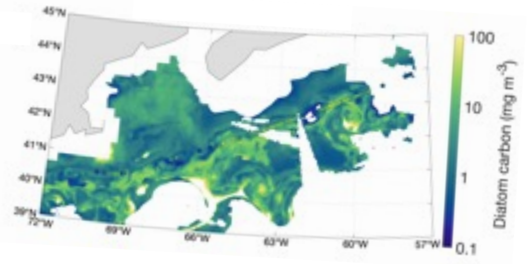
Database of classified plankton imagery



Prepare in situ datasets of environmental measurements & optical properties

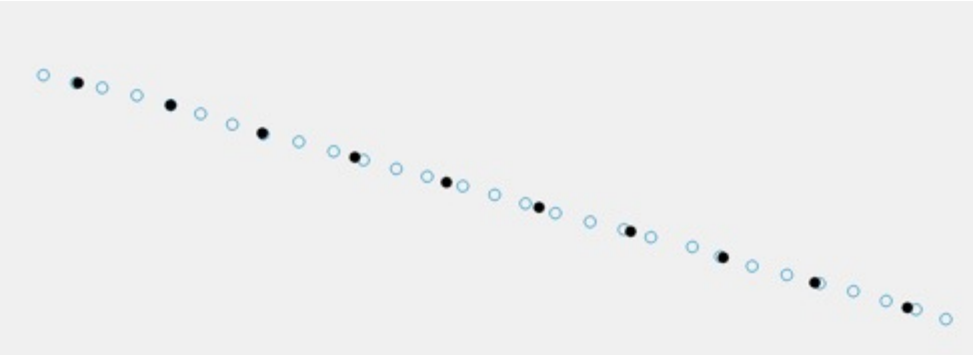
Assess how well data represent global conditions

Train ML networks to estimate diatom biomass from environmental/optical data

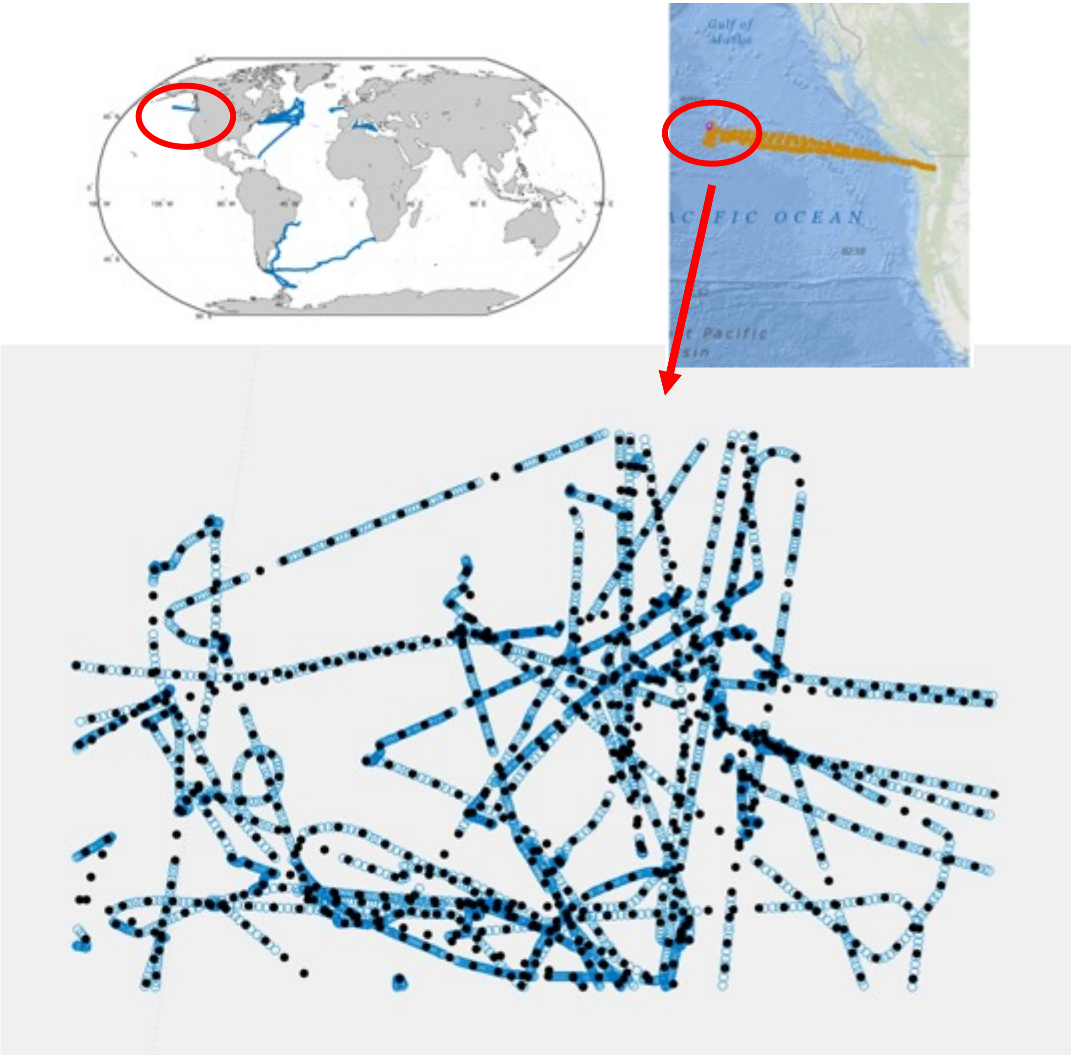


Averaging & merging all input data to a 1-km along-track "grid"

Above: example underway data 1-min binned data
Below: 1-min binned (open blue) & 1-km grid (black)

