Characterizing Phytoplankton Community Composition in the Open and Coastal Waters of the Subarctic Northeast Pacific Using In situ and Remote Sensing Techniques

by

Vishnu Perumthuruthil Suseelan

B.Sc., Mahatma Gandhi University, 2011

M.Sc., Mahatma Gandhi University, 2013

A Dissertation Submitted in Partial Fulfilment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Geography

University of Victoria

We acknowledge and respect the lək^{w}əŋən peoples on whose traditional territory the university stands and the Songhees, Esquimalt and W̱SÁNEĆ peoples whose historical relationships with the land continue to this day.

Characterizing Phytoplankton Community Composition in the Open and Coastal Waters of the Subarctic Northeast Pacific Using In situ and Remote Sensing Techniques

by

Vishnu Perumthuruthil Suseelan

B.Sc., Mahatma Gandhi University, 2011

M.Sc., Mahatma Gandhi University, 2013

Supervisory Committee:

Dr. Maycira Costa, Supervisor (Department of Geography)

Dr. Randall Scharien, Committee Member (Department of Geography)

Dr. Akash Sastri, Committee Member (Fisheries and Oceans Canada)

Abstract

Phytoplankton are the foundation of the marine food web and play a central role in global biogeochemical cycles. The overarching phytoplankton community contains a diverse array of species with distinct ecological and biogeochemical traits and niches. As a result, variations in the composition of phytoplankton communities can significantly affect energy transfer through the marine food web, fisheries production, and the vertical export of carbon into the deep ocean. In the subarctic northeast Pacific (SNEP), phytoplankton plays a distinct role in sustaining fisheries through match-mismatch dynamics. Previous studies in this region have mainly relied on in situ measurements, such as HPLC pigment data, Chemical taxonomy software (CHEMTAX), light microscopy, and physical-biological models, to study phytoplankton communities, but lacking the utilization of satellites. This study employs a comprehensive approach, integrating in situ observations, Earth Observation satellite data, and above-water radiometry, to thoroughly examine the composition of phytoplankton communities. To achieve this goal, first we provide evaluation of the Ocean Land Color Instrument (OLCI)-derived products for Case-1 waters, specifically total chlorophyll-a concentration (TChla; mg/m^3) and remote sensing reflectance (Rrs; sr^{-1}), obtained from the POLYMER (POLYnomial-based algorithm applied to MERIS) and the European Space Agency's (ESA) Baseline Atmospheric Correction (BAC) algorithms. Here, we determined that POLYMER performed better than the BAC algorithm across all bands, with the latter resulting in a loss of approximately 60% of valid data points due to the inability to retrieve products under winter cloud conditions. Additionally, POLYMER successfully captured the regional dynamic range of TChla and aligns well with in situ HPLC TChla data. Next, the study characterizes the composition of phytoplankton communities in Case-1, iron-poor open ocean waters and highly dynamic Case-2 coastal waters in the northeast Pacific using data from the Sentinel 3 OLCI satellite. Utilizing an empirical orthogonal function (EOF) approach, a Sentinel 3 OLCI-based algorithm trained with a regional matchup dataset consisting of phytoplankton groups Chlorophyll-a (Chla) and EOF scores derived from Rrs is employed. Additionally, the robustness of the algorithm is statistically assessed through cross-validation. The Sentinel 3 retrievals from 2019 indicate low biomass for all groups in open ocean Case-1 waters, with minimal variability observed across multiple composites. However, increased biomass is observed for all groups except Hapto/Pelago/Cyano towards the northern Gulf of Alaska (GoA), which aligns with CHEMTAX analysis. Conversely, the

2020 composites reveal a longitudinal gradient from the open ocean to the continental shelf/coast for Diatoms/Dino/Green Algae, but an inverse trend is observed for the Hapto/Pelago/Cyano group. Redundancy analysis (RDA) demonstrates that open ocean flagellate groups are positively correlated with Mixed Layer Depth (MLD) and sea surface salinity (SSS), while Diatoms/Dino/GA are negatively correlated with MLD and SSS. In contrast, satellite-based retrievals for the Case-2 water of the west coast of Canada illustrate seasonal spring and fall blooms dominated by diatoms in the Strait of Georgia (SoG), with the highest biomass concentrated in the central, northern, and Juan de Fuca Strait. During the summer, a similar diatom-dominated trend is observed, with the highest biomass recorded on the southwestern coast of Vancouver Island. Interestingly, raphidophytes exhibit the highest average biomass during the summer, with blooms observed off the southwest coast of Vancouver Island and near the Fraser River mouth. The study further explores the spatial-temporal dynamics of phytoplankton composition within the SoG using EOF retrievals from high-resolution hyperspectral Rrs obtained from autonomous platforms. The results indicate better retrievals for all groups than those obtained using Sentinel 3 OLCI data. Overall, the algorithm successfully predicts the localized, diatom-dominated spring bloom and reveals high variability in flagellate communities, with the highest average biomass occurring in the summer. Furthermore, our algorithm captured strong raphidophyte blooms composed of the harmful algae species, *Heterosigma akashiwo*. Additionally, our high-resolution transect data derived from the algorithm highlighted peaks in raphidophyte and diatom coinciding with a sharp decline in SSS suggesting a frontal region within the Fraser River plume. In summary, using data from in situ measurements, satellites, and autonomous platforms, along with supporting environmental driver data, provides a more comprehensive understanding of the dynamics of individual phytoplankton groups, including harmful algal bloom dynamics with significant implications for water quality monitoring and fisheries management in this region.

Acknowledgements

My six years in the PhD program have been the most significant and transformative period in my academic journey, enriched by the guidance and support of numerous brilliant minds.

First and foremost, I express my deepest gratitude to my Ph.D. supervisor, Dr. Maycira Costa, for her guidance, care, and motivational support throughout these six years. Serving as a mentor, she consistently demonstrated trust in and confidence in my research ideas while providing critical insights and solutions to challenges encountered during my Ph.D., ultimately contributing to successfully completing my research objectives. As my supervisor, she consistently provided financial support through her research projects, extending this assistance even beyond three years. Additionally, she facilitated connections between me and researchers from industry, NGOs, and government institutes, opening doors for collaboration and contributing significantly to successfully completing my research objectives.

I extend my gratitude to my co-authors for their invaluable comments, suggestions, and mentorship, which played a crucial role in successfully completing my research papers. Special acknowledgment goes to Dr. Hongyan Xi and Dr. Astrid Bracher for their expertise in the field, which significantly contributed to the success of my Ph.D. research. I appreciate Justin Del Bel Belluz for his creative ideas, suggestions, and assistance with statistical analysis. Furthermore, I am thankful to Dr. Midhun Sha Hussain for sharing his extensive knowledge of the R programming language, which enhanced my capabilities in data analysis.

I would like to thank Spectral Lab member for their enormous support during my PhD. Special thanks to Dr. Fernanda Giannini, Alena Wachman, and Jackson Joly for their fieldwork support.

I sincerely thank the CFI/BCKDF and NSERC DG to Costa, NSERC NCE MEOPAR - Marine Environmental Observation, Prediction and Response Network, and the Canadian Space Agency for funding my PhD. research. Without their financial support, my studies would not have been possible.

I would like to thank my supervising committee Dr. Randall Scharein and Dr. Akash Sastri, for their guidance, support, and mentorship. I thank the Department of Geography, University of Victoria, for providing all the necessary facilities, especially Rick Sykes for technical support.

Finally, most importantly, I express my heartfelt gratitude to my loving family in India. This journey would not have been possible without their unwavering support and unconditional love. A special acknowledgment goes to my beloved wife, Meenu, for her love, motivation, and care, which have been constant during the challenges that have shaped my path over the years.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AC	Atmospheric Correction
ACC	Alaska Coastal Current
AIC	Akaike Information Criteria
Allo	Alloxanthin
Anther	Antheraxanthin
BAC	Baseline Atmospheric Correction
bbp	Particulate Backscattering Coefficient
BC	British Columbia
BF	19′butanoyloxyfucoxanthin
BFe	Bioavailable Iron
BGC-Argo	Biogeochemical Argo Float
BPAC	Bright Pixel Atmospheric Correction
BRDF	Bidirectional Reflectance Distribution Function
C2RCC	Case 2 Regional Coast Color
CDOM	Colored Dissolved Organic Matter
CHEMTAX	Chemical Taxonomy
Chla	Chlorophyll-a concentration
Chla Allomer	Chlorophyll-a Allomer
Chla′	Chlorophyll-a Prime
Chlb	Chlorophyll-b
$Chlc_{12}$	Chlorophyll c_1c_2
$Chlc_3$	Chlorophyll c_3
Chloro	Chlorophytes
CODA	Copernicus Online Data Access
CODAR	Coastal Ocean Dynamic Application Radar
Cp	Log-transformed CHEMTAX-derived Phytoplankton Groups Chlorophyll-a
Crypto	Cryptophytes
CTD	Conductivity, Temperature, and Depth
CV	Coefficient of Variation
Cyano	Cyanobacteria

CZCS	Coastal Zone Color Scanner
DCM	Deep Chlorophyll Maximum
DFO	Fisheries and Oceans Canada
Diad	Diadinoxanthin
Diat	Diatoxanthin
Dictyo	Dictyochophytes
DIN	Dissolved Inorganic Nitrogen
Dino	Dinoflagellates
DIP	Dissolved Inorganic Phosphate
DMSP	Dimethyl Sulphide Producers
DO	Dissolved Oxygen
DSi	Dissolved Silicate
ENSO	El-Nino Southern Oscillations
EOF	Empirical Orthogonal Function
Es	Total Irradiance
ESA	European Space Agency
Fuco	Fucoxanthin
GA	Green Algae
GF/F	Glass Microfiber Filter
GLM	Generalized Linear Model
GoA	Gulf of Alaska
GPT	Graphic Processing Tool
Gyro	Gyroxanthin-diester
HAB	Harmful Algae Bloom
Hapto	Haptophytes
HF	19′Hexanoyloxyfucoxanthin
HNLC	High-Nitrate, Low-Chlorophyll
HPLC	High-performance Liquid Chromatography
HyperSAS	Manually Operated Hyperspectral Surface Acquisition System
In $situ_i$	In situ-derived
IOPs	Inherent Optical Properties
IYS	International Year of Salmon

Li	Sky Radiance
Lt	Sea Surface Radiance
Lut	Lutein
MERIS	Medium Resolution Imaging Spectrometer
MLD	Mixed Layer Depth
MODIS	Moderate Resolution Imaging Spectroradiometer
N	Number of Observations
NAP	Non-algal Particle
Neo	Neoxanthin
NEP	Northeast Pacific
NIR	Near-infrared
NOC	Ocean Networks Canada
NPDC	North Pacific Drift Current
NPGO	North Pacific Gyre Oscillation
NPQ	Non-photochemical Quenching
OC	Ocean Color
OCI	Ocean Color Instrument
OLCI	Ocean Land Color Instrument
PACE	Plankton Aerosol Cloud and Ecosystem
PDO	Pacific Decadal Oscillation
Pelago	Pelagophytes
Peri	Peridinin
PFT	Phytoplankton Functional Type
POLYMER	POLYnomial-based Algorithm Applied to MERIS
Pras	Prasinoxanthin
Prasino	Prasinophytes
PSC	Phytoplankton Size Classes
PSU	Practical Salinity Unit
PySciDON	Python Scientific Framework for Development of Ocean Network Application
QA	Quality Assurance
QA	Queen of Alberni

QC	Quality Controlled
Raphido	Raphidophytes
RDA	Redundancy Analysis
Rrs	Remote Sensing Reflectance
SAS-ST	Autonomous Solar Tracker HyperSAS
Sat_i	Satellite-derived
SEA	Southeast Alaska
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SERIES	Subarctic Ecosystem Response to Iron Enrichment
SF-Chla	Size-fractionated Chlorophyll-a
SFF	Size-fractionated Filtration
Sigma-t	Potential Density
SNAP	Sentinel Application Platform
SNEP	Subarctic Northeast Pacific
SNPC	Subtropical North Pacific Current
SNR	Signal-to-noise Ratio
SoG	Strait of Georgia
SRF	Spectral Response Function
SSS	Sea Surface Salinity
SST	Sea Surface Temperature
Stn	Station
SVD	Singular Value Decomposition
TChla	Total Chlorophyll-a concentration
TSM	Total Suspended Matter
VIFs	Variance Inflation Factors
VIIRS	Visible Infrared Imaging Radiometer Suite
Viola	Violaxanthin
Ward.D2	Ward's Linkage Method
Zea	Zeaxanthin

List of Symbols

λ	Wavelength
θv	Viewing Zenith Angle
φv	Viewing Sun Azimuth
Sr^{-1}	Steradian
ρs	Fraction of Sky Radiance
U	EOF Scores
Λ	Singular Values
V	EOF Loadings
S	Slope
a	Intercept
MAE	Mean Absolute Error
MDPD	Median Percentage Difference
RMSE	Root Mean Square Error
R^2	Coefficient of Determination
r	Pearson's correlation coefficient
$\sum$	Sum
p	Probability Value
℃	Degree Celsius
<	Less than
>	Greater than
≤	Less than or Equal to
≥	Greater than or Equal to
mg/m^3	Milligram Per Cubic Meter
µm	Micrometer
CO_2	Carbon Dioxide
NO_2	Nitrate
NO_3	Nitrite
PO_4	Phosphate
SiO_4	Silicate
µM/L	Micromol Per Liter
Kg/m^3	Kilogram Per Cubic Meter

ml/L	Milliliters Per Liter
%	Percentage
L	Length
ind/m^3	Individual Per Cubic Meter
μ	Mean
σ	Standard Deviation
$g\ C\ m^{-2}$	Grams of Carbon Per Square Meter
m^{-1}	Per meter
Cells/L	Cells Per Liter
CHL_{FL}	Ferrybox-derived Chlorophyll-a Fluorescence

Preface

This dissertation is an original work by the author.

Chapter 2 has been published in the journal of Remote Sensing:

Vishnu, P.S., Costa, M., 2023. Evaluating the Performance of Sentinel-3A OLCI Products in the Subarctic Northeast Pacific. Remote Sens. 15, 1–22. https://doi.org/10.3390/rs15133244.

Chapter 3 is under review in the Journal of Geophysical Research: Oceans:

Vishnu, P.S., Del, J., Belluz, B., Xi, H., Hunt, B., P.V., Vazhova, A., Hussain, M.S., Pena, A., M., Bracher, A., Costa, M., 2023. Assessing the Spatial-Temporal Variability of Phytoplankton Community Composition and Their Response to Environmental Drivers in the Subarctic Northeast Pacific by Combining In Situ and Satellite Observations. Under Review in JGR:Ocean.

Chapter 4 has been published in Frontiers in Marine Science:

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. Front. Mar. Sci. 1018510, 1–25. https://doi.org/10.3389/fmars.2022.1018510.

Chapter 1: Introduction

1.1 Motivation and Problem Statement

Marine phytoplankton are free-floating photosynthetic organisms that play a crucial role in regulating the Earth's climate through the biological pump, exporting particulate carbon into the deep ocean, fueling secondary production, and sustaining global fisheries (Falkowski and Oliver, 2007; Beaugrand et al., 2010; Finkel et al., 2010; Steinberg and Landry, 2017). Despite only representing < 1% of the global primary producers' biomass, they are responsible for almost half of the global primary production (Falkowski, 2012). Within the overarching phytoplankton community, the contribution of each taxonomic group and phytoplankton size classes (PSCs) to these ecosystem services is distinct. For example, PSC is a key trait regulating critical biological and ecological functions, including photosynthesis (Uitz et al., 2008), nutrient uptake and cycling (Raven, 1998), metabolism (Marañón, 2015), light absorption (Ciotti et al., 2002; Bricaud et al., 2004), export flux (Guidi et al., 2009; Mouw et al., 2016), and energy transfer through trophic levels (Boyce et al., 2015). On the other hand, individual phytoplankton groups also play distinct roles in the elemental cycles; for example, cyanobacteria (e.g., Trichodesmium) play a crucial role in nitrogen cycle by converting atmospheric nitrogen gas into ammonia (e.g., Nair et al., 2008), coccolithophores contribute to the export of particulate inorganic carbon (PIC) (e.g., Balch et al., 2011), diatoms are crucial for silica export into the ocean interior (Falkowski, 2002; Yool and Tyrrell, 2003; Pančić et al., 2019). These different groups are categorized into phytoplankton functional types (PFTs) based on their unique biogeochemical functions rather than their phylogeny (Le Quéré et al., 2005; IOCCG, 2014).

Due to the warming of the global ocean, marine phytoplankton have been reported to be significantly affected by limited nutrient supply resulting from increased vertical stratification, changes in ocean circulation patterns, as well as ocean acidification, and intensifying oxygen minimum zones (Henson et al., 2021). Predictions from ecosystem models indicate that phytoplankton species diversity and biomass in tropical and subtropical regions may experience a significant decrease due to limited nutrient supply as a result of ocean warming (Bopp et al., 2005; Marinov et al., 2010; Bopp et al., 2013; Dutkiewicz et al., 2013; Henson et al., 2021). Whereas, at higher latitude oceans, phytoplankton biomass may increase due to sea ice retreat, shoaling of the mixed layer depth, followed by increasing light availability in the growing season (Marinov et al., 2010; Dutkiewicz et al., 2013; Cabré et al.,

2015; Kwiatkowski et al., 2020). Future projection from the ecosystem model predicts significant changes in the spatial distribution pattern for many PFTs (Henson et al., 2021). For instance, diatom richness in the higher latitude oceans of the northern hemisphere tends to decline due to reduced silicic acid relative to nitrate, potentially leading to the extinction of approximately 30 diatom types (Dutkiewicz et al., 2020). Meanwhile, diazotroph's biogeography pattern showed an inverse trend, shifting more towards the higher latitude ocean due to changing the nitrate-to-iron ratio (which means lower nitrate relative to iron supply rate) (Dutkiewicz et al., 2014). Yet, the potential impact of ocean warming on a regional marine ecosystem is poorly understood and may result in severe negative consequences, such as a change in species diversity, richness, habitat alteration, and shift in community composition.

In the subarctic northeast Pacific (SNEP), phytoplankton plays a distinct role in sustaining fisheries (e.g., Malick et al., 2015; Suchy et al., 2022); however, this region has been undergoing significant warming over the past fifty years (Hobday and Pecl, 2014) with increased frequency of marine heatwaves over the last decade (Di Lorenzo and Mantua, 2016; Amaya et al., 2020), negatively impacting phytoplankton community composition (Peña et al., 2019). Phytoplankton community composition in the high-nitrate-low-chlorophyll (HNLC) waters of the SNEP has been widely studied using microscopy, high-performance liquid chromatography (HPLC)-derived phytoplankton pigment concentrations, and physical-biological models (e.g., Booth et al., 1993; Harrison et al., 2004; Fujiki et al., 2009; Peña et al., 2019; Zhang et al., 2021). Generally, the authors reported low surface chlorophyll-a (hereafter TChla) concentration ($\leq$0.5 mg/m^3) in this region with weak annual and interannual variability; however, a highly diverse phytoplankton community composition comprising haptophytes, pelagophytes, and green algae dominates throughout the year (e.g., Booth et al., 1993; Fujiki et al., 2009; Peña et al., 2019), with iron limiting diatoms growth and bloom (Harrison et al., 2004; Zhang et al., 2021). Besides the iron availability, phytoplankton growth and primary productivity within this region are further limited by sunlight, winter mixing, water column stability, and top-down control (e.g., Harrison et al., 2004; Henson, 2007; Dagg et al., 2009; Aguilar-Islas et al., 2016). Whereas phytoplankton community composition is found to be distinct in the continental shelf/coastal domain of the SNEP region with extensive spring diatom blooms of *Skeletonema costatum*, *Chaetoceros* spp., and *Thalassiosira* spp. (Harrison et al., 1983; Harris, 2001; Strom et al., 2006; Waite and Mueter, 2013). Phytoplankton biomass (TChla) in this region often reaches > 3 mg/m^3

during spring bloom (e.g., Peterson and Harrison, 2012; Waite & Mueter, 2013; Jackson et al., 2015; Strom et al., 2016; Marchese et al., 2022; Pramlall et al., 2023) and they coincide with the growth and development of certain zooplankton species, such as *Neocalanus plumchrus* (Sastri and Dower, 2009). Furthermore, the timing of spring bloom can influence zooplankton community composition through match/mismatch dynamics, which, in turn, can affect Pacific salmon populations (Malick et al., 2015; Perry et al., 2021; Suchy et al., 2022). The timing, onset, and magnitude of spring bloom across this region ultimately depend on multiple physical, chemical and environmental factors, such as sea-surface temperature (SST), sea-surface salinity (SSS), iron-rich freshwater runoff, availability of macro-nutrients, seasonal upwelling, vertical mixing, light, lengthening of days, water column stability, top-down control and large-scale climate oscillation (e.g., Landry et al., 1993; Harrison et al., 1999; Childers et al., 2005; Marchetti et al., 2006; Henson, 2007; Strom et al., 2016; Suchy et al., 2019; Del Bel Belluz et al., 2021; Nemcek et al., 2023). Seasonally, these coastal systems experience a transition from diatom-dominated spring bloom to summer conditions dominated by high diversity flagellate communities comprised mainly of cryptophytes, prasinophytes, dictyochophytes, haptophytes, and dinoflagellates (Del Bel Belluz et al., 2021; Nemcek et al., 2023). Although, the overall phytoplankton biomass, including their seasonality, timing, onset, and regional drivers, has been studied well in this region, there is still a knowledge gap regarding the trends in the phytoplankton community composition over a large spatial-temporal scale within both open ocean and coastal waters of the SNEP.

Over the past decades, satellite-based ocean color instruments have improved the potential for estimating the surface TChla, a proxy for phytoplankton biomass and primary production for the global oceans (Reilly et al., 1998; Maritorena et al., 2002; McClain, 2009; Sathyendranath et al., 2019) with unprecedented spatial-temporal scale often unattainable with ship-based observations. Despite improved algorithms that have been developed to quantify surface TChla concentration from the open ocean (Brewin et al., 2016), relying entirely on TChla is insufficient for a comprehensive understanding of the role of phytoplankton in carbon export and biogeochemical cycling (Bracher et al., 2017). Given this fact, over the last 25 years, Earth Observation satellites with multispectral capability have demonstrated the possibility of retrieving phytoplankton community composition from the world's open ocean waters (see IOCCG, 2014; Bracher et al., 2017; Mouw et al., 2017 for the detailed review). Existing satellite-based algorithms to retrieve phytoplankton community composition and PSCs are categorized into abundance-based (e.g., Uitz et al., 2006; Brewin

et al., 2010; Hirata et al., 2011), absorption-based (Ciotti and Bricaud, 2006; Mouw and Yoder, 2010; Devred et al., 2011), radiance/reflectance-based (e.g., Alvain et al., 2005; Bracher et al., 2009; Ben Mustapha et al., 2014; Werdell et al., 2014; Rêve-Lamarche et al., 2017; Correa-Ramirez et al., 2018; Xi et al., 2020), backscattering-based (e.g., Kostadinov et al., 2009, 2010) and ecological-based methods (e.g., Raitsos et al., 2008; Palacz et al., 2013). However, these methods applied to current multispectral ocean color satellites often suffer from limited spectral bands and, therefore, it is difficult to separate the natural variability in phytoplankton spectral absorption features caused by community composition changes (IOCCG, 2014; Dierssen et al., 2020, 2021). As an alternative to minimize the above issues, algorithms developed using hyperspectral data have been used to retrieve accessory pigments (e.g., Torrecilla et al., 2011; Uitz et al., 2015; Xi et al., 2015; Kramer et al., 2022), separated diatoms from mixed samples (Isada et al., 2015), identified *Karenia brevis* (e.g., Craig et al., 2006), retrieved diatoms and cyanobacteria concentration (e.g., Bracher et al., 2009), and derived PFTs Chla concentration (e.g., Bracher et al., 2020). These studies demonstrated that hyperspectral Rrs with 5 nm resolution or better in the visible spectrum (400–700 nm) can be used as a potential input to extract the spectral signature of PFTs from both Case-1 and Case-2 waters. Developing these techniques, both multispectral and hyperspectral to derive phytoplankton community composition in the SNEP will contribute to our understanding of the dynamics of the trophic link between phytoplankton composition and fisheries, warning of algal blooms (e.g., *Heterosigma akashiwo*), and understanding how climate change alters the surface biology of SNEP.

1.2 Research Objectives

This doctoral research aims to retrieve the spatial-temporal distribution of phytoplankton community composition in the open (Case-1) and coastal (Case-2) waters of SNEP using a combination of in situ HPLC pigment data, ocean color satellite imagery, and hyperspectral radiometry. We used a globally established empirical orthogonal function (EOF)-based algorithm (Bracher et al., 2015) developed using matchups dataset comprising HPLC pigment data and the corresponding EOF scores derived from Rrs to build regression models for each phytoplankton group and applied this model to the entire satellite imagery and hyperspectral Rrs data to derive phytoplankton composition. This study also evaluated the performance of Sentinel-3A OLCI products, such as TChla and Rrs, specifically for the open ocean waters of the Gulf of Alaska to determine which atmospheric correction scheme

provides the best high-quality Rrs product, which was further used as the spectral input in the EOF-based PFT algorithm. The following specific objectives are addressed to achieve the overarching goal of this doctoral thesis:

1. (1) Evaluate the performance of Sentinel-3A OLCI products such as TChla concentration and Rrs derived from the POLYMER and European Space Agency's Baseline Atmospheric Correction algorithms to define which atmospheric correction algorithm provides the most accurate product across the Case-1 open ocean waters of the subarctic northeast Pacific.
2. (2) Harness the potential of in situ and Sentinel-3A OLCI imagery to characterize the phytoplankton community composition and assess their environmental drivers in the Case-1 open ocean waters of the subarctic northeast Pacific. Based on the Rrs retrieval performance from objective 1, we utilized the POLYMER atmospheric correction algorithm to derive Rrs, which was used as spectral input data in the EOF-based phytoplankton community composition algorithm.
3. (3) Based on the successful application of the EOF algorithm from objective 2, we assessed the spatial-temporal dynamics of PFTs from daily Sentinel-3A OLCI ocean color imagery along with CHEMTAX (derived from HPLC pigment data) data across the optically complex Case-2 waters of the Strait of Georgia.
4. (4) Investigate the feasibility of using the EOF-based algorithm on hyperspectral above-water radiometry data to derive highly resolved phytoplankton community composition across the Case-2 waters of the Strait of Georgia.

Overall, this research provides a significant contribution to our understanding of phytoplankton community composition dynamics over a large spatial-temporal scale within the Case-1 and Case-2 waters of the SENP. The outcome of this thesis provides the efficacy of the spectral-based method for characterizing phytoplankton community composition. Our algorithm evaluation and independent validation demonstrated that dominant groups can be retrieved with reasonable accuracy from ocean color satellites. Moreover, better PFT retrieval and more number of individual groups could be obtained utilizing hyperspectral Rrs. Furthermore, our findings suggest the importance of incorporating environmental driver data as prediction variables to assess the spatial-temporal dynamics of phytoplankton composition. Phytoplankton community composition retrieved from satellite imagery and hyperspectral data offers the opportunity to address the link between various phytoplankton composition and zooplankton phenology, which results in better management of fish stock, and

monitoring harmful algal bloom, including *Heterosigma akashiwo*. Finally, the hyperspectral-based PFT retrieval from this thesis provides great significance and confidence to the upcoming PACE mission from NASA, which carries hyperspectral OCI, and the algorithm trained with hyperspectral Rrs data from OCI enables the better quantification of phytoplankton community composition from the coastal and global open ocean waters.

1.3 Thesis Structure

This doctoral thesis comprises different chapters. The first chapter provides an introduction and objectives. The second chapter compares the performance of the POLYMER and ESA BAC algorithms to derive TChla and Rrs from OLCI ocean color imagery (objective 1). The third chapter investigates the utility of using HPLC pigment data and an EOF-based approach on weekly-composite Sentinel-3 OLCI ocean color imagery to derive phytoplankton community composition and assess their environmental drivers in optically clear Case-1 waters (objective 2). The fourth chapter investigates the potential of in situ HPLC pigment data and an EOF-based approach to daily OLCI ocean color imagery to derive PFTs in optically complex Case-2 waters (objective 3). In the fifth chapter, highly resolved phytoplankton community composition is derived from autonomous above-water radiometry for the optically complex Case-2 waters (objective 4). Chapter 6 provides a summary and conclusion. Chapters 2 to 3 are adopted from peer-reviewed research articles, whereas Chapter 4 is under review in a peer-reviewed journal, and Chapter 5 will be submitted to a peer-reviewed journal. Because this doctoral thesis is manuscript-based, redundancy in the introduction and method sections within the thesis chapter can be observed.

Chapter 2: Evaluating the Performance of Sentinel-3A OLCI Products in the Subarctic Northeast Pacific

Remote Sensing, 15(3244), pp. 1-22.

https://doi.org/10.3390/rs15133244

Perumthuruthil Suseelan Vishnu and Maycira Costa

SPECTRAL Remote Sensing Laboratory, Department of Geography, University of Victoria, Victoria, BC V8P 5C2, Canada.

Abstract

The subarctic northeast Pacific (SNEP) is a high-nitrate, low-chlorophyll (HNLC) region in the ocean, where phytoplankton growth and productivity are limited by iron. Moreover, there is a limited application of high spatial (300 m) and temporal resolution (daily) ocean color (OC) satellite imagery in studying the phytoplankton dynamics in this region. To address this issue, we aim to validate the remote sensing reflectance (Rrs; $sr^{-1}(\lambda)$) and total chlorophyll-a (TChla) concentration derived from the POLYMER atmospheric correction algorithm against in situ data for the SNEP obtained during 2019 and 2020. Additionally, we performed qualitative analysis using weekly binned Sentinel 3 TChla maps to determine whether the product reflects the general pattern over a latitudinal and longitudinal domain. We processed the daily Level-1 image using POLYMER and binned them weekly using Graphic Processing Tool (GPT). The validation results indicate that POLYMER algorithm exhibits better retrieval in the blue and green bands and fails to represent in situ Rrs in the red band. Furthermore, the POLYMER slightly over- and underestimates reflectance between 0.0012 and 0.0018 sr^{-1} in the green band. On the other hand, excellent agreement was found between satellite-derived versus in situ TChla in the range between 0.3 to 0.5 mg/m^3, followed by a slight overestimation of in situ TChla in the range from 0.17 to 0.28 mg/m^3. The weekly binned TChla spatial map revealed a spatially homogeneous distribution of surface TChla in Gulf of Alaska, but a substantial increase in TChla ($\geq$0.7 mg/m^3) was recorded toward Southeast Alaska (SEA) and the British Columbia (BC) shelf. Furthermore, TChla derived from latitudinal and longitudinal transects indicates high TChla toward 57° North and −135° West, respectively. Overall, the results of this study emphasize the need to obtain high-quality matchups from under-sampled oligotrophic waters, which are crucial for satellite validation, in addition to highlighting the importance of using high spatial and temporal resolution satellite imagery to study phytoplankton dynamics in the SNEP.

2.1 Introduction

Marine phytoplankton are essential for regulating Earth's climate and account for 50% of the global primary productivity (Field et al., 1998; Le Quéré et al., 2005; Falkowski, 2012). They are highly diverse in size, shape, and taxonomic composition, exhibiting enormous diversity within a single species; for example, phytoplankton are found with different morphological and physiological characteristics and sometimes different life stages and ecological niches (Bouman et al., 2006). Thus, monitoring phytoplankton and assessing their shift in the ocean is paramount to determining the global carbon cycle, ocean biogeochemistry, implications for the marine food web, and how climate change will impact the ocean ecosystem.

This is especially relevant for the open ocean waters of the subarctic northeast Pacific (SNEP; **Figure 1**), a high-nutrient–low-chlorophyll (HNLC) region where iron is the limiting nutrient that regulates phytoplankton biomass and primary productivity (Evans and Parslow, 1985; Miller et al., 1991; Brickley and Thomas, 2004; Strom et al., 2016). Total Chlorophyll-a concentration (hereafter TChla, a proxy for phytoplankton biomass (McClain, 2009)) generally remains low in the Gulf of Alaska at ≤0.5 mg/m^3; nevertheless, sporadic phytoplankton bloom may occur in the HNLC region due to the introduction of iron (Hamme et al., 2010). This HNLC region extended through the central Pacific (Brickley and Thomas, 2004); hence, phytoplankton blooms are generally absent during spring and summer (Strom et al., 2016). However, this regular state may sometimes be altered by short-lived phytoplankton blooms during fall (Evans and Parslow, 1985; Frost, 1991; Banse and English, 1999; Harrison et al., 1999) as a result of the relaxation in grazing pressure from micro- and macrozooplankton, aerial iron deposition, or advection of nutrient-rich water from the shelf (Frost, 1991; Bishop et al., 2002; Harris et al., 2009). The Subarctic Ecosystem Response to Iron Enrichment (SERIES) experiment has clearly shown that iron ultimately limits the utilization of macronutrients and regulates new production during summer in the SNEP (Marchetti et al., 2006). This has important implications for the open ocean waters of the SNEP, which are crucial for sustaining commercially important fisheries and supporting the population of Pacific Salmon originating from Canada, the United States, Japan, Korea, and Russia (Beamish, 2017).

Similarly, the coastal and continental domains of SNEP are known for their high productivity and diverse marine life (Womble and Sigler, 2006; Womble et al., 2009; Johnson et al., 2012; Dragoo et al., 2017; McGowan et al., 2019; Siddon et al., 2019; Budge et al., 2022),

including various marine mammals (Womble and Sigler, 2006; Womble et al., 2009), commercially important fish species (Johnson et al., 2012; Siddon et al., 2019, McGowan et al., 2019; Budge et al., 2022), and seabirds (Dragoo et al., 2017). This high productivity is attributed, in part, to the variability of phytoplankton in this region (Sambrotto and Lorenzen, 1987; Childers et al., 2005). Studies using both in situ measurements and satellite data have revealed high values of TChla (>3 mg m^{-3}) in the coastal waters along with phytoplankton blooms, showing pronounced seasonal, spatial, and interannual variability (Weingartner et al., 2009; Waite and Mueter, 2013).

Monitoring the dynamics of ocean productivity and its implications for the food web, which is based on measuring the concentration and composition of phytoplankton in the ocean, is generally conducted using various in situ techniques, including flow cytometry, size-fractionated filtration (SFF), high-performance liquid chromatography (HPLC) pigment analysis, conventional imaging flow cytometry (Olson and Sosik, 2007), Imaging FlowCytobot, genomics, and next-generation sequencing (Lombard et al., 2019). However, specifically for the SNEP, light-microscopy (Booth et al., 1993; Harrison et al., 2004; Peterson and Harrison, 2012), HPLC pigment analysis (Obayashi et al., 2001; Suzuki et al., 2002; Fujiki et al., 2009; Peterson et al., 2011; Yang et al., 2018; Peña et al., 2019), and physical–biological models (Zhang et al., 2021) have been implemented. These methods are extremely useful but are limited in time and space, and are often time-consuming, labor-intensive, and expensive (Lombard et al., 2019; IOCCG, 2014). Consequently, the availability of in situ phytoplankton data becomes constrained, preventing the use of these dataset for monitoring the large-scale dynamics of phytoplankton and understanding how environmental drivers influence their spatiotemporal distribution.

To address this issue, satellite-based ocean monitoring using optical sensors can monitor the upper ocean ecosystem on a synoptic scale with a high spatial and temporal resolution that is often unattainable for ship-based observation (McClain, 2009). Given the importance of determining phytoplankton dynamics over large spatiotemporal scales, several studies have used ocean color (OC) satellites to elucidate phytoplankton dynamics in the SNEP (Vinogradov et al., 1997; Banse and English, 1999; Brickley and Thomas, 2004; Waite and Mueter, 2013; Strom et al., 2016). Indeed, most of these past studies have used OC sensors that have coarse spatial (4–25 km) resolution, and none have yet explored the potential of the new generation of OC sensors such as Ocean Land Color Instrument (OLCI) onboard

Sentinel-3A satellite with high spatial (300 m) and temporal resolution (<1.4 days) (EUMETSAT, 2018).

Here, we evaluate the performance of satellite-derived remote sensing reflectance (Rrs; $sr^{-1}(\lambda)$) and total chlorophyll-a concentration in the SNEP, specifically during 2019 and 2020. The satellite product was mainly produced using the POLYnomial-based algorithm applied to MERIS (POLYMER) and the European Space Agency Baseline Atmospheric Correction (BAC) algorithm, and its performance was then assessed via statistical validation against in situ measurements. Additionally, we conducted a qualitative analysis to discuss whether the satellite data reflect the region's latitudinal and longitudinal dynamics. The evaluation of the OLCI product derived from multiple processors across the global and regional waters is a critical step in the process of Sentinel-3 OLCI validation, and here we are providing the first reported analysis for the northeast Pacific region.

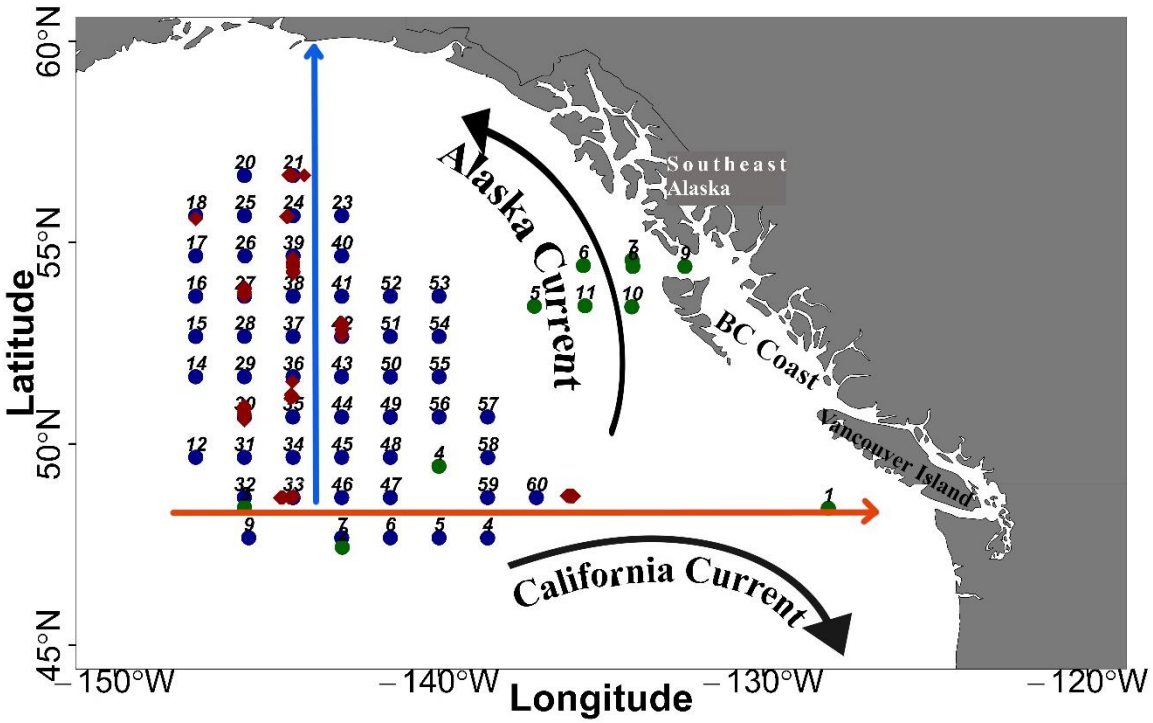

Figure 1. Study area map showing the sampling locations in 2019 and 2020, overlaid with general currents: California Current and Alaska Current. Circles indicate field stations and are shown in blue for 2019 and green for 2020. Red circles indicate the transect for the above-water radiometer operated in 2019. Total Chlorophyll-a concentrations were extracted from each weekly composite imagery over the latitudinal and longitudinal domains represented by the blue and red lines, respectively.

2.2 Materials and Methods

2.2.1 Study Area

The circulation in the Gulf of Alaska (GoA) is primarily driven by the counterclockwise-flowing Subarctic Gyre, consisting of three components: the North Pacific Current, the Alaska Current, and the Alaska Stream (Weingartner, 2006). The Alaska Stream (50–100 cm s^{-1}) is a high-speed current characterized by eddies and meanders, flowing southwestward parallel to the coast of Alaska. It connects to the North Pacific Current at 170° W, completing the Subarctic Gyre (Weingartner, 2006; Weingartner et al., 2009). This oceanographic region (**Figure 1**) is classified as an HNLC area with abundant nitrate concentrations and moderate primary productivity throughout the year due to limited iron availability (Peña and Varela, 2007; Zhang et al., 2021). The primary productivity in the region fluctuates annually between 5 and 18 mol C m^{-2} yr^{-1} (Welschmeyer et al., 1993), whereas the concentrations of TChla in the region exhibit no seasonal or interannual variability, with values ≤ 0.5 mg/m^3 (Peña and Varela, 2007; Zhang et al., 2021). This pattern has been confirmed by remote sensing of the surface TChla and recent observations from the Biogeochemical Argo float (BGC-Argo) deployed in the SNEP, which recorded low variability in TChla concentrations of the upper ocean between 0 and 150 m (Zhang et al., 2021). The dominant phytoplankton community in the open ocean region consists of small-sized nano and picophytoplankton, such as haptophytes, followed by green algae and pelagophytes (Harrison et al., 2004; Fujiki et al., 2009; Peterson et al., 2011; Peña et al., 2019), whereas large-sized microphytoplankton, such as diatoms, contribute little to the total biomass in this region (Fujiki et al., 2009; Peña et al., 2019).

In contrast, the coastal waters of southeast Alaska (SEA) and British Columbia (BC) are highly productive and rich in marine life, providing habitats for commercially important fish species such as Pacific Cod, Walleye Pollock, Pacific Ocean Perch, and Pacific Herring (Johnson et al., 2012; Siddon et al., 2019; Budge et al., 2022), as well as several species of marine mammals (Womble et al., 2009) and seabirds (Dragoo et al., 2017). Despite being a downwelling-dominated region throughout the year, the coastal waters of SEA exhibit exceptional productivity, with annual primary productivity surpassing 300 g C m^{-2} y^{-1} (Sambrotto and Lorenzen, 1987). The Alaska Coastal Current (ACC), driven by wind and freshwater runoff, dominates this region and is responsible for transporting dissolved organic matter, phytoplankton, zooplankton, and other marine organisms northwards to the Bering Sea (Weingartner et al., 2005). On the other hand, the BC coast serves as a transitional zone

between the strong downwelling-dominated region in SEA and the upwelling-dominated coastal region in southern California (Weingartner et al., 2009). This transition zone is demarcated at the shelf break by seasonally varying surface water properties (i.e., salinity, density, and nutrients) and establishing a boundary between oceanic and coastal waters (Favorite et al., 1976; Whitney et al., 1998; Cullen et al., 2009). This region demonstrates substantial variability over the seasonal scale as a result of changes in circulation (Whitney and Freeland, 1999) with significant implications for surface biology. Wind-induced seasonal upwelling is generally weak in this region and therefore the surface water properties were governed mainly by the mixing of coastal and offshore water masses (Ribalet et al., 2010). Field investigations revealed a substantial increase in iron concentration along this transition zone, accompanied by a pronounced gradient in TChla, which aligned with the highest phytoplankton diversity observed in the area, subsequently leading to the observation of a bloom predominantly composed of ultraplankton and nanoplankton communities within this region (Ribalet et al., 2010).

This coastal region serves as the primary marine entry point for many Pacific Salmon species, with the majority of juvenile Sockeye Salmon leaving the Salish Sea during summer and fall and entering the GoA through these coastal waters (Thomson et al., 2012). Phytoplankton biomass undergoes annual cycles, with extensive spring diatom blooms observed along the SEA and BC coasts (Peterson et al., 2007; Waite and Mueter, 2013; Jackson et al., 2015). Diatoms are the dominant bloom-forming community in this region, and TChla concentrations typically exceed 3 mg/m^3 during the spring bloom, dominated by *Skeletonema costatum*, *Chaetoceros* spp., and *Thalassiosira* spp. (Harrison et al., 1999; Peterson et al., 2007; Waite and Mueter, 2013; Strom et al., 2016).

2.3 Dataset

The following dataset was compiled for this study: above-water reflectance measurements were acquired to validate satellite-derived reflectance, and discrete water samples collected from the surface ocean were analyzed for HPLC pigments to validate the satellite-derived TChla product. HPLC-derived TChla is one of the most accurate, widely adopted, and quality-controlled data for validating ocean color product. Additionally, qualitative analysis was conducted using weekly binned surface TChla composites generated from atmospheric-corrected daily OLCI imagery to determine the gradients of the region's latitudinal and

longitudinal dynamics in the data. The following section presents the methods followed in this study.

2.3.1 In situ Data

The in situ dataset used in this study (**Figure 1**) were collected as part of the International Year of Salmon (IYS) expedition to the GoA from 18 February to 21 March 2019 and from 5 March to 5 April 2020. The 2019 expedition was onboard R/V *Professor Kaganovsky*, while the 2020 expedition was onboard R/V *Pacific Legacy*. The dataset derived from these expeditions were utilized to perform one-to-one matchup analysis to determine the accuracy of Sentinel-3 OLCI product over the surface waters of the SNEP region. This section is organized as follows: first, we describe the methods used for processing in situ and satellite data; second, we present the validation of satellite-derived Rrs(λ) and TChla; finally, we qualitatively assess how well the weekly binned Sentinel 3 TChla map reproduces the regional pattern over a latitudinal and longitudinal domain.

2.3.2 Phytoplankton Pigment Analysis

From the expeditions, surface water samples for phytoplankton pigment analysis were collected using CTD/rosette from approximately 5 m. Water samples were then transferred to a 10 L plastic container and transported to the laboratory with immediate initiation of onboard filtration, which was conducted in the dark to avoid potential pigment degradation (Mueller et al., 2003). Surface water samples of 1–2 L from each station, for the duplicate measurement of HPLC pigment analysis, were filtered through a 25 mm Whatman GF/F 0.7 μm glass microfiber filter under low light conditions and low vacuum (≤5 mmHg). After filtration, the filters were flash-frozen in liquid nitrogen and immediately kept at −80 °C until analysis in the laboratory. The pigments were then extracted with methanol and analyzed using the Shimadzu HPLC instrument (Pinckney, 2010). The HPLC pigment analysis was conducted at the University of South Carolina Baruch Institute of Marine and Coastal Science (https://phytoninja.com/lab-protocols/, accessed on), and the method was detailed in Pinckney 2010. Out of the 54 samples, 20% had duplicates, and the uncertainty in these duplicate samples was estimated based on a coefficient of variation (CV) < 20%.

2.3.3 In situ Remote Sensing Reflectance

Above-water spectral reflectance measurements were acquired during good weather conditions using a manually operated radiometer (HyperSAS) to evaluate OLCI-derived

Rrs(λ) (Mobley, 1999). HyperSAS was only operated during the 2019 expedition (N = 43) due to logistic constraints in the 2020 expedition. HyperSAS consists of three hyperspectral radiometers, including sea surface radiance Lt(λ), sky radiance Li(λ), and a third sensor to measure total irradiance Es(λ); from these measurements, Rrs(λ) is derived (Mobley, 1999). HyperSAS was installed on the bow of the R/V *Professor Kaganovsky* at a 13 m height from the sea surface to avoid the infrastructure shadows and spray. Specifically, Lt(λ) and Li(λ) sensors are at a fixed-viewing zenith angle, θv = 50°, and viewing sun azimuth, φv = 120°, to avoid glint effects (Mobley, 1999; Hooker and Morel, 2003). The ideal φv is maintained by looking at the sun's position in the sky. Measurements were acquired between 11 a.m. and 2 p.m. to mimic the time of Sentinel-3A acquisition and optimized sun illumination conditions. Radiometric measurements were limited to a solar zenith angle of less than 60° (Hooker and Morel, 2003) to limit variability in the water-leaving radiance. The data from the HyperSAS were processed using ProSoft version 7.7.19 to calculate Rrs(λ).

The above-water Rrs(λ) was calculated according to Ruddick et al. (2005).

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

where Lt(λ) is the total measured radiance by the sea-viewing sensor and ρs fraction of sky radiance measured by the sea-viewing sensor (Ruddick et al., 2005). ρs is estimated according to the following equations (Ruddick et al., 2006).

$$\rho s = 0.0256 + 0.00039W + 0.000034W^2, \text{when } \frac{\mathrm{Li\ (750nm)}}{\mathrm{Es\ (750nm)}} < 0.05 \quad (2)$$

$$\rho s = 0.0256, \text{when } \frac{\mathrm{Li\ (750nm)}}{\mathrm{Es\ (750nm)}} \geq 0.05 \quad (3)$$

NIR (near-infrared) correction was applied to the reflectance spectra by subtracting the value at 750 nm following the method described in Ruddick et al. (2006). The final Rrs(λ) was derived from the quality-controlled in situ hyperspectral Rrs(λ) (hereafter Rrs in situ) subjected to the spectral response function (SRF) according to the spectral characterization of OLCI (https://sentinel.esa.int/web/sentinel/user-guides/sentinel-3-olci, accessed on).

