LIMNOLOGY
and
OCEANOGRAPHY

ASLO

Limnol. Oceanogr. 61, 2016, 1472–1494

doi: 10.1002/lno.10316

Optical classification and characterization of marine particle assemblages within the western Arctic Ocean

Griet Neukermans,[a]* Rick A. Reynolds, Dariusz Stramski
Marine Physical Laboratory, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California

Abstract

We develop an optical classification of marine particle assemblages from an extensive dataset of particle optical properties collected in the Chukchi and Beaufort Seas. Hierarchical cluster analysis of the spectral particulate backscattering-to-absorption ratio partitioned the dataset into seven optically-distinct clusters of particle assemblages, each associated with different characteristics of particle concentration, composition, and phytoplankton taxonomic composition and size. Three phytoplankton-dominated clusters were identified. One was characterized by small-sized phytoplankton that are typically associated with regenerated production, and comprised samples from the subsurface chlorophyll-*a* maximum in oligotrophic waters of the Beaufort Sea. The other two clusters represented diatom-dominated particle assemblages in turbid shelf waters with differing contributions of photoprotective pigments. Such situations are generally associated with significant new production. Two clusters were dominated by organic nonalgal material, one representing clear waters off the shelf, the other representative of post-diatom bloom conditions in the Chukchi Sea. Another distinct cluster represented mineral-dominated particle assemblages that were observed in the Colville and Mackenzie River plumes and near the seafloor. Finally, samples in a cluster of mixed particle composition were scattered throughout all locations. Optical classification improved performance of predictive bio-optical relationships. These results demonstrate a capability to discriminate distinct assemblages of suspended particles associated with specific ecological conditions from hyperspectral measurements of optical properties, and the potential for identification of ecological provinces at synoptic time and space scales from optical sensors. Analogous analysis of multispectral optical data strongly reduced this capability.

Particles in seawater are a ubiquitous and important component of the marine environment and comprise both living organisms and nonliving material. They directly influence optical properties of seawater that are measurable from various in situ and remote platforms, thus enabling observations of marine particle assemblages over a wide range of temporal and spatial scales. Sensors measuring the inherent optical properties (IOPs) of seawater at a given wavelength of light λ (in vacuo), such as the spectral coefficients for absorption, $a(\lambda)$, scattering, $b(\lambda)$, backscattering, $b_b(\lambda)$, and beam attenuation, $c(\lambda) = a(\lambda) + b(\lambda)$, are now routinely deployed on various in situ platforms including ships, moorings, profilers, and autonomous underwater vehicles. Such data are becoming increasingly important for ocean biogeochemical and ecosystem studies (e.g., Johnson et al. 2009; Fennel et al. 2011). Remote sensing of ocean color from satellites or aircraft, quantified as the spectral remote-sensing reflectance, $R_{rs}(\lambda)$, also allow estimation of seawater IOPs as $R_{rs}(\lambda)$ is, to first order, proportional to the ratio of $b_b(\lambda)$ to $a(\lambda)$ (Gordon et al. 1988; IOCCG 2006). These two IOPs are of direct interest to this study.

Recent technological advances have enabled remote and in situ measurements of ocean optical properties at increasingly higher spectral resolution, from multispectral to hyperspectral acquisition systems (Chang et al. 2004). Examples include the WETLabs ac-s instrument, which provides in situ measurements of $c(\lambda)$ and $a(\lambda)$ at 4 nm spectral resolution, and the Hyperspectral Imager for the Coastal Ocean satellite sensor (HICO), for determining $R_{rs}(\lambda)$ at 5.7 nm spectral resolution. Various space agencies are planning to launch hyperspectral ocean color sensors in the next decade, such as the U.S. National Aeronautics and Space Administration (NASA)

*Correspondence: Griet.Neukermans@takuvik.ulaval.ca

[a]Present address: Takuvik Joint International Laboratory, Université Laval, Québec, Canada

HyspIRI (Hyperspectral Infrared Imager), GEO-CAPE (GEOstationary Coastal and Air Pollution Events), and PACE (current acronym stands for Pre-Aerosol, Clouds, and Ocean Ecosystem) missions, or the German HIS (Hyperspectral Imager) on EnMAP (Environmental Mapping and Analysis Program) mission.

While optical technology has undergone rapid advancement, there is a strong need for improved understanding of the linkages between the magnitude and spectral shape of IOPs and the various optically-significant constituents of seawater. The IOPs of seawater result from the additive contributions of water molecules and the various dissolved and particulate components of seawater, yielding the following operational partitioning of $a(\lambda)$ and $b_b(\lambda)$:

$$a(\lambda)=a_w(\lambda)+a_g(\lambda)+a_p(\lambda), \quad \text{with} \quad a_p(\lambda)=a_d(\lambda)+a_{ph}(\lambda) \tag{1}$$

$$b_b(\lambda)=b_{bw}(\lambda)+b_{bp}(\lambda) \tag{2}$$

where the *w*, *p*, *g*, *ph*, and *d* subscripts denote the respective contributions from water, particles, colored dissolved organic matter (CDOM), phytoplankton, and nonalgal particles. To first order, the magnitude of the particulate absorption and backscattering coefficient at a given wavelength varies with the concentration of particles, with second order effects arising from variability in the composition (i.e., refractive index), size distribution, and shape of suspended particles (e.g., Babin et al. 2003a,2003b; Stramski et al. 2007; Neukermans et al. 2012).

The spectral shape of these IOPs may also reveal information on the characteristics of the particulate assemblage. For instance, the spectral shape of $a_{ph}(\lambda)$ is primarily controlled by the concentration of various photosynthetic and photoprotective pigments and by pigment packaging within algal cells, which depends on cell size and intracellular pigment concentrations (Morel and Bricaud 1981). Analysis of the spectral shape of $a_{ph}(\lambda)$ can therefore reveal information on taxonomic composition and dominant size of the algal population (e.g., Ciotti et al. 2002; Bricaud et al. 2004). The wavelength dependency of the $b_{bp}(\lambda)$ coefficient is commonly modeled as a power law (Gordon and Morel 1983), which, according to Mie scattering theory, steepens with increasing slope of the particle size distribution (e.g., Morel 1973; Kostadinov et al. 2009). A limited number of field observations have provided general support for these results (Loisel et al. 2006; Slade and Boss 2015). Lastly, the spectral shape of $a_d(\lambda)$ is commonly modeled as an exponentially decreasing function of wavelength and it has been hypothesized that the slope parameter of $a_d(\lambda)$ depends on the composition of the nonalgal matter (Babin and Stramski 2004; Estapa et al. 2012).

Much effort has been devoted to advancing our understanding of optical variability with changes in the characteristics of the phytoplankton assemblage, such as community composition, physiology, and size structure, which reflect differences in biogeochemical functioning of the phytoplankton community (IOCCG 2014). Retrieval of these characteristics is of critical importance to studies in ecological modeling and biogeochemical cycling, from which the concept of phytoplankton functional types (PFTs) emerged. The term PFT defines a group of phytoplankton taxa that have a common biogeochemical role (Le Quéré et al. 2005; IOCCG 2014). Optical differentiation among a few PFTs has been achieved in open ocean waters using multispectral ocean color data (Alvain et al. 2005; Bricaud et al. 2012) and it is expected that hyperspectral satellite sensors will enhance the capability of PFT identification (e.g., Lubac et al. 2008; IOCCG 2014).

Even though the PFT concept provides an important bridge between marine optics and biogeochemistry, it does not account for nonalgal particles, which also contribute to optical variability and represent an essential component of marine ecosystem and biogeochemical models (Fasham et al. 1990; Dunne et al. 2005; Stukel et al. 2014). For example, waters characterized by the same PFT but with increasing contributions of detrital material may reflect progressive stages of a phytoplankton bloom. In a previous study (Neukermans et al. 2014), we demonstrated the potential for optical differentiation of peak diatom bloom from post-diatom bloom conditions in Arctic waters based on spectral changes in $b_{bp}(\lambda)$ and $a_p(\lambda)$. Other particle assemblages representative of distinct ecological conditions, such as a regime of regenerated production or waters influenced by resuspension of mineral particles, were also differentiated optically. These findings demonstrated that the spectral coefficients of $b_{bp}(\lambda)$ and $a_p(\lambda)$ provide complementary information on the characteristics of the particle assemblage owing to their different sensitivities to particle concentration, bulk composition, and size distribution, which are important indicators of ecological conditions.

In the present study, we investigate the synergistic use of $b_{bp}(\lambda)$ and $a_p(\lambda)$ for differentiation of marine particle assemblages representative of distinct ecological conditions. We use an extensive set of field data collected in the western Arctic Ocean comprising measurements of IOPs in conjunction with a detailed characterization of the particulate assemblage. We utilize a cluster analysis of the spectral shape of the particulate backscattering-to-absorption ratio, $b_{bp}(\lambda)/a_p(\lambda)$, to partition the dataset into optical classes. The $b_{bp}(\lambda)/a_p(\lambda)$ ratio largely removes first order dependency on particle concentration, and is thus expected to mainly reflect changes in bulk composition and size distribution of the particle assemblage. It represents the proportion of photons backscattered relative to photons absorbed by particles and can be thought of as the approximate contribution of particles to the ocean color remote-sensing reflectance.

We demonstrate that each optical class represents a particle assemblage characterized by a distinct set of indicators of particle concentration, composition, and phytoplankton taxonomic composition and size, and that these optically-derived particle assemblages (hereafter referred to as OPAs) are representative of distinct ecological conditions. We further illustrate

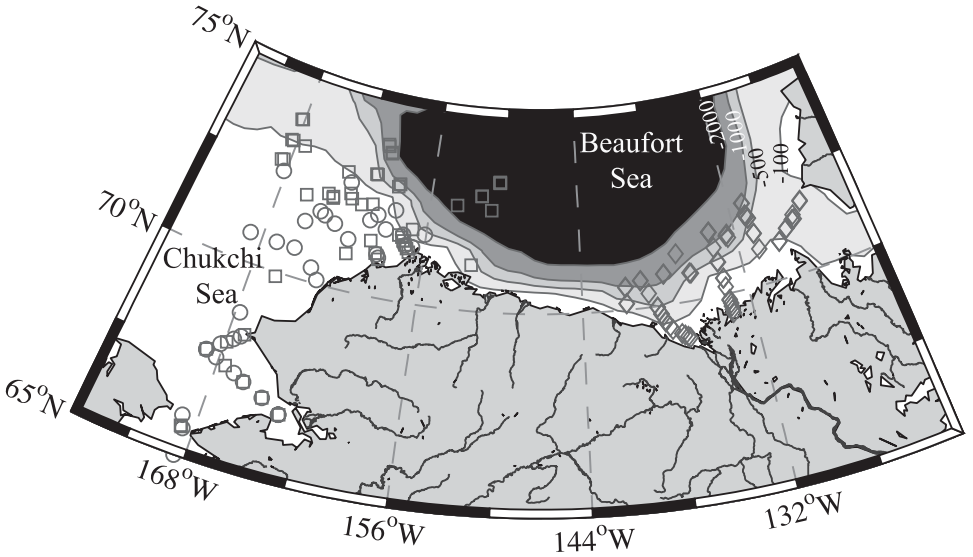

Fig. 1. Map of the Chukchi and Beaufort Seas showing the location of stations sampled during the MALINA campaign in 2009 (diamonds) and the ICESCAPE campaigns in 2010 (circles) and 2011 (squares). Isobaths with depth in meters.

that differentiation of OPA's allows improved performance of bio-optical relationships. The OPA concept provides an extension to the PFT concept by accounting for the coexistence of algal and nonalgal material in natural waters. Here, we examine an application of the OPA concept in optically-complex waters of the western Arctic Ocean. We anticipate that this concept can be applied to other regions as well.

Methodology

Study area

Measurements were collected during three field campaigns in coastal and offshore waters in the western Arctic Ocean (Fig. 1). The MALINA (MAckenzie LIght aNd cArbon) project cruise onboard the *CCGS Amundsen* occurred in the southeastern Beaufort Sea from 31 July 2009 through 24 August 2009 and encompassed several transects from the Mackenzie River delta to the southernmost limit of the pack ice off the continental shelf. Two cruises associated with the NASA ICESCAPE (Impacts of Climate on the Eco-Systems and Chemistry of the Arctic Pacific Environment) project onboard the *USCGC Healy* were performed in the Chukchi Sea and western Beaufort Sea from 18 June 2010 through 16 July 2010 and from 28 June 2011 through 24 July 2011 (Arrigo et al. 2012, 2014). Sampling on these cruises included transects from ice-free to ice-covered waters on the Chukchi Sea shelf. For the three cruises, a total of 106 stations were sampled, 27 for MALINA and 79 for ICESCAPE, which included measurements of backscattering and absorption and a detailed characterization of the particulate assemblage.

Measurements of particulate backscattering

At each station, a profiling instrument package with optical sensors was deployed from the sea surface to a few meters above the seafloor, or to a maximum depth of 300 m. The spectral backscattering coefficient, $b_b(\lambda)$, was determined with a combination of multispectral Hydroscat-6 and single wavelength a-Beta sensors (HOBI Labs, Inc.). These instruments provide a measurement of the spectral volume scattering function $\beta(\psi)$ at a fixed scattering angle in the backward direction $\psi = 140°$ (Maffione and Dana 1997). During MALINA, a single Hydroscat-6 and two a-Betas provided measurements in eight spectral bands with nominal center wavelengths of 420, 442, 470, 510, 550, 589, 620, and 671 nm. For both ICESCAPE cruises, two Hydroscat-6 instruments provided measurements in 11 spectral bands with nominal center wavelengths of 394, 420, 442, 470, 510, 532, 550, 589, 649, 730, and 852 nm. The 420 nm band on the MALINA cruise and the 442 nm band on the 2010 ICESCAPE cruise failed to operate correctly, and these data were excluded from subsequent analyses.

The values of $b_{bp}(\lambda)$ were calculated based on a method described originally by Maffione and Dana (1997), and a detailed description of the procedure is provided in Stramski et al. (2008). Briefly, for each spectral band, measurements of β were corrected for the attenuation of light along the optical pathlength using the so-called σ-correction (Maffione and Dana 1997), utilizing the parameter $K_{bb}(\lambda) = a_p(\lambda) + a_g(\lambda) + 0.4b_p(\lambda)$ (HOBI Labs Hydrosoft user manual). For MALINA, the coefficients of $a_p(\lambda)$, $a_g(\lambda)$, and $b_p(\lambda)$ were obtained from measurements with a WETLabs ac-9 instrument. For ICESCAPE, $K_{bb}(\lambda)$ was calculated from relationships between the beam attenuation coefficient of particles and CDOM at 660 nm, $c_{pg}(660)$, measured by a WETLabs C-Star beam transmissometer, and measurements of $a_p(\lambda)$, $a_g(\lambda)$, and $b_p(\lambda)$ on discrete water samples.