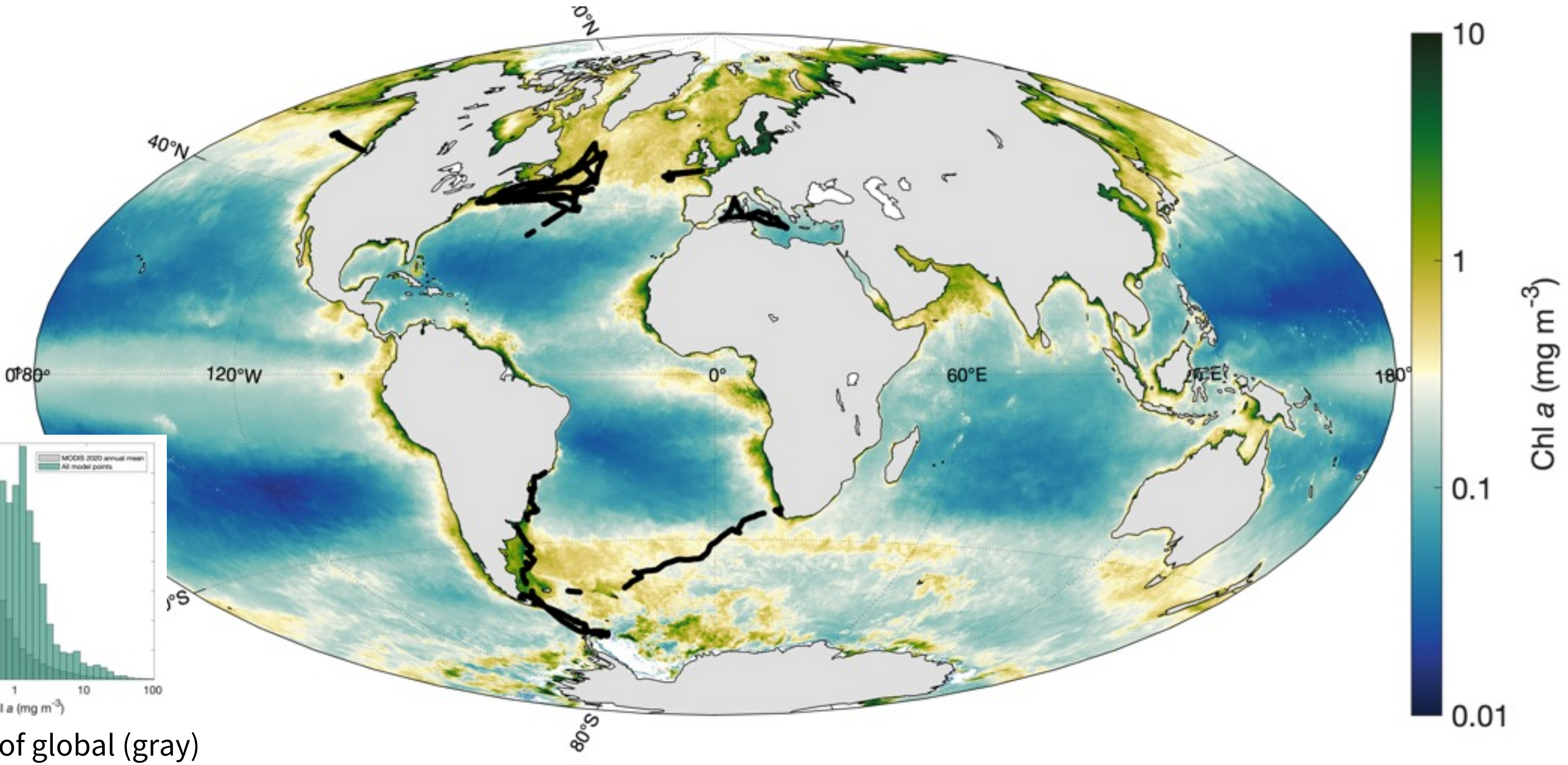


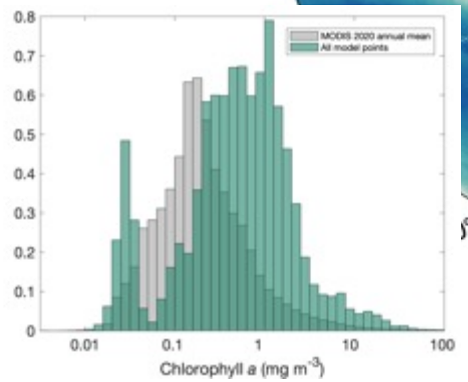
- Mean and standard deviation are stored during averaging for subsequent error propagation and uncertainty analysis
- Various types of datasets can be easily compared based on grid indices



Example when the ship is focusing on one location or feature in the ocean

How well do our model training data represent global conditions?