2.3.4 Sentinel-3A OLCI Data and Atmospheric Correction

The Ocean Land Color Instrument (OLCI) onboard the Sentinel-3A satellite provides high spatial (300 m) and temporal resolution ocean color data, and it operates in 10 spectral bands (400, 412, 443, 490, 510, 560, 620, 665, 674, and 681 nm) within the visible spectrum (Donlon et al., 2012). The sensor also features an improved signal-to-noise ratio (SNR) and an off-nadir pointing of the instrument swath to minimize the sunglint effect (Giannini et al.,

2021). Therefore, for this study, OLCI Sentinel-3A Level-1 full-resolution data acquired during 2019 and 2020 (18 February to 21 March 2019, and from 5 March to 5 April 2020, corresponding to the fieldwork period) were downloaded from Sentinel-3 Marine CODA (Copernicus Online Data Access) web service and processed using POLYMER version 4.10 to obtain Rrs(λ) and TChla (mg/m^3) following Steinmetz et al. (2011). POLYMER is a spectral matching algorithm that employs two models. Firstly, a polynomial function utilizes the full spectrum to model the spectral reflectance of the atmosphere and the residual sunglint contribution (Steinmetz et al., 2011). Secondly, the water reflectance model is based on TChla and a coefficient (*f*b), which scales the total backscattering coefficient of marine particles (bbp) (Steinmetz et al., 2011). The parameters in the atmospheric and water reflectance models are optimized through an iterative process to achieve the best spectral fit for the total reflectance received by the sensor (Steinmetz et al., 2011). TChla is a direct output of POLYMER and is derived as part of the optimization scheme (Steinmetz et al., 2011). Specifically, TChla in the water reflectance models is derived from known measurements from various oceanic regions, rather than through a semi-analytical or bio-optical model (Steinmetz et al., 2011). It has been shown that POLYMER water reflectance is applicable in both Case-1 and Case-2 water types (Steinmetz et al., 2011). This algorithm was originally developed for Medium Resolution Imaging Spectrometer (MERIS) and later adapted to other OC sensors such as Moderate Resolution Imaging Spectroradiometer (MODIS), Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Visible Infrared Imaging Radiometer Suite (VIIRS), and OLCI. In this study, daily Level-1 data from the 2019 and 2020 were downloaded and processed with POLYMER and then binned into four composites using SNAP (version 6.0) in batch mode using GPT (Graphics Processing Tool). While binning, the following quality flags of "Cloud", "Invalid", "Negative Backscattering (BB)", "Out-of-bounds", "Exception", "High Air Mass", and "Thick Aerosol" were applied based on the recommendation from Steinmetz et al. (2016) following Giannini et al. (2021), and the median Rrs(λ) calculated from each band was used for the analysis. Overall, eight composites were generated for 2019 and 2020. To obtain higher spatial coverage OC data for the open ocean, spatially and temporally binned 8-day composite maps were employed (Scott and Werdell, 2019); this is especially valid for the iron-poor open ocean waters of GoA, where TChla concentration exhibited subtle variability throughout our investigation (0.38 ± 0.08 mg/m^3) (Peña and Varela, 2007; Peña et al., 2019; Zhang et al., 2021). Along with POLYMER (version 4.10), we tested the 8-day composite map of Level-2 Rrs(λ) and TChla distributed by the European

Space Agency (ESA) to define the best processor for this region. The ESA Level-2 Rrs(λ) product is derived using the BAC algorithm (EUMETSAT, 2017), which was initially developed for MERIS by Antoine and Morel, (1999) and assumes there is no reflectance in the NIR part of the spectrum. Hence, the TChla and Rrs(λ) products are only valid for the global open ocean waters, where phytoplankton pigments dominate the signal. Later, BAC was integrated with the Bright Pixel Atmospheric Correction (BPAC) proposed by Moore et al. (1999) and Moore et al. (2017). BPAC can account for high scattering waters dominated by high TChla and total suspended matter (TSM), where NIR reflectance is not zero. However, the initial validation results of satellite-derived Rrs(λ) (hereafter Rrs-satellite) showed poor radiometric performance at 443 (r = −0.42, Mean Absolute Error = 0.003 sr^{-1}, Median Percentage Difference = 44.99%, BIAS = 12.51%), 560 (r = −0.23, Mean Absolute Error = 0.0006 sr^{-1}, Median Percentage Difference = 33.55%, BIAS = 11.67%), and 665 nm (r = 0.00, Mean Absolute Error = 0.0002 sr^{-1}, Median Percentage Difference = 59.93%, BIAS = 52.04%). Similarly, TChla derived from the ESA Level-2 product for the 2019 dataset also showed poor statistical performance (r = 0.3, MAE = 0.08 mg/m^3, MDPD = 19.31%, BIAS = −2.74%) compared to POLYMER. In addition, we observed a significant reduction (about 60%) of the valid data points when utilizing the ESA Level-2 product. Furthermore, due to clouds, the spatial coverage was substantially reduced during the expedition in the middle of winter in GoA. However, we also noticed that the POLYMER-processed Level-2 image provided much better spatial coverage compared to the ESA Level-2 product. Our observations are consistent with the latest study by Tilstone et al. (2021), where POLYMER-derived TChla reportedly showed the highest number of valid pixels and higher spatial coverage than other AC algorithms across the Atlantic Ocean, and that for the northeast Pacific by Giannini et al. (2021). Due to the design of the POLYMER processor, it is capable of retrieving the signal under sky glint conditions and thin clouds (Steinmetz et al., 2016), thus allowing for extensive spatial coverage. As a result, we decided to conduct the entire validation utilizing POLYMER. Finally, the Level-2 processed imagery from POLYMER was cropped to the region of interest using Graph Processing Tool (GPT) of the Sentinel Application Platform (SNAP), specifically the “Subset” function. The resulting cropped imagery was then visualized using the “Raster” package in R (Hijmans, 2020).

2.3.5 Matchup Analysis: OLCI Rrs and TChla

The performance of the POLYMER AC algorithm was evaluated by conducting a one-to-one matchup analysis between satellite and in situ data. The Rrs satellite from the daily imagery was extracted from a 5×5-pixel window centered on the exact location of the Rrs in situ. Mean, median, and coefficient of variation (CV) were calculated for each matchup using the following criteria: ±5 h, valid pixels must be ≥17/25, CV at 560 nm must be ≤20%, and the median value from each band was used to avoid outlier (Bailey and Werdell, 2006; Mograne et al., 2019; Pramlall et al., 2023).

The performance of OLCI-TChla retrievals was assessed based on the combined data from 2019 and 2020. For each in situ sample, satellite TChla was extracted from the 8-day composite imagery, and the mean, median, and CV were calculated for the TChla matchup using similar criteria as for the Rrs in situ (Bailey and Werdell, 2006; Mograne et al., 2019; Pramlall et al., 2023). We achieved a higher number of points using the 8-day composite imagery for both 2019 and 2020. In addition, the statistical performance was also highly comparable to that of the daily matchups.

The statistical error metrics such as r, mean absolute error (MAE), median percentage error (MDPD), and Bias (BIAS) were obtained to evaluate the performance of POLYMER AC algorithms (Giannini et al., 2021). The equation of MAE, MDPD, and BIAS are expressed as follows.

$$\mathrm{MAE} = \sum_{i=1}^{i=N} |\mathrm{Sat_i} - \mathrm{In\ situ_i})| \quad (4)$$

$$\mathrm{MDPD} = \mathrm{Median\ value\ of}\ 100 \times \left| \frac{\mathrm{Sat_i} - \mathrm{In\ situ_i}}{\mathrm{In\ situ_i}} \right| \quad (5)$$

$$\mathrm{Bias} = \sum_{i=1}^{i=N} \frac{(\mathrm{Sat_i} - \mathrm{In\ situ_i})}{\mathrm{In\ situ_i}} \times 100 \quad (6)$$

where $\mathrm{Sat_i}$ is the satellite-derived measurements, $\mathrm{In\ situ_i}$ is the in situ-derived measurements, and N is the number of observations.

Finally, a weekly trend analysis was conducted to evaluate the accuracy of satellite-derived TChla over a broad spatial scale. The trends are extracted across a latitudinal and longitudinal domain from a 5×5 window within each 8-day composite imagery, and the corresponding standard deviation is presented.

2.4 Results

2.4.1 Remote Sensing Reflectance Validation

Rrs in situ and Rrs satellite demonstrated a typical open ocean water spectrum (Tilstone et al., 2021), characterized by a maximum blue reflectance that decreased toward longer wavelengths and a TChla fluorescence peak near 680 nm (**Figure 2**). Rrs in situ and Rrs satellite for various bands (443, 560, and 665 nm) showed comparable dynamic ranges (**Figure 3).** Specifically, the average values at the 443, 560, and 665 nm bands were comparable between Rrs in situ and Rrs satellite, with values of 0.0049 ± 0.0009, 0.0048 ± 0.0005, 0.0018 ± 0.0003 and 0.0016 ± 0.0003, 0.00024 ± 0.00007, 0.00025 ± 0.00003, respectively. A detailed quantitative analysis demonstrated that the accuracy of POLYMER AC varies with wavelength, where the correlation decreases with increasing wavelength (**Figure 4**). A highly significant ($p < 0.001$) correlation of 0.73 was obtained for 443 nm, followed by the best statistical performance (BIAS = 0.22%, MDPD = 7.97%; **Figure 4**). Similarly, a highly significant ($p < 0.001$) correlation of 0.50 was found for 560 nm, followed by a moderate statistical performance (BIAS = −13.09%, MDPD = 17.08%; **Figure 4**). For Rrs satellite at 560 nm, values < 0.0015 sr^{-1} were slightly overestimated, and Rrs between 0.0015 and 0.0018 sr^{-1} were underestimated. Finally, a poor correlation of −0.2 and decreased statistical metrics were obtained for 665 nm (BIAS = 11.82%, MDPD = 25.47%; **Figure 4**). Overall, Rrs satellite demonstrated the highest overestimation at 665 nm (BIAS = 11.82%), followed by the highest underestimation at 560 nm (BIAS = −13.09%). Similarly, the highest MDPD was observed at 665 nm (25.47%), followed by 560 nm (17.08%) and 443 nm (7.97%). On the other hand, the MAE was lower for all three bands (<4×10^{-4} sr^{-1}).

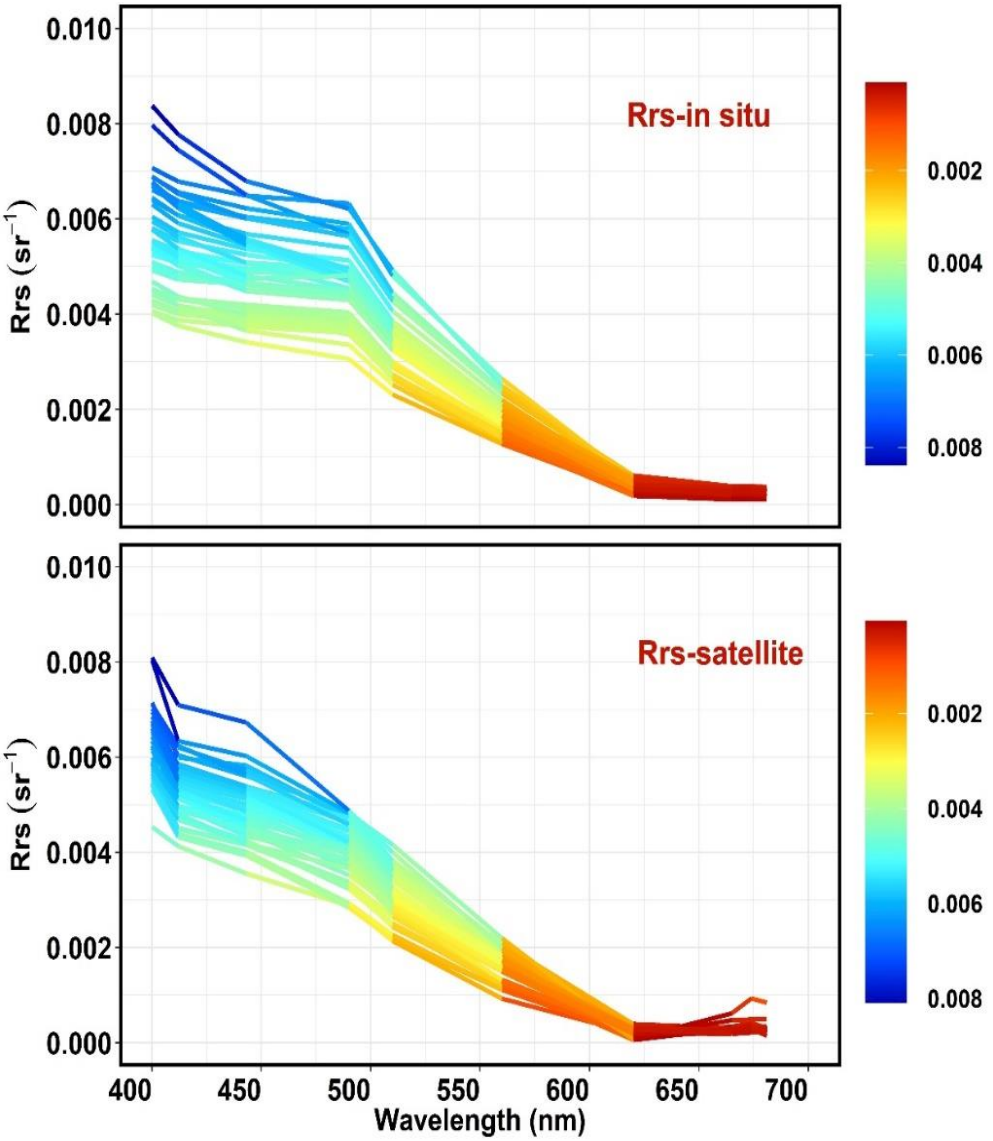

Figure 2. The diagram illustrates the simulated remote sensing reflectance spectra from the in situ hyperspectral remote sensing reflectance (Rrs in situ) in the top panel and the corresponding OLCI remote sensing reflectance spectra (Rrs satellite) in the bottom panel. Rrs satellite represents the median value from a 5×5-pixel window.

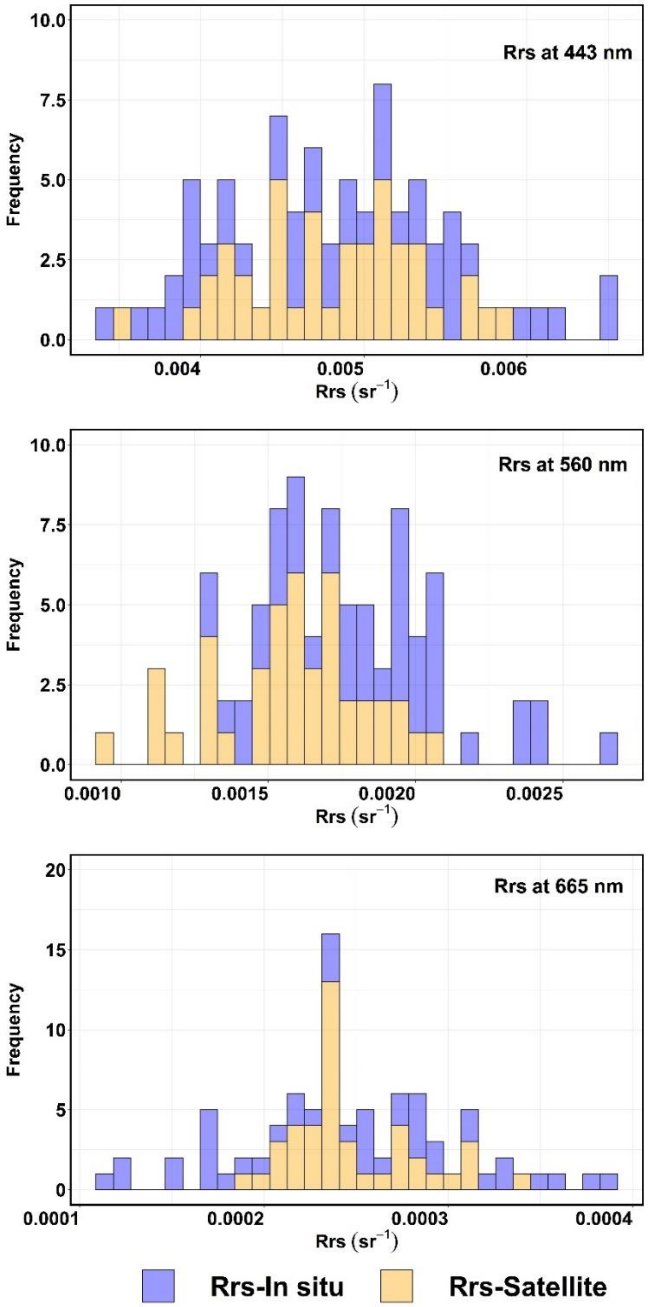

Figure 3. The histogram illustrates the simulated remote sensing reflectance spectra (Rrs in situ) at 443, 555, and 665 nm as well as the corresponding OLCI remote sensing reflectance spectra (Rrs satellite).

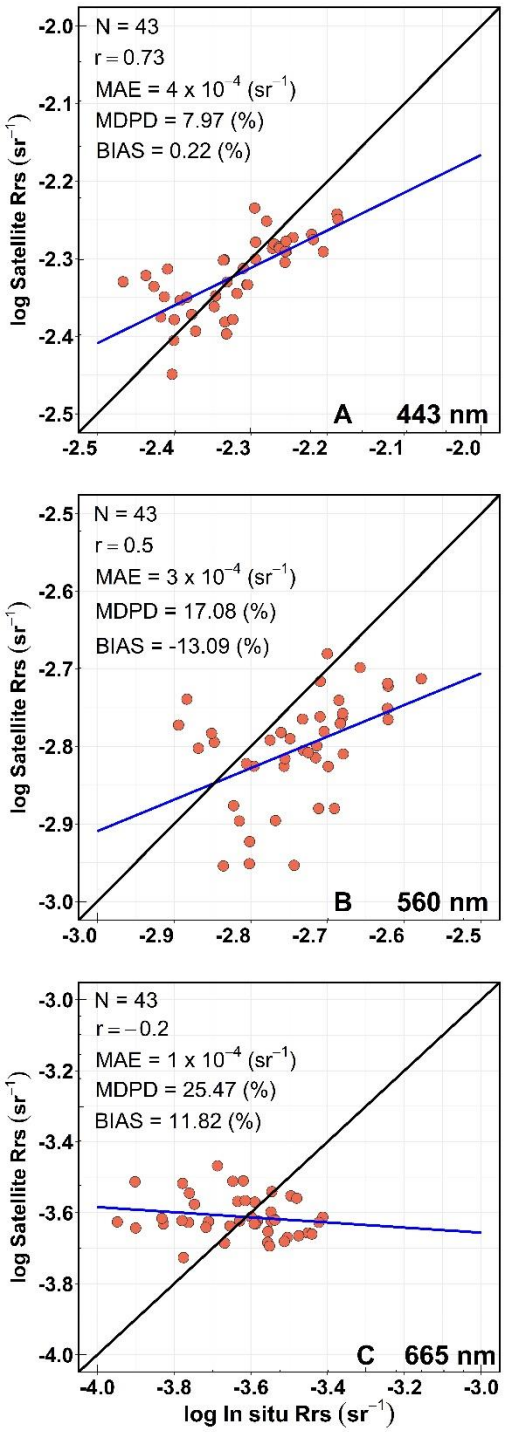

Figure 4. Regression between in situ (*x*-axis) versus satellite-derived (*y*-axis) remote sensing reflectance at 443, 560, and 665 nm. The black line represents the 1:1 line, and the colored line is the linear regression fit for the atmospheric correction algorithm. Satellite Rrs represents the median value from the 5×5 pixel-window.

2.4.2 Total Chlorophyll-a Matchup Analysis

The dynamic range of in situ and OLCI-derived TChla corresponds to typical conditions for the region (Peña and Varela, 2007). The in situ TChla ranged from 0.17 to 0.52 mg/m^3 (0.35 ± 0.08 mg/m^3) (**Figure 5**), while the satellite-derived TChla ranged from 0.21 to 0.59 mg/m^3 (0.38 ± 0.08 mg/m^3) (**Figure 5**). One-to-one matchup analysis (**Figure 6**) showed excellent agreement between in situ and satellite-derived TChla (r = 0.78, MAE = 0.05 mg/m^3, MDPD = 9.67%, BIAS = 10.38%; **Figure 6**). This indicates the ability of the POLYMER AC algorithm to retrieve TChla concentration within the expected range for this region. However, a slight overestimation of in situ TChla concentration by POLYMER was noted in the range between 0.17 and 0.28 mg/m^3. The matchup analysis showed that, except for a few points, most of the points fall close to the 1:1 line, indicating a strong correlation between in situ and satellite-derived TChla, which results in better statistical metrics (**Figure 6**). Therefore, the matchup analysis noted a low MDPD (9.67%) and BIAS (10.38%). At the same time, MAE depends on the concentration, which was comparatively lower (MAE = 0.05 mg/m^3).

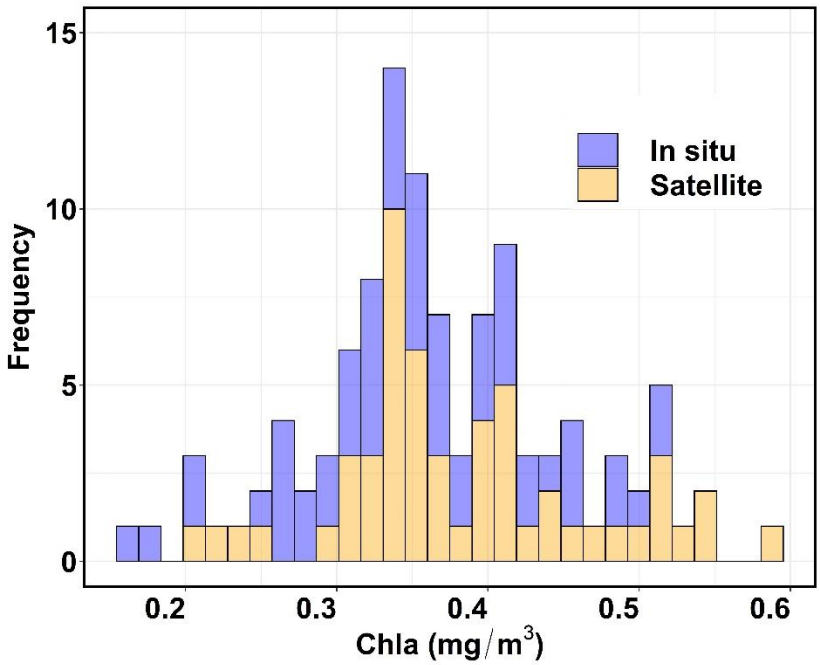

Figure 5. Histogram showing the in situ total chlorophyll-a concentration and the corresponding satellite-derived total chlorophyll-a concentration for the combined dataset of 2019 and 2020.

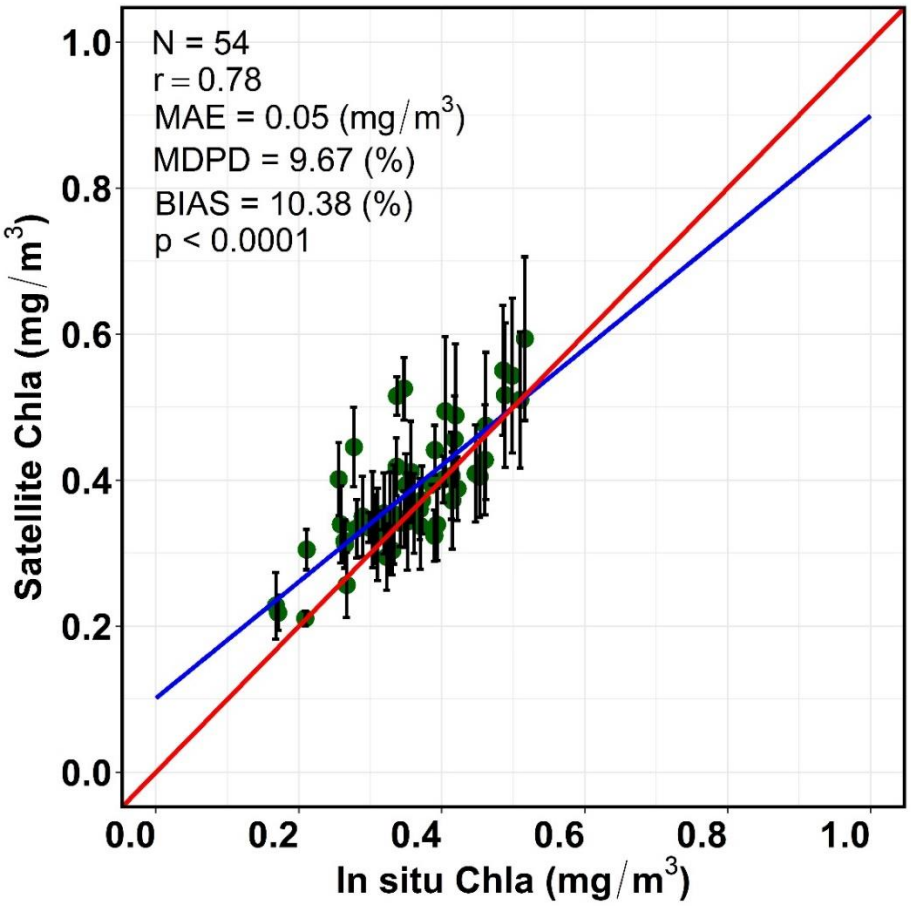

Figure 6. The comparison between in situ versus satellite-derived total chlorophyll-a concentration (TChla) for the combined dataset of 2019 and 2020. The red line shows the 1:1 line and the blue line is the linear regression fit for TChla derived from the POLYMER atmospheric correction algorithm. The standard deviation of valid pixels within a 5×5-pixel window centered on the location of the in situ observation is represented by vertical bars. The satellite-derived TChla concentration were obtained from weekly-composite imagery.

2.4.3 Spatial Variability of Sentinel-3A OLCI-Derived Total Chlorophyll-a

The quality of the satellite-derived TChla was further evaluated to determine its ability to capture the known pattern of surface TChla in SNEP over a spatial–temporal domain. Overall, the weekly composites indicated the spatially homogeneous distribution of surface TChla concentration in the iron-poor open ocean waters of GoA (**Figure 7**). Temporally, the average TChla from the two expeditions remained almost invariant (avg < 0.6 mg/m^3; **Figure 7**) throughout the low productive winter waters, with values ranging from 0.3 to 0.7 mg/m^3. In addition, no clear spring bloom signal was noticeable in this region; instead, patches of moderate TChla values (~0.6 mg/m^3) were observed. However, these values fall within the dynamic range observed for this region. On the other hand, a substantial increase in concentration (≥0.7 mg/m^3) was recorded toward SEA and the BC coast (**Figure 8**). For

example, the latitudinal trend of TChla derived from each week (**Figure 8**) showed elevated TChla at 57°N. Similarly, TChla derived from the longitudinal transect also indicated a consistently high value toward −135°W throughout the weeks.

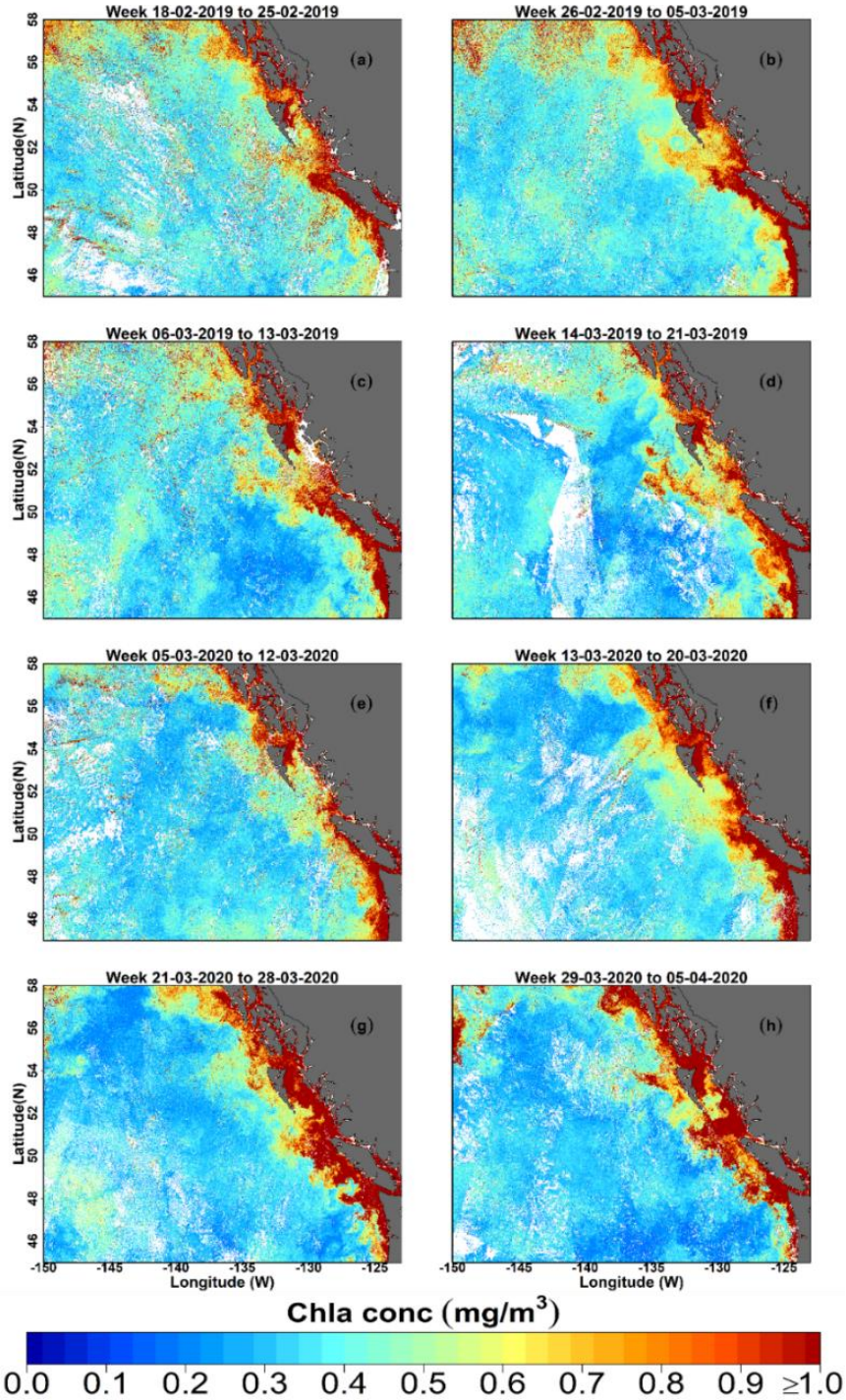

Figure 7. Spatial distribution of the weekly binned surface total chlorophyll-a concentration over the subarctic northeast Pacific for 2019 and 2020. (a-d) spatially and temporally binned surface total chlorophyll-a for the week of 18-2-2019 to 21-3-2019. (e-h) spatially and temporally binned surface total chlorophyll-a for the week of 05-3-2020 to 05-04-2020. White regions in the map represent clouds.

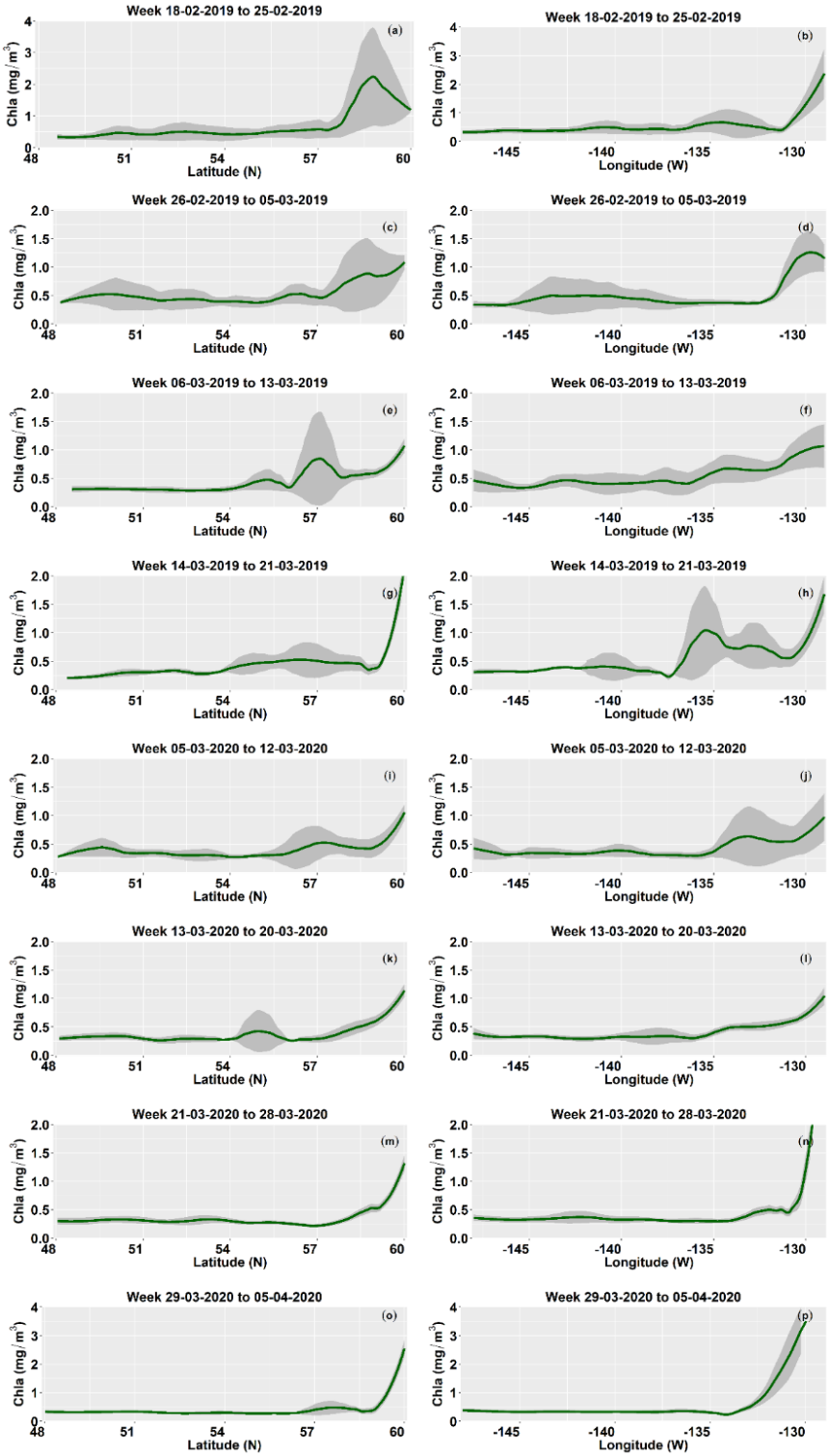

Figure 8. Latitudinal and longitudinal distribution of surface total chlorophyll-a concentration extracted from each weekly composite imagery for 2019 and 2020. The TChla concentration is extracted from a 5×5-pixel window, and the standard deviation is plotted as a gray-shaded region. (a-h) latitudinal and longitudinal distribution of surface total chlorophyll-a for the week of 18-2-2019 to 21-3-2019. (i-p) latitudinal and longitudinal distribution of surface total chlorophyll-a for the week of 05-3-2020 to 05-04-2020. The corresponding longitudinal and latitudinal lines are displayed in **Figure 1**.

Spatially, regions of high TChla concentration and pronounced seasonal spring blooms were observed along the coast from 2019 to 2020 (**Figure 7**). For instance, patches of elevated TChla concentrations were marked in northern and SEA, the east coast of Haida Gwaii, and the west coast of Vancouver Island. In the 2019 composites, high TChla concentration patches were defined in the third week of February (18 to 25 February), mainly along northern and SEA, the east coast of Haida Gwaii, northern Vancouver Island, and the west coast of Vancouver Island. Regions of high TChla concentration persisted during the final week of February (26 February to 5 March 2019), with wider spatial distribution mainly along northern and SEA, the east coast of Haida Gwaii, the north and west coast of Vancouver Island, and the Strait of Georgia (SoG), with the highest recorded TChla of 5.8 mg/m^3. Similarly, localized regions of high TChla concentration with larger spatial dispersion were prominent along the west coast from the second (6 to 13 March 2019) to the third (13 to 21 March 2019) weeks of March, with the highest recorded TChla of 6.9 mg/m^3 observed during the third week of March.

On the other hand, during the first week of March 2020 (5–12 March), high TChla concentrations were observed in multiple regions, including SEA, the east coast of Haida Gwaii, the west coast of Vancouver Island, and SoG, with a peak TChla concentration of 6.6 mg/m^3. However, in the third week (13–20 March), a more extensive distribution of high TChla regions was noted, exhibiting a pattern similar to that observed during the first week. Additionally, the highest TChla value (8.9 mg/m^3) among the four composites was recorded that week. Subsequently, from the fourth week of March (21–28 March) to the first week of April (29 March–5 April), a widespread distribution of high TChla values was observed along the coast from SEA to the BC coast. Furthermore, the highest TChla concentrations recorded during the fourth week of March and the first week of April were at 8.0 and 8.3 mg/m^3, respectively.

2.5 Discussion

This study provides the first validation of Sentinel-3 OLCI Rrs(λ) and TChla products in relation to in situ measurements of the SNEP region. The validation results of the Rrs satellite revealed superior radiometric performance for the blue and green bands and inferior performance for the red band. However, excellent agreement between in situ and satellite-derived TChla concentration was demonstrated in TChla validation. Furthermore, satellite-derived weekly TChla composites displayed relatively low (<0.5 mg/m^3) and spatially

homogenous distribution in the ion-poor open ocean waters, while elevated TChla values were (≥0.7 mg/m^3) observed toward SEA and BC. The following sections present the analysis of our main findings.

2.6 Evaluation of the Satellite-Derived Products

2.6.1 Remote Sensing Reflectance

We evaluated the ability of the POLYMER and ESA BAC algorithm to retrieve Rrs Satellite for determining TChla in the SNEP using Rrs in situ and TChla data. Our findings indicate that Rrs satellite and Rrs in situ follow spectrally similar shapes and have highly comparable average values, indicating the ability of POLYMER AC schemes to derive Rrs in optically Case-1-type water. Additionally, the validation effort to determine the retrieval accuracy of POLYMER AC demonstrated that retrievals improved with decreasing wavelength. For example, POLYMER showed better radiometric performance in the blue and green spectral regions, whereas an inferior radiometric performance was obtained in the red region (**Figure 4).** Moreover, the POLYMER slightly over and underestimated Rrs values between 0.0012 and 0.0018 sr^{-1} in the green band, whereas it fails to represent Rrs in the red band. In their latest study, Tilstone et al. (2021) observed that the residual error in POLYMER-derived Rrs values varies across different spectral bands, with relatively low mean residual error at the blue (6%) and green (15%) bands, followed by high mean residual error at the red band (60%), which is consistent with our observation. This variation in residual error is likely due to imperfections in AC and uncertainty of in situ measurements. For instance, satellite sensors often detect weak signals at longer wavelengths, such as at 560 and 665 nm from clear Case-1 waters (Hu et al., 2013); thus, AC remains highly challenging in this region compared to other spectral regions, eventually leading to over- and underestimation of the lower reflectance values (Zibordi et al., 2012; Gilerson et al., 2022). Until now, there has been no established method to remove the residual error from satellite-derived Rrs, which for the clear open ocean waters is observed to be more pronounced in the red band, accounting for up to 60%, followed by the green band (15%) (Tilstone et al., 2021). Consistent with the above findings, the latest study by Gilerson et al. (2022) assessed uncertainties in satellite-derived Rrs from the open ocean and found the highest CV of 1.5–2.0 in the red band compared to the blue band.

Beyond the performance of satellite-derived products, the quality of in situ data is crucial for the accurate validation of satellite data. However, in situ data carry substantial uncertainty for

multiple reasons, including inherent sensor noise combined with experimental and environmental factors (Mobley, 1999; Hooker and Morel, 2003; Ruddick et al., 2019; Vabson et al., 2019a; Vabson et al., 2019b; Tilstone et al., 2020). For instance, uncertainties in above-water radiometry can arise from sensor calibration, ship shadow and superstructure, environmental conditions, and parameterization of sky glint (Hooker and Morel, 2003). Nevertheless, strict guidelines are followed to acquire high-quality radiometric measurements, such as regarding the use of calibrated sensors, performing data collection only during solar noon conditions, and a setup to eliminate ship shadow and superstructure (Mobley, 1999; Hooker and Morel, 2003). According to a recent study, environmental variability (e.g., wind speed, sun zenith, sky condition, wave height, and sunglint) is the primary source of uncertainty across all wavelengths (410, 490, 550, and 620 nm), accounting for up to 80% of the total (Lin et al., 2022), wherein the environmental variability, sky condition, wind speed, and the sun zenith angle contribute to the highest uncertainty (Lin et al., 2022). In our study, despite conducting data acquisition during solar noon, we encountered several days with overcast conditions during the cruise, which could have contributed considerable uncertainty to the Rrs(λ) measurements. This has been documented in a recent study, which found that uncertainties of up to 80% can occur in the blue (410 and 490 nm) and green (550 nm) bands and up to 60% in the red band (620 nm) under overcast conditions (Lin et al., 2022). These conditions, characterized by scattered clouds and limited sun illumination, cause substantial variability in sunlight, leading to significant deviations in Li(λ) (Lin et al., 2022). In addition, Lt(λ) is also affected, as it receives a significant amount of reflectance from different parts of the sky, and its measurements can be unpredictable due to variations in wave height over time (Lin et al., 2022). Finally, radiometric measurements were also affected by the viewing and illumination geometry, which can be addressed by correction for bidirectional effects (BRDF) (Morel and Gentili, 1996). Therefore, it is important to take caution and follow strict guidelines while acquiring high-quality in situ radiometry data for validating OC products.

2.6.2 Total Chlorophyll-a Concentration

Despite OLCI slightly overestimating lower-range TChla values (0.17 to 0.28 mg/m3) in comparison to in situ measurements, the validation results demonstrate the excellent performance of the POLYMER AC algorithm in this region, with dynamic ranges consistent with in situ values (Marchese et al., 2022). Our observation is supported by Giannini et al.

(2021), who evaluated the performance of POLYMER and other AC schemes in the northeast Pacific region and found that POLYMER provided the best results for TChla and that the dynamic range was within the observed range (Peña et al., 2019). This finding is also substantiated by Marchese et al. (2022), who showed that the monthly climatology of TChla derived from OLCI using POLYMER recorded low biomass (≤0.5 mg/m3) in offshore waters of GoA. In contrast to these studies, our findings were limited to the winter and early spring of 2019 and 2020, while Giannini et al. (2021) utilized seasonal TChla climatology data (spring, summer, fall) from 2016 to 2018, and Marchese et al. (2022) utilized similar climatological map from 2016 to 2020. Similar to our findings, Tilstone et al. (2021) found that POLYMER is the most promising AC processor for retrieving TChla concentration between 0.01 and 1.00 mg/m^3 across the oligotrophic Atlantic Gyre, particularly under thin clouds and sunglint, which may affect the TChla retrieval.

Nonetheless, the slight discrepancy observed between satellite and in situ-derived TChla values may arise from multiple sources. For instance, in our study, the satellite data typically represent the median value within the 5×5-pixel window extracted from weekly composite imagery, whereas in situ data represent point location. Satellite matchups using weekly composite extracted from a 5×5-pixel window are accepted in oligotrophic waters due to the low variability in phytoplankton biomass within 1500×1500 m spatial resolution (Bailey and Werdell, 2006). This is relevant for the SNEP because of only subtle variability in the phytoplankton community of the GoA compared to those of the highly dynamic SEA and BC shelf regions (Zhang et al., 2021). Furthermore, we also found that spatial-temporal mismatch might not have played a substantial role, as we did not observe any notable difference between the average values of satellite-derived TChla from the weekly composite (0.38 ± 0.08 mg/m^3) versus the daily imagery (0.39 ± 0.08 mg/m^3). Another potential discrepancy arises from the tendency of satellite-based measurements to generally overestimate the surface TChla values when a deep chlorophyll maximum (DCM) is present (Sá et al., 2015). In a recent study, it was found that TChla values integrated over the first optical depth are 30% higher than surface values (Sá et al., 2015), a finding relevant for the SNEP, but only observed during summer conditions (Meyer et al., 2022). However, this discrepancy may not have any impact on our study, which is focused on winter conditions.

Beyond the performance of the atmospheric correction algorithm, the validation performance was impacted by the uncertainty associated with the analytical quality of in situ HPLC

pigment measurements. Despite the widespread use of HPLC-derived TChla for the validation of satellite-derived TChla (Bracher et al., 2017), these measurements are often subject to multiple sources of errors (Stanford and Hooker, 2011). For instance, a round-robin experiment conducted by using multiple HPLC methods to quantify the uncertainty budget revealed a 7% error for TChla and a 21% error for other pigments. In this study, the overall uncertainty associated with the HPLC pigment data, including sampling methods, filtration, extraction, and HPLC analysis, based on evaluating duplicates of samples, resulted in a CV lower than 20%. However, as in many studies (Claustre et al. 2004; Ribalet et al., 2010), it is difficult to quantify any remaining source of error attributed to the in situ TChla measurements.

2.7 Spatial–Temporal Dynamics of Phytoplankton Biomass

Although matchup statistics were obtained, the evaluation of the overall quality of satellite-derived TChla remains challenging due to the sources of uncertainty. Here, we further assessed the quality of the satellite-derived TChla products by comparing the latitudinal and longitudinal gradient from weekly composite imagery with the previous literature. Previous remote sensing observations used the Coastal Zone Color Scanner (CZCS), Moderate Resolution Imaging Spectroradiometer (MODIS-Aqua), and Sentinel 3A imagery, which indicated low concentrations (<0.5 mg/m^3) in the iron-poor open ocean waters of the GoA, with no significant seasonal or interannual variability noted; the average TChla value fell well within the range reported for satellite observations (Vinogradov et al., 1997; Banse and English, 1999; Brickley and Thomas, 2004; Jackson et al., 2015; Westberry et al., 2016; Marchese et al., 2022). Additionally, simulations from the coupled Regional Ocean Modeling System (ROMS) model corroborated our finding of low TChla values within the less productive GoA waters, with a discernible gradient in biomass observed toward the continental shelf domain (Peña et al., 2019). Furthermore, our results aligned with recent model observations, indicating that TChla concentration in the iron-poor GoA waters remained consistently low (≤ 0.4 mg/m^3) throughout the year, with an annual mean peak occurring only in the fall (Zhang et al., 2021). In line with satellite and model findings, field-level observations also revealed low temporal variability in the upper TChla concentration in this region, without apparent seasonal or interannual variability (Peña and Varela, 2007; Peña et al., 2019). These findings were confirmed by the latest observation using the BGC-Argo float, which indicated less variability in TChla concentration in the upper 150 m, consistent

with the satellite observation (Zhang et al., 2021). As an HNLC region, iron primarily governs the bloom dynamics, and iron limitation restricts the occurrence of diatom bloom, which is a common phenomenon for the shelf region (Ribalet et al., 2010). Another reason for the low TChla concentration in this region is the strong coupling between the phytoplankton and zooplankton communities, where the latter effectively grazes on the former, resulting in a relatively constant TChla concentration throughout the year (Landry et al., 1993).

Unlike the open waters, our weekly composite maps showed high TChla concentrations (i.e., between 5.4 and 8.9 mg/m^3) and pronounced spring bloom signals along SEA to the BC coast (**Figure 7**). Consistent with our findings, previous studies reported relatively high TChla levels (>3.0 mg/m^3) along with pronounced spring bloom characterized by marked seasonal, spatial, and interannual variability in this region (Brickley and Thomas, 2004; Waite and Mueter, 2013; Jackson et al., 2015; Strom et al., 2016). The latest satellite-based observations also substantiate these findings by recording high spring average TChla (>2.0 mg/m^3) across SEA to the BC coast (Westberry et al., 2016; Marchese et al., 2022; Pramlall et al., 2023).

Generally, our composite imagery from 2019 displayed a larger spatial extent of variation in TChla concentration along the coast from the last week of February to the third week of March. Conversely, the imagery from 2020 revealed a larger spatial extent of higher TChla concentration along the coast from the second week of March to the second week of April. These findings are consistent with recent satellite-based observations of the BC coast, which have found that, in terms of timing, the spring bloom along the west coast of Vancouver typically starts in March (Peterson et al., 2007; Ribalet et al., 2010; Westberry et al., 2016). The latest study utilizing high spatial and temporal resolution Sentinel 3A imagery found that the seasonal increase in average TChla concentration from 2016 to 2020 across SEA to the BC coast region occurs between late February and early April (Marchese et al., 2022), consistent with our findings. In addition, our weekly composite imagery revealed regions of high TChla values during spring, which were concentrated in northern and SEA, the east coast of Haida Gwaii, and along the north, south, and west coasts of Vancouver Island, as well as in the SoG. These zones were identified as bioregions in the northeast Pacific, confirming our findings and demonstrating that they are characterized by high TChla values (>2.4 mg/m^3) and exhibit an annual peak in spring (Marchese et al., 2022). The mechanism by which elevated TChla levels and spring blooms are sustained across the coastal waters of SEA and the BC is attributed to several processes, including iron-rich coastal freshwater

runoff, strong upwelling, tidal mixing, high levels of irradiance, the lengthening of daylight hours, and stratification of the water column (e.g., Landry et al., 1993; Harrison et al., 1999; Childers et al., 2005; Marchetti et al., 2006; Strom et al., 2006; Henson, 2007; Ribalet et al., 2010; Strom et al., 2016; Suchy et al., 2019).

2.8 Conclusion

In this study, we aimed to investigate the performance of Sentinel-3A OC products, specifically Rrs(λ) and TChla concentrations, against in situ data. Additionally, we conducted a qualitative assessment with reference to previous studies to determine whether the weekly binned surface TChla concentration reflects the general latitudinal and longitudinal trends in this region. The in situ and satellite-derived Rrs(λ) revealed a Case-1 water spectrum with maximum reflectance in the blue band that exponentially decreased toward longer wavelengths, followed by TChla fluorescence at 680 nm. The validation of the Rrs satellite against Rrs in situ showed there was a good correlation and better statistical performance in the blue and green regions, with a degraded performance recorded in the red band. Moreover, POLYMER tends to overestimate or underestimate Rrs levels in the green band, specifically within the range of 0.0012 to 0.0018 sr^{-1}, where POLYMER fails to represent the Rrs in the red bands. TChla matchup analysis showed excellent agreement between in situ and satellite-derived TChla. Furthermore, POLYMER slightly overestimates the lower ranges of in situ TChla from 0.17 to 0.28 mg/m^3. The spatial and temporal dynamics of weekly binned surface TChla maps show low biomass in the iron-poor open ocean water; however, a substantial increase in TChla (≥0.7 mg/m^3) was recorded toward the SEA and BC shelf. Furthermore, the latitudinal and longitudinal distribution of TChla reveals high biomass toward 57° North and −135° West, respectively.

High spatial and temporal resolution TChla maps are essential for this region to better understand the dynamics of surface phytoplankton, which are the foundation of the marine food web (Malick et al., 2015) reported that the spring bloom phenology in the northeast Pacific Ocean is significantly correlated with Pacific salmon productivity in Alaska and BC. Similarly, Suchy et al. (2022) demonstrated the connection between phytoplankton bloom dynamics and zooplankton phenology in the coastal waters of BC, revealing that early and late spring blooms result in mismatches between phytoplankton and zooplankton phenology, with implications for the food web. In this context, our findings could help fishery management by providing spatial and temporal data on phytoplankton bloom in the SNEP

region, where Pacific salmon spend most of their adult life (Beamish, 2017), enabling the better management of fishery resources across this region. In addition, our study represents the first attempt to utilize high spatial (300 m) and temporal (daily) resolution satellite imagery to validate Sentinel-3 OLCI products in the undersampled oligotrophic region of the SNEP. The results show the potential of OLCI to retrieve the OC products as well as the spatial–temporal pattern of TChla concentration over a synoptic scale, which would not have been possible with discrete water samples.

Chapter 3: Assessing the Spatial-Temporal Variability of Phytoplankton Community Composition and Their Response to Environmental Drivers in the Subarctic Northeast Pacific by Combining In Situ and Satellite Observations

Perumthuruthil Suseelan Vishnu[1], Justin Del Bel Belluz[2], M. Angelica Peña[3], Brian P.V. Hunt[4], Anna Vazhova[5], Midhun Shah Hussain[6], Hongyan Xi[7], Astrid Bracher[7,8], Maycira Costa[1]

[1]SPECTRAL Remote Sensing Laboratory, Department of Geography, University of Victoria, Victoria, British Columbia, Canada.

