JGR Biogeosciences

RESEARCH ARTICLE
10.1029/2025JG008956

Highly Resolved Surface Phytoplankton Community Composition Along the British Columbia Coast, Derived From In Situ Hyperspectral Radiometry

Perumthuruthil Suseelan Vishnu[1,2], Justin Del Bel Belluz[3], Hongyan Xi[4], Midhun Shah Hussain[5], Astrid Bracher[4,6], and Maycira Costa[1]

[1]Spectral Remote Sensing Laboratory, University of Victoria, Victoria, BC, Canada, [2]Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway, [3]Hakai Institute, Victoria, BC, Canada, [4]Helmholtz Centre for Polar and Marine Research, Phytooptics Group, Physical Oceanography of Polar Seas, Climate Sciences, Alfred Wegener Institute, Bremerhaven, Germany, [5]Department of Marine Biology, Microbiology and Biochemistry, Cochin University of Science and Technology, Kochi, India, [6]Department of Physics and Electrical Engineering, Institute of Environmental Physics, University of Bremen, Bremen, Germany

Key Points:

- A hyperspectral-based algorithm was tested to quantitatively retrieve phytoplankton groups in a turbid estuary
- Algorithm trained with hyperspectral data outperformed multispectral data in retrieving all phytoplankton groups
- Diatom blooms occurred in the high-salinity western region, while raphidophytes were concentrated near the Fraser River plume

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

P. S. Vishnu,
psvishnu2014@gmail.com

Citation:

Vishnu, P. S., Del Bel Belluz, J., Xi, H., Shah Hussain, M., Bracher, A., & Costa, M. (2025). Highly resolved surface phytoplankton community composition along the British Columbia Coast, derived from in situ hyperspectral radiometry. *Journal of Geophysical Research: Biogeosciences*, *130*, e2025JG008956. https://doi.org/10.1029/2025JG008956

Received 20 MAR 2025
Accepted 22 SEP 2025

Abstract Quantitative measurements of phytoplankton community composition (PCC) are essential for understanding fisheries production, ocean nutrient cycling, and the export of particulate carbon to the ocean interior. However, these measurements are constrained in dynamic coastal waters due to the spatial-temporal constraints of in situ sampling, difficulty quantifying communities, and the challenges of deriving community compositions via satellites. Here, we work to address these issues by using highly resolved in situ hyperspectral radiometry, along a ship of opportunity track through Case-2 waters of the Strait of Georgia (SoG) British Columbia, to derive phytoplankton community composition. First, an empirical orthogonal function (EOF)-based algorithm was developed using HPLC CHEMTAX-derived phytoplankton group-level chlorophyll-a (Chla) and Total Chla (TChla) concentrations and corresponding principal components derived from hyperspectral remote sensing reflectance. Second, the outputs were evaluated using cross-validation, showing good retrievals for TChla and the regionally dominant phytoplankton groups: diatoms, cryptophytes, green algae, and raphidophytes, which followed expected spatial-temporal trends with diatom-dominated spring blooms and succession to high diversity flagellate-dominated summer conditions. Furthermore, the outputs captured fine spatial scale trends including strong harmful raphidophyte blooms over the narrow transition to low salinity Fraser River plume influenced waters. These findings highlight the potential of using highly resolved hyperspectral radiometry to derive fine-scale trends in phytoplankton group level community composition in optically dynamic coastal waters. Coupled with additional measures, this method could provide valuable information on phytoplankton dynamics in the SoG, which is a critical habitat for a high diversity of pelagic fish species, including Pacific salmon.

Plain Language Summary Deriving phytoplankton community composition is crucial for understanding their distinct roles in biogeochemical cycles, fisheries production, and carbon export to the ocean interior. In this study, we use continuous hyperspectral radiometric measurements alongside in situ phytoplankton data collected from a passenger ferry operating across the dynamic Strait of Georgia to develop an algorithm for retrieving information on different phytoplankton groups. Our hyperspectral algorithm outperforms previous satellite-based observations trained on the same data set, demonstrating improved retrieval accuracy for all phytoplankton groups. Specifically, the algorithm captured the diatom-dominated spring bloom, which was primarily observed in the high-salinity western portion of the transect, followed by a community shift toward smaller phytoplankton species by summer. Additionally, the algorithm successfully retrieved the concentration of toxic raphidophyte blooms, which were predominantly located near the Fraser River plume waters. Overall, our findings highlight the potential of algorithms trained with continuous radiometric measurements to resolve fine-scale phytoplankton dynamics that would not be captured by conventional ship-based sampling.

1. Introduction

Phytoplankton contribute to more than half of the planet's primary production, playing a central role in the marine food web and driving global biogeochemical cycling (e.g., Field et al., 1998; Le Quéré et al., 2005). Within the

overarching phytoplankton grouping, many species exist with distinct ecological and biogeochemical traits and niches. As a result, variability in phytoplankton community composition often has significant implications on the transfer of energy through the marine food web, influencing higher trophic levels, including fisheries production and vertical export of carbon and nutrients to the deep ocean (Friedland et al., 2021; Guidi et al., 2009; Le Quéré et al., 2005). Phytoplankton communities are strongly linked and quickly respond to physical (e.g., temperature, salinity, and sunlight), chemical (e.g., inorganic nutrients), and biological (e.g., grazing pressure and viral infection) drivers (Evans & Brussaard, 2012; Margalef, 1978), making them key indicators of environmental and anthropogenic change. Therefore, understanding phytoplankton community dynamics is important for furthering our knowledge of marine systems in light of a changing climate (Henson et al., 2021).

Numerous in situ techniques have been developed to assess phytoplankton composition both quantitatively and qualitatively, including microscopy, cell imaging using flow cytometry (Dashkova et al., 2017; Olson & Sosik, 2007), high-performance liquid chromatography (HPLC) pigment analysis, molecular sequencing techniques (Y. Lin et al., 2019), and digital inline holographic microscopy (MacNeil et al., 2022). These techniques have advantages and limitations: many are labor-intensive, time-consuming, expensive, and sparse in space and time (see IOCCG, 2014; Lombard et al., 2019, for a detailed review). These limitations can be largely addressed through the use of remote sensing satellites, which enable the synoptic-scale monitoring of the world surface ocean at high spatial and temporal resolutions, often unattainable with ship-based observations. For the last four decades, satellite-based observations have allowed quantitative assessments of surface ocean total chlorophyll-a concentrations (hereafter TChla) used as a proxy for phytoplankton biomass (e.g., Dai et al., 2023; McClain, 2009). Total chlorophyll-a does not provide direct insights into community compositions, and contemporary research has strived to develop methods to derive compositions and size classes via remote sensing. These methods are either categorized as abundance-based (e.g., Brewin et al., 2010; Hirata et al., 2011; Uitz et al., 2006), ecological-based (e.g., Mouw, 2019; Palacz et al., 2013; Raitsos et al., 2008), reflectance\radiance-based (e.g., Alvain et al., 2005; Ciotti & Bricaud, 2006; Devred et al., 2006), and backscattering-based (e.g., Kostadinov et al., 2009). A comprehensive description of these methodologies can be found in IOCCG (2014), Bracher et al. (2017), and Mouw et al. (2017). These methods have been hindered by current multispectral ocean color satellites being spectrally limited, making it difficult or impossible to detect narrow spectral absorption features associated with pigment differences between phytoplankton species groupings (Dierssen et al., 2021; Kramer et al., 2022). Recently, promise has been shown with the accurate quantification of three phytoplankton groups using high spectral resolution atmospheric sensors (Bracher et al., 2009; Losa et al., 2017).

Existing studies have shown that hyperspectral remote sensing reflectance (Rrs; $sr^{-1}(\lambda)$) with 5 nm resolution or better in the visible spectrum (400–700 nm) is necessary to extract the spectral signature of phytoplankton groups in both Case-1 and Case-2 waters (Kramer et al., 2022; Wolanin et al., 2016; Xi et al., 2017). For example, hyperspectral optical data have been used to identify spectral peaks associated with biomarker pigments (e.g., Torrecilla et al., 2011; Uitz et al., 2015), to retrieve accessory pigment concentrations on a global scale (Kramer et al., 2022), to detect the toxic harmful dinoflagellate, *Karenia brevis* (e.g., Craig et al., 2006), and to derive chlorophyll-a concentrations (hereafter Chla) of multiple phytoplankton groups across the Atlantic Ocean (e.g., Bracher et al., 2020) and North Pacific (Konik et al., 2024). Besides the exploitation of atmospheric sensors with large pixels (>5 km; Bracher et al., 2021), recently launched hyperspectral sensors, such as DESIS, PRISMA, and EnMAP, have shown initial promise over inland and coastal waters (Bracher et al., 2021; Storch et al., 2023) but are limited by spatial and temporal coverage and low sensitivity for dark waters (Giardino et al., 2020). The recent Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission from NASA, launched in February 2024, is the first global ocean hyperspectral satellite equipped with an imaging spectro-radiometer Ocean Color Instrument (OCI) with a 5-nm resolution over the visible spectrum (Werdell et al., 2019). When compared to existing multispectral sensors, this resolution will advance the global retrieval of phytoplankton composition through improved inherent optical properties (IOPs) inversions (Cael et al., 2020) and extraction of accessory pigment spectral peaks (Chase et al., 2013; Kramer et al., 2022). Nonetheless, the spatial resolution of this sensor is likely too coarse (1.2 km) to accurately derive phytoplankton composition from highly dynamic coastal settings.

The Strait of Georgia (hereafter SoG; Figure 1) is a productive and dynamic estuary located on the west coast of Canada between Vancouver Island and the Mainland of British Columbia (BC). Primary productivity is relatively high (280 g C m^{-2}) throughout the SoG, with elevated levels often found directly outside the Fraser River, which is the dominant source of freshwater to the estuary. Phytoplankton dynamics in the SoG follow seasonal trends comparable to many temperate estuarine systems. In winter, phytoplankton are light-limited, and low biomass

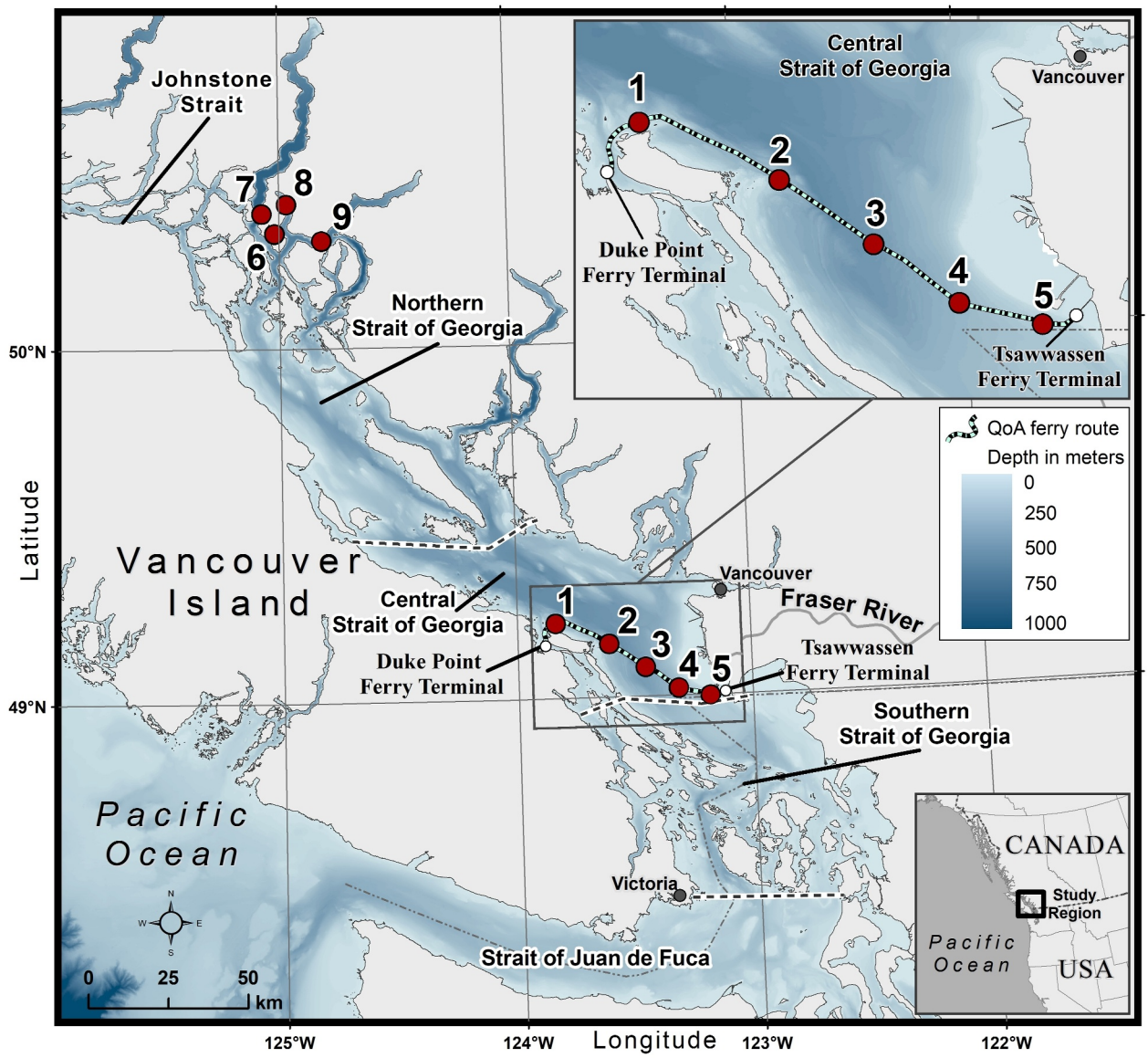

Figure 1. Study area map and in situ data collection locations for sensor data along the ferry line highlighted in gray and for discrete water samples marked as red circles in the central and northern Strait of Georgia. The running lines represent the Queen of Alberni ferry's track from Duke Point, Nanaimo on Vancouver Island to Tsawwassen on Mainland, British Columbia. The color scale shows depth in meters. The dashed lines spanning the Strait of Georgia delineate the boundaries between the four main components: the Strait of Juan de Fuca and the southern, central and northern regions.

conditions are observed, which are dominated by small flagellates (Harrison et al., 1983; Nemcek et al., 2023). In the spring, high biomass (>15 mg/m^3) diatom-dominated blooms develop occurring between late February and mid-April (Marchese et al., 2022; Suchy et al., 2022) with bloom initiation timing driven by reduced wind velocity, mixing and cloud cover working to lift light limitation to surface phytoplankton (Allen & Wolfe, 2013; Collins et al., 2009; Peña et al., 2016). The spring bloom is often terminated by nutrient depletion, and in late spring, stratification develops, limiting surface nutrient availability and phytoplankton transition to moderate biomass flagellate-dominated (cryptophytes and green algae) communities (Del Bel Belluz et al., 2021; Nemcek et al., 2023; Peña et al., 2016). During this season, large harmful raphidophyte blooms, typically consisting of *Heterosigma akashiwo*, can sporadically form with detrimental effects on fisheries and aquaculture (Esenkulova et al., 2021). In autumn, increased wind mixing renews nutrients to the surface layer, and periodically, large diatom blooms are observed prior to the return of strong winds and low biomass winter conditions (Del Bel Belluz et al., 2021). While considerable knowledge of broad phytoplankton community dynamics exists for the region,

much of this knowledge comes from spatially rich, but temporally limited, seasonal cruises (e.g., Nemcek et al., 2023) or temporally rich, but spatially limited, time series (e.g., Del Bel Belluz et al., 2021). Hyperspectral remote sensing has potential to fill in knowledge gaps in these variable and dynamic waters.

Similar to many estuarine systems, the SoG is optically complex, complicating the use of remote sensing to derive phytoplankton community compositions (Vishnu et al., 2022). The Fraser River is one of the main sources of optically active constituents to the central SoG, and during peak summer discharge, high concentrations of colored dissolved organic matter (0.007–5 m^{-1}) and total suspended matter (0.82–65.63 mg/L) are observed in the vicinity of plume waters. These river influenced waters are also highly productive with transitional waters directly outside of the plume showing reduced light attenuation, lifting light limitation, and high nutrients, which elevate primary productivity and phytoplankton biomass (Johannessen et al., 2021; Sutton et al., 2013). The Fraser River has less influence in the northern SoG and optical signatures are more phytoplankton driven (Del Bel Belluz et al., 2021; Loos & Costa, 2010). Less optical characterization has been performed in these more northern waters where smaller watersheds and large fjord systems contribute to the area's fresh water and may introduce localized optical complexity. The optically complex nature of the SoG necessitates the use of hyperspectral methods to untangle the signatures of different phytoplankton groups.

Here, we aim to derive surface phytoplankton community dynamics from hyperspectral radiometry aboard a ship of opportunity regularly traversing the SoG on the west coast of British Columbia, Canada. The utilized algorithm was adapted from Bracher et al. (2020), employing an empirical orthogonal function (EOF)-based hyperspectral method trained on HPLC TChla and CHEMTAX phytoplankton group-level outputs. Cross-validation techniques were employed to quantitatively assess the robustness of the fitted regression model and compare the outputs with the expected regional phytoplankton trends. Finally, the corresponding sea surface salinity data obtained from the ship's ferrybox were used to investigate the influence of Fraser River plume waters on the derived phytoplankton community composition. This work provides an approach to systematically derive highly resolved phytoplankton composition from a ship of opportunity, which could fill knowledge gaps required for fisheries management of the many regional culturally and economically valuable species and aid in the development of algorithms for hyperspectral satellites such as PACE.

2. Materials and Methods

2.1. Data Acquisition

This section outlines the methodology for data acquisition and processing of physical and biogeochemical variables, including sea surface salinity (SSS), total chlorophyll-a concentration (TChla), total suspended matter (TSM) concentration, and absorption by colored dissolved organic matter (aCDOM), which were used to describe regional bio-optical properties. Additionally, the acquisition and processing of the above-water hyperspectral data are described, and a regional EOF-based hyperspectral algorithm to retrieve the Chla concentration of phytoplankton group composition is presented.

2.1.1. Ferrybox Data

All physical and chemical oceanographic properties, such as SSS and Chla derived from fluorescence (CHL_{FL}), were obtained using a ferrybox system installed on the commercial BC ferry Queen of Alberni (QoA), which runs across the SoG from Duke Point, Nanaimo (Vancouver Island), to Tsawwassen, Vancouver (Mainland) (stations 1–5; Figure 1), facilitated by Ocean Networks Canada (ONC) (https://data.oceannetworks.ca/DataSearch; Travers-Smith et al., 2021). The ferrybox was installed below the main deck, drawing seawater through a 1.5-m pipe from approximately 2 m depth (Travers-Smith et al., 2021). Specifically, SSS was acquired using a Seabird SBE19 thermosalinograph, whereas CHL_{FL} was estimated using a WET Labs ECO Triplet fluorometer and corrected for both biofouling and nonphotochemical quenching (NPQ) (Owens et al., 2022; Travers-Smith et al., 2021). ONC maintained and cleaned these sensors weekly or biweekly, depending on the season, followed by applying a comprehensive data quality control process (Owens et al., 2022). Only fully quality-controlled data were used for analysis. The frequency of the data acquisition was 1 Hz, and the quality-checked sensor-acquired data were binned at 1-min intervals to match the in situ water samples.