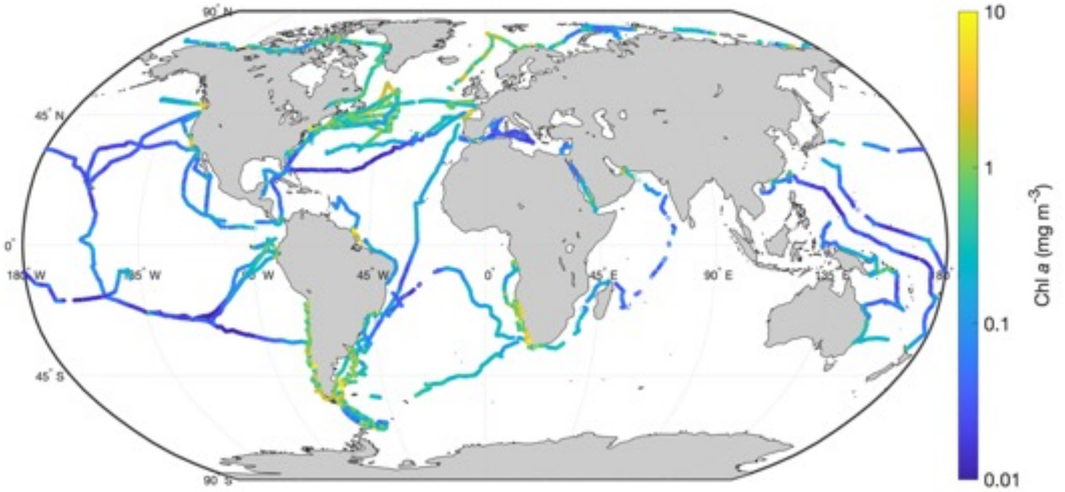


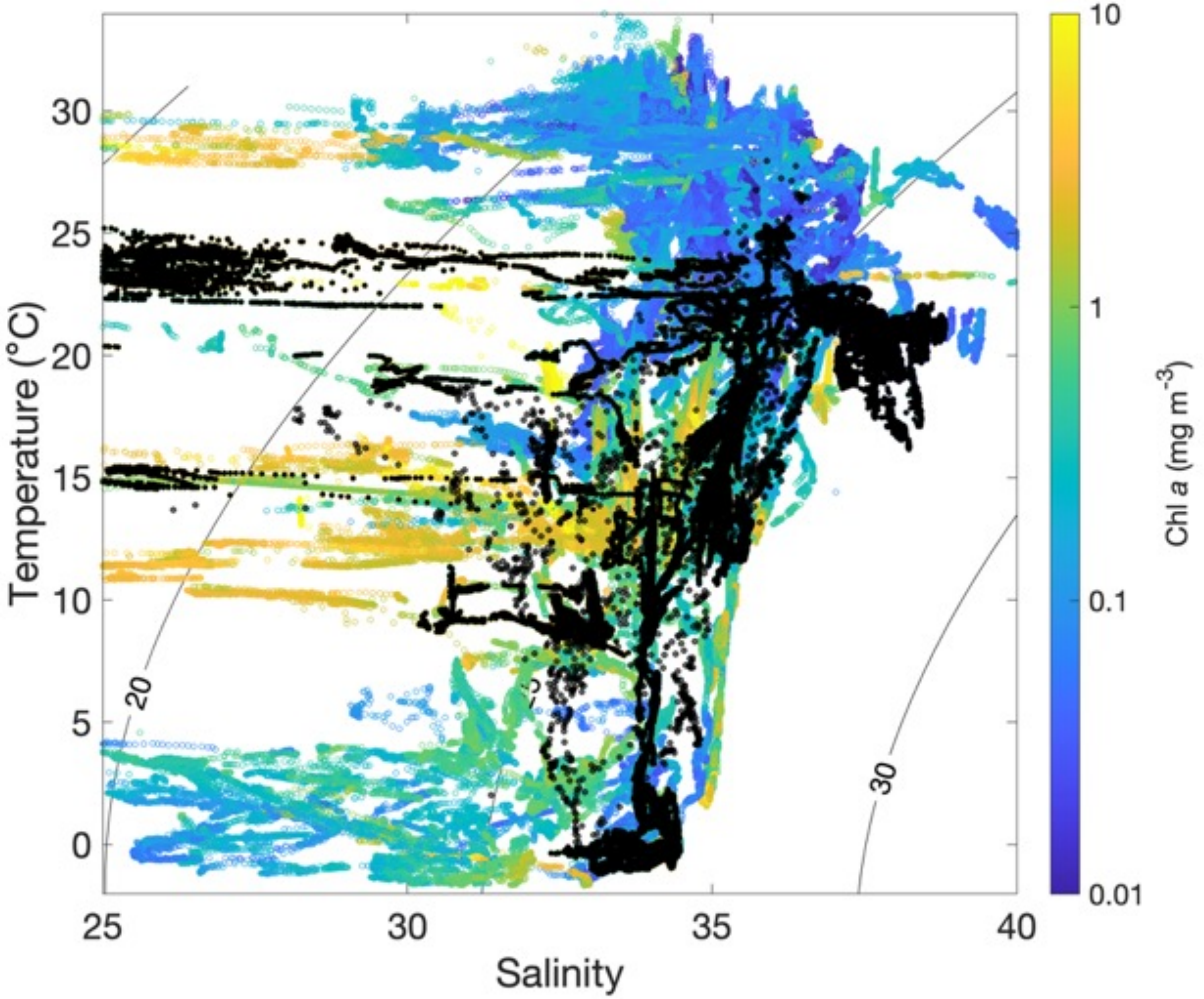
Distributions of global (gray) and algorithm training (green) chlorophyll *a*

NASA MODIS Aqua 2020 annual mean

Dots show locations with both plankton imagery data & model input variables

How well do our model training data represent large-scale variability?

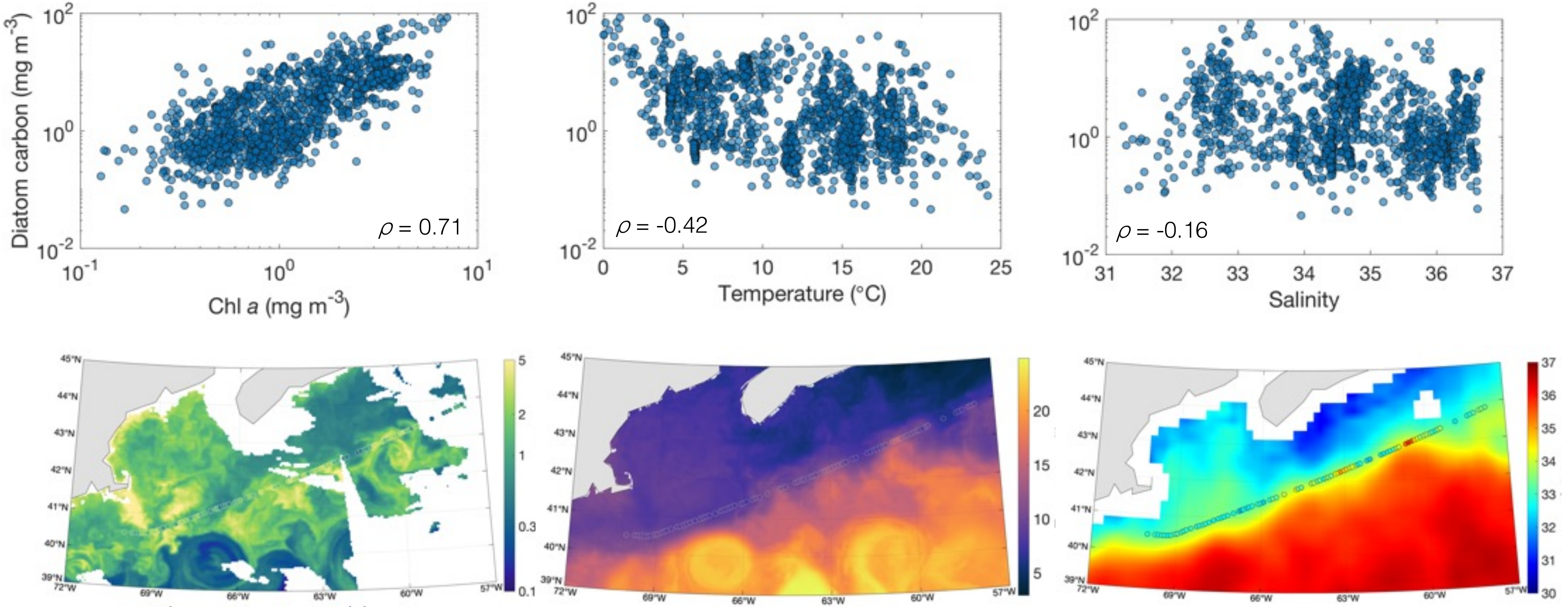




- Black dots at right represent cruises with both input data and plankton imagery and thus can be used for network training