[2]Hakai Institute, Victoria, British Columbia, Canada.

[3]Fisheries and Oceans Canada, Institute of Ocean Sciences, Sidney, British Columbia, Canada.

[4]Institute for the Oceans and Fisheries and Department of Earth, Ocean, and Atmospheric Sciences, University of British Columbia, Vancouver, British Columbia, Canada.

[5]Pacific Branch of the Federal State Budget Scientific Institution of Fisheries and Oceanography, Vladivostock, Russia.

[6]Department of Marine Biology, Microbiology, and Biochemistry, Cochin University of Science and Technology, Kerala, India.

[7]Phytooptics Group, Physical Oceanography of Polar Seas, Climate Sciences, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany.

[8]Department of Physics and Electrical Engineering, Institute of Environmental Physics, University of Bremen, Bremen, Germany.

Abstract

The subarctic Northeast Pacific (SNEP) is characterized as a high nutrient, low chlorophyll (HNLC) region where primary productivity and phytoplankton growth are limited by bioavailable iron (BFe). Here, we investigate the spatial-temporal distribution and drivers of surface phytoplankton community composition within this HNLC region in winter using in situ environmental data, chemotaxonomic analysis (CHEMTAX) of phytoplankton pigment samples, and Sentinel-3A ocean color imagery. Latitudinal gradients were observed with warm, low-salinity subtropical water to the south and colder, high-salinity subarctic water to the north. Satellite and CHEMTAX-derived phytoplankton community composition showed homogeneous communities across this gradient dominated by haptophytes, green algae, and pelagophytes; however, distinct cryptophyte increases were observed at the northern extents of the study region in 2019, potentially driven by the presence of the Alaska coastal current and/or Sitka and Yakutat eddies. In 2020, when shelf waters were also sampled, longitudinal gradients were observed with increased contributions of diatoms in the coastal/shelf waters off of Haida Gwaii. Redundancy analysis showed negative correlations between diatoms and sea surface salinity and mixed layer depth suggesting coastal influences with potentially higher BFe. Cryptophytes were correlated with increased dissolved silicate and reduced sea surface temperature, but these trends were likely a result of similar latitudinal gradients rather than driver-species relationships. This study revealed the potential of Sentinel-3 Ocean Land Color Instrument (OLCI) to provide higher temporal and spatial resolution data on phytoplankton community composition required for understanding trophic linkages, match/mismatch dynamics, and better management of fisheries resources.

3.1 Introduction

Phytoplankton play crucial roles in the functioning of the global ocean as the base of the marine food web, exporting carbon to the deep ocean and thereby driving the biological carbon pump, nutrient cycling, and elemental exchange with the atmosphere (e.g., Le Quéré et al., 2005; Hood et al., 2006). These biogeochemical functions were strongly coupled with the composition and size structure of phytoplankton communities (Le Quéré et al., 2005). For example, microphytoplankton such as diatoms were the predominant silicifiers in the ocean responsible for transporting carbon, nitrogen, and silicate into the deep waters (IOCCG, 2014). Furthermore, phytoplankton size structure often exhibits a direct link with fisheries; in the Benguela upwelling ecosystem, for example, small-sized flagellate communities supported the growth of sardines, whereas diatoms supported the growth of anchovies (Cury et al., 2008). As a result, shifts in phytoplankton community composition and size structure can have significant implications for oceanic elemental cycling and the transfer of energy to higher trophic communities, including commercially and culturally important fish populations (e.g., Legendre, 1990; Vanni & Findlay, 1990; Beardall & Raven, 2004; Le Quéré et al., 2005; Beardall & Stojkovic, 2006; Falkowski & Oliver, 2007; IOCCG, 2014; Stock et al., 2014; Petrou et al., 2016; Dutkiewicz et al., 2020). Therefore, monitoring phytoplankton community composition in global and regional ocean waters is paramount to inform how phytoplankton respond to changing climate and their potential implication on fisheries production and ocean biogeochemistry.

Understanding phytoplankton community dynamics is especially important for the subarctic northeast Pacific (SNEP), which is considered a high nutrient, low chlorophyll (HNLC) region where phytoplankton growth is limited by bioavailable iron (BFe) (Harrison et al., 2004; Peña & Varela, 2007). In this region, total Chlorophyll-a (hereafter TChla, a proxy for phytoplankton biomass) is generally low (< 0.5 mg/m^3) and invariant, despite marked seasonality in nutrient concentrations and primary productivity (e.g., Evans & Parslow, 1985; Vinogradov et al., 1997; Peña & Varela, 2007; Westberry et al., 2016; Peña et al., 2019); however, sporadic blooms can occur due to the introduction of BFe into the mixed layer (Hamme et al., 2010). Beyond the availability of BFe, phytoplankton growth and production in the subarctic gyre largely depend on irradiance, vertical mixing, and grazing pressure (Harrison et al., 2004; Henson, 2007; Dagg et al., 2009; Aguilar-Islas et al., 2016). In winter, strong vertical mixing brings inorganic nutrients (nitrate and silicate) into the mixed layer,

but at the same time, transports phytoplankton below the euphotic depth and therefore decreases primary productivity (Harrison et al., 2004; Peña & Varela, 2007).

Annual primary productivity in the SNEP is estimated to be 170 g C m^2/y (Welschmeyer et al., 1993), and the community composition that makes up this productivity has been investigated using various methods. For example, studies have utilized microscopy, high-performance liquid chromatography (HPLC)-derived phytoplankton pigment concentrations and corresponding chemical taxonomy (CHEMTAX), and physical-biological models to elucidate phytoplankton community composition in this region (e.g., Booth et al., 1993; Harrison et al., 2004; Fujiki et al., 2009; Peterson et al., 2011; Peterson & Harrison, 2012; Peña et al., 2019; Zhang et al., 2021). Among the numerous in situ techniques used in the SNEP, conventional light microscopy excels at identifying phytoplankton to either genus or species level and is effective for enumeration and identification of microphytoplankton (> 10 µm), but generally misses smaller-sized species that are regionally important (Goela et al., 2014; IOCCG, 2014; Brito et al., 2015). Conversely, in the SNEP, HPLC-derived phytoplankton pigment data and corresponding CHEMTAX analysis have been widely used to characterize phytoplankton community composition at a broad functional group level (Fujiki et al., 2009; Peterson & Harrison, 2012; Peña et al., 2019; Taves et al., 2022) and can resolve groups containing small nano- and pico-sized species (< 10 µm) (Vaillancourt et al., 2018; Del Bel Belluz et al., 2021). Despite the utility of these in situ methods, they are generally time-consuming, labor-intensive, expensive, and, most importantly, limited in space and time (IOCCG, 2014). Thus, it is necessary to combine earth observation satellites and in situ techniques to address the spatially and temporally limited phytoplankton community composition data gaps in the SNEP. This has been done relatively effectively on the west coast of Canada, where CHEMTAX and daily Sentinel-3 OLCI imagery were employed for in situ and satellite-based estimates of phytoplankton group-level Chla, respectively (Vishnu et al., 2022).

Data acquired by earth observation satellites offers an opportunity for retrieving phytoplankton biomass from the upper global ocean over high spatial and temporal resolutions. For example, over the last four decades, ocean color satellites have enabled the quantification of surface TChla (e.g., McClain, 2009) on unprecedented spatial and temporal scales. Additionally, satellite-based methods have been developed for determining phytoplankton groups and phytoplankton size classes in the ocean using abundance-based

(e.g., Uitz et al., 2006; Hirata et al., 2011; Brewin et al., 2010, 2015), radiance/reflectance-based (e.g., Alvain et al., 2005; Bracher et al., 2009; Werdell et al., 2014; Xi et al., 2020), and backscattering-based methods (e.g., Kostadinov et al., 2009, 2010). An alternative satellite-based approach employs Empirical Orthogonal Function (EOF) analysis on remote sensing reflectance Rrs(λ) and derived accessory pigment concentrations (e.g., Lubac & Loisel, 2007; Craig et al., 2012; Bracher et al., 2015; Soja-Woźniak et al., 2017), Chla concentrations of multiple phytoplankton groups (e.g., Xi et al., 2020; Vishnu et al., 2022), or phytoplankton community abundances (e.g., Lange et al., 2020). Previous studies have used EOF algorithms to retrieve phytoplankton groups and accessory pigments from multispectral reflectance data using a compilation of global dataset (e.g., Xi et al., 2020, 2021); however, these global algorithms often have poor performance in specific geographical regions (Xi et al., 2020). Therefore, geographically focused studies require regionally tuned algorithms combined and validated with in situ data to better understand the fine-scale spatial and temporal variability of phytoplankton groups (Vishnu et al., 2022).

This study combines various techniques to investigate the spatial-temporal dynamics of multiple phytoplankton groups in the surface waters of the SNEP in winter. The specific objectives of this research were to: (1) train an EOF-based algorithm to characterize phytoplankton community composition from Sentinel-3 OLCI imagery and validate the derived outputs against independent CHEMTAX-derived phytoplankton community composition; (2) use a combination of field data and validated Sentinel-3 OLCI outputs to characterize the spatial and temporal patterns in phytoplankton community composition; and (3) investigate the physical and chemical drivers of the observed trends in the SNEP. This work is part of the International Year of Salmon (IYS) expedition, which aimed to investigate the environmental drivers determining the winter foraging conditions and the survival of Pacific Salmon in the GoA. Beyond spatial-temporal TChla concentrations (a proxy for phytoplankton biomass; McClain, 2009), characterizing phytoplankton community composition in the SNEP is crucial as this region provides critical habitat for pelagic fisheries, including Pacific Salmon (Beamish, 2017). More recently, Suchy et al. (2022) demonstrated that early/late spring blooms on the west coast of Canada resulted in a mismatch between phytoplankton and phenology of larger, energy-rich crustacean zooplankton. This mismatch resulted in lower zooplankton abundance and biomass available for the higher trophic levels. Therefore, the results of this study and the derived satellite methods are a baseline for future work investigating the link between phytoplankton

community composition, fish recruitment, and the factors that affect the survival of Pacific Salmon in the SNEP.

3.2 Materials and Methods

3.2.1 Specific to the Study Area

The SNEP is generally considered north of 45˚N and includes the open ocean basin, continental shelf, and coastal margins of Alaska and British Columbia (BC) (**Figure 9**). Within the SNEP, the ocean circulation and phytoplankton communities in the open ocean region of the GoA are distinct from the coastal and continental shelf margin of Alaska and BC (e.g., Booth et al., 1993; Peña & Varela, 2007; Fujiki et al., 2009; Weingartner et al., 2009; Peña et al., 2019). For example, the Alaska Gyre dominates the circulation in the open ocean, a component of the broad subarctic gyre in the North Pacific Ocean (Weingartner et al., 2002; Weingartner, 2006). This cyclonic gyre consists of the eastward-flowing North Pacific Current, the northwest-flowing Alaska Current, and the westward-flowing Alaska Stream (Stabeno et al., 2004; Weingartner et al., 2009). The circulation of the gyre is driven by counter-clockwise winds, which initiate surface Ekman transport leading to upwelling in the central GoA and downwelling towards the continental shelf margin (Weingartner, 2006). Generally, this region shows high phytoplankton diversity largely comprised of small (< 10 μm) haptophytes, green algae, and pelagophytes (e.g., Booth et al., 1993; Fujiki et al., 2009; Peña et al., 2019) that can survive in the iron-poor open ocean waters (Peña & Varela, 2007). Furthermore, the SNEP is crucial for commercial fisheries (Whitney et al., 2005), including Pacific Salmon originating from Canada, the United States, Russia, Japan, and Korea (Beamish, 2017). Productivity in this area is variable and linked to mesoscale eddies (Crawford, 2005), particularly Haida and Sitka eddies transporting heat, iron-rich coastal waters, and marine organisms to the open ocean waters of the GoA (Ueno et al., 2023). Moreover, climate forcings, such as El-Nino Southern Oscillations (ENSO) (Schwing et al., 2002), Pacific Decadal Oscillation (PDO), and the North Pacific Gyre Oscillation (NPGO) (Matua & Hare, 2002; Chavez et al., 2003; Di Lorenzo et al., 2008) play an important role in regulating year-to-year variability in nutrients and primary production.

Unlike the open ocean component, the continental shelf is dominated by the Alaska Coastal Current (ACC), which is driven by wind and freshwater runoff and is responsible for the advection of dissolved nutrients, phytoplankton, and other organisms across the shelf (Royer,

1983; Stabeno et al., 1995, 2016). This region is highly productive, with annual primary productivity exceeding 300 g C/m^2/y (e.g., Sambrotto & Lorenzen, 1987; Siddon et al., 2019). This high primary productivity supports numerous commercially important fish species, including Pacific Cod (*Gadus macrocephalus*), Walleye Pollock (*Gadus chalcogrammus*), Pacific Ocean Perch (*Sebastes alutus*), Pacific herring (*Clupea pallasii*), Arrowtooth Flounder (*Atheresthes stomias*) (Johnson et al., 2012; McGowan et al., 2019; Siddon et al., 2019; Budge et al., 2022), and various marine mammals (Womble & Sigler, 2006; Womble et al., 2009), and sea birds (Dragoo et al., 2017). Large phytoplankton, such as diatoms, dominate the coastal and continental shelf domains of the GoA and BC, with seasonal spring blooms reaching TChla concentrations > 3 mg/m^3 (e.g., Peterson & Harrison, 2012; Waite & Mueter, 2013; Jackson et al., 2015; Strom et al., 2016; Peña et al., 2019; Peña & Nemcek, 2020; Giannini et al., 2021; Marchese et al., 2022; Pramlall et al., 2023). Additionally, this region is highly dynamic, showing strong seasonal, spatial, and interannual variability in TChla (e.g., Brickley & Thomas, 2004; Waite & Mueter, 2013), with spring bloom timing varying spatially and interannually (Marchese et al., 2022).

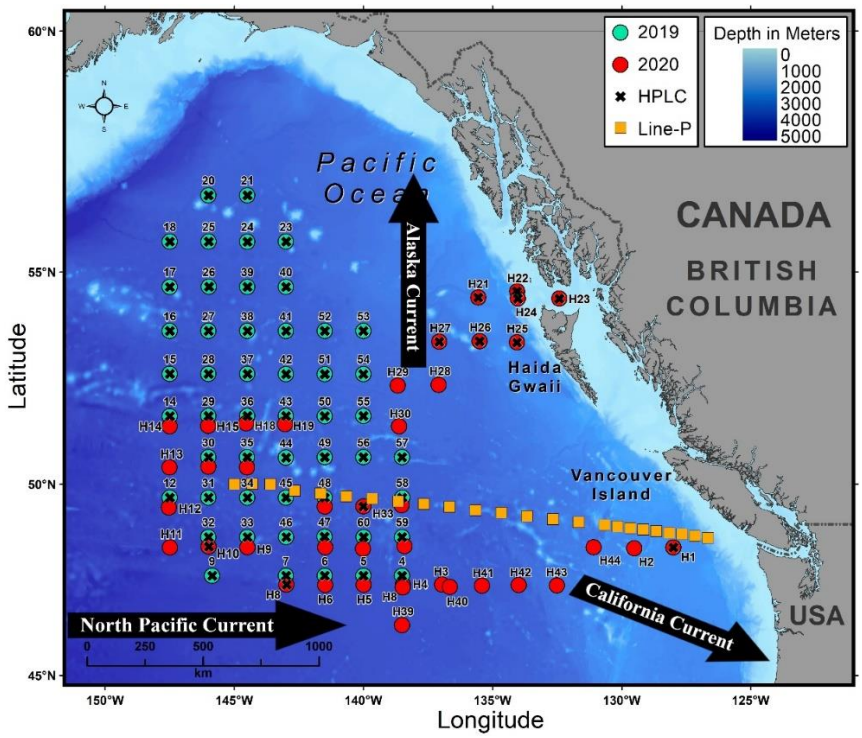

Figure 9. Map showing the study area and sampling locations in the subarctic northeast Pacific with a general view of the major currents: North Pacific Current, California Current, and Alaska Current. Green and red circles represent field stations for all physical and chemical variables in 2019 and 2020, respectively. The symbol "x" indicates high performance liquid chromatography (HPLC)-pigments field stations for 2019 and 2020. Yellow squares denote the Line-P monitoring stations sampled by Fisheries and Oceans Canada (DFO), Canada. Bathymetry is shown in meters.

3.3 Field Data

Data were acquired as part of the International Year of Salmon (IYS) Expedition to the GoA onboard the Research Vessel *Professor Kaganovsky* (18th February to 21st March 2019) and the Research Vessel *Pacific Legacy* (5th March to 5th April 2020). At each station, conductivity-temperature-depth (CTD) profiles were performed, and surface (~1–5 m depth) water samples for phytoplankton pigments macronutrient concentrations were collected using a CTD/rosette system equipped with Niskin bottles. The following section describes the methods for processing CTD, macronutrients, phytoplankton pigments, and satellite data. Then, the training and implementation of the EOF-based algorithms to retrieve phytoplankton community composition from weekly composite imagery are described along with the statistical methods performed to investigate their physical and chemical drivers.

3.3.1 Environmental Variables

Continuous CTD profiles for both expeditions were performed with a Seabird SBE 911plus system with dual temperature and conductivity (salinity) sensors and oxygen sensor. All the sensors were calibrated before the commencement of each expedition. Prior to analysis, the CTD data were aggregated into one-meter-depth bins. Potential density was calculated at each station from temperature and salinity using the Gibbs Sea Water Function from the R *gsw* package (Kelley & Richards, 2021). Potential density was then used to derive the mixed layer depth (MLD) by calculating the depth at which the potential density was 0.1 kg/m^3 greater than the surface value (Plant et al., 2016).

Samples for the determination of nitrate (NO_2), nitrite (NO_3), phosphate (PO_4), and silicate (SiO_4) were collected at each station and transferred into 500 ml polyethylene bottles until analysis was performed. During the 2019 campaign, analysis was conducted onboard using a Shimadzu UV-1800 spectrophotometer following standard protocols (Grasshoff et al., 1999). NO_2 and NO_3 concentrations in the submicromolar range were obtained using Wood et al. (1967) and phosphate, as described by Murphy & Riley (1962). During the 2020 campaign, samples for inorganic nutrients were stored at -20° C until analysis was performed onshore in the laboratory using a Lachat QuikChem 8500 Series 2 Flow Injection Analysis System, and analysis followed similar methods as in 2019.

3.3.2 Phytoplankton Pigments Concentration

Samples for the determination of HPLC-derived phytoplankton pigments were collected during both expeditions (N = 51 and N = 11 for 2019 and 2020, respectively). After collection, sample water was transferred to opaque 10 L plastic bottles and directly filtered under dim light to avoid pigment degradation (Mueller et al., 2003). Two liters of sample water was vacuum filtered (≤ 5 mmHg) through 25 mm Whatman glass microfibre filters (GF/F) with 0.7 μm pore size, and following filtration, the filters were folded inwards, placed in plastic vials, and flash frozen in liquid nitrogen. The samples were stored at -80˚ C until laboratory analysis was conducted at the University of South Carolina Baruch Institute of Marine and Coastal Sciences (https://phytoninja.com/lab-protocols/) following Pinckney (2010). Twenty-two pigments were identified, including TChla, chlorophyll c_3 ($chlc_3$), chlorophyll c_1c_2 ($chlc_{12}$), peridinin (peri), 19′butanoyloxyfucoxanthin (BF), fucoxanthin (fuco), 19′Hexanoyloxyfucoxanthin (HF), neoxanthin (neo), prasinoxanthin (pras), violaxanthin (viola), diadinoxanthin (diad), antheraxanthin (anther), alloxanthin (allo), diatoxanthin (diat), lutein (lut), zeaxanthin (zea), gyroxanthin-diester (gyro), chlorophyll-b (chlb), chlorophyll-a allomer (chla allomer), chlorophyll-a prime (chla′), α-carotene, and β-carotene. To assess uncertainty, duplicates were collected and analyzed for 20% (n = 11) of the samples (only available for 2019) with a coefficient of variation (CV) < 20%.

The HPLC-derived TChla and accessory pigment concentrations were further used as inputs for CHEMTAX analysis. CHEMTAX optimizes input accessory pigment:TChla ratios to those of the field data through an iterative process to derive estimates of the Chla concentration of multiple phytoplankton groups (e.g., Mackey et al., 1996). This approach has been used successfully in the region to investigate the distribution of dominant phytoplankton groups (Fujiki et al., 2009; Peterson et al., 2011; Peterson & Harrison, 2012; Peña et al., 2019). Before the CHEMTAX analysis, all of the HPLC-pigments stations were clustered based on their pigment:TChla ratios (**Figure S1,** stations input as rows and ratios as columns) to derive groupings of stations with similar pigment ratios (method described in section 3.3.5). These similar stations were then analyzed together to account for potential differences in pigment ratios due to light and environmental variability across the dataset that can skew outputs (Swan et al., 2016). Our cluster analysis resulted in three clusters: cluster-1 comprised 51 stations within the iron-poor open ocean waters that were entirely from 2019 and had relatively high BF and HF ratios with TChla; cluster-2 showed only 4 stations from

2020 that were from the southern open ocean portion of the study region and, when compared to cluster-1, had higher TChla ratios for most pigments, but still showed the highest ratios with HF and BF; cluster-3 encompassed 7 sites from the 2020 campaign representing continental shelf influenced stations that had notably lower HF and BF to TChla ratios (**Figure S2** and **Figure S3**). Due to the low number of samples in cluster-2, we were forced to combine it with cluster-1 for CHEMTAX analysis, but kept cluster-3 separate due to the distinct differences in pigment:TChla ratios. The following phytoplankton groups were added to the CHEMTAX analysis: cyanobacteria (Cyano), haptophytes (Hapto), pelagophytes (Pelago), prasinophytes (Prasino), chlorophytes (Chlor), Cryptophytes (Crypto), dinoflagellates (Dino), and diatoms (Diatoms). The pigments utilized in the analysis were chlc3, $chlc_1c_2$, peri, BF, fuco, HF, pras, viola, allox, zea, lut, and chlb. These groups and the output ratios used for each group are provided in the supporting information (**Table S1–S2**). Input groups and pigment ratios were the same as those used by Peña et al. (2019) and were taken from Higgins et al. (2011). Using these groups and ratios was important as our results were compared to more recent Line-P results derived using these inputs.

3.3.3 Sentinel-3A OLCI Data

The Ocean Land Color Instrument (OLCI) onboard the Sentinel-3A satellite is a sensor with high spatial (300 m) and temporal resolution (daily), capable of acquiring water reflectance in 10 bands (400, 412, 443, 490, 510, 560, 620, 665, 674, and 681 nm) within the visible spectrum, along with an improved signal-to-noise ratio and the off-nadir pointing of the sensor swath to eliminate sun glint (Donlon et al., 2012). Level-1 full-resolution OLCI data from February 18th to March 21st, 2019, and March 5th to April 5th, 2020, were downloaded from the Sentinel-3 Marine CODA (Copernicus Online Data Access) web service and processed using POLYMER version 4.10 following Giannini et al. (2021) and Vishnu et al. (2022). POLYMER is an atmospheric correction algorithm that separates the radiometric contribution of the seawater from the atmospheric and residual sunglint using spectral matching techniques (Steinmetz et al., 2011). To achieve this, the atmospheric and residual sunglint contribution is modeled using a simple polynomial function, and the water reflectance model is achieved using the known values of Chla concentration and particle backscattering coefficient (bbp) from various oceanic regions (Steinmetz et al., 2011). The water reflectance model in POLYMER was obtained from Park and Ruddick (2005) and is valid in both Case-1 and Case-2 water types, including the bidirectional reflectance

(Steinmetz et al., 2011). POLYMER processed daily Rrs(λ) images were binned using SNAP (version 6.0) in batch mode using GPT (Graphics Processing Tool), and a total of eight composites (including 2019 and 2020) were obtained. While binning, the following quality flags were applied: "Cloud," "Invalid," "Negative Backscattering" (BB), "Out-of-bounds," "Exception," "High Air Mass," and "Thick Aerosol" (Steinmetz et al., 2016) following the recommendation of Giannini et al. (2021). Finally, the median Rrs(λ) calculated from each band was used for the analysis. Temporally binned weekly composite maps were chosen for studying the phytoplankton dynamics since spatial coverage is much increased as compared to the daily data; using weekly resolved data sets is generally accepted since it matches the temporal scale of phytoplankton phenology in the open ocean (Scott & Werdell, 2019; Lange et al., 2020). Specifically, in the Alaska Gyre, no abrupt temporal and spatial changes in the phytoplankton biomass are typically observed, with minimum seasonal or inter-annual variability (Evans & Parslow, 1985; Vinogradov et al., 1997; Peña & Varela, 2007; Westberry et al., 2016). Therefore, our study used the weekly binned OLCI imagery for further analysis.

3.3.4 EOF-Based Algorithm

The EOF-based algorithm was initially developed by Bracher et al. (2015) and globally applied by Xi et al. (2020), and here, this approach was adapted to the SNEP waters to retrieve the Chla concentration of phytoplankton groups. Before running the EOF analysis, HPLC pigment ratios used for the CHEMTAX analysis were clustered by ratio (i.e., pigment:TChla ratios input as rows and station as columns) using a novel data-driven approach to derive pigment groupings that showed similar trends (i.e., pigment collinearity) and, as a result, may not have been separable via satellite retrievals (Catlett & Siegel, 2018; Kramer & Siegel, 2019; Kramer et al., 2020). Methods for this clustering are described in section 3.3.5. Clustering resulted in three broad groupings with pigments characteristic of: (1) Hapto/Pelago/Cyano; (2) Diatoms/Dino/GA; and (3) Crypto (**Figure 10**). For example, pigment:TChla ratios for HF, BF, chlc3, and zea clustered together suggesting that haptophytes (chlc3, HF), pelagophytes (chlc3, BF), and cyanobacteria (zeaxanthin) covaried in the environment and may have been difficult to derive independently via satellite. Here, we combine the closely related phytoplankton subgroups, such as chlorophytes and prasinophytes, into a broader green algae (GA) category for better retrieval from the satellite. Based on these results, the corresponding CHEMTAX output groups were summed into these

broader groupings to improve accuracy, signal-to-noise ratios, and limit misclassifications of OLCI-derived phytoplankton groups.

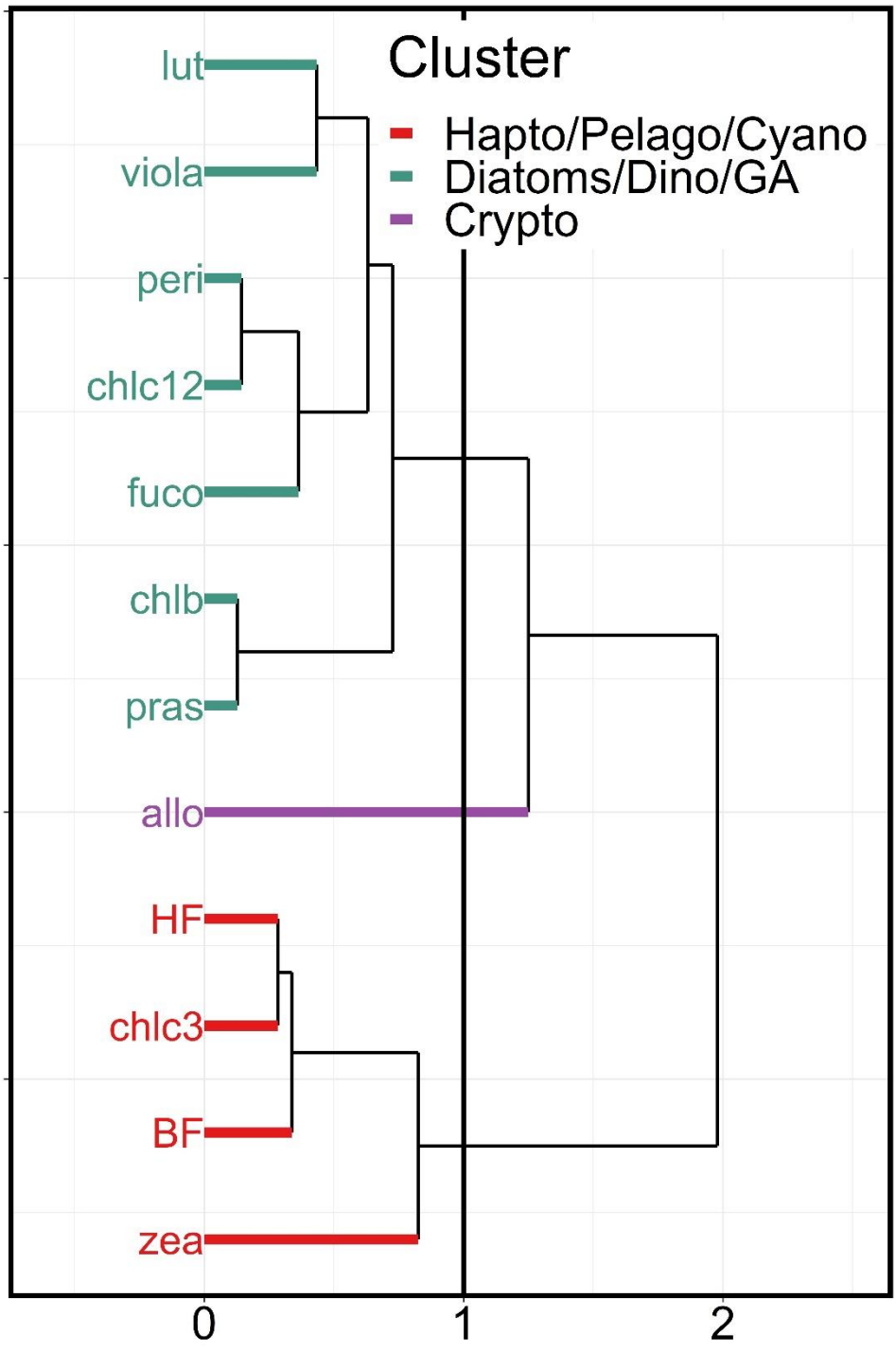

Figure 10. Hierarchical clustering of pigment:total Chlorophyll-a (TChla) ratios using the Pearson correlation distance and Ward's linkage method following Catlett and Siegal (2017), Kramer and Siegal (2019), and Kramer et al. (2020). The cophenetic correlation coefficient for the clustering was 0.74, indicating that the clustering was a good representation of the data (Kramer & Siegal, 2019). A linkage distance of 1 was used following Kramer and Siegal (2019) and Kramer et al. (2020) to define clusters. Different colors are used to show the three clusters, which are made up of pigments representing different phytoplankton groups.

Next, the regression model was built for the Hapto/Pelago/Cyano, Diatoms/Dino/GA, and Crypto groups using the EOF scores derived from the weekly composites of OLCI Rrs(λ) and the corresponding Chla contributions. From each weekly composite image, matchups of satellite Rrs at the 10 OLCI bands were extracted from 5x5 pixel windows corresponding to the location of the in situ sampling region for a total of 61 matchups. For each 5x5 pixel window, the following quality control criteria were adopted: valid pixels were ≥ 17/25, the CV at 560 nm was ≤ 20%, and the median value from each band was used to avoid outliers (Werdell & Bailey, 2005; Mograne et al., 2019). Similarly, all pixels that did not satisfy μ–1.5σ, μ+1.5σ were excluded from the analysis (Xi et al., 2020), where μ and σ are the mean and the standard deviation, respectively.

The regression model training with the matchup dataset showed reliable performance for Hapto/Pelago/Cyano and Diatoms/Dino/GA (e.g., R^2 = 0.45, 0.42, respectively; $p < 0.0001$). A cross-validation procedure was performed to assess the robustness of the fitted regression model (Bracher et al., 2015; Xi et al., 2020). The performance of the regression model was evaluated by calculating the coefficient of determination (R^2), mean absolute error (MAE; mg/m^3), median percentage difference (MDPD; %), and bias (%). The equation for these statistical parameters is provided by Xi et al. (2020) and Vishnu et al. (2022). Finally, the regression model was applied to each weekly composite, and phytoplankton group spatial maps were generated. The statistical performance of the regression model was obtained using a full-fit and cross-validation procedure provided in **Table S3**.

The spatial distribution of the OLCI-derived phytoplankton group Chla was quantitatively assessed by comparing the result against the temporally closest 2019 and 2020 CHEMTAX results from DFO Line-P monitoring program (Peña & Nemcek, 2021). The matchup extraction criteria and the statistical metrics used for the validation were the same as followed in the algorithm training. However, it is important to note that there was a temporal mismatch between the Line-P observations and our phytoplankton group retrieval from OLCI due to the limitation of valid satellite data. For example, the 2019 OLCI phytoplankton groups were retrieved between February 18th to March 21st, and the corresponding Line-P data were obtained between February 6th to February 12th. Similarly, 2020 OLCI retrievals were between March 5th and April 5th, and the Line-P data was mainly conducted from February 10th to February 16th. Furthermore, while the pigment ratios used in the CHEMTAX analysis of the Line-P dataset were comparable to those in our study, the HPLC phytoplankton

pigments samples were analyzed via a different protocol (see Nemcek & Peña, 2014) at the Institute of Ocean Sciences.

Finally, the reliability of the satellite retrievals over a broader spatial and temporal domain was investigated. For this analysis, the composite with the highest spatial coverage was selected from 2019 and 2020. Spatial-temporal trends were extracted from a 5x5 pixel window within the weekly composite imagery across a latitudinal and longitudinal domain, and the corresponding standard deviation is presented.

3.3.5 Statistical Analyses

All statistical analyses were performed using R version 1.0.153 software (R Core Team, 2021; https://www.r-project.org/), and figures were created using ggplot2 (Wickham, 2016). Spatial maps for surface water properties and nutrients were created by interpolating discrete CTD stations using the bilinear interpolation method with the "Raster" package and the "disaggregate" functions. Additionally, contours were created by applying the "contour" function from the "Raster" package to the interpolated raster data. Following the spatial map, hierarchical clustering was done using the "stats" package and "dist" and "hclust" functions. Firstly, the clustering to derive the CHEMTAX groupings was done by station (i.e., station as rows and pigment ratios as columns) using euclidean distances and Ward's linkage method (Ward.D2) following Swan et al. (2016). Selection of clusters was performed via visual inspection of the dendrogram and the Silhouette method visualized using the "fviz_nbclust" function from the "factoextra" package. Secondly, the Hierarchical clustering was done to assess pigment colinearity and derive broad satellite groupings was done by pigment ratio (i.e., pigment ratios as rows and stations as columns) using the correlation distance (1-R, where R is the Pearson correlation between the input pigment ratios) and Ward's linkage method (Ward.D2) (Catlett & Siegel, 2018; Kramer & Siegel, 2019; Kramer et al., 2020). In this hierarchical clustering, a cutoff linkage of 1 was used to derive clusters representing similar and covariant phytoplankton communities (Kramer et al., 2020). For both clustering approaches, the quality of the clustering was assessed using the cophenetic correlation coefficient derived using the "stats" package and the "cophenetic" and "cor" functions. Cophenetic correlations range from 0–1 with values close to 1 suggesting that the dendrogram accurately represented the distances between the input parameters (Legendre & Legendre, 1998).

Following the cluster analysis, redundancy analysis (RDA) was performed to investigate the physicochemical drivers of the CHEMTAX and OLCI-derived phytoplankton groups. This method is commonly used in ecological studies, and several authors have used it to elucidate the role of environmental drivers on CHEMTAX-derived phytoplankton groups across various environmental conditions (e.g., Vaillancourt et al., 2018; Del Bel Belluz et al., 2021). Redundancy analysis was performed independently on the CHEMTAX dataset (N = 79) and the OLCI outputs (N = 79). The response variables of the first analysis were the Chla contributions of the eight CHEMTAX phytoplankton groups (Cyano, Hapto, Pelago, Prasino, Chlor, Crypto, Dino, and Diatoms). The second RDA analysis used the three significant OLCI phytoplankton groups (Hapto/Pelago/Cyano, Diatoms/Dino/GA, and Crypto). Explanatory variables for both analyses were the same: sea surface temperature (SST), sea surface salinity (SSS), potential density (sigma-t), mixed layer depth (MLD), surface dissolved oxygen (DO), dissolved inorganic nitrogen (DIN), dissolved inorganic phosphate (DIP), dissolved silicate (DSi), N/P, Si/P, and Si/N ratios. Prior to each RDA analysis, the response variables were transformed using the Hellinger transformation ("Vegan" package and "decostand" function) which makes taxonomic data more suitable for the linear RDA method and gives groups with low or zero contributions reduced weights (Legendre & Gallagher, 2001; Ramette, 2007). Furthermore, collinearity of explanatory variables, which can reduce the quality of RDA analysis, was checked using variance inflation factors (VIFs) (Vaillancourt et al., 2018) via the "vif.cca" function from the "vegan" package. Explanatory variables with the highest VIFs were removed until all remaining variables had VIF scores < 10 (Mojica et al., 2015; Vaillancourt et al., 2018). Once the transformations and removals were complete, the RDAs were run using the "rda" function (vegan package) on the response and explanatory dataset to test for global significance (Ramette, 2007). After it was determined that the data were appropriate and the global models were significant (CHEMTAX: $p < 0.001$; OLCI: $p < 0.05$), forward selection of the explanatory variables was performed using the "OrdiR2step" function (vegan package) and Monte Carlo permutation (49,999) to determine and remove variables that did not significantly ($p > 0.05$) contribute to the variance (Ramette, 2007; Blanchet et al., 2008; Vaillancourt et al., 2018). The significant variables for the CHEMTAX data were SSS, MLD, DSi, and SST. While the global model for the OLCI phytoplankton groups was significant, no explanatory variables were found to be significant and this analysis was not used.

Lastly, significant differences between the physical and chemical variables between 2019 and 2020 were determined using the nonparametric Wilcoxon Rank-Sum test using “wilcox.test” function.

3.4 Results

3.4.1 Physicochemical Properties

In general, most of the physicochemical properties showed latitudinal (north to south) and longitudinal (east to west) gradients (**Figure 11** and **Figure 12**). During 2019, SST (6.57 ± 0.75°C; **Table 1**) varied latitudinally with the 7° C isotherm demarcating the boundary between the subarctic water mass to the northwest and the subtropical water mass to the southeast (**Figure 11A**; **Table 1**). Similarly, SSS (**Figure 11C**; **Table 1**) reflected the temperature distribution with the boundary between subarctic and subtropical waters identifiable by the 32.4 PSU isohaline (**Figure 11C**). This year, sigma-t did not exhibit a latitudinal or longitudinal gradient (**Figure 11E**); however, a low sigma-t feature was observed south of 55° N. While surface DO was largely consistent throughout the study region, varying from 6.73 to 7.11 (ml/L), elevated levels were observed in conjunction with the low sigma-t feature. Finally, the MLD in 2019 varied from 70 to 115 m with the deepest depth in the southwestern portion of the study region (**Figure 11G**).

In 2020, differences in water masses were again demarcated by the 7° C SST isotherm and the 32.4 PSU isohaline. The longitudinally expanded study region in this year showed continental shelf influences with lower salinities observed towards Vancouver Island and considerably lower values of Haida Gwaii. Furthermore, both sigma-t and MLD tended to follow the trends observed in SSS with lower sigma-t and MLD towards the shelf and increasing values towards the subarctic water mass (**Figure 11**). Finally, surface DO was generally consistent across the study region (6.49 ± 0.15 ml/L; **Table 1**) showing a slight longitudinal trend with lower values towards Vancouver Island and a low point on the southern edge of the study region.

In agreement with the water properties, surface inorganic nutrients generally showed latitudinal and longitudinal gradients (**Figure 12**), with higher concentrations recorded within the subarctic water mass. Specifically, during 2019, surface concentrations of DIN, DIP, and DSi varied from 9.15 to 18.71 (13.53 ± 2.38 μM/L; **Table 1**), 0.77 to 1.31 (1 ± 0.12 μM/L; **Table 1**), and 6.67 to 27.96 (18.64 ± 4.78 μM/L; **Table 1**), respectively. In 2020, DIN, DIP,

and DSi ranged from 4.77 to 20.65 (10.90 ± 3.38 µM/L; **Table 1**), 0.56 to 2.22 (0.98 ± 0.31 µM/L; **Table 1**), and 2.70 to 32.59 (11.06 ± 6.82 µM/L; **Table 1**), respectively. Similar to 2019, the surface concentration of all of the nutrients was generally higher offshore in the subarctic water mass, but the expanded sampling in this year showed higher concentrations in the continental shelf waters near Haida Gwaii and lower concentrations in the subtropical water mass off Vancouver Island. Unlike 2019, an elevated DIP feature was observed near 50° N (**Figure 12D**). Finally, a significant (p < 0.05, Wilcoxon Rank-Sum test) annual variation was found for all physical and chemical variables between 2019 and 2020.

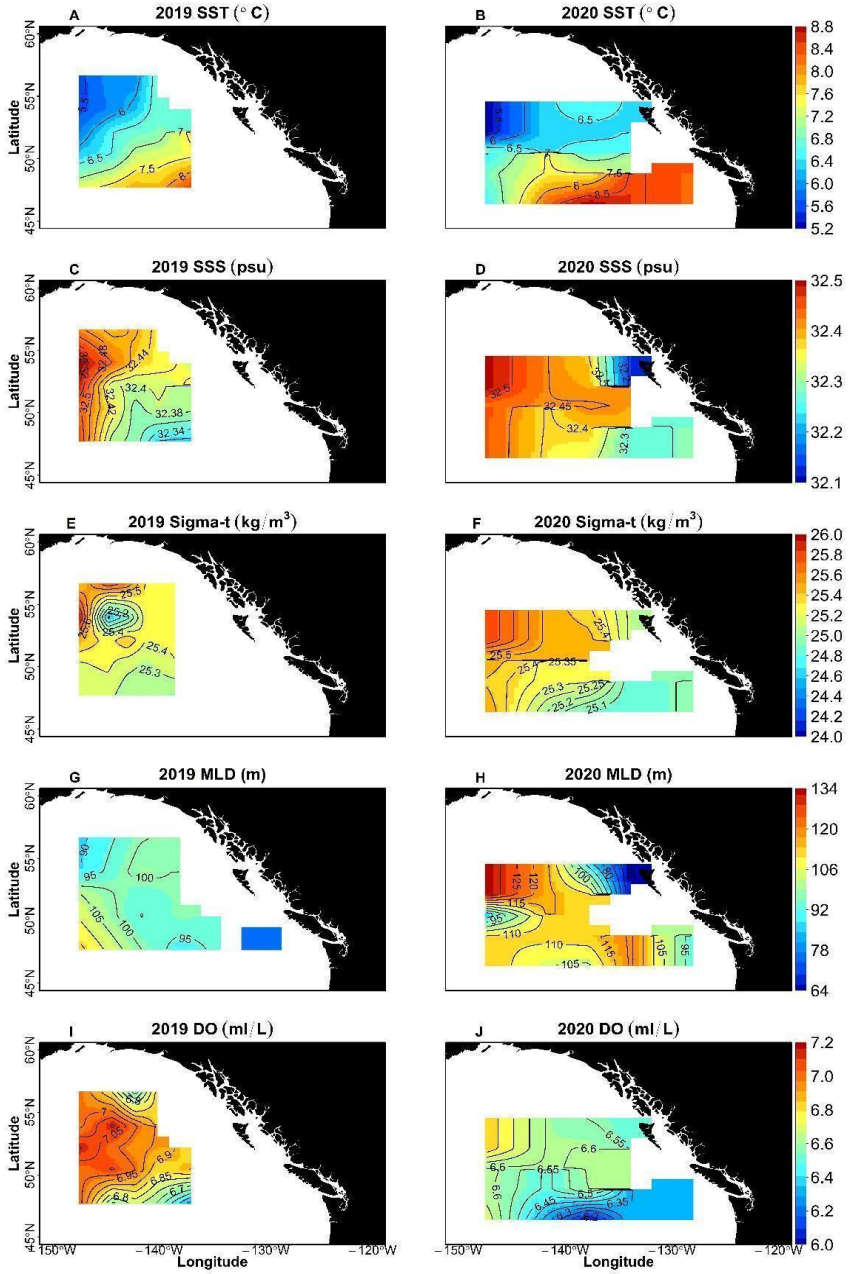

Figure 11. Distribution of surface water properties of the subarctic northeast Pacific for 2019 and 2020: (**A** to **B**) sea surface temperature (SST; ˚C), (**C** to **D**) sea surface salinity (SSS;

PSU), (**E** to **F**) potential density (sigma-t; kg/m^3), (**G** to **H**) mixed layer depth (MLD; m), and (**I** to **J**) surface dissolved oxygen (DO; ml/L). The 2019 dataset covers from 18th February to 21st March, while the 2020 spans from 5th March to 5th April.

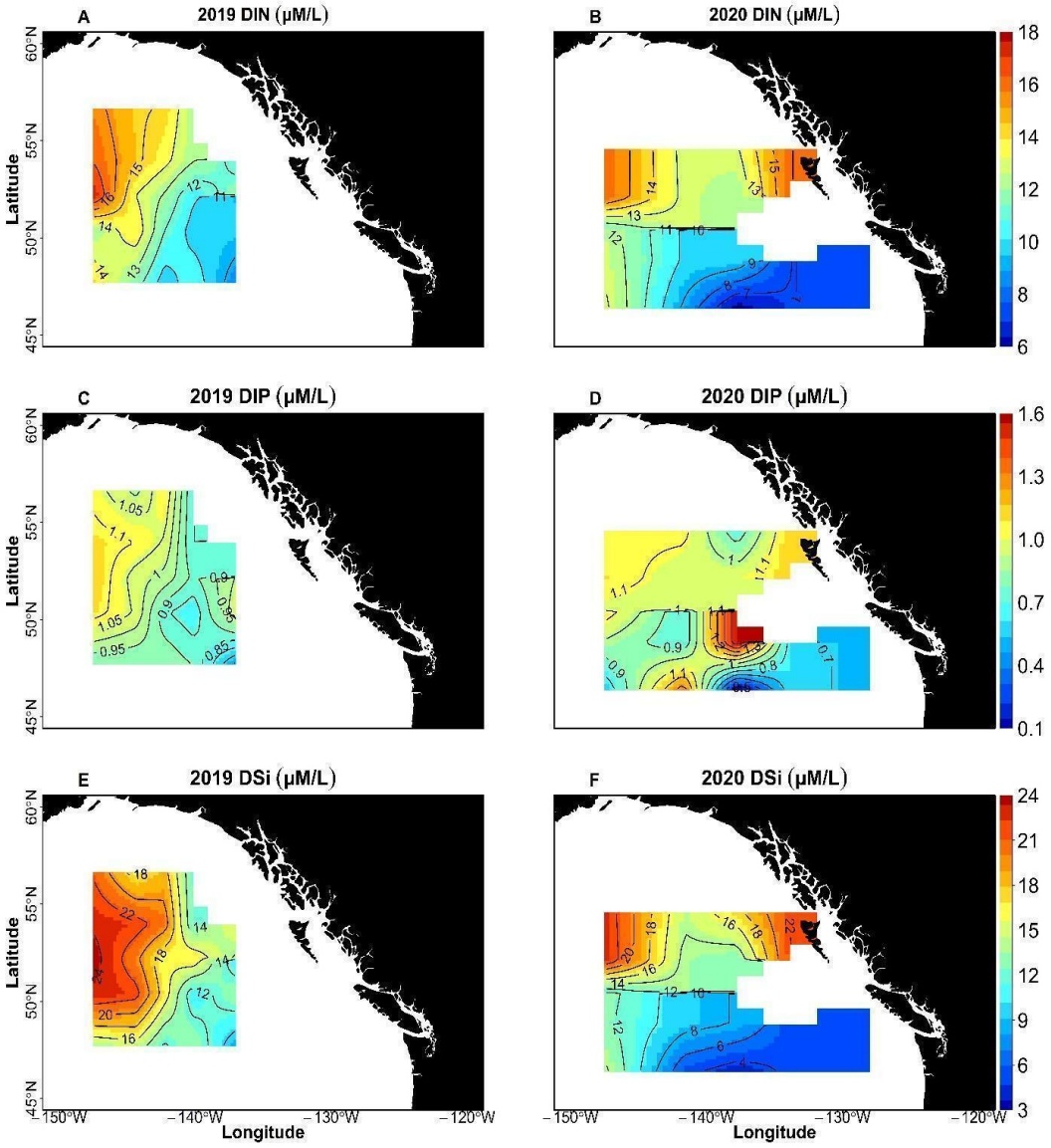

Figure 12. Spatial distribution of dissolved inorganic nutrients in the surface waters of the subarctic northeast Pacific for 2019 and 2020: (**A** to **B**) dissolved inorganic nitrogen (DIN), (**C** to **D**) dissolved inorganic phosphate (DIP; µM/L), and (**E** to **F**) dissolved inorganic silicate (DSi; µM/L). The 2019 dataset covers from 18th February to 21st March, while the 2020 spans from 5th March to 5th April.

Table 1. Distribution of physicochemical properties, CHEMTAX-derived phytoplankton community composition, and nutrient ratios during 2019 and 2020. Minimum and maximum values are shown in square brackets, average in bold and standard deviation (± SD) in parenthesis. The 2019 dataset covers from 18th February to 21st March, while the 2020 spans from 5th March to 5th April.