The spectral particulate backscattering coefficient was obtained from the measured $\beta(140°, \lambda)$ using the relation

$$b_{bp}(\lambda)=2\pi\chi_p(\beta(140^\circ,\ \lambda)-\beta_w(140^\circ,\ \lambda)) \quad (3)$$

where β_w represents the volume scattering function of pure seawater, which was calculated according to the procedure described in Twardowski et al. (2007) using measured water temperatures as input to the equations of Buiteveld et al. (1994) and a depolarization ratio for water molecules of 0.039 (Jonasz and Fournier 2007). A spectrally constant value of 1.13 was used for the parameter χ_p (Dana and Maffione 2002). Data of $b_{bp}(\lambda)$ were filtered to remove spikes and averaged into 0.5 or 1 m depth bins for analysis and correlation with data from discrete water samples.

Measurements of the spectral particulate backscattering coefficient were fit to the relationship (Gordon and Morel 1983):

$$b_{bp}(\lambda) = b_{bp}(\lambda_o)\left[\frac{\lambda}{\lambda_o}\right]^{\zeta} \quad (4)$$

where λ_o is a reference wavelength and ζ represents the dimensionless spectral slope parameter of $b_{bp}(\lambda)$ which was obtained by least squares linear regression on the $\log_{10}$-transformed data. Measurements from the 442 nm band were excluded when fitting spectral relationships for ICESCAPE to avoid possible effects associated with the anomalous dispersion of refractive index of particles, which can occur in spectral regions exhibiting strong pigment absorption (van de Hulst 1957; Morel and Bricaud 1981), and for consistency with ICESCAPE 2010 measurements in which this channel failed to operate. This band was included in spectral fitting of MALINA data, however, because fewer blue wavelengths were available for this cruise and values of both chlorophyll-*a* concentration and the particulate absorption coefficient at 442 nm were generally lower compared to ICESCAPE.

Measurements of particulate absorption

Water samples were collected at discrete depths from a CTD-Rosette equipped with Niskin bottles immediately before or after in situ optical measurements. Samples were taken at the surface (~1–3 m nominal depth), at additional depths associated with features such as maxima in chlorophyll fluorescence, beam attenuation, or backscattering, and within 3–5 m of the bottom for regions on the shelf. For determinations of light absorption by suspended particulate matter, particles were collected onto a 25 mm diameter Whatman GF/F filter and stored in liquid nitrogen until measurement. Filtration volume was adjusted based on particle concentration and ranged from 30 mL to 11 L.

The spectral absorption coefficient of particles, $a_p(\lambda)$, was determined at 1 nm intervals over the spectral range 300–850 nm for ICESCAPE cruises and 300–800 nm for the MALINA cruise in a dual-beam spectrophotometer (Lambda 18, Perkin Elmer) equipped with a 15 cm integrating sphere (Labsphere). Filters were placed inside the integrating sphere for measurement and scanned in two different orientations to subsample different portions of the filter, whereafter results were averaged (e.g., Röttgers and Gehnke 2012; Stramski et al. 2015). A correction for the pathlength amplification factor was applied to the data (Stramski et al. 2015).

Additional absorption measurements were made on the sample filters following pigment extraction with methanol to partition total $a_p(\lambda)$ into the contributions of phytoplankton, $a_{ph}(\lambda)$, and nonphytoplankton, $a_d(\lambda)$, components (Kishino et al. 1985). In a few cases pigment extraction was incomplete or the partitioning method failed to provide reliable estimates of $a_{ph}(\lambda)$ when the phytoplankton contribution to particle absorption was extremely small (see, for example, Woźniak et al. 2010). All other spectra of $a_p(\lambda)$ and $a_d(\lambda)$ were smoothed using a moving average filter with size ranging between 3 nm and 9 nm depending on the general spectral shape (Stramski et al. 2015). From 305 absorption measurements, 291 spectra of $a_d(\lambda)$ and 269 spectra of $a_{ph}(\lambda)$ were reliably determined. Spectra of absorption by nonalgal matter were further fit to an exponential relationship $a_d(\lambda)=a_d(\lambda_0)\exp(-S_{NAP}(\lambda-\lambda_0))$ using data between 380 nm and 730 nm, excluding the 400–480 nm and 620–710 nm ranges to avoid residual pigment absorption effects.

Bulk measures of particle mass concentration and composition

Measurements of the mass concentrations of particulate matter, organic carbon, and phytoplankton pigments were performed by filtration of collected seawater and were used to characterize the bulk particle assemblage. All filtrations were made with duplicate filters and results were averaged. Determination of the mass concentration of dried suspended particulate matter per unit volume of water, SPM, was accomplished by filtering samples under low vacuum onto precombusted (5 h at 450°C), prerinsed, and preweighed 25 mm diameter glass-fiber filters (Whatman GF/F). Following filtration, filter and filter edges were rinsed multiple times with deionized water to remove residual sea salt. Filters were dried at 60°C and stored sealed until analysis. SPM was determined gravimetrically by weighing sample filters with a Mettler-Toledo MT5 microbalance with 1 μg precision, subtracting the blank filter weight, and dividing by the volume of seawater filtered.

Particulate organic carbon concentration (POC) was obtained using a method consistent with established protocols (e.g., Knap et al. 1996). Water samples were filtered onto precombusted 25 mm GF/F filters; sample filters were then transferred to clean glass scintillation vials, dried at 60°C, and stored until post cruise analysis. Following acidification to remove inorganic carbon, a standard high temperature combustion technique was used to determine POC on each filter. For MALINA, POC was measured from combustion of the same filters used in SPM determination. Several

unused filters from each lot of precombusted filters were used to quantify the background carbon content of filters and subtracted from the sample measurements.

Samples for phytoplankton pigments were filtered onto 25 mm GF/F filters under low light and flash frozen in liquid nitrogen until analysis by High Performance Liquid Chromatography (HPLC). MALINA samples were analyzed using the analytical procedure described in Ras et al. (2008) and ICESCAPE samples were determined using the method of Van Heukelem and Thomas (2001). In this study, we use the HPLC-determined total chlorophyll-*a* as the measure of chlorophyll-*a* concentration, Chl*a*, which represents the summed concentrations of mono- and divinyl chlorophyll-*a*, chlorophyllide-*a*, and the allomeric and epimeric forms of chlorophyll-*a*.

Estimation of phytoplankton community composition and size

Phytoplankton taxa or groups are generally characterized by specific pigments, rendering HPLC-determined phytoplankton pigment concentrations an indicator of phytoplankton composition. However, deriving phytoplankton composition from phytoplankton pigments is not straightforward (e.g., Jeffrey and Vesk 1997; Roy et al. 2011). Whereas certain diagnostic pigments can serve as unambiguous markers for some phytoplankton classes (e.g., peridinin in dinoflagellates), many pigments are shared by more than one algal taxon (e.g., fucoxanthin in diatoms, chromophytes, and nanoflagellates). Despite these limitations, the ratios of the concentrations of specific pigments to Chl*a* provide a useful indication of contributions of certain phytoplankton classes to a mixed phytoplankton population because these ratios can differ between taxonomic groups (Gieskes et al. 1988; Jeffrey and Vesk 1997; Vidussi et al. 2001). In addition to total chlorophyll-*a*, the following pigments were included in our analysis: total chlorophyll-*b* (Chl *b* = the summed contributions of mono- and divinyl chlorophyll-*b*, and the epimeric forms of chlorophyll *b*), the summed contributions of chlorophyll *c*1 and chlorophyll *c*2, chlorophyll *c*3, the summed contributions of α- and β-carotenes, alloxanthin, diadinoxanthin, diatoxanthin, fucoxanthin, 19′-hexanoyloxyfucoxanthin, 19′-butanoyloxyfucoxanthin, neoxanthin, prasinoxanthin, violaxanthin, zeaxanthin, peridinin, pheophorbide-*a*, pheophytin-*a*, lutein, and prasinoxanthin.

In addition to providing information on the composition of the algal assemblage, phytoplankton pigments may also be used to estimate cell size characteristics of the phytoplankton population. The approach is based on seven diagnostic pigments assumed to represent biomarkers of specific taxonomic groups, which are then assigned to different size classes, including microphytoplankton (cell diameter $D > 20$ μm), nanophytoplankton (2 μm $\leq D \leq$ 20 μm), and picophytoplankton (0.2 μm $\leq D <$ 2 μm). We used the method of Uitz et al. (2006), which utilizes the relative weighted concentrations of diagnostic pigments to estimate the fractional contribution f of the three size classes to Chl*a*. Limitations of this method have been noted previously, for example certain diagnostic pigments are shared by multiple phytoplankton groups and some phytoplankton groups may encompass a wide size range (Vidussi et al. 2001; Uitz et al. 2006). In this study, we incorporated the modification of the Uitz et al. (2006) approach proposed by Devred et al. (2011), which corrects for the potential overestimation of the microphytoplankton size fraction which can result from the presence of fucoxanthin in other taxa belonging to the nanophytoplankton size class. It should be noted that this approach has been developed and parameterized using a large pigment dataset from nonpolar oceanic waters, and has not been validated in the Arctic where some phytoplankon taxa (e.g., cyanobacteria) are absent or exhibit low abundance. The estimated fractional contributions of micro-, nano-, and picophytoplankton were then combined to provide a single indicator of the dominant cell size of the algal population, SI (in units of μm), defined as (Bricaud et al. 2004):

$$\mathrm{SI} = 1 \times f_{\mathrm{pico}} + 5 \times f_{\mathrm{nano}} + 50 \times f_{\mathrm{micro}} \tag{5}$$

Lastly, an optical indicator of dominant phytoplankton cell size was derived. The dimensionless size parameter, S_f, was calculated from each normalized phytoplankton absorption spectrum, $a_{\mathrm{ph}}(\lambda)$, according to Ciotti et al. (2002). The S_f size parameter covers a continuum from a value of 0 for a population composed exclusively of microphytoplankton to a value of 1 for a pure picophytoplankton assemblage.

Hierarchical cluster analysis

To obtain optical partitioning of the dataset, a hierarchical cluster analysis (HCA) of the spectral shape of the particulate backscattering-to-absorption ratio, $b_{\mathrm{bp}}(\lambda)/a_{\mathrm{p}}(\lambda)$, was performed. Prior to classification, measurements of $b_{\mathrm{bp}}(\lambda)$ and $a_{\mathrm{p}}(\lambda)$ were quality controlled and measurements of $b_{\mathrm{bp}}(\lambda)$ were spectrally interpolated to match the spectral resolution of $a_{\mathrm{p}}(\lambda)$ measurements using Eq. 4. Spectra of $b_{\mathrm{bp}}(\lambda)/a_{\mathrm{p}}(\lambda)$ were then normalized by the mean value over the spectral range 400−690 nm:

$$\left\langle \frac{b_{\mathrm{bp}}(\lambda)}{a_{\mathrm{p}}(\lambda)} \right\rangle = \frac{1}{291} \int_{400}^{690} \frac{b_{\mathrm{bp}}(\lambda)}{a_{\mathrm{p}}(\lambda)} d\lambda \tag{6}$$

The spectral domain was limited to 690 nm because particulate absorption becomes very small for longer wavelengths, resulting in high noise in $b_{\mathrm{bp}}(\lambda)/a_{\mathrm{p}}(\lambda)$. We note that with this procedure, the relative magnitude of $b_{\mathrm{bp}}(\lambda)$ and $a_{\mathrm{p}}(\lambda)$ to each other as well as the spectral shape of each component influences the normalized spectrum. For example, the spectral shape of $b_{\mathrm{bp}}(\lambda)$ exerts a stronger influence on the $b_{\mathrm{bp}}(\lambda)/a_{\mathrm{p}}(\lambda)$ ratio in the case of mineral-dominated particle assemblages compared to more organic particle assemblages because of generally higher $b_{\mathrm{bp}}(\lambda)$ relative to $a_{\mathrm{p}}(\lambda)$. Because spectra of b_{bp} are derived from multispectral measurements fitted to the power-law relationship in Eq. 4, the b_{bp} spectra are featureless.

The HCA method classifies objects, i.e., normalized spectra of $b_{bp}(\lambda)/a_p(\lambda)$, into groups or clusters of similar objects. We used Ward's (1963) minimum variance method because it minimizes the within-cluster variance and is less sensitive to outliers than other methods such as single linkage or complete linkage (Kamvar et al. 2002). Ward's method uses an agglomerative (bottom–up) hierarchical clustering algorithm, proceeding sequentially by binary group fusions, at each step creating bigger clusters until all the objects in the original data set are linked together in a hierarchical cluster tree or dendrogram. At each clustering step going from K to $K-1$ clusters, Ward's method merges the pair of objects or clusters whose fusion results in the smallest increase of the "total within-cluster sum of squares" over all K clusters formed so far. This sum is defined as the sum of the squares of the Euclidean distances between all objects in the cluster and the centroid of the cluster:

$$\sum_{k=1}^{K}\sum_{i=1}^{n_k}\|x_{ik}-m_k\|^2 \tag{7}$$

where n_k is the number of objects in cluster k and $\|x_{ik}-m_k\|^2$ is the squared Euclidean distance between an object x_{ik} in cluster k and the centroid of the cluster to which it belongs, $m_k=\frac{1}{n_k}\sum_{i=1}^{n_k}x_{ik}$. Merging the cluster pair that minimizes the total within-cluster sum of squares is equivalent to merging the pair of clusters c_k and c_l that minimizes the following distance measure (Legendre and Legendre 2012):

$$D(c_k,c_l)=\sqrt{\frac{2n_kn_l}{n_k+n_l}}\|m_k-m_l\| \tag{8}$$

where $\|m_k-m_l\|$ is the Euclidean distance between the centroids of clusters k and l, and n_k and n_l are the number of objects in clusters k and l, respectively. This distance measure is commonly referred to as the linkage distance for Ward's method. The multiplicative factor of 2 in Eq. 8 is used in the MATLAB implementation of Ward's algorithm so that the distance between two singleton clusters is the same as the Euclidean distance between them.

Each level in the dendrogram represents a particular grouping of the data into a certain number of distinct clusters. It is up to the user to decide which level represents a natural clustering or equivalently, which number of clusters is optimal. To aid this decision, a plot of linkage distance as a function of number of clusters was examined. For agglomerative HCA, the linkage distance between merged clusters increases monotonically with each step in the hierarchical procedure as increasingly dissimilar clusters are being merged. Steeper increases in the linkage distance are associated with a stronger increase of the total within-cluster variance and finding the appropriate number of clusters can be thought of as locating the point of maximum curvature in the relationship between linkage distance and number of clusters. The L-method of Salvador and Chan (2004) was used to identify the point of maximum curvature as the boundary between the pair of straight lines that most closely fit the curve of linkage distance vs. number of clusters. This method is computationally efficient and robust to local trends and outliers. All cluster pairs with linkage distance below the cutoff distance were then merged.

To evaluate the assignment of an object to a cluster, we computed the object's silhouette value S_i which is a measure of how close object i is to objects in its own cluster k compared to objects in other clusters l (Rousseeuw 1987):

$$S_i=\frac{\min(d_{il})-d_i}{\max(d_i,\min(d_{il}))} \tag{9}$$

where d_i is the average distance from object i to the other objects in the same cluster, and d_{il} is the average distance from object i to objects in cluster l, which is then minimized over all clusters $l\neq k$. Values of S_i range from -1 to 1, with negative values for objects that are more similar to objects in other clusters than to objects in their own cluster. Conversely, objects with positive S_i are more similar to objects in their own cluster than to objects in other clusters. The range of S_i thus comprises a continuum representing misclassified to well classified objects, with zero values for intermediate cases. The distance metric in the calculation of S_i used here was the squared Euclidean distance, consistent with the between-object distance metric utilized by Ward's clustering algorithm.