- Some portions of the TS/Chl space are underrepresented

Addition of ancillary environmental data



Daily MODIS Aqua Chl *a*

Daily MUR SST product

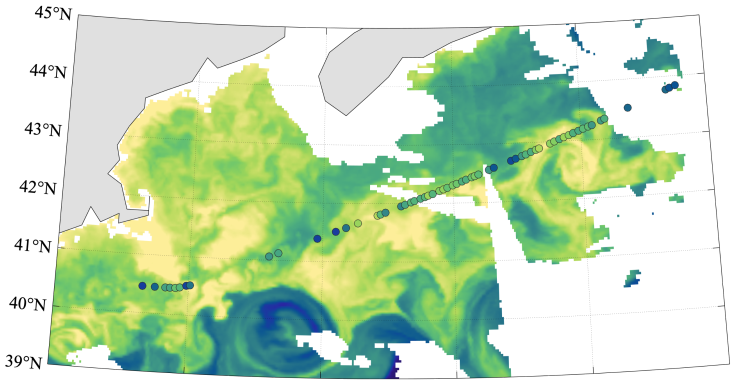
Monthly SMAP SSS

- Diatom carbon and environmental variables are correlated but with high variability
- Chl *a*, temperature, and salinity are all available from satellite

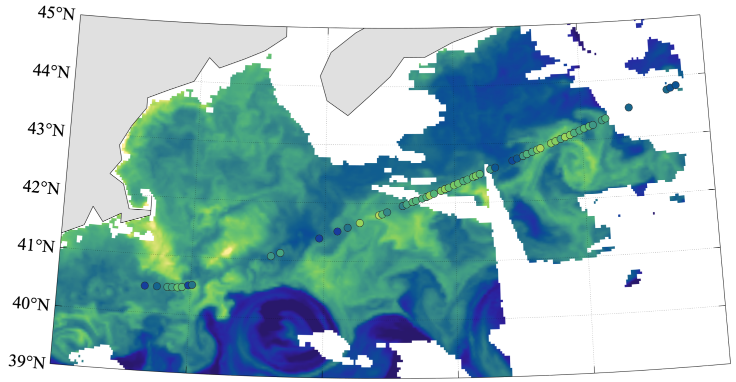
Chase et al. 2022

Environmental data + plankton imagery + machine learning

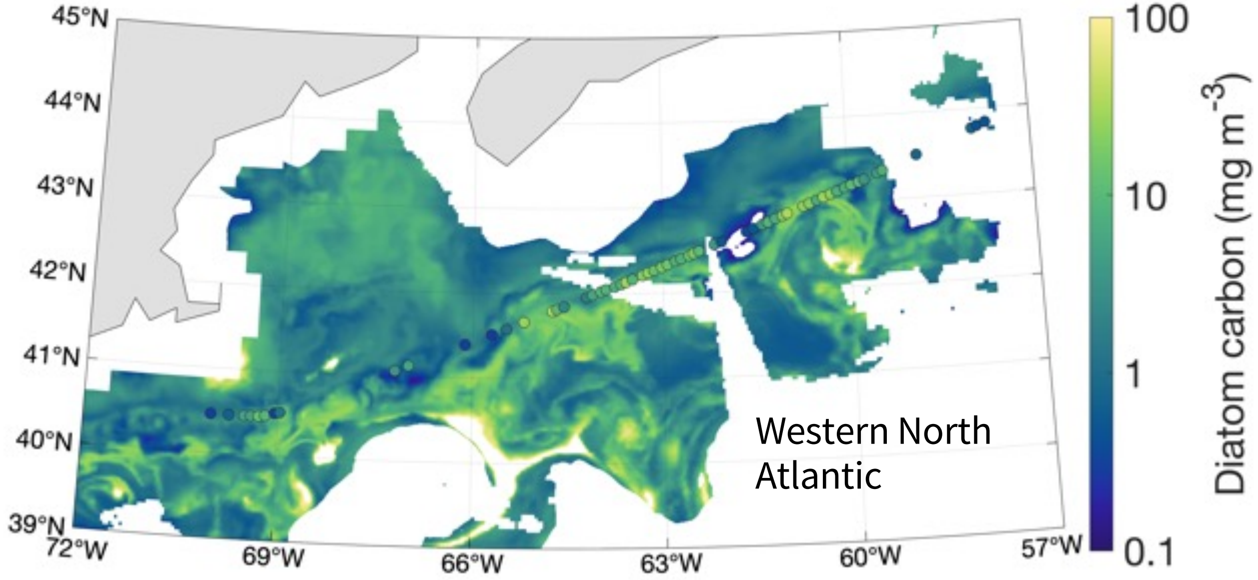
Previous Chl a-based method (Hirata et al. 2011)



Updated Chl a-based method



Neural network-based method



Chase et al. 2022

- Different spatial patterns are observed compared to estimates of diatoms solely from Chl *a*
- Full error propagation: uncertainty in diatom carbon = 65%

Chase et al. 2022

Uncertainty calculations are necessary!

Cell biovolume estimate

Diatom ID accuracy

$$\mathbf{Unc_{data}} = \sqrt{\mathbf{0.17^2 + 0.18^2 + 0.1^2 + 0.29^2}} = \mathbf{0.39},$$

Statistical counting error

Chl *a* uncertainty

$\mathbf{Unc_{data}}$

Neural network uncertainty

$$\mathbf{Unc_{NN}} = \sqrt{\mathbf{0.39^2 + 0.52^2}} = \mathbf{0.65},$$

At low estimated diatom carbon values, the absolute error dominates over the relative error, and thus Unc_{NN} = max(1.05 mg m^{-3}, 65%)

Is it good enough???

Overcoming the Challenges of Ocean Data Uncertainty

In oceanography, as in any scientific field, the goal is not to eliminate uncertainty in data, but instead to better quantify and clearly communicate its size and nature.