Variables	Years	
	2019	2020
SST	[5.04 – 8.20]	[5.25 – 9.31]
(˚C)	**6.57** (0.75)	**7.25** (0.94)
Salinity	[32.34 – 32.63]	[32.14 – 32.54]
(PSU)	**32.43** (0.06)	**32.37** (0.21)
Potential density	[24.29 – 25.81]	[24.50 – 25.70]
(kg/m^3)	**25.40** (0.28)	**25.31** (0.21)
MLD	[70 – 115]	[24 – 130]
(m)	**98** (10)	104 (22)
DO	[6.73 – 7.11]	[6.08 – 6.83]
(ml/L)	**6.97** (0.10)	**6.49** (0.15)
NO_3+NO_2	[9.15 – 18.71]	[4.77 – 20.65]
(µmol/L)	**13.53** (2.38)	**10.90** (3.38)
PO_4	[0.77 – 1.31]	[0.56 – 2.22]
(µmol/L)	**1.00** (0.12)	**0.98** (0.31)
SiO_4	[6.67 – 27.96]	[2.70 – 32.59]
(µmol/L)	**18.64** (4.78)	**11.06** (6.82)
Cyano	[0.01 – 0.04]	[0 – 0.01]
(mg/m^3)	**0.02** (0.01)	**0.01** (0.005)
Hapto	[0.06 – 0.21]	[0.06 – 0.20]
(mg/m^3)	**0.14** (0.04)	**0.10** (0.04)
Pelago	[0.02 – 0.09]	[0.02 – 0.06]
(mg/m^3)	**0.05** (0.02)	**0.04** (0.01)
Chlor	[0 – 0.08]	[0 – 0.09]
(mg/m^3)	**0.01** (0.01)	**0.03** (0.04)
Prasino	[0.01 – 0.07]	[0.01 – 0.15]

(mg/m^3)	**0.04** (0.01)	**0.08** (0.03)
Crypto	[0.01 – 0.11]	[0 – 0.06]
(mg/m^3)	**0.03** (0.02)	**0.03** (0.02)
Dino	[0.01 – 0.05]	[0.04 – 0.11]
(mg/m^3)	**0.03** (0.02)	**0.07** (0.02)
Diatoms	[0 – 0.04]	[0 – 0.14]
(mg/m^3)	**0.01** (0.01)	**0.03** (0.04)
TChla	[0.17 – 0.50]	[0.21 – 0.52]
(mg/m^3)	**0.35** (0.07)	**0.37** (0.12)

3.4.2 Phytoplankton Community Composition

3.4.2.1 CHEMTAX-Derived Phytoplankton Community Composition

The CHEMTAX-derived phytoplankton community composition for 2019 and 2020 are shown in **Figure 13**. Overall, low surface TChla concentrations were observed throughout the study period ranging from 0.17 to 0.52 mg/m^3, with an average of 0.35 mg/m^3. Overall, a diverse phytoplankton community was present across the region. In 2019, haptophytes were the dominant group (0.16 ± 0.04 mg/m^3, 47 ± 6%; **Figure 13A**, **13G**), followed by pelagophytes (0.06 ± 0.02 mg/m^3, 17 ± 3%) and green algae (0.05 ± 0.01 mg/m^3, 15 ± 6%). Within the green algae community, prasinophytes were the predominant group (0.04 ± 0.01 mg/m^3, 11 ± 3%). Diatoms, dinoflagellates, cryptophytes, and cyanobacteria were the less abundant groups throughout the region (2 ± 3%, 6 ± 1%, 6 ± 3%, and 6 ± 3%, respectively). In general, the community compositions were very similar throughout the study region in 2019 with the exception of increased contributions by cryptophytes in the northern portion of the subarctic water mass **(Figure 13D**; stations 18–24).

In 2020, the phytoplankton community composition data extended from the open ocean to shelf stations, making comparison with 2019 difficult (see **Figure 13C** to **13K**). These shelf stations showed increased TChla (0.51 mg/m^3) concentrations, and compared to the open ocean, these stations showed reduced contributions by haptophytes (18%) and pelagophytes (8%), and with the station within Dixon entrance showed no contribution of haptophytes and considerably higher diatom contributions (31%, **Figure 13G**, **13H**). Interestingly, the few offshore stations collected in 2020 showed higher dinoflagellate contributions (22%) and lower haptophyte contributions compared to those in 2019.

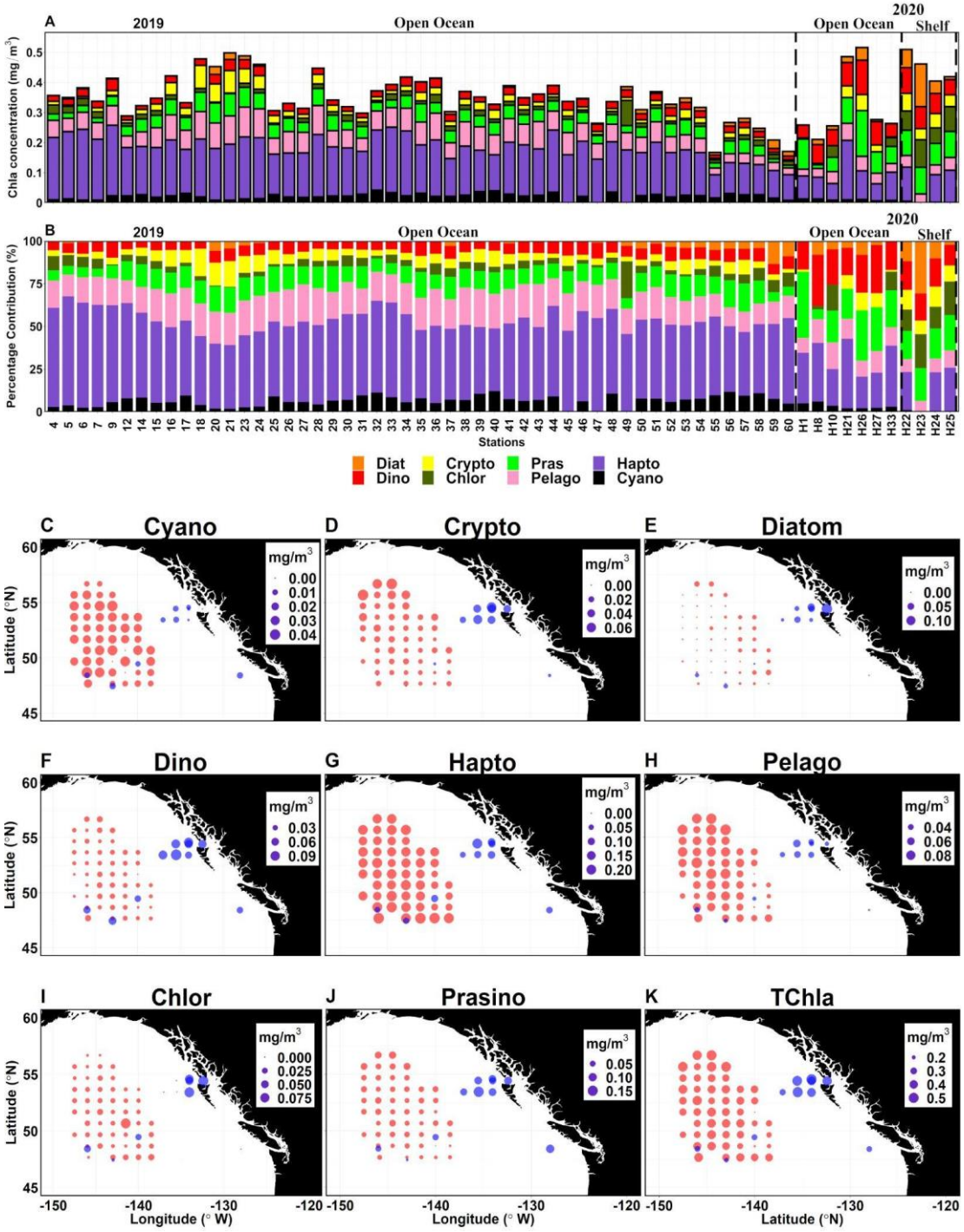

Figure 13. Spatial distribution of CHEMTAX-derived phytoplankton community composition in the surface waters of the subarctic northeast Pacific for 2019 and 2020. (**A**) Chla concentration of cyanobacteria, cryptophytes, diatoms, dinoflagellates, haptophytes, pelagophytes, chlorophytes, prasinophytes, and total chlorophyll-a concentration. (**B**) percentage contribution of each group. (**C** to **K**) spatial distribution of each of these groups. The red and blue circles represent 2019 and 2020, respectively. The 2019 dataset covers from 18th February to 21st March, while the 2020 spans from 5th March to 5th April.

3.4.2.2 OLCI-Derived Phytoplankton Community Composition

Figure 14 and **Figure 15** show the spatial distribution of Crypto, Hapto/Pelago/Cyano, Diatoms/Dino/GA, and TChla derived from the weekly composite OLCI imagery across the mostly iron-poor open ocean waters of the SNEP during 2019 and 2020. These groups were selected as a result of their statistically better performance (section 3.3.4) and their better agreement with the independent in situ data (**Figure 17**). In 2019, a homogeneous low Chla distribution was generally noted for all groups (**Figure 14**; **Table S4**). Consistent with the in situ CHEMTAX data, the highest average contributions in 2019 were from Hapto/Pelago/Cyano, followed by Diatoms/Dino/GA and then Crypto (0.21, 0.12, 0.03, mg/m^3, respectively). However, intermittent patches of increased Chla levels were observed for cryptophytes throughout the region, with higher concentrations predominantly found towards the north, which mirrored the in situ CHEMTAX data. Finally, the average TChla ranged from 0.34 to 0.36 mg/m^3 (**Table S4**) across the 2019 composites, with patches of increased concentration found within the composites notably in the northern GoA (see **Figure 14** and **Figure 16**).

Similar to 2019, a spatially homogeneous distribution was found for all groups in the open ocean in 2020 (**Figure 15**). However, a patch of increased contributions was recorded for Crypto, Diatoms/Dino/GA, and TChla on the continental shelf off of Haida Gwaii that was matched by decreased Hapto/Pelago/Cyano contributions. The highest average concentration from 2020 was by Hapto/Pelago/Cyano, followed by Diatoms/Dino/GA and then Crypto (0.17, 0.12, 0.02, mg/m^3, respectively). The average TChla concentration remains the same (0.32 mg/m^3) between the weekly composite, with a notable increase in concentration found towards the Haida Gwaii shelf (see **Figure 15**).

The OLCI-derived distribution of Chla concentrations of the major groups and TChla mirrored the CHEMTAX-derived data when considering latitudinal (2019) and longitudinal (2020) transects extracted from satellite data (**Figure 16**). Similar to the CHEMTAX results, all groups, including TChla, were invariant in the HNLC region; however, elevated concentrations were found for cryptophytes, Diatoms/Dino/GA, and TChla towards the north in 2019. Also, a small cryptophyte peak was discernible in the south of the study region; however, it was similar to the range observed by the in situ CHEMTAX. Similar to the CHEMTAX results in 2020, increased Diatoms/Dino/GA and TChla concentrations were

observed towards the Haida Gwaii shelf with corresponding decreases in Hapto/Pelago/Cyano (see **Figure 16**).

The comparison between OLCI retrieved TChla and Hapto/Pelago/Cyano against independent Line-P CHEMTAX-derived phytoplankton community composition data showed that a good agreement was observed despite a notable temporal mismatch of the dataset (**Figure 17**). Of the considered groups, Hapto/Pelago/Cyano demonstrated a strong agreement between dataset (R^2 = 0.32, MAE = 0.03 mg/m^3, BIAS = 8.5%, MDPD = 9.9%, p < 0.001; **Figure 17B**), followed by cryptophytes (R^2 = 0.27, MAE = 0.03 mg/m^3, BIAS = 26.37%, MDPD = 43.49%, p < 0.05; **Figure 17A**). In contrast, no agreement was observed for Diatoms/Dino/GA (R^2 = 0, MAE = 0.02 mg/m^3, BIAS = -1.35%, MDPD = 14.5%; **Figure 17C**). Finally, TChla showed the best agreement between OLCI and in situ CHEMTAX data (R^2 = 0.53; MAE = 0.01 mg/m^3, BIAS = -17.39%, MDPD = 21.06%, p < 0.001; **Figure 17D**).

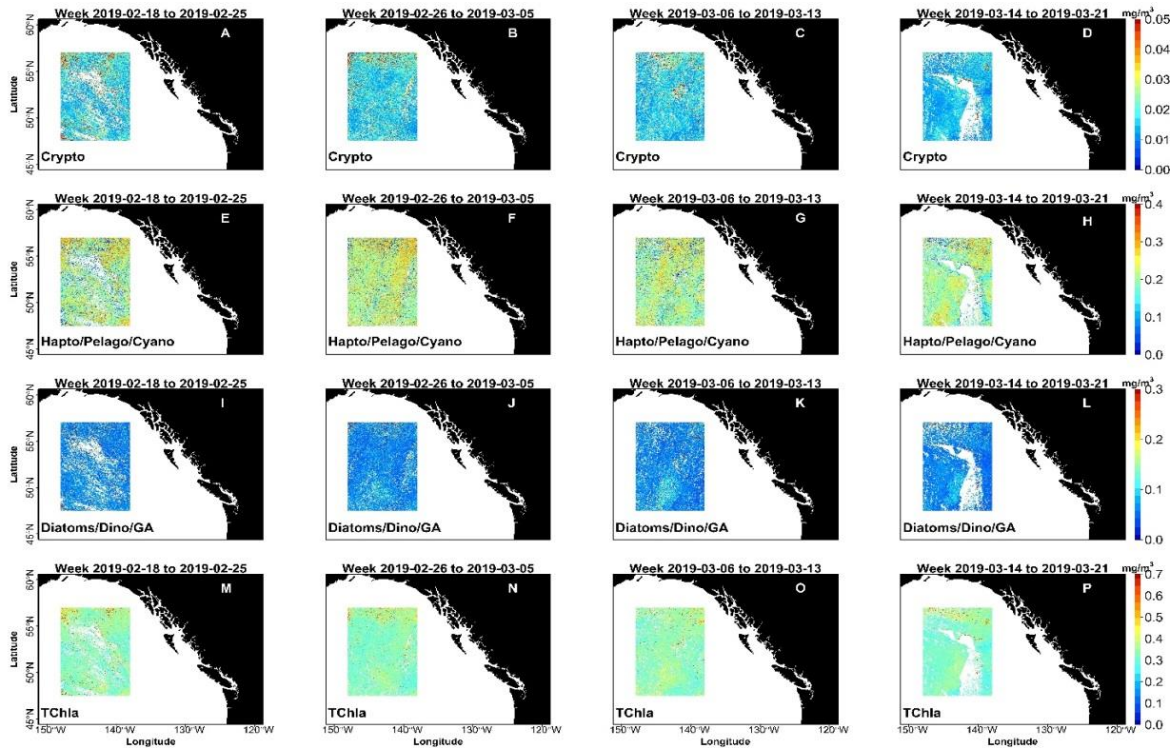

Figure 14. Spatial distribution of cryptophytes, haptophytes/pelagophytes/cyanobacteria, diatoms/dinoflagellates/green algae, and TChla concentration retrieved over the subarctic northeast Pacific from the 2019 weekly binned Ocean Land Color Imagery (OLCI) imagery. The spatial extent of the phytoplankton community composition map derived from OLCI was limited to the coverage of the matchup dataset, which primarily consisted of open ocean samples. Extending the algorithm's application to the region outside the matchup area may result in unreliable retrievals; therefore, the spatial extension is constrained to match the coverage of the matchup dataset from 2019. White regions on the map represent clouds. This dataset covers from 18th February to 21st March.

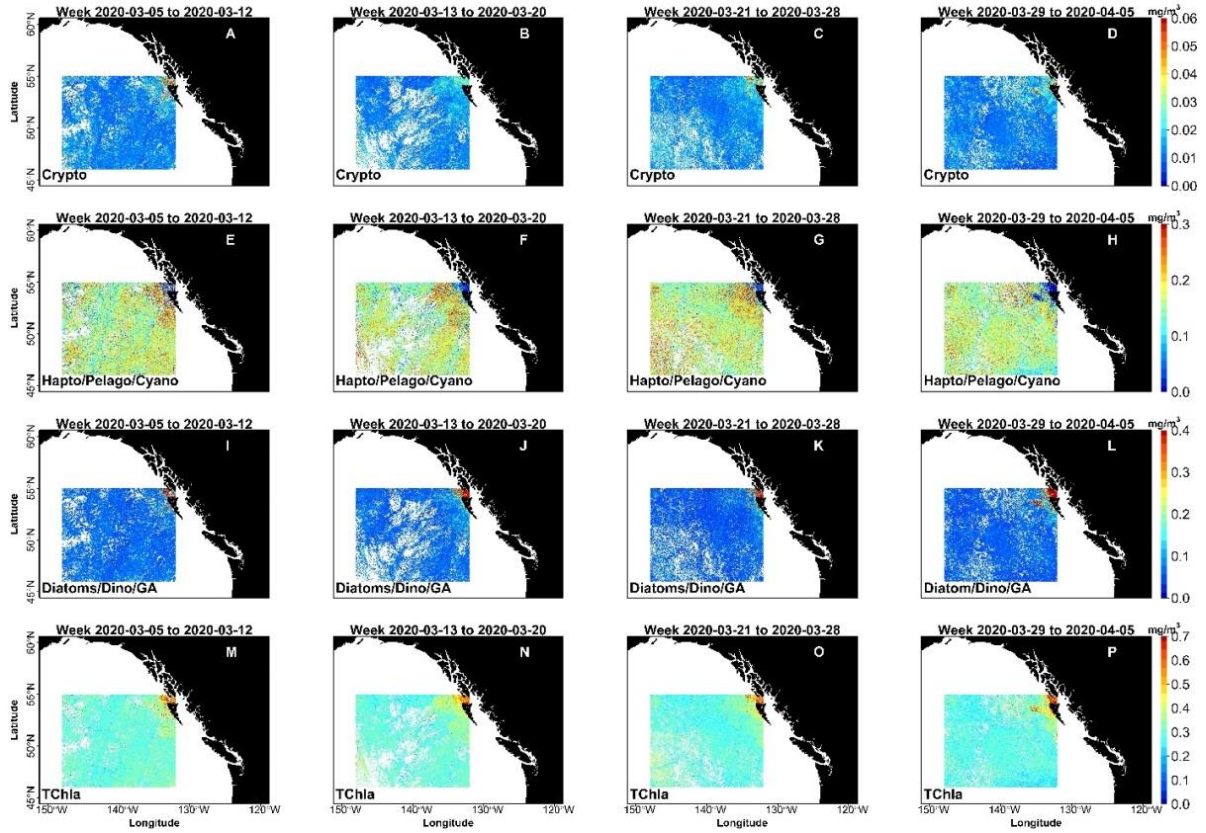

Figure 15. Spatial distribution of cryptophytes, haptophytes/pelagophytes/cyanobacteria, diatoms/dinoflagellates/green algae, and TChla concentration retrieved over the subarctic northeast Pacific from the 2020 weekly binned Ocean Land Color Imagery (OLCI) imagery. The spatial extent of the phytoplankton community composition map derived from OLCI was limited to the coverage of the matchup dataset, which primarily consisted of open ocean areas with a few samples from continental shelf waters off of Haida Gwaii. Extending the algorithm's application to the entire shelf may result in unreliable retrievals; therefore, the spatial extension is constrained to match the coverage of the matchup dataset from 2020. White regions in the open ocean represent clouds. This dataset covers from 5th March to 5th April.

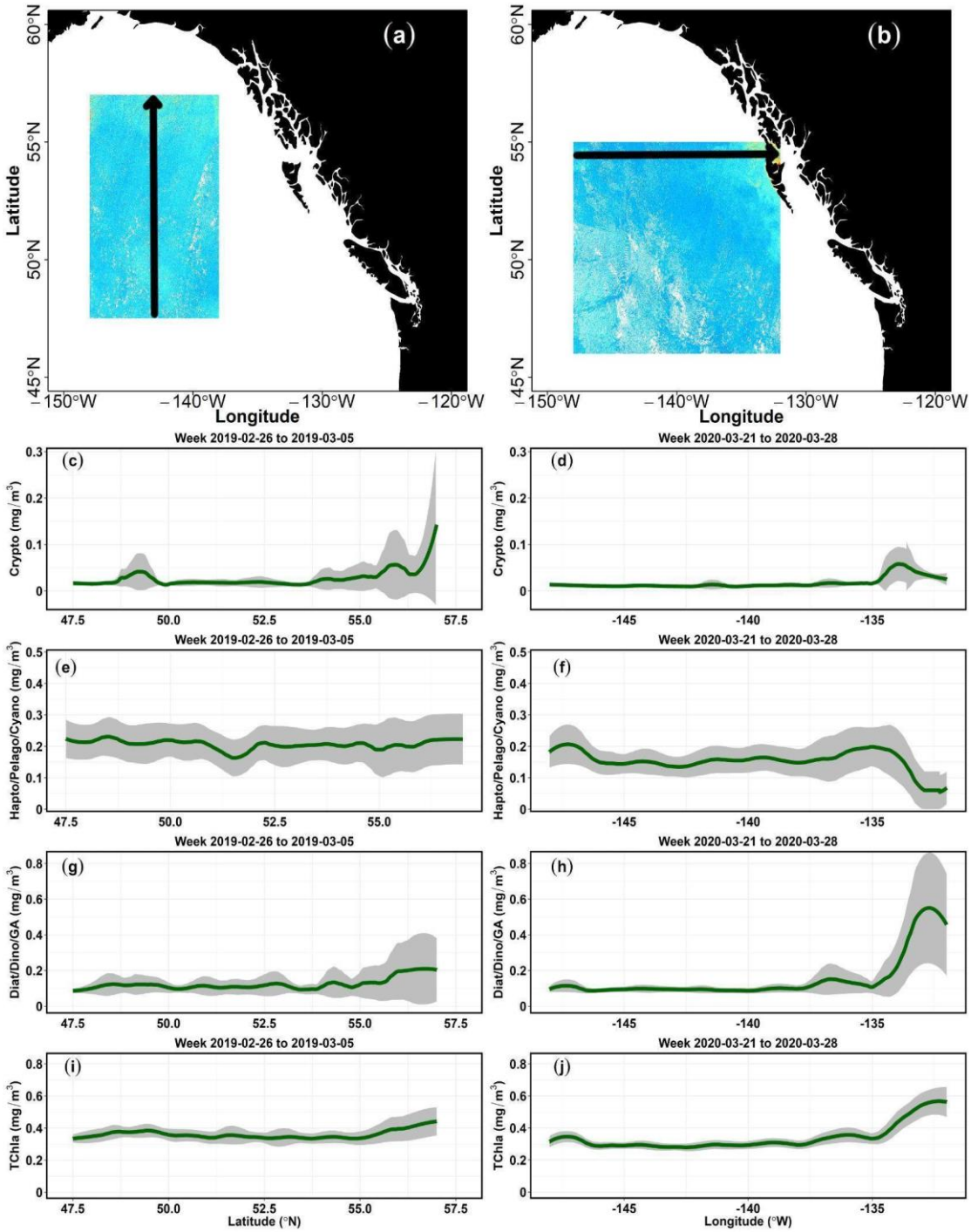

Figure 16. Latitudinal (2019) and longitudinal (2020) transects of Ocean Land Color Instrument derived surface chlorophyll-a (Chla; mg/m^3) concentration for cryptophytes, Hapto/Pelago/Cyano, Diatoms/Dino/GA, and total chlorophyll-a (TChla; mg/m^3) concentration extracted from the least cloud affected composites from 2019 and 2020. The Chla concentration was extracted from a 5x5-pixel window, and the standard deviation is plotted as a gray-shaded region. (a–b) represent the latitudinal and longitudinal transect lines for 2019 and 2020, respectively. (c,e,g,i) latitudinal distribution of surface Chla concentration for the major groups and TChla from 26th February 2019 to 05th March 2019. (d,f,h,j) Longitudinal distribution of OLCI derived surface Chla concentration for the major groups, including TChla, from 21st to 28th March 2020.

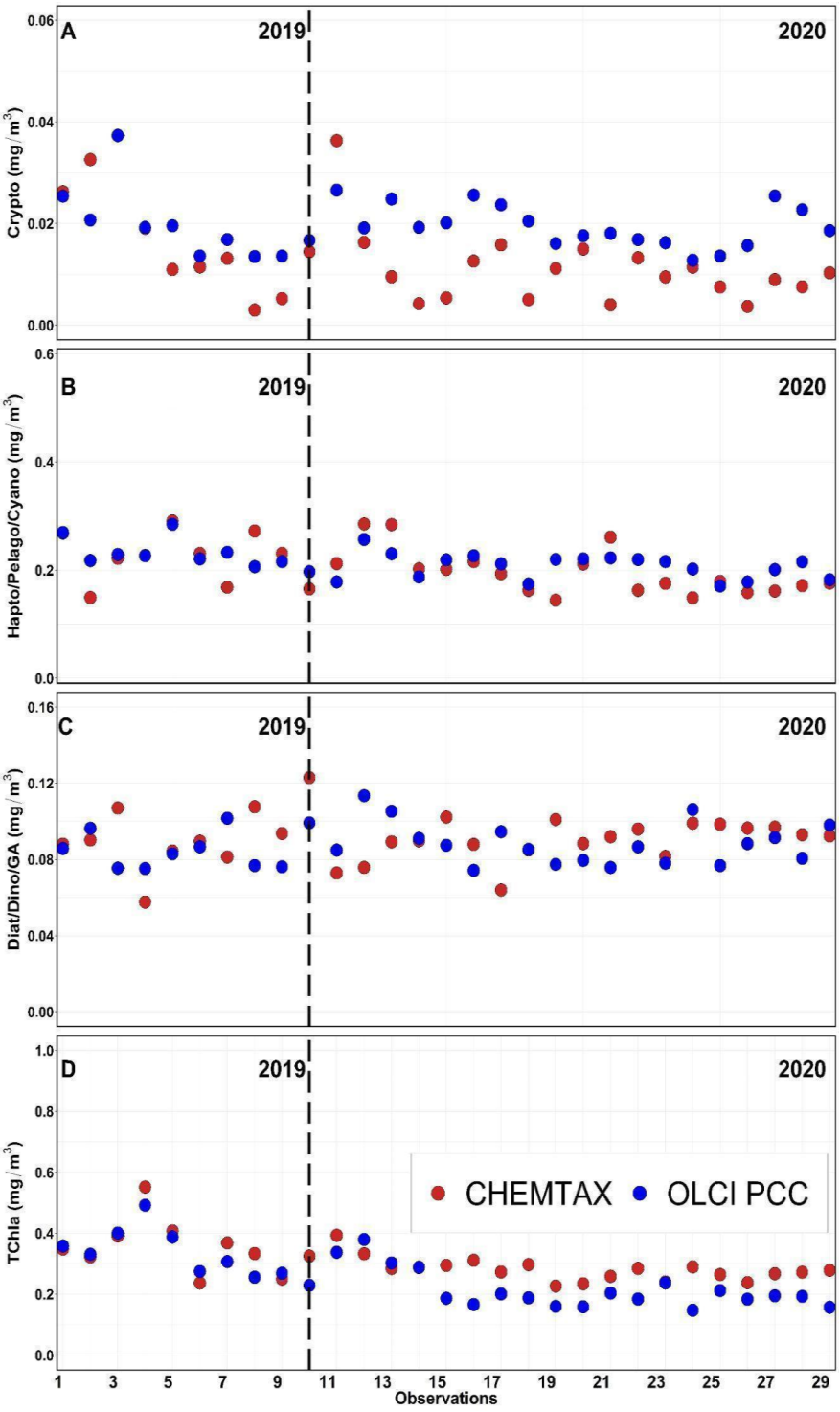

Figure 17. Comparison of phytoplankton group concentrations derived from the 2019 and 2020 Ocean Land Color Instrument (OLCI) weekly composite along the Line-P transect with the corresponding independent in situ Line-P CHEMTAX-derived phytoplankton community composition concentrations. The 2019 OLCI phytoplankton groups were retrieved between February 18th to March 21st, and the corresponding Line-P data were obtained between February 6th to 12th. Similarly, 2020 OLCI retrievals were between March 5th and April 5th, and the Line-P data was mainly conducted from February 10th to 16th. Red and blue circles represent CHEMTAX and OLCI-derived phytoplankton community composition, respectively.

3.4.2.3 Environmental Drivers

Redundancy analysis considering the output of the 8 CHEMTAX-derived phytoplankton community composition and matching environmental variables had an R^2 = 0.42 (p = 0.001) with RDA1 explaining 36% of the variance (p = 0.001) and RDA2 6% of the variance (p = 0.01) (**Figure 18**). The forward-selected significant variables were SSS (28% of variance, p = 0.0001), MLD (6%, p = 0.01), DSi (5%, p = 0.002), and SST (3%, p = 0.01). The RDA triplot (Scaling 2) showed large differences between years representing longitudinal open ocean-coastal variability along RDA1, with 2019 (dominantly open ocean samples) having higher contributions by haptophytes, prasinophytes, cyanobacteria, and pelagophytes; and 2020 (mostly shelf/coastal samples) having higher contributions by diatoms, dinoflagellates, and chlorophytes. Furthermore, the 2019 open ocean groups were strongly positively correlated with MLD and SSS while the 2020 shelf/coastal groups were negatively correlated with these drivers and groups. In turn, RDA2 was largely driven by latitudinal differences with the subarctic water mass samples showing higher DSi and cryptophyte contributions and an inverse correlation with SST and the more subtropical water mass samples. Results of RDA analysis between environmental drivers and the OLCI-derived phytoplankton groups found no significant forward selected variables.

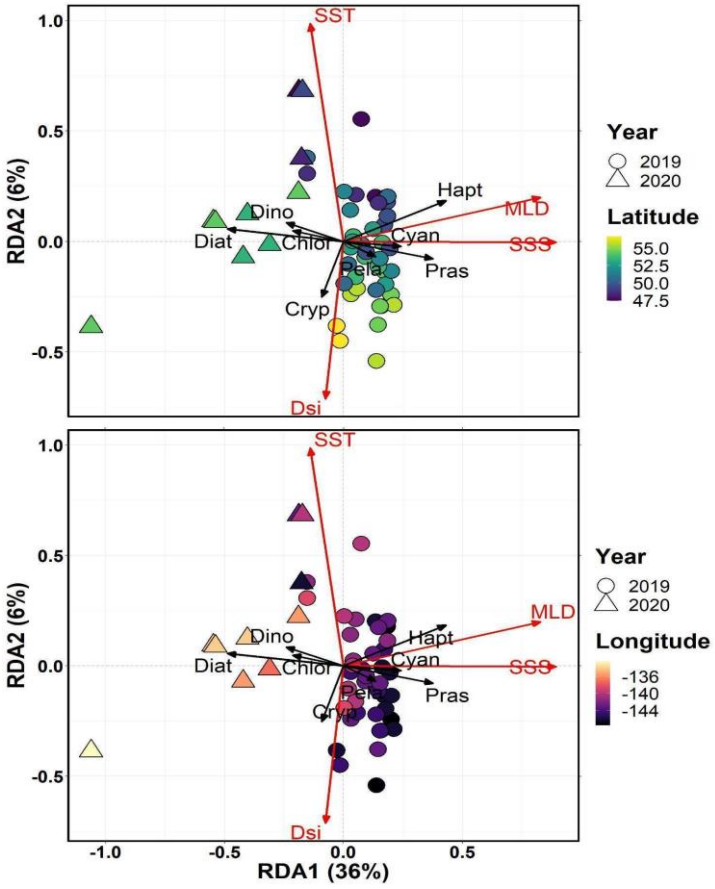

Figure 18. The redundancy analysis triplot (Hellinger transformation, scaling 2) illustrates the relationship between response and explanatory variables. The red arrows indicate statistically significant explanatory variables, sea surface temperature (SST), sea surface salinity (SSS), mixed layer depth (MLD), and dissolved silicate (DSi), and the black arrows show the CHEMTAX-derived phytoplankton community composition. The shape of each point represents the collection year with circles for 2019 and triangles for 2020. The color of the points represents the latitude (a) and longitude (b) of the sample collection location. The global model was significant with $p < 0.001$, and the explanatory variables explaining 42% of the constrained variance in the CHEMTAX-derived phytoplankton community composition. Scaling 2 was used for the triplot with the angle between arrows, both response and explanatory, indicating the linear correlation between the variables. For example, an angle of 90° shows no correlation, while an angle less than 90° indicates a positive correlation, and an angle greater than 90° indicates a negative correlation.

3.5 Discussion

This research used CHEMTAX and OLCI-derived phytoplankton composition data to provide an overview of the spatial-temporal distribution of the phytoplankton groups and investigates their potential environmental drivers in the SNEP. Our dataset is limited to the winter conditions of 2019 and 2020. Our findings showed distinct water masses in the region with a 7° C isotherm demarcating the boundary between the subarctic water mass to the northwest and the subtropical water mass to the southeast. Additionally, the continental shelf region of Haida Gwaii, which was sampled only in 2020, was dominated by the ACC and showed coastal influences with implications for the phytoplankton community composition in this region. Broadly, all physical and chemical variables tended to follow the regionally distinct water masses. By combining CHEMTAX and OLCI-derived phytoplankton community composition data, the dominance of small flagellates such as haptophytes, green algae, and pelagophytes in the iron-poor open ocean waters and increased diatom contributions on the shelf was demonstrated. Furthermore, the higher temporal and spatial resolution OLCI retrievals showed weak spatial-temporal variability for the significant phytoplankton groups within the iron-poor open ocean waters. Correlations between SSS, MLD, and diatoms were largely driven by the 2020 samples from the shelf and suggest coastal influenced water with increased BFe. In turn, correlations between SST, DSi, and cryptophytes in the open ocean regions were likely a result of similar latitudinal gradients (i.e., increased northern cryptophytes in 2019) rather than driver-species relationships. The following sections provide a discussion of the major findings of this study.

3.5.1 Total Chlorophyll-a Concentration in the Surface Waters of SNEP

Total chlorophyll-a (TChla) has been observed to be markedly low (< 0.5 mg/m^3) in the SNEP throughout the year (Brickley & Thomas, 2004; Childers et al., 2005) with little seasonal or inter-annual variability (Evans & Parslow, 1985; Vinogradov et al., 1997; Peña & Varela, 2007). Both the in situ HPLC and OLCI-derived TChla showed low (≤ 0.52 mg/m^3) values in the iron-poor open ocean waters in both 2019 and 2020, in agreement with previous studies in this region based on in situ data (e.g., Suzuki et al., 2002; Harrison et al., 2004; Peña & Varela, 2007; Peña et al., 2019; Peña & Nemcek, 2021). These had previously been demonstrated by satellite studies from the CZCS (Vinogradov et al., 1997; Banse & English, 1999) and subsequent recent studies utilizing the MODIS-Aqua (Westberry et al., 2016). A recent study showed subtle variability in surface Chla concentrations within iron-poor open ocean waters for the years 2019 and 2020, utilizing weekly composite imagery from Sentinel-3 OLCI (Vishnu & Costa, 2023).

3.5.2 Phytoplankton Community Composition Derived from CHEMTAX and Sentinel-3 OLCI

This study showed a diverse but spatially and temporally similar phytoplankton community composition throughout the SNEP. Specifically, CHEMTAX-derived results showed the dominance of haptophytes (avg 43 ± SD 11%), followed by green algae (18 ± 7%) and pelagophytes (16 ± 4%). All of these groups are generally made up of small-sized (< 10 μm) phytoplankton that have been found to regionally dominate the phytoplankton community in the SNEP throughout the year (Boyd & Harrison, 1999; Harrison et al., 2004); and are similar to previous findings using comparable pigment-based CHEMTAX approaches (e.g., Obayashi et al., 2001; Suzuki et al., 2002; Fujiki et al., 2009; Royer et al., 2010; Yang et al., 2018). Similar to our findings, the studies by Peña et al. (2019) and Peña & Nemcek (2021) show that haptophytes (also derived from the CHEMTAX/HPLC technique) dominate this region, mostly by even > 50% of the total phytoplankton community in the Alaska Gyre from the winters in 2013 to 2016 and more recently for February 2019 and 2020, respectively. Within the haptophyte group, *Phaecocystis pouchetti* has been found to be a dominant species and is likely the region's chief primary producer (Booth et al., 1993). In addition, Fujiki et al. (2009) showed that green algae contributed up to 30% to the total phytoplankton biomass throughout the Alaska Gyre and while ubiquitous, pelagophytes contributed < 20% to the total phytoplankton biomass during the winter of 2019 and 2020 (Peña & Nemcek, 2021). In turn, large-size (> 10 μm) phytoplankton in the SNEP are mainly composed of diatoms,

which are rarely seen in high abundance due to iron limitation restricting their growth (Peña & Varela, 2007). In agreement with this observation, our CHEMTAX results showed that diatoms contributed little (3 ± 5%) to the total phytoplankton biomass in the sampled open ocean waters but up to 30% over the Haida Gwaii shelf.

Apparently, CHEMTAX and OLCI preserved the overall spatial pattern of phytoplankton community composition well, with the highest concentrations found to be associated with the Hapto/Pelago/Cyano group. Furthermore, the EOF-based regression algorithm showed statistically significant retrievals based on the cross-validation result for Hapto/Pelago/Cyano, Diatoms/Dino/GA and TChla; however, independent validation using Line-P CHEMTAX data showed strong comparability only for the most abundant Hapto/Pelago/Cyano group and TChla, which well followed the observed Line-P concentrations. However, both the Crypto and Diatoms/Dino/GA groups exhibited poor comparability with the Line-P data, potentially due to their low abundance and small relative contributions (6% and 29%, respectively), as observed in the Line-P data. These influences would have exacerbated errors from atmospheric correction reducing comparability. Furthermore, temporal mismatches between our matchup dataset as well as its quality and quantity may have played a role in reducing comparability. When compared to other studies using comparable methods, our OLCI retrievals agree well with similar concentrations for haptophytes and green algae obtained from ten years (2002–2012) (Xi et al., 2020) and 20 years (2002–2022) (Konik et al., submitted) of GlobColor merged satellite imagery in the SNEP.

Despite the limitation of the broad phytoplankton grouping, the OLCI provides a more comprehensive spatial assessment of phytoplankton community composition than the spatially and temporally limited CHEMTAX output. However, fine scale spatial features and latitudinal and longitudinal trends are more apparent in the CHEMTAX data than in the OLCI outputs (see **Figure 13**). This loss of information is potentially a factor of the OLCI-derived phytoplankton community composition representing the median value from a 5x5 pixel window, whereas the HPLC pigment samples represent 1–4 L of samples within this pixel matchup. Furthermore, OLCI imagery undergoes data gaps from cloud cover and smoothening of trends as a result of spatial and temporal binning. Nevertheless, our results are in line with previous findings on phytoplankton community composition in the area showing the dominance of haptophytes, lack of diatoms, and largely invariant communities and biomass across much of the BFe-limited offshore region (Booth et al., 1993; Fujiki et al., 2009; Peña et al., 2019; Peña & Nemcek, 2021). As expected, our OLCI also showed

increases in Diatoms/Dino/GA contributions towards the Haida Gwaii Shelf (Waite & Mueter, 2013) and decreases in haptophytes (Peña & Nemcek, 2021), which were mirrored in the CHEMTAX data. Furthermore, the notable increases in cryptophyte contributions towards the northern GoA were observed in both the CHEMTAX and OLCI outputs. Considering this comparability, and despite the shortcomings, our results show the potential of EOF-based algorithms in retrieving phytoplankton community composition from OLCI over a large spatial-temporal scale. Characterizing phytoplankton community composition and their seasonality in the SNEP is important for better understanding match/mismatch dynamics, assessing environmental pressures on salmon stocks, and predicting their survival (Myers et al., 2016). Our results suggest that EOF methods can provide valuable information on phytoplankton community composition over spatial and temporal scales not achievable via ship-based methods. These findings are significant given the declining populations of many fish species in this region, including Pacific Salmon (Burger et al., 2015; Bendriem et al., 2019).

3.5.3 Influence of Environmental Drivers on Phytoplankton Community Composition

The study region is characterized by distinct water masses (Pakhomov et al., 2022), with the southern GoA being influenced by the Subtropical North Pacific Current (SNPC) and the northern region dominated by North Pacific Drift Current (NPDC) and the Alaska Gyre (Yuan & Talley, 1996; Whitney et al., 2007; Thomson & Krassovski, 2010). The SNPC is generally characterized by warm, low salinity, and nutrient-poor waters, while the NPDC is marked by low salinity and nutrient-rich cold waters (Thomson & Krassovski, 2010). In addition, the continental shelf domain of southeast Alaska (sampled only in 2020) is strongly influenced by the ACC, and a divergence zone exists between these domains (Pakhomov et al., 2022). Following expected trends, this study found that the 7° C isotherm and the corresponding 32.45 PSU isohaline frontal structure demarcated the boundary between the NPDC and the SNPC with higher and lower nutrients, respectively (**Figure 11**) (Yuan & Talley, 1996). Furthermore, the sampled shelf and coastal waters off of Haida Gwaii displayed markedly low salinities, shallow MLDs, and higher nutrients and showed the strongest differences in phytoplankton communities (see **Figure 15**). Outside of the differences observed between the open ocean and coastal waters, few differences in phytoplankton compositions were observed across the distinct open ocean gradients with exception to increased cryptophytes at the very northern extents of the study area in 2019.

Generally, in the Alaska Gyre dominated offshore waters, phytoplankton biomass, primary productivity, and seasonal differences in species composition are largely dependent on BFe availability (Miller et al., 1991; Zhang et al., 2021). Evidence of iron limitation has been provided by the Subarctic Ecological Response to Iron Enrichment Study (SERIES) and observations of aerial deposition of BFe resulting in increased phytoplankton biomass largely made up of > 20 μm diatoms (Boyd et al., 1998; Bishop et al., 2002; Marchetti et al., 2006; Hamme et al., 2010; Nishioka et al., 2021). Given the natural low BFe in these waters, diatoms are limited, and small-flagellates, such as haptophytes, dominate throughout the year (Yang et al., 2018; Peña et al., 2019; Zhang et al., 2021). While iron deficiency is the dominant regional driver of phytoplankton communities, other variables, such as sunlight, SST, MLD, inorganic nutrients, and grazing, also play a notable role in controlling regional phytoplankton growth and composition (e.g., Harrison et al., 2004; Henson, 2007; Peña & Varela, 2007; Aguilar-Islas et al., 2016; Li et al., 2018; Peña et al., 2019; Wang et al., 2019; Wu et al., 2020; Zhang et al., 2021). In particular, light plays a critical role in determining phytoplankton production in offshore waters during winter (Maldonado et al., 1999; Strom et al., 2010) as low irradiance and deep mixing (MLD = 130 m) results in the dilution of phytoplankton population, reducing the light available for their growth and photosynthesis (Zhang et al., 2021).

Latitudinally, our RDA analysis indicated that DSi (rather than DIN) and SST were strong drivers of the variability in the 2019 offshore waters, and cryptophyte contributions increased with DSi towards the northern portion of the study region; however, increased cryptophyte contributions were largely constrained to the northern borders of the study area while DSi increased further south following the prevailing water mass (see **Figure 12E**). It has been shown that in marine environments, algal species are found to be correlated with inorganic nutrient concentrations (e.g., Alexander et al., 2015; Endo et al., 2018), and nutrients play a significant role in shaping the phytoplankton community composition (Margalef, 1995). However, nutrient concentrations were found to be above limiting conditions throughout the region and cryptophytes are not generally associated with DSi uptake (Xin et al., 2023). Alternatively, it is possible that the increased cryptophyte contributions and TChla observed in the northern extremes of the 2019 study region, may have been a result of two influences: (1) the presence of the ACC in this region in this year, and (2) the presence of two large eddies in the northern GoA (Stabeno et al., 2016; Pakhomov et al., 2022; Ross & Peña, 2022). Transition zones have been found to be important regions for increased nanoplankton

production through increases in nutrients, BFe, and the influences of selective grazing (Juranek et al., 2020). For example, the transition zone between the Alaska Gyre and coastal waters off of the Aleutian Islands has been shown to be an important region for the production of nano-sized phytoplankton, which cryptophytes generally fall into, and could have important implications for regional CO_2 export (Palevsky et al., 2013). Furthermore, increased cryptophyte contributions have been found in the transition zone between coastal and oceanic waters in the California Current, and nanoplankton biomass is relatively high across the north Pacific subtropical-subpolar transition zone (Chavez et al., 2000; Peterson et al., 2011; Juranek et al., 2020). As a result, the extended distribution of the ACC in 2019 may have resulted in an extended transition zone transporting coastal influenced waters with higher BFe further offshore into the northern SNEP. Second, eddies in the northern GoA have been shown to entrain nutrients and BFe-rich coastal waters and carry them outwards into the HNLC region (Crawford et al., 2005). In 2019, two large eddies (Sitka and Yakutat eddies) were observed directly to the north of the study region with the notable feature observed in sigma-t south of 55° N potentially associated with these eddies (Pakhomov et al., 2022; Ross & Peña 2022). Therefore, the presence of these eddies could have driven the observed increase in cryptophyte and phytoplankton biomass by introducing coastal influenced waters with higher BFe and DSi (Batten & Crawford, 2005). These observations suggest that the position of the ACC and/or the presence of eddies (and the interactions between them) could have far ranging implications on phytoplankton biomass and community structure even during light-limited winter conditions. These trends and features were captured with the utilized satellite algorithms highlighting that the methods derived in this study could provide valuable high spatial-temporal information for further investigation of cryptophyte contributions and their relation to the position of the ACC and eddies in the northern regions of the GoA.

Eddies may have also played a role in 2020, with satellite-based sea level anomaly maps showing the presence of two large Haida eddies intersecting the study region this year (Ross & Peña, 2022). In this year, our phytoplankton pigment sampling was limited in the regions where the eddies were observed, making the investigation of their influences difficult. When compared to 2019, the available data suggest reduced contributions and biomass of many of the main groups in the vicinity of the elevated DIP feature. Yet, the satellite outputs from this year (**Figure 15**), which have high spatial resolution, did not show notable trends associated with the eddies. Regional eddies have the potential to drive large changes in phytoplankton

community composition and biomass; however, these changes are generally observed following winter conditions when light limitation has been lifted and there is reduced mixing (Peterson et al., 2011). As a result, the lack of observable influence by the eddies in 2020 could have been a factor of the early sampling period and changes being too small to be picked up via the satellite algorithms (Ueno et al., 2023). Adding to this lack of sensitivity is that high cloud cover in winter meant that there were many cloud-affected pixels and in conjunction, the weekly binning of the imagery may have diluted small- and fine-scale spatial features that were present during the low biomass winter conditions. It is possible that daily imagery obtained later in the season would have been more successful at discerning the changes in phytoplankton community composition and biomass driven by these eddies.

Beyond mesoscale eddies and other environmental limitations, zooplankton can exert top-down control, inhibiting regional phytoplankton biomass accumulation (Landry et al., 1993; Miller, 1993) and shaping the community composition via selective grazing (Strom & Welschmeyer, 1991; Landry et al., 1993). Pakhomov et al. (2022) noted that the highest zooplankton abundances (~400 ind/m^3) were found within this region in both years, with 2020 having notably higher zooplankton (~650 ind/m^3) and micronekton abundances. In 2019, reduced contributions and biomass of TChla, pelagophytes, and cryptophytes were observed in the vicinity of increased zooplankton abundances (Pakhomov et al., 2022). Interestingly, the few 2020 open ocean phytoplankton samples collected in this study were focused in southern regions and tended to show reduced biomass when compared to 2019, notably for haptophytes and pelagophytes. Therefore, the latitudinal spatial gradients observed in many phytoplankton community composition within the offshore waters may have been partly driven by increased grazing in the southern subtropical water mass. Additional studies combining high spatial-temporal phytoplankton community composition satellite imagery with zooplankton data could provide further insights into the role of top-down forcing in shaping the phytoplankton community composition.

Phytoplankton community composition and biomass showed strong longitudinal trends in 2020 with increased contribution by diatoms and chlorophytes on the continental shelf. These groups were negatively correlated with SSS and MLD (observed in the RDA), with the shelf waters displaying marked reductions in salinity, likely driving stratification and reduced MLDs and indicative of coastal/freshwater influences. Unlike flagellates, diatoms prefer highly dynamic nutrient-abundant coastal regions due to their high sinking rate and affinity to

BFe and silicate, which are essential for their growth (Wong & Matear, 1999; Marchetti et al., 2006; Nishioka et al., 2021). The coastal waters of southeast Alaska are enriched with iron and silicate sourced from bottom waters, river runoff, and aerial dust deposition (Johnson et al., 2005; Whitney et al., 2005; Strom et al., 2016), and as a result, would have provided the conditions necessary for the observed diatom growth. The shelf and coastal waters off of Haida Gwaii and southeast Alaska are characterized by seasonal spring blooms dominated by diatoms where TChla often increases to > 3 mg/m^3 (e.g., Brickley & Thomas, 2004; Ribalet et al., 2010; Waite & Mueter, 2013; Jackson et al., 2015; Strom et al., 2016; Du et al., 2022; Marchese et al., 2022; Vishnu et al., 2022); however, our observed contributions were considerably lower (< 0.51 mg/m^3) likely due to pre-bloom conditions driven by light limitation. Furthermore, despite reduced contributions, small flagellates were still important within the coastal/shelf waters, showing their adaptability to a broad range of conditions. Small flagellates can generally thrive under numerous conditions due to various factors, giving them unique competitive advantages over larger species such as diatoms (Gregg & Casey, 2007). For example, small flagellates are efficient light harvesters due to their small pigment packaging making them more competitive in highly mixed and low-light winter conditions (Margalef, 1995; Marañón et al., 2012; Alexander et al., 2015; Endo et al., 2018). Furthermore, these small species have higher surface-to-volume ratios allowing them to thrive under low nutrient conditions where they are able to absorb nutrients more efficiently (Marañón et al., 2012). For example, haptophytes are particularly well suited to highly mixed, low light, and BFe deficient conditions (Endo et al., 2018), potentially explaining their observed dominance across the latitudinal offshore environmental gradients and within the coastal/shelf waters where BFe was likely available. The high spatial and temporal resolution OLCI-derived phytoplankton community composition data in this study were able to discern the observed latitudinal gradients and highlight increased shelf diatom contributions. As a result, these dataset will greatly aid in studying the varying influences of coastal and offshore waters on phytoplankton dynamics.

3.5.4 Limitations in Our In Situ and Satellite Data

Multiple sources of uncertainty are associated with in situ pigment/CHEMTAX and OLCI methods for deriving phytoplankton community composition in ocean waters. In terms of HPLC phytoplankton pigment measurements, differences in analytical methods used between studies have been shown to introduce uncertainties as high as 7 and 21% for TChla and accessory pigments, respectively (Claustre et al., 2004). These uncertainties are important to

consider when the methods used in this study are applied to other dataset and regions where HPLC pigment data were analyzed using different approaches (Claustre et al., 2004).