Different clustering algorithms may produce different dendrograms because clustering methods impose different models onto the data. We tested alternative distance and clustering algorithms and compared results using various criteria: the cophenetic correlation coefficient (Sokal and Rohlf 1962), a measure of how precisely a dendrogram preserves the pairwise distance between data objects, statistics of silhouette values, and best professional judgment.

Differences in particle characteristics among optical clusters

Differences in characteristics of the particulate assemblage among the optical clusters were examined using a one-way Analysis of Variance (ANOVA). To test which pairs of cluster means are significantly different from each other, multiple comparison tests were done at a significance level of 0.05. Particle characteristics that were tested included bulk measurements of particle concentration, composition, and phytoplankton composition and size indicators. If necessary, values were log-transformed prior to analysis to better approximate a normal distribution.

Statistical evaluation of bio-optical relationships among optical clusters

We examined if a priori knowledge of optical cluster membership improved commonly used bio-optical

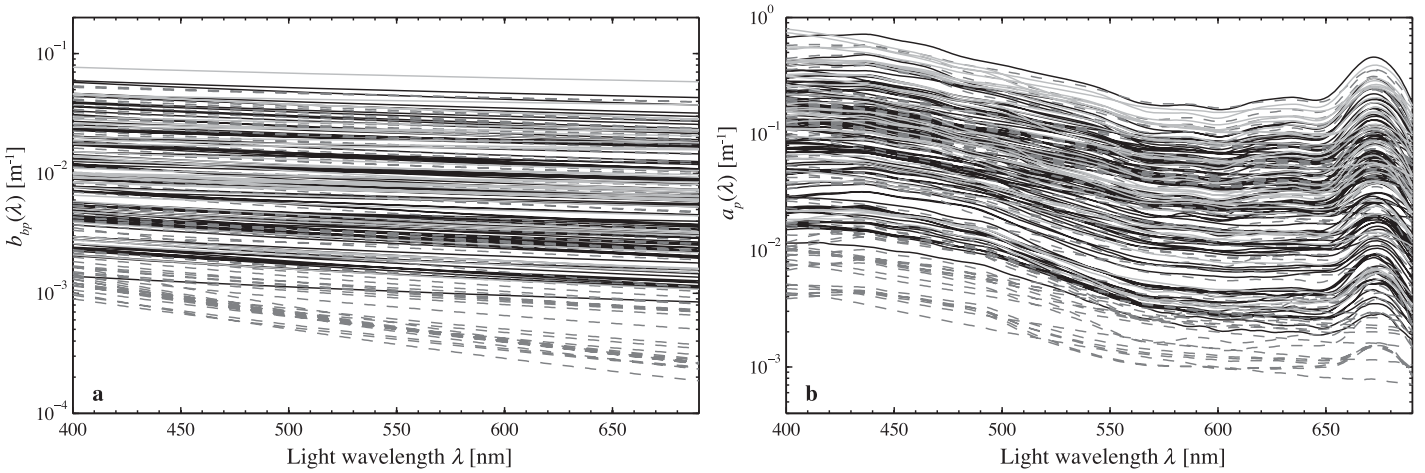

Fig. 2. Spectra of **(a)** the backscattering coefficient of particles, $b_{bp}(\lambda)$, and **(b)** the absorption coefficient of particles, $a_p(\lambda)$, for MALINA (dashed lines) and ICESCAPE (solid lines, 2010 in grey, 2011 in black) samples. Measurements of $b_{bp}(\lambda)$ were made at 11 wavelengths, and subsequently fitted to the relationship in Eq. 4.

relationships, such as the relationship between SPM and the particulate backscattering coefficient or between Chl*a* and the phytoplankton absorption coefficient. Ordinary least squares regression analysis on $\log_{10}$-transformed data of selected optical properties and measurements of particle concentration were performed for the entire dataset and for each optical cluster separately. Prior to regression analysis, data outliers were removed according to methods described in the Web Appendix of Neukermans et al. (2012).

Goodness-of-fit statistics of regression models are reported using the root-mean-square-error (RMSE),

$$\mathrm{RMSE}=\sqrt{\frac{1}{n-2}\sum_{k=1}^{n}(\log y_k-\log \hat{y}_k)^2} \qquad (10)$$

where y_k is the observed value, and $\hat{y}_k$ is the model-derived value. We also report the median prediction error of the regression model, MPE, calculated as the median of the absolute percent difference between model-derived and observed values, $100\times|\hat{y}_k-y_k/y_k|$. For comparison between OPA-specific models and the general model we report the median ratio, MR, of model-derived to observed values $\hat{y}_k/y_k$, and the median bias, MB, calculated as the median difference between model-derived and observed values $\hat{y}_k-y_k$ for each OPA-specific model and the general model.

Results

For the three cruises, 216 water samples which comprised concurrent measurements of particulate backscattering, absorption partitioned into algal and nonalgal contributions, phytoplankton pigments, POC, and SPM were used in our analysis. The wide range in spectral shapes and magnitudes of $b_{bp}(\lambda)$ and $a_p(\lambda)$ spanned by this dataset is illustrated in Fig. 2. We first describe optical partitioning of this dataset and then examine differences in characteristics of suspended particles among optical clusters.

Optical classification of samples

Figure 3a depicts the dendrogram obtained from a HCA of normalized $b_{bp}(\lambda)/a_p(\lambda)$ spectra. The cophenetic correlation coefficient for this dendrogram is 0.56, indicating good preservation of the Euclidean distances between pairs of objects. The optimal number of clusters was derived from the "knee" in a diagram of between-cluster distances along the dendrogram shown in Fig. 3b. The L-method of Salvador and Chan (2004) finds the knee in this curve by an iterative procedure that fits a straight line through the first m points and another line through the remaining points. The knee is found where the root-mean-square-error of the fit to this curve by the pair of lines is minimal. The L-method suggests that the knee is located at $k=16$ clusters (Fig. 3b). However, the L-method may position the knee to the right of the actual knee when there are too many points on the right side of the actual knee (Salvador and Chan 2004). This is the case for our dataset. The L-method was therefore reapplied to a restricted range of number of clusters so that the initial knee estimate is in the middle of this range (Fig. 3c). We obtained a refined estimate of the knee at $k=6$ clusters. We further refined this estimate by close examination of the curve in Fig. 3c which exhibits a first steep increase in the linkage distance along the dendrogram at $k=7$ clusters. This suggests an optimal partitioning of the dataset into seven clusters. Such a partitioning corresponds to a cutoff distance of 8, represented by the dashed line in Fig. 3a. Clusters are labeled TMin, TMix, COD, MOD, COP, TOP1, and TOP2. The first character of the label refers to overall water turbidity with "T" for turbid ($c_p(660)\geq 0.3$ m^{-1}), "M" for moderately turbid (0.1 m$^{-1}<c_p(660)<0.3$ m^{-1}), and "C" for clear waters ($c_p(660)\leq 0.1$ m^{-1}). Subsequent characters refer to

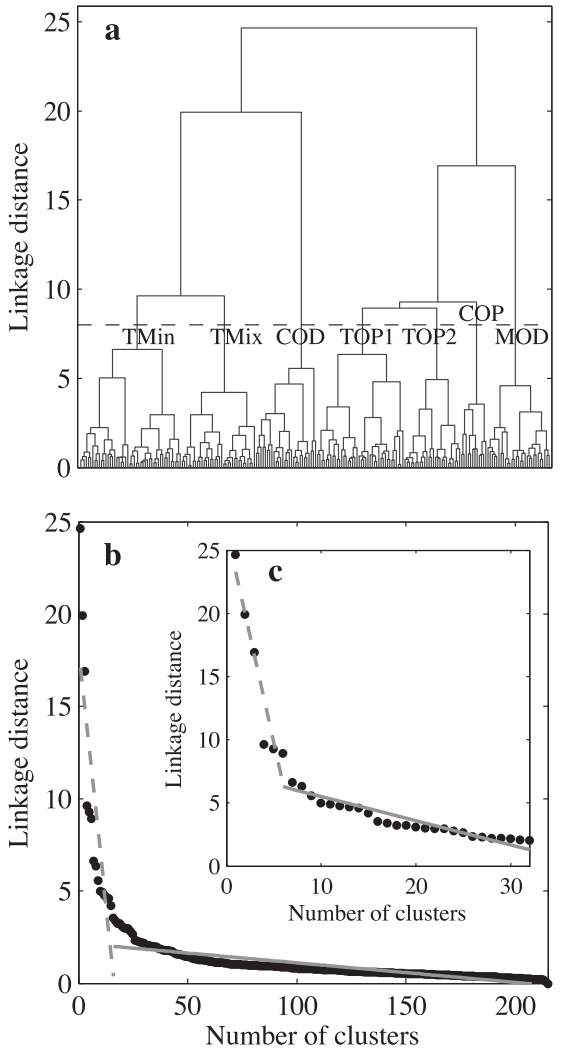

Fig. 3. (a) Dendrogram obtained from HCA of 216 normalized spectra of $b_{bp}(\lambda)/a_p(\lambda)$. **(b)** Linkage distance as function of number of clusters obtained from the dendrogram in (a). The lines intersect at the first estimate of the knee of the curve as obtained from the L-method of Salvador and Chan (2004). **(c)** Similar to (b) but with a restricted range for the number of clusters and a refined estimate of the location of the knee. The dashed line corresponds to a linkage distance of 8, which divides the dataset into seven clusters.

bulk composition of the particle assemblage and indicate the dominance of minerals "Min," organic material "O," phytoplankton "P," or nonalgal organic material (also referred to as detritus) "D," while "Mix" refers to particle assemblages of mixed composition. These characteristics are discussed in detail in the next section.

The resulting partitioning of 216 normalized spectra of $b_{bp}(\lambda)/a_p(\lambda)$ in seven clusters obtained from HCA is illustrated in Fig. 4. Some spectra exhibited negative silhouette values indicative of poor classification (grey dashed lines in Fig. 4a–g). These 28 spectra were removed from subsequent analyses. Silhouette values for the majority of the remaining spectra varied from 0.10 to 0.77, with a median value of 0.49. Strong differences in spectral shape of $b_{bp}(\lambda)/a_p(\lambda)$ were observed among clusters, especially in blue and red spectral regions (Fig. 4h). For example, the coefficient of variation in $b_{bp}(\lambda)/a_p(\lambda)$ of the cluster means was 27% at 400 nm, 5% at 530 nm, and 30% at 675 nm. Variability in the spectral shape of $b_{bp}(\lambda)/a_p(\lambda)$ was mostly driven by absorption with comparatively less contribution to variability from backscattering (Figs. 2, 5a,b).

Spectral features associated with the presence of phytoplankton pigment absorption bands in blue and the red wavelengths were easily discernible in spectra of $b_{bp}(\lambda)/a_p(\lambda)$ and $a_p(\lambda)$ in all clusters. These features were least pronounced in cluster TMin, indicating a much weaker contribution of algal particles to $a_p(\lambda)$ in this cluster (Figs. 4a, 5a). Indeed, the ratio of nonalgal to total particulate absorption at a wavelength of 400 nm, $a_d(400)/a_p(400)$, in cluster TMin averaged 0.86 ($\pm$0.07) (mean $\pm$ standard deviation), significantly higher than all other cluster means (Table 1). Particulate absorption spectra for samples in cluster TMin were nearly featureless with $a_p(\lambda)$ decreasing with increasing wavelength and with significant absorption in the near-infrared spectral region (not shown). Such absorption features are typically associated with mineral-dominated particle assemblages (e.g., Babin and Stramski 2004; Stramski et al. 2007). The strong optical distinction of cluster TMin was corroborated by its members' silhouette values, which were on average highest for this cluster (Table 1).

Phytoplankton pigment absorption features were most strongly pronounced in $a_p(\lambda)$ spectra of clusters COP, TOP1, and TOP2 (Fig. 5a). Correspondingly, cluster mean values of $a_d(400)/a_p(400)$ were below 0.5, suggesting a dominant contribution of phytoplankton to the particulate assemblage in these clusters (Table 1). The spectral shape of $a_{ph}(\lambda)$ differed strongly among these clusters reflecting changes in phytoplankton communities, pigment composition, pigment packaging, or some combination of these processes. The pigment package effect describes the reduction in absorption per unit pigment concentration when pigments are contained in cells compared with the absorption potential for the same amount of pigment in solution (Duysens 1956; Morel and Bricaud 1981; Kirk 1996). An increase in pigment packaging occurs either as cell size increases or as absorption of intracellular material increases (Morel and Bricaud 1981). This results in a flattening of the $a_{ph}(\lambda)$ spectrum and a reduction in the blue-to-red absorption ratio for wavelengths corresponding to chlorophyll-*a* absorption peaks. To first order, the influence of pigment packaging can be assessed in the red absorption band of chlorophyll-*a*, a spectral region in which the contribution to absorption by accessory pigments is generally small. The significantly lower mean value of $a_{ph}(676)$/Chl*a* of the TOP clusters compared to cluster COP suggests a stronger degree of pigment packaging in turbid phytoplankton-dominated waters (Table 1). Pigment