By **Shane Elipot**, **Kyla Drushka**, **Aneesh Subramanian**, and **Mike Patterson** 12 January 2022

> *An ocean data set may otherwise be of the highest scientific quality, but if quantified uncertainties do not accompany it, it will not be useful to scientists or other stakeholders.*

> *Some concepts that are applicable to bench measurements are difficult to translate to the oceanographer's laboratory—the ocean—because the ocean and the climate system in which it is embedded are constantly changing.*

This view from the International Space Station shows sea ice floes and eddy currents near the coast of Russia's Kamchatka Peninsula. Credit: NASA JSC Earth Science and Remote Sensing Unit

https://eos.org/opinions/overcoming-the-challenges-of-ocean-data-uncertainty

(near) real-time use of optics to locate features of interest

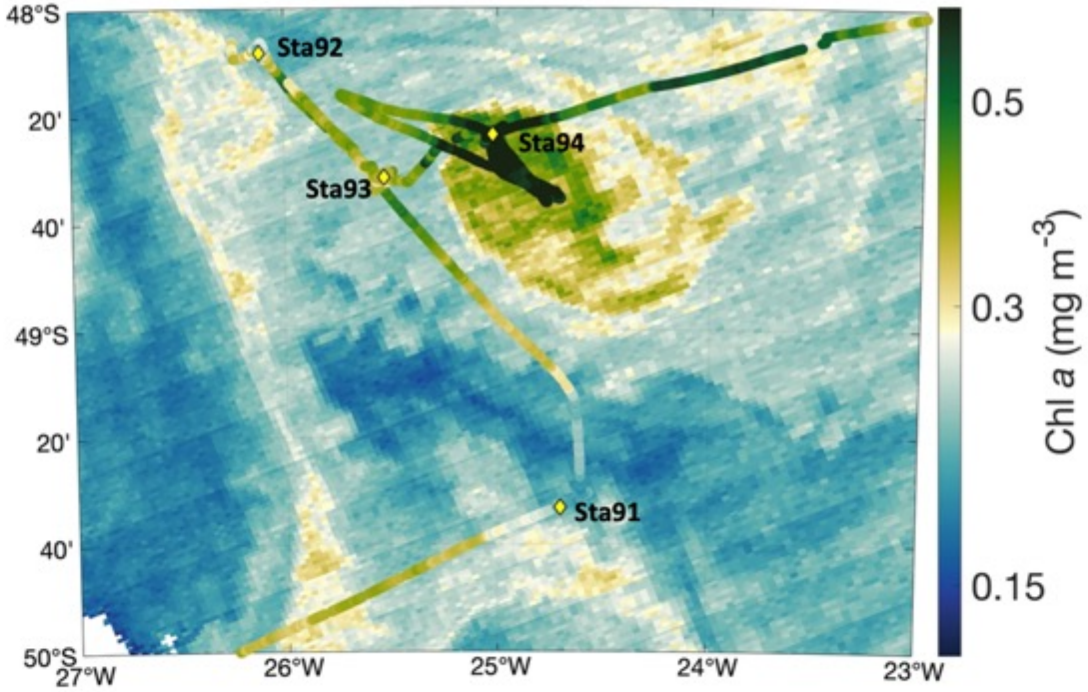
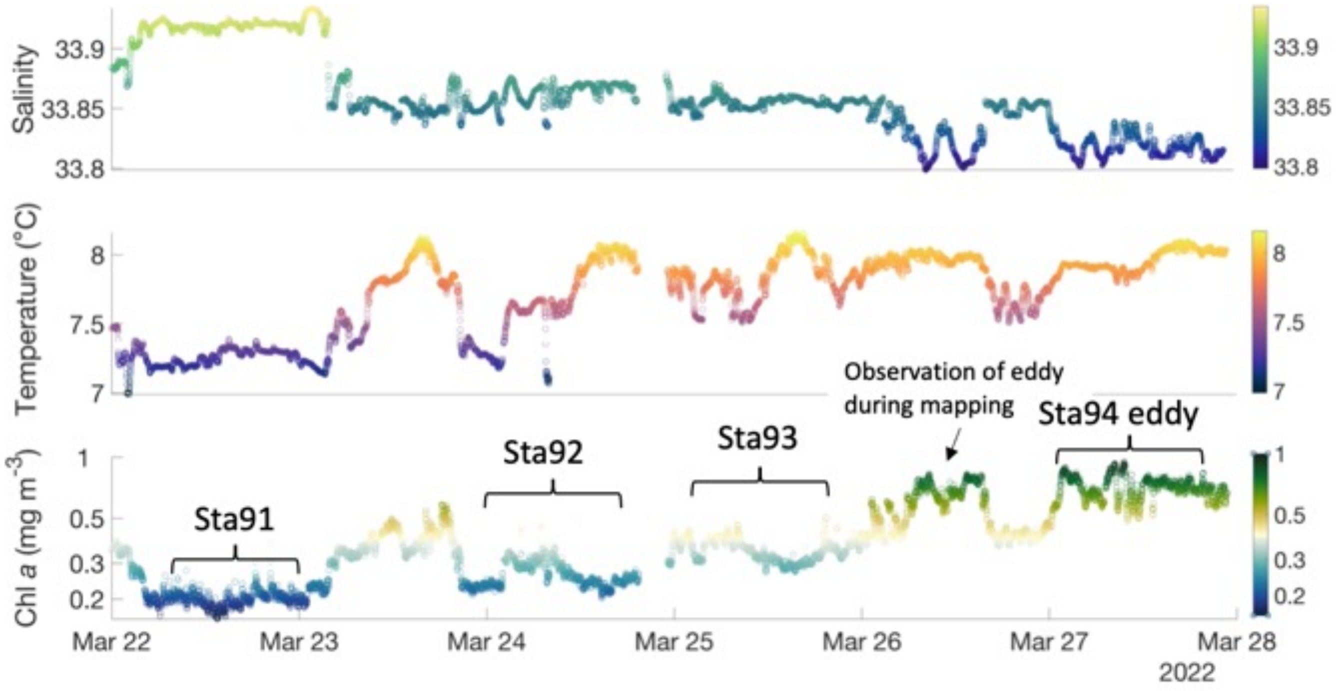
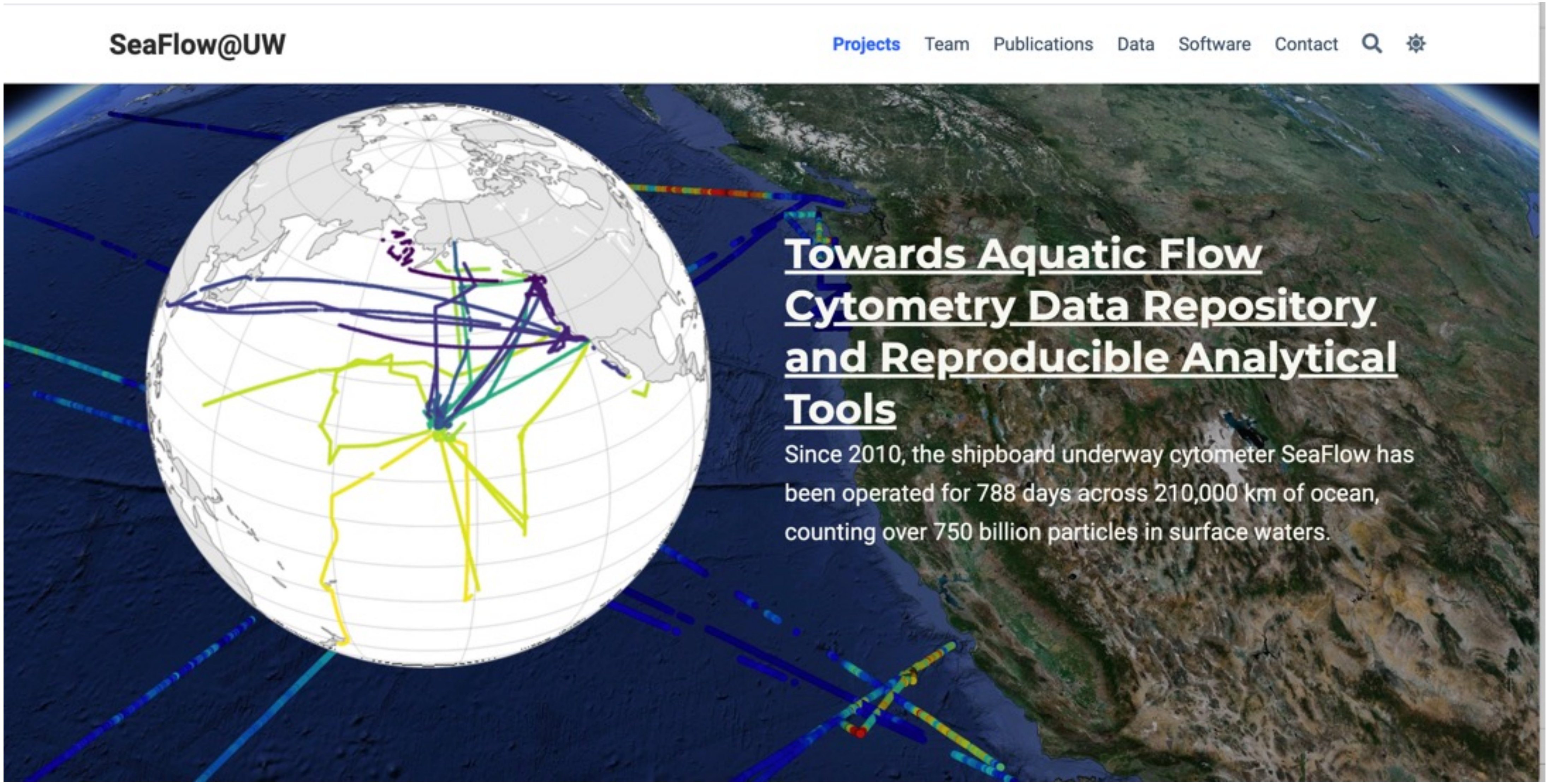