Further, uncertainty can also be introduced by CHEMTAX analysis, which relies on assumptions that are often violated (Catlet & Siegal, 2017; Moore & Brown, 2020). For example, CHEMTAX assumes that pigment/TChla ratios remain constant across time and space; however, recent studies have demonstrated that these ratios can vary within a species due to genetic differences and their response to the environment (Higgins et al., 2011). Additionally, CHEMTAX assumes that the abundance of accessory pigments is not correlated; yet, recent findings have shown that they are strongly and highly correlated with each other (Catlet & Siegal, 2017; Kramer et al., 2019). These uncertainties are generally not well resolved as there is no "gold-standard" approach to measure phytoplankton communities, and other approaches have their own unique uncertainties and shortfalls, making comparison difficult (Kramer & Siegal, 2019; Chase et al., 2022). In terms of CHEMTAX, the analysis is highly dependent on the correct selection of input phytoplankton groups and ratios, and thus, strong *a priori* knowledge of the studied phytoplankton communities is required. In this study, no comparison with taxonomic data was available; however, care was taken to select groups and ratios that have been used successfully by other authors within the region and the results were in line with expected conditions (Peña et al., 2019). Second, CHEMTAX assumes that pigment ratios remain consistent through the studied conditions and observed communities; however, pigment ratios can be variable within the study area due to various factors, including light and nutrient conditions (Swan et al., 2016). Because our study was carried out in winter, when variations in surface light and nutrient conditions were low, this assumption was likely not violated. Here, care was taken to limit this issue by grouping data with similar pigment ratios for CHEMTAX analysis. Third, error can be introduced when complex communities are present with many species having similar pigment profiles (e.g., haptophytes, pelagophytes, silicoflagellates, and some dinoflagellates can show HF and BF as dominant pigments), cryptic species that show comparable pigment profiles to broader groupings (e.g., Bolidophytes and diatoms), and/or mixotrophic species that contain endosymbionts (e.g., mesodinium rubrum and cryptophyte-like pigments) (Vaillancourt et al., 2018). Under these conditions, CHEMTAX outputs may have misclassifications, and biomass from unconsidered species will be assigned to the most comparable CHEMTAX grouping. Finally, there is evidence that satellite algorithms may not be able to discern CHEMTAX groupings with low contributions and/or that covary within the

environment (e.g., Vishnu et al., 2022). To deal with this and some of the above issues, novel data-driven approaches have been developed that cluster species with similar pigment profiles or collinearity into fewer groupings (< 5) such as “small flagellates” (Catlett & Siegal, 2018; Kramer & Siegal, 2019; Kramer et al., 2020). Hence we followed this approach for the satellite application and limited the output of our study to three distinct groupings. Yet, of note, these broad groupings may need to be revised for further analysis including temporally resolved data. For example, the derivation of the diatom/dinoflagellate/green algae grouping was a consequence of these groups showing increased importance in the coastal samples collected under winter conditions; however, these groups would likely not have clustered together if spring and summer coastal data were included in the clustering analysis as they show contrasting seasonal trends. For example, regional coastal waters generally show diatom dominated spring blooms followed by succession to flagellate dominated conditions including important contributions by green algae and dinoflagellates (Del Bel Belluz et., 2021; Nemcek et al., 2023). Consequently, the application of diatom/dinoflagellate/green algae grouping to imagery from these seasons would likely result in the obscuration of important trends in phytoplankton community compositions, notably in coastal regions.

Additional uncertainty can arise from the development of the regression models used to derive phytoplankton community composition from the OLCI imagery. Here, this error was quantified via the cross-validation method (**Table S3**). The results from the cross-validation procedure were comparable with the full-fit statistics indicating that the number of matchup points was sufficient to develop a robust regression model. In this study, the number of matchup points (59–61) was in line with those used by other authors using similar approaches (e.g., Bracher et al., 2015; Lange et al., 2020; Xi et al., 2020; Vishnu et al., 2022).

OLCI Rrs(λ) are also a source of uncertainty that could propagate into the final phytoplankton community composition estimates derived from the satellite. For example, Pahlevan et al. (2021) demonstrated that instrument noise and atmospheric correction uncertainty could lead to 25–70% errors in products derived from multispectral sensors. Furthermore, satellite-derived Rrs(λ) retrievals from clear Case-1 oligotrophic waters have been shown to have 5% uncertainty in the blue band with increases towards the red bands (Hu et al., 2013). In agreement with these findings, our in situ and OLCI-derived Rrs(λ) daily matchups showed better radiometric performance in the blue band (r = 0.73, BIAS = 0.22%), followed by the green band (r = 0.50, BIAS = -13%), and then inferior radiometric

performance in the red band (r = -0.2, BIAS = 11%). These results have been further substantiated by more recent studies using similar methods (Tilstone et al., 2021). Finally, Xi et al. (2021) quantified per-pixel uncertainty of global satellite-derived phytoplankton group estimates from multiple sources and found similar uncertainty distribution patterns for diatoms, dinoflagellates, green algae, and haptophytes that were lower in oligotrophic waters and increased towards the coast. On a global scale, the overall uncertainty obtained for haptophyte (the base 10 logarithmic uncertainty) varied from 0.12 to 0.58; for green algae, it ranged between 0.18 to 0.63. Diatoms and dinoflagellates showed the highest uncertainties, with values ranging from 0.20 to 0.66 and 0.27 to 0.71, respectively.

Finally, another limitation of our study was the absence of per-pixel uncertainty for our OLCI output. To our knowledge, Brewin et al. (2017) and Xi et al. (2021) are the only studies to report per-pixel uncertainty for satellite-derived phytoplankton groups. These studies propagated (except Brewin et al., 2017) the input satellite-derived Rrs uncertainty when assessing the retrieval accuracy of phytoplankton community composition derived from ocean color sensors. Unfortunately, uncertainty is unavailable in POLYMER-derived OLCI Rrs(λ), making it impossible to investigate here. The derivation of this uncertainty is currently in the development stage and as a result, future studies will be able to investigate the effect of this error when using POLYMER-derived OLCI Rrs(λ) to estimate phytoplankton community composition.

3.6 Conclusion

Our study region in winter exhibited distinct water masses, with warm, low-salinity subtropical water to the south and cold, high-salinity, nutrient-rich subarctic water to the north. Furthermore, strong gradients were observed between shelf/coastal and open ocean waters with lower salinity and MLDs and higher nutrients towards Haida Gwaii. In general, our data showed homogeneity of phytoplankton community compositions in the iron-poor open ocean regions across the strong latitudinal gradients observed in the physicochemical data. In contrast to this homogeneity was the distinct increase in cryptophyte contributions at the northern extremes of the study region which were potentially driven by the presence of the ACC and/or Sitka and Yakutat eddies. Furthermore, strong longitudinal gradients were observed with the coastal/shelf waters directly off of Haida Gwaii showing increased contributions, notably of diatoms strongly driven by MLD and SSS, and suggesting the influence of coastal water with increased BFe.

The use of broad phytoplankton groupings from the CHEMTAX-derived data for the development of the satellite-based EOF algorithm showed potential for deriving multiple phytoplankton groups from space. Statistically significant retrievals were obtained for the Hapto/Pelago/Cyano, Crypto and Diatoms/Dino/GA groups including TChla concentrations. Additionally, the satellite retrievals showed similar pattern to the in situ CHEMTAX-derived data with the dominant contribution and homogeneous distribution of the Hapto/Pelago/Cyano group across the open ocean region, increased Diatoms/Dino/GA and TChla towards the coast, and increased Crypto contributions in the northern extents of the study region and off of Haida Gwaii (**Figure 15**). When compared to Line-P CHEMTAX-derived phytoplankton community composition data, which was used as an independent validation source, there was strong agreement for the Hapto/Pelago/Cyano group and TChla. These results show promise of the developed algorithms for monitoring phytoplankton community composition in the SNEP and further work across broader spatial-temporal scales (including the productive season) will likely highlight more prominent trends.

Deriving phytoplankton community composition using satellite data is extremely important for regions such as SNEP as they provide habitat for commercially, culturally, and economically important fish species including Pacific Salmon originating from Canada, the United States, Japan, Russia, and Korea. In situ data collected using conventional ship-based surveys are essential but often limited in space and time, and this can be addressed only with ocean color satellites. Many recent studies have shown important links between phytoplankton and higher trophic levels in British Columbia waters (i.e., Malick et al., 2015 and Suchy et al., 2022); however, most of these studies are constrained to coastal regions and little is known about these relationships in the open ocean. The methods developed in this study could help to fill these gaps and provide valuable information to fisheries and ecosystem managers.

Chapter 4: 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 Sciences, 9(1018510), pp. 1-22.

https://doi.org/10.3389/fmars.2022.1018510

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

[1]SPECTRAL Remote Sensing Laboratory, Department of Geography, University of Victoria, Victoria, British Columbia, Canada.

[2]Phytooptics Group, Physical Oceanography of Polar Seas, Climate Sciences, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany.

[3]Hakai Institute, Victoria, British Columbia, Canada.

[4]Department of Marine Biology, Microbiology, and Biochemistry, Cochin University of Science and Technology, Kochi, Kerala, India.

[5]Department of Physics and Electrical Engineering, Institute of Environmental Physics, University of Bremen, Bremen, Germany.

Abstract

Monitoring the spatial distribution and seasonal dynamics of phytoplankton functional types (PFTs) in coastal oceans is essential for understanding fisheries production, changes in water quality, and carbon export to the deep ocean. The launch of new generation ocean color sensors such as OLCI (Ocean Land Color Instrument) onboard Sentinel 3A provides an unprecedented opportunity to study the surface dynamics of PFTs at high spatial (300 m) and temporal (daily) resolution. Here we characterize the seasonal dynamics of the major PFTs over the surface waters of the west coast of Canada using OLCI imagery and Chemical Taxonomy (CHEMTAX, v1.95) software. The satellite-based approach was adapted from a previously proven Empirical Orthogonal Function (EOF)-based algorithm by using a local matchup dataset comprising CHEMTAX model output and EOF scores derived from OLCI remote sensing reflectance. The algorithm was developed for the following PFTs: diatoms, dinoflagellates, dictyochophytes, haptophytes, green algae, cryptophytes, cyanobacteria, raphidophytes, and total chlorophyll-a (TChla) concentration. Of these PFTs, first level evaluation of the OLCI-derived retrievals showed reliable performance for diatoms and raphidophytes. The second level of validation showed that TChla had the best performance, and green algae, cryptophytes, and diatoms followed seasonal trends of a high temporal resolution in situ CHEMTAX time-series. Somewhat reduced correspondence was observed for raphidophytes. Due to their combined low contribution to the phytoplankton community (26%) and low range of variation, weak performance was noted for haptophytes, dictyochophytes, cyanobacteria, and dinoflagellates. The EOF-based PFT maps from daily OLCI imagery showed seasonal spring and fall diatom blooms with succession from spring blooms to high diversity flagellate dominated summer conditions. Furthermore, strong localized summer raphidophyte blooms (*Heterosigma akashiwo*) were observed, which are a regionally important harmful species. Overall, this study demonstrates the potential of the OLCI in deriving the surface dynamics of major PFTs of the SoG, a critical habitat for the juvenile Pacific Salmon.

4.1 Introduction

Phytoplankton represent only 0.2% of global primary producer biomass; however, they are responsible for more than 90% of the ocean's primary productivity (Chassot et al., 2010), including half of the biosphere's net primary productivity (Field et al., 1998). These ubiquitous single-celled autotrophs are the base of the marine food web (Falkowski, 2012), and are capable of fixing ~50-130 million tons of atmospheric carbon dioxide (CO_2) daily into the form of organic materials, playing a central role in the global biogeochemical cycles (Behrenfeld et al., 2006). Numerous in situ methods are available to quantify size fractions and community composition of phytoplankton in the ocean, yet these observations are time-consuming, labor-intensive, and sparse in space and time (IOCCG, 2014; Chase et al., 2017).

Space-based optical sensors measure water-leaving radiance in multiple wavebands, allowing the retrieval of the surface total chlorophyll-a (hereafter TChla), a proxy for phytoplankton biomass, with high spatial and temporal resolutions (McClain, 2009). Nevertheless, TChla alone is not sufficient to describe the ecological and biogeochemical processes in the ocean (Bracher et al., 2017). Extracting information about phytoplankton functional types (PFTs), and phytoplankton size classes (PSCs), is essential for understanding the linkage between phytoplankton composition and fisheries, and in determining the role of the ocean in a changing climate (e.g., Falkowski et al., 1998; Le Quere et al., 2005; IOCCG, 2014). Considering the highly diverse nature (e.g., taxonomical, morphological, and physiological) of marine phytoplankton, climatologists and marine biogeochemical modelers have defined PFTs according to their specific function or role in the biogeochemical cycles irrespective of their taxonomical affiliations (Le Quere et al., 2005; Nair et al., 2008). For example, diatoms are considered to be essential silicifiers, cyanobacteria as nitrogen fixers, Haptophytes as calcifiers, and Dinoflagellates/Haptophytes/Green Algae as dimethyl sulphide producers (DMSP) (IOCCG, 2014).

Given the importance of determining phytoplankton from space, several satellite-based algorithms have been developed to identify the significant PFTs (e.g., Alvain et al., 2005; Navarro et al., 2014; Correa-Ramirez et al., 2018), derive the chlorophyll-a concentration (hereafter Chla) of PFTs (e.g., Bracher et al., 2009; Sadeghi et al., 2012; Moisan et al., 2017), retrieve PSCs (e.g., Ciotti et al., 2002; Kostadinov et al., 2009; Mouw and Yoder, 2010; Brewin et al., 2010; Devred et al., 2011; Hirata et al., 2011; Brotas et al., 2013; Roy et al., 2013; Lamont et al., 2018; Moore and Brown, 2020; Liu et al., 2021), and retrieve accessory

pigment concentrations (e.g., Pan et al., 2010; Bracher et al., 2015; Wang et al., 2018; El Hourany et al., 2019; Stock and Subramaniam, 2020; Kramer et al., 2022). These satellite-based algorithms explore specific signatures, such as anomalies in the reflectance spectra, inherent optical properties (IOPs), and TChla concentration (see IOCCG, 2014; Mouw et al., 2017; Bracher et al., 2017 for a detailed review). Of the many approaches developed, the Empirical Orthogonal Function (EOF) approach has been successfully used among these algorithms (e.g., Craig et al., 2012). The EOF approach is a statistical technique that reduces the high dimensionality of spectral data into dominant modes, explaining the variance of structures in the data. Specifically, the EOF approach allows the extraction of the spectral signature associated with specific PFTs, thus enabling the development of empirical models for various PFTs (e.g., Xi et al., 2020). This approach has been successful in retrieving the concentration of different accessory pigments from Medium Resolution Imaging Spectrometer (MERIS) and hyperspectral data (Bracher et al., 2015), Chla concentration of multiple PFTs using GlobColor merged and OLCI data (Xi et al., 2020), and cell abundance of *Prochlorococcus*, *Synechococcus*, and autotrophic picoeukaryotes based on multispectral and hyperspectral data (Lange et al., 2020). However, most of these studies were performed in open oceans or optically Case-1 type waters, with less application to coastal systems (e.g., Craig et al., 2012).

The coastal ocean of British Columbia, Canada, specifically the Strait of Georgia, is a spatially and temporally dynamic Case-2 water body (Phillips and Costa, 2017) crucial for Pacific salmon, herring, and hake fisheries (Beamish et al., 1994; Beamish et al., 2012; Thomson et al., 2012), which are of vital importance to the regional economy and First Nations (Nesbitt and Moore, 2016). In these waters, phytoplankton have been shown to be associated with the growth and survival of these species (Cushing, 1990) through their bottom-up control of key zooplankton species such as *Neocalanus plumchrus* (e.g., Sastri and Dower, 2009; Perry et al., 2021; Suchy et al., 2022). Hence, spatial-temporal satellite-derived PFTs will allow for a more coherent understanding of the function/role of multiple PFTs on the regional food web.

Here, we adapted an EOF-based algorithm developed by Xi et al. (2020) to the Strait of Georgia, British Columbia (BC), Canada, to retrieve TChla concentration and the Chla concentration of multiple PFTs. The model was developed for each PFT using the matchup between CHEMTAX-derived PFTs and EOF scores derived from the OLCI remote sensing reflectance (Rrs; $sr^{-1}(\lambda)$. The robustness of the fitted regression model was assessed

statistically by the cross-validation procedure, and independent validation of the major derived PFTs was performed against an independent in situ CHEMTAX time-series. Further, the seasonal and spatial dynamics of major PFTs were determined by applying the regression model to daily OLCI imagery.

4.2 Materials and Methods

4.2.1 *Study Area*

The Strait of Georgia (SoG) is the largest semi-enclosed estuarine system in the Salish Sea, located on the west coast of Canada (**Figure 19**). This partially enclosed region has a total length of 222 km, a maximum width of 28 km, and a depth of about 350 m within its central region (Masson and Peña, 2009), which is generally divided into three marine areas: the southern, central, and northern Strait. The SoG is connected to the Pacific Ocean through the Johnstone Strait in the northern region and Juan de Fuca Strait in the southern boundary (Pawlowicz, 2002). An estuarine circulation typically drives the SoG regional dynamics, characterized by low saline surface waters from the adjacent Fraser River and highly saline nutrient-rich subsurface waters from the Pacific Ocean (Li et al., 2000; Masson and Cummins, 2004). The largest freshwater discharge to the SoG is from the Fraser River, which contributes 75% of the freshwater to the Strait, and is primarily responsible for the water column stratification and estuarine-like circulation in the central and southern Strait (Harrison et al., 1983; Johannessen et al., 2003; Pawlowicz et al., 2007). Incoming seawater dominantly enters the SoG through the southern Juan de Fuca Strait (Masson, 2002), bringing nutrient-rich water from the Pacific Ocean through the deep water inflow and is eventually upwelled into the surface waters by rigorous tidal mixing in the Haro Strait (Li et al., 2000). In the summer, the northern Strait becomes highly stratified, limiting surface nutrient renewal. This stratification is thought to be a function of high summer surface heating, low summer winds, and localized freshwater discharge, and can periodically be disrupted by strong wind events (Evans et al., 2019).

This region is one of the most productive estuaries in the Pacific North East with an average annual primary productivity of about 280 g C m^{-2}, which is the highest in the frontal regions and close to the Fraser River (Harrison et al., 1983; Sutton et al., 2013; Johannessen et al., 2021). Seasonally, SoG surface TChla greatly varies with winter concentrations <1 mg m^{-3} and increases to > 15 mg m^{-3} during spring (Harrison et al., 1983; Jackson et al., 2015; Peña

et al., 2016; Suchy et al., 2019). As in many temperate coastal environments, spring blooms are dominated by diatoms with *Skeletonema* sp., *Thalassiosira* spp., and *Chaetoceros* spp. being the dominant bloom-forming species. By the end of spring, diatoms are usually replaced by high diversity communities constituted by flagellates such as cryptophytes, prasinophytes, and haptophytes (Harrison et al., 1983; Haigh and Taylor, 1991; Del Bel Belluz et al., 2021).

Optically, the SoG is a dynamic Case-2 water body characterized by high and spectrally dependent light attenuation with estuarine waters comprised of high loads of suspended inorganic sediment and dissolved organic matter from the Fraser River discharge (Komick et al., 2009; Loos and Costa, 2010; Phillips and Costa, 2017). For these highly attenuating waters, particulate scattering varies from approximately 0.2 to 16.0 m^{-1} in the blue spectral region (Loos and Costa, 2010), and colored dissolved organic matter (CDOM) absorption at 443 nm varies from 0.007 to 3.07 m^{-1} (Phillips and Costa, 2017). On the other hand, total attenuation in the northern SoG is predominantly controlled by absorption by CDOM and TChla (Loos and Costa, 2010; Loos et al., 2017; Phillips and Costa, 2017).

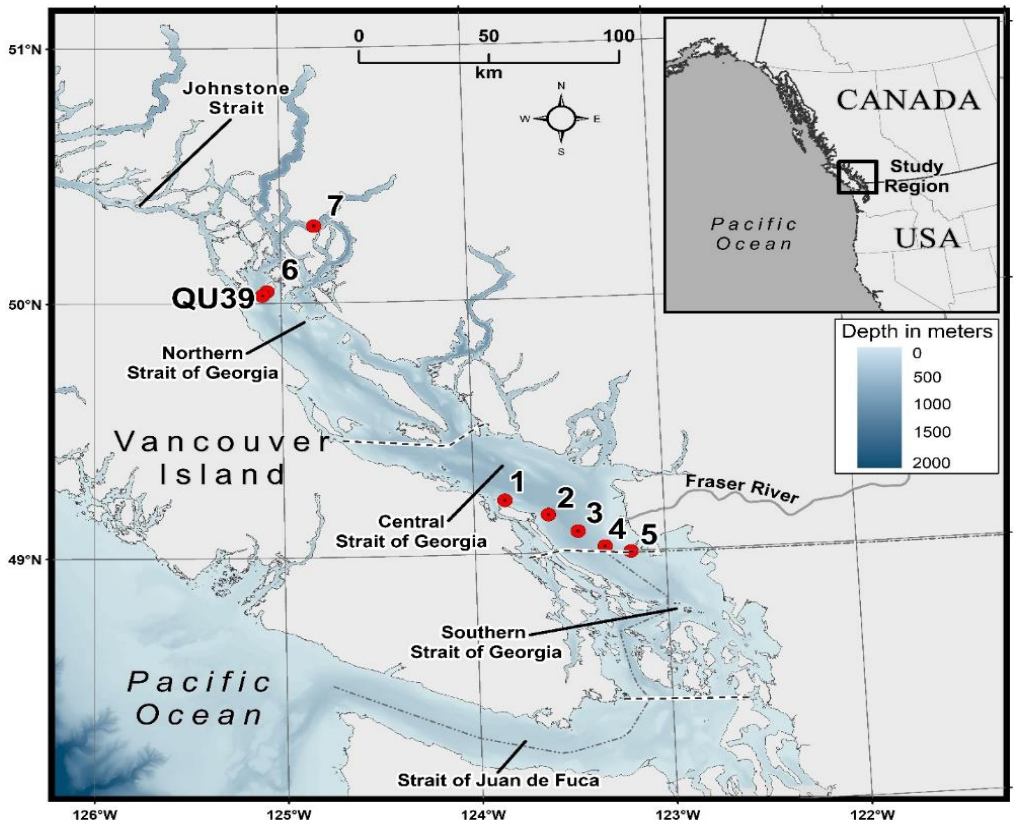

Figure 19. Map showing the study area. The red circles indicate the sampling locations in the central and northern Strait of Georgia (SoG). QU39 represents the Hakai time-series station. Bathymetry is shown in meters.

4.3 Data Acquisition

4.3.1 In situ Data: Water Sampling and Analysis

Surface water samples were collected at two regions of the SoG. For the first location, samples (N=108) were acquired along the BC ferry route aboard the *Queen of Alberni* (QA) crossing from Duke Point (Nanaimo) to Tsawassen (Vancouver; Stations 1-5, **Figure 19**) during the spring (March to May) and summer (June to September) of 2018 and 2019. Water samples were collected using a seawater intake pump installed on the ferry, representing a well-mixed layer of <3 m (Halverson and Pawlowicz, 2013; Wang et al., 2019; Travers-Smith et al., 2021). Triplicate samples from each station were collected for the different analyses (HPLC pigment concentration, CDOM absorption, and total suspended matter (TSM) concentration), kept in a dark environment, and immediately filtered onboard to minimize any potential compositional and concentration changes to phytoplankton pigments and dissolved organic matter (Mueller et al., 2003). For the second location, the northern SoG (Stations 6-7, **Figure 19**), surface water samples (~ 2 m depth) (N=19) were collected during summer (July) as part of an algal bloom monitoring component at aquaculture sites. Triplicate water samples were collected using a 5 L Niskin water sampler, transferred to a 10 L dark container, transported to the laboratory in the dark, and immediately filtered for HPLC pigments, CDOM absorption, and TSM. In addition, at the northern SoG location, single samples were collected for microscopic identification and enumeration of phytoplankton species (N = 8). Water for these samples was immediately transferred to 250 mL amber glass bottles and fixed using Lugol's solution (1% concentration); samples were kept cool and dark until analysis.

Surface water samples for the TSM were filtered through precombusted and pre-weighed 47 mm Whatman GF/F 0.7 µm filter paper. Precombustion of the filter paper increases the efficiency by retaining more particulate matter as it shrinks the pore size to 0.3 µm during combustion (Nayar and Chou, 2003). Based on the color of the seawater, filtered volumes varied from 0.5 – 2 L. After the sample filtration, the filter paper was washed with 200 ml of distilled water to remove sea salt (Stavn et al., 2009), and immediately stored in dry ice, transported to the laboratory, and kept under -80°C until analysis. The filter paper was dried at 60°C for 24 hours in a clean oven and subsequently reweighed to 6 digits of precision, at

which point TSM was calculated (Röttgers et al., 2014). The final TSM concentration at each station was the average of the triplicates with a coefficient of variation (CV) < 25%.

Surface water samples to analyze CDOM were filtered through 47 mm Whatman 0.2 µm membrane filter paper. The filtrate was stored in acid-washed, precombusted (550°C for 1 hour) borosilicate glass bottles (Ferrari et al., 1996). Glass bottles were triple rinsed with sample water before the collection, and samples were kept in a cold environment, transferred to the laboratory and kept under 4°C until analysis. Prior to analysis, samples were allowed to reach room temperature to minimize temperature differences between the samples and the blank. The absorption of CDOM was measured using an Ocean Optics USB 4000 spectrophotometer (300-850 nm, 10 cm path length) against Milli-Q water as blank. The sample absorption was converted to CDOM absorption using Beer-Lambert's Law equation (Ferrari et al., 1996).

$$a_{CDOM}(\lambda) = \frac{2.303 * A(\lambda)}{L} (m^{-1}) \quad (1)$$

A(λ) is the spectral absorption of the sample, and L is the cuvette path length in meters. The final a_{CDOM} (λ) for each station is the average of the triplicate with a CV < 25%.

Phytoplankton microscopy was performed by LCJL Marine Ecological Services using the Utermöhl method. Phytoplankton were identified and enumerated using 50 mL settling chambers and phase contrast microscopy with an inverted light microscope (Utermöhl, 1958). Count data were reported in cells/L. Counts from species within PFT groups used for analysis were summed to make comparisons with pigment-based estimates.

Surface water samples to measure HPLC pigments were filtered through 25 mm Whatman GF/F 0.7 µm glass microfiber filters under a low vacuum (≤5 mmHg). Based on the color of the seawater, filtered volumes varied from 0.5 – 2 L. Filters were immediately put in dry ice and kept under -80°C until laboratory analysis. TChla and phytoplankton accessory pigments were analyzed using HPLC analysis done at the University of South Carolina Baruch Institute (https://phytoninja.com/lab-protocols/), following the methods described in Pinckney (2010). HPLC samples were collected in duplicates, and the uncertainty in the HPLC pigment dataset was computed from the duplicate measurements for 20% of the samples, resulting in a CV < 20%.

The HPLC pigment dataset was used as the input in the CHEMTAX software (v1.95) to derive the Chla concentration of each PFT towards the TChla concentration of the sample

(Mackey et al., 1996). The selection of PFT and input pigment ratio were largely based on those in Del Bel Belluz et al. (2021), which were defined via in-depth analysis of pigments and pigment ratios, microscopic observations, and literature review of phytoplankton species in the SoG.

Unlike Del Bel Belluz et al. (2021), a raphidophyte group was included in the analysis to account for the extensive 2018 *Heterosigma akashiwo* blooms observed via microscopy at comparable sampling locations and times as those investigated in this study (Esenkulova et al., 2021). Additionally, our 2018 observations showed high violaxanthin concentration events where relationships with TChla (generally driven by diatom blooms), fucoxanthin (diatoms), Chlb (green algae), and zeaxanthin (typically only present in low concentrations and related to green algae or cyanobacteria) diverged from those through the rest of the dataset (**Figure S4**). During these high concentration events, fucoxanthin, violaxanthin and zeaxanthin ratios to TChla were comparable to those found for raphidophytes in the literature (**Table S5**, Lewitus et al., 2005; Higgins et al., 2011). It is unlikely that other violaxanthin containing species would bloom to such high violaxanthin and TChla concentrations in the SoG. For reference, through the four-year time-series presented in Del Bel Belluz et al. (2021), which spanned multiple high TChla bloom events, violaxanthin was largely < 0.10 mg/m^3 and reached a maximum concentration of 0.41 mg m^{-3} (TChla = 6.18 mg/m^3), an order of magnitude lower than observed during this study. For the central SoG data, pigment ratios for the raphidophyte group were taken from Lewitus et al. (2005), derived from *Heterosigma akashiwo*. In the northern SoG, pigment and microscopic observations showed that raphidophytes were absent. As a result, this group was not included in the northern SoG analysis. With these considerations, the following groups were considered in CHEMTAX analysis: cyanobacteria-1 (Cyano), haptophytes (Hapto), green algae (GA), cryptophytes (Crypto), dinoflagellates (Dino), raphidophytes (Raphido, central SoG only), dictyochophytes (Dictyo), and diatoms (Diat). Input ratios for each taxonomic group are given in **Table S6** and **Table S7**. While quantitive evaluation of CHEMTAX is often lacking, many studies have utilized similar groupings to describe phytoplankton community compositions in coastal waters (e.g., Lu et al., 2018; Ferreira et al., 2020). Here, it was decided to attempt to maximize the retrievable groups; however, it is important to note that following the methods of Catlett and Siegel (2018), clustering of the HPLC-derived pigment dataset suggests that only four PFT groups may be separatable and optically discernable (**Figure S8**). As a result,

CHEMTAX model outputs are supported by additional information, including microscopy (northern SoG) and data from complimentary studies done at similar locations and times.

CHEMTAX analysis was run on the northern and central SoG data separately to account for regional differences in pigment ratios as a result of differing oceanographic conditions (i.e., stratification, nutrients, and light; Suchy et al., 2019). An initial seeding run was first performed on the full dataset from each region to optimize the literature-derived ratios to the SoG (Armbrecht et al., 2015; Del Bel Belluz et al., 2021). Then, the central SoG data were binned into different groups based on TSM concentrations to account for the high variability in light conditions along the ferry track and between surveys. The following TSM bins were used: Low (TSM < 5 mg/L), Medium (TSM 5 – 11 mg/L), and High (TSM > 11 mg/L), based on bio-optical classes derived by Phillips and Costa (2017) for the central SoG. CHEMTAX analysis on the binned data utilized the output ratios from the initial seeding run performed on the full central SoG dataset. For the northern SoG, TSM conditions were comparable across the surveys, and HPLC data were run as a single group. Similar to the secondary runs done on the clustered data above, a second run was performed, using the output ratios from the initial run to ensure stabilized ratios and to reduce RMSE.

Beyond consideration of environmental conditions, we evaluated pigment colinearity, which can affect CHEMTAX results (Higgins et al., 2011; Catlett and Siegel, 2018). In this study, positive correlations did exist between pigments used for analysis (**Figure S5**). When standardized to TChla, many of these correlations were still strong (e.g., Chlb and prasinoxanthin, **Figure S5**); however, no correlation of pigments with TChla. Careful consideration was taken in the CHEMTAX design to limit error; however, pigment colinearity can add bias to the CHEMTAX model output and complicate retrieval via satellite (Higgins et al., 2011; Catlett and Siegel, 2018). As independent validation of CHEMTAX model outputs was not performed here, it was difficult to quantify this error, and this is further discussed in section 4.6.1. Care should be taken to extend results until extensive validation has been performed.

Following the CHEMTAX analysis, we also performed a Shannon-Weaver Index (Lohrenz et al., 2003; Latasa et al., 2010) on the CHEMTAX model output to assess the diversity of phytoplankton community composition:

$$H = \sum_{i=1}^{N} P_i \ [P_i] \qquad (2)$$

where, P_i is the proportion of each phytoplankton group's Chla relative to the TChla concentration (%), and N is the number of groups.

4.3.2 Satellite Data

High spatial (300 m) and temporal resolution (daily) OLCI data were used in this study as this sensor has improved signal-to-noise ratios and an off-nadir tilted view to minimize the effect of sun glint (Donlon et al., 2012). Therefore, OLCI Level-1 Full-Resolution data were downloaded from the Sentinel 3 Marine CODA (Copernicus Online Data Access) web service, and processed with POLYMER (POLYnomial-based algorithm applied to MERIS) version 4.10 (Steinmetz et al., 2011) to obtain Rrs. POLYMER is an atmospheric correction (AC) algorithm that separates the atmospheric and sunglint spectral reflectance from the water reflectance using a spectral matching technique (Steinmetz et al., 2011). To do so, it uses two models: a simple polynomial used to model the spectral reflectance from atmospheric and residual sunglint, and a model that delineates ocean water reflectance using TChla concentration and the particle backscattering coefficient (b_{bp}) (Park and Ruddick, 2005). The water reflectance model used in this algorithm is valid in coastal and open ocean waters, including the bidirectional reflectance, and it provides better spatial coverage retrievals because it accounts for sky glint and thin cloud conditions (Steinmetz et al., 2011). Quality flags such as "Cloud," "Invalid," "Negative Backscattering (BB)," "Out-of-bounds," "Exception," and "High Air Mass" (Steinmetz et al., 2016) were applied following Giannini et al. (2021) recommendations for this region. According to the technical document (Steinmetz et al., 2016), the "Inconsistency" flag should not invalidate the results; however, our Rrs matchup analysis showed artifacts with pixels associated with this flag. As a result, data with this flag were removed from analysis. Giannini et al. (2021) demonstrated that POLYMER-derived products such as TSM and TChla showed the best performance compared to the Case-2 Regional CoastColor (C2RCC) processor for the coastal waters of BC and Northeast Alaska. This outcome further validates the efficacy of utilizing POLYMER-derived reflectance as spectral inputs in EOF-based algorithms.

4.4 Data Analysis: PFT Satellite-based Method and Validation

The Chla concentration of PFTs was derived from OLCI satellite imagery using the EOF-based algorithm by assessing the variance or spectral shape in the OLCI Rrs spectra (Bracher et al., 2015). This algorithm was recently tested globally by Xi et al. (2020), and we further adapted it considering the local matchups between OLCI Rrs and the CHEMTAX-derived

PFTs Chla. A schematic workflow of the EOF-PFT algorithm is shown in **Figure 20**. The general steps of this approach are explained as follows:

Step 1: Each Rrs (400, 412, 443, 490, 510, 560, 620, 665, 674, 681 nm) spectrum was standardized by subtracting the spectral average and normalized by the spectral standard deviation (Xi et al., 2020). The standardized Rrs was then subjected to a singular value decomposition (SVD) to derive the EOF scores (U), singular values (Λ), and EOF loadings (V).

$$X = U \Lambda V^{T} \quad (3)$$

The dataset used in step 1 corresponds to matchup satellite data obtained from a 3×3 pixel window (900×900 m) centered around the in situ sampling region and meeting the following criteria: ± 3 hours time difference, valid pixels are ≥5/9, CV at 560 nm ≤15 %, and the median of the 3×3-pixel window at each band was used to avoid outliers (Bailey and Werdell, 2006; Mograne et al., 2019). Furthermore, to minimize the spatial-temporal mismatch between the in situ water sampling and the satellite overpass in a dynamic waterbody such as the SoG, samples collected from the Fraser River plume were relocated according to the surface current direction derived from the Ocean Networks Canada CODAR system (Coastal Ocean Dynamic Application Radar) (Halverson and Pawlowicz, 2016), following Giannini et al. (2021).

Step 2: A generalized linear model (GLM) was built for each PFT, considering the log-transformed CHEMTAX-derived PFT Chla and the corresponding EOF scores subset. Following Xi et al. (2020), insignificant EOF modes were excluded, and a stepwise routine was then adapted to search for the most significant regression model through the minimization of the Akaike information criteria (AIC). AIC is an estimator of prediction error, which uses the number of scores to determine the best model among the total in a given dataset (Bracher et al., 2015). Through the AIC criteria, only significant EOF scores were used in the regression model, and, this may vary for each PFT. As a result, the regression model for the PFT Chla prediction was expressed as:

$$\ln(C_p) = a + b_1 u_1 + b_2 u_2 + \cdots . . b_n u_n \quad (4)$$

Where C_p is the Chla concentration of each PFT, a is the intercept, $b_{1,2,\ldots n}$ are the regression coefficients, and $u_{1,2,\ldots n}$ are the n number of EOF scores.

Step 3: As part of a level-1 validation, a cross-validation was performed following Bracher et al. (2015) and Xi et al. (2020). The cross-validation technique was applied to assess the robustness of the fitted regression model by splitting the whole dataset into two subsets; 80% of the data were used for model training and 20% for model validation. This procedure was run for 500 permutations, and during each permutation, EOF predicted, observed PFT Chla, and the statistical metrics for each permutation were recorded for evaluation.

The regression model performance was assessed by calculating the coefficient of determination (R^2), Mean-Absolute Error (MAE), Median Percentage Difference (MDPD), and Bias. R^2, slope (S), and intercept (a) were obtained using the log-scaled (in natural logarithm) predicted versus log-scaled (in natural logarithm) observed PFT Chla, while the MAE, MDPD, and Bias were based on the non-log transformed data. The equations of these statistical metrics were given in Xi et al. (2020), except MAE, expressed as follows:

$$\mathrm{MAE} = \sum_{i=1}^{i=M} \left|(\mathrm{Chla_{pi}} - \mathrm{Chla_{oi}})\right| \quad (5)$$

where $\mathrm{Chla_{pi}}$ is the predicted (using the regression model) PFT Chla, $\mathrm{Chla_{oi}}$ is the CHEMTAX observed PFT Chla, and M is the number of observations. Moreover, the mean value of all the statistical metrics for the cross-validation was recorded (R^2cv, MAEcv, MDPDcv) to evaluate the robustness of the fitted regression model. This step comprises the first level of evaluation of the EOF-based PFT retrievals.

Step 4: The EOF-based PFT algorithm was applied to the OLCI imagery to map the spatial distribution of PFTs. To derive PFT Chla from satellite imagery for which we do not have any CHEMTAX or HPLC pigment data, we projected the standardized OLCI Rrs data onto the EOF loading and derived a new EOF score set. This new set of EOF scores was used in Equation 4 to retrieve the spatial distribution of PFTs. The regression coefficient and intercepts were taken from Equation 4. Finally, spatial maps were produced for TChla and the following PFTs: diatoms, cryptophytes, green algae, and raphidophytes. PFT selection was based on the level-2 validation derived using a high-resolution independent CHEMTAX time-series from the Hakai Institute (called Hakai time-series hereafter) acquired between spring 2017 and summer 2020 (N=26) at station QU39, located in the northern part of the SoG (—125.099°W, 50.0307°N; **Figure 19**). The independent matchup extraction criteria performed in the level-2 validation were the same as in the model training. The CHEMTAX result used for this validation was analyzed the same way as the CHEMTAX used in the model development, following Del Bel Belluz et al. (2021). Beyond the level-1 and level-2

quantitative validation, we also conducted a qualitative evaluation of the seasonal PFT maps in comparison with the literature.

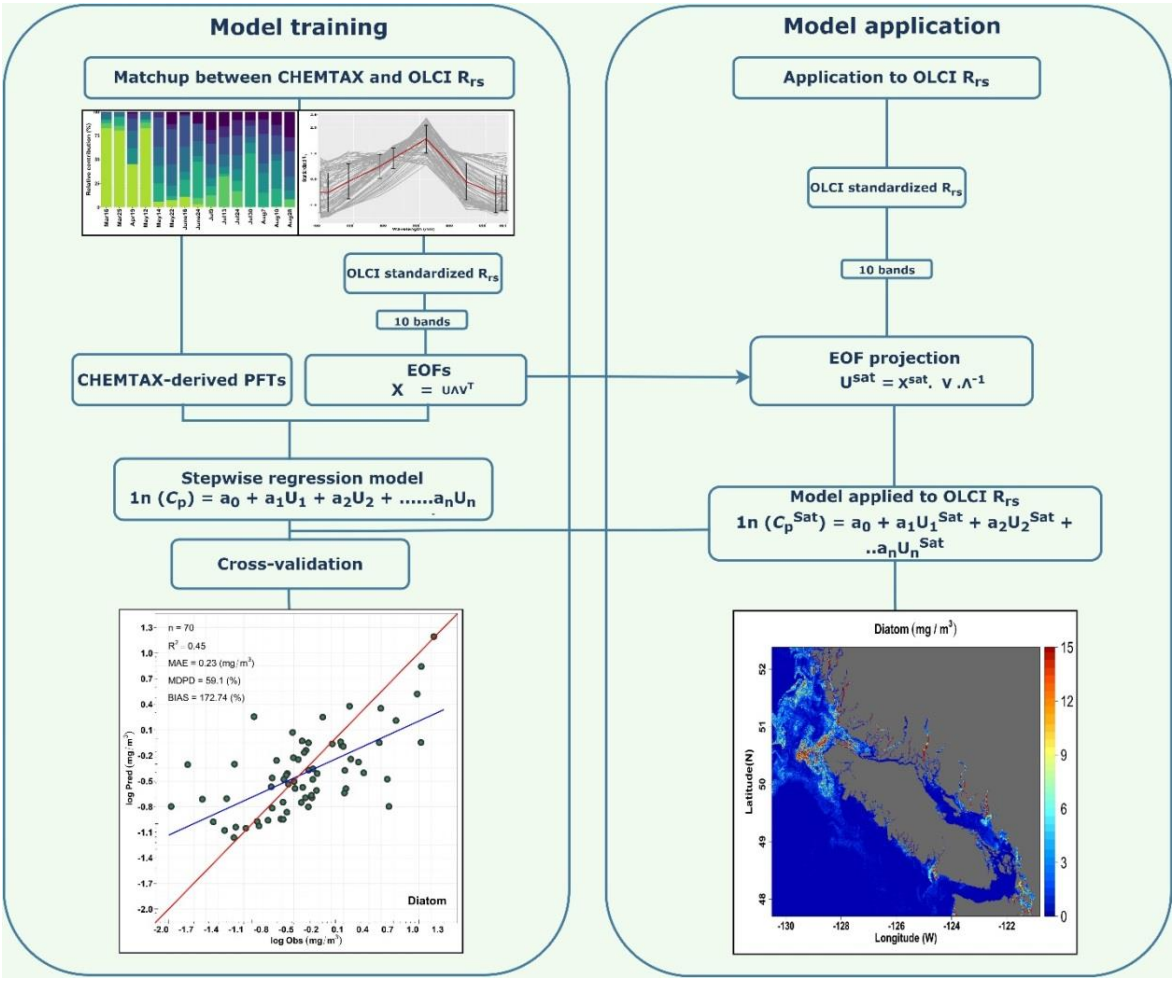

Figure 20. Schematic flowchart illustrating the EOF-based algorithm for predicting multiple Phytoplankton Functional Types (PFTs). The left side demonstrates the model training part with matchup comprised of CHEMTAX model output and Ocean Land Color Instrument (OLCI) remote sensing reflectance (Rrs; sr^{-1}). The right side indicates the model application to OLCI imagery to derive the spatial distribution of PFTs.

4.5 Results

4.5.1 Biogeochemical Characterization

Surface TChla was highly variable, showing remarkable seasonal and spatial variation (**Table 2**, **Figure 21A**), ranging from 0.39 to 28.82 mg/m^3 with an exceptionally high value of 120.11 mg/m^3 recorded at Station 4 during the summer of 2018. Outside of this value, all of the highest TChla were observed in 2018 with Stn 4 (28.82 mg/m^3) and Stn 2 (21.08 mg/m^3) showing high magnitudes during the summer and Stn 1 (18.26 mg/m^3), Stn 2 (21.56 mg/m^3), and Stn 3 (12.80 mg/m^3) during the spring. Similarly, high seasonal variability was noticeable

for TSM (**Table 2**, **Figure 21B**), especially during spring 2018 with values ranging from 1.69 to 65.63 mg/L. Persistently high TSM was observed mainly at stations close to the Fraser River; for example, Stn 4 (45.90 mg/L), Stn 2 (17.01 mg/L), and Stn 5 (13.48 mg/L) showed high TSM during spring 2018. In agreement with the TSM, CDOM 443 also demonstrated high seasonal variation (**Table 2**, **Figure 21C**) during spring 2018, ranging from 0.21 to 3.45 m^{-1}. Relatively higher (>1.14 m^{-1}) CDOM was noted at stations close to Fraser River, such as Stn 2, Stn 3, Stn 4, and Stn 5, during spring 2018 and 2019.

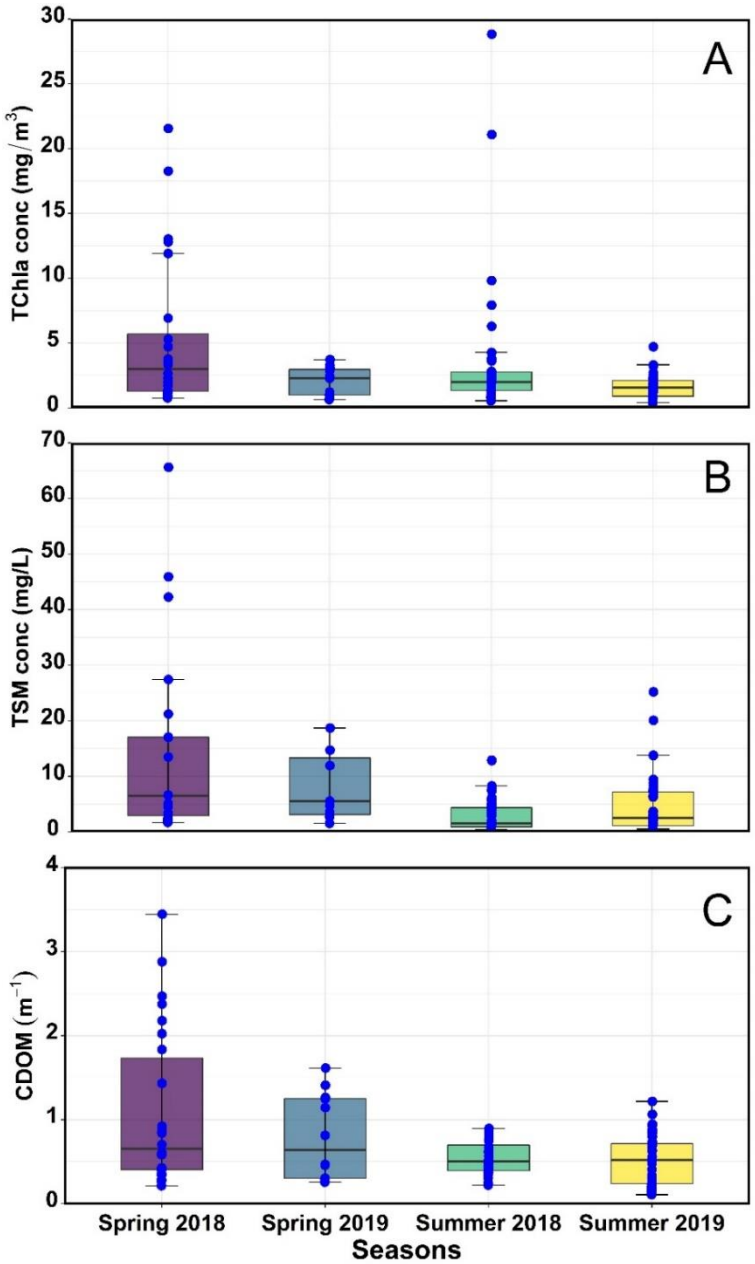

Figure 21. Seasonal variation of total chlorophyll-a concentration (TChla, mg/m^3), total suspended matter concentration (TSM, mg/L), and the absorption by colored dissolved organic matter (CDOM (m^{-1})) in the Strait of Georgia (SoG) during 2018 and 2019. The upper and lower edge of the box plot represents the first and third quartile, respectively.

Outside of the boxes, the whiskers on both quartiles indicate standard error. Blue points represent data points, and points outside the third quartile indicate extreme values.

Table 2. Seasonal surface biogeochemical variables in the Strait of Georgia (SoG) during 2018 and 2019. Minimum and maximum in square brackets, average in bold, and standard deviation in parentheses.

Years	Seasons	Variables		
		TChla (mg/m^3)	TSM (mg/L)	CDOM (m^{-1})
2018	Spring	[0.76 – 21.56] **5.25** (5.85)	[1.69 – 65.63] **13.86** (16.77)	[0.21 – 3.45] **1.07** (0.94)
2019	Spring	[0.63 – 3.72] **2.01** (1.14)	[1.56 – 18.69] **8.34** (6.18)	[0.26 – 1.61] **0.80** (0.51)
2018	Summer	[0.54 – 120.11] **6.49** (19.20)	[0.36 – 12.92] **3.12** (3.05)	[0.22 – 0.90] **0.54** (0.20)
2019	Summer	[0.39 – 4.71] **1.62** (0.91)	[0.54 – 25.20] **8.84** (5.46)	[0.11 – 1.22] **0.52** (0.30)

Surface phytoplankton community composition derived from CHEMTAX for the central SoG (Stn 1 to 5) during 2018 and 2019 is shown in **Figure 22**. Generally, during spring (March to May) 2018, diatoms dominated the community composition in almost all the stations (54 ± 32%; average±SD). At this time of the year, the relative contribution of flagellates, including haptophytes, green algae, cryptophytes, and dictyochophytes including dinoflagellates, was low (<14%) and followed by cyanobacteria and raphidophytes (< 7%). A seasonal succession from diatoms to flagellates persisted across the sampling stations during the summer (June to August) of 2018 (**Figure 22**), when the phytoplankton community composition was highly dominated by raphidophytes (24 ± 23%), followed by green algae (23 ± 10%) and cryptophytes (21 ± 16%). In 2018, high TChla concentration events dominated by raphidophytes were observed towards the western portion of the transect (Station 2) in June and concentrated near the mouth of the Fraser River (Stations 4 and 5) in July. Outside of these events, raphidophyte contributions often persisted through low TChla conditions.

Unlike the spring of 2018, in 2019, we observed the predominance of cryptophytes (37 ± 13%), followed by diatoms (28 ± 17%). A very similar trend was evident during summer, when cryptophytes (23 ± 16%) and diatoms (22 ± 25%) dominated the community composition, followed by green algae (20 ± 11%). Nevertheless, the relative contribution of other flagellates, such as haptophytes, dinoflagellates, raphidophytes, and dictyochophytes, towards the total phytoplankton concentration was minimal (< 8%). Similar to 2018,

raphidophyte contributions were small, but somewhat persistent through low TChla conditions.

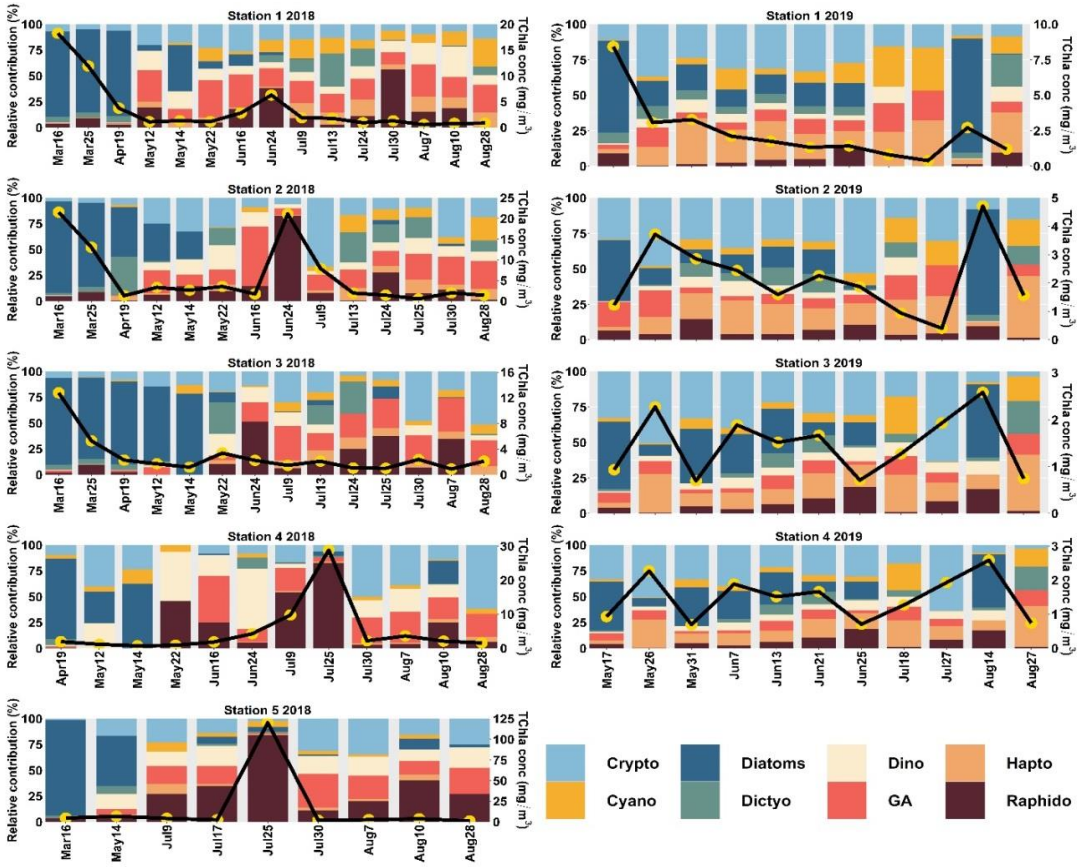

Figure 22. CHEMTAX-derived Phytoplankton Functional Types (PFTs) in the central (Station 1 to 5) Strait of Georgia (SoG) for 2018 and 2019. The primary y-axis shows the relative contribution (%), and the secondary y-axis (black line with yellow circles) indicates the total chlorophyll-a (TChla; mg/m^3) concentration.