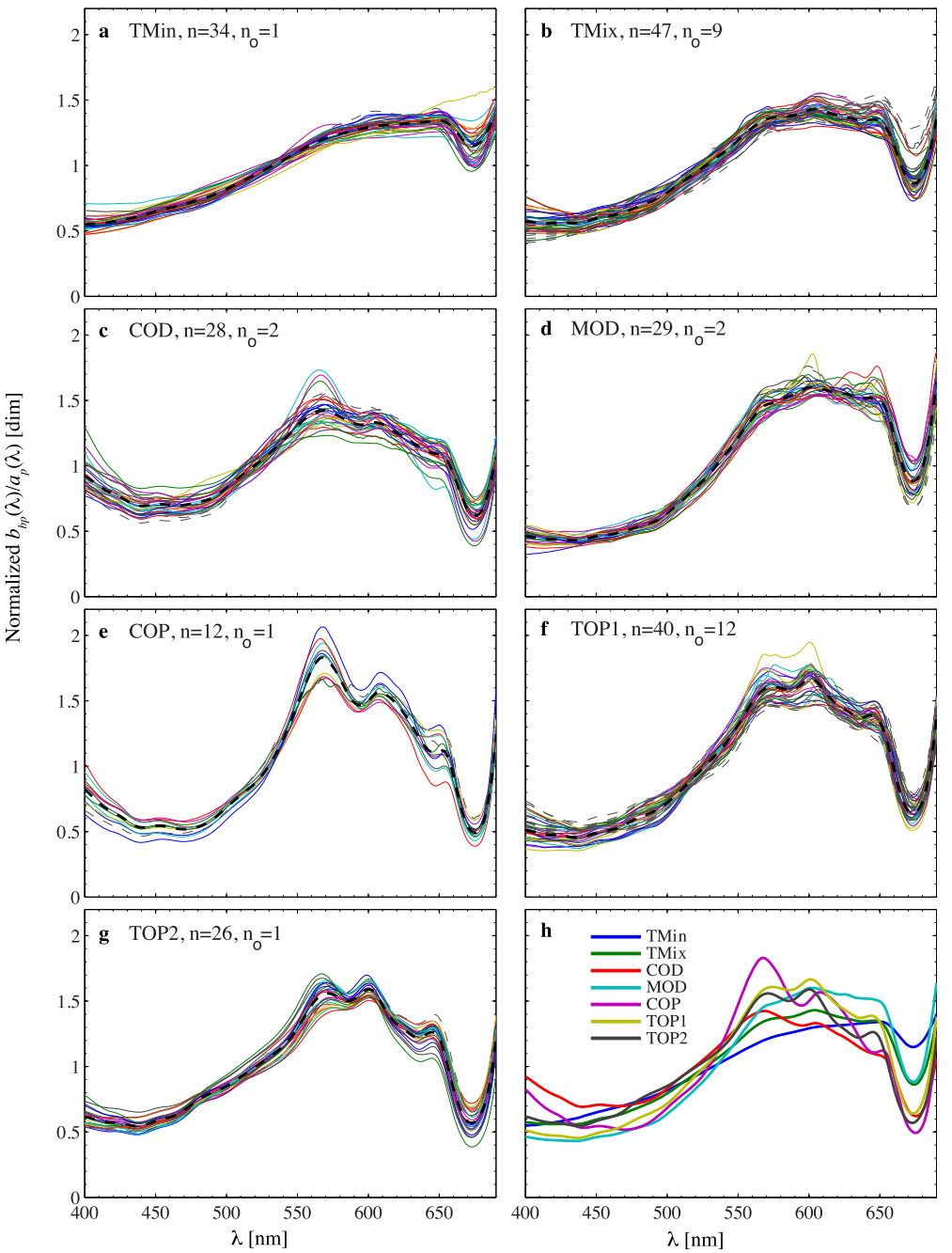

Fig. 4. (a–g) Normalized spectra of $b_{bp}(\lambda)/a_p(\lambda)$ associated with each of the seven clusters obtained from HCA. *n* is the number of objects in each cluster of which n_0 have negative silhouette value, indicative of poor classification. Spectra with negative silhouette values are shown as grey dashed lines and are not included in the calculation of the mean spectrum for each cluster, which is represented by the black dashed line in each panel. **(h)** Mean spectrum of normalized $b_{bp}(\lambda)/a_p(\lambda)$ for each cluster.

packaging also resulted in a strongly reduced $a_{ph}(443)/a_{ph}(676)$ ratio in cluster TOP2, whereas this effect appears to be largely offset by increased relative contribution of ancillary pigments to absorption in cluster TOP1.

Cluster COP exhibited high blue-to-red phytoplankton absorption ratio, $a_{ph}(443)/a_{ph}(676)$, a small degree of pigment packaging effect, and very steep backscattering spectra suggesting the dominance of small-sized particles and phytoplankton (Table 1). Samples assigned to cluster TOP2, on the other hand, exhibit much weaker spectral dependence of backscattering and a cluster mean $a_{ph}(443)/a_{ph}(676)$ value of 1.33 (±0.11), significantly lower than all other clusters (Table 1). These observations suggest that cluster TOP2 comprises particle assemblages dominated by large-sized

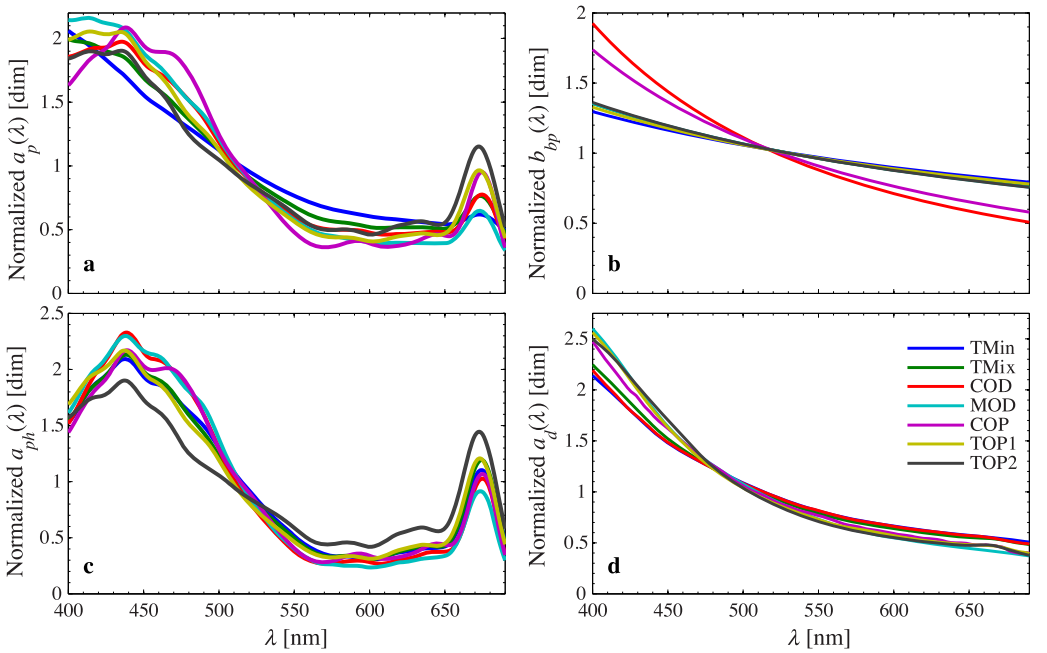

Fig. 5. Cluster mean spectra of the normalized coefficients of **(a)** particle absorption, $a_p(\lambda)$, **(b)** particle backscattering $b_{bp}(\lambda)$, **(c)** phytoplankton absorption, $a_{ph}(\lambda)$, and **(d)** nonalgal particle absorption, $a_d(\lambda)$. The legend in panel d applies to all four panels.

Table 1. Optical and optically-derived properties associated with each cluster. n is the number of objects in each cluster of which n_0 had silhouette values, S_i, that were negative and subsequently discarded from analysis. For each cluster, we tabulate the 5th, 50th, and 95th percentiles of S_i and the mean (± standard deviation) of other quantities. The spectral slope of particle backscattering, ζ, is calculated from Eq. 4. The particulate backscattering ratio, $\tilde{b}_{bp} = b_{bp}/b_p$, is given at a wavelength of 660 nm. The ratio of the nonalgal absorption coefficient, a_d, to the total particle absorption coefficient, a_p, is provided at a wavelength of 400 nm. S_{NAP} represents the spectral slope of the nonalgal absorption coefficient. Phytoplankton optical properties include the blue-to-red absorption ratio of the phytoplankton absorption coefficient, a_{ph}, and the phytoplankton chlorophyll-*a* specific absorption coefficient at 676 nm. The phytoplankton size parameter, S_f, represents the relative contribution of picophytoplankton to absorption (Ciotti et al. 2002).

	TMin	TMix	COD	MOD	COP	TOP1	TOP2
Observations	$n = 34$, $n_0 = 1$	$n = 47$, $n_0 = 9$	$n = 28$, $n_0 = 2$	$n = 29$, $n_0 = 2$	$n = 12$, $n_0 = 1$	$n = 40$, $n_0 = 12$	$n = 26$, $n_0 = 1$
P5-50-95(S_i)	0.15, 0.68, 0.80	0.07, 0.41, 0.63	0.13, 0.49, 0.70	0.34, 0.64, 0.79	0.09, 0.48, 0.72	0.01, 0.33, 0.56	0.19, 0.53, 0.75
ζ(dim)	−0.90(±0.29)	−1.06(±0.38)	−2.45(±0.35)	−1.07(±0.18)	−2.02(±0.46)	−0.97(±0.49)	−1.08(±0.35)
$\tilde{b}_{bp}$ (dim)	0.022(±0.007)	0.015(±0.006)	0.011(±0.005)	0.012(±0.003)	0.009(±0.005)	0.011(±0.005)	0.010(±0.008)
$a_d(400)/a_p$ (400) (dim)	0.86(±0.07)	0.72(±0.06)	0.60(±0.15)	0.65(±0.09)	0.29(±0.09)	0.45(±0.11)	0.41(±0.09)
S_{NAP}(nm^{-1})	0.0060(±0.0004)	0.0066(±0.0007)	0.0061(±0.0009)	0.0082(±0.0009)	0.0075(±0.0016)	0.0081(±0.0011)	0.0080(±0.0008)
$a_{ph}(443)/a_{ph}$ (676) (dim)	1.92(±0.37)	1.86(±0.52)	2.32(±0.50)	2.56(±0.36)	2.02(±0.14)	1.87(±0.50)	1.33(±0.11)
$a_{ph}(676)$/Chl *a* (m^2 mg^{-1})	0.028(±0.013)	0.023(±0.011)	0.029(±0.015)	0.025(±0.009)	0.038(±0.017)	0.019(±0.009)	0.017(±0.009)
S_f (dim)	0.31(±0.20)	0.33(±0.25)	0.51(±0.21)	0.53(±0.10)	0.40(±0.06)	0.33(±0.24)	0.06(±0.07)

phytoplankton. Cluster TOP1 is also characterized by weak spectral dependence of backscattering but with higher average $a_{ph}(443)/a_{ph}(676)$ ratio than observed for cluster TOP2. This suggests the dominance of large-sized phytoplankton with increased contribution of accessory pigments, which offsets the effects of pigment packaging at blue wavelengths. We note, however, that in contrast to clusters COP and TOP2 the spectral shape of phytoplankton absorption varies widely within cluster TOP1, with $a_{ph}(443)/a_{ph}(676)$ ranging from 1.2 to 3.0 and S_f ranging from 0 to 0.8 (Table 1).

Finally, the median silhouette value of samples assigned to cluster TOP1 is lowest indicating strong optical similarity with other clusters, particularly with cluster TOP2.

Clusters TMix, COD, and MOD encompass samples with intermediate contribution of nonalgal matter to absorption (Table 1). Cluster COD is characterized by the steepest backscattering spectra and high $a_{ph}(443)/a_{ph}(676)$, suggesting the presence of mainly small-sized algal and nonalgal particles (Table 1). In contrast, cluster TMix exhibits much weaker spectral dependence of backscattering and much lower $a_{ph}(443)/a_{ph}(676)$, as could be expected from a mixture of large-sized algal and nonalgal particles. Cluster MOD represents samples which combine a rather flat backscattering spectrum with a phytoplankton absorption spectrum characterized by a cluster mean value of $a_{ph}(443)/a_{ph}(676) = 2.6$, significantly higher than all other cluster means.

Variations in spectral dependency of absorption by nonalgal particles are small, but nonetheless two groups with significantly different spectral dependencies can be distinguished (Table 1). Weak spectral dependency is observed for clusters TMin, TMix, and COD ($S_{NAP} = 0.0062 \pm 0.0007$ nm^{-1}), while the remaining clusters exhibit stronger exponential decreases with increasing wavelength ($S_{NAP} = 0.0080 \pm 0.0010$ nm^{-1}). It has been suggested that S_{NAP} may correlate with organic content of the particulate assemblage (Babin et al. 2003a), and this hypothesis is supported by limited field observations in the Louisiana coastal environment (Estapa et al. 2012), and by our observations. We note that our observations of S_{NAP} are at the lower end of observations in European coastal waters (Bowers et al. 1996; Babin et al. 2003a).

Characterization of particle assemblages associated with each optical cluster

In this section, we investigate to what extent the optically-classified samples with distinct spectral shape of $b_{bp}(\lambda)/a_p(\lambda)$ represent distinct assemblages of suspended particles. We investigated differences among optical clusters for a comprehensive set of parameters characterizing concentration and composition of the particulate matter (Fig. 6), and the pigment composition and size of the phytoplankton population (Fig. 7).

Figure 6 illustrates variability in particle concentration and composition among optical clusters and provides an overview of the diversity of particle assemblages covered by our dataset. Particle concentration, expressed in terms of SPM, spans over two orders of magnitude from less than 0.04 g m^{-3} typical of oligotrophic waters to 7.06 g m^{-3} for turbid nearshore waters (Fig. 6a). Bulk particle composition is characterized by combinations of three measured parameters. The ratios of POC and Chl*a* to SPM provide indicators for the respective contributions of organic and algal matter to the total particulate assemblage. The ratio POC/SPM can be used to classify samples into composition-related groups; following Woźniak et al.

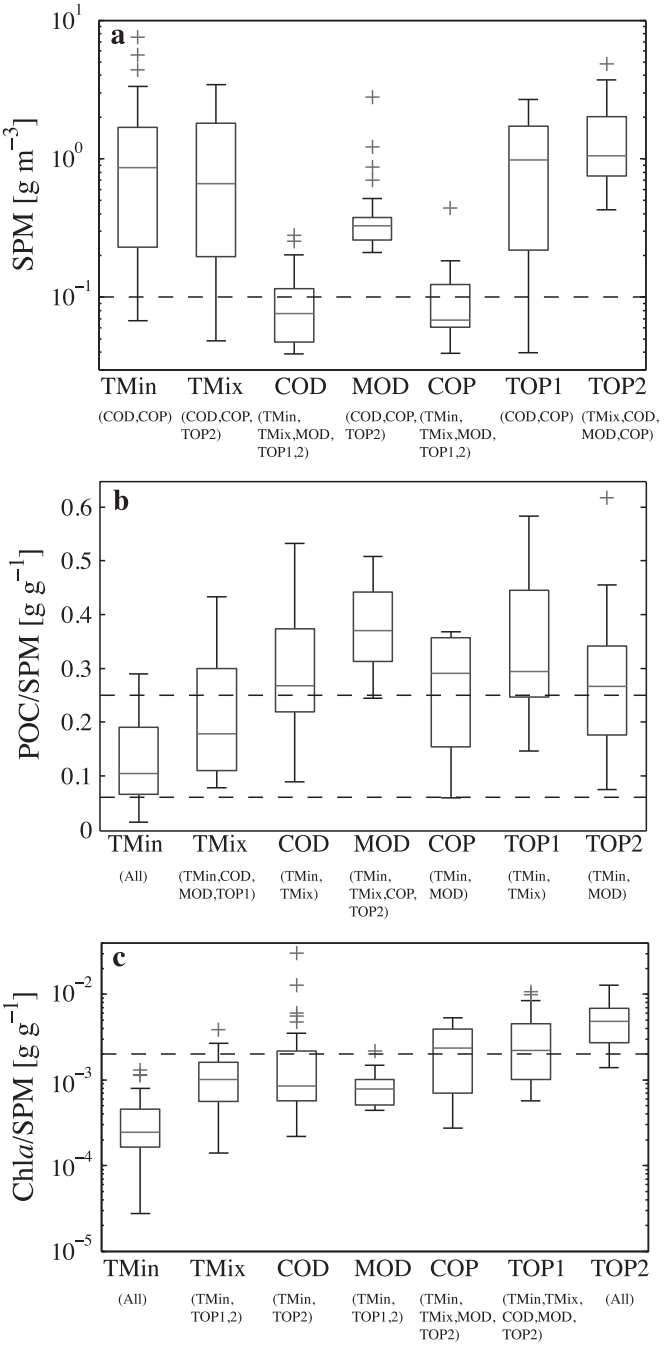

Fig. 6. Boxplots of **(a)** SPM, **(b)** POC/SPM, and **(c)** Chl*a*/SPM for the seven clusters obtained from HCA of $b_{bp}(\lambda)/a_p(\lambda)$. Names of clusters below cluster labels indicate which other cluster means are significantly different from the labeled cluster mean as obtained from pairwise *t*-tests. Boxes extend from the 25th to the 75th percentile with the line indicating the median. Observations further than 1.5 times the length of the box away from the box are considered outliers and indicated by crosses. Whiskers cover the range of values without outliers. Dashed lines indicate threshold values for (a) clear water, (b) compositional groups as defined in Woźniak et al. (2010), (c) phytoplankton-dominated assemblages.