In the northern SoG (Figure 1), SSS was acquired using a calibrated Fisherbrand traceable salinity meter pen (accuracy ±2% Full-scale +1 digit).

2.1.2. In Situ Water Samples

In situ data were collected from two locations in the SoG. In the central SoG (Figure 1; $N = 108$), discrete water samples were collected in triplicate at approximately 2 m depth (Travers-Smith et al., 2021) from the ferry flow-through system for the determination of HPLC pigments, TSM, and CDOM. These samples were collected during the 2018 and 2019 spring (March–April) and summer (May–September) seasons. Once collected samples were immediately filtered and stored in the dark to avoid degradation (Mueller et al., 2003).

In the northern SoG (Figure 1; stations 6–9; $N = 18$), water samples for the same sample suite were collected from a small vessel from ∼2 m depth using a 5-L Niskin bottle.

2.1.2.1. Absorption by Colored Dissolved Organic Matter

CDOM samples collected from each station were filtered through 0.2 µm GF/F membrane filters and stored at 4°C in acid-cleaned, precombusted amber bottles until analysis, which occurred within 24 hr. The spectral absorption coefficient of the CDOM was measured using a USB400 Ocean Optics spectrophotometer, with the absorption spectra being corrected by using freshly prepared deionized (DI) water as blank. The CDOM absorption coefficient (m^{-1}) was subsequently calculated from the sample absorbance values, applying the Beer-Lambert Law (Ferrari et al., 1996).

$$\text{aCDOM}\ (\lambda) = \frac{2.3030 * A\ ()}{L} \tag{1}$$

A (λ) is the spectral absorbance of the sample, and L is the path length of the cuvette (0.1 m). The final aCDOM(λ) for each station was calculated as the triplicate average with the coefficient of variation (CV) < 25% for all sample sets.

2.1.2.2. Total Suspended Matter Concentration

For TSM analysis, 0.5–2 L of water were filtered onto precombusted and preweighed 0.7 µm GF/F filters. After filtration, deionized (DI) water was passed through the filters to minimize sea-salt retention (Stavn et al., 2009). The filters were then stored on dry ice, transported to the laboratory, and kept at −80°C until being dried for 24 hr at 60°C and weighed with the final TSM concentrations estimated following Röttgers et al. (2014). Final TSM concentrations were determined as the averages of triplicate measurements, with a CV < 25% for all sample sets.

2.1.2.3. HPLC Pigment Concentrations

Duplicate water samples (0.5–2 L) were filtered through 25 mm GF/F 0.7 µm filters under low pressure (≤5 mmHg) for phytoplankton pigment analysis. After the filtration, filters were immediately frozen on dry ice and kept in a dark environment, transported to the laboratory, and stored at −80°C until extraction. Pigment analysis was performed using a Shimadzu HPLC system at the University of South Carolina Baruch Institute of Marine and Coastal Sciences, following Pinckney (2010).

The HPLC-derived phytoplankton pigment data set was used to derive the Chla concentration of the multiple phytoplankton groups based on the CHEMTAX matrix factorization method (v 1.95) (Mackey et al., 1996). Among the 23 HPLC identified accessory pigments, the following pigments were used in the CHEMTAX analysis: chlorophyll c_1c_2 ($Chlc_1c_2$), peridinin (Perid), 19′butanoyloxyfucoxanthin (19′BF), fucoxanthin (Fuco), 19′hexanoyloxyfucoxanthin (19′HF), prasinoxanthin (Prasino), violaxanthin (Viola), alloxanthin (Allox), zeaxanthin (Zea), lutein (Lut), chlorophyll-b (Chlb), and total Chlorophyll-a (TChla). Input ratios and phytoplankton groups used for CHEMTAX analysis are shown in Tables S1 and S2 in Supporting Information S1. Utilized groups were cyanobacteria-1 (Cyano), haptophytes (Hapto), green algae (GA), cryptophytes (Crypto), dinoflagellates (Dino), raphidophytes (Raphido, central SoG only), dictyochophytes (Dictyo), and diatoms (Diat). These input ratios and groups were based on Del Bel Belluz et al. (2021) and the same analysis and output data were used by Vishnu et al. (2022) to develop similar algorithms to this study, but with Sentinel-3A satellite data: additional information on CHEMTAX analysis can be found there.

2.2. In Situ Remote Sensing Reflectance Spectra

Two different data acquisition approaches were conducted for the two sampling regions due to equipment setup logistics, with the central SoG ferry data acquired using a (a) autonomous solar tracker HyperSAS (hereafter SAS-ST) and the northern SoG using a (b) manually operated HyperSAS (hereafter SAS) operated from a small vessel; both systems use the same OCR hyperspectral radiometers from the manufacturer Seabird/Satlantic, Inc. The SAS-ST system was deployed on the BC ferry QoA to collect total radiance and irradiance data, as described in Wang and Costa (2022) and Giannini et al. (2021). The system is equipped with three hyperspectral radiometers: one for measuring the total radiance (i.e., Lt(λ)) from the sea-viewing sensor, another for capturing sky radiance (i.e., Li(λ)), and a third for assessing total irradiance (i.e., Es(λ)). Using these measurements, the remote sensing reflectance, Rrs(λ), was calculated according to the methodology outlined by Mobley (1999). The ship runs from Duke Point, Nanaimo to Tsawassen (Vancouver) at a speed of 10 m/s and crosses a distance of 45 km (Figure 1). The SAS-ST was installed 14 m above the water surface to avoid shadowing effects from the ship, superstructure, and spray (Wang & Costa, 2018), and the instrument was fully calibrated by the manufacturer (Satlantic, Inc.). The Lt (λ) and Li (λ) sensors were installed with fixed viewing zenith angles, $\theta v = 45°$, and viewing-sun azimuth, $\varphi v = 120°$, to minimize the influence of direct sun glint (Hooker & Morel, 2003). The optimal φv was achieved using a "steeper motor platform" (Satlantic Inc), which autonomously adjusted the sensor to the required position based on the ship heading and sun azimuth. The ship's superstructure (ship wall) had a notable impact on above-water radiometry data as it reflected the signal onto the field of view measured by the Lt(λ) sensor. We addressed this superstructure influence on derived Rrs(λ) by implementing the method thoroughly outlined in Giannini et al. (2021) and Wang and Costa (2022), in accordance with the approach of Hooker and Morel (2003). Consequently, we apply a correction factor of 0.00005 for the QoA (Giannini et al., 2021).

Radiometry data were collected between 1:00 p.m. and 2:30 p.m. on the ferry track run from the Duke Point ferry terminal to the Tsawwassen ferry terminal (west to east) as this traverse coincided with optimal sun conditions (Nasiha et al., 2022). The data from the SAS-ST were processed with PySciDON (Python Scientific Framework for Development of Ocean Network application) to calculate Rrs(λ) (Vandenberg et al., 2017). Data were acquired every 2 s and binned into 1-min intervals during the processing to match the discrete water samples and ferrybox. The mean, median, and CV were calculated for each 1-min bin, ensuring that each 1-min Rrs(λ) value had a CV of <10% following Nasiha et al. (2022).

Rrs(λ) was determined using the equation below.

$$\mathrm{Rrs}(\lambda) = \frac{\mathrm{Lt}(\lambda) - \rho s \mathrm{Li}(\lambda)}{\mathrm{Es}(\lambda)}, \left(\mathrm{sr}^{-1}\right) \tag{2}$$

where ρ_s is the fraction of sky radiance that is reflected by a wavy sea surface into the sensor's field of view (Mobley, 1999), obtained from the look-up table provided by Mobley (1999) based on sky condition, viewing geometry, sea state, and wind speed (Mobley, 1999). Wind data were acquired from the meteorological stations on Entrance Island, Sandhead CS, and Tsawwassen Ferry Auto (https://climate.weather.gc.ca/historical_data/search_historic_data_e.html). All the Rrs(λ) data from the SoG were corrected for the bidirectional reflectance distribution function (BRDF) effect using the code by Wang and Costa (2022) following the method of Lee et al. (2011).

For the northern SoG, the SAS system was installed on the bow of the small sampling vessel. Above-water reflectance acquired from this region was recorded using SatView™ software and processed with Prosoft™ software 7.7.16. The Rrs(λ) calculation was the same as for the SAS-ST.

2.3. Regionally Tuned EOF-Based Hyperspectral Algorithm

The EOF-based algorithm to derive phytoplankton group Chla concentration utilizes in situ hyperspectral Rrs(λ) across the visible spectrum (400–700 nm) with a 1-nm interval (Bracher et al., 2020). A similar approach has been applied globally to estimate phytoplankton taxa using multispectral Rrs(λ) data (Xi et al., 2020). Here, we adapted the Xi et al. (2020) algorithm to the SoG data by considering 103 valid matchup data sets comprising SAS-ST and SAS Rrs(λ), and the CHEMTAX-derived phytoplankton community composition. A schematic workflow of the regression model development and validation is shown in Figure 2. Following Xi et al. (2020), the hyperspectral SAS-ST and SAS Rrs(λ) were subject to singular value decomposition (SVD) to extract the different modes of

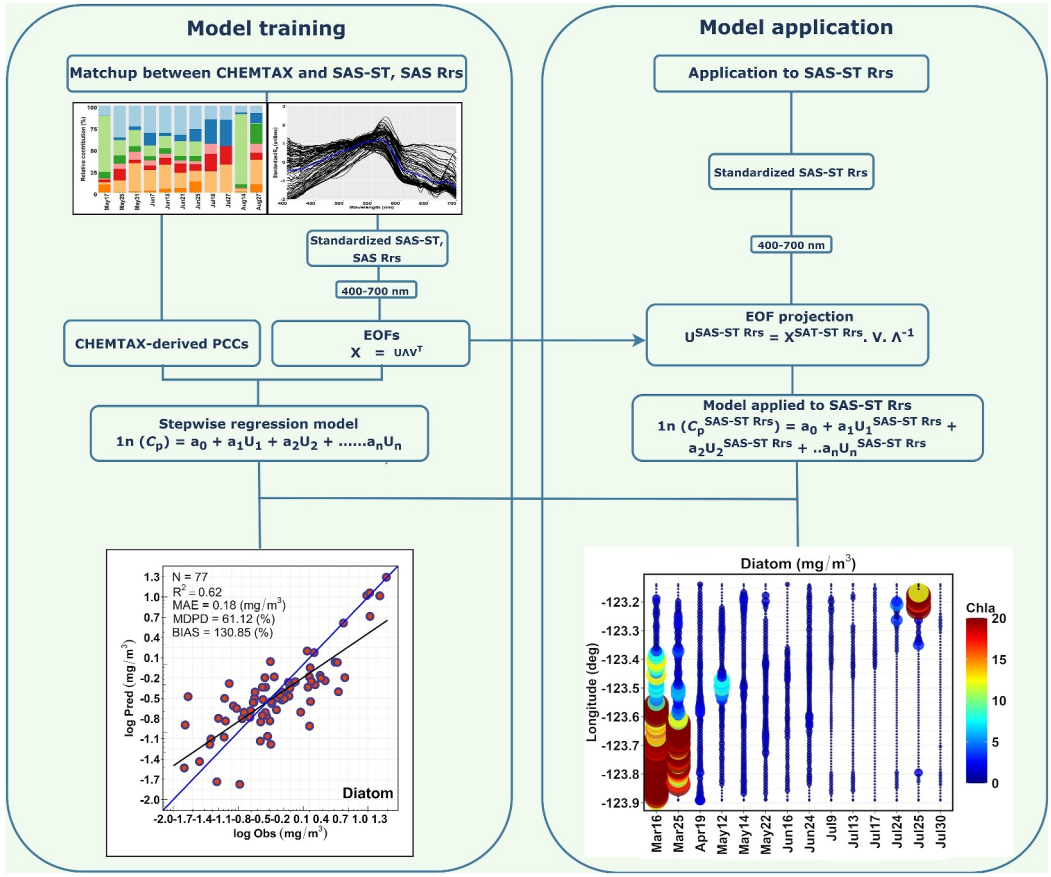

Figure 2. The flowchart illustrates the method to derive phytoplankton community composition using the empirical orthogonal function-based algorithm on hyperspectral remote sensing reflectance. The left section represents the algorithm training process, utilizing a matchup data set that integrates phytoplankton composition estimates from CHEMTAX with corresponding hyperspectral reflectance measurements. On the right, the trained algorithm is applied to hyperspectral reflectance data to generate spatial maps of different phytoplankton groups.

variability (i.e., EOF modes) in the reflectance spectrum. The key steps of this methodology are outlined as follows.

Step 1: Before extracting the EOF modes, SAS-ST and SAS Rrs(λ) were standardized by the approach outlined by Xi et al. (2020). First, the standardized Rrs(λ) data, denoted as matrix X, were collocated to the respective phytoplankton group Chla. Then, the standardized SAS-ST and SAS Rrs(λ) underwent SVD to obtain the dominant mode of variability within the spectral data.

$$X = U \Sigma V^T \tag{3}$$

where matrix U is the principal component scores (eigenvalues), matrix V^T is the identified spectral patterns (also known as eigenvectors), and vector Σ, singular values representing the importance of each EOF mode.

The reflectance data from Step 1 were then binned into 1-min bins to match the discrete water samples and ferrybox data. The following criteria were followed for the matchup: <1-hr difference and the coefficient of variation (CV) across the visible spectrum (400–700 nm) being <10%. Median values were used to avoid outliers (Bailey & Werdell, 2006).

Step 2: A generalized linear regression model (GLM) was developed using log-transformed CHEMTAX-derived phytoplankton group Chla (Cp) and a selected subset of principal components. The principal components were chosen through a stepwise process, following Xi et al. (2020), to ensure the model's efficiency by

minimizing the Akaike information criterion (AIC). Only the most significant components were included in the final model. The regression algorithm for the phytoplankton group Chla prediction is as follows:

$$ln(Cp) = a + b_1u_1 + b_2u_2 + \cdots..b_nu_n \quad (4)$$

where a is the intercept, $b_{1,2\ldots n,}$ are the regression coefficients for the selected principal components, and $b\,u_{1,2,\ldots n}$ are the n number of principal components (U).

Step 3: A cross-validation (or model evaluation) technique was then performed to assess the fitted regression model robustness using 500 permutations, employing an 80/20 data split for training and validation. The training data set (80%) included X^{train} (standardized SAS-ST and SAS Rrs(λ)) and C^{train} (CHEMTAX phytoplankton groups Chla), and the validation data set (20%) included X^{val} and C^{val}. A regression model selected through a stepwise routine was fitted between ln (C^{train}) and corresponding principal components (U^{train}) for each permutation. A pair of the EOF-predicted and CHEMTAX-derived phytoplankton group Chla concentration (C_p^{val} and C_o^{val}, respectively) for the 500 iterations were recorded for the evaluation of the regression model.

As part of the model evaluation, the predictive performance of the fitted regression model was assessed using key statistical metrics, such as coefficient of determination (R^2), median absolute error (MDAE), median percentage difference (MDPD), and bias. The R^2, along with the slope (S) and intercept (a) of the GLM, was derived from the logarithmic relationship between predicted ln ($Chla_{pi}$) versus observed ln ($Chla_{oi}$) phytoplankton group Chla concentration. In contrast, the MDAE, MDPD, and bias were computed using nonlog transformed data. These statistical formulas were provided by Xi et al. (2020), except MDAE, which is expressed as follows:

$$\text{MDAE} = \text{Median value of } |(Chla_{pi} - Chla_{oi})| \quad (5)$$

where ($Chla_{pi}$) is the model-predicted Chla concentration of each phytoplankton group, ($Chla_{oi}$) is the CHEMTAX-derived phytoplankton group Chla concentration, and M is the number of observations.

Of these model assessment metrics, MDAE represents absolute differences between predicted and observed values and reflects the typical magnitude of prediction error while being less sensitive to extreme values compared to the mean. MDPD represents the median of the absolute percentage differences between predicted and observed values, providing a robust estimate of the typical relative error in model predictions. Finally, bias provides an important metric for model performance and is extensively used in this study. Positive bias indicates systematic overestimation, negative bias indicates underestimation, and a bias of 0% reflects perfect average agreement between predicted and observed values.

Step 4: The final step was to apply the regional EOF-based algorithm to the hyperspectral data acquired along the ferry track to derive highly resolved phytoplankton group Chla. This step was not performed on the Northern SoG SAS data due to its discrete nature. To do this, we projected the standardized SAS-ST onto the eigenvector to obtain the new principal component (U) sets. Then, this new pair of principal components (U^{sat}) was fed into Equation 4 to derive the Chla concentration of multiple phytoplankton groups, where the intercepts (a) and the regression coefficients ($b_{1,2\ldots n}$) were obtained from Equation 4. Finally, an intercomparison was performed for TChla utilizing the QA/QC ferrybox-derived CHL_{FL} data obtained during 2018 and 2019. The matchup extraction criteria used were the same as those used in the model training. It is worth noting that this intercomparison was restricted to only EOF-derived TChla due to the limited availability of independent CHEMTAX-derived phytoplankton group Chla data that could be spatially and temporally matched with the ferry track samples. Following the full-fit and cross-validation steps, the phytoplankton group Chla retrieved from the algorithm underwent additional qualitative evaluation based on previous literature.