In the northern SoG (**Figure 23**), diatoms were the predominant group at Stn 6 (66 ± 19%) and 7 (24 ± 20%). At station 6, the average contribution of all other groups was minimal (< 34%). In contrast, Stn 7 showed higher contributions from dictyochophytes (17 ± 14%), cryptophytes (17 ± 5%), haptophytes (17 ± 8%), and green algae (16 ± 2%).

After the spring bloom, a diverse phytoplankton community comprised of flagellates evolved and is summarized by the Shannon-Weaver Index (**Figure S6)**. In general, the central SoG data showed increased diversity during the summer (1.5±0.3) when diatom contributions were lowest (12±21%). Similarily, in the northern SoG, diversity was higher at Stn 7 (1.6±0.9,) which had lower diatom contributions (24±20%) (**Figure S7**).

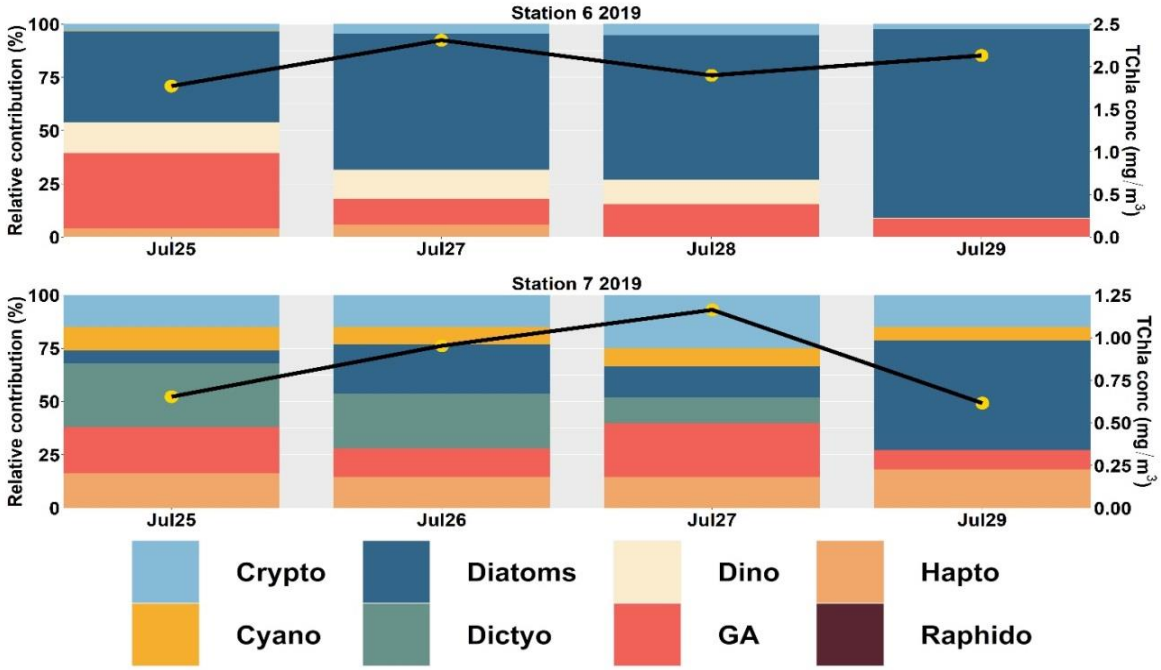

Figure 23. CHEMTAX-derived Phytoplankton Functional Types (PFTs) in the northern (Station 6 to 7) Strait of Georgia (SoG) during 2019. The primary y-axis shows the relative contribution (%), and the secondary y-axis (black line with yellow circles) indicates the total chlorophyll-a (TChla; mg/m^3) concentration.

The comparison of the CHEMTAX-derived PFT biomass output with group-level phytoplankton abundance (abundances of species within the broad CHEMTAX-derived PFT groupings were summed to derive a comparable measure) from microscopy showed significant relationships for diatoms (R^2 = 0.87, $p < 0.001$, n = 8) and dinoflagellates (R^2 = 0.52, $p < 0.05$, n = 8). This is an important result that validates the CHEMTAX analysis; however, it is important to note that the sample size for these comparisons was small. Further, pico-sized species likely constituted many of the groups showing high contributions and these would have largely been missed by microscopy and many groups showed low variability in TChla contributions (small dynamic range).

4.5.2 OLCI Derived PFTs

4.5.2.1 EOF-based Analysis of the OLCI Rrs Data

The spectral shapes and magnitude of OLCI Rrs and the corresponding standardized Rrs at 10 bands (400, 412, 443, 490, 510, 560, 620, 665, 674, 681 nm) are given in **Figure 24**. These OLCI Rrs matchup data were then decomposed using SVD to derive nine EOF modes (**Table 3**), in which the first four modes explained 99.48% of the variance. The spectral distribution of each EOF and the spectral features associated with several seawater constituents are shown

in **Figure 25**. In general, EOF-1 explained 60.33% of the total variability, and exhibited a high reflectance peak at 560 nm and a 680 nm feature likely associated with phytoplankton and TChla fluorescence, respectively. EOF-2 explained 38.33% of the total variability and exhibited a sharp decrease from 500 to 400 nm, indicating the influence of high CDOM absorption in the blue region, and a broad peak from 560 to 700 nm, indicating high backscattering from inorganic suspended sediments. Moreover, EOF-3 explained 0.50% of the variability, showing a decrease in spectral structure from 400 to 450 nm, likely associated with absorption by different accessory pigments and a spectral peak at 560 and 680 nm. The higher EOFs, such as EOF-4, EOF-5, EOF6, EOF7, EOF8, and EOF9, only showed variance < 0.32%.

The ΔAIC (**Table S8**) showed that EOF-4 was the most significant for TChla, diatoms, and dictyochophytes, and higher EOFs were selected for most PFTs (**Table S8**). At the same time, the second most important EOFs differed among PFTs (**Table S8**). For cyanobacteria and dinoflagellates < 3 EOFs were selected to build the regression model (**Table S8**).

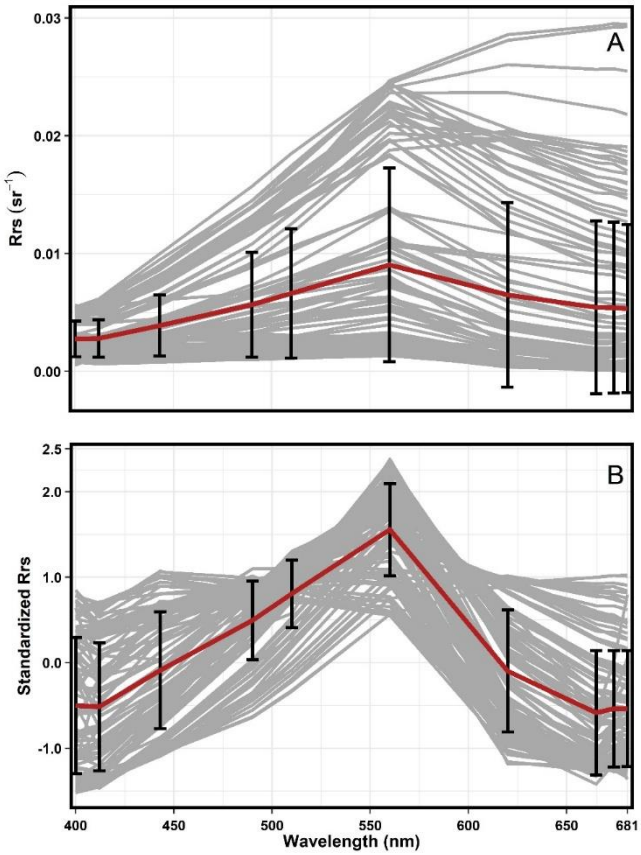

Figure 24. (A) Remote sensing reflectance (Rrs; $sr^{-1}(\lambda)$) and (B) corresponding standardized Rrs subjected to the Empirical Orthogonal Function (EOF) analysis. The grey lines show the median spectra from a 3×3 pixel window with the corresponding average spectra in red, and the standard deviation at each band is shown in black.

Table 3. Percentage of variance explained by Empirical Orthogonal Function (EOF) modes derived from the matchup between Ocean Land Color Instrument (OLCI) remote sensing reflectance (Rrs; $sr^{-1}(\lambda)$) and CHEMTAX-derived Phytoplankton Functional Types (PFTs). This percentage of variance obtained from the matchup between OLCI Rrs and total chlorophyll-a (TChla; mg/m^3).

% of variance	EOF1	EOF2	EOF3	EOF4	EOF5	EOF6	EOF7	EOF8	EOF9
OLCI Rrs	60.33	38.33	0.5	0.32	0.20	0.11	0.06	0.04	0.01

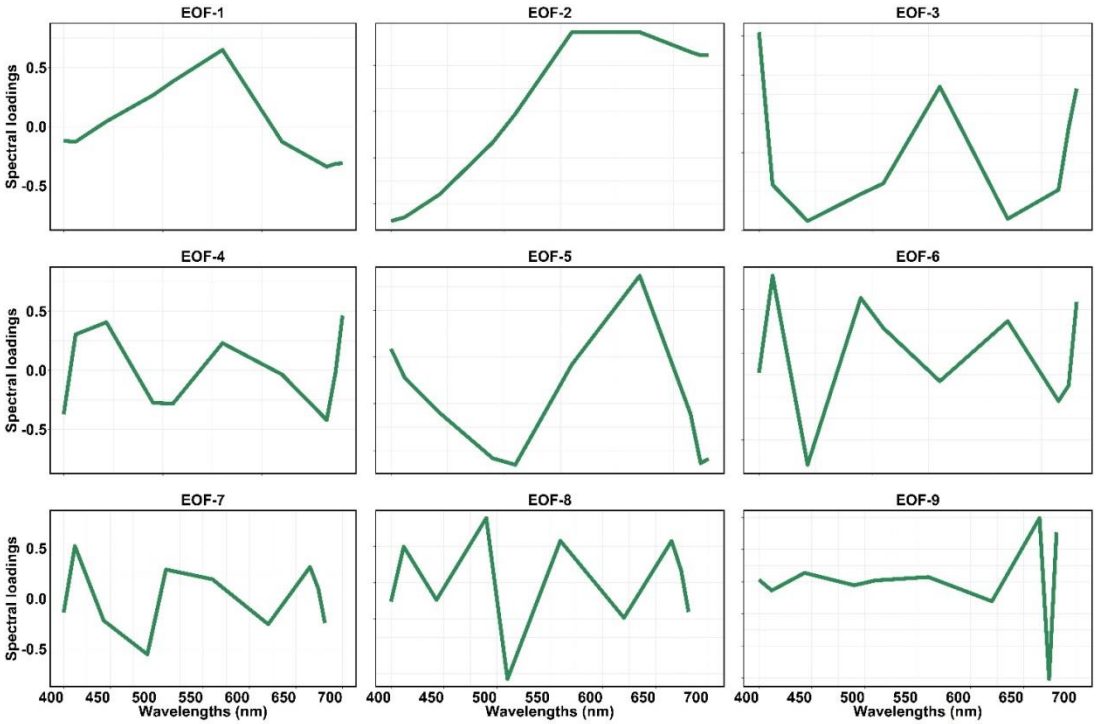

Figure 25. Spectral loadings for the nine Empirical Orthogonal Function (EOF) modes derived from the Ocean Land Color Instrument (OLCI) remote sensing reflectance (Rrs; $sr^{-1}(\lambda)$. This spectral loading was obtained from the matchup between OLCI Rrs and total chlorophyll-a (TChla; mg/m^3) concentration.

4.5.2.2 Model Evaluation

Figure 26 shows the regression between the CHEMTAX-derived PFTs *versus* TChla concentrations and the Chla concentrations of the eight PFTs retrieved from OLCI. Overall, the regression models showed moderate and highly significant retrieval (p <0.0001) for diatoms and raphidophytes (R^2 = 0.45, 0.37, respectively; **Table 4**). In addition, the regression models for dictyochophytes, haptophytes, green algae, and cryptophytes were also significant (p <0.001; **Table 4**); however, relatively lower R^2 (≤0.30) were observed. Poor performances were observed for dinoflagellates and cyanobacteria (R^2 = 0.07, 0.11, respectively; **Table 4**), with the best performance (R^2 = 0.51; p <0.0001) for TChla retrievals.

The lowest MDPDs were observed for TChla, green algae, haptophytes, and cyanobacteria (34.26%, 41.18%, 44.18%, 44.29%, respectively), whereas the highest MDPDs were noted for raphidophytes, followed by dinoflagellates (66.78%, 60.27%, respectively; **Table 4**). MAE was relatively higher for diatoms (MAE = 0.23 mg/m^3); however, it was <0.23 mg/m^3 for all other PFTs.

As part of the level-1 evaluation**,** the robustness of all PFT regression models was further assessed with the cross-validation approach (**Table 4**). The statistical parameters obtained from the cross-validation procedure were poorer than the entire dataset statistics. For example, a poor R^2_{CV} was obtained for almost all PFTs, except TChla (0.43). In general, R^2_{CV} ranged from 0 to 0.43 (**Table 4**). Whereas, $MDPD_{cv}$ ranged from 39.12 to 74.17%, in which the highest $MDPD_{cv}$ was obtained for raphidophytes (MDPD = 74.17%), followed by diatoms (65.77%). However, a low $MDPD_{cv}$ (39.12%) was observed for TChla.

The level-2 evaluation (**Figure 27**) was conducted for OLCI-derived outputs against the independent Hakai time-series from station QU39 (N = 26, **Figure 19**). Of the considered groups, green algae, cryptophytes, and diatoms followed similar trends through the Hakai time-series, resulting in significant (p <0.05) retrieval (R^2 = 0.33, 0.23, 0.20, MDPD = 163, 46.86, 88.59%, respectively) (**Figure 27**; **Table 5**). The best retrieval was noted for TChla, which highly corresponded with the Hakai time-series (R^2 = 0.68, MDPD = 52.25%) (**Figure 27**; **Table 5**). Except for TChla, diatoms recorded the highest MAE (0.58 mg/m^3; **Table 5**), and MAE for all other groups was < 0.58 mg/m^3. The MDPD was the highest for green algae (163%) and raphidophytes (130%), whereas MDPD was relatively lower for cryptophytes (46.86%). Overall, OLCI-derived PFTs, including TChla, over-estimated the lower ranges and under-estimated higher ranges of in situ values. Raphidophytes showed poor correspondence (R^2 = 0.08, MDPD = 130%) with the Hakai time-series (**Figure 27**; **Table 5**); however, it is important to note that raphidophytes were not observed via microscopy on any of the considered dates, highlighting misclassifications by both CHEMTAX and OLCI-derived outputs under low TChla conditions.

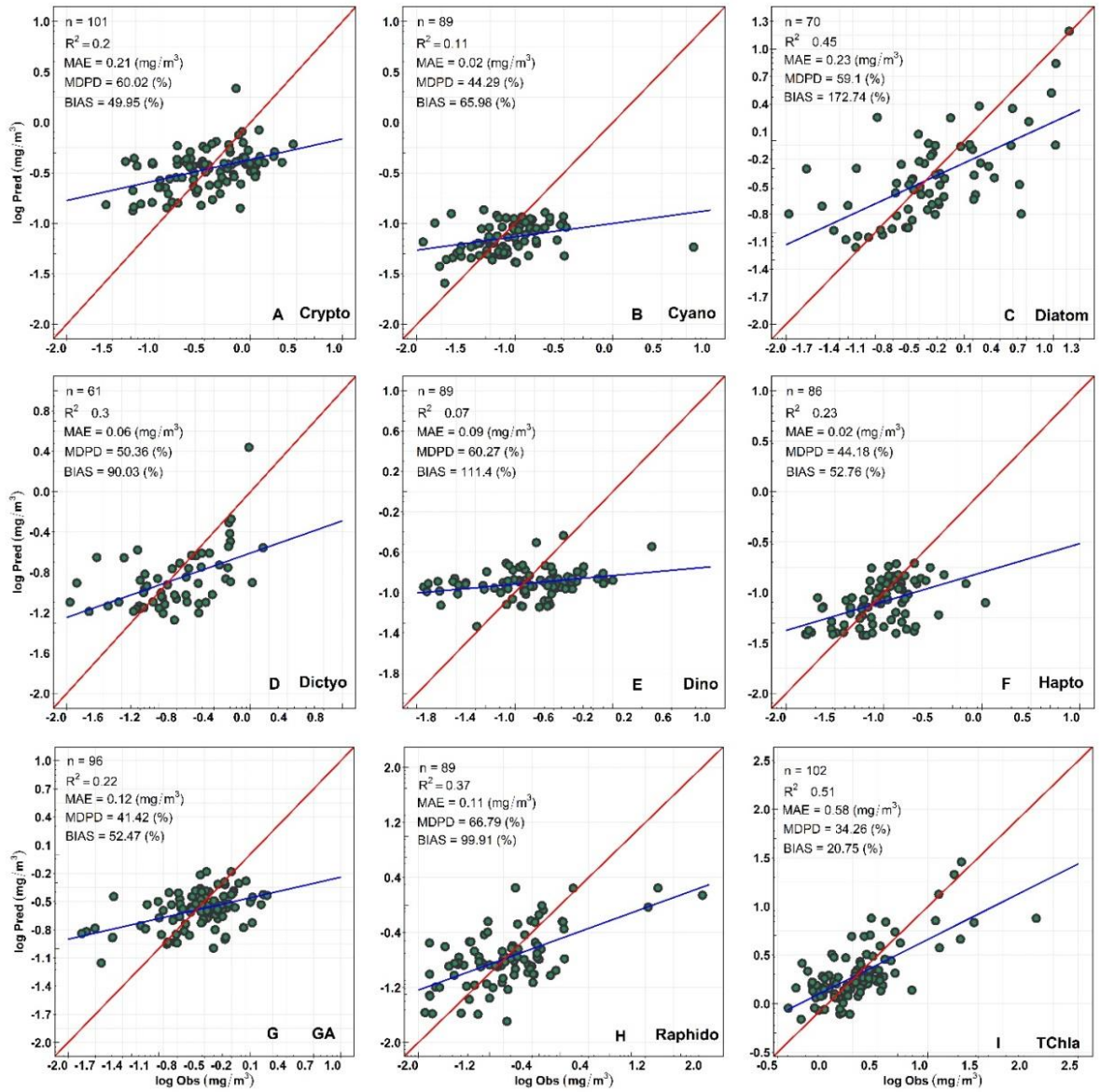

Figure 26. Regression between CHEMTAX-derived (x-axis) *versus* OLCI-retrieved (y-axis) Chla concentration of (A) Cryptophytes, (B) Cyanobacteria, (C) Diatoms, (D) Dictyochophytes, (E) Dinoflagellates, (F) Haptophytes, (G) Green Algae, (H) Raphidophytes, and (I) total chlorophyll-a (TChla).

Table 4. Full-fit statistics and Cross-validation statistics derived from the EOF-based algorithm, developed using the matchup between Ocean Land Color Instrument (OLCI) remote sensing reflectance (Rrs; $sr^{-1}(\lambda)$) and CHEMTAX-derived Phytoplankton Functional Types (PFTs). R^2, Slope, and Intercept were calculated on a log scale (in natural logarithm), whereas mean absolute error (MAE; mg/m^3) and median percentage difference (MDPD; %) were calculated on a non-log scale.

	N	R^2	MAE	MDPD	Slope	Intercept	R^2cv	MAEcv	MDPDcv	*p*-value
TChla	102	0.51	0.58	34.26	0.51	0.34	0.43	2.76	39.12	2.2×10^{-16}
Diatoms	70	0.45	0.23	59.09	0.46	-0.49	0.30	1.34	65.77	1.3×10^{-10}
Dictyo	61	0.30	0.06	50.36	0.30	-1.43	0.11	0.20	61.34	4.6×10^{-6}
Raphido	89	0.37	0.11	66.78	0.37	-1.13	0.25	1.90	74.17	2.2×10^{-10}
Hapto	86	0.23	0.02	44.18	0.23	-1.96	0.14	0.07	47.13	2.9×10^{-6}
GA	96	0.22	0.12	41.41	0.22	-1.06	0.08	0.22	49.14	1.5×10^{-6}
Crypto	101	0.20	0.21	60.02	0.20	-0.85	0.10	0.46	63.54	2.2×10^{-6}
Dino	89	0.07	0.09	60.27	0.07	-1.93	0.00	0.17	63.25	0.01
Cyano	89	0.11	0.02	44.29	0.12	-2.34	0.08	0.13	48.51	0.001

Table 5. Validation statistics were obtained from the regression between OLCI-derived Phytoplankton Functional Types (PFTs), including total chlorophyll-a (TChla; mg/m^3) *versus* Hakai time-series data. R^2, slope, and intercept were calculated on a log scale (in natural logarithm), whereas mean absolute error (MAE; mg/m^3) and median percentage difference (MDPD; %) were calculated on a non-log scale.

	N	R^2	MAE	MDPD	Slope	Intercept	*p*-value
TChla	26	0.68	0.73	52.25	0.65	0.35	2.52×10^{-7}
Diatoms	22	0.20	0.58	88.59	0.22	-0.63	0.03
Crypto	26	0.23	0.21	46.86	0.23	-1.18	0.01
GA	25	0.33	0.12	163	0.16	-1.07	0.002
Raphido	19	0.08	0.10	130	0.19	-1.53	0.23

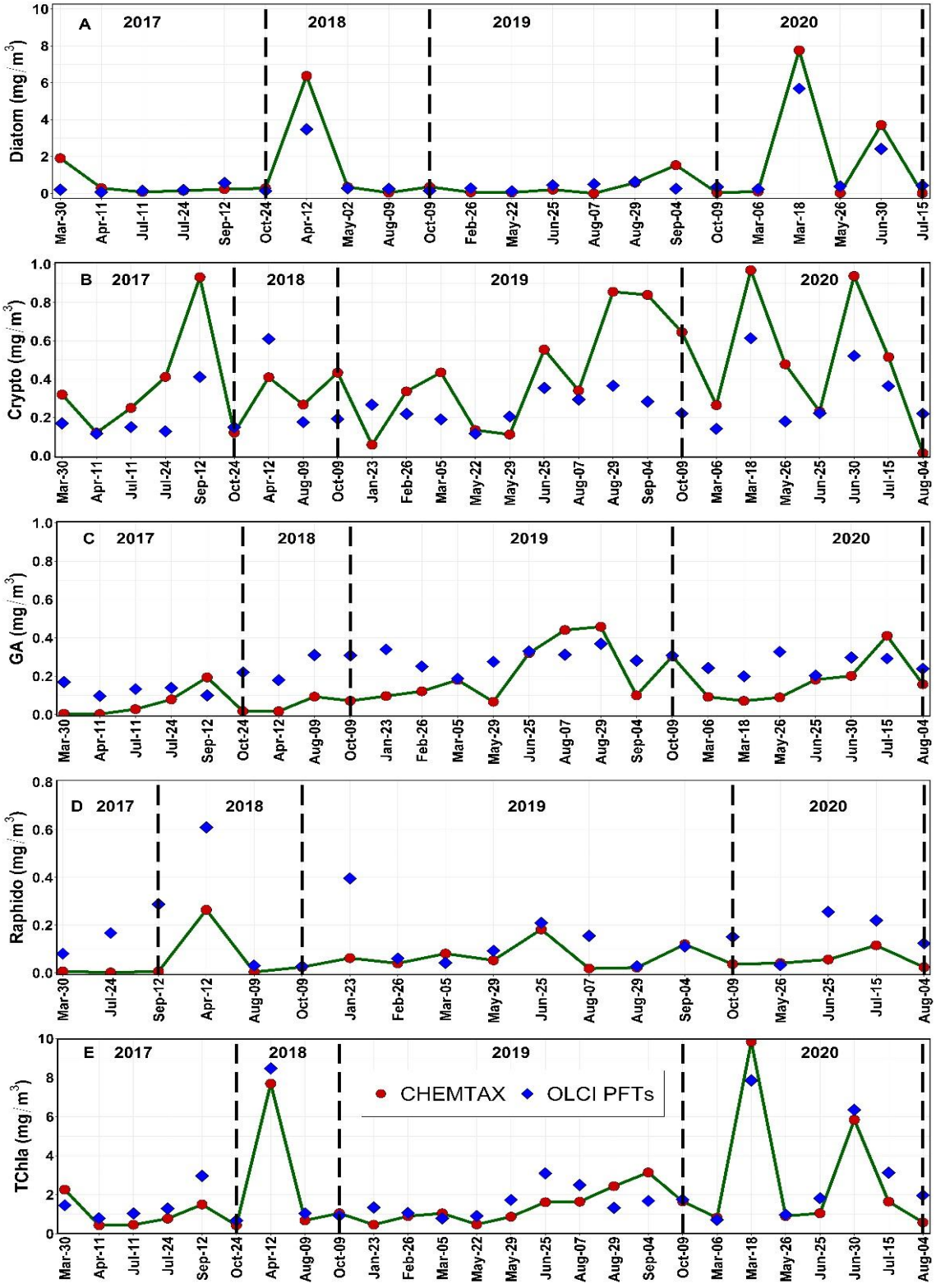

Figure 27. Comparison of the performance of the OLCI-derived (A) Diatoms, (B) Cryptophytes, (C) Green Algae, (D) Raphidophytes, and (E) total chlorophyll-a (TChla) concentration obtained from the daily cloud-free images between spring 2017 to summer 2020 (blue diamonds) against the Hakai time-series data at station QU39 located in the northern Strait of Georgia (SoG) (red circles). The location of station QU39 is given in **Figure 19**. The vertical lines on each panel separate different years, and note that the

matchup data points differ among Phytoplankton Functional Types (PFTs), mainly because all PFT may not be present at all stations. For example, Raphidophytes showed fewer points (19) out of the 26 matchup points, followed by Diatoms (22). Hence, the dates on the x-axis might not be similar for all PFTs.

4.5.3 PFT Mapping in SoG From Daily OLCI Imagery

After the model evaluation, we applied the EOF-based regression algorithm to daily OLCI imagery acquired during cloud-free days between March and September 2018 to provide a qualitative evaluation based on the spatial distribution and seasonal dynamics of PFTs over the SoG. The spatial maps (**Figure 28**) were produced for diatoms, cryptophytes, green algae, and raphidophytes since they accounted for 74% of the total phytoplankton concentration. Overall, the PFT spatial maps showed a distinct spring and fall bloom in the SoG, which was mostly comparable with the Hakai time-series. During spring, diatoms showed the highest average concentration (avg 5.15 mg/m^3) with high concentrations observed near the northern and central SoG, and the Juan de Fuca Strait (**Figures 28-A1** to **28-A2**), similar to the Hakai time-series data (diatom Chla ranging from 0.23 to 6.00 mg/m^3) in the northern SoG. In comparison, the spatial distribution of cryptophytes and raphidophytes for the spring followed a similar spatial distribution to diatoms, but with relatively lower concentrations (avg 0.70, 1.30 mg/m^3, respectively; **Figures 28-B1** to **28-B2, Figures 28-D1** to **28-D2**). Furthermore, the spatial distribution of green algae for the spring showed a comparatively low average concentration (0.32 mg/m^3) and tended to show inverse spatial trends to diatoms. Hakai's data for this week of the year also showed comparable concentrations for cryptophytes, green algae, and raphidophytes in the northern SoG.

Following the spring period, the summer average diatom concentration decreased (4.37 mg/m^3), but strong diatom bloom events were still observed off of the southwestern coast of Vancouver Island, with patches of high contributions in the central and northern SoG near the beginning of summer (June 16th – **Figure 28-A3**). By July 25th, diatom contributions had decreased considerably (**Figure 28-A4**). Consistent with these observations, Hakai's data for a similar period showed low diatom concentrations in the northern SoG, ranging from 0.12 to 0.14 mg/m^3. Furthermore, comparable with the Hakai time-series, OLCI-derived green algae showed a relatively higher average concentration (0.37 mg/m^3; **Figure 28-C3**) than spring with localized areas of high cryptophyte concentrations in the central SoG and Juan de Fuca Strait (**Figure 28-B3**). Interestingly, raphidophytes recorded the highest average concentration during the summer (2.24 mg/m^3), and blooms were observed southwest coast

of Vancouver Island and close to the Fraser River mouth (**Figure 28-D3**, **28-D4**). Of note, on June 16th, both diatoms and raphidophytes showed high concentrations off of the southwest coast of Vancouver Island; however, concentrations between the two groups was decoupled within the central and northern SoG, where diatom concentrations were high, and raphidophyte contributions were low. Finally, in fall, diatoms still constituted the highest average concentration (1.35 mg/m^3), especially near the central and northern SoG and off the south of Vancouver Island (**Figure 28-A6**). During this season and, similar to the Hakai time-series, the spatial distribution of cryptophytes was similar to the distribution of diatoms but with a relatively low average concentration (0.49 mg/m^3). In turn, green algae showed relatively high concentrations in the central and northern SoG and off of southern Vancouver Island (**Figure 28-C6**) resulting in the highest average seasonal concentration (0.41 mg/m^3) for this PFT. Overall, TChla showed high concentrations throughout the season with high concentration regions localized in the central and northern SoG, near the mouth of the Fraser River, within the Juan de Fuca Strait, and off the south coast of Vancouver Island (**Figure 28-E1**, **28-E2**, **28-E3**, **28-E4**, **28-E5**, **28-E6**).

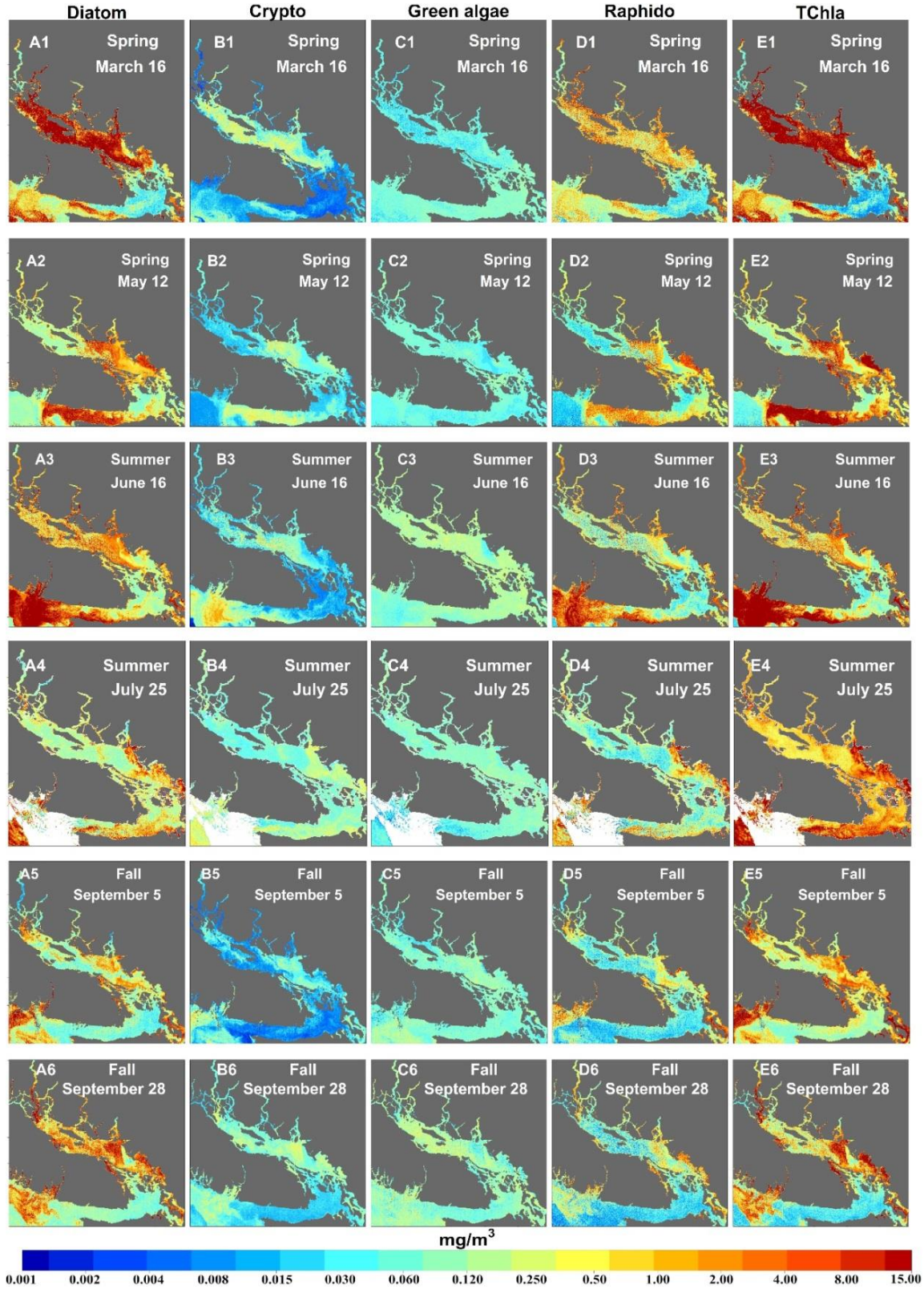

Figure 28. Spatial and seasonal distribution of Diatoms, Cryptophytes, Green Algae, Raphidophytes, and total chlorophyll-a (TChla), retrieved from Ocean Land Color Instrument (OLCI) over the surface waters of the Strait of Georgia (SoG). Only images with >80% coverage were selected for the final map. OLCI imagery was acquired on cloud-free days from spring to fall of 2018, and the specific days were given in the spatial map. White regions in the maps show the clouds.

4.6 Discussion

4.6.1 EOF-based PFT Retrievals and Source of Uncertainties

EOF analysis has successfully been used to derive phytoplankton cell abundance and the concentration of accessory pigments and multiple PFTs from the global open ocean and regional waters (e.g., Craig et al., 2012; Bracher et al., 2015; Lange et al., 2020; Xi et al., 2020); however, few of these studies have addressed coastal Case-2 waters (e.g., Craig et al. 2012). In these waters, the relationship between EOFs to specific PFTs can be hindered by high concentrations of water constituents such as CDOM, detritus, and TSM, which can largely define the reflectance spectra and mask PFT signals (Mouw and Yoder, 2010; Craig et al., 2012; Mouw et al., 2017). In this study, high variability of these constituents were observed (**Table 2**), thus significantly contributing to the reflectance signal and EOF modes (**Figure 25**), as observed in Phillips and Costa (2017) and Craig et al. (2012). Furthermore, EOF modes from oceanic waters (Bracher et al., 2015; Xi et al., 2020) show high dissimilarity to those found in this study, likely due to the absence of high concentrations of non-phytoplankton constituents. As a result, the derivation of PFTs in this study was likely hindered by the high bio-optical complexity characteristic of the SoG (Loos and Costa, 2010; Loos et al., 2017; Phillips and Costa, 2017; Gianinni et al., 2021). Our study showed high seasonal and spatial variability in CDOM 443 and TSM (see **Table 2**) and this would have significantly affected the satellite retrieval for both high-biomass bloom forming groups as well as the low biomass flagellates.

Beyond the inherent role of the Case-2 waters' different optical constituents on the EOF retrievals, the biomass contribution of a specific PFT to the total phytoplankton biomass is known to be relevant for accurate PFT retrievals from spectral data. Of the two levels of validation, the first validation showed good retrievals for diatoms and raphidophytes, and the second validation showed that green algae, cryptophytes, and diatoms followed the seasonal trends of their corresponding in situ values from a high temporal resolution Hakai time-series (**Figure 27**). This is likely associated with their larger contribution (62%) to the total concentration, as demonstrated in the CHEMTAX model output (**Figure 22**). These results are encouraging with regard to the remote sensing of PFTs over optically complex Case-2 waters, where previous studies have documented that PFT retrieval over Case-2 waters is challenging (Craig et al., 2012; IOCCG, 2014; Mouw et al., 2017). At the same time, the poorest retrievals obtained for haptophytes, dictyochophytes, cyanobacteria, and

dinoflagellates was likely related to their low concentration (26%), resulting in low contributions to the Rrs signal. Similarly, Xi et al. (2020) also observed poor performance for low biomass prokaryotes (R^2cv = 0.11; MDPDcv = 55.08%) and *Prochlorococcus* (R^2cv = 0.18; MDPDcv = 42.68%), and Bracher et al. (2015) showed slightly poorer retrievals (R^2cv = 0.28; MDPD = 31%) for the pigment zeaxanthin (found in low concentrations), a diagnostic pigment for cyanobacteria.

Other factors, such as the quantity and quality of in situ and satellite matchup data also impact the performance of satellite-derived PFTs. For the EOF-based PFT algorithm, the number of matchup data points and their representation of the range of phytoplankton groups in the region are crucial to enhance the accuracy of the retrievals (Bracher et al., 2015; Xi et al., 2020). Bracher et al. (2015) suggested that at least 50 valid matchup data points are necessary for statistically significant pigment estimation, specifically for open ocean waters. Xi et al. (2020) employed a larger number of matchup data points (52-394) for global open ocean waters, with an even higher number (483) in a more recent publication (Xi et al., 2021). Lange et al. (2020) employed 73-78 matchup points to develop an EOF-based algorithm to retrieve phytoplankton cell abundance from MODIS across the Atlantic Ocean. However, Craig et al. (2012) showed that a reliable EOF-hyperspectral-based model could be built with 15 valid data points, specifically for one season and at a local scale. In our study, the total number of matchup data points was between 61 and 102, which is within the range suggested by the above authors. However, given the optical complexity of our waters, a higher number of matchups would likely improve the PFT retrievals.

The spatial and temporal variability of the matchup samples, which has implications on data quality, is also an important factor in evaluating the EOF-based PFT retrievals, especially in Case-2 dynamic coastal waters such as the SoG. This is a common problem in validating satellite-derived products of coastal waters (e.g., Mahadevan and Campbell, 2002; Moses et al., 2016; Barnes et al., 2019; Tilstone et al., 2022). Specifically in the SoG, Nasiha et al. (2022) have shown large spatial variability of in situ above-water reflectance within the 300m spatial resolution of OLCI. Notably, most of this variability was observed at the frontal region of the Fraser River and within transitional oceanic waters and likely has considerable implications on the quality of in situ data for satellite validation purposes. Further, Giannini et al. (2021) have shown the importance of the temporal mismatch between in situ and OLCI data in the SoG due to the dynamic nature of this water system. Here, an attempt was made to

reduce these spatial issues by rejecting matchups corresponding to transitional waters, and further, to minimize temporal mismatch by relocating the sample's position using surface current maps generated using a CODAR system (Halverson and Pawlowicz, 2016). This method was successfully adopted by Giannini et al. (2021). Out of the 102 matchup samples utilized in this study, 20 were successfully relocated using the surface current map. Indeed, after the relocation, we observed improved performance in the Rrs matchup analysis (not shown). However, it is still challenging to detect and quantify any remaining source of uncertainties associated with this spatial and temporal water mass mismatch.

Uncertainties associated with CHEMTAX model outputs should also be considered in evaluating the EOF-based PFT retrievals. For example, CHEMTAX outputs are more prone to errors in environments with complex phytoplankton community composition, notably when local ratios are unknown, and many species have similar pigment composition (Lewitus et al., 2005); this resembles some of the conditions in the Salish Sea. For instance, the haptophyte, *Phaeocystis pouchetti*, has been shown to periodically lack 19′-hexanoyloxyfucoxanthin (HF) and have relatively high 19′-butanoyloxyfucoxanthin (BF) concentrations; in these conditions, the utilized CHEMTAX ratios may result in misclassification of the haptophyte as a dictyochophyte (Del Bel Belluz et al., 2021). In addition, raphidophytes, such as *Heterosigma akashiwo*, often bloom regionally and have similar pigment ratios to diatoms, with the separation of raphidophytes dependent on their high violaxanthin to TChla ratios. However, violaxanthin concentrations are variable and broadly present among groups (Lewitus et al., 2005; Laza-Martinez et al., 2007; Higgins et al., 2011). In this study, violaxanthin to TChla ratios were noticeably high during raphidophyte blooms, and CHEMTAX appeared to be able to distinguish these events. Yet, under lower TChla concentration conditions, raphidophyte contributions often persisted, likely representing misclassifications and underestimations of other groups. These misclassifications are evidenced by the small CHEMTAX and satellite-based raphidophyte contributions shown over the Hakai time-series comparison (**Figure 27**); no raphidophytes were observed via microscopy during much of this time-series (Del Bel Belluz et al., 2021). As such, errors from CHEMTAX may have been perpetuated to the satellite models.

To address this issue, novel "data driven" approaches have been developed to help account for PFTs present in low contributions and also, to deal with issues in CHEMTAX analysis such as misclassifications of groups with similar pigment profiles and pigment collinearity

(Catlett and Siegel, 2018; Kramer and Siegel, 2019; Kramer et al., 2020). These approaches, which are not hindered by co-correlation between pigments, use both cluster and EOF analysis to group species with similar pigment ratios. Outputs from these methods typically have fewer PFT groups (< 5) that combine communities that covary in the environment. Using the same clustering methods, the pigment data in this study clustered into 4 pigment groups: diatoms, dinoflagellates, haptophytes/pelagophytes, and green algae/cryptophytes/cyanobacteria (**Figure S8**). The use of this more generalized approach in our study may have allowed for a more accurate retrieval of a broad "small flagellate" group during high diversity low TChla summer conditions, similar to those used effectively by regional particulate fatty acid studies (Costalago et al., 2020; McLaskey et al., 2022). However, the results of this clustering approach grouped violaxanthin with green algae pigments suggesting that this group drove its variability, which is generally the case under non-raphidophyte bloom conditions. As a result, this approach would not have allowed for the separation of raphidophytes, a toxic species with considerable regional impacts on fish (Esenkulova et al., 2021), but likely helps explain the misclassifications described above.

In order to fully assess the above error and decide on the best approaches, independent validation of in situ pigment-based PFT results are needed: this is a major shortcoming of many pigment-based studies (e.g., Chase et al., 2020; Kramer et al., 2020). In our study, regional CHEMTAX results corresponded well with expected trends in phytoplankton communities (discussed below), qualitative microscopy comparisons (Del Bel Belluz et al., 2021), and particulate matter fatty acids and size-fractionated Chla (SF-Chla) (Costalago et al., 2020; McLaskey et al., 2022); however, no direct cross comparisons have yet been performed. This lack of comparison is partially due to the difficulty in comparing pigment-based approaches with other measures of phytoplankton composition (Higgins et al., 2011). In particular, microscopy is skewed to larger species and cannot resolve pico-sized species known to often dominate summer conditions in the SoG (Del Bel Belluz et al., 2021). This trend is evidenced by the small and limited CHEMTAX validation dataset used here, which showed correspondence with only more easily preserved diatoms. Additional sources of validation, using flow cytometry, size-fractionated filtration (SFF), conventional imaging flow cytometry (Olson and Sosik, 2007), Imaging FlowCytobot, genomics, and next-generation sequencing (see IOCCG, 2014 and Lombard et al., 2019 for a detailed review) are needed to evaluate pigment-based community methods and satellite outputs.

Finally, the quality of atmospheric correction and its effect on the OLCI-measured signal is a consideration in evaluating the EOF-based PFT retrievals. For oligotrophic oceanic gyre waters, Rrs uncertainties in blue wavelengths are above 5% (Hu et al., 2013), and are even higher in optically complex waters (Moore et al., 2015). A recent study by Giannini et al. (2021) assessed the radiometric qualities of OLCI Rrs in northeast Pacific coastal waters, showing the highest errors associated with the blue bands (400, 412, 443 nm) when compared with autonomous above-water in situ reflectance measurements (Wang and Costa, 2022). A significant Rrs underestimate in the blue and red bands of OLCI was also reported in the European coastal and inland waters (Zibordi et al., 2018; Bi et al., 2018; Li et al., 2019; Alikas et al., 2020; Shen et al., 2020; Vanhellemont and Ruddick, 2021). Although POLYMER has been proven to outperform other methods to retrieve OLCI Rrs for the coastal waters of BC, Giannini et al. (2021) argue that uncertainties exist and are likely associated with the optimization of the model, which considers TChla and a coefficient (*f*b) that scales the particulate backscattering in the water column, but does not consider the contribution of CDOM (Steinmetz et al., 2016). As previously discussed, CDOM is a significant optical constituent affecting the light spectra in the SoG (Loos and Costa, 2010; Phillips and Costa, 2017, **Table 2**), with CDOM in our study ranging from 0.11 to 3.45 m^{-1} (**Table 2**). Since the launch of the Sentinel 3A, the uncertainties found between the products derived from OLCI using different processing approaches, especially for coastal waters, have been under critical evaluation and are still an active area of research (e.g., Zibordi et al., 2018; Gossn et al., 2019; Hieronymi, 2019; Mograne et al., 2019; Giannini et al., 2021; Tilstone et al., 2021; Vanhellemont and Ruddick, 2021).

4.6.2 Dynamics of PFTs in the SoG

The OLCI-derived PFT maps showed strong spring diatom blooms with the highest concentrations observed in the central and northern SoG (**Figure 28-A1**). These findings are in-line with TChla concentrations and CHEMTAX, which showed that diatoms dominated the phytoplankton community in the spring of 2018 (**Figure 22**). For the same period, in situ observations (microscopy, CHEMTAX, particulate matter fatty acids, and SF-Chla) collected across the SoG showed a similar spring bloom timing with the community dominated by the centric diatoms *Thalassiosira*, *Skeletonema*, and *Chaetoceros* (Nemcek et al., 2018; Del Bel Belluz et al., 2021; Esenkulova et al., 2021; McLaskey et al., 2022). Additionally, the SalishSeaCast model accurately predicted the timing of seasonal spring bloom in 2018, which occurred between March 13th to March 23rd, consistent with our findings (Allen et al., 2018).

In the SoG, diatom-dominated spring blooms are well documented in the literature (e.g., Harrison et al., 1983; Allen and Wolfe, 2013; Suchy et al., 2019; Del Bel Belluz et al., 2021). As in many temperate systems, diatoms are the primary group that quickly respond to increased nutrient and light conditions resulting in rapid accumulation and bloom conditions (Allen and Wolfe, 2013). Both satellite-derived time-series data (Suchy et al., 2019) and coupled-biophysical model results (Collins et al., 2009) have shown that spring bloom timing and strength are related to early water column stratification due to positive anomalies in SST, Fraser River discharge, and reduced wind stress. These conditions promote the relatively high average surface TChla concentration in the SoG, which is one of the most productive regions on the west coast of North America (Jackson et al., 2015).

During summer, EOF-based PFT retrievals generally showed reduced TChla concentrations and diatom contributions, with exception to localized events off of the southwest coast of Vancouver Island and in the central and southern SoG. Furthermore, the northern SoG typically showed reduced TChla and diatom contributions compared to the central and southern regions **(Figure 28-A4** to **28-E4**). This succession to lower concentration summer conditions is expected for the region. For instance, a monthly climatology of satellite-derived surface TChla over 13 years (2003 to 2016) showed relatively lower TChla during summer and reduced concentrations focused in the northern SoG (Suchy et al., 2019). Spatially, the northern SoG is more stable and stratified and prone to limiting nitrate concentrations, whereas the southern SoG is exposed to higher vertical mixing bringing nutrient-rich water from the subsurface to the sunlit layer (Peña et al., 2016). In contrast, higher phytoplankton concentrations found to the south of Vancouver Island have been linked to wind-induced seasonal upwelling, freshwater discharge from Fraser and Columbia Rivers, and the outflow of nutrient-rich low salinity water from the Juan de Fuca Strait (Masson and Cummins, 2000; Hickey and Banas, 2008; Thomson et al., 2014). Indeed, the brackish water outflow from the Juan de Fuca Strait is a significant source of nitrate to the continental shelf of Vancouver Island and northern California Current (Foreman et al., 2008; Hickey and Banas, 2008). These nutrient-rich waters supply adequate nitrate concentration to the mixed layer depth promoting increased phytoplankton biomass and primary productivity to the southwest of Vancouver Island during summer (MacFadyen and Hickey, 2010).

The summer succession to flagellate-dominated communities in the SoG was apparent in the in situ CHEMTAX model output and less so in the OLCI-derived spatial maps. In agreement with our observations, many authors using a variety of approaches have reported a similar

seasonal succession from diatoms to small flagellates and noted that the northern SoG has a more diverse flagellate dominated community when compared to the central SoG (e.g., Harrison et al., 1983; Haigh and Taylor, 1991; Del Bel Belluz et al.,2021; McLaskey et al., 2022). The Shannon-Weaver Index calculated from our CHEMTAX model output supports the literature showing higher diversity in the summer when small flagellates were dominant (such as green algae and cryptophytes, followed by dinoflagellates, haptophytes, and dictyochophytes), and relatively higher diversity in the northern SoG than the central SoG (**Figure S7**).

The strong raphidophyte blooms observed in summer 2018 were an exception to the seasonal succession to a diverse range of small flagellates. These blooms were observed by multiple other studies in the region (Nemcek et al., 2018; Pawlowicz et al., 2020). In particular, Esenkulova et al. (2021) found June and July 2018 to be an exceptional year in terms of *Heterosigma akashiwo*, with abundance reaching 11,000 cells mL^{-1}, compared to data from 2015-2017 (150 cell mL^{-1}). Similar to the results of this study, these authors showed that in the SoG, *Heterosigma akashiwo* are typically observed from May to September, with the summer 2018 blooms having a distinct spatial distribution concentrated near the Fraser River mouth (**Figure 28-D4**). The location of these blooms near the Fraser River suggests that *Heterosigma akashiwo* is strongly associated with highly stratified low salinity waters with low nitrate and phosphate concentrations and high silicate concentrations (Esenkulova et al., 2021). In addition, and similar to the results of this study, *Heterosigma akashiwo* was observed off of the southwest coast of Vancouver Island (**Figure 28-D3**) on June 16th, 2018 by Gower and King (2018) using the Maximum Chlorophyll Index (MCI) method on satellite imagery and these results were corroborated with microscopy. In the SoG, *Heterosigma akashiwo* are thought to have negative impacts on wild Pacific salmon populations and significantly impact the fin-fish and salmon aquaculture industries driving large fish kill events (Rensel et al., 2010; Haigh and Esenkulova, 2014). As a result, the ability to monitor these species using EOF-based PFT maps would provide a mechanism for better understanding and mitigating the negative effects of these harmful species.