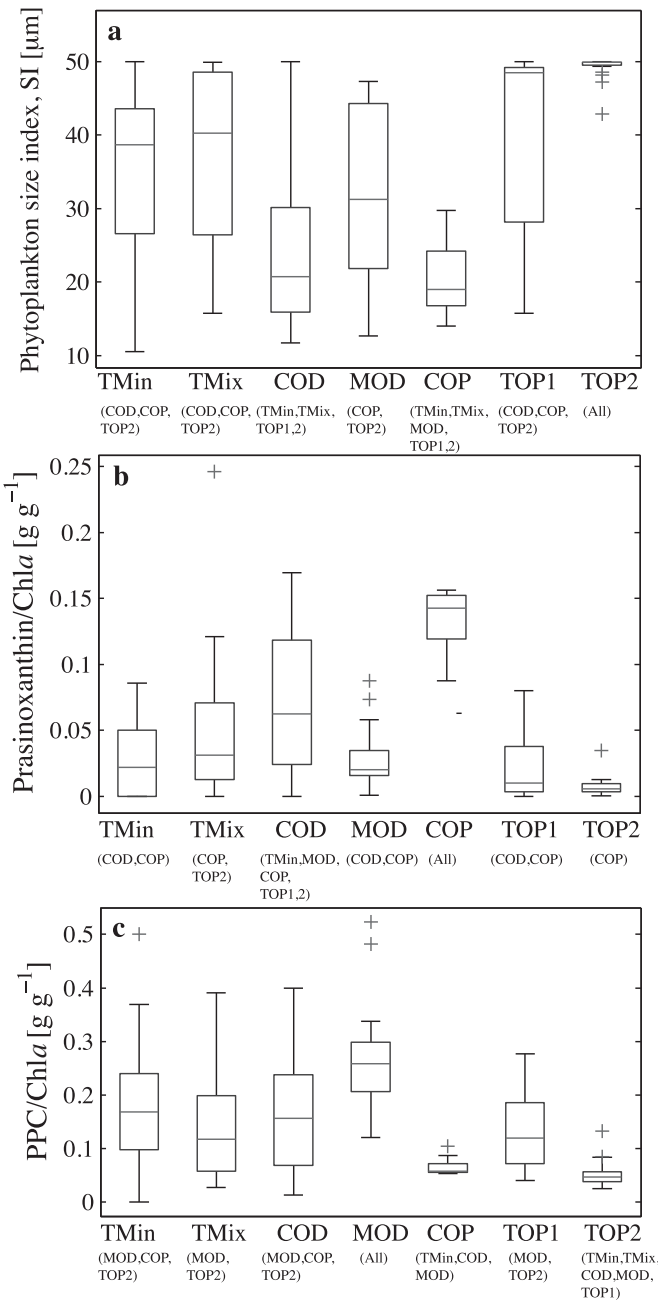

Fig. 7. Boxplots of **(a)** the phytoplankton size index of Bricaud et al. (2004) calculated from Eq. 5 and the concentration ratios of **(b)** prasinoxanthin and **(c)** photoprotective carotenoids (PPC) to chlorophyll-*a* for the seven clusters obtained from HCA of $b_{bp}(\lambda)/a_p(\lambda)$. PPC is the sum of alloxanthin, diadinoxanthin, diatoxanthin, zeaxanthin, and alpha- and beta-carotene concentrations. Names of clusters below cluster labels indicate which other cluster means are significantly different as obtained from pairwise *t*-tests.

(2010), we defined assemblages as "organic-dominated" (POC/SPM > 0.25), "mineral-dominated" (POC/SPM < 0.06), and "mixed composition" for intermediate values (0.06−0.25). In our dataset, POC/SPM varies from 0.01 to 0.62, thus covering all compositional groups (Fig. 6b). The ratio Chl*a*/SPM spans three orders of magnitude, from 2.7×10^{-5} for a particle assemblage dominated by nonalgal matter to 3.0×10^{-2} for a phytoplankton dominated assemblage.

Significant differences in particle concentration and composition are observed among clusters. Cluster TMin is distinct from all other clusters in terms of particle composition. The cluster mean values of Chl*a*/SPM = 0.37 (±0.34) $\times 10^{-4}$ and POC/SPM = 0.13 (±0.08) are lowest and this cluster includes all mineral-dominated samples in the dataset. These observations suggest that the relative contribution of mineral material for samples in this cluster is strongest, and we refer to this cluster as the mineral-dominated particle assemblage.

At the other end of the compositional spectrum is cluster TOP2, which is characterized by the highest mean Chl*a*/SPM = 5.42 (±3.19) $\times 10^{-3}$, indicating the dominance of phytoplankton in samples assigned to this cluster. Cluster TOP2 is further characterized by high particle concentration (Fig. 6a) and a majority of the samples classified as organic-dominated (Fig. 6b). Cluster TOP1 is similar to cluster TOP2 in terms of bulk particle concentration and composition, but with slightly but significantly lower mean Chl*a*/SPM = 3.20 (±2.81) $\times 10^{-3}$. Both clusters represent turbid waters dominated by organic material with large contributions of phytoplankton, hence the label "TOP." The main difference between both TOP clusters lies in the characteristics of the phytoplankton community.

The phytoplankton community in samples assigned to cluster TOP2 is almost exclusively composed of microphytoplankton, which accounted for at least 96% of Chl*a* for 90% of the samples. The dominance of microphytoplankton in this cluster derives entirely from fucoxanthin, a pigment indicative of the presence of diatoms, after an adjustment correcting for its co-occurrence in some nanophytoplankton species (Devred et al. 2011). Cluster TOP1 is also dominated by microphytoplankton derived mainly from fucoxanthin, resulting in a median phytoplankton cell size index similar to the one of cluster TOP2, but with stronger variability (Fig. 7a) and with a generally higher contribution to absorption and to Chl*a* by picophytoplankton (Table 1). Cluster TOP1 has significantly higher proportion of pigments devoted to photoprotection. In this cluster, we observed a factor of three increase in the ratio of photoprotective pigments to total chlorophyll-*a* which mainly originates from the pigment diadinoxanthin, indicative of diatoms acclimated to higher light levels (Fig. 7c).

A third cluster with similarly high contributions of algal material to the total particulate assemblage was identified (Fig. 6c). In contrast to the TOP clusters, this cluster is representative of clear waters with particle concentrations generally below 0.1 g m^{-3} (Fig. 6a), and is thus labeled "COP" to denote clear waters, dominated by organic material, mainly phytoplankton. In accordance with the high mean value of Chl*a*/SPM = 2.47 (±1.81) $\times 10^{-3}$ observed for this cluster,

the average contribution of nonalgal matter to particulate absorption was the lowest of all clusters, $a_d(400)/a_p(400) = 0.29$ (± 0.09). In contrast to the TOP clusters, the phytoplankton assemblage in cluster COP exhibited larger pigment diversity and was dominated by smaller-sized cells (Fig. 7a) with average fractional contributions from nano- and picophytoplankton to total chlorophyll-*a* of 0.42 (± 0.13) and 0.21 (± 0.10), respectively. The contribution by nanophytoplankton was predominantly associated with the pigment 19′-hexanoyloxyfucoxanthin, indicative of prymnesiophytes, while picoplankton contributions to total chlorophyll-*a* were derived almost entirely from total chlorophyll-*b*, indicative of the presence of chlorophytes and prasinophytes. Chlorophytes also contain the pigment zeaxanthin, but the ratio of zeaxanthin to total chlorophyll-*a* was small for this cluster, in contrast to the ratio of prasinoxanthin, a pigment unique to prasinophytes, to total chlorophyll-*a* (Fig. 7b). This suggests that prasinophytes had a dominant contribution to the picoplankton pool in the COP cluster. The predominance of small-sized phytoplankters in this cluster is consistent with our observations of steeper particulate backscattering spectra and weak influence of pigment packaging on phytoplankton absorption (Table 1).

Clusters COD and MOD exhibit strong similarities in particle composition: they are characterized by elevated contributions of POC to SPM, but contributions of Chl*a* to SPM are lower than in clusters TOP1, TOP2, and COP (Fig. 6b,c). This suggests increased importance of organic nonalgal particles in clusters COD and MOD, which is corroborated by observations of increased contribution of nonalgal matter to total particulate absorption (Table 1). However, clusters COD and MOD strongly differ in particle concentration. Whereas cluster COD is representative of clear waters with particle concentrations generally below 0.1 g m^{-3}, cluster MOD represents moderately turbid waters with an average value of SPM that is four times higher (Fig. 6a).

The fractional contributions of the micro-, nano-, and picophytoplankton size classes to Chl*a* in cluster COD are similar to those found for cluster COP. COD and COP clusters are dominated by smaller-sized phytoplankton resulting in similarly small mean SI values of 25.7 (± 11.9) μm and 20.8 (± 5.2) μm, respectively (Fig. 7a). However, differences in phytoplankton taxonomic composition were found between these clear water clusters. Whereas the pigment alloxanthin, indicative of the presence of nanoplanktonic cryptophytes, was absent in samples belonging to cluster COP, its presence was detected in cluster COD, albeit in low concentrations. The remaining pigments associated with nanoplanktonic taxonomic groups (19′-butanoyloxyfucoxanthin, 19′-hexanoyloxyfucoxanthin, and fucoxanthin attributed to nanoflagellates) all had significantly lower relative concentrations in cluster COD. This suggests an increased contribution of cryptophytes and a decreased contribution of chromophytes and nanoflagellates to the nanoplankton pool in cluster COD. Also, the picophytoplankton pool in cluster COD was less influenced by prasinophytes (Fig. 7b).

The phytoplankton population of samples in cluster MOD is characterized by intermediate contributions of micro- ($f_{micro} = 0.64 \pm 0.24$) and picophytoplankton ($f_{pico} = 0.22 \pm 0.17$) to Chl*a*, resulting in an intermediate phytoplankton size index of 32.8 (± 11.5) μm (Fig. 7a). The phytoplankton population in cluster MOD significantly differs from those in other clusters in terms of relative amount of photoprotective pigments, which is highest in this cluster (Fig. 7c). Major contributions to photoprotection were derived from diadinoxanthin (44%), α and β carotenoids (19%), and zeaxanthin (14%). This cluster also has the strongest contribution of dinoflagellates to the microplankton pool of 8% on average, derived from relative concentrations of the pigment peridinin. Furthermore, the ratio of prasinoxanthin to total chlorophyll-*a* was low (Fig. 7b) and the ratio of zeaxanthin, a pigment found in chlorophytes but not in prasinophytes, to total chlorophyll-*a* was highest of all clusters. This suggests that chlorophytes contributed more strongly to the picoplankton pool than prasinophytes in this cluster. Despite intermediate phytoplankton cell size observed for this cluster, the S_f size parameter and the $a_{ph}(443)/a_{ph}(676)$ ratio were high due to increased absorption by photoprotective pigments.

Finally, in cluster TMix indicators of bulk particle composition take values intermediate between phytoplankton-dominated clusters, COP, TOP1, and TOP2, and the mineral-dominated cluster TMin (Fig. 6b–d). This suggests that cluster TMix represents particle assemblages in which neither organic nor inorganic particles dominate, but comprise a more balanced mixture. The phytoplankton community in cluster TMix was dominated by diatoms, with average contribution of 72 (± 25) % to Chl*a*, similar to cluster TMin. Intermediate relative concentrations of photoprotective pigments result in intermediate S_f size parameter and intermediate $a_{ph}(443)/a_{ph}(676)$ ratios.

Improved bio-optical relationships from optical classification

In this section we examine whether optical classification allows improved performance of bio-optical models. We focus on two commonly used predictive models, one for SPM from $b_{bp}(650)$ and one for Chl*a* from $a_{ph}(676)$.

Figure 8a illustrates the relationship between SPM and $b_{bp}(650)$ with observations identified by OPA membership. Statistics of prediction models set up for each OPA and for the entire dataset are given in Table 2. The regression model for the entire dataset predicted the majority of the data with relative errors ranging between 6% and 78%, with a median prediction error, MPE, of 30% (Table 2). The MPE for the entire dataset is up to three times larger than in other studies (Boss et al. 2009; Neukermans et al. 2012), which may be caused by a comparatively large number of observations in clear waters

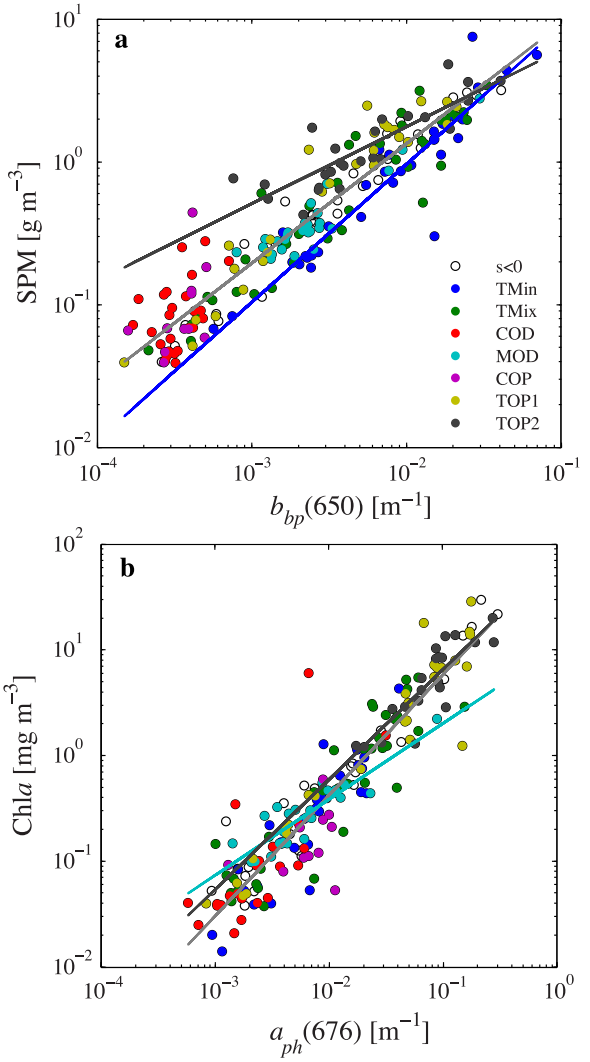

Fig. 8. Scatter plot of **(a)** particle dry mass concentration and particle backscattering coefficient at 650 nm and **(b)** total chlorophyll-*a* pigment concentration and phytoplankton absorption coefficient at 676 nm. Samples are color coded according to cluster membership. White circles indicate samples with negative silhouette values, indicative of poor classification, which are shown only for the sake of completeness. Grey lines were obtained from robust linear regression on $\log_{10}$-transformed data using all samples. Other colored lines represent linear regression obtained from $\log_{10}$-transformed data belonging to the corresponding cluster. Regression coefficients and statistics of the fits can be found in Tables 2, 3.