3. Results

3.1. Physical and Biogeochemical Properties

Surface physical and biogeochemical variables demonstrated patterns characteristic of optically complex coastal waters showing high spatial and seasonal variability (Figure 3; Table 1). In general, SSS showed the lowest

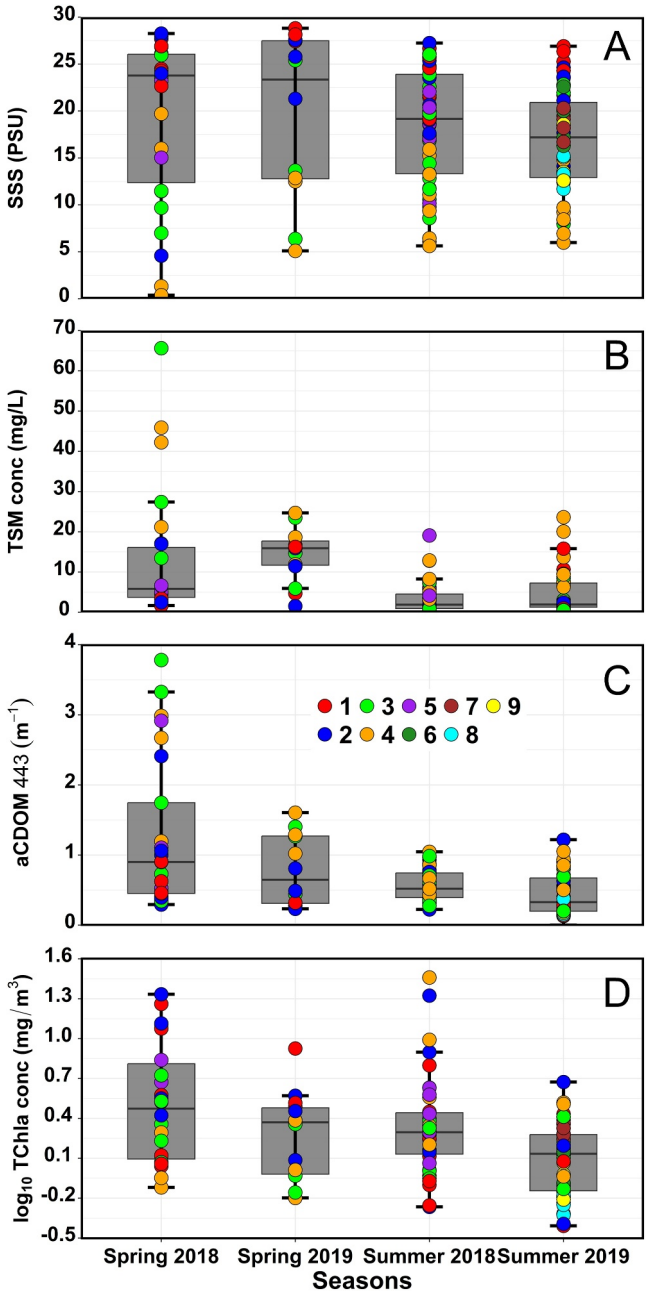

Figure 3. Seasonal variability of (a) sea surface salinity (PSU), (b) total suspended matter (TSM; mg/L) concentration, (c) absorption by colored dissolved organic matter (aCDOM 443; (m^{-1})) at 443 nm, (d) total chlorophyll-a (TChla; mg/m^3, shown on a $\log_{10}$ scale for consistency) concentration. The box plot boundaries correspond to the first and third quartiles, while the whiskers extending beyond the boxes indicate the standard error for each quartile. Colored points denote data from various stations (see Figure 1), with data points outside the third quartile indicating extreme values. During the summer of 2018, a very high TChla concentration (120 mg/m^3) was measured at Station 5, which is not shown in the figure. Note that the range of variability for each biogeochemical variable represents spring and summer of 2018 and 2019 from the central and northern Strait of Georgia. However, samples from the northern Strait of Georgia represent only the summer of 2019 (mainly July).

salinities coinciding with the Fraser River plume and the highest salinities on the Vancouver Island (western) side of the transect. Annually, spring 2018 showed the lowest SSS magnitudes (0.38 PSU) at Station 4, and 2019 showed the highest magnitudes (28.82 PSU) at Station 1, located on the western side of the transect. TSM and aCDOM 443 always showed the highest magnitudes within the Fraser River plume, with maximum values in the spring of 2018, coinciding with the lowest observed SSS and less variability and lower magnitudes in the spring of 2019. In both years, TSM and aCDOM 443 showed lower magnitudes and less variability (4.61 ± 5.30 mg/L, 0.57 ± 0.23 m^{-1}, respectively) in summer than in spring conditions (12.89 ± 15.98 mg/L, 1.40 ± 1.28 m^{-1}, respectively). TChla (Figure 3; Table 1) ranged from 0.39 to 28.82 mg/m^3 and showed higher values in 2018 than in 2019. During the summer of 2018, a very high TChla concentration (120.11 mg/m^3) was measured at Station 5 (not shown in the table or the figure as it was far outside the range of the other values) close to the Fraser River plume, highlighting summer bloom conditions in this year. The low concentrations observed in spring 2019 were likely a result of the later sampling period missing the spring bloom this year. Furthermore, 2019 showed low TChla concentrations throughout most of the summer season.

3.2. EOF Analysis on Hyperspectral Remote Sensing Reflectance

Figure 4 shows the hyperspectral Rrs(λ) associated with their corresponding in-water constituents. The above-water Rrs(λ) spectra showed a broad peak observed between 550 and 700 nm, primarily due to the high backscattering from inorganic suspended sediments. The low reflectance found in the blue (400 nm) region was driven by high aCDOM 443 absorption, and the high green (550 nm) reflectance coupled with the Chla fluorescence peak at 680 nm was representative of highly varying phytoplankton biomass, typical of the Case-2 waters along the BC coast (e.g., Giannini et al., 2021; Komick et al., 2009; Phillips & Costa, 2017).

Application of SVD to the Rrs(λ) spectra identified nine significant EOF modes (Figure 5), with the first three modes accounting for 99.43% of the total variance. The ΔAIC results applied to these significant EOFs to develop regression models for each phytoplankton group are shown in Table 2. Among the nine significant EOFs, EOF-3 emerged as the most important predictor for TChla, diatoms, dictyochophytes, raphidophytes, and cryptophytes. In contrast, EOF-5 was the key predictor for haptophytes and prasinophytes, while EOF-8 was the most important term for predicting dinoflagellates and cyanobacteria.

The spectral shapes of each EOF were regulated by the particles within the surface layer (Figure 5). EOF-1 (Figure 5), which explained 67.65% of the variance, had a spectral shape primarily driven by phytoplankton with a characteristic reflectance peak at 550 nm and a Chla fluorescence peak at 680 nm. EOF-2 explained 30.06% of the total variance, characterized by a sharp decrease from 550 to 400 nm, driven by CDOM absorption, and a broad 550–700 nm reflectance peak, indicative of backscattering by inorganic suspended sediments. EOF-3, which explained 1.72% of the overall variability, exhibited a pronounced trough between 500 and 400 nm, indicative of light attenuation by specific phytoplankton pigments, followed by a secondary peak near 660 nm. The higher EOF modes (i.e., EOF-4 and above) only explained a nominal portion of the total variance (≤0.30%).

Table 1
Seasonal Distribution of Biogeochemical Variables Such as Sea Surface Salinity (SSS; PSU), Total Suspended Matter Concentration (TSM; mg/L), Absorption by Colored Dissolved Organic Matter at 443 nm (aCDOM 443; (m^{-1})), and Total Chlorophyll-a Concentration (TChla; (mg/m^3)), During 2018 and 2019

		Physical and biogeochemical variables			
Year	Season	SSS (PSU)	TSM (mg/L)	aCDOM 443 (m^{-1})	TChla (mg/m^3)
2018	Spring	[0.38–28.26]	[1.69–65.63]	[0.29–4.92]	[0.76–21.56]
		19.41 (8.84)	**12.89** (15.98)	**1.40** (1.28)	**5.22** (5.71)
2019	Spring	[5.11–28.82]	[1.56–24.71]	[0.23–1.61]	[0.63–3.72]
		19.59 (8.91)	**14.68** (6.85)	**0.79** (0.51)	**2.09** (1.17)
2018	Summer	[5.63–27.24]	[0.50–19.12]	[0.22–1.05]	[0.54–28.82]
		18.39 (6.10)	**3.52** (3.70)	**0.57** (0.23)	**3.57** (5.46)
2019	Summer	[5.99–26.91]	[0.50–23.62]	[0.13–1.22]	[0.39–4.71]
		16.83 (5.49)	**4.61** (5.30)	**0.45** (0.29)	**1.46** (0.89)

Note. Minimum and maximum are given in square brackets; the average is in bold, and the standard deviation is in parentheses. Note that the range of variability for each biogeochemical variable represents spring and summer of 2018 and 2019 from the central and northern Strait of Georgia. However, samples from the northern Strait of Georgia represent only the summer of 2019 (mainly July).

3.3. Algorithm Evaluation

Regressions between TChla and the Chla concentration of the eight phytoplankton groups derived using CHEMTAX (x-axis) and the EOF-based hyperspectral algorithm (y-axis) are shown in Figure 6. Overall, 103 valid matchup samples were utilized for the regression model prediction; however, the number of valid matchup samples differed among phytoplankton groups mainly because some groups were not present in some samples (i.e., Chla = 0 mg/m^3). For example, out of the 103 valid matchup samples, dictyochophytes ($N = 66$) and diatoms ($N = 77$) contributed to relatively low numbers of samples. Overall, correlation coefficients using the full-fit statistical results showed that TChla and phytoplankton group Chla were well predicted with the EOF-based hyperspectral model. The best results ($p < 0.0001$) were observed for TChla ($R^2 = 0.81$; Slope = 0.84; MDAE = 0.46 mg/m^3; MDPD = 24.88%), followed by diatoms ($R^2 = 0.61$; Slope = 0.71; MDAE = 0.18 mg/m^3; MDPD = 61.12%), raphidophytes ($R^2 = 0.57$; Slope = 0.69; MDAE = 0.13 mg/m^3; MDPD = 65.01%), green algae ($R^2 = 0.57$; Slope = 0.66; MDAE = 0.09 mg/m^3; MDPD = 35.45%), and cryptophytes ($R^2 = 0.53$; Slope = 0.59; MDAE = 0.13 mg/m^3; MDPD = 40.14%; Table 3). Moderate retrieval accuracy ($p < 0.0001$) was found for haptophytes ($R^2 = 0.47$; Slope = 0.50; MDAE = 0.03 mg/m^3; MDPD = 43.88%; Table 3). Weaker correlations ($p < 0.0001$) were noted for dinoflagellates ($R^2 = 0.24$; Slope = 0.36; MDAE = 0.08 mg/m^3; MDPD = 49.68%), dictyochophytes ($R^2 = 0.36$; Slope = 0.42; MDAE = 0.07 mg/m^3; MDPD = 47.71%), and cyanobacteria ($R^2 = 0.38$; Slope = 0.43; MDAE = 0.03 mg/m^3; MDPD = 46.30%; Table 3). The MDPD ranged from 24.88% to 65.01%, corresponding to TChla (24.88%) and raphidophytes (65.01%), respectively. MDAE was calculated on nonlog transformed data; the highest MDAE was recorded for diatoms (0.18 mg/m^3) and TChla (0.46 mg/m^3), followed by raphidophytes (0.13 mg/m^3). In contrast, the lowest MDAE was noted for haptophytes and cyanobacteria (0.03 mg/m^3).

Cross-validation statistics were used to evaluate the robustness of the fitted regression model (Table 3). For the diatom, raphidophyte, green algae, cryptophyte, and haptophyte groups, the statistics obtained from the cross-validation ($R^2_{cv} = 0.60$; $MDAE_{cv} = 0.21$ mg/m^3; $MDPD_{cv} = 68.10\%$, $R^2_{cv} = 0.41$; $MDAE_{cv} = 0.13$ mg/m^3; $MDPD_{cv} = 71.87\%$, $R^2_{cv} = 0.45$; $MDAE_{cv} = 0.11$ mg/m^3; $MDPD_{cv} = 42.48\%$, $R^2_{cv} = 0.45$; $MDAE_{cv} = 0.14$ mg/m^3; $MDPD_{cv} = 44.90\%$, $R^2_{cv} = 0.32$; $MDAE_{cv} = 0.04$ mg/m^3; $MDPD_{cv} = 49.56\%$, respectively) were comparable with the full-fit statistics, except for the dictyochophyte, cyanobacteria, and dinoflagellate groups, which showed considerably lower performance ($R^2_{cv} = 0.22$; $MDAE_{cv} = 0.09$ mg/m^3; $MDPD_{cv} = 53.10\%$, $R^2_{cv} = 0.20$; $MDAE_{cv} = 0.04$ mg/m^3; $MDPD_{cv} = 53.10\%$, $R^2_{cv} = 0$; $MDAE_{cv} = 0.09$ mg/m^3; $MDPD_{cv} = 58.94\%$, respectively). Finally, the highest $MDPD_{cv}$ was obtained for raphidophytes ($MDPD_{cv} = 71.87\%$), followed by diatoms ($MDPD_{cv} = 68.10\%$), and the lowest $MDPD_{cv}$ was recorded for TChla ($MDPD_{cv} = 29.70\%$).

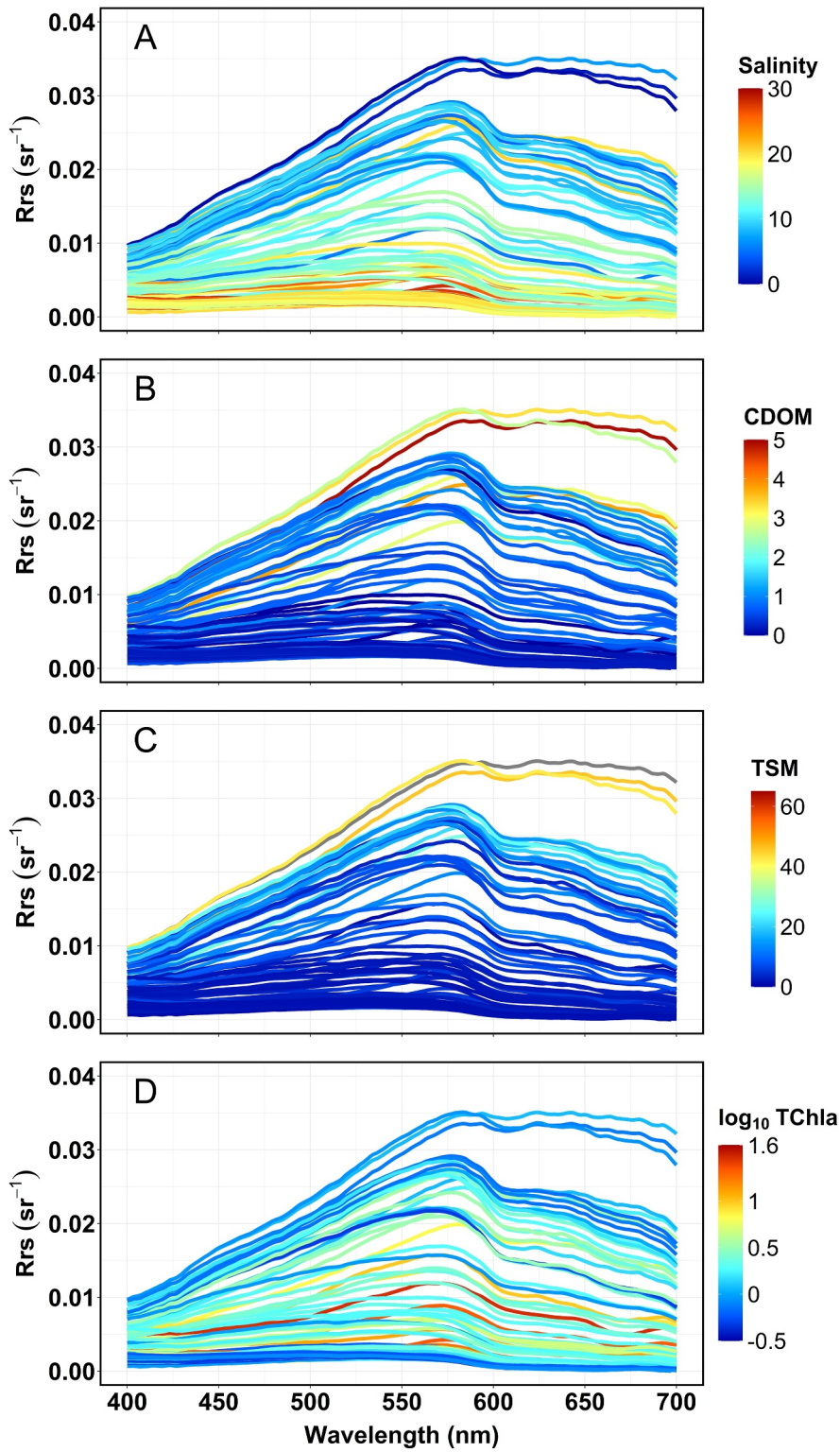

Figure 4. Hyperspectral remote sensing reflectance (Rrs; sr^{-1}) across the visible spectrum (400–700 nm) with each spectrum colored by the corresponding (a) sea surface salinity (SSS, PSU), (b) absorption by colored dissolved organic matter at 443 nm (aCDOM 443; (m^{-1})), (c) total suspended matter concentration (TSM; mg/L), and (d) total chlorophyll-a concentration (TChla; mg/m^3, shown on a $\log_{10}$ scale for consistency).

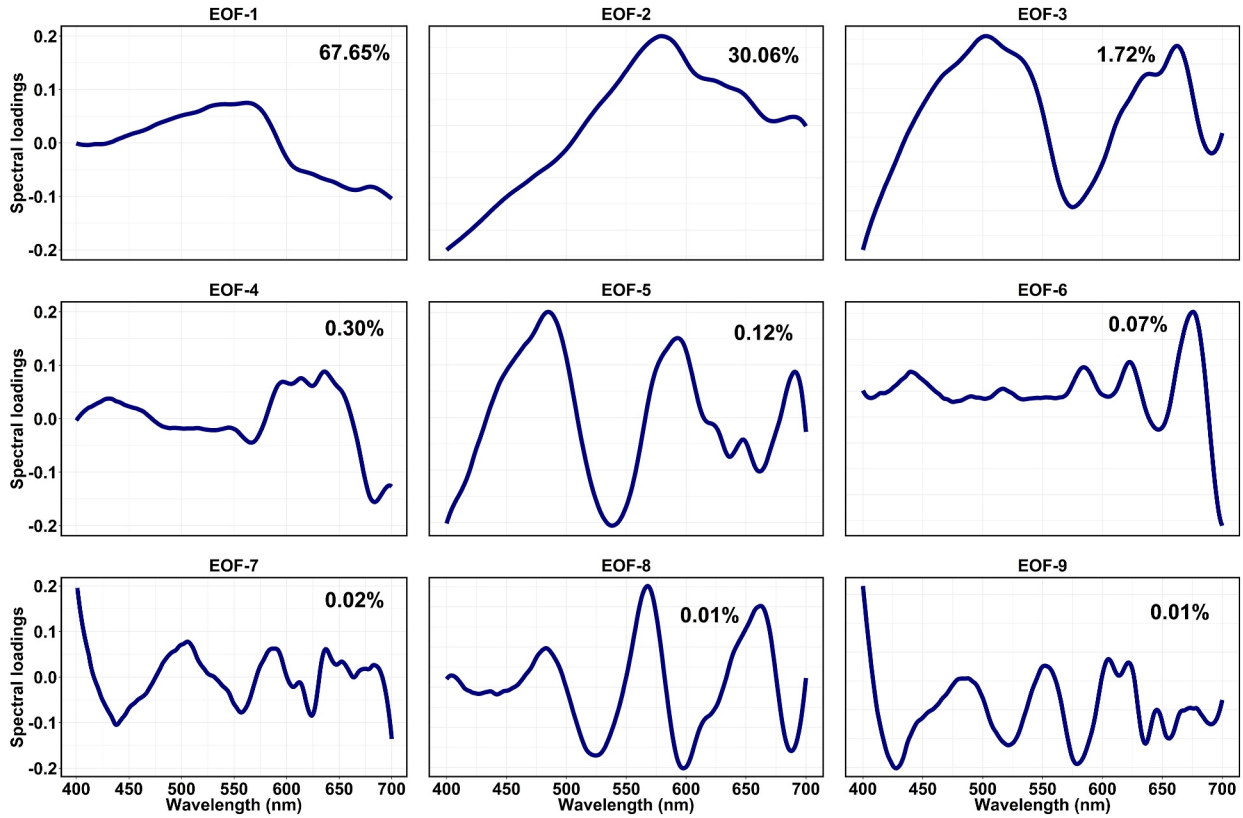

Figure 5. Distribution of the spectral loadings for the nine EOF modes derived from SAS-ST and SAS Rrs(λ) using the singular value decomposition. These spectral loadings were obtained from the matchup between SAS-ST, SAS Rrs(λ), and the total chlorophyll-a concentration (TChla; mg/m^3).