Lastly, a fall bloom dominated by diatoms is evident in our spatial maps, which is typical for this region (Nemcek et al., 2018; Del Bel Belluz et al., 2021). This bloom was largely constituted by *Pseudo-nitzschia* and was observed spanning much of the northern SoG (Nemcek et al., 2018; Esenkulova et al., 2021; Del Bel Belluz et al., 2021). On a broader temporal scale, Jackson et al. (2015), Peña and Nemcek (2018), and Suchy et al. (2019) have

reported fall blooms on the Vancouver Island coast. Generally, the primary driver for the 2018 fall bloom in the SoG was the reintroduction of nutrients into the sunlit layer following high summer stratification and grazing (Del Bel Belluz et al., 2021). In turn, autumn blooms to the south of Vancouver Island have been attributed to the seasonal wind-induced upwelling followed by freshwater discharge from the Columbia River or Fraser River and the nutrient-rich water outflow from the Juan de Fuca Strait (Jackson et al., 2015).

4.7 Conclusion and Outlook

This is the first study to characterize the spatial-temporal dynamics of the major PFTs in the SoG, an optically complex Case-2 system on the west coast of Canada, with satellite data using an EOF-based algorithm combined with CHEMTAX model outputs. As a result of the complexity of these waters, many of the low concentration groups such as haptophytes, dictyochophytes, dinoflagellates, and cyanobacteria showed poor performance. However, the first level of validation demonstrated reliable performance for diatoms and raphidophytes, and the second level of validation with an independent Hakai time-series showed the best retrieval for TChla and green algae, cryptophytes, and diatoms showed similar seasonal trends as their corresponding in situ values. In this comparison, raphidophyte outputs by both CHEMTAX and satellite retrievals suggested misclassifications during the observed non-raphidophyte bloom conditions over the Hakai time-series when no raphidophytes were observed via microscopy.

Qualitative validations of the spatial-temporal 2018 EOF-based PFT generally agreed with expected patterns for this region with strong spring bloom highly dominated by diatoms, and primarily located in the central and northern SoG and Juan de Fuca Strait. During summer, diatoms showed high concentration off of the southwest coast of Vancouver Island and in patchy regions within the central SoG. As expected based on regional studies, summer TChla and diatom contributions decreased with the lowest contributions observed towards the northern SoG. In turn, patches of increased cryptophyte contributions were noticeable in the central SoG and Juan de Fuca Strait from spring through summer. Green algae, which generally shows its highest concentrations in the summer through autumn, showed low concentrations during spring throughout the central and northern SoG. Furthermore, similar to the CHEMTAX model output and literature, the EOF-based PFT maps revealed 2018 summer raphidophyte blooms (*Heterosigma akashiwo*) off of the southwest coast of Vancouver Island, within the Juan de Fuca Strait and close to the Fraser River mouth. Finally,

a fall bloom was evident in our spatial maps in which diatoms displayed high concentration near the central and northern SoG and south of Vancouver Island. Overall, diatoms demonstrated high average concentrations throughout the study period, and relatively low average concentrations were observed from green algae, cryptophytes, and raphidophytes.

The results of this study suggest that retrieving PFTs in the SoG is highly challenging, likely as a result of the strong contributions of CDOM and particulate suspended matter to the satellite-measured reflectance signal. Further work is required to improve the algorithm and should rely on the following: better sensor vicarious calibration and atmospheric correction of the measured radiance signal; considerations of in situ data quality, including analytical errors; use of broader data-driven pigment methods and quantitative in situ validation utilizing a variety of methods; representation of the phytoplankton group's dynamic range, and; improving the spatial-temporal quality of the matchups. Also, considering that the retrieval algorithm is an emerging technique, further research is necessary to investigate the over-estimation of the lower ranges and under-estimation of higher ranges of PFT Chla, and to promote the introduction of non-linear fitting in the regression model. Finally, another limitation of our EOF-based PFT analysis is the lack of estimation of per-pixel uncertainty, which to our knowledge has only been reported by Brewin et al. (2017) and Xi et al. (2021). In our study, we used OLCI Rrs as the spectral input data derived through the POLYMER algorithm, and the uncertainty for the POLYMER derived Rrs is currently under development (Steinmetz, personal communication). Hence, performing per-pixel uncertainty of our OLCI-derived PFTs is currently not feasible. Besides, the advance of POLYMER is still an ongoing field, so future studies that aim to elucidate PFTs using POLYMER Rrs should be able to justify the per-pixel uncertainty estimates of PFTs using an updated version of POLYMER, especially over coastal waters.

Despite the shortcomings and uncertainties of our satellite product, our study shows the potential of OLCI for deriving major PFTs over a large spatiotemporal scale, which are not achievable via ship-based measurements. With continued research, the methods deployed here will greatly help to further our understanding of phytoplankton dynamics in complex coastal regions such as SoG and aid in the management of critical species such as migrating Pacific Salmon.

Chapter 5: Highly resolved surface phytoplankton community composition on the West Coast of Canada derived from hyperspectral radiometry

Perumthuruthil Suseelan Vishnu[a*], Justin Del Bel Belluz[b], Hongyan Xi[c], Midhun shah Hussain[d], Astrid Bracher[b,e], Maycira Costa[a]

[a]SPECTRAL Remote Sensing Laboratory, University of Victoria, Victoria, British Columbia, Canada.

[b]Hakai Institute, Victoria, British Columbia, Canada.

[c]Phytooptics Group, Physical Oceanography of Polar Seas, Climate Sciences, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany.

[d]Department of Marine Biology, Microbiology and Biochemistry, Cochin University of Science and Technology, Kerala, India.

[c,e]Department of Physics and Electrical Engineering, Institute of Environmental Physics, University of Bremen, Bremen, Germany.

*Corresponding author. Vishnu Perumthuruthil Suseelan[a]

E-mail address: psvishnu2014@gmail.com

Abstract

Phytoplankton form the foundation of the marine food web and play a distinct role in fisheries production and global biogeochemical cycles. Therefore, quantitatively characterizing phytoplankton community composition is essential, particularly in a changing climate. However, knowledge of phytoplankton dynamics is lacking, notably in dynamic coastal waters, due to the spatial-temporal constraints of in situ sampling and difficulties in deriving community compositions via satellite. Here, we work to address these issues by using highly resolved hyperspectral radiometry to derive phytoplankton community composition along a regularly traveled ship of opportunity ferry track in optically Case-2 waters on the west coast of Canada. First, an empirical orthogonal function (EOF)-based algorithm was developed using CHEMTAX-derived phytoplankton group chlorophyll-a (Chla) concentrations and corresponding EOF scores derived from hyperspectral remote sensing reflectance. Second, a cross-validation was performed to quantitatively assess the robustness of the fitted regression model for each phytoplankton group and total chlorophyll-a (TChla; mg/m^3) concentration. The Level-1 evaluation of our algorithm showed good retrievals for diatoms, cryptophytes, green algae, raphidophytes and TChla, moderate retrievals for haptophytes and poor performance for dictyochophytes, cyanobacteria, and dinoflagellates. Furthermore, independent validation showed good agreement with HPLC-derived TChla and ferrybox-derived Chla fluorescence (CHL_{FL}). Overall, our EOF-based results followed regionally expected spatial-temporal trends with diatom-dominated spring blooms and succession to high diversity flagellate-dominated summer conditions. Furthermore, our algorithm captured strong raphidophyte blooms composed of the harmful algae species, *Heterosigma akashiwo*. Additionally, our high-resolution transect plot highlighted peaks in raphidophyte and diatom coinciding with a sharp decline in SSS suggesting a frontal region within the Fraser River plume, which has been associated with the highest regional primary productivity due to increased nutrients. These features would not have been easily captured via in situ sampling and may provide valuable information on the drivers of phytoplankton community composition. Therefore, our study highlights the potential of using highly-resolved hyperspectral radiometry-derived phytoplankton community composition outputs in optically dynamic coastal waters like the Strait of Georgia (SoG), a critical habitat for diverse pelagic fish species, including Pacific salmon.

5.1 Introduction

Phytoplankton are responsible for 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 many having 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 the vertical export of carbon and nutrients to the deep ocean (Le Quéré et al., 2005; Guidi et al., 2009; Friedland et al., 2021). Variability in phytoplankton community composition is largely driven by various physical (e.g., temperature, salinity, light, density), chemical (e.g., inorganic nutrients), and biological (e.g., grazing pressure and viral infection) factors (Margalef, 1978; Evans and Brussaard, 2012). Phytoplankton communities quickly respond to these various drivers, and as a result, they are key indicators of environmental and anthropogenic change. Therefore, understanding the links between phytoplankton and their various drivers 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 conventional light-microscopy, size-fractionated filtration (SFF) of Chlorophyll-a, imaging flow cytometry (Dashkova et al., 2017), imaging flow cytobot (Olson and Sosik, 2007), high-performance-liquid chromatography (HPLC) pigment analysis, and next-generation sequencing (Lin et al., 2019). These techniques have advantages and limitations; however, they are often 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, space-based observations have allowed quantitative assessment of surface ocean phytoplankton biomass dynamics in terms of total chlorophyll-a concentrations (hereafter TChla) (e.g., McClain, 2009; Dai et al., 2023). However, TChla does not provide direct insight into community compositions, and contemporary research has strived to develop methods to derive compositions and size classes via remote sensing. These have included approaches using abundance (e.g., Uitz et al., 2006; Brewin et al., 2010; Hirata et al., 2011; Brotas et al., 2013; Lamont et al., 2018; Gittings et al., 2019), ecological (e.g.,

Raitsos et al., 2008; Palacz et al., 2013; Mouw, 2019), reflectance\radiance (e.g., Alvain et al., 2005; Ciotti and Bricaud, 2006; Devred et al., 2006; Bracher et al., 2009; Ben Mustapha et al., 2014; Rêve-Lamarche et al., 2017; Xi et al., 2020; Liu et al., 2021; Vishnu et al., 2022), and backscattering-based methods (e.g., Kostadinov et al., 2009). A detailed review of these different approaches can be found in IOCCG (2014), Bracher et al. (2017), and Mouw et al. (2017). However, 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 (IOCCG, 2014; Mouw et al., 2017; Dierssen et al., 2021; Kramer et al., 2022).

In order to derive information on phytoplankton community composition, 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 (Lee et al., 2007). For example, hyperspectral optical data has been used to identify spectral peaks associated with biomarker pigments (e.g., Torrecilla et al., 2011; Uitz et al., 2015; Xi et al., 2015), to retrieve accessory pigments concentration on a global scale (Kramer et al., 2022), to detect toxic harmful dinoflagellates, *Karenia brevis* (e.g., Craig et al., 2006), and to derive Chlorophyll-a (hereafter Chla) concentrations of multiple phytoplankton groups across the Atlantic Ocean (e.g., Bracher et al., 2020) and North Pacific (Konik et al, submitted). However, no ocean color sensors with hyperspectral capability are currently operational. The upcoming Plankton Aerosol Cloud and Ecosystem (PACE) mission from NASA will be the first ocean hyperspectral satellite carrying a hyperspectral imaging spectro-radiometer Ocean Color Instrument (OCI) that acquires visible light spectrum with 5 nm resolution, and will thus allow for considerable improved characterization of global phytoplankton composition when compared to current multispectral sensors (Werdell et al., 2019). Specifically, OCI is expected to improve the inversion of IOPs that can yield information about phytoplankton community structure (Cael et al., 2020) and extraction of accessory pigments spectral peaks (Chase et al., 2013; Kramer et al., 2022), thereby improving the global retrieval of phytoplankton composition (Wolanin et al., 2016; Xi et al., 2017; Werdell et al., 2018; Cael et al., 2020).

Here, we aim to derive Chlorophyll-a (Chla) of phytoplankton community composition using an empirical orthogonal function (EOF) based hyperspectral algorithm, applied to above-water hyperspectral radiometry, and assess their seasonal and spatial variability within the

Strait of Georgia on the west coast of Canada. This algorithm was adapted from Xi et al. (2020), and the regression model for each phytoplankton type was built considering CHEMTAX-derived phytoplankton community composition data and corresponding EOFs score subsets derived from the hyperspectral remote sensing reflectance. In addition, cross-validation techniques were used to quantitatively determine the robustness of the fitted regression model, and independent validation of the TChla retrievals was performed utilizing ferry box-derived fluorescence data. Transect analysis was conducted on high-resolution phytoplankton community composition data, paired with their corresponding sea surface salinity data obtained from ferrybox data, to investigate the influence of Fraser River plume waters. This work provides a methodological baseline for the upcoming hyperspectral PACE satellite (Werdell et al., 2019).

The Strait of Georgia is a highly dynamic Case-2 estuarine system that provides habitat for various economically and culturally valuable species, such as Pacific Salmon, Herring, and Hake, and offers opportunities for recreational fisheries and aquaculture activities (Ketchen et al., 1983; Beamish et al., 1994; Beamish et al., 2012; Thomson et al., 2012). In this region, phytoplankton have been linked to the survival and growth of zooplankton species such as *Neocalanus plumchrus*, an important food for juvenile salmon (Sastri and Dower, 2009; Perry et al., 2021). Further, Suchy et al. (2022) demonstrated that differences in spring bloom timing resulted in mismatches between phytoplankton and larger crustacean zooplankton species abundance, having significant implications for upper trophic levels; however, relatively few studies have investigated long-term and spatially resolved links between phytoplankton community composition and food web dynamics. Furthermore, large regional harmful algae blooms (HABs) have been shown to have detrimental effects on various fisheries and aquaculture species, and a lack of monitoring capacity has resulted in the permanent closure of shellfish beds used for thousands of years by local First Nations (Rensel et al., 2010; Haigh and Esenkulova, 2014; Esenkulova et al., 2021). Therefore, highly resolved phytoplankton composition data derived from hyperspectral remote sensing would fill knowledge gaps in understanding regional links between phytoplankton and zooplankton communities and their implications for higher trophic levels. In addition, these data would allow for better environmental monitoring of the negative effects of HABs and the understanding of how anthropogenic influences will change phytoplankton dynamics in the future.

5.2 Materials and Methods

5.2.1 Study Area

The Strait of Georgia (SoG; **Figure 29**), located on the west coast of Canada between Vancouver Island and the Mainland of British Columbia (BC), is a highly productive region in the northeast Pacific (NEP). This region receives substantial freshwater inputs from the Fraser River, resulting in outward estuarine flow of surface waters above incoming high salinity nutrient-rich deep water from the Pacific Ocean (Li et al., 2000; Masson and Cummins, 2004; Phillips and Costa, 2017). The SoG is connected to the Pacific Ocean through the Juan de Fuca Strait in the south and Johnstone Strait in the north (Harrison et al., 1983); however, the dominant estuarine exchange with the open ocean occurs through the Juan de Fuca Strait (Dosser et al., 2021). As observed in the other optically Case-2 water systems, the distribution of biogeochemical variables and optically active constituents in the SoG is highly driven by the Fraser River dynamics, which peak discharge happens during the summer (Peña et al., 2016). Water near the Fraser River plume show high concentration of CDOM and TSM, limiting light for primary productivity, whereas waters directly outside the plume show elevated levels of Chla (Johannessen et al., 2006; Loos and Costa, 2010). Recent studies demonstrated that CDOM and TSM in this region are highly variable during spring, between 0.21 to 3.45 m^{-1} and 1.60 to 65.63 mg/L, respectively (Loos and Costa, 2010; Phillips and Costa, 2017; Vishnu et al., 2022). In agreement with this observation, several studies recorded high spatial-temporal variation in phytoplankton biomass and primary productivity in the Strait (e.g., Harrison et al., 1983; Mackas and Harrison, 1997; Yin et al., 1997; Masson and Peña, 2009; Suchy et al., 2019; Del Bel Belluz et al., 2021; Nemcek et al., 2023). Seasonal phytoplankton variability in this region is primarily limited by solar radiation during winter, while during the summer, limitation arises from nitrate drawdown and zooplankton grazing (Harrison et al., 1983; Allen and Wolfe, 2013; Del Bel Belluz et al., 2021; Nemcek et al., 2023). In addition, primary productivity in the SoG is relatively high (280 g C m^{-2}) compared to other regions in the northeast Pacific, with elevated levels often found at frontal regions and near the Fraser River plume (Harrison et al., 1983; Sutton et al., 2013; Johannessen et al., 2021). In terms of the phytoplankton biomass, Chla concentration in SoG varies from < 1 mg/m^3 in winter to > 15 mg/m^3 during spring (Harrison et al., 1983; Peña et al., 2016; Suchy et al., 2019; Giannini et al., 2021). Diverse phytoplankton communities are often observed in the SoG, but seasonal succession generally follows expected trends for temperate coastal waters, with diatoms being the dominant bloom-

forming species during the spring bloom and periodical autumn blooms. Bloom-forming diatoms are typically *Thalassiosira* spp., *Skeletonema costatum*, and *Chaetoceros* spp., but *Pseudo-nitzschia* are also known to form periodic blooms (Harrison et al., 1983; Esenkulova et al., 2021; Del Bel Belluz et al., 2023). By the end of spring, diatoms are generally replaced by small flagellates comprised of cryptophytes, prasinophytes, dinoflagellates, and dictyochophytes (Del Bel Belluz et al., 2021; Vishnu et al., 2022; Nemcek et al., 2023), which rarely reach blooming levels.

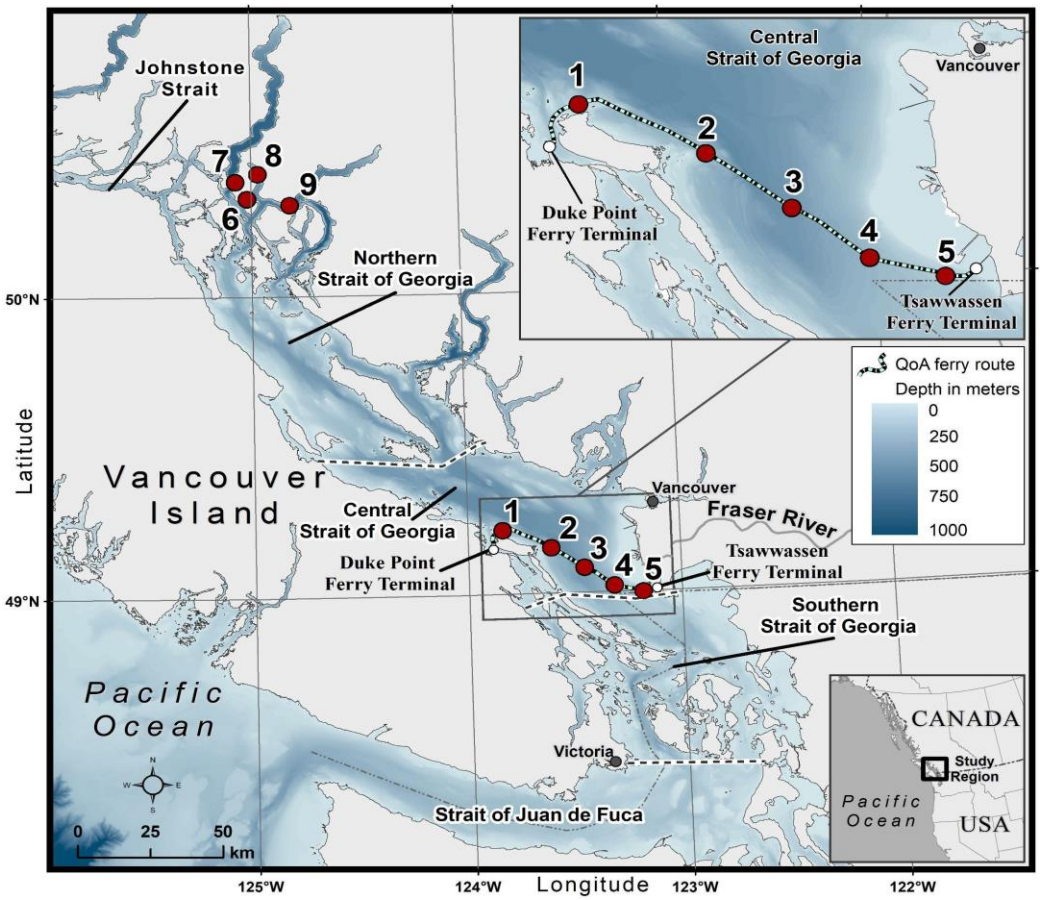

Figure 29. Study area map and in situ data collection locations in the central and northern Strait of Georgia. The running lines represent the Queen of Alberni ferry's track from Duke Point, Nanaimo on Vancouver Island to Tsawassen on Mainland, British Columbia. The color scale shows depth in meters.

5.3 Data Acquisition

In this section, we describe the 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 (CDOM), which were used to describe regional bio-optical properties. Additionally, the acquisition and processing of the above-water hyperspectral

radiometer are described, and we present a regional EOF-based hyperspectral algorithm to retrieve the Chla concentration of phytoplankton group composition.

5.3.1 Ferrybox Data

All physical and chemical oceanographic properties, such as SSS, and Chla 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) (station 1 to 5; **Figure 29**), facilitated by Ocean Networks Canada (ONC) (https://data.oceannetworks.ca/DataSearch; Traver-Smith et al., 2021). Specifically, SSS were acquired using a Seabird SBE19 thermosalinograph, whereas CHL_{FL} was estimated using a WET Labs ECO Triplet fluorometer and corrected for both biofouling and non-photochemical quenching (NPQ) (Travers-Smith et al., 2021; Owens et al., 2022). ONC maintains and cleans these sensors weekly or bi-weekly, depending on the season, followed by a comprehensive data quality control process (refer to Owens et al., 2022 for the detailed QA/QC). Only fully quality-controlled data were used for analysis. The frequency of the sensor measurement was 1 Hz, and the quality-checked sensor-acquired data were binned into 1-minute intervals to match with the in situ water samples. The ferrybox is located below the main deck and pumps seawater through a 1.5 m long pipe from approximately 2 m depth (Travers-Smith et al., 2021).

In the northern SoG (**Figure 29**), SSS was acquired using a calibrated Fisherbrand traceable salinity meter pen (accuracy ±2% Full-scale +1 digit).

5.3.2 In situ Water Samples

Surface water samples were collected from two locations in SoG. In the central SoG (**Figure 29**; N=108)**,** surface water samples for analysis of HPLC TChla and phytoplankton pigment concentrations, TSM, and CDOM were collected during the 2018 and 2019 spring (March-April) and summer (May-September) seasons. Samples were collected at different stations using a seawater intake pump installed on QoA. Triplicate water samples for the different analyses were collected from approximately 2m depth (Halverson and Pawlowicz, 2013; Wang et al., 2019; Travers-Smith et al., 2021) from each station, kept under a dark environment and immediately filtered onboard to minimize compositional changes to phytoplankton pigments and dissolved organic matter (Mueller et al., 2003).

In the northern SoG (**Figure 29**; stations 6 to 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.

5.3.2.1 Absorption by Colored Dissolved Organic Matter

Surface water samples for CDOM measurements from each station were filtered through 47 mm Whatman 0.2 µm membrane filter paper and collected in acid-washed, precombusted (550° C for 1 hour) borosilicate glass bottles (Ferrari et al., 1996). Before the sample collection, the acid-washed glass bottles were triple rinsed with sample water, then kept in a cold environment, and stored under 4°C until the analysis in the laboratory. Before the analysis, the discrete samples were allowed to reach room temperature to minimize the temperature difference between the samples and the blank. The absorption of CDOM was measured using an Ocean Optics USB 4000 spectrophotometer (300-850 nm, 0.2 nm resolution) against Milli-Q water as blank. The sample absorption was converted to CDOM absorption using Beer-Lambert's Law equation (Ferrari et al., 1996).

$$\mathrm{aCDOM}\ (\lambda) = \frac{2.3030 * \mathrm{A}\ (\lambda)}{\mathrm{L}} \qquad (1)$$

A (λ) is the spectral absorption of the sample, and L is the cuvette path length in meters (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.

5.3.2.2 Total Suspended Matter Concentration

Surface water samples of approximately 0.5–2 L were filtered through precombusted and pre-weighed 47 mm Whatman GF/F 0.7 µm Glass microfiber filter (GF/F) paper to quantify the TSM concentration. The precombustion process reduces the pore size of the glass fiber filter to approximately 0.3 µm, thereby increasing efficiency by retaining more particulate matter (Nayar and Chou, 2003). Following sample filtration, 200 ml of distilled water was filtered to minimize sea salt retention (Stavn et al., 2009). After filtration, filters were kept on dry ice, transported to the laboratory, and stored at -80° C until the analysis. In the laboratory, each filter was dried at 60° C for 24 hours in a clean oven and reweighted to 6 degrees of precision (Röttgers et al., 2014). TSM was calculated by following Röttgers et al. (2014). The final TSM concentration at each station was the average of the triplicates with a CV < 25% for all sample sets.

5.3.2.3 HPLC Pigment Concentrations

Surface water samples for the duplicate measurement of HPLC pigments from each station were filtered, under low pressure (≤5mmHg) through 25 mm Whatman GF/F 0.7 µm filters. Samples were filtered to color with filtered volumes ranging from 0.5–2 L. Filters were immediately placed in dry ice and then stored at -80˚C until analysis. HPLC analysis was performed at the University of South Carolina Baruch Institute of Marine and Coastal Sciences following Pinckney (2010).

The HPLC-derived phytoplankton pigment dataset 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_3 (Chl c_3), Chlorophyll c_1c_2 (Chlc_1c_2), Peridinin (Perid), 19'Butanoyloxyfucoxanthin (BF), Fucoxanthin (Fuco), 19′Hexanoyloxyfucoxanthin (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 **Table S9** and **Table S10**. 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, groups and output results are identical to those used in Vishnu et al. (2022), which were used for similar purposes on Sentinel 3A satellite data and where extensive analysis and description were provided rationalizing their use (Vishnu et al., 2022). In general, groups and ratios are comparable to those of Del Bel Belluz et al. (2021), which were guided by microscopy and pigment analysis, but here, raphidophytes were included in the central SoG data as a result of large observed *Heterosigma* blooms (evidence and rationale for the inclusion of this group is discussed in Vishnu et al. (2022). Furthermore, CHEMTAX analysis was performed independently on the northern and central SoG data to account for regional differences in pigment ratios. Specifically for the central SoG data, CHEMTAX was run in bins delineated by TSM concentrations to account for differences in pigment ratios driven by high variability in light conditions along the ferry track, which traverses the Fraser River plume (Loos and Costa, 2010).

5.3.3 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 (1) autonomous solar tracker HyperSAS (hereafter SAS-ST) and the northern SoG using a (2) manually operated HyperSAS (hereafter SAS) operated from a small vessel. The SAS-ST was installed on the BC ferry QoA, to measure radiances and irradiance with solar tracking capability that permits autonomous operation following the protocols in Wang and Costa, (2022) and Giannini et al. (2021). The system comprises three hyperspectral radiometers, where Lt(λ) represents the total radiance measured by the sea-viewing sensor, Li(λ) denotes the sky radiance, and the third sensor, Es(λ), measures the total irradiance. From these measurements, remote sensing reflectance, Rrs, was derived following Mobley (1999). The SAS-ST was installed 14 meters above the water surface to avoid shadowing effects from the ship, spray, and sun glint (Wang and Costa, 2018) and calibrated by the manufacturer (Satlantic, Inc.). The Lt (λ) and Li (λ) sensors were installed with fixed viewing zenith angles, θv = 45, and viewing-sun azimuth, φv = 120^0, to avoid the effects of glint (Hooker and Morel, 2003). The ideal φv was maintained using an autonomous "steeper motor platform" (Satlantic Inc), which autonomously adjusted the sensor to the required position according to the ship heading and sun azimuth.

Radiometric data was acquired 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) and Sentinel 3A and 3B imagery acquisitions. The data from the SAS-ST were processed with PySciDON (Python Scientific Framework for Development of Ocean Network application) to calculate Rrs (Vandenberg et al., 2017). The data, acquired every 2 seconds were binned into 1-minute intervals during processing, corresponding to a spatial coverage of 679 meters, aligning with the discrete water samples and ferrybox measurements. The mean, median, and CV were calculated from each station, and each Rrs represents CV < 10%.

Remote sensing reflectance was calculated using the following equation.

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

Where Lt(λ) is the total radiance measured by the sea-viewing sensor, Li (λ) is the sky radiance, Es (λ) is the total irradiance, and ρ_s is the fraction of sky radiance (Mobley, 1999).

The value for ρ_s was obtained from the look-up table provided by Mobley (1999), according to sky condition, viewing geometry, sea state, and wind speed (Mobley, 1999, Hooker and Morel, 2003, Brando et al., 2016). Wind data were acquired from the meteorological stations on Entrance Island, Sandhead CS, and Tsawwassen Ferry Auto. All the remote sensing reflectance data from the SoG were corrected for the bidirectional effect (BRDF) using the code by Wang and Costa (2022) following the method of Lee et al. (2011).

For the NSoG, 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 calculation of Rrs was the same as for the SAS-ST.

5.3.4 Regionally-tuned EOF-based Hyperspectral Algorithm

The EOF-based hyperspectral algorithm to derive the Chla concentration of multiple phytoplankton groups made use of the in situ hyperspectral Rrs across the visible spectrum (400 to 700 nm) with a 1 nm interval, similar to Bracher et al. (2020) and Lange et al. (2020), and this EOF-based regression model was recently tested globally by Xi et al. (2020) using the multispectral Rrs.

Here, we adapted Xi et al. (2020) algorithm to the SoG data by considering 103 valid matchup datasets comprising SAS-ST, SAS Rrs, and the CHEMTAX-derived phytoplankton community composition. A schematic workflow of the regression model development and validation is shown in **Figure 30**. Following Xi et al. (2020), the hyperspectral SAS-ST and SAS Rrs were subject to singular value decomposition (SVD) to derive the different EOF modes (dominant signals) that best described the variance in the reflectance spectrum. The following are the general steps of this approach.

Step 1: Before deriving the EOF modes, the SAS-ST and SAS Rrs were standardized by subtracting the spectral average and then normalized by the spectral standard deviation. We used SAS-ST and SAS Rrs spanning 400-700 nm with a 1 nm interval for the EOF analysis. First, the standardized Rrs data, denoted as matrix X, were collocated to the respective phytoplankton group Chla. Then, the standardized SAS-ST and SAS Rrs were subjected to a singular value decomposition to derive the EOF modes.

$$X = U \Sigma V^T \quad (3)$$

Where matrix U was the EOF scores (eigenvalues), matrix V^T was the spectral loading (also known as eigenvector), and vector Σ, singular values of matrix X in decreasing diagonal order.

The reflectance data from Step 1 was then binned into 1-minute bins to match the discrete water samples and ferrybox data. The following criteria were followed for the matchup: < 1 hour differecne and the Coefficient of Variation (VC) across the visible spectrum (400 to 700 nm) being < 10%. Median values are used to avoid outliers (Bailey and Werdell, 2006).

Step 2: A generalized linear model (GLM) was then built considering the log-transformed CHEMTAX-derived phytoplankton groups Chla (Cp) and the corresponding subset of EOF scores (U). Based on Bracher et al. (2015) and Xi et al. (2020), insignificant EOFs were excluded, and a stepwise routine was implemented to search for the most significant EOF score by minimizing the Akaike information criteria (AIC). Through AIC criteria, only significant EOF scores (U) were fed into the regression model. The regression algorithm for the phytoplankton groups Chla prediction is :

$$\ln(\mathrm{Cp}) = \mathrm{a} + \mathrm{b}_1\mathrm{u}_1 + \mathrm{b}_2\mathrm{u}_2 + \cdots \mathrm{b}_\mathrm{n}\mathrm{u}_\mathrm{n} \quad (4)$$

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

Step 3: A level-1 cross-validation (or model evaluation) technique was then performed following Bracher et al. (2015) and Xi et al. (2020) to assess the fitted regression model robustness using 500 permutations and by splitting the dataset into two subsets: 80% of the data were used for regression model training and 20% for the regression model validation. The training dataset (80%) included X^{train} (Standardized SAS-ST and SAS Rrs), C^{train} (CHEMTAX phytoplankton groups Chla), and the validation dataset (20%) included X^{val} and C^{val}. A regression model selected through a stepwise routine was fitted between ln (C^{train}) and corresponding (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}) for the 500 permutations were recorded for the regression model assessment.

As part of the level- evaluation, the regression model assessment was performed by calculating the statistical metrics such as coefficient of determination (R^2), Mean Absolute Error (MAE), Median Percentage Difference (MDPD), and Bias. R^2, slope (*S*), and intercept (*a*) of the GLM were obtained using the log-scaled predicted ln($Chla_{pi}$) *versus* log-scaled

observed ln($Chla_{oi}$) phytoplankton group Chla concentration, while the MAE, MDPD, and Bias were based on the non-log transformed data. The equation of these statistical metrics was given by Xi et al. (2020) with the exception of MAE, which is expressed as follows:

$$\mathrm{MAE} = \sum_{i=1}^{i=M} \left| (\mathrm{Chla_{pi}} - \mathrm{Chla_{oi}}) \right| \quad (5)$$

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

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. While used for model development, this step was not performed on the SAS data due to its discrete nature. To do this, we projected the standardized SAS-ST onto EOF loading (V) to derive the new EOF scores (U) set. Then, this new set of EOF scores (U^{sat}) was fed into equation 4 to derive the spatial distribution of multiple phytoplankton groups. Where the intercepts (a) and the regression coefficients ($b_{1,2....n}$) were obtained from equation 4. Finally, an independent level-2 validation was performed for TChla utilizing the QA/QC ferrybox-derived CHL_{FL} data obtained during 2018 and 2019. The matchup extraction criteria used in the independent validation were the same as those used in the model training. It is worth noting that the level-2 validation was restricted to only HPLC-derived TChla due to the limited availability of independent phytoplankton groups Chla data that could be spatially and temporally matched with the ferry track samples. Following both level-1 and level-2 quantitative validation, the phytoplankton group Chla retrieved from the algorithm underwent additional qualitative evaluation based on the previous literature.

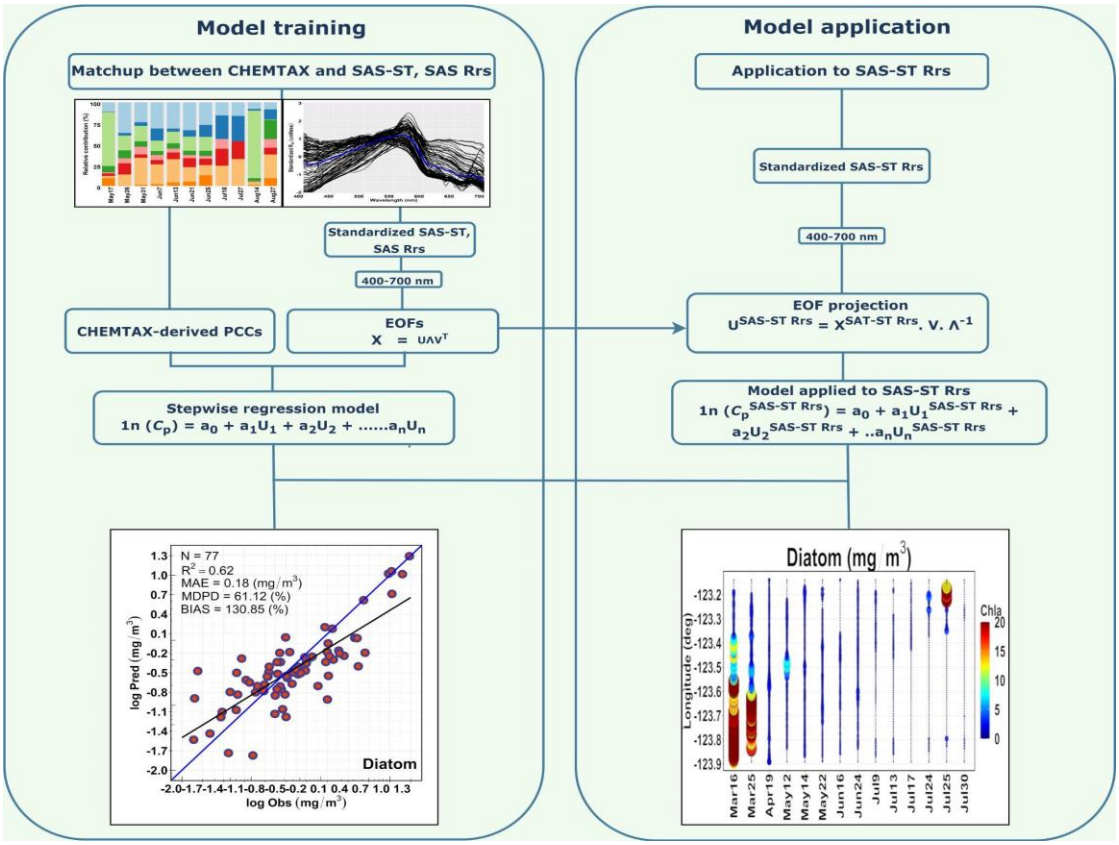

Figure 30. The flowchart illustrates the method to derive phytoplankton community composition using the empirical orthogonal function-based algorithm on hyperspectral remote sensing reflectance. The left side demonstrates the algorithm training section with a matchup dataset comprised of CHEMTAX-derived phytoplankton community composition and hyperspectral reflectance. The right side shows the algorithm application part to hyperspectral reflectance to derive the spatial distribution of multiple phytoplankton groups.

5.4 Results

5.4.1 Physical and Biogeochemical Properties

Surface physical and biogeochemical variables demonstrated trends typical of Case-2 waters, with high variability and seasonal variation throughout the year (**Figure 31**; **Table 6**). In general, SSS was highly variable, with the lowest salinities always observed near the Fraser River and the highest salinities on the Vancouver Island (western) side of the transect. Annually, spring 2018 showed the lowest SSS magnitudes (0.38 PSU) and 2019 showed the highest magnitudes (28.82 PSU) near the Fraser River and Vancouver Island sides of the transect, respectively. Similarly, TSM and CDOM always showed their highest magnitudes near the Fraser River, with their maxima 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 CDOM showed lower magnitudes and less variability in summer when compared to spring conditions. TChla was highly variable (**Figure 31**; **Table 6**), ranging

from 0.39 to 28.82 mg/m^3 and showing higher values in 2018 than in 2019. During the summer of 2018, an anomalous value of 120.11 mg/m^3 was observed (not shown in the table or the figure as it was far outside the range of other values) close to the Fraser River plume, highlighting summer bloom conditions 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 through much of the summer season.

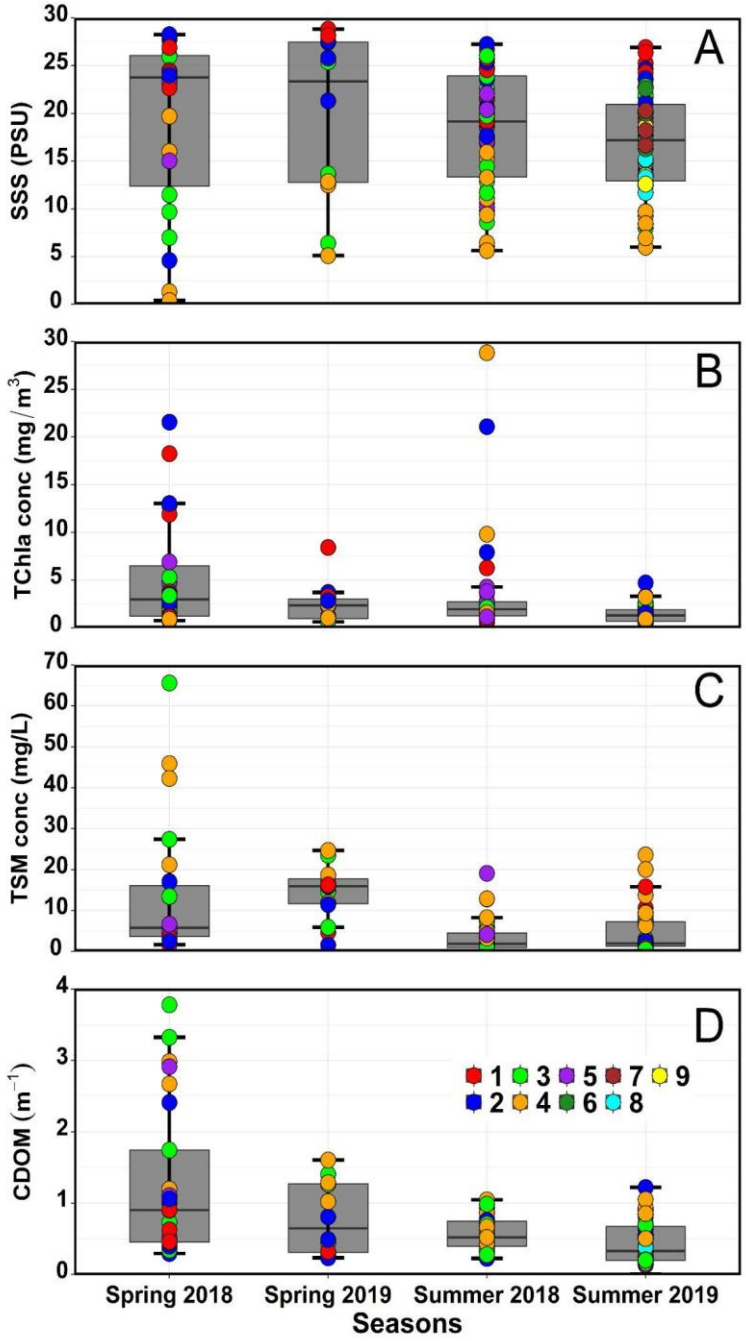

Figure 31. Seasonal variation of (**A**) sea surface salinity (SSS, PSU), (**B**) total chlorophyll-a (TChla; mg/m^3) concentration, (**C**) total suspended matter (TSM: mg/L) concentration, and (**D**) absorption by colored dissolved organic matter (CDOM; m^{-1}). The upper and lower end

of the box blot represents first and third quartile, respectively. Outside of the boxes, the whiskers on both quartiles indicate standard error. Colored points denote data from various stations (see **Figure 29**), with points outside the third quartile indicate extreme values.

Table 6. Seasonal distribution of biogeochemical variables such as SSS, TChla concentration, TSM, and CDOM during 2018 and 2019. Minimum and maximum were given in square brackets; the average was bold, and the standard deviation was in parentheses.

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

5.4.2 EOF Analysis on Hyperspectral Remote Sensing Reflectance

Figure 32 shows the spectral shape and magnitude of hyperspectral Rrs associated with their corresponding in-water constituents. The spectral shape of the above-water Rrs showed typical Case-2 water spectra with the broad peak observed between 550 to 700 nm due to the high backscattering from inorganic suspended sediments, the low reflectance found in the blue (400 nm) region driven by high absorption by CDOM, and the high reflectance in the green (550 nm) along with the Chla fluorescence peak at 680 nm representative of highly varying phytoplankton biomass (e.g., Komick et al., 2009; Carswell et al., 2017; Phillips and Costa, 2017; Kratzer et al., 2020). Application of SVD on the Rrs spectra resulted in nine significant EOF modes (**Table 7**), with the first three modes explaining 99.43 % of the total variance. The results of the ΔAIC applied to the significant EOFs to derive regression models for each phytoplankton group are shown in **Table 8.** Among the nine significant EOFs, EOF3 was the most important term for predicting TChla, diatom, dictyochophytes, raphidophytes, and cryptophytes. The other groups showed different EOFs, which were the most important term for predicting their concentration, indicating that AIC helped retain the statistically significant higher EOFs.

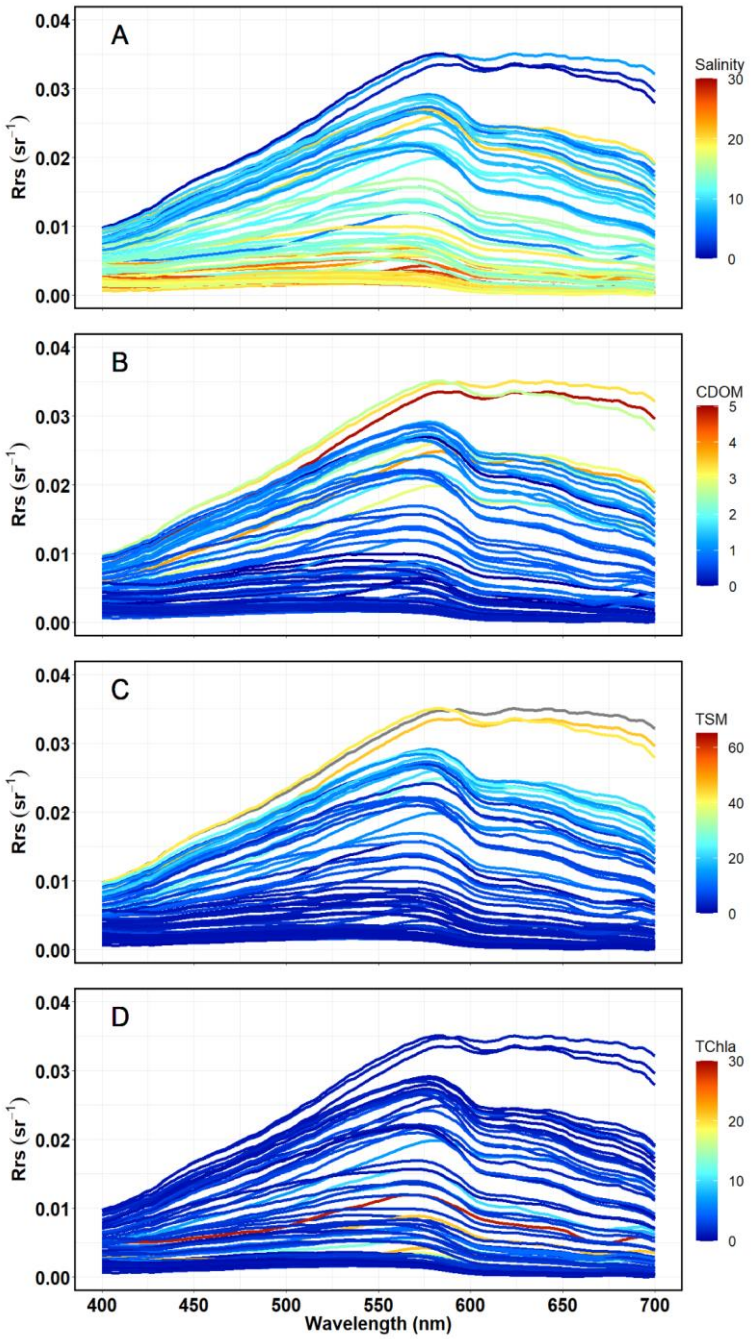

Figure 32. Hyperspectral remote sensing (Rrs; sr^{-1}) reflectance across the visible spectrum (400 to 700 nm) with each spectra colored by their corresponding (A) salinity (PSU), (B) absorption by colored dissolved organic matter (CDOM; m^{-1}), (C) total suspended matter concentration (TSM; mg/L), and (D) total chlorophyll-a concentration (TChla; mg/m^3).

Table 7. The percentage of variance explained by the first nine EOFs derived from the matchup dataset comprised of SAS-ST, SAS Rrs, and phytoplankton composition.

% of variance	EOF1	EOF2	EOF3	EOF4	EOF5	EOF6	EOF7	EOF8	EOF9
SAS-ST, SAS Rrs	67.65%	30.06%	1.72%	0.30%	0.12%	0.07%	0.02%	0.01%	0.01%

Table 8. Δ AIC derived from the EOF analysis on the matchup dataset between SAS-ST, the total chlorophyll-a concentration (TChla; mg/m^3) and the eight CHEMTAX phytoplankton groups.

SAS-ST, SAS Rrs	EOF1	EOF2	EOF3	EOF4	EOF5	EOF6	EOF7	EOF8	EOF9
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

The spectral shapes of each EOF were regulated by the optically active constituents (OACs) in the water column (**Figure 33**). The first EOF mode (**Figure 33**), which explained 67.65% (see **Table 7**) 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. The second EOF explained 30.06% of the total variance and showed a different shape than EOF1, with a sharp decrease from 550 to 400 nm driven by CDOM absorption and a broad 550 to 700 nm reflectance peak, indicative of backscattering by inorganic suspended sediments. In turn, EOF3, which explained 1.72% of the total variance, exhibited a trough from 500 to 400 nm, likely associated with phytoplankton variability and a Chla fluorescence peak near 680 nm. EOF4 explained 0.30% of the total variance and did not resemble the first three EOFs or show characteristic features. Finally, EOF5 explained 0.12% of the total variance and showed a trough from 486 to 400 nm and a trough between 486 to 543 nm, followed by a spectral peak at 600 nm and 680 nm. The higher EOF modes ≥ 6 only explained a nominal portion of the total variance (≤ 0.07%).

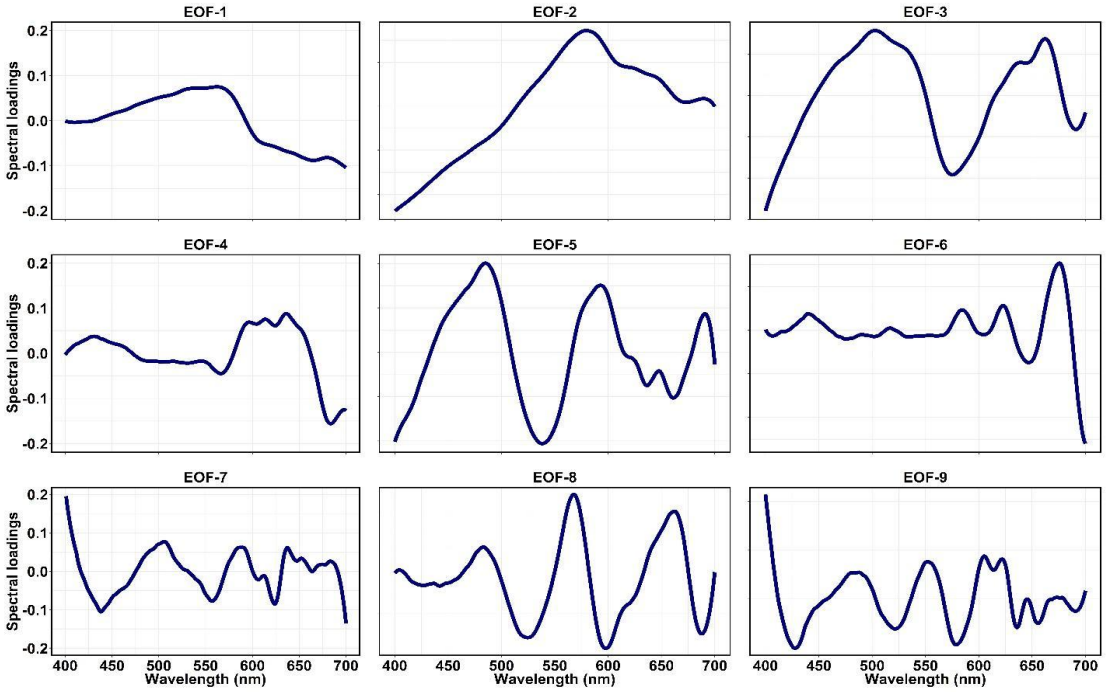

Figure 33. Spectral loadings of the nine Empirical Orthogonal Function (EOF) modes derived from SAS-ST and SAS Rrs using the singular value decomposition. These spectral loadings were obtained from the matchup between SAS-ST, SAS Rrs, and the total chlorophyll-a (TChla; mg/m^3) concentration.