(SPM < 0.1 g m^{-3}) where measurement uncertainties are larger, or the larger size and particle diversity of our dataset.

Significant improvement in the prediction of SPM from $b_{bp}(650)$ was achieved when OPA membership was accounted for. For example, the MPE of the TOP1-specific prediction model was reduced to 17% and to 12% for TMin and MOD-specific models. Furthermore, OPA-specific models reduced the bias in the prediction of SPM obtained from the general model based on the entire dataset. For example, the general model overestimated SPM for samples in cluster TMin by 0.21 g m^{-3} on average, giving a median ratio of model-derived to observed values, MR, of 1.43. These biases were strongly reduced by the TMin-specific prediction model (Table 2). Another example is the underestimation of SPM by the general model for samples in cluster TOP1 by 0.33 g m^{-3} on average, giving a median MR value of 0.654, which was strongly reduced by the TOP1-specific prediction model. Our findings are consistent with observations of the effect of particle composition on the mass-specific particulate backscattering coefficient, b_{bp}/SPM, which is up to three times higher for mineral-dominated particle assemblages compared to phytoplankton-dominated particle assemblages (Neukermans et al. 2012).

OPA-specific SPM prediction models showed better performance for all clusters, except for the clear water clusters. For cluster COP, which contains only 11 observations, no significant correlation was found between b_{bp} and SPM. For cluster COD, the correlation was poor ($r = 0.44 \pm 0.38$), and regression coefficients estimates and SPM predictions exhibited large uncertainties (Table 2). This illustrates the difficulties of predicting SPM from $b_{bp}(650)$ in clear waters, which may be partly caused by larger uncertainties in measurements of particle concentration and optical properties, but possibly also due to a higher relative contribution of submicron-sized particles to $b_{bp}(650)$ which may have significant contribution to light backscattering and are not included in the estimation of particulate dry mass (Stramski and Kiefer 1991; Stramski and Woźniak 2005).

A similar analysis was conducted for the relationship between Chl*a* and $a_{ph}(676)$, illustrated in Fig. 8b. Again, OPA-specific prediction models allowed significant improvement in Chl*a* estimation with overall lower MPE, median ratios closer to 1, and reductions in median bias (Table 3). For example, the general model underestimated Chl*a* by 0.06 mg m^{-3} and 0.48 mg m^{-3} on average for observations in clusters TOP1 and TOP2, giving respective median ratios of 0.81, and 0.84. The TOP1 and TOP2-specific models allowed for significantly improved predictions with median ratios of 0.95 and 1.01, respectively. This is in correspondence with observations of low Chl*a*-specific phytoplankton absorption coefficients for these clusters (Table 1). In contrast, Chl*a* was overestimated by a factor 1.85 for samples in cluster COP. Similar to the relationship between $b_{bp}(650)$ and SPM for this cluster, the limited number of observations and small dynamic range of variability precluded establishment of a significant relationship between Chl*a* and $a_{ph}(676)$ (Table 3).

These examples illustrate the capability of optical classification for significant improvement in the estimation of particle mass concentrations such as SPM and Chl*a* from IOPs. The improvement is particularly pronounced for clusters with significantly different mass-specific IOPs, which reflect differences in particle composition in the case of b_{bp} vs.

Table 2. Coefficient estimates and goodness-of-fit statistics of linear robust regression analysis between $\log_{10}$ transformed SPM and $b_{bp}(650)$ for each optical cluster. n is the number of observations, n_x is the number of outliers as determined by the MATLAB robustfit.m routine with Talwar weighting. Outliers are removed from further analysis. If significant ($p < 0.001$), Pearson's product moment correlation coefficient r is shown with 95% confidence interval Δ. The slope and offset of the linear regression are given with their standard error estimates, Δ. The median prediction error, MPE (%), and the RMSE of the regression models are given. The median ratio of model-derived to observed values, MR (dim), and the median bias, MB (g m^{-3}), calculated as the median difference between model-derived and observed values are given for each class-specific model (OPA subscript) as well as for the general model (GM subscript). ns denotes not significant.

Cluster	n	n_x	$r \pm \Delta$	slope $\pm \Delta$	offset $\pm \Delta$	RMSE	MPE	MR_{GM}	MR_{OPA}	MB_{GM}	MB_{OPA}
TMin	33	4	0.99 ± 0.01	0.97 ± 0.03	1.92 ± 0.07	0.08	12	1.43	1.00	0.21	0.00
TMix	38	0	0.94 ± 0.06	0.84 ± 0.05	1.76 ± 0.13	0.20	27	1.13	1.02	0.06	0.01
COD	26	0	0.44 ± 0.38	0.80 ± 0.33	1.69 ± 1.15	0.22	40	1.10	1.15	0.01	0.01
MOD	27	0	0.93 ± 0.08	0.74 ± 0.06	1.52 ± 0.15	0.09	12	1.08	0.99	0.03	0.00
COP	11	0	ns	–	–	–	–	1.00	–	0.00	–
TOP1	28	2	0.98 ± 0.03	0.94 ± 0.04	2.11 ± 0.11	0.12	17	0.85	0.90	−0.10	−0.02
TOP2	25	0	0.85 ± 0.16	0.54 ± 0.07	1.32 ± 0.16	0.15	21	0.65	1.05	−0.33	0.05
General	188	4	0.94 ± 0.02	0.84 ± 0.02	1.80 ± 0.06	0.19	30	1.06	–	0.01	–

Table 3. Coefficient estimates and goodness-of-fit statistics of linear robust regression analysis between $\log_{10}$ transformed Chl*a* and $a_{ph}(676)$ for each optical cluster. Notations as in Table 2. MB in units of mg m^{-3}.

Cluster	n	n_x	$r \pm \Delta$	slope $\pm \Delta$	offset $\pm \Delta$	RMSE	MPE	MR_{GM}	MR_{OPA}	MB_{GM}	MB_{OPA}
TMin	27	2	0.94 ± 0.07	1.26 ± 0.10	2.11 ± 0.22	0.20	24	1.02	1.00	0.01	0.00
TMix	36	0	0.91 ± 0.08	1.04 ± 0.08	1.71 ± 0.17	0.30	37	0.95	0.98	−0.01	−0.01
COD	22	2	0.89 ± 0.15	0.96 ± 0.14	1.36 ± 0.36	0.20	31	0.92	0.88	−0.01	−0.01
MOD	26	0	0.89 ± 0.12	0.72 ± 0.07	1.02 ± 0.17	0.14	25	0.96	1.01	−0.01	0.00
COP	11	0	ns	–	–	–	–	1.85	–	0.12	−0.01
TOP1	27	2	0.99 ± 0.02	1.14 ± 0.04	1.97 ± 0.08	0.15	23	0.81	0.95	−0.06	−0.02
TOP2	21	0	0.87 ± 0.17	1.04 ± 0.14	1.86 ± 0.16	0.20	24	0.84	1.01	−0.48	0.03
General	170	7	0.96 ± 0.02	1.14 ± 0.03	1.91 ± 0.06	0.23	32	0.93	–	−0.01	–

SPM, or phytoplankton composition in the case of $a_{ph}(676)$ vs. Chl*a*.

Discussion

We examined an extensive in situ dataset of spectral backscattering and absorption properties of marine particles in conjunction with measurements of particle concentration, composition, and phytoplankton pigments in the Chukchi and Beaufort Seas. Using hierarchical cluster analysis of the spectral shape of the particulate backscattering-to-absorption ratio, $b_{bp}(\lambda)/a_p(\lambda)$, we developed an optically-based classification of suspended particle assemblages. Seven distinct clusters were identified, which we refer to as OPAs. Each OPA is characterized by a specific set of indicators of particle concentration, composition, and phytoplankton taxonomic composition and size. Here we provide a summary of optical and particle characteristics of each OPA (Table 4), and discuss spatial distribution (Fig. 9; Table 4) and ecological conditions associated with each OPA.

Three OPAs were dominated by phytoplankton. One was representative of clear waters dominated by small-sized phytoplankton, cluster COP. The predominance of small-sized phytoplankters in this cluster is consistent with our optical observations of very steep b_{bp} spectra, a large blue-to-red absorption ratio of a_{ph}, and weak influence of pigment packaging on phytoplankton absorption. Pigment information indicated a diverse assemblage of phytoplankton taxonomic groups, with a strong contribution of prasinophytes, likely *Micromonas pusilla* (1–2 μm diameter) (Lovejoy et al. 2007; Coupel et al. 2015). This species is known to persist throughout all seasons in the Canadian Arctic and is considered to be the major component of the photosynthetic picoeukaryotic community in Arctic waters (Lovejoy et al. 2007; Tremblay et al. 2009; Balzano et al. 2012).

Samples in this OPA were uniquely found at the subsurface Chl*a* maximum at the Beaufort Sea shelf break, located at a depth of 59 ± 10 m (Fig. 9b; Table 4). The Beaufort Sea shelf break is characterized by the presence of a deep and strong halocline, limiting nutrient replenishment in the

Table 4. Descriptive summary of the characteristics of the optically-derived particle assemblages, OPAs, and the physical environment associated with each OPA, synthesizing information given in Figs. 6, 7, 9. Photoprotection is evaluated based on the ratio of photoprotective carotenoids to Chl*a*, PPC/Chl *a*, as given in Fig. 7c. Positioning in the water column is given in three depth categories, near-surface (NS), near-bottom (NB), and at the subsurface Chl*a* maximum (SCM). M, N, P denote micro-, nano-, and picophytoplankton size classes, respectively.

			TMin	TMix	COD	MOD	COP	TOP1	TOP2
Characteristics of particle assemblage	Particle concentration		High	High	Low	Intermediate	Low	High	High
	Particle bulk composition		Rather mineral	Mixed	Detritus-dominated	Detritus-dominated	Phytoplankton-dominated	Phytoplankton-dominated	Phytoplankton-dominated
	Dominant phytoplankton size class (with major taxonomic group)		M (diatoms), wide range	M (diatoms), wide range	P and N (cryptophytes)	M (diatoms, some dinoflagellates), wide range	N (prymnesiophytes) and P (prasinophytes)	M (mostly diatoms), wide range	M (exclusively diatoms)
	Photoprotection		Intermediate	Intermediate	Intermediate	High	Low	Intermediate	Low
Physical environment	Proportion of samples at various positions in water column (%)	NS	48	39	27	89	0	64	12
		SCM	22	40	73	11	100	32	72
		NB	30	21	0	0	0	4	16
	Proportion of shelf samples (%)		73	82	23	100	9	86	96
	Proportion of samples with proportional ice-cover > 0.2 (%)		21	5	8	19	0	39	24

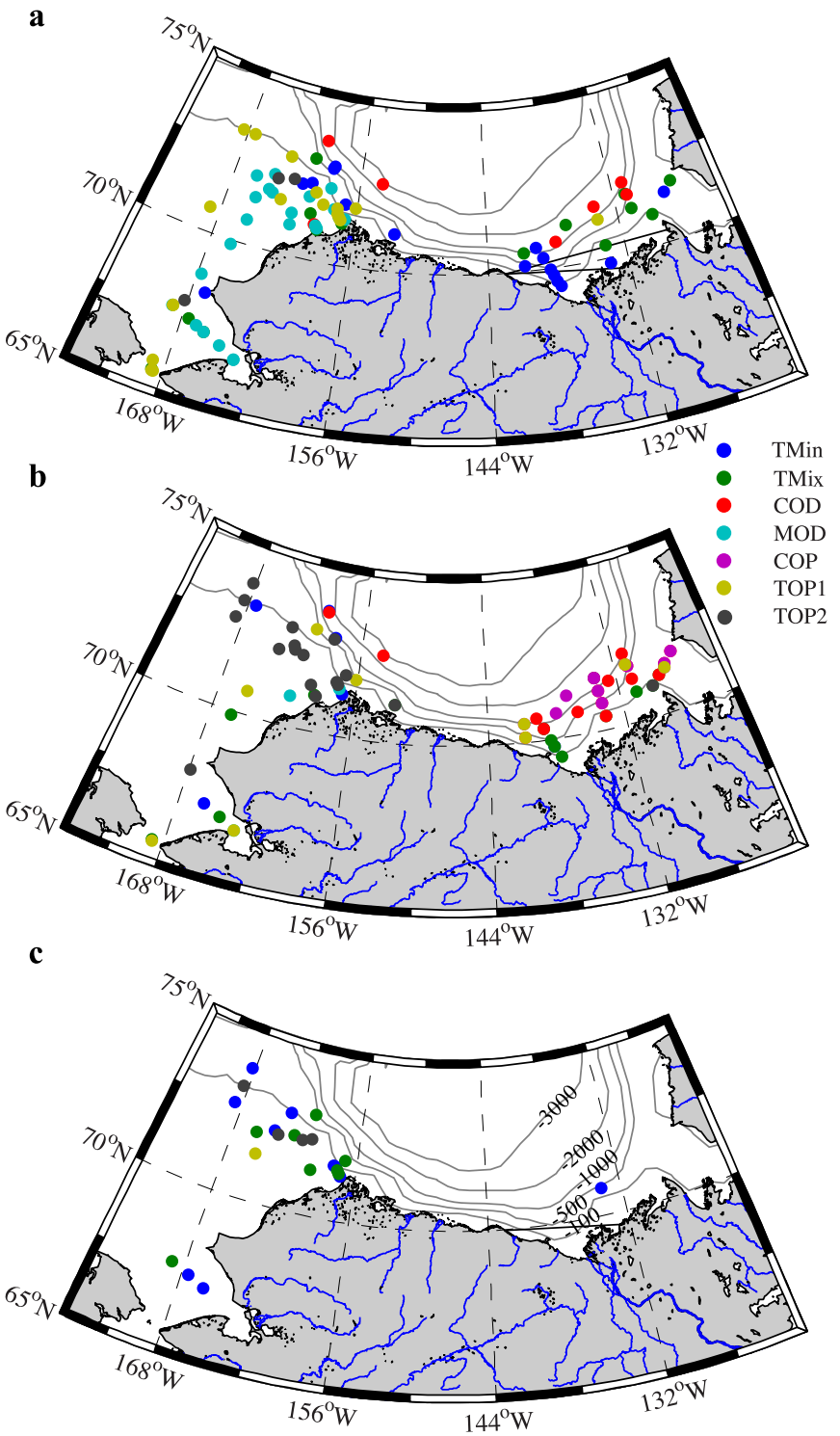

Fig. 9. Maps depicting cluster membership of samples collected **(a)** near-surface, **(b)** at the subsurface Chl*a* maximum, and **(c)** near-bottom during three cruises in the Beaufort and Chukchi Seas. Isobaths with depth in meters.

upper water column (e.g., McLaughlin and Carmack 2010; Coupel et al. 2015). These environmental conditions result in low phytoplankton biomass and productivity, and favor small flagellates such as prymnesiophytes and prasinophytes (Li et al. 2009; Tremblay et al. 2009). Such a community recycles energy and carbon in longer food chains with less

efficient transfer to higher trophic levels or export to the benthos (Tremblay et al. 2009, 2012).