As an independent evaluation, we intercompared our HPLC and EOF-derived TChla estimates with flow-through ferryBox CHL_F measurements (Figure 7). Both estimates showed strong and statistically significant correlations with CHL_F, with R^2 values of 0.62 ($p \leq 0.0001$) and 0.70 ($p \leq 0.0001$) corresponding to the HPLC and EOF-derived TChla estimates, respectively. However, CHL_F was found to significantly underestimate HPLC-derived TChla ($p < 0.001$, BIAS $= -28.30\%$; Figures 7a and 7c). The MAE for this regression (Figure 7a) was 0.82 mg/m^3, and the MDPD was 44%.

Table 2
ΔAIC Derived From the EOF Analysis on the Matchup Data Set Between SAS-ST, the Total Chlorophyll-a (TChla; mg/m^3) and the Eight CHEMTAX Phytoplankton Groups

SAS-ST, SAS Rrs	EOF-1	EOF-2	EOF-3	EOF-4	EOF-5	EOF-6	EOF-7	EOF-8	EOF-9
TChla	3.67	24.46	**145.07**	NA	2.17	13.04	13.66	NA	6.28
Diatom	5.32	9.77	**44.24**	0.37	NA	NA	15.22	4.26	22.37
Dictyo	2.23	3.35	**16.94**	4.45	NA	NA	0.96	1.93	NA
Raphido	NA	12.62	**40.04**	NA	6.84	4.09	9.80	NA	NA
Hapto	6.23	13.75	NA	4.11	**18.25**	12.14	1.37	8.02	9.23
Prasino	NA	NA	0.85	10.20	**31.40**	NA	18.50	8.33	29.07
Crypto	NA	23.56	**34.29**	3.73	9.59	NA	10.91	8.68	NA
Dino	1.53	NA	2.10	NA	NA	NA	NA	**13.05**	1.50
Cyano	6.23	8.02	NA	NA	1.12	13.23	11.82	**14.38**	NA

Note. The bold numbers indicate the EOF that emerged as the most important predictor for each phytoplankton group.

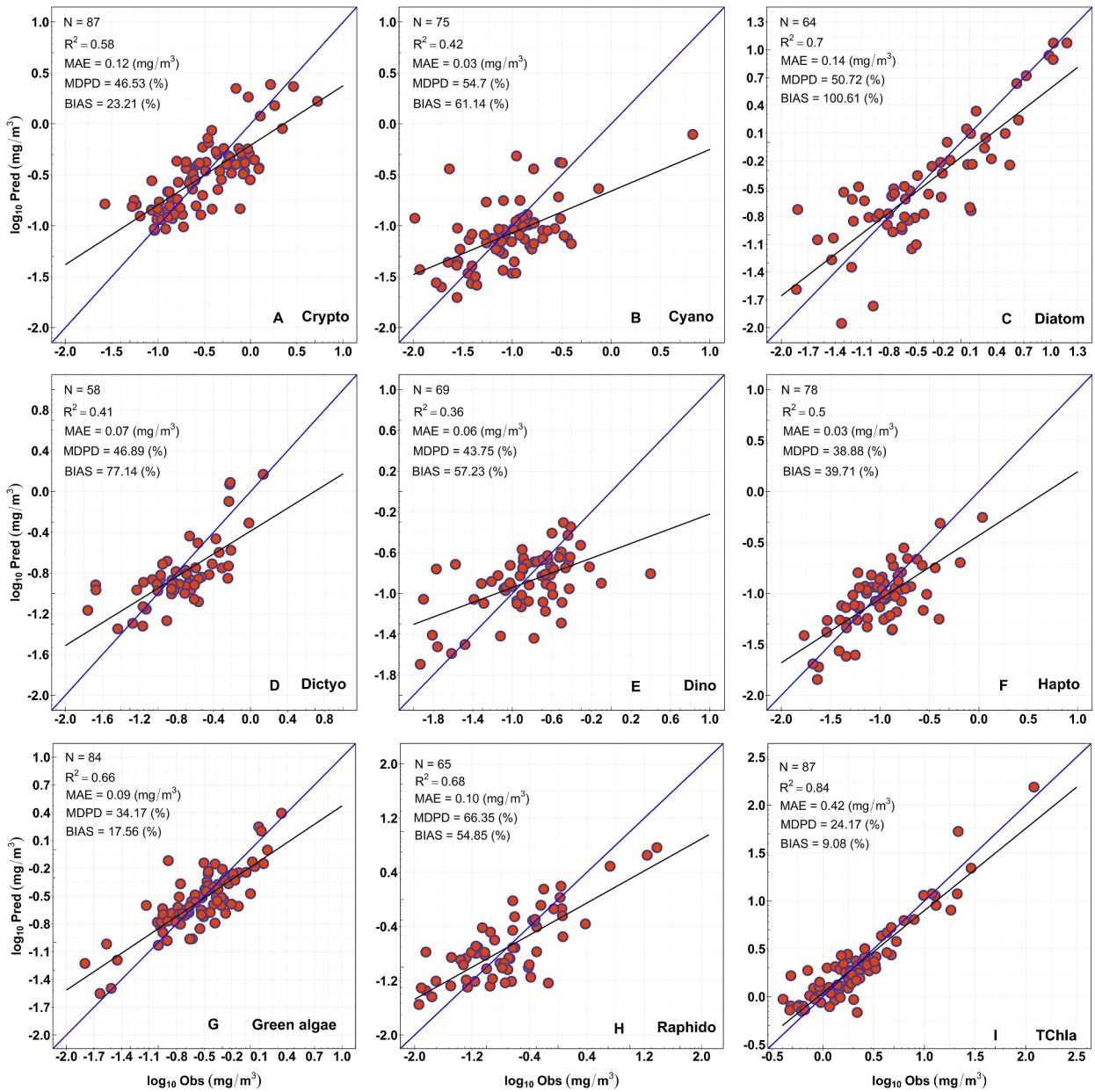

Figure 6. Regression between CHEMTAX-derived (*x*-axis) and the corresponding empirical orthogonal function (EOF)-predicted (*y*-axis) phytoplankton group chlorophyll-a (mg/m^3) concentration for cryptophytes, cyanobacteria, diatoms, dictyochophytes, dinoflagellates, haptophytes, green algae, raphidophytes, and total chlorophyll-a (TChla; mg/m^3, shown on a $\log_{10}$ scale for consistency) concentration.

3.4. Application of EOF-Based Algorithm to SAS-ST Rrs Data

After the model evaluation (i.e., full fit and cross-validation), we applied the EOF-based algorithm to independent SAS-ST Rrs(λ) data to elucidate highly resolved spatial-temporal dynamics of phytoplankton composition along the ferry track (Figures 8 and 9). Based on the statistical performance, the algorithm was applied to TChla and five phytoplankton groups (diatoms, raphidophytes, haptophytes, green algae, and cryptophytes) (see Table 3). In March 2018, the seasonal spring diatom bloom was clearly observed across much of the western and central portions of the ferry track (Figure 8a), with the highest concentrations (20.00 mg/m^3) occurring on March 16th and the bloom declining by April 19th. During this bloom, TChla concentrations were low (≤1.39 mg/m^3) in the eastern portions of the track near the Fraser River plume. In addition, the first captured date of the spring bloom on March 16th showed increased concentrations of cryptophytes, green algae, and raphidophytes near the central

Table 3
Performance Metrics for Total Chlorophyll-a Concentration (TChla; mg/m^3) and the Eight Phytoplankton Groups, Including Full-Fit and Cross-Validation Statistics

	N	R^2	Slope	MDAE	MDPD	R^2cv	MDAEcv	MDPDcv	*p*-value
TChla	103	0.81	0.84	0.46	24.88	0.73	0.49	29.70	2.210^{-16}
Diatom	77	0.61	0.71	0.18	61.12	0.60	0.21	68.10	2.210^{-16}
Raphido	77	0.57	0.69	0.13	65.01	0.41	0.13	71.87	2.870^{-15}
GA	96	0.57	0.66	0.09	35.45	0.45	0.11	42.48	2.210^{-16}
Crypto	102	0.53	0.59	0.13	40.14	0.45	0.14	44.90	2.210^{-16}
Hapto	87	0.47	0.50	0.03	43.88	0.32	0.04	49.56	1.876^{-13}
Cyano	88	0.38	0.43	0.03	46.30	0.20	0.04	53.10	1.413^{-11}
Dictyo	66	0.36	0.42	0.07	47.71	0.22	0.09	54.18	8.160^{-7}
Dino	82	0.24	0.36	0.08	49.68	0	0.09	58.94	2.917^{-6}

Note. R^2, slope, and intercepts were computed on a logarithmic scale, while median absolute error (MDAE; mg/m^3) and median percentage difference (MDPD; %) were calculated on a linear scale.

portions of the transect (see also Figure 10a), but the contributions (avg 3.18 mg/m^3) of these groups to total concentration remained relatively low compared to that of diatoms. Haptophytes showed marginally increased concentrations during the bloom conditions. Following March, diatom concentrations showed the second highest peak on July 25th (20.00 mg/m^3, see Figure 8a).

Beyond diatoms, raphidophytes also showed high concentration events. Specifically, increased raphidophyte concentrations were observed directly offshore of the Fraser River on 24 July 2018 (15.14 mg/m^3) and much higher concentrations on 25 July 2018 (25 mg/m^3) (Figure 8b). Other flagellate groups, including cryptophytes, green algae, and haptophytes, also showed increased concentrations during this observed raphidophyte bloom. Outside of the spring diatom and raphidophyte blooms, the nonblooming flagellate groups tended to show their highest contributions to total concentration during the summer. Green algae (Figure 8e) and cryptophytes (Figure 8f), the groups after diatoms and raphidophytes contributing most to TChla, displayed similar trends, with elevated concentrations observed across much of the transect through June and transitioning toward the Fraser River in July. Also, sporadic peaks were observed, including the highest green algae and cryptophyte concentrations recorded on July 9th (3.12 mg/m^3, and 3.18 mg/m^3, respectively). Finally, haptophytes concentrations remained low (0.72 mg/m^3) throughout the year.

In 2019, the later start of the sampling period missed the spring bloom, and TChla and phytoplankton group concentrations were comparatively lower throughout the late spring and summer compared to 2018 (Figure 9). In general, diatoms (1.05 mg/m^3), green algae (0.73 mg/m^3), and cryptophytes (0.67 mg/m^3) were dominant during spring. Following spring, diatom contributions generally decreased, whereas cryptophytes and green algae contributions persisted throughout summer, but with more of a ubiquitous presence than in 2018. During summer, sporadic and localized biomass peaks were observed on June 7th and August 14th. On June 7th, raphidophytes, cryptophytes, and green algae dominated the community. In turn, August 14th showed increased contributions of diatoms, raphidophytes, and cryptophytes in the central portions of the transect. Interestingly, a spike in green algae (2.81 mg/m^3) was observed on July 27th near the Fraser River plume and was reflected in the TChla data (3.77 mg/m^3). Of the lower concentration groups, haptophytes tended to show their highest concentrations toward the end of August in the western and central portions of the transect.

3.5. EOF-Derived Phytoplankton Composition Transects and Variability Along Salinity Gradients

To highlight the above trends and showcase the potential of the high spatially resolved data, two transects of high-resolution phytoplankton group concentration, with accompanying continuous ferrybox SSS, are shown under contrasting conditions in Figure 10. During the spring bloom on 16 March 2018, diatoms showed high (>30.00 mg/m^3) but variable concentrations along much of the western portions of the transect, where SSS was relatively high and invariant (27 PSU, Figure 10a). Within this region, centered at −123.6°W, was a narrow but prominent peak in green algae with concentrations reaching blooming magnitudes upward of 10 mg/m^3. Moving

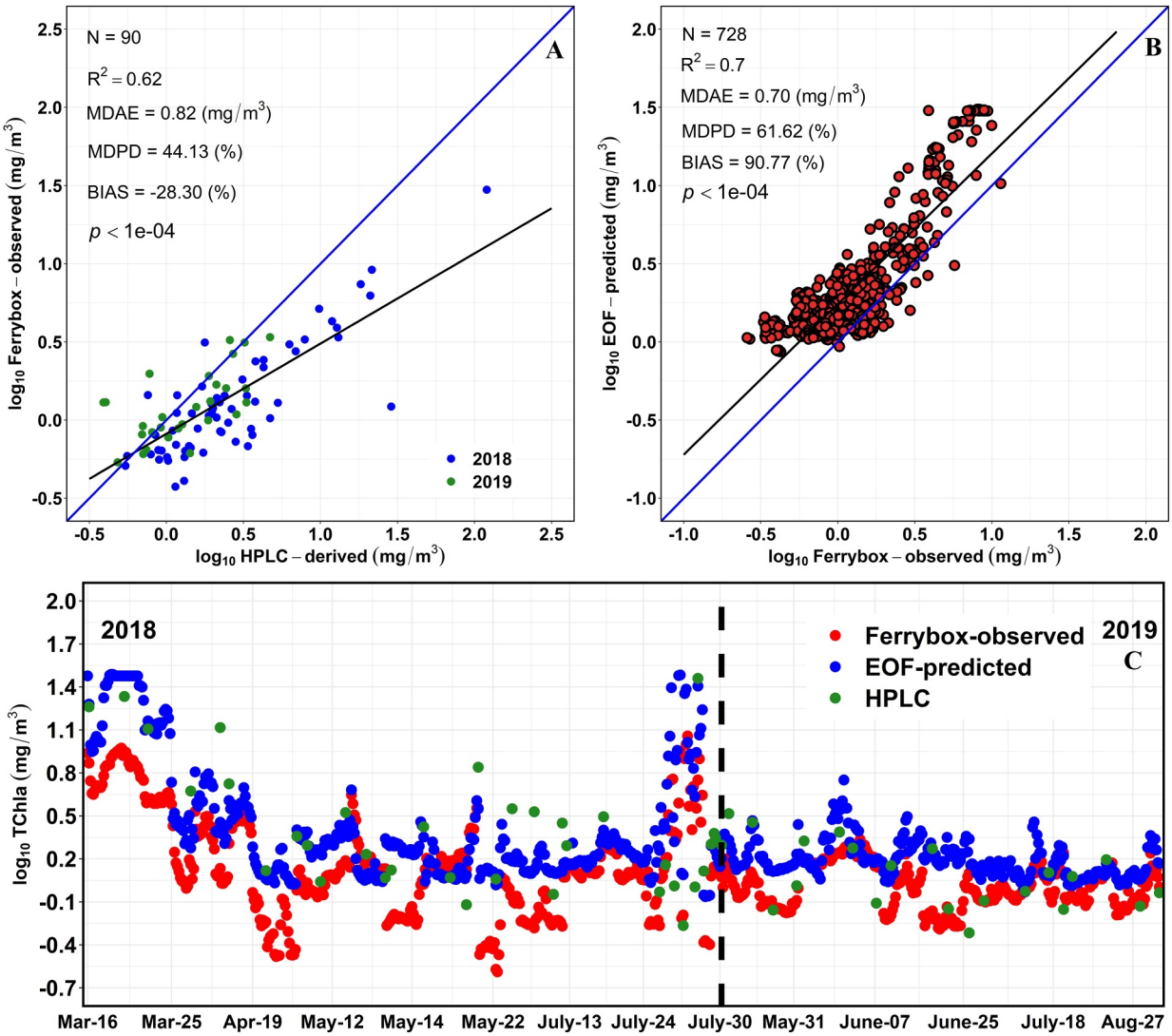

Figure 7. (a) Regression between HPLC-observed and the corresponding ferrybox-derived (*y*-axis) CHL_{FL} and (b) regression between ferrybox-observed and the corresponding empirical orthogonal function (EOF)-predicted (*y*-axis) total chlorophyll-a (TChla; mg/m^3, shown on a $\log_{10}$ scale for consistency) concentration. (c) Spatial-temporal variability of ferrybox CHL_{FL} and the corresponding EOF-predicted TChla concentration and HPLC TChla. For the 2018 and 2019 independent validation, samples span from March 16th to July 30th and from May 31st to August 27th, respectively, consistent with the model application depicted in Figures 8 and 9.

eastward, green algae contributions declined and diatoms regained dominance with a sharp decline in diatom contributions coinciding with a slight drop in SSS (25 PSU, −123°W), but concentrations remained at blooming magnitudes (approx. 10 mg/m^3). Further east, another sharp decline was observed in SSS (22.2 PSU) toward the eastern portion of the transect near the Fraser River and diatom concentrations dropped to low levels (0.97 mg/m^3). 14 August 2019 showed comparatively lower SSS across the transect, but similar to 2018, variability in phytoplankton concentration and compositions coincided with changes in salinity. In particular, the western portions of the transect displayed the highest SSS (23 PSU) and low phytoplankton concentrations. Over this part of the transect, moderate increases in diatom contributions were observed over a small area of decreased salinity (18.3 PSU), but then declined with increasing salinity moving eastwards. Toward the central portions of the transect, a sharp decline in SSS was observed, coinciding with increased phytoplankton concentration largely driven by raphidophytes, diatoms, cryptophytes, and green algae. Salinity then remained relatively invariant over a short portion of the transect, and, within this region, diatom and raphidophyte biomass peaked and then declined to their minimum and stayed relatively low moving eastward under low, but variable SSS.