Figure 5. Satellite (MODIS AQUA) chlorophyll a (Chl a) from March 27. Ship-track from March 22-27 overlain with points colored by Chl a derived from absorption spectra, and with four stations labeled.



Open-source data repository of picoplankton in the open ocean



https://seaflow.netlify.app/

https://simonscmap.com/

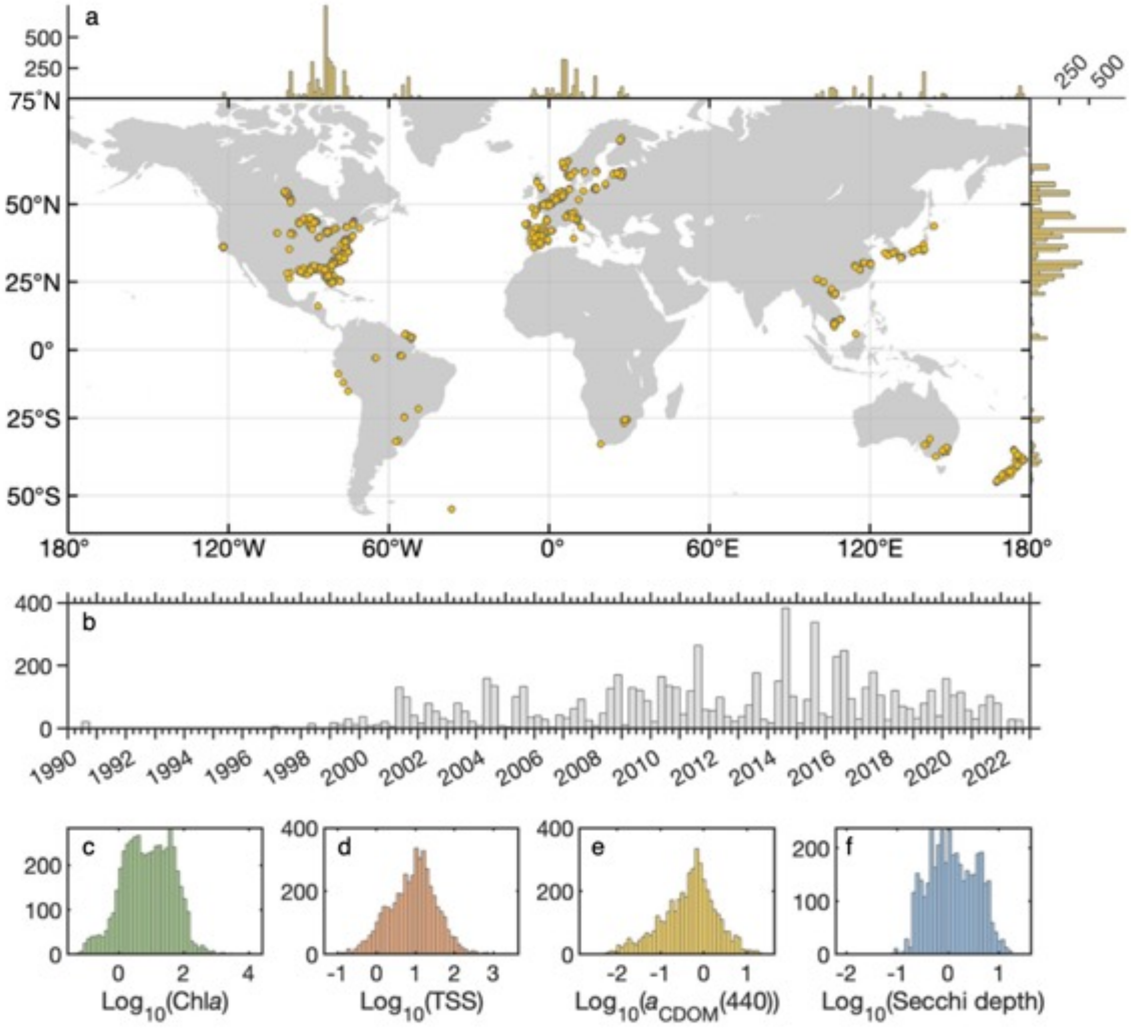
scientific **data**

OPEN

DATA DESCRIPTOR

GLORIA - A globally representative hyperspectral *in situ* dataset for optical sensing of water quality

Moritz K. Lehmann *et al.**



https://www.nature.com/articles/s41597-023-01973-y

How can we navigate the push-pull of the untapped potential in PCC algorithms and the inherent challenges?