The other two OPAs that were dominated by phytoplankton, clusters TOP1 and TOP2, represented waters with high phytoplankton biomass, in contrast to cluster COP. Pigment information revealed that clusters TOP1 and TOP2 were dominated by diatoms. The predominance of large-sized phytoplankters in these clusters are consistent with our optical observations of the small contribution of nonalgal matter to particulate absorption, of weaker spectral dependence of $b_{bp}(\lambda)$, and of increased influence of pigment packaging on $a_{ph}(\lambda)$.

Samples in the TOP clusters were mainly found on the Chukchi Sea shelf (Fig. 9), a shallow shelf sea supported by high nutrient supply from the Pacific Ocean. Such physical conditions generally result in high phytoplankton biomass typically dominated by diatoms, which are characteristic of new production with short food chains and an efficient transfer of carbon to higher trophic levels and the seafloor (e.g., Ardyna et al. 2011; Tremblay et al. 2012). Diatom communities in clusters TOP1 and TOP2 differed in contributions of photoprotective pigments. Diatom communities with increased photoprotection (cluster TOP1) were predominantly found in surface waters, whereas communities with reduced photoprotection (cluster TOP2) were found deeper in the water column on the Chukchi Sea shelf (Fig. 9; Table 4). Observations on the contribution of photoprotective pigments are in accordance with values of the blue-to-red absorption ratio of a_{ph}, which were higher for the surface diatom communities (cluster TOP1) than for the deeper water communities (cluster TOP2). The pigment that contributed most strongly to photoprotection in these diatom communities was diadinoxanthin, a carotenoid synthesized by diatoms to protect the cell when light levels become saturating or photoinhibitory for photosynthesis (e.g., Johnsen et al. 1994).

Two OPAs represented particle assemblages dominated by organic nonalgal material (clusters COD and MOD), which differed in terms of turbidity and phytoplankton community composition and size characteristics. Cluster COD is representative of clear waters with mainly small-sized phytoplankton with strong contributions of cryptophytes to the nanophytoplankton assemblage and chlorophytes to the picophytoplankton assemblage. Such phytoplankton communities are associated with regenerated production as expected in the stratified oligotrophic Beaufort Sea, where COD samples were mainly found (Hill et al. 2005; Ardyna et al. 2011; Coupel et al. 2012, 2015).

Cluster MOD, on the other hand, is representative of moderately turbid waters containing a diverse phytoplankton community with a slight dominance of microphytoplankton in which diatoms and dinoflagellates coexist and where chlorophytes contribute strongly to the pigment-based class of picophytoplankton. Also, the contribution of photoprotective phytoplankton pigments was highest in this cluster, as could be expected for phytoplankton communities in surface waters exposed to high irradiance or for communities under nutrient stress (Brunet et al. 1996; Staehr et al. 2002). The majority of samples assigned to cluster MOD were found in the near-surface waters of the Chukchi Sea (Table 4; Fig. 9a), a shallow nutrient-rich shelf sea, where extensive diatom blooms were observed beneath sea ice during both ICESCAPE cruises (Arrigo et al. 2012, 2014). Observations assigned to cluster MOD were sampled on average 28 ± 10 days after sea ice break-up and may thus reflect a postbloom situation, in correspondence with dominance of organic nonalgal particles and a planktonic community that is shifting toward larger contributions of dinoflagellates and smaller taxonomic groups promoting increased recycling of carbon and materials within the water column.

The significantly different spectral slope of absorption by nonalgal particles, S_{NAP}, between these two nonalgal-dominated OPAs suggests strong differences in the characteristics of the nonalgal material. This can be expected as the characteristics of the algal material that is broken down also significantly differ between these OPAs. Because small-sized phytoplankton dominates in cluster COD, nonalgal particles resulting from the breakdown or grazing of such phytoplankton are likely also smaller than in cluster MOD, which is consistent with observations of very steep particulate backscattering spectra for this cluster. Furthermore, cluster MOD is more likely to contain nonalgal material that had previously sunk to the seafloor, brought back to the surface by resuspension processes typical for the shallow Chukchi shelf sea (e.g., Carmack et al. 2006). However, microbial heterotrophs also contribute to nonalgal particulate absorption, so the observed differences in S_{NAP} could indicate different planktonic communities, namely different proportions of bacteria or heterotrophic flagellates.

Cluster TMin, associated with turbid mineral-dominated particle assemblages, is easiest to optically differentiate from the other clusters owing largely to its distinct spectral shape of absorption typically associated with mineral particles (Babin and Stramski 2004; Stramski et al. 2007). The contribution of nonalgal matter to the total particulate absorption is highest of all clusters (Fig. 5a; Table 1), as is the particulate backscattering ratio, $\tilde{b}_{bp}=b_{bp}/b_p$. The value of $\tilde{b}_{bp}$ obtained at 660 nm, 0.022 ($\pm$0.007), is in good agreement with previous observations for suspensions of inorganic particles (Boss et al. 2009; Neukermans et al. 2012). Samples belonging to cluster TMin were mainly found in the plumes of the Colville and Mackenzie Rivers and near the seafloor, as would be expected from riverine transport of mineral-rich material and resuspension of bottom sediments (Fig. 9a,c; Table 4). Finally, mineral-dominated near-surface samples were also found further offshore at the Chukchi Sea shelf in the vicinity of melting sediment-laden sea ice (Fig. 9a).

Lastly, particle assemblages of mixed composition and relatively high concentration are represented by cluster TMix, with observations scattered throughout all locations (Fig. 9; Table 4). This cluster had intermediate values for optical properties and particle characteristics from the other clusters.

The OPAs presented in this work are derived from HCA of the spectral shape of $b_{bp}(\lambda)/a_p(\lambda)$. An analogous analysis based on $a_p(\lambda)$ alone partitioned the dataset in seven clusters with stronger differences in phytoplankton characteristics, but reduced differences in particle concentration and composition, compared to the partitioning based on $b_{bp}(\lambda)/a_p(\lambda)$ (not shown). We also performed an analogous analysis on multispectral $b_{bp}(\lambda)/a_p(\lambda)$ data obtained through spectral integration of $b_{bp}(\lambda)/a_p(\lambda)$ over the spectral bands of the MODIS ocean color sensor in the 400–690 nm range with central wavelengths at 412, 443, 488, 531, 555, 667, and 678 nm. Hierarchical cluster analysis of multispectral data partitioned the dataset into four optical classes (not shown). Samples from clusters TOP1 and TOP2 were merged in one cluster. Similarly, samples from clusters COD and COP were merged in a single cluster. The third multispectral-based cluster comprised all TMin samples, but also samples from clusters MOD and TMix, while the fourth cluster comprised most of the TMix samples, and samples from clusters MOD and a few from cluster TOP1. These results demonstrate that the capability to optically differentiate distinct assemblages of suspended particles was strongly reduced by sub-sampling the high spectral resolution data of $b_{bp}(\lambda)/a_p(\lambda)$ with multispectral resolution.

Conclusions

Our results demonstrate a capability to optically differentiate assemblages of marine particles (OPAs), representative of distinct ecological conditions, based on hyperspectral measurements of the particle absorption and backscattering coefficients (the latter obtained from a fit to multispectral measurements). Analogous analysis based on multispectral optical data strongly reduced this capability. This capability has potential application to many autonomous and remote-sensing observations. Sensors for in situ hyperspectral measurements of $a_p(\lambda)$ and multispectral measurements of $b_{bp}(\lambda)$ are commercially available and can be deployed from various in situ platforms including profilers and autonomous underwater vehicles. The recent development of improved models to partition $a_p(\lambda)$ into phytoplankton and nonphytoplankton components enable estimation of these component coefficients (e.g., Zheng and Stramski 2013; Zheng et al. 2015). Such a partitioning could provide parallel estimates of the composition of the particulate assemblage from $a_d(400)/a_p(400)$ and provide indicators of the characteristics of the algal and nonalgal particle community. Furthermore, these coefficients are obtainable from satellite measurements of ocean color using a combination of existing inverse models (e.g., Loisel and Stramski 2000; Lee et al. 2002; Maritorena et al. 2002), among which one has been evaluated in Arctic waters (Zheng et al. 2014), and models that partition the absorption coefficient of seawater $a(\lambda)$ into particulate and dissolved component coefficients. A class-matching procedure would then determine assignment of a measured spectrum of $b_{bp}(\lambda)/a_p(\lambda)$ to an optical class, in analogy with existing approaches for ocean color remote sensing reflectance (Mélin et al. 2011; Vantrepotte et al. 2012; Moore et al. 2014). Class-specific SPM and Chl*a* algorithms, such as discussed in this study, could then be applied to improve estimates of biogeochemical stocks. Further, identification of OPAs at synoptic spatial and temporal scales will improve our understanding of the spatiotemporal dynamics of marine particles and ecological processes in the ocean.

References

Alvain, S., C. Moulin, and Y. Dandonneau. 2005. Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. Deep-Sea Res. **52**: 1989–2004. doi:10.1016/j.dsr.2005.06.015

Ardyna, M., M. Gosselin, C. Michel, M. Poulin, and J.-É. Tremblay. 2011. Environmental forcing of phytoplankton community structure and function in the Canadian High Arctic: Contrasting oligotrophic and eutrophic regions. Mar. Ecol. Prog. Ser. **442**: 37–57. doi:10.3354/meps09378

Arrigo, K. R., and others. 2012. Massive phytoplankton blooms under Arctic sea ice. Science **336**: 1408–1408. doi: 10.1126/science.1215065

Arrigo, K. R., and others. 2014. Phytoplankton blooms beneath the sea ice in the Chukchi Sea. Deep-Sea Res., Part II **105**: 1–16. doi:10.1016/j.dsr2.2014.03.018

Babin, M., D. Stramski, G. M. Ferrari, H. Claustre, A. Bricaud, G. Obolensky, and N. Hoepffner. 2003a. Variations in the light absorption coefficients of phytoplankton, non-algal particles, and dissolved organic matter in coastal waters around Europe. J. Geophys. Res. **108**: 3211. doi:10.1029/2001JC000882

Babin, M., A. Morel, V. Fournier-Sicre, F. Fell, and D. Stramski. 2003b. Light scattering properties of marine particles in coastal and oceanic waters as related to the particle mass concentration. Limnol. Oceanogr. **48**: 843–859. doi:10.4319/lo.2003.48.2.0843

Babin, M., and D. Stramski. 2004. Variations in the mass-specific absorption coefficient of mineral particles suspended in water. Limnol. Oceanogr. **49**: 756–767. doi: 10.4319/lo.2004.49.3.0756

Balzano, S., D. Marie, P. Gourvil, and D. Vaulot. 2012. Composition of the summer photosynthetic pico and nanoplankton communities in the Beaufort Sea assessed by T-RFLP and sequences of the 18S rRNA gene from flow cytometry sorted samples. ISME J. **6**: 1480–1498. doi: 10.1038/ismej.2011.213

Boss, E., and others. 2009. Comparison of inherent optical properties as a surrogate for particulate matter concentration in coastal waters. Limnol. Oceanogr.: Methods **7**: 803–810. doi:10.4319/lom.2009.7.803

Bowers, D. G., G. E. L. Harker, and B. Stephan. 1996. Absorption spectra of inorganic particles in the Irish sea and their relevance to remote sensing of chlorophyll. Int. J. Remote Sens. **17**: 2449–2460. doi:10.1080/01431169608948782

Bricaud, A., H. Claustre, J. Ras, and K. Oubelkheir. 2004. Natural variability of phytoplankton absorption in oceanic waters: influence of the size structure of algal populations. J. Geophys. Res. **109**: C11010. doi:10.1029/2004JC002419

Bricaud, A., A. M. Ciotti, and B. Gentili. 2012. Spatial-temporal variations in phytoplankton size and colored detrital matter absorption at global and regional scales, as derived from twelve years of SeaWiFS data (1998-2009). Global Biogeochem. Cycles **26**: GB1010. doi:10.1029/2010GB003952

Buiteveld, H., J. H. M. H. M. Hakvoort, and M. Donze. 1994. The optical properties of pure water. Proc. SPIE **2258**: 174. doi:10.1117/12.190060

Brunet, C., D. Davoult, and R. Casotti. 1996. Physiological reactions to a change in light regime in cultured Skeletonema costatum (Bacillariophyta): Implications to estimation of phytoplankton biomass. Hydrobiologia **333**: 87–94. doi:10.1007/BF00017571

Carmack, E., D. Barber, J. Christensen, R. Macdonald, B. Rudels, and E. Sakshaug. 2006. Climate variability and physical forcing of the food webs and the carbon budget on panarctic shelves. Progr. Oceanogr. **71**: 145–181. doi:10.1016/j.pocean.2006.10.005

Chang, G. C., and others. 2004. The new age of hyperspectral oceanography. Oceanography **17**: 16–23. doi:10.5670/oceanog.2004.43

Ciotti, A.M., M. R. Lewis, and J. J. Cullen. 2002. Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. Limnol. Oceanogr. **47**: 404–417. doi:10.4319/lo.2002.47.2.0404

Coupel, P., and others. 2012. Phytoplankton distribution in unusually low sea ice cover over the Pacific Arctic. Biogeosciences **9**: 4835–4850. doi:10.5194/bg-9-4835-2012

Coupel, P., and others. 2015. The impact of freshening on phytoplankton production in the Parcific Arctic Ocean. Progr. Oceanogr. **131**: 113–125. doi:10.1016/j.pocean.2014.12.003

Dana, D. R., and R. A. Maffione. 2002. Determining the backward scattering coefficient with fixed-angle backscattering sensors—revisited. In: Proceedings of Ocean Optics XVI, Santa Fe, New Mexico.

Devred, E., S. Sathyendranath, V. Stuart, and T. Platt. 2011. A three component classification of phytoplankton absorption spectra: Application to ocean-color data. Remote Sens. Environ. **115**: 2255–2266. doi:10.1016/j.rse.2011.04.025

Dunne, J. P., R. A. Armstrong, A. Gnanadesikan, and J. L. Sarmiento. 2005. Empirical and mechanistic models for the particle export ratio. Global Biogeochem. Cycles **19**: GB4026. doi:10.1029/2004GB002390

Duysens, L. M. N. 1956. The flattening effect of the absorption spectra of suspensions as compared to that of solutions. Biochem. Biophys. Acta **19**: 1–12.