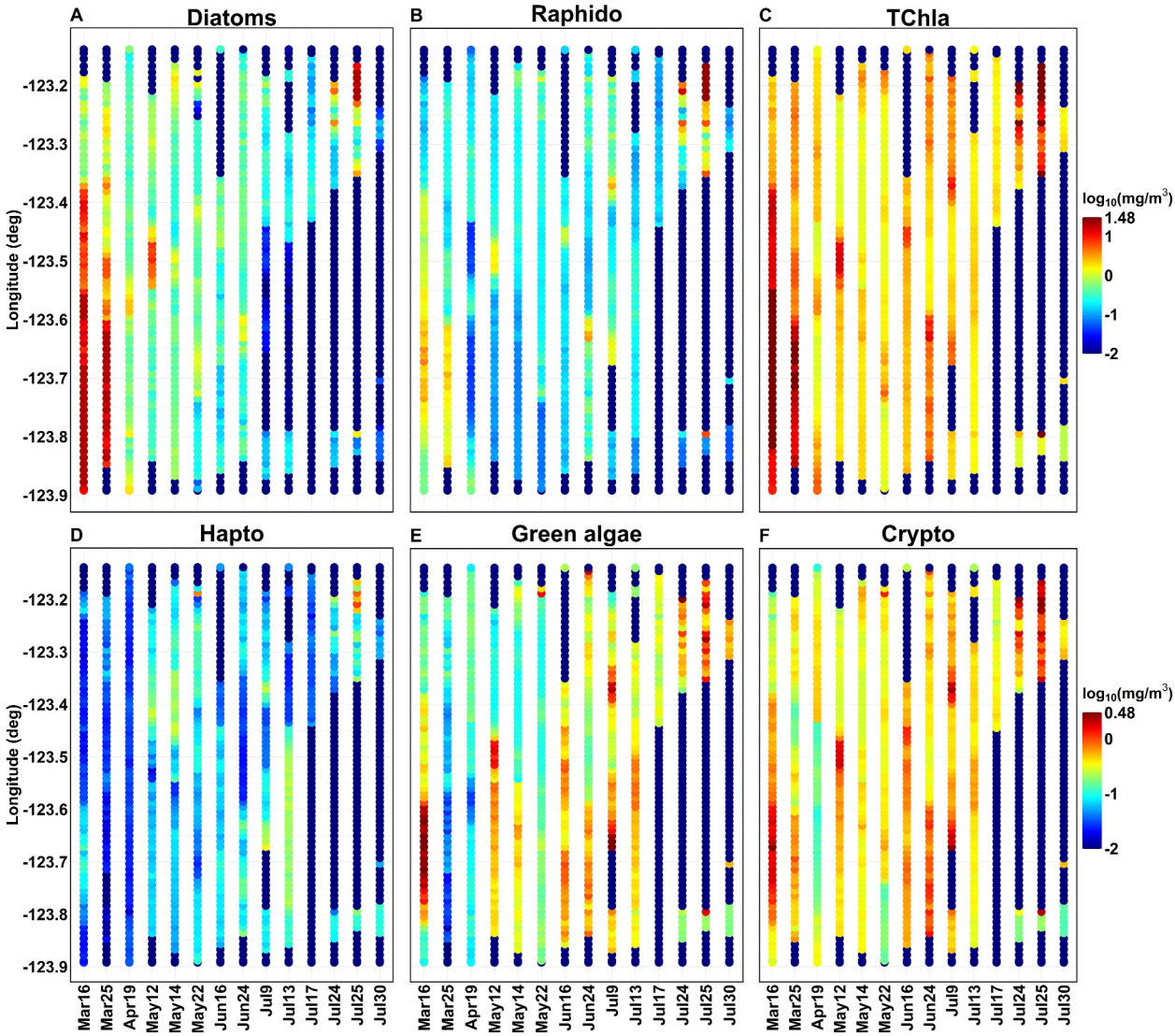

Figure 8. Spatial-temporal distribution of highly resolved 2018 phytoplankton composition along the ferry track derived from the EOF-based hyperspectral algorithm. The ferry track is shown in Figure 1.

4. Discussion

This study aimed to obtain high-resolution information on TChla and phytoplankton community composition within the SoG on the British Columbia coast. This was successfully achieved by adapting the EOF-based hyperspectral algorithm of Bracher et al. (2015) to the above-water hyperspectral radiometry data from Case-2 waters. Generally, the best retrievals were obtained for TChla, diatoms, raphidophytes, green algae, and cryptophytes, moderate retrievals for haptophytes, and finally poor performance for dictyochophytes, cyanobacteria, and dinoflagellates. Validation using an independent CHL_{FL} data set showed statistically significant retrievals with low error and a good correlation ($R^2 = 0.70$; $N = 728$, $p \leq 0.0001$) for TChla. Furthermore, the EOF-derived phytoplankton composition spatial maps showed expected regional trends, demonstrating that diatoms dominated the spring bloom, particularly in the western portions of the central SoG and notably in 2018. Following the spring bloom, diatom concentrations tapered off, followed by flagellate-dominated summer conditions and large sporadic raphidophyte blooms associated with the Fraser River plume. The subsequent sections present a comprehensive analysis of our key findings.

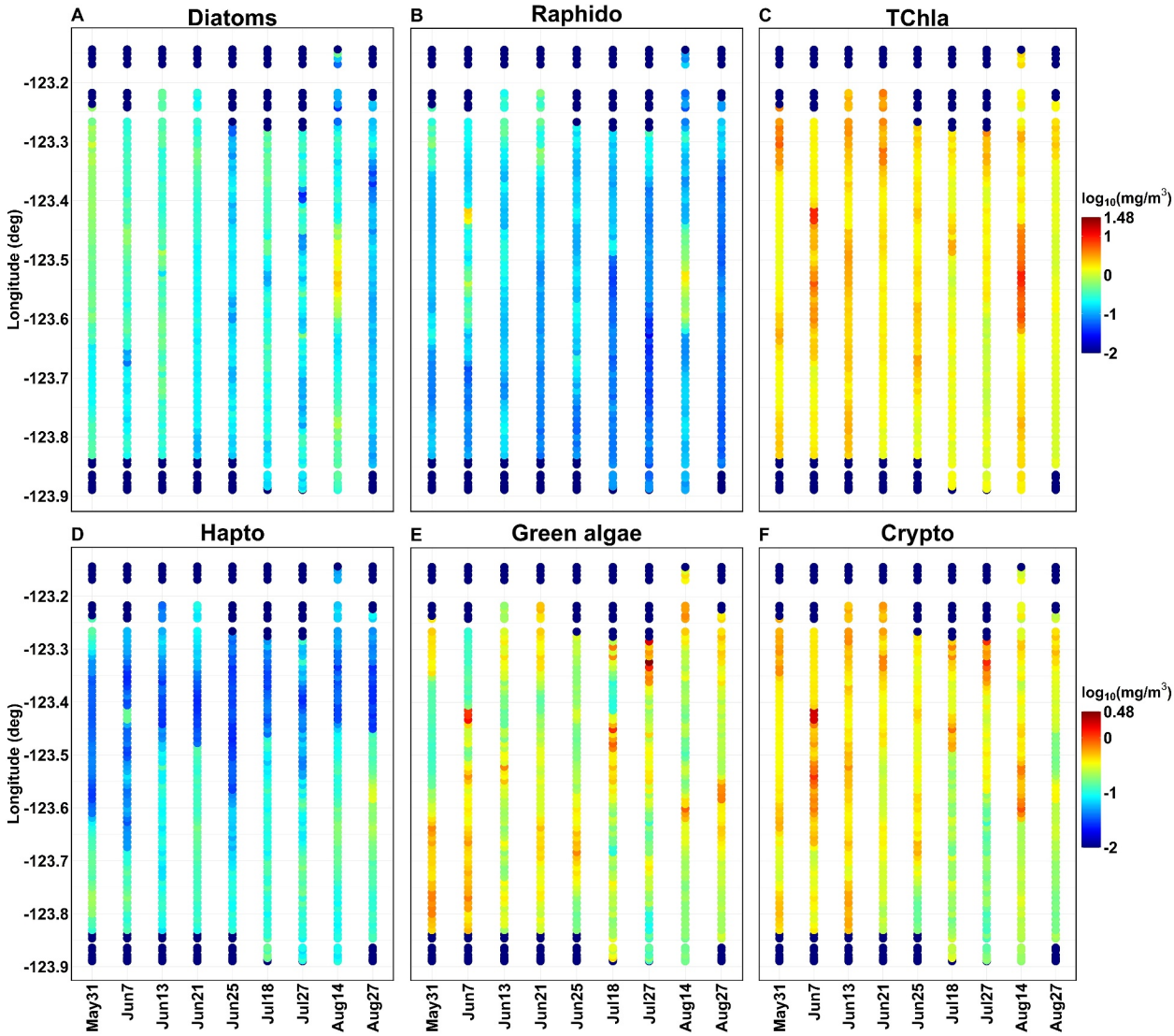

Figure 9. Spatial-temporal distribution of highly resolved 2019 phytoplankton composition along the ferry track derived from the EOF-based hyperspectral algorithm. The ferry track is shown in Figure 1.

4.1. EOF-Based Phytoplankton Group Retrievals From Hyperspectral Rrs

Considerably, more work has gone into deriving and evaluating TChla, as compared to phytoplankton functional types (hereafter PFTs), from multispectral data and there are multiple regional studies focused on TChla that can be compared with the results of this study. Here, the continuous EOF-derived TChla and the ferrybox CHL_{FL} showed strong comparability mirroring spatial-temporal trends across the transects, but with high bias (90.77%) and MDPD (61.62%) likely representing well-documented issues with fluorescence-based data (Roesler et al., 2017). TChla retrievals were generally better than those derived from regional satellite-based models and showed high comparability with the in situ HPLC-derived training data (R^2_{cv} = 0.73) and only a moderate overestimation (9%). In comparison, a regional study using the OC3M algorithm to derive TChla from MODIS-Aqua imagery showed the most comparable results to our hyperspectral retrieval (R_{log} = 0.81, MAD = 64%, MRD = 18%) (Carswell et al., 2017). Yet, a subsequent study using the same algorithm demonstrated considerably poorer agreement (R^2 = 0.47, RMSE = 0.23 $\log_{10}$ mg/m^3) (Hilborn & Costa, 2018). Recent retrievals of TChla from Sentinel-3 OLCI imagery using the POLYMER algorithm for the entire coast of British Columbia produced inferior statistics (R = 0.51, Bias = 0.99, MDPD = 1.82) (Giannini et al., 2021). Pramlall et al. (2023), using coarse resolution (4 km) combined and interpolated data from different ocean color satellite missions, also

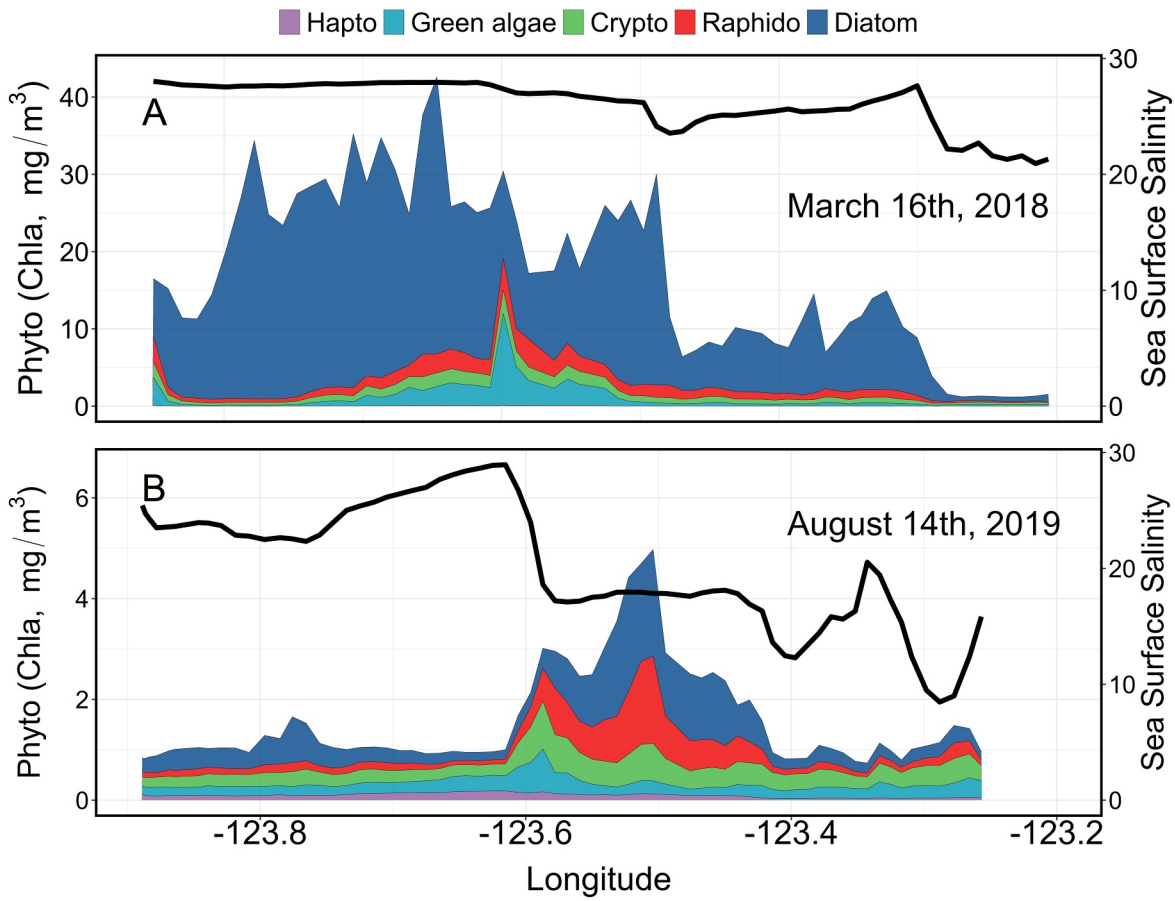

Figure 10. Transects of high spatial resolution EOF-derived phytoplankton composition for spring bloom conditions (16 March 2018) and summer raphidophyte bloom conditions (14 August 2019). For 16 March 2018, the corresponding continuous ferrybox sea surface salinity (SSS; PSU) is displayed on the right *y*-axis in (a). Comparable plots are shown for 14 August 2019 in (b).

achieved inferior results (Ocean Color Climate Change Initiative (OC-CCI) product—R^2 = 0.62, Bias = 0.78, MDAD = 1.44, RMSE = 0.26; GlobColor interpolated product—R^2 = 0.47, Bias = 0.96, MDAD = 1.57, RMSE = 0.30). Unsurprisingly, these results highlight that above-water hyperspectral radiometry data and its comparatively higher spatial resolutions and signal-to-noise ratios considerably improved the retrieval of TChla when compared to multispectral satellite retrievals. These factors will be discussed further following the comparison of the PFT results with other studies.

In contrast to TChla, fewer studies have been done attempting to characterize PFTs from both multi and hyperspectral data using the EOF approach (e.g., Bracher et al., 2020; Konik et al., 2024; Xi et al., 2020, 2021; Vishnu et al., 2022). The results of this study add to a growing field of research and demonstrate the potential of the EOF approach to derive phytoplankton group composition from hyperspectral Rrs(λ) under dynamic optically complex coastal conditions. Here, we provide a comparison of existing studies using comparable methods on both multi and hyperspectral data. To facilitate easier interpretation, the characteristics of these studies and their main findings and validation statistics are summarized in Table 4.

Our results were in line with several studies that have used a multispectral-based EOF approach to derive phytoplankton cell abundance (Lange et al., 2020), accessory pigment concentration (Bracher et al., 2015), and, Chla concentration of PFTs from both global ocean (Xi et al., 2020, 2021), Case-1 waters of the subarctic Pacific (Konik et al., 2024) and coastal waters of the SoG (Vishnu et al., 2022). In open ocean Case-1 waters, the results of Xi et al. (2020) were most comparable to those of this study providing PFT Chla outputs and showing good performance for diatoms, dinoflagellates, haptophytes, and green algae. More recently, these authors significantly improved these statistics (averages R^2cv = 0.64, MDPDcv = 52%) by including sea surface temperature (SST) in

Table 4
Comparison of Phytoplankton Group Statistics Retrieved Using EOF-Based Algorithms Built on Multispectral and Hyperspectral Reflectance Data

Multi/Hyper	Study	Water type	Phytoplankton groups				
			Diatoms	Dinoflagellates	Haptophytes	Green algae	Cyanobacteria
Multispectral	Vishnu et al. (2022)	Case-2	R^2cv = 0.30, MDPDcv = 65.77%	R^2cv = 0, MDPDcv = 63.25%	R^2cv = 0.14, MDPDcv = 47.13%	R^2cv = 0.08, MDPDcv = 49.14%	R^2cv = 0.08, MDPDcv = 48.51%
Multispectral	Xi et al. (2021)	Case-1	R^2cv = 0.76, MDPDcv = 57.71%	R^2cv = 0.60, MDPDcv = 54.68%	R^2cv = 0.69, MDPDcv = 43.63%	R^2cv = 0.51, MDPDcv = 52.96%	R^2cv = 0.38, MDPDcv = 45.44%
Multispectral	Xi et al. (2020)	Case-1	R^2cv = 0.65, MDPDcv = 74.74%	R^2cv = 0.62, MDPDcv = 57.29%	R^2cv = 0.64, MDPDcv = 48.62%	R^2cv = 0.51, MDPDcv = 56.26%	R^2cv = 0.15, MDPDcv = 55.08%
Hyperspectral	Bracher et al. (2020)	Case-1	R^2cv = 0.68, MDPDcv = 53%	R^2cv = 0.59, MDPDcv = 43%	R^2cv = 0.57, MDPDcv = 44%	R^2cv = 0.55, MDPDcv = 48%	R^2cv = 0.54, MDPDcv = 29%
Hyperspectral	Current Study	Case-2	R^2cv = 0.60, MDPDcv = 68.10%	R^2cv = 0 MDPDcv = 58.94%	R^2cv = 0.32, MDPDcv = 49.56%	R^2cv = 0.45, MDPDcv = 42.48%	R^2cv = 0.20, MDPDcv = 53.10%

their regression model (Xi et al., 2021). Comparatively, few studies have been done on Case-2 waters. Uniquely, Vishnu et al. (2022) used multispectral Sentinel-3A OLCI imagery for the same coastal SoG waters studied here during the spring and summer of 2018, which allowed us to evaluate the retrieval of phytoplankton groups from both multispectral and hyperspectral Rrs(λ) data as inputs to the EOF method. This multispectral-based analysis showed statistically significant retrievals for only diatoms and raphidophytes, which were prominent high biomass bloom-forming groups (see Table 3 in Vishnu et al., 2022). This limited number of groups was considerably less than the hyperspectral results of this study, which showed highly significant ($p < 0.001$) retrievals for all groups and the best performance for all of the dominant groups: diatoms, raphidophytes, green algae, and cryptophytes, which generally constituted 73% of the total phytoplankton community (see Table 3). However, similar to Vishnu et al. (2022), haptophytes, dictyochophytes, cyanobacteria, and dinoflagellates showed moderate to poor performance, which is in contrast to other hyperspectral-based studies. For example, Bracher et al. (2020) used the most similar EOF-based methods; however, they were applied to underwater hyperspectral apparent optical properties (AOP) obtained from the Case-1 Atlantic waters. Their results produced statistically significant retrievals for diatoms, haptophytes, cyanobacteria, green algae, and dinoflagellates (Table 4). The retrievals from our study were comparable for diatoms and green algae, but lower for haptophytes and poorer for cyanobacteria, and dinoflagellates (Table 4). There are multiple factors likely driving the poor performance of these groupings in this study.