Different questions will have different data needs. Consider when a given data product is applicable, and when it is not. ***What*** *do you want to know, and* ***why****?*

→ Consider scales of spatial and temporal variability
→ Remember that uncertainties "complete the data"

What are the major challenges in PCC algorithm work?

- Sensitivity of methods to the uncertainties in measured products and/or intermediate derived products

- Target variables (PCC groups) are often defined by proxy, ultimately limiting algorithm refinement

- Sufficient datasets for model development and testing are not trivial to collect

- Linking products to what is needed by end users (e.g., climate & ecosystem modelers, water quality management & HAB detection)

What are the exciting opportunities in PCC algorithm work?

- Advancements in data collection technology for assessing in situ PCC
- Hyperspectral satellite remote sensing & UAV data
- Increased application of machine learning and computing power advancements
- Incorporation of additional/ancillary data, both in situ and via combing data from multiple satellite platforms
- Improved models and data collection that in turn provide insights into finer spatial and temporal scale properties of ocean dynamics

Ocean Optics Summer Course, 2011
Darling Marine Center, UMaine


Thank you!

A few favorite resources for GitHub, python, and machine learning

Git – the simple guide:
https://rogerdudler.github.io/git-guide/

Data Analysis in python for oceanographers:
https://currents.soest.hawaii.edu/ocn_data_analysis/index.html

Recommendation from Patrick:
https://www.pythonlikeyoumeanit.com/

Tools for satellite data analysis designed by Patrick:
https://github.com/patrickcgray/open-geo-tutorial

Set of four videos that explain neural networks and deep/shallow learning:
https://www.youtube.com/watch?v=aircAruvnKk&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi

This website lets you play around with number of layers and neurons in a neural network and visualize the effects:
https://playground.tensorflow.org

General resource for clear explanations of math terms and concepts:
https://betterexplained.com/

References

Alvain, S., C. Moulin, Y. Dandonneau, and F.M. Bréon. 2005. "Remote Sensing of Phytoplankton Groups in Case 1 Waters from Global SeaWiFS Imagery." *Deep Sea Research Part I: Oceanographic Research Papers* 52 (11): 1989–2004. https://doi.org/10.1016/j.dsr.2005.06.015.

Bracher, A., M. Vountas, T. Dinter, J. P. Burrows, R. Röttgers, and I. Peeken. 2009. "Quantitative Observation of Cyanobacteria and Diatoms from Space Using PhytoDOAS on SCIAMACHY Data." *Biogeosciences* 6 (5): 751–64. https://doi.org/10.5194/bg-6-751-2009 .

Bracher, A., H. Bouman, R. J. Brewin, A. Bricaud, A. M Ciotti, Le. Clementson, E. Devred, et al. 2017. "Obtaining Phytoplankton Diversity from Ocean Color: A Scientific Roadmap for Future Development." *Frontiers in Marine Science* 4 (March): 1–15. https://doi.org/10.3389/fmars.2017.00055 .

Brewin, R. J. W., S. Sathyendranath, T.Hirata, S. J. Lavender, R. M. Barciela, and N. J. Hardman-Mountford. 2010. "A Three-Component Model of Phytoplankton Size Class for the Atlantic Ocean." *Ecological Modelling* 221 (11): 1472–83. https://doi.org/10.1016/j.ecolmodel.2010.02.014

Cael, B. B., Alison Chase, and Emmanuel Boss. 2020. "Information Content of Absorption Spectra and Implications for Ocean Color Inversion." *Applied Optics* 59 (13): 3971. https://doi.org/10.1364/ao.389189

Catlett, D., and D. A. Siegel. 2018. "Phytoplankton Pigment Communities Can Be Modeled Using Unique Relationships With Spectral Absorption Signatures in a Dynamic Coastal Environment." *Journal of Geophysical Research: Oceans*, 246–64. https://doi.org/10.1002/2017JC013195.

Chase, A., E. Boss, R. Zaneveld, A. Bricaud, H. Claustre, J. Ras, G. Dall'Olmo, and T.K. Westberry. 2013. "Decomposition of in Situ Particulate Absorption Spectra." *Methods in Oceanography* 7. https://doi.org/10.1016/j.mio.2014.02.002.

Chase, A. P., E. Boss, I. Cetinić, and W. Slade. 2017. "Estimation of Phytoplankton Accessory Pigments From Hyperspectral Reflectance Spectra: Toward a Global Algorithm." *Journal of Geophysical Research: Oceans* 122 (12): 9725–43. https://doi.org/10.1002/2017JC012859

Chase, A. P., Boss, E. S., Haëntjens, N., Culhane, E., Roesler, C., & Karp- Boss, L. 2022. "Plankton imagery data inform satellite-based estimates of diatom carbon". *Geophysical Research Letters*, *49*, e2022GL098076. https://doi.org/10.1029/2022GL098076

Chekalyuk, Alexander, and Mark Hafez. 2013. "Next Generation Advanced Laser Fluorometry (ALF) for Characterization of Natural Aquatic Environments: New Instruments." *Optics Express* 21 (12): 14181–201. https://doi.org/10.1364/OE.21.014181.

Dekker, Arnold G., and Nicole Pinnel (Eds). 2018. "Feasibility Study for an Aquatic Ecosystem Earth Observing System," 195. https://ceos.org/observations/documents/Feasibility-Study-for-an-Aquatic-Ecosystem-EOS-v.2-hi-res_05April2018.pdf

Dierssen, Heidi, Astrid Bracher, Vittorio Brando, Hubert Loisel, and Kevin Ruddick. 2020. “Data Needs for Hyperspectral Detection of Algal Diversity across the Globe.” *Oceanography*. Vol. 33. https://doi.org/10.5670/oceanog.2020.111.

Dierssen, Heidi M, Steven G Ackleson, Karen E Joyce, Erin L Hestir, Alexandre Castagna, Samantha Lavender, and Margaret A. McManus. 2021. “Living up to the Hype of Hyperspectral Aquatic Remote Sensing: Science, Resources and Outlook.” *Frontiers in Environmental Science* 9 (June): 1–26. https://doi.org/10.3389/fenvs.2021.649528

Giardino, C., V. E. Brando, P. Gege, N. Pinnel, E. Hochberg, E. Knaeps, I. Reusen, et al. 2019. “Imaging Spectrometry of Inland and Coastal Waters: State of the Art, Achievements and Perspectives.” *Surveys in Geophysics* 40 (3): 401–29. https://doi.org/10.1007/s10712-018-9476-0.