Estapa, M., E. Boss, L. M. Mayer, and C. S. Roesler. 2012. Role of iron and organic carbon in mass-specific light absorption by particulate matter from Louisiana coastal waters. Limnol. Oceanogr. **57**: 97–112. doi:10.4319/lo.2012.57.1.0097

Fasham, M. J. R., H. W. Ducklow, and S. M. McKelvie. 1990. A nitrogen-based model of plankton dynamics in the oceanic mixed layer. J. Mar. Res. **48**: 591–639. doi:10.1357/002224090784984678

Fennel, K., I. Cetinic, E. D'Asaro, C. Lee, and M. J. Perry. 2011. Autonomous data describe North Atlantic spring bloom. EOS **92**: 465–466. doi:10.1029/2011EO500002

Gieskes, W. W. C., G. W. Kraay, A. Nontji, D. Setiapermana, and D. Sutomo. 1988. Monsoonal alternation of a mixed and a layered structure in the phytoplankton of the euphotic zone of the Banda Sea (Indonesia): A mathematical analysis of algal pigment fingerprints. Neth. J. Sea Res. **22**: 123–137. doi:10.1016/0077-7579(88)90016-6

Gordon, H. R., and A. Morel. 1983. Remote assessment of ocean color for interpretation of satellite visible imagery. A review. Springer-Verlag.

Gordon, H. R., O. B. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker, and D. K. Clark. 1988. A semianalytical radiance model of ocean color. J. Geophys. Res. **93**: 10909–10924. doi:10.1029/JD093iD09p10909

Hill, V., G. Cota, and D. Stockwell. 2005. Spring and summer phytoplankton communities in the Chukchi and Eastern Beaufort Seas. Deep-Sea Res., Part II **52**: 3369–3385. doi:10.1016/j.dsr2.2005.10.010

IOCCG. 2006. Remote sensing of inherent optical properties: fundamentals, tests of algorithms, and applications. In Z.-P. Lee [ed.], Reports of the International Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada.

IOCCG. 2014. Phytoplankton functional types from space. *In* S. Sathyendranath [ed.], Reports of the International Ocean-Colour Coordinating Group, No. 15, IOCCG, Dartmouth, Canada.

Jeffrey, S.W., and M. Vesk. 1997. Introduction to marine phytoplankton and their pigment signatures, p. 37–84. *In* S. W. Jeffrey, R. F. C. Mantoura, and S. W. Wright [eds.], Phytoplankton pigments in oceanography. UNESCO.

Johnsen, G., O. Samset, L. Granskog, and E. Sakshaug. 1994. In vivo absorption characteristics in 10 classes of bloom-forming phytoplankton: Taxonomic characteristics and responses to photoadaptation by means of discriminant

and HPLC analysis. Mar. Ecol. Prog. Ser. **105**: 149–157. doi:10.3354/meps105149

Johnson, K. S., and others. 2009. Observing biogeochemical cycles at global scales with profiling floats and gliders: Prospects for a global array. Oceanography **22**: 216–225. doi:10.5670/oceanog.2009.81

Jonasz, M., and G. R. Fournier. 2007. Light scattering by particles in water—theoretical and experimental foundations. Elsevier.

Kamvar, S. D., D. Klein, and C. D. Manning. 2002. Interpreting and extending classical agglomerative clustering algorithms using a model-based approach. ICML 2002 Proceedings of the Nineteenth International Conference on Machine Learning, p. 283–290.

Kirk, J. T. O. 1996. Light and photosynthesis in aquatic ecosystems, 2nd ed. Cambridge Univ. Press.

Kishino, M., M. Takahashi, N. Okami, and S. Ichimura. 1985. Estimation of spectral absorption coefficients of phytoplankton in the sea. Bull. Mar. Sci. **37**: 634–642.

Knap, A., A. Michaels, A. Close, H. Ducklow, and A. Dickson. 1996. Protocols for the Joint Global Ocean Flux Study (JGOFS) Core Measurements, JGOFS Report Nr. 19 (Reprint of the IOC Manuals and Guides No. 29, UNESCO, Paris, 1994), 170 pp.

Kostadinov, T. S., D. A. Siegel, and S. Maritorena. 2009. Retrieval of the particle size distribution from satellite ocean color observations. J. Geophys. Res. **114**: C09015. doi:10.1029/2009JC005303

Le Quéré, C., and others. 2005. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. Glob. Chang. Biol. **11**: 2016–2040. doi:10.1111/j.1365-2486.2005.01004.x

Lee, Z. P., K. L. Carder, and R. A. Arnone. 2002. Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. Appl. Opt. **41**: 5755–5772. doi:10.1364/AO.41.005755

Legendre, P., and L. Legendre. 2012. Numerical ecology, 3rd. ed. Elsevier.

Li, W. K. W., F. A. McLaughlin, C. Lovejoy, and E. C. Carmack. 2009. Smallest algae thrive as the Arctic Ocean freshens. Science **326**: 539–539. doi:10.1126/science.1179798

Loisel, H., and D. Stramski. 2000. Estimation of the inherent optical properties of natural waters from the irradiance attenuation coefficient and reflectance in the presence of Raman scattering. Appl. Opt. **39**: 3001–3011. doi:10.1364/AO.39.003001

Loisel, H., J. M. Nicolas, A. Sciandra, D. Stramski, and A. Poteau. 2006. Spectral dependency of optical backscattering by marine particles from satellite remote sensing of the global ocean. J. Geophys. Res. **111**: C09024. doi:09010.01029/02005JC003367

Lovejoy, C., and others. 2007. Distribution, phylogeny, and growth of cold-adapted picoprasinophytes in Arctic seas. J. Phycol. **43**: 78–89. doi:10.1111/j.1529-8817.2006.00310.x

Lubac B., H. Loisel, N. Guiselin, R. Astoreca, X. Mériaux, and F. Artigas. 2008. Hyperspectral and multispectral ocean color inversions to detect *Phaeocystis globosa* in coastal waters. J. Geophys. Res. **113**: C06026. doi:10.1029/2007JC004451

Maffione, R. A., and D. R. Dana. 1997. Instruments and methods for measuring the backward-scattering coefficient of ocean waters. Appl. Opt. **36**: 6057–6067. doi:10.1364/AO.36.006057

Maritorena, S., D. A. Siegel, and A. R. Peterson. 2002. Optimization of a semianalytical ocean color model for global-scale applications. Appl. Opt. **41**: 2705–2714. doi:10.1364/AO.41.002705

McLaughlin, F. A., and E. C. Carmack. 2010. Deepening of the nutricline and chlorophyll maximum in the Canada Basin interior, 2003–2009. Geophys. Res. Lett. **37**: L24602. doi:10.1029/2010GL045459

Mélin, F., V. Vantrepotte, M. Clerici, D. D'Alimonte, G. Zibordi, and J. F. Berthon. 2011. Multi-sensor satellite time series of optical properties and chlorophyll a concentration in the Adriatic Sea. Progr. Oceanogr. **31**: 229–244. doi:10.1016/j.pocean.2010.12.001

Moore, T. S., M. D. Dowell, S. Bradt, and A. V. Verdu. 2014. An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. Remote Sens. Environ. **143**: 97–111. doi:10.1016/j.rse.2013.11.021

Morel, A. 1973. Diffusion de la lumière par les eaux de mer; résultats expérimentaux et approche théorique. [The scattering of light by sea water: Experimental results and theoretical approach]. In Optics of the sea, interface, and in-water transmission and imaging. NATO AGARD Lect. Ser. **61**, pp. 3.1.1–3.1.76.

Morel, A., and A. Bricaud. 1981. Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. Deep-Sea Res., Part A **28**: 1375–1393. doi:10.1016/0198-0149(81)90039-X

Neukermans, G., H. Loisel, X. Meriaux, R. Astoreca, and D. McKee. 2012. In situ variability of mass-specific beam attenuation and backscattering of marine particles with respect to particle size, density, and composition. Limnol. Oceanogr. **57**: 124–144. doi:10.4319/lo.2011.57.1.0124

Neukermans, G., R. A. Reynolds, and D. Stramski. 2014. Contrasting inherent optical properties and particle characteristics between an under-ice phytoplankton bloom and open water in the Chukchi Sea. Deep Sea Res., Part II **105**: 59–73. doi:10.1016/j.dsr2.2014.03.014

Ras, J., J. Uitz, and H. Claustre. 2008. Spatial variability of phytoplankton pigment distributions in the Subtropical South Pacific Ocean: comparison between in situ and modelled data. Biogeosciences **5**: 353–369. doi:10.5194/bg-5-353-2008

Röttgers, R., and S. Gehnke. 2012. Measurement of light absorption by aquatic particles: Improvement of the

quantitative filter technique by use of an integrating sphere approach. Appl. Opt. **51**: 1336–1351. doi:22441480

Rousseeuw, P. J. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Comp. Appl. Math. **20**: 53–65. doi:10.1016/0377-0427(87)90125-7

Roy, S., C. Llewellyn, E. S. Egeland, and G. Johnsen. 2011. Phytoplankton pigments: Characterization, chemotaxonomy and applications in oceanography. Cambridge Univ. Press.

Salvador, S., and P. Chan. 2004. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms, p. 576–584. In Proceedings of 16th IEEE International Conference on Tools with AI.

Slade, W., and E. Boss. 2015. Spectral attenuation and backscattering as indicators of average particle size. Appl. Opt. **54**: 7264–7277. doi:10.1364/AO.54.007264

Sokal, R. R., and F. J. Rohlf. 1962. The comparison of dendrograms by objective methods. Taxon **11**: 33–40. doi:10.2307/1217208

Staehr, P. A., P. Henriksen, and S. Markager. 2002. Photoacclimation of four marine phytoplankton species to irradiance and nutrient availability. Mar. Ecol. Prog. Ser. **238**: 47–59. doi:10.3354/meps238047

Stramski, D., and D. A. Kiefer. 1991. Light scattering by microorganisms in the open ocean. Prog. Oceanogr. **28**: 343–383. doi:10.1016/0079-6611(91)90032-H

Stramski, D., and S. B. Woźniak. 2005. On the role of colloidal particles in light scattering in the ocean. Limnol. Oceanogr. **50**: 1581–1591. doi:10.4319/lo.2005.50.5.1581

Stramski, D., M. Babin, and S. B. Woźniak. 2007. Variations in the optical properties of terrigenous mineral-rich particulate matter suspended in seawater. Limnol. Oceanogr. **52**: 2418–2433. doi:10.4319/lo.2007.52.6.2418

Stramski, D., and others. 2008. Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern South Pacific and eastern Atlantic Oceans. Biogeosciences **5**: 171–201. doi:10.5194/bg-5-171-2008

Stramski, D., R. A. Reynolds, S. Kaczmarek, J. Uitz, and G. Zheng. 2015. Correction of pathlength amplification in the filter-pad technique for measurements of particulate absorption coefficient in the visible spectral region. Appl. Opt. **54**: 6763–6782. doi:10.1364/AO.54.006763

Stukel, M. R., K. A. S. Mislan, M. Décima, and L. Hmelo. 2014. Detritus in the pelagic ocean. Eco-DAS IX Symposium Proceedings, ASLO. doi:10.4319/ecodas.2014.978-0-9845591-3-8.49

Tremblay, G., C. Belzile, M. Gosselin, M. Poulin, S. Roy, and J. E. Tremblay. 2009. Late summer phytoplankton distribution along a 3500 km transect in Canadian Arctic waters: strong numerical dominance by picoeukaryotes. Aquat. Microb. Ecol. **54**: 55–70. doi:10.3354/ame01257

Tremblay, J. É., and others. 2012. Carbon and nutrient fluxes. *In* D. G. Barber [ed.], On the edge: from action to knowledge during the fourth International Polar Year. Arboriginal Issues Press.

Twardowski, M. S., H. Claustre, S. A. Freeman, D. Stramski, and Y. Huot. 2007. Optical backscattering properties of the 'clearest' natural waters. Biogeosciences **4**: 1041–1058. doi:10.5194/bg-4-1041-2007

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. J. Geophys. Res. **111**: C08005. doi:08010.01029/02005JC003207

van de Hulst, H. C. 1957. Light scattering by small particles. John Wiley & Sons.

Van Heukelem, L., and C. S. Thomas. 2001. Computer-assisted high-performance liquid chromatography method development with applications to the isolation and analysis of phytoplankton pigments. J. Chromatogr. A **910**: 31–49. doi:10.1016/S0378-4347(00)00603-4

Vantrepotte, V., H. Loisel, D. Dessailly, and X. Meriaux. 2012. Optical classification of contrasted coastal waters. Remote Sens. Environ. **123**: 306–323. doi:10.1016/j.rse.2012.03.004

Vidussi, F., H. Claustre, B. B. Manca, A. Luchetta, and J. C. Marty. 2001. Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. J. Geophys. Res. **106**: 19939–19956. doi:10.1029/1999JC000308

Ward, J. H. 1963. Hierarchical groupings to optimize an objective function. J. Am. Stat. Assoc. **58**: 234–244. doi:10.1080/01621459.1963.10500845

Woźniak, S. B., and others. 2010. Optical variability of seawater in relation to particle concentration, composition, and size distribution in the nearshore marine environment at Imperial Beach, California. J. Geophys. Res. **115**: C08027. doi:10.1029/2009JC005554

Zheng, G., and D. Stramski. 2013. A model for partitioning the light absorption coefficient of suspended marine particles into phytoplankton and nonalgal components. J. Geophys. Res.: Oceans **118**: 2977–2991. doi:10.1002/jgrc.20082

Zheng, G., D. Stramski, and R. A. Reynolds. 2014. Evaluation of the Quasi-Analytical Algorithm for estimating the inherent optical properties of seawater from ocean color: Comparison of Arctic and lower-latitude waters. Rem. Sens. Environ. **155**: 194–209. doi:10.1016/j.rse.2014.08.020

Zheng, G., D. Stramski, and P. M. DiGiacomo. 2015. A model for partitioning the light absorption coefficient of natural waters into phytoplankton, nonalgal particulate, and colored dissolved organic components: A case study for the Chesapeake Bay. J. Geophys. Res.: Oceans **120**: 2601. doi:10.1002/2014JC010604

Acknowledgments

We thank D. Doxaran, J. Ehn, J. Tatarkiewicz, S. Watanabe, and G. Zheng for assistance in field data collection and processing. We are grateful to S. Bélanger for providing the MALINA particulate absorption data, A. Matsuoka for the MALINA and ICESCAPE CDOM absorption data, and J. Ras for MALINA HPLC analyses. We thank the captain and crews of the USCGC Healy and CCGS Amundsen, and fellow scientists on the ICESCAPE and MALINA cruises for assistance and supporting data. K. Arrigo and his coworkers are thanked for providing ICESCAPE HPLC data. The University of Maryland's Horn Point Analytical Services Laboratory is acknowledged for determinations of pigment concentrations and the Marine Science Institute Analytical Laboratory, University of California at Santa Barbara for determinations of particulate organic carbon. We also thank I. Cetinic and one anonymous reviewer for their comments and suggestions. This work was supported by grants from the NASA Cryospheric Sciences Program (NNX07AR20G) and Ocean Biology and Biogeochemistry Program (NNX10AG05G). The MALINA project was co-funded by ANR (Agence Nationale de la Recherche), INSU-CNRS (Institut National des Sciences de l'Univers – Centre National de la Recherche Scientifique), CNES (Centre National d'Etudes Spatiales), and ESA (European Space Agency).

Submitted 4 October 2015
Revised 17 February 2016
Accepted 30 March 2016

Associate editor: David Antoine