Likely of most importance, this study was performed in highly Case-2 waters, and under these conditions, uncertainties arise from the high contributions of other optical constituents such as CDOM, nonalgal particle (NAP) absorption, and inorganic suspended sediments (Craig et al., 2012; Mouw & Yoder, 2010). These nonalgal constituents are introduced via river discharge (e.g., here, the Fraser River is a major driver of bio-optical variability (Loos & Costa, 2010)), bottom resuspension, and tidal mixing and their variability is often independent of TChla (IOCCG, 2014; Mouw et al., 2017). The presence of these nonalgal constituents negatively impacts Rrs-based PFT/PSC retrievals (Craig et al., 2012; Mouw & Yoder, 2010) by masking subtle reflectance signatures from phytoplankton pigments. Combined, these constituents influence the entire visible spectrum with CDOM and detrital materials dominating in the blue and inorganic sediments driving high scattering in the red and near-infrared regions (IOCCG, 2000; Ruddick et al., 2006). The influence of these materials was evident within our reflectance signals and this result is in line with other studies done on Case-2 waters. For example, our study reported high variability in aCDOM 443 and TSM (see Table 1) across the SoG, largely defining the spectral shape of the hyperspectral Rrs(λ) (see Figure 4), and likely resulting in substantial contributions to the spectral loading of the first few EOF modes (see Figure 5). Vishnu et al. (2022), Phillips and Costa (2017), and Craig et al. (2012) showed comparable spectral shapes and EOF structures driven by high concentrations of TChla, TSM, and CDOM for comparable Case-2 waters. Specifically, EOF-1 (60.33% of the variability) and EOF-2 (38.33%) from Vishnu et al. (2022) on Sentinel-3 OLCI (but using the same input CHEMTAX data) exhibited a similar spectral structure to our EOF-1 (67.65% of the data variability) and EOF-2 (30.06% of the data variability). Additionally, EOF-2 from Craig et al. (2012) showed a similar spectral structure to our EOF-1, with peaks at 560 and 680 nm (TChla fluorescence), while their EOF-1 was comparable to our EOF-2. Furthermore,

our EOF-2 showed a similar spectral feature to the EOF-1 from Phillips and Costa (2017) (94.5% of the total variance). In contrast, the first EOFs derived by multiple studies from Case-1 waters (e.g., Bracher et al., 2015; Lange et al., 2020; Xi et al., 2020) were dissimilar to those found in this study with spectral shapes primarily defined by phytoplankton pigments and seawater absorption. Consequently, in this study, the high observed scattering from inorganic suspended sediments, combined with the strong absorption by CDOM, likely masked signals from the less prominent phytoplankton groups, making their retrieval more difficult.

Under the observed Case-2 conditions, the groups with the lowest contributions showed the poorest performance. Specifically, poor retrievals were observed for CHEMTAX-derived haptophytes (7%), cyanobacteria (~6% contribution), dictyochophytes (~6%), and dinoflagellates (~8%; Figure S1 in Supporting Information S1). In the Salish Sea, these groups rarely represent high contributions to total phytoplankton biomass (Del Bel Belluz et al., 2021; Nemcek et al., 2023) even during low-moderate TChla flagellate dominated conditions. Consequently, due to the masking effects described above, the low spectral signals from these groups were likely difficult to separate from the dominant contributions of other phytoplankton groups and nonalgal optical constituents (Mouw et al., 2017). In contrast, under Case-1 conditions, these groups often form more prominent components of phytoplankton communities. For example, across the Atlantic Ocean, Bracher et al. (2020) showed cyanobacteria and haptophyte contributions as high as 32% and 42% of TChla (see Table 5 of Bracher et al., 2020), respectively, and having a greater dynamic range than observed in this study; however, their dinoflagellate contributions were comparable to this study despite showing stronger retrievals. Consequently, it was likely that under Case-1 conditions, both the higher contributions of certain groups and the reduced influence of nonalgal optical constituents led to better retrievals for both prominent and nonprominent groups. Nonetheless, it should be highlighted that when compared to the multispectral Vishnu et al. (2022) results, the hyperspectral methods deployed here allowed for good retrievals of the dominant groups under lower biomass conditions (i.e., cryptophytes and green algae) rather than just high biomass blooming groups (i.e., diatoms and raphidophytes), demonstrating a valuable improvement when using above-water hyperspectral data.

The high spectral resolution of the hyperspectral data used in this study was likely the main factor driving the improved results when compared to multispectral methods. Under high optical complexity, increased spectral resolution can positively influence the number of groups retrieved from the Rrs(λ). For example, earlier studies have shown that hyperspectral data with higher spectral resolution (i.e., 5 nm or better) produce improved phytoplankton composition retrievals (e.g., Wolanin et al., 2016; Xi et al., 2017), which was further illustrated through comparison with Vishnu et al. (2022). This improvement occurs despite recent findings showing that on a global scale, phytoplankton absorption spectra and hyperspectral Rrs(λ) typically contain only ~5–6 degrees of freedom, representing the number of independent spectral components available to resolve phytoplankton pigments and relate them to the phytoplankton community structure (Cael et al., 2020; Kramer et al., 2022). This strong performance makes hyperspectral data especially pertinent under certain conditions in Case-2 waters, where CDOM, NAP, and TSM dominated the optical signal, thereby reducing the discernibility of phytoplankton pigment signatures from other optical constituents. The promising results of this study highlight that hyperspectral satellites such as PACE, that have even higher spectral resolutions than the instruments used here (e.g., PACE OCI has a 5 nm spectral resolution compared to the 10 nm in this study) could considerably improve spaceborne PFT retrieval in Case-2 waters when compared to multispectral data (Werdell et al., 2019). Specifically, it is expected that there will be a significant improvement when using hyperspectral Rrs(λ) with 5 nm spectral resolution for deriving phytoplankton pigments (Kramer et al., 2022) or phytoplankton groups (Wolanin et al., 2016) with decreased performance at coarser spectral resolutions. Nonetheless, there are other considerations, such as the optical complexity of the region, likely contributed to the strong PFT retrievals observed in this study and may hinder hyperspectral satellite-based retrievals.

In addition to spectral resolution, the high spatial resolution and close temporal matchups of the validation data used in this study likely considerably improved results when compared to lower spatial resolution satellite-based multispectral data (and presumably future satellite-based hyperspectral data). In this study, the in situ hyperspectral Rrs(λ) data were continuously collected but were interpolated to 1-min intervals along the ferry track, ensuring relatively close matchups between radiometry and discrete water samples used to develop the hyperspectral algorithms. In contrast, EOF algorithms developed with multispectral data (Vishnu et al., 2022; Xi et al., 2020) generally utilize imagery obtained from $\pm$3 hr of discrete water samples and use a median Rrs(λ) value from a 3 $\times$ 3 pixel window (for reference, the Sentinel-3A spatial resolution is 300 m per pixel). This spatial-temporal averaging of satellite data often adds a considerable error through pixel mixing, where high variability

in IOPs and PFTs and their spectral signatures are averaged into a single value and are effectively diluted. This influence is pronounced in highly dynamic coastal waters such as SoG, where the Fraser River plume can introduce high variability in IOPs over short distances (Loos & Costa, 2010) and PFTs can vary greatly over space and time (e.g., results of this study, Figures 8 and 10). Specifically, a recent study by Nasiha et al. (2022) showed that in situ hyperspectral Rrs(λ) varies significantly within the 300-m pixel of the Sentinel-3A OLCI, with the most marked variability occurring in the frontal region of the Fraser River plume (also shown here in Figure 10). Yet, it should be noted that the 1-min averaging performed in this study would have also resulted in averaging the subtle differences in PFT changes. This averaging would have added error to our model development and validation statistics, notably for the low contribution groups displaying subtle signals, and could further explain the better retrievals observed in more homogenous Case-1 waters. Nonetheless, the finer spatial resolution of our in situ hyperspectral Rrs(λ) data allowed for more precise quantification of phytoplankton communities, which is less achievable with the broader, spatially averaged Sentinel-3 OLCI data and likely even those retrievable via the PACE satellite.

Finally, beyond optical complexity and spectral resolution, the quality and quantity of matchup points may play an essential role in defining the accuracy of phytoplankton group retrievals. Bracher et al. (2015) observed that approximately 45–50 matchup points are necessary to construct an effective regression model to predict the concentration of accessory pigments from multispectral data. Furthermore, Xi et al. (2020) followed the same method and employed 52–394 training points and, more recently, used an even higher number of points (483) (Xi et al., 2021). These high numbers of matchup points resulted in the highly successful derivation of PFTs from global ocean Case-1 waters. In comparison, for the Case-2 waters studied here, Vishnu et al. (2022) used the same EOF-based methods and employed 61–102 samples to retrieve PFTs from Sentinel-3 OLCI data. In this study, the total number of matchup points (103) was within the range of matchup points suggested by the above authors; however, given the optically complex nature of our region, more samples may have improved the accuracy of our retrievals, especially for lower concentration groups that showed a low dynamic range. Dinoflagellates and dictyochophytes periodically form large blooms and haptophytes and cyanobacteria can sometimes constitute more dominant proportions of the phytoplankton community (Del Bel Belluz et al., 2021; Nemcek et al., 2023), and the inclusion of these points in the model may have helped to improve the retrievals of these generally low contribution groups.

4.2. Seasonal and Spatial Patterns in Phytoplankton Community Composition

In this study, autonomous ferry-based hyperspectral radiometry provided highly resolved phytoplankton composition data that followed expected regional trends and provided unique spatial insights not available from other methods. Specifically, in spring 2018 (March), the hyperspectral retrievals showed blooming diatom concentrations (see Figure 8) that were corroborated by the ferrybox-derived CHL_{FL} ($\sim$10 mg/m^3). This observation is in line with spring bloom conditions commonly observed across temperate coastal regions including the SoG (e.g., Allen & Wolfe, 2013; Del Bel Belluz et al., 2021; Gianini et al., 2021; Marchese et al., 2022; Nemcek et al., 2023; Pramlall et al., 2024). Specifically, the mid-March onset of the spring bloom is within the SoG late February through the mid-April window defined by Collins et al. (2009), close to the MODIS Aqua 2003–2016 time-series mean initiation date defined by Suchy et al. (2022), and was comparable to the 20 March 2018 initiation date observed by Del Bel Belluz et al. (2021). In contrast, the 2019 sampling period was outside the spring bloom window for the SoG explaining why spring bloom conditions were not observed in this year.

Spatially, our data showed that the 2018 spring bloom was concentrated in the higher salinity waters along the western portions of the transect likely least affected by the Fraser River, which is typical and has been observed by other authors including those using CHL_{FL} along the same transect (Halverson & Pawlowicz, 2013; Suchy et al., 2019). Compositionally, the high dominance of diatoms under the observed 2018 spring bloom conditions follows expected trends both within the SoG and more broadly in temperate regions (Carstensen et al., 2015; Del Bel Belluz et al., 2021). Unlike other studies, the high spatial resolution spring bloom transect from 16 March 2018 (Figure 10a) showed a narrow band of high green algae biomass (appox. 10 mg/m^3), which is a novel observation as species within this group do not commonly form blooms, especially of this magnitude, in the SoG. Due to the uniqueness of this observation, this result should be taken with caution despite green algae having a pigment unique to only this group (Chlb), showing relatively strong validation statistics and capturing seasonal trends outside of this event. Considering its spatial and temporal transience (it was gone by 25 March 2018), it is possible that similar blooms have been regularly missed by less resolved methods such as discrete ship-based sampling. During spring, Chla is highly variable along this transect (Halverson & Pawlowicz, 2013), and if

coupled with more comprehensive physical and biogeochemical data, the models derived here could provide valuable insights into the variability and drivers of both phytoplankton biomass and composition.

Following spring bloom conditions, phytoplankton communities in the SoG transition to somewhat persistent moderate biomass communities dominated by cryptophytes and green algae (Del Bel Belluz et al., 2021; Nemcek et al., 2023). The results of this study followed this expected succession with TChla concentrations declining and the EOF and CHEMTAX-based PFTs displaying high cryptophyte and green algae contributions. Typically, these groups become dominant with the onset of stratified nutrient-limiting surface conditions and increased zooplankton grazing (McEwan et al., 2023; Peña et al., 2016). Under these conditions, species from these groups have competitive advantages arising from their smaller cell size and corresponding high surface-to-volume ratios and mixotrophic capacities, making them better suited to warm, low-nutrient stratified surface waters when compared to diatoms (Margalef, 1978). Yet, there were departures from these characteristic moderate biomass conditions. Our EOF-based retrievals showed a strong raphidophyte bloom in the summer of 2018 and a weaker bloom in 2019, with both occurring in proximity to the low SSS waters of the Fraser River plume (Figure 8b, July 25th). The large 2018 bloom was also observed by multiple in situ studies showing high abundances of *Heterosigma akashiwo* adjacent to the Fraser River plume during the same period (Esenkulova et al., 2021; Vishnu et al., 2022). This species has been found to have an affinity for low salinity, highly stratified waters with high nutrient concentrations (Esenkulova et al., 2021). Interestingly, our high-resolution transect plot from 14 August 2019 (Figure 10) highlighted peaks in raphidophyte, but also diatom and cryptophyte concentrations, coinciding with a sharp decline in SSS moving eastward along the transect. These trends suggest the transect crossed the Fraser River plume, which is a frontal region that has been associated with some of the highest regional primary productivity due to increased nutrients (Peña et al., 2016). Without the high resolution methods developed here, this feature and its spatial progression would have been challenging to capture. *Heterosigma akashiwo* is one of the most prolific toxic algae globally (Hallegraeff et al., 2004) and, in BC waters, they are the leading cause of farmed salmon mortality and severely impact the survival of wild herring and salmon (Rensel, 2007). As a result, highly resolved data on *Heterosigma akashiwo* dynamics from hyperspectral radiometry data, either in situ or satellite-based, may improve the understanding of their spatial-temporal distribution and environmental drivers. This knowledge would help to mitigate the negative effects on aquaculture salmon and other commercially important fish species in this region.

4.3. Limitations of the Study

Uncertainties associated with the HPLC and CHEMTAX data are important to consider, as these data sets were used to train the EOF regression models. Analysis of duplicate HPLC samples (20% of total samples) showed a low CV (<20%), suggesting relatively high precision and low uncertainty for TChla and the pigments input into CHEMTAX analysis. However, CHEMTAX has inherent uncertainties, and in complex coastal waters, species with similar pigment profiles, mixotrophs with endosymbionts, or the inclusion of small uncharacterized species into broad groupings can lead to misclassification or error (Lewitus et al., 2005). This error can be difficult to assess as the quantitative validation of CHEMTAX outputs is challenging (Kramer & Siegel, 2019). New statistical programs such as phytoclass show promise in reducing group classification errors through improved convergence on optimal pigment ratios, but remain relatively untested with CHEMTAX comparisons done using synthetic data sets and limited field applications beyond the Southern Ocean (Hayward et al., 2023). Applicable to both CHEMTAX and phytoclass, recent analysis has shown the covariation of groups, which often happens in coastal waters, complicates their separation and retrieval via radiometry (Catlett & Siegel, 2018; Kramer et al., 2020). For example, in the SoG, green algae and cryptophytes tend to follow similar temporal trends (Del Bel Belluz et al., 2021). To address this issue, data-driven approaches have been developed to cluster covarying pigment-ratio based groupings into broad classifications such as "small flagellates" that, in the SoG, would group green algae and cryptophytes from the above example. Here, we decided to use CHEMTAX outputs as, despite its known issues, this program has been broadly applied and considerable regional efforts have been done to fine-tune analysis (e.g., Del Bel Belluz et al., 2021; Nemcek et al., 2023). Our results followed expected regional trends derived using other methods (MacNeil et al., 2024) providing confidence in our results even for covarying groups such as cryptophytes and green algae. Importantly, data-driven approaches (e.g., Kramer & Siegel, 2019) made it difficult to separate raphidophytes, which include toxic *Heterosigma akashiwo*, which has considerable detrimental impacts on fish and was important to characterize (Esenkulova et al., 2021). Further research should investigate new methods such as phytoclass, which may reduce errors when compared to CHEMTAX, and

continued validation of pigment-based methods, potentially using molecular approaches able to quantify small species, is necessary. Currently, no single "gold standard" method exists to quantify the full range of phytoplankton communities and each has inherent uncertainty.

Finally, uncertainties originating from the in situ hyperspectral Rrs(λ) must also be considered. Errors from spectral data obtained from above-water platforms occur from multiple sources, including instrument calibration, inherent sensor noise and ship shadow and superstructure influences. To address these issues, we followed rigorous quality control criteria, including factory precalibration of instruments and well-defined regional methods to remove ship-based influences (Nasiha et al., 2022; Wang & Costa, 2018). Yet, some of the largest sources of error for above-water radiometry come from environmental factors (Mobley, 1999) including wind speed, sun zenith angle, sky glint, and wave height. These sources can contribute up to 80% of the variability in Rrs(λ) across all wavelengths and this variability becomes worse under overcast conditions (J. Lin et al., 2022). Within these environmental factors, ρ_s contributes the most (50%) to Rrs(λ) variability (J. Lin et al., 2022). To limit this source of error, we performed our radiometric measurements only during solar noon under ideal atmospheric conditions (Nasiha et al., 2022) and, following Mobley (1999), we calculated the value of ρ_s using wind speed data, solar zenith angle, and viewing geometry. As a result, our final Rrs(λ) showed CV < 10% across the visible bands. In addition, we conducted a BRDF correction for our above-water Rrs(λ) that were tailored explicitly for Case-2 water types (Lee et al., 2011; Wang & Costa, 2022). While these strict protocols were deployed, the final level of uncertainty in our data was difficult to report directly. Studies performed with fiducial reference measurements (e.g., Alikas et al., 2020; Tilstone et al., 2020; Zibordi et al., 2012) have reported that, on average, Rrs(λ) uncertainties are most significant at 443 nm (within 3.5%), decline at 560 nm (1.0%), and increase at 665 nm (3.0%), with Ed measurements contributing to the highest level of variance (Tilstone et al., 2020). In addition to the above uncertainties, there are broader challenges associated with hyperspectral data that are important considerations for future work. For example, measurements taken at close wavelengths are highly correlated, and as a result, the effectiveness of the individual wavelengths is weakened (Cael et al., 2020). In conclusion, uncertainties are inherent to radiometric measurements, but we acquired data following restrictive protocols during optimal environmental conditions and conducted corrections specifically tailored for optically Case-2 water types, effectively limiting variability and errors as much as possible.