Gray, Patrick Clifton, Gregory D. Larsen, and David W. Johnston. 2021. “Drones Address an Observational Blind Spot for Biological Oceanography.” *Frontiers in Ecology and the Environment* 1–9. https://doi.org/10.1002/fee.2472.

Hirata, T., N. J. Hardman-Mountford, R. J. W. Brewin, J. Aiken, R. Barlow, K. Suzuki, T. Isada, et al. 2011. “Synoptic Relationships between Surface Chlorophyll-*a* and Diagnostic Pigments Specific to Phytoplankton Functional Types.” *Biogeosciences* 8 (2): 311–27. https://doi.org/10.5194/bg-8-311-2011.

Kramer, Sasha J., and David A. Siegel. 2019. “How Can Phytoplankton Pigments Be Best Used to Characterize Surface Ocean Phytoplankton Groups for Ocean Color Remote Sensing Algorithms?” *Journal of Geophysical Research: Oceans* 124 (11): 7557–74. https://doi.org/10.1029/2019JC015604.

Kramer, Sasha J., David A. Siegel, Stéphane Maritorena, and Dylan Catlett. 2022. “Modeling Surface Ocean Phytoplankton Pigments from Hyperspectral Remote Sensing Reflectance on Global Scales.” *Remote Sensing of Environment* 270 (December 2021). https://doi.org/10.1016/j.rse.2021.112879

Morel, André, and Louis Prieur. 1977. “Analysis of Variations in Ocean Color.” *Limnology and Oceanography* 22 (4): 709–22. https://doi.org/10.4319/lo.1977.22.4.0709

Mouw, C. B., N. J. Hardman-Mountford, S. Alvain, A. Bracher, R. J. W. Brewin, A. Bricaud, A. M. Ciotti, et al. 2017. “A Consumer’s Guide to Satellite Remote Sensing of Multiple Phytoplankton Groups in the Global Ocean.” *Frontiers in Marine Science* 4 (February). https://doi.org/10.3389/fmars.2017.00041.

O’Reilly, John E., and P. Jeremy Werdell. 2019. “Chlorophyll Algorithms for Ocean Color Sensors - OC4, OC5 & OC6.” *Remote Sensing of Environment* 229 (April): 32–47. https://doi.org/10.1016/j.rse.2019.04.021.

Organelli, Emanuele, Annick Bricaud, David Antoine, and Julia Uitz. 2013. “Multivariate Approach for the Retrieval of Phytoplankton Size Structure from Measured Light Absorption Spectra in the Mediterranean Sea (BOUSSOLE Site).” *Applied Optics* 52 (11): 2257–73. http://www.ncbi.nlm.nih.gov/pubmed/23670753.

Pahlevan, Nima, Brandon Smith, Caren Binding, Daniela Gurlin, Lin Li, Mariano Bresciani, and Claudia Giardino. 2020. “Hyperspectral Retrievals of Phytoplankton Absorption and Chlorophyll-a in Inland and Nearshore Coastal Waters.” *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2020.112200.

Ryan, John P., Curtiss O. Davis, Nicholas B. Tufillaro, Raphael M. Kudela, and Bo Cai Gao. 2014. "Application of the Hyperspectral Imager for the Coastal Ocean to Phytoplankton Ecology Studies in Monterey Bay, CA, USA." *Remote Sensing* 6 (2): 1007–25. https://doi.org/10.3390/rs6021007

Smith, Marié E, and Stewart Bernard. 2020. "Satellite Ocean Color Based Harmful Algal Bloom Indicators for Aquaculture Decision Support in the Southern Benguela." *Frontiers in Marine Science* 7 (February): 1–13. https://doi.org/10.3389/fmars.2020.00061.

Uitz, J., H.Claustre, A. Morel, and S. B. Hooker. 2006. "Vertical Distribution of Phytoplankton Communities in Open Ocean: An Assessment Based on Surface Chlorophyll." *Journal of Geophysical Research* 111 (C8): C08005. https://doi.org/10.1029/2005JC003207 .

Uitz, Julia, Dariusz Stramski, Rick A Reynolds, and Jean Dubranna. 2015. "Assessing Phytoplankton Community Composition from Hyperspectral Measurements of Phytoplankton Absorption Coefficient and Remote-Sensing Reflectance in Open-Ocean Environments." *Remote Sensing of Environment* 171: 58–74. https://doi.org/http://dx.doi.org/10.1016/j.rse.2015.09.027

Vandermeulen, Ryan A., Antonio Mannino, Aimee R. Neeley, P. Jeremy Werdell, and Robert Arnone. 2017. "Determining the Optimal Spectral Sampling Frequency and Uncertainty Thresholds for Hyperspectral Remote Sensing of Ocean Color." *Optics Express* 25 (16): 785–97. https://doi.org/10.1364/OE.25.00A785

Vandermeulen, Ryan A., Antonio Mannino, Susanne E. Craig, and P. Jeremy Werdell. 2020. "150 Shades of Green: Using the Full Spectrum of Remote Sensing Reflectance to Elucidate Color Shifts in the Ocean." *Remote Sensing of Environment* 247 (May): 111900. https://doi.org/10.1016/j.rse.2020.111900

Werdell, P. J., C.. S Roesler, and J. I. Goes. 2014. "Discrimination of Phytoplankton Functional Groups Using an Ocean Reflectance Inversion Model." *Applied Optics* 53 (22): 4833–49. http://www.ncbi.nlm.nih.gov/pubmed/25090312.

Werdell, P. Jeremy, Lachlan I.W. McKinna, Emmanuel Boss, Steven G. Ackleson, Susanne E. Craig, Watson W. Gregg, Zhongping Lee, et al. 2018. "An Overview of Approaches and Challenges for Retrieving Marine Inherent Optical Properties from Ocean Color Remote Sensing." *Progress in Oceanography* 160 (January): 186–212. https://doi.org/10.1016/j.pocean.2018.01.001.

Wolanin, Aleksandra, Mariana Soppa, and Astrid Bracher. 2016. "Investigation of Spectral Band Requirements for Improving Retrievals of Phytoplankton Functional Types." *Remote Sensing* 8 (10): 871. https://doi.org/10.3390/rs8100871