To address some of the above issues with hyperspectral radiometry, IOPs could be incorporated into future studies. Specifically, measurement of the phytoplankton spectral absorption coefficient (aph; m^{-1}) has also been used to infer accessory pigments through statistical and analytical techniques including the decomposition model (Bidigare et al., 1990), spectral derivative analysis (Bidigare et al., 1989), and Gaussian decomposition or absorption-based pigment concentration models (Chase et al., 2013; Hoepffner & Sathyendranath, 1991, 1993; X. Sun et al., 2022, D. Sun et al., 2025; Teng et al., 2022; Ye et al., 2019; Zhang et al., 2021). A key advantage of these methods is that aph is specific to phytoplankton pigments excluding the influence of other optical constituents that complicate Rrs(λ) retrievals (X. Sun et al., 2022). However, for operational applications, aph is often derived from radiometric data via bio-optical inversions, which introduce further uncertainties to the aph estimations, therefore hindering the capability of differentiation of phytoplankton groups (Xi et al., 2015). Despite these strengths, hyperspectral Rrs is the primary observable quantity in ocean color remote sensing and is readily retrievable from satellite sensors, making our algorithm directly applicable to current hyperspectral satellite missions, such as PACE, PRISMA, and EnMAP. Inclusion of both apparent optical properties (AOPs) and IOPs could provide synergistic outputs, providing additional validation of the models derived here.

5. Conclusion and Outlook

This study contributes to the small amount of research conducted using hyperspectral Rrs(λ) to derive the Chla concentration of phytoplankton community composition under optically complex Case-2 waters. Here, highly resolved hyperspectral Rrs(λ) data from an above-water radiometer installed on a ship of opportunity (QoA ferry) enabled highly resolved phytoplankton community composition across the SoG. Our results demonstrated a substantial improvement over satellite-based multispectral retrievals, with statistically significant retrievals for most phytoplankton groups. The full-fit model evaluation demonstrated the best retrievals for diatoms, cryptophytes, green algae, raphidophytes, and TChla concentrations. Moderate retrievals were observed for haptophytes, while dictyochophytes, dinoflagellates, and cyanobacteria showed poor retrievals. The reduced performance of these groups was likely a result of their low dynamic range and small contribution toward the total phytoplankton community under Case-2 conditions that would have masked their optical signatures.

The highly resolved phytoplankton community composition outputs retrieved from the EOF-based algorithm agreed well with regional trends highlighting spring diatom blooms and the transition to moderate biomass flagellate-dominated summer conditions. Spatially, diatoms tended to be more prevalent in the more oceanic western portions of the transect that had higher SSS and that were least influenced by the Fraser River, which follows findings from other studies using underway fluorescence and satellite-based methods. Uniquely, we also observed a narrow spatial band of very high green algae contributions, with this group rarely showing bloom conditions in the SoG and more broadly in other regions. Furthermore, the results of this study captured large harmful raphidophyte blooms (*Heterosigma akashiwo*), which showed expected trends being observed adjacent to the Fraser River plume. Specifically, in 2018, strong raphidophyte bloom conditions were observed, alongside diatoms, at the front region with the Fraser River plume with this feature often associated with high phytoplankton biomass. Beyond these general trends, this study has shown the value of having highly resolved phytoplankton community composition data (Figure 10), capturing features in PFTs that would not have been captured using discrete samples.

The results of this study are promising and further development of the methods along with additional sampling measures such as water column stratification, nutrient conditions, and zooplankton could provide unique insights, not achievable via discrete sampling, into the drivers of regional phytoplankton dynamics and links with the Fraser River plume. This insight would provide valuable knowledge for ecosystem-based applications, including water quality and harmful algal bloom monitoring, export production, and fisheries management. The findings of this study have important implications for the recently launched NASA PACE mission, which carries OCI with 5 nm spectral resolution, offering the potential for synoptic-scale monitoring and assessment of phytoplankton communities.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Acknowledgments

The authors wish to express their gratitude to the members of the Spectral Lab for their invaluable assistance during fieldwork. We extend our appreciation to BC Ferries Corporation for granting access and providing technical support. We thank Ocean Network Canada (ONC) for facilitating ferrybox data acquisition and for supporting the Ferry Ocean Color Observation Systems (FOCOS). We are thankful to Cermaq Canada for permitting field investigations at their aquaculture sites and for their logistical and sampling support. Special thanks go to Keith Holmes for his contributions to graphics. This research was supported by funding from NSERC NCE MEOPAR—Marine Environmental Observation, Prediction, and Response Network, as well as by the Canadian Space Agency (FAST 18FAVICB09). Additional financial support for instrumentation on BC Ferries was provided by CFI/BCKDF and NSERC-DG grants awarded to Costa. Contributions from Hongyan Xi and Astrid Bracher were funded through the Copernicus Marine Service GLOPHYTS project (21036L05B-COP-INNO SCI-9000), implemented by Mercator Ocean under a delegation agreement with the European Union. In addition, Astrid Bracher's contribution was partially funded by the German Federal Ministry of Economic Affairs and Climate Action through the TypSynSat (FKZ 50 EE 1915) and EnMAP CalVal Water (FKZ 50 EE 1923) projects.

Data Availability Statement

The ferrybox and above-water radiometry data set utilized in this study is publicly accessible through online repositories (Data Search: Ocean Networks Canada (2024)—Oceans 3.0). In situ data sets, including CHEMTAX-derived phytoplankton composition, bio-optical properties, (https://doi.org/10.5281/zenodo.15051890, Vishnu, 2024a), and hyperspectral reflectance data (https://doi.org/10.5281/zenodo.15051957, Vishnu, 2024b) are available on Zenodo. Data analysis was performed using R programming and the script to estimate phytoplankton composition is available at https://github.com/vishnu-star/Mapping-Phytoplankton-Functional-Types.

References

Alikas, K., Vabson, V., Ansko, I., Tilstone, G. H., Olmo, G. D., Nencioli, F., et al. (2020). Comparison of above-water Seabird and TriOS radiometers along an Atlantic meridional transect. *Remote Sensing*, *12*(10), 1–29. https://doi.org/10.3390/rs12101669

Allen, S. E., & Wolfe, M. A. (2013). Hindcast of the timing of the spring phytoplankton bloom in the Strait of Georgia, 1968–2010. *Progress in Oceanography*, *115*, 6–13. https://doi.org/10.1016/j.pocean.2013.05.026

Alvain, S., Moulin, C., Dandonneau, Y., & Bréon, F. M. (2005). Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep Sea Research Part I: Oceanographic Research Papers*, *52*(11), 1989–2004. https://doi.org/10.1016/j.dsr.2005.06.015

Bailey, S. W., & Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment*, *102*(1–2), 12–23. https://doi.org/10.1016/j.rse.2006.01.015

Bidigare, R. R., Morrow, J. H., & Kiefer, D. A. (1989). Derivative analysis of spectral absorption by photosynthetic pigments in the western Sargasso Sea. *Journal of Marine Research*, *47*(2), 323–341. https://doi.org/10.1357/002224089785076325

Bidigare, R. R., Ondrusek, M. E., Morrow, J. H., & Kiefer, D. A. (1990). In vivo absorption properties of algal pigments. *Ocean Optics X*, *1302*, 290–302.

Bracher, A., Bouman, H. A., Brewin, R. J. W., Bricaud, A., Brotas, V., Ciotti, A. M., et al. (2017). Obtaining phytoplankton diversity from ocean color: A scientific roadmap for future development. *Frontiers in Marine Science*, *4*, 55. https://doi.org/10.3389/fmars.2017.00055

Bracher, A., Soppa, M. A., Gege, P., Losa, S. N., Silva, B., Steinmetz, F., & Dröscher, I. (2021). Extension of atmospheric correction Polymer to hyperspectral sensors: Application to HICO and first results for DESIS data. In *2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1237–1240). https://doi.org/10.1109/IGARSS47720.2021.9553568

Bracher, A., Taylor, M. H., Taylor, B., Dinter, T., Röttgers, R., & Steinmetz, F. (2015). Using empirical orthogonal functions derived from remote-sensing reflectance for the prediction of phytoplankton pigment concentrations. *Ocean Science*, *11*(1), 139–158. https://doi.org/10.5194/os-11-139-2015

Bracher, A., Vountas, M., Dinter, T., Burrows, J. P., Röttgers, R., & Peeken, I. (2009). Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences*, *6*(5), 751–764. https://doi.org/10.5194/bg-6-751-2009

Bracher, A., Xi, H., Dinter, T., Mangin, A., Strass, V., Von Appen, W., & Wiegmann, S. (2020). High-resolution water column phytoplankton composition across the Atlantic Ocean from ship-towed vertical undulating radiometry. *Frontiers in Marine Science*, *7*, 235. https://doi.org/10.3389/fmars.2020.00235

Brewin, R. J. W., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., & Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecological Modelling*, *221*(11), 1472–1483. https://doi.org/10.1016/j.ecolmodel.2010.02.014

Cael, B. B., Chase, A., & Boss, E. (2020). Information content of absorption spectra and implications for ocean color inversion. *Applied Optics*, *59*(13), 3971–3984. https://doi.org/10.1364/ao.389189

Carstensen, J., Klais, R., & Cloern, J. E. (2015). Phytoplankton blooms in estuarine and coastal waters: Seasonal patterns and key species. *Estuarine, Coastal and Shelf Science*, *162*, 98–109. https://doi.org/10.1016/j.ecss.2015.05.005

Carswell, T., Costa, M., Young, E., Komick, N., Gower, J., & Sweeting, R. (2017). Evaluation of MODIS-Aqua atmospheric correction and chlorophyll products of western North American coastal waters based on 13 years of data. *Remote Sensing*, *9*(10), 1–24. https://doi.org/10.3390/rs9101063

Catlett, D., & Siegel, D. A. (2018). Phytoplankton pigment communities can be modeled using unique relationships with spectral absorption signatures in a dynamic coastal environment. *Journal of Geophysical Research: Oceans*, *123*(1), 246–264. https://doi.org/10.1002/2017JC013195

Chase, A., Boss, E., Zaneveld, R., Bricaud, A., Claustre, H., Ras, J., et al. (2013). Decomposition of in situ particulate absorption spectra. *Methods in Oceanography*, *7*, 110–124. https://doi.org/10.1016/j.mio.2014.02.002

Ciotti, A. M., & Bricaud, A. (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil. *Limnology and Oceanography: Methods*, *4*(7), 237–253. https://doi.org/10.4319/lom.2006.4.237

Collins, A. K., N, S. E. A., & Pawlowicz, R. (2009). The role of wind in determining the timing of the spring bloom in the Strait of Georgia. *Canadian Journal of Fisheries and Aquatic Sciences*, *66*(9), 1597–1616. https://doi.org/10.1139/F09-071

Craig, S. E., Jones, C. T., Li, W. K. W., Lazin, G., Horne, E., Caverhill, C., & Cullen, J. J. (2012). Deriving optical metrics of coastal phytoplankton biomass from ocean colour. *Remote Sensing of Environment*, *119*, 72–83. https://doi.org/10.1016/j.rse.2011.12.007

Craig, S. E., Lohrenz, S. E., Lee, Z. P., Mahoney, K. L., Kirkpatrick, G. J., Schofield, O. M., & Steward, R. G. (2006). Use of hyperspectral remote sensing reflectance for detection and assessment of the harmful alga *Karenia brevis*. *Applied Optics*, *45*(21), 5414–5425. https://doi.org/10.1364/AO.45.005414

Dai, Y., Yang, S., Zhao, D., Hu, C., Xu, W., Anderson, D. M., et al. (2023). Coastal phytoplankton blooms expand and intensify in the 21st century. *Nature*, *615*(7951), 280–284. https://doi.org/10.1038/s41586-023-05760-y

Dashkova, V., Malashenkov, D., Poulton, N., Vorobjev, I., & Barteneva, N. S. (2017). Imaging flow cytometry for phytoplankton analysis. *Methods*, *112*, 188–200. https://doi.org/10.1016/j.ymeth.2016.05.007

Del Bel Belluz, J., Peña, M. A., Jackson, J. M., & Nemcek, N. (2021). Phytoplankton composition and environmental drivers in the Northern Strait of Georgia (Salish Sea), British Columbia, Canada. *Estuaries and Coasts*, *44*(5), 1419–1439. https://doi.org/10.1007/s12237-020-00858-2

Devred, E., Sathyendranath, S., Stuart, V., Maass, H., Ulloa, O., & Platt, T. (2006). A two-component model of phytoplankton absorption in the open ocean: Theory and applications. *Journal of Geophysical Research*, *111*(C3), C03010. https://doi.org/10.1029/2005JC002880

Dierssen, H. M., Ackleson, S. G., Joyce, K. E., Hestir, E. L., Castagna, A., Lavender, S., & McManus, M. A. (2021). Living up to the hype of hyperspectral aquatic remote sensing: Science, resources, and outlook. *Frontiers in Environmental Science*, *9*, 1–26. https://doi.org/10.3389/fenvs.2021.649528

Esenkulova, S., Suchy, K. D., Pawlowicz, R., Costa, M., & Pearsall, I. A. (2021). Harmful algae and oceanographic conditions in the Strait of Georgia, Canada based on citizen science monitoring. *Frontiers in Marine Science*, *8*, 725092. https://doi.org/10.3389/fmars.2021.725092

Evans, C., & Brussaard, C. P. D. (2012). Viral lysis and microzooplankton grazing of phytoplankton throughout the Southern Ocean. *Limnology & Oceanography*, *57*(6), 1826–1837. https://doi.org/10.4319/lo.2012.57.6.1826

Ferrari, G. M., Dowell, M. D., Grossi, S., & Targa, C. (1996). Relationship between the optical properties of chromophoric dissolved organic matter and total concentration of dissolved organic carbon in the southern Baltic Sea region. *Marine Chemistry*, *55*(3–4), 299–316. https://doi.org/10.1016/S0304-4203(96)00061-8

Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science*, *281*(5374), 237–240. https://doi.org/10.1126/science.281.5374.237

Friedland, K. D., Moisan, J. R., Maureaud, A. A., Brady, D. C., Davies, A. J., Bograd, S. J., et al. (2021). Trends in phytoplankton communities within large marine ecosystems diverge from the global ocean. *Canadian Journal of Fisheries and Aquatic Sciences*, *78*(11), 1–12. https://doi.org/10.1139/cjfas-2020-0423

Giannini, F., Hunt, B. P. V., Jacoby, D., & Costa, M. (2021). Performance of OLCI Sentinel-3A satellite in the Northeast Pacific coastal waters. *Remote Sensing of Environment*, *256*, 1–22. https://doi.org/10.1016/j.rse.2021.112317

Giardino, C., Bresciani, M., Braga, F., Fabbretto, A., Ghirardi, N., Pepe, M., et al. (2020). First evaluation of PRISMA Level 1 data for water applications. *Sensors*, *20*, 1–16. https://doi.org/10.3390/s20164553

Guidi, L., Stemmann, L., Jackson, G. A., Ibanez, F., Claustre, H., Legendre, L., et al. (2009). Effects of phytoplankton community on production, size, and export of large aggregates: A world-ocean analysis. *Limnology & Oceanography*, *54*(6), 1951–1963. https://doi.org/10.4319/lo.2009.54.6.1951

Hallegraeff, G. M., Anderson, D. M., & Cembella, A. D. (2004). *Manual on harmful marine microalgae, monographs on oceanographic methodology*. UNESCO. Paris.

Halverson, M. J., & Pawlowicz, R. (2013). High-resolution observations of chlorophyll-a biomass from an instrumented ferry: Influence of the Fraser River plume from 2003 to 2006. *Continental Shelf Research*, *59*, 52–64. https://doi.org/10.1016/j.csr.2013.04.010

Harrison, P. J., Fulton, J. D., Taylor, F. J. R., & Parsons, T. R. (1983). Review of the biological oceanography of the Strait of Georgia: Pelagic environment. *Canadian Journal of Fisheries and Aquatic Sciences*, *40*(7), 1064–1094. https://doi.org/10.1139/f83-129

Hayward, A., Pinkerton, M. H., & Gutierrez-Rodriguez, A. (2023). phytoclass: A pigment-based chemotaxonomic method to determine the biomass of phytoplankton classes. *Limnology and Oceanography: Methods*, *21*(4), 220–241. https://doi.org/10.1002/lom3.10541

Henson, S. A., Cael, B. B., Allen, S. R., & Dutkiewicz, S. (2021). Future phytoplankton diversity in a changing climate. *Nature Communications*, *12*, 1–8. https://doi.org/10.1038/s41467-021-25699-w

Hilborn, A., & Costa, M. (2018). Applications of DINEOF to satellite-derived chlorophyll-a from a productive coastal region. *Remote Sensing*, *10*(9), 1449. https://doi.org/10.3390/rs10091449

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J. W., Aiken, J., Barlow, R., Suzuki, K., et al. (2011). Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences*, *8*(2), 311–327. https://doi.org/10.5194/bg-8-311-2011

Hoepffner, N., & Sathyendranath, S. (1991). Retrieval of phytoplankton pigment composition from their in vivo absorption spectra. In *Marine Ecology Progress Series* (Vol. 73, pp. 11–23).

Hoepffner, N., & Sathyendranath, S. (1993). Determination of the major groups of phytoplankton pigments from the absorption spectra of total particulate matter. *Journal of Geophysical Research*, *98*(C12), 22789–22803. https://doi.org/10.1029/93jc01273

Hooker, S. B., & Morel, A. (2003). Platform and environmental effects on above-water determinations of water-leaving radiances. *Journal of Atmospheric and Oceanic Technology*, *20*(1), 187–205. https://doi.org/10.1175/1520-0426(2003)020<0187:paeeoa>2.0.co;2

IOCCG. (2000). Remote sensing of ocean colour in coastal, and other optically-complex, waters. In S. Sathyendranath (Ed.), *Reports of the International Ocean Color Coordinating Group (No. 3)*. IOCCG.

IOCCG. (2014). Phytoplankton functional types from space. In S. Sathyendranath & V. Stuart (Eds.), *Reports of the International Ocean Color Coordinating Group (No. 15)*. IOCCG.

Johannessen, S. C., Macdonald, R. W., & Strivens, J. E. (2021). Has primary production declined in the Salish Sea? *Canadian Journal of Fisheries and Aquatic Sciences*, *78*(3), 312–321. https://doi.org/10.1139/cjfas-2020-0115

Komick, N. M., Costa, M. P. F., & Gower, J. (2009). Bio-optical algorithm evaluation for MODIS for western Canada coastal waters: An exploratory approach using in situ reflectance. *Remote Sensing of Environment*, *113*(4), 794–804. https://doi.org/10.1016/j.rse.2008.12.005

Konik, M., Peña, M. A., Hirawake, T., Hunt, B. P. V., Suseelan, V. P., Eisner, L. B., et al. (2024). Bioregionalization of the subarctic Pacific based on phytoplankton phenology and composition. *Progress in Oceanography*, *228*, 1–22. https://doi.org/10.1016/j.pocean.2024.103315

Kostadinov, T. S., Siegel, D. A., & Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *Journal of Geophysical Research*, *114*(C9), 1–22. https://doi.org/10.1029/2009JC005303

Kramer, S. J., & Siegel, D. A. (2019). How can phytoplankton pigments be best used to characterize surface ocean phytoplankton groups for ocean color remote sensing algorithms? *Journal of Geophysical Research: Oceans*, *124*(11), 7557–7574. https://doi.org/10.1029/2019JC015604

Kramer, S. J., Siegel, D. A., & Graff, J. R. (2020). Phytoplankton community composition determined from co-variability among phytoplankton pigments from the NAAMES field campaign. *Frontiers in Marine Science*, *7*, 1–15. https://doi.org/10.3389/fmars.2020.00215

Kramer, S. J., Siegel, D. A., Maritorena, S., & Catlett, D. (2022). Modeling surface ocean phytoplankton pigments from hyperspectral remote sensing reflectance on global scales. *Remote Sensing of Environment*, *270*, 1–14. https://doi.org/10.1016/j.rse.2021.112879

Lange, P. K., Werdell, P. J., Erickson, Z. K., Dallamo, G., Brewin, R. J. W., Zubkov, M. V., et al. (2020). Radiometric approach for the detection of picophytoplankton assemblages across oceanic fronts. *Optics Express*, *28*(18), 25682–25705. https://doi.org/10.1364/OE.398127

Lee, Z. P., Du, K., Voss, K. J., Zibordi, G., Lubac, B., Arnone, R., & Weidemann, A. (2011). An inherent-optical-property-centered approach to correct the angular effects in water-leaving radiance. *Applied Optics*, *50*(19), 3155–3167. https://doi.org/10.1364/AO.50.003155

Le Quéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., et al. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, *11*, 2016–2040. https://doi.org/10.1111/j.1365-2486.2005.01004.x

Lewitus, A. J., White, D. L., Tymowski, R. G., Geesey, M. E., Hymel, S. N., & Noble, P. A. (2005). Adapting the CHEMTAX method for assessing phytoplankton taxonomic composition in southeastern U.S. estuaries. *Estuaries*, *28*(1), 160–172. https://doi.org/10.1007/BF02732761

Lin, J., Olmo, G., Tilstone, G., Brewin, R., Vabson, V., Ansko, I., et al. (2022). Derivation of uncertainty budgets for continuous above-water radiometric measurements along an Atlantic Meridional Transect. *Optics Express*, *30*(25), 45648–45675. https://doi.org/10.1364/oe.470994

Lin, Y., Gifford, S., Ducklow, H., Schofield, O., & Cassar, N. (2019). Towards quantitative microbiome community profiling using internal standards. *Applied and Environmental Microbiology*, *85*(5), 1–14. https://doi.org/10.1128/aem.02634-18

Lombard, F., Boss, E., Waite, A. M., Uitz, J., Stemmann, L., Sosik, H. M., et al. (2019). Globally consistent quantitative observations of planktonic ecosystems. *Frontiers in Marine Science*, *6*, 1–21. https://doi.org/10.3389/fmars.2019.00196

Loos, E. A., & Costa, M. (2010). Inherent optical properties and optical mass classification of the waters of the Strait of Georgia, British Columbia, Canada. *Progress in Oceanography*, *87*(1–4), 144–156. https://doi.org/10.1016/j.pocean.2010.09.004

Losa, S., Soppa, M. A., Dinter, T., Wolanin, A., Brewin, R. J. W., Bricaud, A., et al. (2017). Synergistic exploitation of hyper- and multispectral precursor Sentinel measurements to determine Phytoplankton Functional Types at best spatial and temporal resolution (SynSenPFT). *Frontiers in Marine Science*, *4*, 203. https://doi.org/10.3389/fmars.2017.00203

Mackey, M. D., Mackey, D. J., Higgins, H. W., & Wright, S. W. (1996). CHEMTAX—A program for estimating class abundances from chemical markers: Application to HPLC measurements of phytoplankton. *Marine Ecology Progress Series*, *144*, 265–283. https://doi.org/10.3354/meps144265

MacNeil, L., Costa, M., & LaRoche, J. (2024). Glimpsing the 2020 spring bloom in the Strait of Georgia (Canada) with autonomous ferry-based sensors. *Marine Ecology Progress Series*, *736*, 181–187. https://doi.org/10.3354/meps14574

MacNeil, L., Desai, D. K., Costa, M., & LaRoche, J. (2022). Combining multi-marker metabarcoding and digital holography to describe eukaryotic plankton across the Newfoundland Shelf. *Scientific Reports*, *12*, 1–11. https://doi.org/10.1038/s41598-022-17313-w

Marchese, C., Hunt, B. P. V., Giannini, F., Ehrler, M., & Costa, M. (2022). Bioregionalization of the coastal and open oceans of British Columbia and Southeast Alaska based on Sentinel-3A satellite-derived phytoplankton seasonality. *Frontiers in Marine Science*, *9*(968470), 1–22. https://doi.org/10.3389/fmars.2022.968470

Margalef, R. (1978). Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanologica Acta*, *1*, 493–509. https://doi.org/10.1007/BF00202661

McClain, C. R. (2009). A decade of satellite ocean color observations. *Annual Review of Marine Science*, *1*, 19–42. https://doi.org/10.1146/annurev.marine.010908.163650

McEwan, N., Pawlowicz, R., Pakhomov, E., & Maldonado, M. T. (2023). Seasonality of modeled planktonic food web structure in the Strait of Georgia, Canada. *Ecological Modelling*, *482*, 1–14. https://doi.org/10.1016/j.ecolmodel.2023.110402

Mobley, C. D. (1999). Estimation of the remote-sensing reflectance from above-surface measurements. *Applied Optics*, *38*(36), 7442–7455. https://doi.org/10.1364/ao.38.007442

Mouw, C. B., Ciochetto, A. B., & Yoder, J. A. (2019). A satellite assessment of environmental controls of phytoplankton community size structure. *Global Biogeochemical Cycles*, *33*(5), 540–558. https://doi.org/10.1029/2018GB006118

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. J. W., Bricaud, A., et al. (2017). A consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Frontiers in Marine Science*, *4*, 1–19. https://doi.org/10.3389/fmars.2017.00041

Mouw, C. B., & Yoder, J. A. (2010). Optical determination of phytoplankton size composition from global SeaWiFS imagery. *J. Geophys. Res. Ocean*, *115*, 1–20. https://doi.org/10.1109/TWC.2006.1611067

Mueller, J. L., Bidigare, R. R., Trees, C., Balch, W. M., Dore, J., Drapeau, D. T., et al. (2003). Ocean optics protocols for satellite ocean color sensor validation, Revision 5: Biogeochemical and bio-optical measurements and data analysis protocols. Greenbelt, US.

Nasiha, H. J., Wang, Z., Giannini, F., & Costa, M. (2022). Spatial variability of in situ above-water reflectance in coastal dynamic waters: Implications for satellite match-up analysis. *Frontiers in Remote Sensing*, *3*, 1–17. https://doi.org/10.3389/frsen.2022.876748

Nemcek, N., Hennekes, M., Sastri, A., & Perry, R. I. (2023). Seasonal and spatial dynamics of the phytoplankton community in the Salish Sea, 2015–2019. *Progress in Oceanography*, *217*, 1–19. https://doi.org/10.1016/j.pocean.2023.103108

Ocean Networks Canada. (2024). Ferrybox and above-water radiometry dataset available via Oceans 3.0 data portal [Dataset]. Retrieved from https://www.oceannetworks.ca/data-tools/data-access/oceans-3-0

Olson, R. J., & Sosik, H. M. (2007). A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot. *Limnology and Oceanography: Methods*, *5*(6), 195–203. https://doi.org/10.4319/lom.2007.5.195

Owens, D., Abeysirigunawardena, D., Biffard, B., Chen, Y., Conley, P., Jenkyns, R., et al. (2022). The oceans 2.0/3.0 data management and archival system. *Frontiers in Marine Science*, *9*, 1–33. https://doi.org/10.3389/fmars.2022.806452

Palacz, A. P., St. John, M. A., Brewin, R. J. W., Hirata, T., & Gregg, W. W. (2013). Distribution of phytoplankton functional types in high-nitrate, low-chlorophyll waters in a new diagnostic ecological indicator model. *Biogeosciences*, *10*(11), 7553–7574. https://doi.org/10.5194/bg-10-7553-2013

Peña, M. A., Masson, D., & Callendar, W. (2016). Annual plankton dynamics in a coupled physical–biological model of the Strait of Georgia, British Columbia. *Progress in Oceanography*, *146*, 58–74. https://doi.org/10.1016/j.pocean.2016.06.002

Phillips, S. R., & Costa, M. (2017). Spatial-temporal bio-optical classification of dynamic semi-estuarine waters in western North America. *Estuarine, Coastal and Shelf Science*, *199*, 35–48. https://doi.org/10.1016/j.ecss.2017.09.029

Pinckney, J. L. (2010). The fourth SeaWiFS HPLC analysis round-robin experiment (SeaHARRE-4): Chapter 11 the USC method. In *NASA Technical Memorandum*.

Pramlall, S., Jackson, J. M., Konik, M., & Costa, M. (2023). Merged multi-sensor ocean colour chlorophyll product evaluation for the British Columbia coast. *Remote Sensing*, *15*(3), 1–17. https://doi.org/10.3390/rs15030687

Pramlall, S., Jackson, J. M., Marchese, C., Suchy, K. D., Hunt, B. P. V., & Costa, M. (2024). Mapping phenoregions and phytoplankton seasonality in Northeast Pacific marine coastal ecosystems via a satellite-based approach. *Progress in Oceanography*, *228*, 1–17. https://doi.org/10.1016/j.pocean.2024.103336

Raitsos, D. E., Lavender, S. J., Maravelias, C. D., Haralabous, J., Richardson, A. J., & Reid, P. C. (2008). Identifying four phytoplankton functional types from space: An ecological approach. *Limnology & Oceanography*, *53*(2), 605–613. https://doi.org/10.4319/lo.2008.53.2.0605

Rensel, J. E. J. (2007). *Fish kills from the harmful alga Heterosigma Akashiwo in Puget Sound: Recent blooms and review* (p. 58). National Oceanic and Atmospheric Administration Center for Sponsored Coastal Ocean Research (CSCOR). Washington, DC.

Roesler, C., Uitz, J., Claustre, H., Boss, E., Xing, X., Organelli, E., et al. (2017). Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: A global analysis of WET Labs ECO sensors. *Limnology and Oceanography: Methods*, *15*(6), 572–585. https://doi.org/10.1002/lom3.10185

Röttgers, R., Heymann, K., & Krasemann, H. (2014). Suspended matter concentrations in coastal waters: Methodological improvements to quantify individual measurement uncertainty. *Estuarine, Coastal and Shelf Science*, *151*, 575–582. https://doi.org/10.1016/j.ecss.2014.10.010

Ruddick, K. G., De Cauwer, V., Park, Y., & Moore, G. (2006). Seaborne measurements of near-infrared water-leaving reflectance: The similarity spectrum for turbid waters. *Limnology & Oceanography*, *51*(2), 1167–1179. https://doi.org/10.4319/lo.2006.51.2.1167

Stavn, R. H., Rick, H. J., & Falster, A. V. (2009). Correcting the errors from variable sea salt retention and water of hydration in loss on ignition analysis: Implications for studies of estuarine and coastal waters. *Estuarine, Coastal and Shelf Science*, *81*(4), 575–582. https://doi.org/10.1016/j.ecss.2008.12.017

Storch, T., Honold, H.-P., Chabrillat, S., Habermeyer, M., Tucker, P., Brell, M., et al. (2023). The EnMAP imaging spectroscopy mission towards operations. *Remote Sensing of Environment*, *294*, 113682. https://doi.org/10.1016/j.rse.2023.113632

Suchy, K. D., Le Baron, N., Hilborn, A., Perry, R. I., & Costa, M. (2019). Influence of environmental drivers on spatio-temporal dynamics of satellite-derived chlorophyll a in the Strait of Georgia. *Progress in Oceanography*, *176*, 1–17. https://doi.org/10.1016/j.pocean.2019.102134

Suchy, K. D., Young, K., Galbraith, M., Perry, R. I., & Costa, M. (2022). Match/mismatch between phytoplankton and crustacean zooplankton phenology in the Strait of Georgia, Canada. *Frontiers in Marine Science*, *9*, 1–22. https://doi.org/10.3389/fmars.2022.832684

Sun, D., Chen, Y., Li, Z., Wang, S., Chen, P., & Huan, Y. (2025). Global satellite observations of marine phytoplankton pigments over the past two decades. *Limnology & Oceanography*, *70*(5), 1167–1180. https://doi.org/10.1002/lno.70017

Sun, X., Shen, F., Brewin, R. J. W., Li, M., & Zhu, Q. (2022). Light absorption spectra of naturally mixed phytoplankton assemblages for retrieval of phytoplankton group composition in coastal oceans. *Limnology & Oceanography*, *67*(4), 946–961. https://doi.org/10.1002/lno.12047

Sutton, J. N., Johannessen, S. C., & Macdonald, R. W. (2013). A nitrogen budget for the Strait of Georgia, British Columbia, with emphasis on particulate nitrogen and dissolved inorganic nitrogen. *Biogeosciences*, *10*(11), 7179–7194. https://doi.org/10.5194/bg-10-7179-2013

Teng, J., Zhang, T., Sun, K., & Gao, H. (2022). Retrieving pigment concentrations based on hyperspectral measurements of the phytoplankton absorption coefficient in global oceans. *Remote Sensing*, *14*(15), 1–27. https://doi.org/10.3390/rs14153516

Tilstone, G., Olmo, G. D., Hieronymi, M., Ruddick, K., Beck, M., Ligi, M., et al. (2020). Field intercomparison of radiometer measurements for ocean colour validation. *Remote Sensing*, *12*, 1587. https://doi.org/10.3390/rs12101587

Torrecilla, E., Stramski, D., Reynolds, R. A., Millán-Núñez, E., & Piera, J. (2011). Cluster analysis of hyperspectral optical data for discriminating phytoplankton pigment assemblages in the open ocean. *Remote Sensing of Environment*, *115*(10), 2578–2593. https://doi.org/10.1016/j.rse.2011.05.014

Travers-Smith, H., Giannini, F., Sastri, A. R., & Costa, M. (2021). Validation of non-photochemical quenching corrections for chlorophyll-a measurements aboard ships of opportunity. *Frontiers in Marine Science*, *8*, 686750. https://doi.org/10.3389/fmars.2021.686750

Uitz, J., Claustre, H., Morel, A., & Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *Journal of Geophysical Research*, *111*(C8). https://doi.org/10.1029/2005JC003207

Uitz, J., Stramski, D., Reynolds, R. A., & Dubranna, J. (2015). Assessing phytoplankton community composition from hyperspectral measurements of phytoplankton absorption coefficient and remote-sensing reflectance in open-ocean environments. *Remote Sensing of Environment*, *171*, 58–74. https://doi.org/10.1016/j.rse.2015.09.027

Vandenberg, N., Costa, M., Coady, Y., & Agbaje, T. (2017). *PySciDON: A Python scientific framework for development of ocean network applications*. IEEE.

Vishnu, P. S. (2024a). CHEMTAX-derived phytoplankton composition and bio-optical properties from the British Columbia coast [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.15051890

Vishnu, P. S. (2024b). Hyperspectral reflectance data collected from the British Columbia coast [Dataset]. https://doi.org/10.5281/zenodo.15051957

Vishnu, P. S., Xi, H., Del, J., Belluz, B., Hussain, M. S., Bracher, A., & Costa, M. (2022). Seasonal dynamics of major phytoplankton functional types in the coastal waters of the west coast of Canada derived from OLCI Sentinel 3A. *Frontiers in Marine Science*, *10*, 1018510. https://doi.org/10.3389/fmars.2022.1018510

Wang, Z., & Costa, M. (2018). Autonomous shipborne in situ reflectance data in optically complex coastal waters for validation of Sentinel-3 imagery: A case study of the Salish Sea, Canada. In *XXIV Ocean optics conference*. Dubrovnik.

Wang, Z., & Costa, M. (2022). Autonomous shipborne in situ reflectance data in optically complex coastal waters: A case study of the Salish Sea, Canada. *Frontiers in Remote Sensing*, *3*, 867570. https://doi.org/10.3389/frsen.2022.867570

Werdell, P. J., Behrenfeld, M. J., Bontempi, P. S., Boss, E., Cairns, B., Davis, G. T., et al. (2019). The plankton, aerosol, cloud, ocean ecosystem mission status. *Bulletin of the American Meteorological Society*, *100*(9), 1775–1794. https://doi.org/10.1175/BAMS-D-18-0056.1

Wolanin, A., Soppa, M. A., & Bracher, A. (2016). Investigation of spectral band requirements for improving retrievals of phytoplankton functional types. *Remote Sensing*, *8*(10), 871. https://doi.org/10.3390/rs8100871

Xi, H., Hieronymi, M., Krasemann, H., & Röttgers, R. (2017). Phytoplankton group identification using simulated and in situ hyperspectral remote sensing reflectance. *Frontiers in Marine Science*, *4*, 272. https://doi.org/10.3389/fmars.2017.00272

Xi, H., Hieronymi, M., Röttgers, R., Krasemann, H., & Qiu, Z. (2015). Hyperspectral differentiation of phytoplankton taxonomic groups: A comparison between using remote sensing reflectance and absorption spectra. *Remote Sensing*, *7*(11), 14781–14805. https://doi.org/10.3390/rs71114781

Xi, H., Losa, S. N., Mangin, A., Garnesson, P., Bretagnon, M., Demaria, J., et al. (2021). Global chlorophyll a concentrations of phytoplankton functional types with detailed uncertainty assessment using multisensor ocean color and sea surface temperature satellite products. *Journal of Geophysical Research: Oceans*, *126*(5), 127. https://doi.org/10.1029/2020JC017127

Xi, H., Losa, S. N., Mangin, A., Soppa, M. A., Garnesson, P., Demaria, J., et al. (2020). Global retrieval of phytoplankton functional types based on empirical orthogonal functions using CMEMS GlobColour merged products and further extension to OLCI data. *Remote Sensing of Environment*, *240*, 111704. https://doi.org/10.1016/j.rse.2020.111704

Ye, H., Zhang, B., Liao, X., Li, T., Shen, Q., Zhang, F., et al. (2019). Gaussian decomposition and component pigment spectral analysis of phytoplankton absorption spectra. *Journal of Oceanology and Limnology*, *37*(5), 1542–1554. https://doi.org/10.1007/s00343-019-8079-z

Zhang, Y., Wang, G., Sathyendranath, S., Xu, W., Xiao, Y., & Jiang, L. (2021). Retrieval of phytoplankton pigment composition from their in vivo absorption spectra. *Remote Sensing*, *13*(24), 5112. https://doi.org/10.3390/rs13245112

Zibordi, G., Ruddick, K., Ansko, I., Moore, G., Kratzer, S., Icely, J., & Reinart, A. (2012). In situ determination of the remote sensing reflectance: An inter-comparison. *Ocean Science*, *8*(4), 567–586. https://doi.org/10.5194/os-8-567-2012