5.4.3 Algorithm Evaluation

Regressions between TChla and the Chla concentration of the eight multiple phytoplankton groups derived using CHEMTAX (x-axis) *versus* the EOF-based hyperspectral algorithm (y-axis) are shown in **Figure 34**. 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 the sample. For example, out of the 103 valid samples, dictyochophytes (N = 66) and diatoms (N = 77) showed relatively low numbers of samples. Overall, phytoplankton groups Chla and TChla concentration is well predicted with our EOF-based hyperspectral model, with the best retrievals noted for diatoms, raphidophytes, green algae, and cryptophytes (R^2 = 0.61, 0.57, 0.57, 0.53, respectively; **Table 9**), and moderate retrievals were observed for haptophytes (R^2 = 0.47; **Table 9**). However, poor correlations were observed for dinoflagellates, dictyochophytes, and cyanobacteria (R^2 = 0.24, 0.36, 0.38, respectively; **Table 9**). The best performance was observed for TChla (R^2 = 0.81). The MDPD ranged from 24.88 to 65.01 %, with the lowest value recorded for TChla (24.88 %) and the highest for raphidophytes (65.01 %). MAE was calculated on non-log transformed data and varied according to the Chla concentration of each group. For example, the highest MAE was recorded for diatoms (49.59

mg/m^3) and TChla (13.67 mg/m^3), followed by raphidophytes (4.91 mg/m^3), whereas the lowest MAE was noted for haptophytes (0.07 mg/m^3).

As part of the level-1 evaluation, the robustness of the fitted regression model was investigated using cross-validation statistics (**Table 9**). The statistics obtained from the cross-validation were comparable with the full-fit statistics except for the dictyochophytes, cyanobacteria, and dinoflagellate groups, which showed considerably lower performance (R^2_{cv} = 0.22, R^2_{cv} = 0.20, R^2_{cv} = 0, 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%).

The level-2 evaluation or an independent evaluation restricted to the EOF-based TChla against the flowthrough ferrybox CHL_{FL}, showed a highly significant correlation (R^2 = 0.70; N=728, p≤0.0001), with the predicted TChla capturing the trends and features observed in the ferrybox CHL_{FL} (see **Figure 35B**). However, it is important to note that the EOF-based hyperspectral algorithm considerably overestimated (BIAS = 90.77%) the ferry box CHL_{FL}, especially at the high TChla concentrations. In turn, the MAE for this relationship was lower (2.01 mg/m^3), and the MDPD was roughly 61.62%.

Table 9. Full-fit and cross-validation statistics for the total chlorophyll-a concentration (TChla; mg/m^3) and the eight phytoplankton groups.

	N	R^2	MAE	MDPD	R^2cv	MAEcv	MDPDcv	*p*-value
TChla	103	0.81	0.46	24.88	0.73	13.67	29.70	2.210^{-16}
Diatom	77	0.61	0.18	61.12	0.60	49.59	68.10	2.210^{-16}
Raphido	77	0.57	0.13	65.01	0.41	4.91	71.87	2.870^{-15}
GA	96	0.57	0.09	35.45	0.45	0.18	42.48	2.210^{-16}
Crypto	102	0.53	0.12	40.14	0.45	0.39	44.90	2.210^{-16}
Hapto	87	0.47	0.03	43.88	0.32	0.07	49.56	1.876^{-13}
Cyano	88	0.38	0.03	46.30	0.20	0.16	53.10	1.413^{-11}
Dictyo	66	0.36	0.07	47.71	0.22	0.93	54.18	8.160^{-7}
Dino	82	0.24	0.08	49.68	0	0.33	58.94	2.917^{-6}

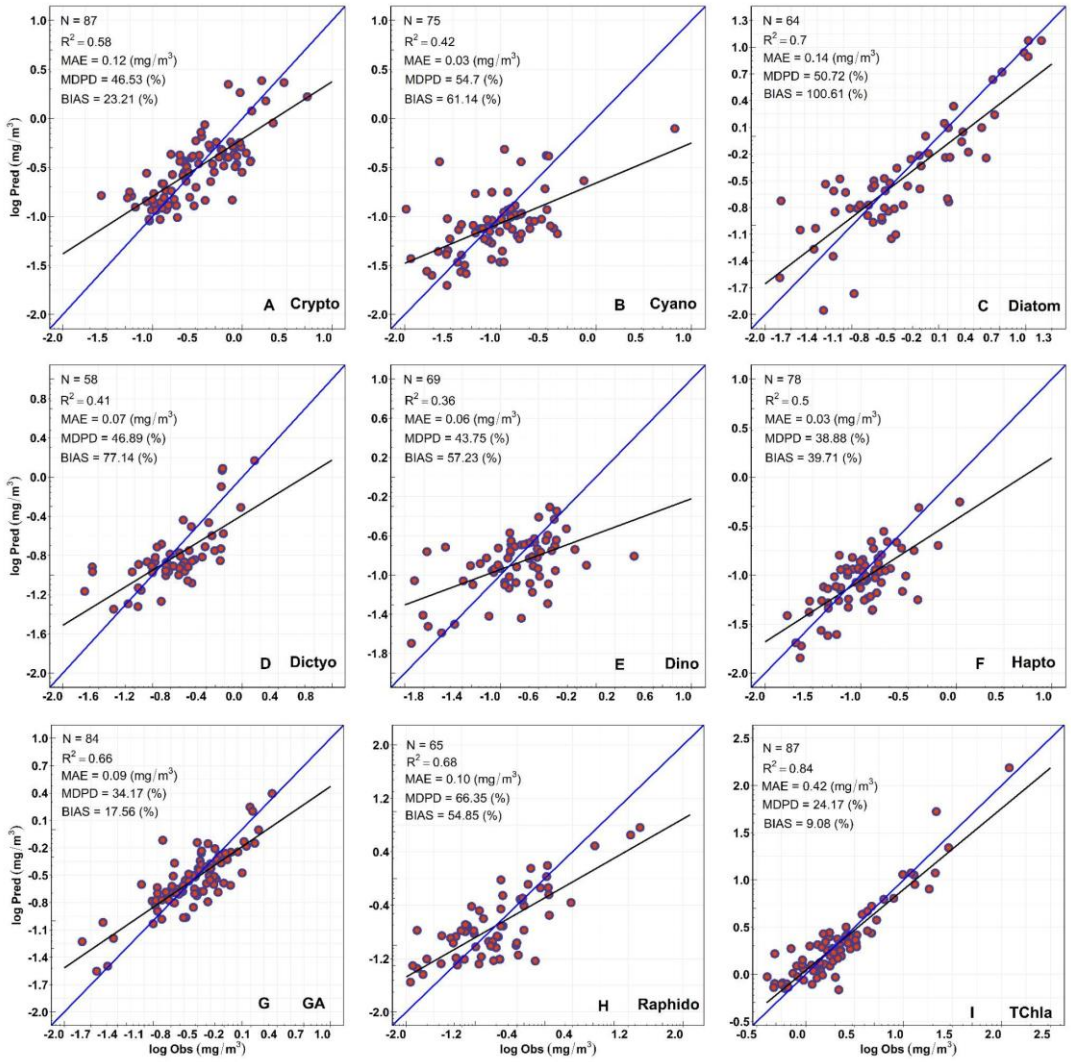

Figure 34. Regression between CHEMTAX derived (x-axis) and the corresponding Empirical Orthogonal Function (EOF) predicted (y-axis) phytoplankton composition such as cryptophytes, cyanobacteria, diatom, dictyochophytes, dinoflagellates, haptophytes, green algae, raphidophytes, and total chlorophyll-a (TChla) concentration.

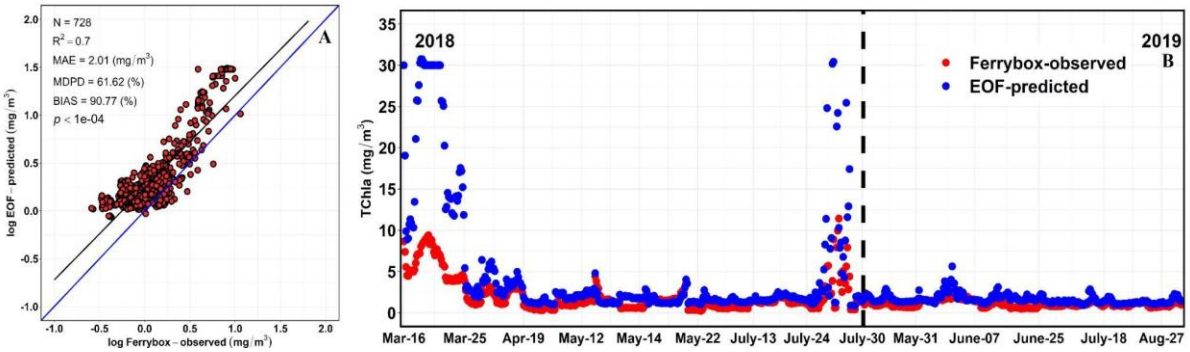

Figure 35. (A) Regression between ferrybox-observed (x-axis) and the corresponding Empirical Orthogonal Function (EOF) predicted (y-axis) total chlorophyll-a (TChla) concentration. (B) Spatial-temporal variability between the ferrybox observed *versus* the corresponding EOF-predicted TChla concentration. 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 **Figure 36** and **Figure 37**.

5.4.4 Application of EOF-based Algorithm to SAS-ST Rrs Data

After the level-1 and level-2 evaluations, we applied the EOF-based algorithm to independent SAS-ST Rrs data to elucidate the spatial-temporal dynamics of highly resolved phytoplankton composition along the ferry track (**Figures 36 and 37**). The algorithm was applied to all 8 of the phytoplankton groups, including dictyochophytes, cyanobacteria, and dinoflagellates, which showed poor retrievals; however, these groups were included in the spatial-temporal group analysis to evaluate if they followed expected trends. In March 2018, a clear seasonal spring diatom bloom was observed across much of the western and central portions of the ferry track (**Figure 36A**), with the highest concentrations (20.00 mg/m^3) occurring on March 16th, and the bloom declining by March 25th. During this bloom, TChla concentrations mirrored those of diatoms, with low concentrations of both measures (≤1.39 mg/m^3) were observed in the eastern portions of the track near the Fraser River plume. In addition, this spring bloom event showed increased concentrations of cryptophytes, dictyochophytes, green algae, and dinoflagellates, but the contributions (avg 3.18 mg/m^3) of these groups to total concentration remained relatively low compared to that of diatoms. The other groups did not show 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 36A**); however, there were other notable high TChla bloom events with low diatom contributions.

Beyond diatoms, raphidophytes also showed high concentration events. Specifically, increased raphidophyte concentrations were observed directly offshore of the Fraser River on July 24th (15.14 mg/m^3) and much higher concentrations on July 25th (25 mg/m^3) (**Figure 36B**). Many of the flagellate groups, including cryptophytes, green algae, dictyochophytes, haptophytes and dinoflagellates also showed increased concentrations during this observed raphidophyte bloom. Outside of the spring diatom and raphidophyte blooms, the non-blooming flagellate groups tended to show their highest contributions to total concentration during the summer. Cryptophytes and green algae, the dominant groups outside diatoms and raphidophytes, displayed similar trends throughout the transect, with the highest concentrations generally located close to the Fraser River-influenced portions of the transect. Also, regions with increased peaks in their concentration were observed across the transect multiple times during summer (see **Figure 36**). For example, the highest cryptophytes and

green algae concentrations were recorded on July 9th (3.18 mg/m^3, 3.12 mg/m^3, respectively). In turn, dictyochophytes, and dinoflagellates showed generally comparable spatial distributions with low concentrations outside high TChla events. Finally, haptophytes and cyanobacteria concentrations remained low (0.72, 0.86 mg/m^3, respectively) throughout the year.

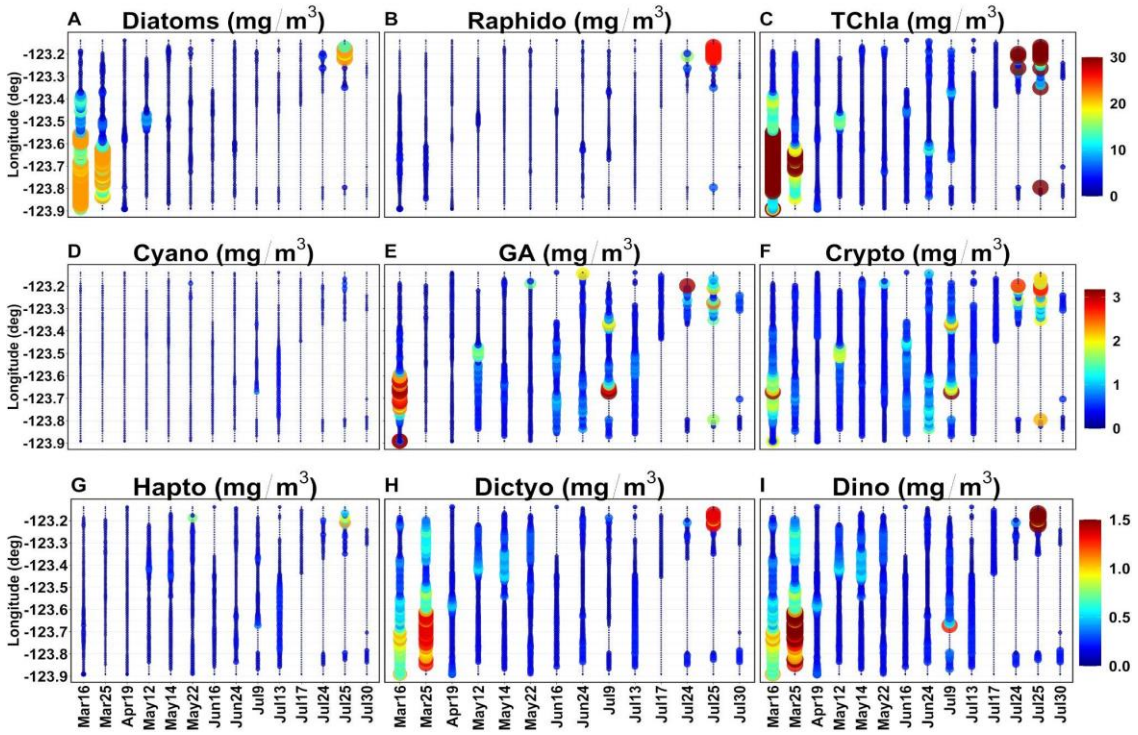

Figure 36. Spatial-temporal distribution of highly-resolved 2018 phytoplankton composition along the ferry track derived from the EOF-based hyperspectral algorithm. The ferry track is shown in **Figure 29**.

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 37**). In general, diatoms (1.05 mg/m^3) and cryptophytes (0.67 mg/m^3) were dominant during spring (**Figure 37**), while the low biomass flagellate groups increased their contributions in summer. Similar to 2018, sporadic and localized increased concentration events were observed in 2019, albeit at much lower levels. For instance, the highest TChla events were observed on June 7th and August 14th. On June 7th, diatoms were not prominent, and 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 (**Figure 38**). Interestingly, a spike in green algae (2.81 mg/m^3) and dinoflagellates 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,

dictyochophytes and haptophytes tended to show their highest concentrations in July and toward the end of August in the western and central portions of the transect. Other groups, such as cyanobacteria, remained invariant throughout the transect, with a maximum concentration of 0.21 mg/m^3 observed on June 7th.

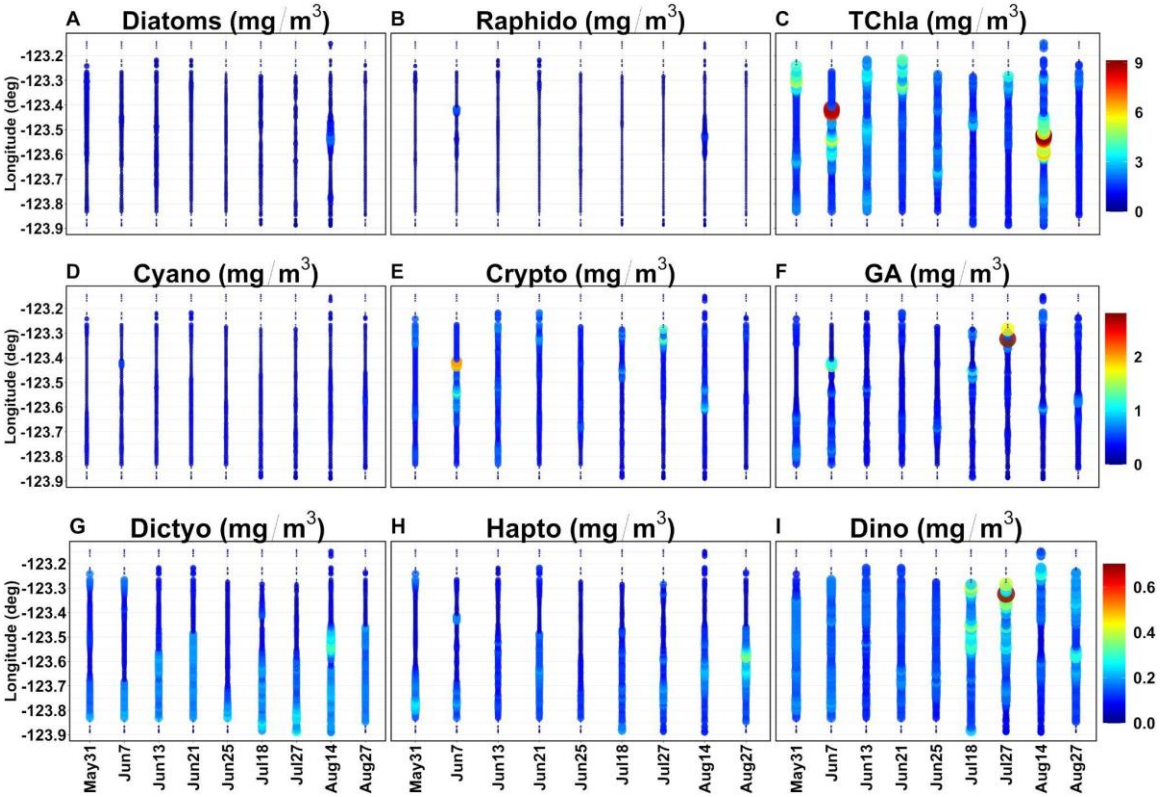

Figure 37. Spatial-temporal distribution of highly-resolved 2019 phytoplankton composition along the ferry track derived from the EOF-based hyperspectral algorithm. The ferry track is shown in **Figure 29**.

5.4.5 Environmental Drivers and EOF-derived Phytoplankton Composition

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 38**. During the spring bloom conditions observed on March 16th, 2018, diatoms showed high (> 50.00 mg/m^3) but variable concentrations along much of the western portions of the transect where SSS was relatively invariant (27 PSU, **Figure 38A**). Moving eastward, a slight drop in SSS (25 PSU) coincided with diatom concentrations decreasing by half (20.00 mg/m^3), but remaining high. Another sharp decline was observed in SSS towards the eastern portion of the transect, and diatom concentrations dropped to low levels (0.97 mg/m^3). In turn, August 14th, 2019 showed comparatively lower SSS across the transect, but variability in phytoplankton concentration and compositions also appeared to be driven by 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 westwards. Towards the central portions of the transect, a sharp decline in SSS was observed, coinciding with increased phytoplankton concentration largely driven by raphidophytes, diatoms, and cryptophytes. 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 minimum and stayed relatively low moving westwards under low, but variable SSS.

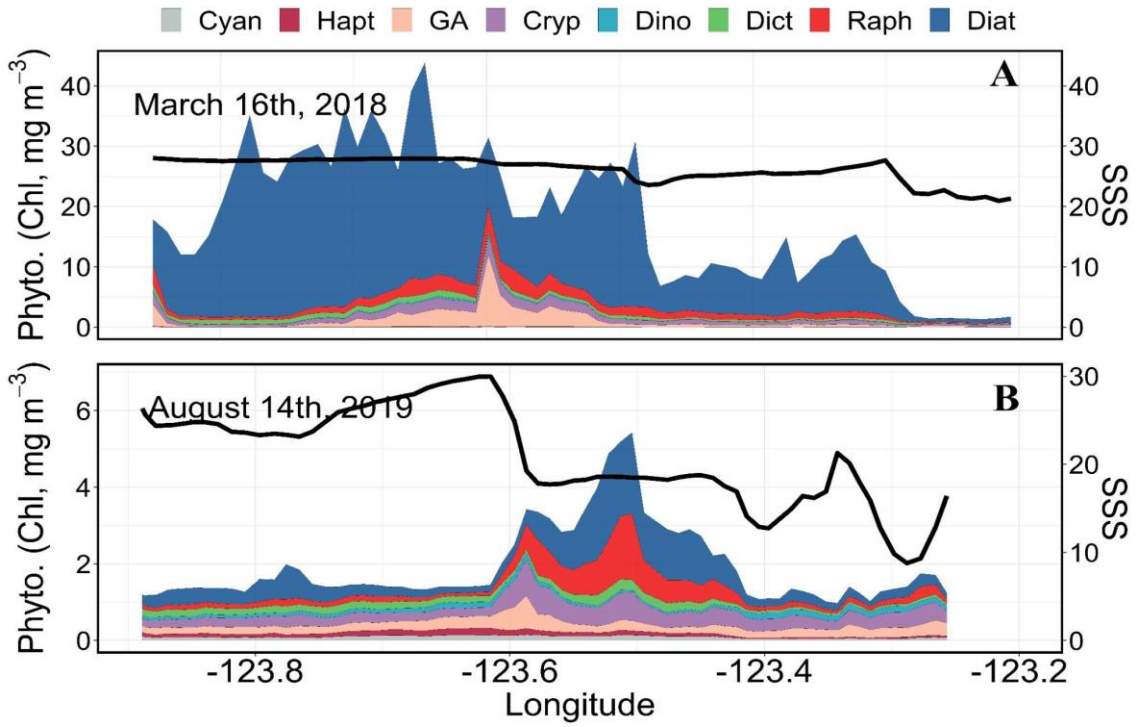

Figure 38. Transects of high spatial resolution EOF-derived phytoplankton composition for spring bloom conditions (March 16th, 2018) and summer raphidophyte bloom conditions (August 14th, 2019). For March 16th, 2018, the corresponding continuous ferrybox SSS is displayed on the right y-axis in A. Comparable plots are shown for August 14th, 2019 in B.

5.5 Discussion

The goal of this research was to derive TChla and phytoplankton community composition Chla contributions using an EOF-based hyperspectral algorithm, applied to above-water hyperspectral radiometry within the SoG on the west coast of Canada. The level-1 evaluation showed the best performance for TChla, diatoms, raphidophytes, green algae, and cryptophytes, moderate retrievals for haptophytes and finally poor performance for

dictyochophytes, cyanobacteria, and dinoflagellates. The level-2 validation using independent samples restricted to TChla, showed statistically significant retrievals and followed similar trends as the ferrybox CHL_{FL}, but with a large overestimation at the high Chla concentrations. The EOF-derived phytoplankton composition spatial maps demonstrated that diatoms dominated the spring bloom, particularly in western portions of the central SoG and notably in 2018. Following the spring bloom, diatom concentrations tapered off and by the summer were succeeded by a high diversity of small flagellate groups and a large raphidophyte bloom close to the Fraser River. The following sections provide a discussion of our main findings.

5.5.1 EOF-based Phytoplankton Group Retrievals From Hyperspectral Rrs

The regional algorithm evaluation demonstrated the potential of the EOF approach to derive phytoplankton group composition from hyperspectral Rrs under dynamic optically complex coastal conditions. Our results were in line with several studies that have used the multispectral-based EOF approach to derive phytoplankton cell abundance (Lange et al., 2020), accessory pigment concentration (Bracher et al., 2015), and Chla concentration of phytoplankton functional types (PFTs) from global ocean waters (Xi et al., 2020). Recently, covering Case-1 waters, Xi et al. (2020) utilized the same method on satellite-based multispectral Rrs and obtained good retrievals for diatoms (R^2cv = 0.63, MDPDcv = 74.74%), dinoflagellates (R^2cv = 0.59, MDPDcv = 57.29%), haptophytes (R^2cv = 0.61, MDPDcv = 48.62%) and green algae (R^2cv = 0.48, MDPDcv = 56.26%) and more recently, the authors slightly improved these statistics through the inclusion of SST into their regression model (Xi et al., 2021). For a similar type of Case-1 waters, Bracher et al. (2020), performed comparable methods on underwater hyperspectral Rrs data and showed good retrievals for diatoms (R^2cv = 0.68, MDPDcv = 53%), haptophytes (R^2cv = 0.57, MDPDcv = 44%), cyanobacteria (R^2cv = 0.54, MDPDcv =29%), green algae (R^2cv = 0.55, MDPDcv =48%), and dinoflagellates (R^2cv = 0.59, MDPDcv =43). The retrievals of these studies were better than the ones derived in this study; for instance, our retrieval for diatoms (R^2cv = 0.60, MDPDcv =68.10%) haptophytes (R^2cv = 0.32, MDPDcv = 49.56%), cyanobacteria (R^2cv = 0.20, MDPDcv =49.56%), green algae (R^2cv = 0.45, MDPDcv =42.48%), and dinoflagellates (R^2cv = 0, MDPDcv =58.94%) retrievals were relatively poorer. Again, here it is important to note that these studies used datasets from optically Case-1 waters.

Comparatively fewer studies have assessed these methods in Case-2 coastal waters (Vishnu et al., 2022), and even fewer using hyperspectral Rrs (e.g., Craig et al., 2012; Bracher et al.,

2020). Specifically, our hyperspectral-based EOF results produced improved phytoplankton composition retrievals (see **Table 9**) when compared to Vishnu et al. (2022). The authors used similar methods, but were limited to the spectral bands of the Sentinel-3A OLCI and the model evaluation was statistically significant only for diatoms and raphidophytes. When compared with Vishnu et al. (2022), the hyperspectral-EOF algorithms derived in this study showed highly significant (p<0.001) retrievals for all groups (see **Table 9**), with the best performance recorded for diatoms, raphidophytes, green algae, and cryptophytes, which were the dominant observed groups generally constituting 73% of the total phytoplankton community. However, similar to Vishnu et al. (2022) haptophytes, dictyochophytes, cyanobacteria, and dinoflagellates showed moderate to poor performance, likely as a result of their low contributions to the phytoplankton community making it difficult to separate the optical signals from these groups.

For TChla, our EOF-derived TChla model showed high comparability with the in situ HPLC-derived training data (R^2_{cv} = 0.73) with a moderate level of overestimation (9%). Additionally, independent validation between the continuous EOF-derived TChla and the ferrybox CHL_{FL} showed strong spatial-temporal comparability with the EOF retrievals capturing similar trends as the fluorometer; however, this comparison showed high Bias (90.77%) and MDPD (61.62%). The fact that the EOF-CHL_{FL} comparison showed such high Bias and the EOF-HPLC comparison showed considerably lower Bias suggests that the Chlorophyll-a fluorescence measurements may have been underestimating TChla. Of note, Chlorophyll-a fluorescence can be highly variable under differing nutrient, light and phytoplankton community conditions (Roesler et al., 2017). While the fluorometer was corrected for non-photochemical quenching (Traver-Smith et al., 2021), these corrections in themselves may have contained uncertainties and other factors such as variable fluorescence quantum yield between different phytoplankton communities may have added uncertainty to the EOF-CHL_{FL} relationship. These results suggest that the EOF outputs are more representative of TChla when compared to CHL_{FL}, thus highlighting the value of hyperspectral radiometry on ships of opportunity (Wang et al., 2019) and the potential of PACE hyperspectral sensor (Werdell et al., 2019).

Further, the TChla retrievals obtained in this study were generally better than those derived from regional satellite-based models. For instance, a regional study using standard atmospheric correction techniques (MUMM+SWIR) and the OC3M algorithm to derive

TChla from MODIS-Aqua imagery showed most comparable results to our EOF 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 and Costa, 2018). Similarly, poorer results when compared to the retrievals in this study were derived from recent analysis deriving TChla from Sentinel 3 OLCI imagery using the POLYMER algorithm (R = 0.51, Bias = 0.99, MDPD = 1.82) (Giannini et al., 2021), the Ocean Color Climate Change Initiative (OC-CCI) product (R^2 = 0.62, Bias = 0.78, MDAD = 1.44, RMSE = 0.26) and the GlobColor interpolated product (R^2 = 0.47, Bias = 0.96, MDAD = 1.57, RMSE = 0.30) (Pramlall et al., 2023). Therefore, the stronger TChla retrievals derived in this study, when compared to regional models done with multispectral satellite data, further support the use of hyperspectral Rrs for dynamic Case-2 waters, again supporting the significance of PACE hyperspectral OCI.

Despite the promising retrievals, the results of this study showed that even with hyperspectral Rrs, uncertainties arise under Case-2 conditions from the high contributions of other optical constituents such as CDOM, non-algal particle (NAP) absorption, and high backscattering from inorganic suspended sediments (IOCCG, 2000; Ruddick et al., 2006; Mouw and Yoder, 2010; Craig et al., 2012; Mouw et al., 2017; Aurin et al., 2018). Various optical constituents define the shape of reflectance signals and, consequently, impact the phytoplankton composition retrievals from spectral data. For example, our study reported high spatial-temporal variability in CDOM and TSM (see **Table 6**) across the SoG, and this variability largely defined the spectral shape of our hyperspectral Rrs (see **Figure 32**), and likely resulted in substantial contributions to the spectral loading of the first few EOF modes (see **Figure 33**). Similar to these results, 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 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 nm 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 pigment and seawater absorption.

The described high contributions of non-algal optical constituents to the spectral signature likely explains the poor retrieval performance of low biomass contribution groups. Poor retrievals were observed for cyanobacteria (~6% contribution), dictyochophytes (~6%), and dinoflagellates (~8%) and in the Salish Sea, these groups are overarchingly considered minor groups, rarely representing high contributions to total phytoplankton biomass (Del Bel Belluz et al., 2021; Nemcek et al., 2023). Consequently, the low spectral signals from these groups were likely difficult to separate from the dominant contributions of other phytoplankton groups and non-algal optical constituents (Mouw et al., 2017). Therefore, these results underscore the challenges associated with remote sensing retrievals in high optical complexity waters and that high concentrations of TSM and CDOM can contribute to significant uncertainties in retrievals of low biomass communities using spectral data.

Despite the optical complexity, increased spectral resolution can positively influence the number of groups retrieved from the remote sensing reflectance. For example, earlier studies have demonstrated 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; Werdell et al., 2018; Cael et al., 2020). This has been corroborated in our study by comparing results derived via hyperspectral methods here and those done on the same data using multispectral Sentinel-3A OLCI data (Vishnu et al., 2022). As reported earlier, hyperspectral Rrs offers more degrees of freedom to model various phytoplankton pigments and associate them with phytoplankton composition (Wolanin et al., 2016; Werdell et al., 2018; Cael et al., 2020). Yet, there are challenges associated with hyperspectral data; 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). Here, we used hyperspectral data acquired at 2 nm spectral resolution and 1 nm spectral intervals, which is higher but somewhat comparable to the resolution of the OCI onboard the upcoming PACE mission (with 5 nm spectral resolution) (Werdell et al., 2019). It is expected that there will be minimal loss of both qualitative and quantitative information when using hyperspectral Rrs with either 1 nm or 5 nm spectral resolution for deriving phytoplankton pigments (Kramer et al., 2022). As a result, the application of the algorithms developed in this study to the PACE OCI is possible, and may lead to improved phytoplankton composition retrievals at a global scale (Cael et al., 2020) compared to the same retrievals using existing multispectral sensors.

Finally, beyond optical complexity and spectral resolution, the quality and quantity of matchup points, are also critical factors in determining the accuracy of the phytoplankton group retrievals. For instance, 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, Vishnu et al. (2022) used the same EOF-based methods and employed 61-102 samples to retrieve PFTs from Sentinel 3 OLCI data on a local scale (Case-2) as opposed to the above global studies. In this study, the total number of matchup points 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 low dynamic range. For instance, dinoflagellates and dictyochophytes periodically form large blooms where they constitute a dominant proportion of the phytoplankton community (Del Bel Belluz et al., 2021) and the inclusion of these points in the model may help to improve the retrievals of these generally low contribution groups.

5.5.2 Seasonal and Spatial Patterns in Phytoplankton Composition

Data from the ferry-based autonomous hyperspectral radiometry has provided highly-resolved phytoplankton composition data that follow expected regional trends, including diatom-dominated spring blooms, succession to moderate concentration summer flagellate-dominated communities and, large summer *Heterosigma akashiwo* blooms. Noticeably, the hyperspectral retrievals showed blooming diatom concentrations (see **Figure 36**) along the ferry track in the spring of 2018. These findings were corroborated by the ferrybox-derived CHL_{FL} (~10 mg/m^3) and are in line with the spring bloom conditions commonly observed across temperate coastal regions and within this region at this time of the year (e.g., Allen and Wolfe, 2013; Del Bel Belluz et al., 2021; Esenkulova et al., 2021; Gianinni et al., 2021; Marchese et al., 2022; McLaskey et al., 2022; Suchy et al., 2022; McEwan et al., 2023; Nemcek et al., 2023; Vishnu et al., 2022; Pramlall et al., submitted). Multiple factors have been attributed to the development of the spring bloom in the central SoG, including stratification from freshwater discharge, SST, a positive/negative North Pacific Gyre

Oscillation (NPGO) phase, grazing and seasonal reductions in wind and clouds (Yin et al., 1997; Collins et al., 2009; Masson and Peña, 2009; Allen and Wolfe, 2013; Suchy et al., 2019). While the focus of this study was not to determine drivers of spring bloom timing, the high spatial resolution outputs provide valuable information on spring bloom dynamics. Specifically, the spring blooms in 2018 were focused on the western portions of the ferry track, with lower concentrations associated with lower SSS moving eastward, representing Fraser River waters. This spatial trend suggests that light limitation as a result of high inorganic turbidity from the Fraser River or nutrient limitation due to increased stratification may have driven the low concentration conditions (Stockner et al., 1979; Yin et al., 1997; Masson and Peña, 2009; Loos and Costa, 2010; Del Bel Belluz et al., 2021).

Following the 2018 spring blooms, our EOF-based retrievals showed strong raphidophyte blooms that typically occurred in the low SSS waters within the Fraser River plume. The strongest of these blooms was matched by increased HPLC TChla, ferrybox CHL_{FL}, and corroborated by multiple in situ studies showing high abundances of *Heterosigma akashiwo* in the vicinity of the Fraser River during that period (Pawlowicz et al., 2020; Esenkulova et al., 2021; Vishnu et al., 2022). This species has been found to have an affinity for low salinity, highly stratified waters with low nitrate and phosphate and high silicate concentrations (Esenkulova et al., 2021). Interestingly, our high-resolution transect plot from August 18th, 2018 (**Figure 38**) 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 fish-killing algae in the world (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 data, either in situ radiometry or satellite-based, may improve the understanding of their spatial-temporal distribution and environmental drivers, helping to mitigate the negative effects on aquacultured salmon and other commercially important fish species in this region.

In comparison to 2018, the spring data from 2019 did not show high diatom bloom conditions, likely as a result of the later sampling period missing the bloom. Following spring conditions this year (and in 2018 outside of the raphidophyte blooms), both EOF-derived diatom and TChla concentrations declined along the ferry track and high diversity flagellate-dominated conditions were observed. These results align with expected seasonal trends in the SoG with the succession to lower concentration flagellate-dominated conditions during summer (Del Bel Belluz et al., 2021; Nemcek et al., 2023). As reported in numerous studies, diatom concentrations typically decline following the spring bloom largely due to stratified nutrient-limiting surface conditions and increased zooplankton grazing (Peña et al., 2016; McEwan et al., 2023). Under these conditions, diatoms are replaced by high diversity conditions dominated by flagellates, including cryptophytes, green algae, dictyochophytes, haptophytes, and dinoflagellates (Harrison et al., 1983; Del Bel Belluz et al., 2021; McEwan et al., 2023; Nemcek et al., 2023). These species have high surface-to-volume ratios, and many display mixotrophic capacities, making them better suited to warm, low nutrient stratified surface waters when compared to diatoms.

5.6 Limitations of the Study

Uncertainties associated with the HPLC and CHEMTAX data are important to consider, as these datasets 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 associated with the data input into CHEMTAX, or with TChla, used in the model. However, CHEMTAX has inherent uncertainty, notably in regions with dynamic phytoplankton community compositions (Lewitus et al., 2005). Under these conditions, misclassifications generally arise from species having variable pigment profiles (i.e., *Phaeocystis pouchetii* have been misclassified regionally due to variable HF presence), having similar ratios to other groups (here, raphidophytes showed small misclassifications outside of bloom conditions due to similarity to diatoms and their marker, violaxanthin, shared with prasinophytes), concentrations from cryptic pico-sized species being attributed to the most similar CHEMTAX input group (i.e., pelagophytes and bolidophytes into dictyochophytes and diatoms, respectively), and the endosymbionts of mixotrophic species causing overestimations of the endosymbiont species (i.e., cryptophyte alloxanthin incorporation within *Mesodinium rubrum* (ciliate) and *Dinophysis* (dinoflagellate)). Furthermore, pigment collinearity can result in CHEMTAX error and arises from groups

following similar spatial-temporal trends within the environment (Catlett and Siegel, 2018; Kramer and Siegel, 2019; Kramer et al., 2020). Novel "data-driven" approaches have been developed using both cluster and EOF analysis to group species with similar pigment ratios to address pigment collinearity and some of the other misclassification issues. Outputs from these methods typically have fewer PFT groups (< 5) than CHEMTAX as they combine communities into broader classes such as "small flagellates". The use of these methods in this study may have improved retrievals, notably during diverse low TChla summer conditions. Yet, we decided to use CHEMTAX outputs as considerable work has been done in this region using CHEMTAX to define the dominant regional groups and their input ratios (Del Bel Belluz et al., 2021, Vishnu et al., 2022, Nemcek et al., 2023). Furthermore, the data-driven approaches would not have allowed for the separation of raphidophytes, a toxic species with considerable regional impacts on fish (Esenkulova et al., 2021). Based on these factors, the CHEMTAX method was the best approach to define the dominant phytoplankton groups. Still, outputs certainly contained uncertanties, which should be considered with further application of the models derived in this study.

Finally, uncertainties originating from the in situ hyperspectral Rrs must also be considered. Errors from spectral data obtained from above-water platforms occur from multiple sources, including instrument calibration, inherent sensor noise, ship shadow and superstructure influences, environmental factors, and parameterization of the sky-glint (Mobley, 1999). To address these issues, we followed rigorous quality control criteria, including factory pre-calibration of instruments, data acquisition only during solar noon and clear sky conditions, and methods to remove ship shadow and superstructure influences and the effects of sun glint (Mobley 1999, Hooker and Morel, 2003, Wang and Costa, 2018; Nasiha et al., 2022). Although characterization of sources of error is important to define uncertainties associated with the retrieved Rrs (Zibordi et al., 2012; Ruddick et al., 2019; Vabson et al., 2019a, 2019b), these errors are challenging to quantify due to the variable nature of the sources (e.g. Zibordi et al., 2012; Alikas et al., 2020; Tilstone et al., 2020). As such, here we consider uncertainty estimates from other studies focused on fiducial reference measurements (e.g., Zibordi et al., 2012; Alikas et al., 2020; Tilstone et al., 2020). For example, previous studies have reported that, on average, Rrs uncertainties are most significant at 443nm (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, recent studies have shown that environmental factors, including wind speed, sun zenith angle, sky glint, and wave

height, contribute 80% of the variability in Rrs across all wavelengths and the variability becomes worse under overcast conditions (Lin et al., 2022). Within the environmental factors, ρ_s contributes the most (50%) to Rrs variability (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 showed CV<10% across the visible bands. In addition, we conducted a BRDF correction for our above-water Rrs that were tailored explicitly for Case-2 water types (Lee et al., 2011; Wand and Costa, 2022). 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 error as much as possible.

5.7 Conclusion and Outlook

This study contributes to the small number of research conducted using hyperspectral Rrs to derive the Chla concentration of phytoplankton community composition under optically complex Case-2 waters. Here, highly-resolved hyperspectral Rrs 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 level-1 evaluation demonstrated the best retrievals for diatom, cryptophyte, green algae, raphidophyte 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 therefore small contribution toward total phytoplankton concentration. Despite the observed overestimation, the level-2 independent validation exhibited a good agreement between EOF-derived TChla and ferrybox-derived $CHL_{FL,}$ adding further confidence to our results.

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. An exception was the

observation of increased diatom contributions in the reduced SSS frontal zone at the boundary of the Fraser River plume. In turn, flagellates were generally more prevalent in the low SSS Fraser River plume influenced waters. Furthermore, the results of this study captured large harmful raphidophyte blooms (*Heterosigma akashiwo*), which showed expected trends and a strong association with the Fraser River plume. Beyond these general trends, the results of this study demonstrated the value of having highly resolved phytoplankton community composition data. For example, transects (**Figure 38)** showed high spatial variability and distinct features in phytoplankton group concentrations, likely driven by fronts, 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 insight, 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 results obtained from this study hold significant implications in the context of the upcoming NASA PACE mission, which carries a hyperspectral instrument with 5 nm spectral resolution, offering the potential for synoptic-scale monitoring and assessment of phytoplankton communities.

Chapter 6: Summary and Conclusion

6.1 Thesis overview

The main objective of this doctoral research is to characterize phytoplankton community composition using in situ and remote sensing techniques for Case-1 and Case-2 waters of the SNEP. In addition, we also evaluated the performance of POLYMER and Level-2 BAC derived Rrs and TChla concentration through matchup analysis using in situ data to determine which atmospheric correction algorithm provides better quality Rrs. We define four research objectives to achieve this goal.

Chapter 2 aimed to evaluate the performance of POLYMER and Level-2 BAC-derived Rrs and TChla through matchup analysis using in situ data from the Case-1 waters of the SNEP. Additionally, the reliability of satellite retrievals over a broader spatial and temporal scale has been investigated. Our initial validation of comparing the performance of Rrs and TChla derived through POLYMER and Level-2 BAC showed better performance for POLYMER. Furthermore, we also find a substantial loss of (~60%) valid data points when utilizing Level-2 BAC products due to their inefficiency in retrieving products under cloud conditions. Validation of our POLYMER-derived Rrs showed better radiometric performance in the blue, followed by green and inferior radiometric performance in the red band. Similarly, POLYMER successfully retrieved the regional TChla dynamic range and showed excellent agreement with in situ HPLC data. The weekly-binned surface TChla analysis showed low TChla and weak variability (0.3 to 0.7 mg/m^3) within the Alaska Gyre. However, elevated concentration ranges were observed towards the SEA and BC coast. Within this region, patches of increased TChla concentration were found towards 57°N and -135° W on a latitudinal and longitudinal domain, respectively. The result obtained from this objective provided the objective evidence to select POLYMER-derived Rrs as the spectral input in the EOF-based PFT algorithm in Chapter 3.

Chapter 3 utilizes the potential of the in situ HPLC pigment dataset and OLCI ocean color imagery to characterize phytoplankton community composition in the open ocean waters of the SNEP and investigates their environmental drivers. The satellite-based approach utilizes a previously established EOF-based algorithm (Bracher et al., 2015) trained using a matchup dataset of CHEMTAX and EOF scores derived from the collocated OLCI Rrs data. Evaluation of the algorithm showed reliable performance for Hapto/Pelago/Cyano, Diatoms/Dino/GA, and Crypto groups. Our CHEMTAX and satellite retrievals confirmed the

dominance of small flagellates such as haptophytes, pelagophytes, and green algae in the iron-poor open ocean waters. In contrast, diatoms were dominant in the iron-rich Haida Gwaii shelf. Our satellite spatial analysis over a latitudinal and longitudinal domain showed an invariant phytoplankton community within the iron-poor open ocean waters. A peak in cryptophyte concentration was recorded in OLCI imagery towards northern Alaska in 2019, consistent with the CHEMTAX and attributed to the intrusion of ACC and Sitka and Yakutat eddies. Similarly, increased Diatoms/Dino/GA concentrations observed towards the Haida Gwaii shelf with corresponding decreases in Hapto/Pelago/Cyano were also consistent with the CHEMTAX. Finally, our environmental analysis revealed that SSS and MLD positively drove open ocean groups such as haptophytes and green algae distribution. Meanwhile, diatoms, dominant in coastal/shelf waters, were negatively correlated with SSS and MLD.

Chapter 4 characterizes PFTs in the Case-2 waters of the west coast of Canada using CHEMTAX and the application of an EOF-based algorithm on daily OLCI imagery. Evaluation of the algorithm showed reliable performance only for diatoms and raphidophytes. Our OLCI-derived PFT distribution showed diatom-dominated seasonal spring and fall bloom with distinct spatial distribution patterns. During spring, diatoms showed peaks, with the highest Chla concentration in the central and northern SoG and the Juan de Fuca Strait. The lowest biomass groups, including cryptophytes and raphidophytes, mirrored the distribution of diatoms, while green algae exhibited an opposite trend, displaying a lower average concentration. During summer, we observed the largest raphidophyte bloom, mainly composed of *Heterosigma akashiwo*, located off of the southwest coast of Vancouver Island and close to the mouth of the Fraser River. This bloom has been validated against in situ HPLC pigment and microscopy data. Still, diatom concentration remains high during summer, with distinct peaks found off of the southwestern coast of Vancouver Island and central and northern SoG. Finally, by fall, diatom still constituted the highest concentration of total biomass, with peaks identified in central and northern SoG and the southwest coast of Vancouver Island. The PFT retrieved with limited multispectral data using an EOF-based approach was successful, and we proceeded to test the same EOF approach on high-resolution hyperspectral data, as discussed in Chapter 5, to explore its potential further.

Chapter 5, utilizing the same EOF-based algorithm built considering CHEMTAX-derived phytoplankton groups Chla concentration and EOF scores derived from hyperspectral Rrs, to characterize phytoplankton community composition within the central SoG. Our model

evaluation performed relatively well for all phytoplankton groups compared to the Sentinel 3 OLCI retrievals discussed in Chapter 4, which utilized the same HPLC pigment dataset. Overall, our EOF-based hyperspectral algorithm successfully predicted the seasonal diatom-dominated spring bloom, which is highly localized within the western portion of the transect, mostly associated with high salinity oceanic waters. Meanwhile, flagellate communities were prevalent across the region, with the highest average biomass recorded in the summer and they were found near low-saline Fraser River plume waters. Furthermore, our algorithm retrieved the largest *Heterosigma akashiwo* bloom in 2018, found close to the Fraser River plume associated with low salinity conditions. Additionally, our high-resolution transect plot highlighted peaks in raphidophyte, diatom and cryptophytes coinciding with a sharp decline in SSS suggesting a frontal region within the Fraser River plume. Therefore, the highly-resolved phytoplankton groups derived from the autonomous platform data have large potential in understanding harmful algal bloom dynamics, water quality monitoring, and fisheries management in this region.

Overall, our EOF-based algorithm evaluation for the Case-1 and Case-2 waters showed better retrieval only for dominant groups, with Case-1 waters generally resulting in improved statistics (i.e., MAE, Bias, and MDPD). However, retrievals from Case-2 waters were affected by other optical constituents such as absorption by CDOM and TSM scattering (Mouw et al., 2017), but they were comparable to the existing literature. The application of this EOF-based algorithm to daily and weekly composite OLCI imagery well predicted the surface dynamics of major groups within the west coast of Canada and in the HNLC region of SNEP. An improved retrieval was obtained for all phytoplankton groups when the EOF algorithm trained with hyperspectral Rrs obtained along a regularly traveled ship of opportunity ferry track in optically complex Case-2 waters of SoG. In turn, the application of this algorithm to hyperspectral Rrs data is able to derive highly resolved phytoplankton community composition data, which would not have been possible with discrete water samples.

6.2 Limitations and Recommendations for Future Work

Several limitations exist within this research, most of which have already been discussed in the previous chapters. First, limitations arose from the EOF-based regression model being a statistical algorithm rather than a mechanistic or IOPs-based algorithm. Here, the performance of our EOF-based algorithm ultimately depends on the training dataset (i.e., the

number of points and the representation of the dynamic range), and therefore, the application of this algorithm outside the range of its training dataset, including geographical region, may lead to uncertainties and unreliable retrievals, and should be avoided. For detecting long-term shifts in phytoplankton community composition in the global ocean, mechanistic or IOPs-based algorithms are more recommended since they are built on physical principles established using radiative transfer modeling, which minimizes the empiricism in the algorithm (Mouw et al., 2017).

Second, the HPLC pigment data are widely used to calibrate and develop satellite-based algorithms for deriving phytoplankton community composition that suffers from pigment collinearity and correlation issues (Kramer and Siegel, 2019). As a result, the phytoplankton taxonomic resolution level obtained from HPLC pigment data over a global scale is limited to four groups, while more number of taxonomic groups is detected on a local scale (Kramer and Siegel, 2019; Kramer et al., 2020). In turn, the CHEMTAX program also assumes that biomarker pigments are uncorrelated with each other (Mackey et al., 1996); however, recent observations showed that the concentration of these pigments is highly and significantly correlated (Catlett and Siegel, 2018; Kramer and Siegel, 2019; Kramer et al., 2020). Additionally, it assumes that pigment:TChla ratios for each taxonomic group remain the same across the spatial-temporal scale and among species within the same pigment class (Mackey et al., 1996; Higgins et al., 2011). However, they have been shown to vary substantially due to genetics and physiological differences, as well as between species of the same pigment class (Higgins et al., 2011). Therefore, efforts have been made to limit this error by grouping phytoplankton pigment data based on the same environmental condition when utilized in the CHEMTAX (Swan et al., 2016). Still, this approach does not entirely eliminate interspecific or intraspecific genetics or small-scale differences in physiological variations. Therefore, care must be taken to extend the taxonomic resolution offered by HPLC pigment data and be aware of the uncertainties associated with CHEMTAX data unless extensive validation has been performed. Pigment collinearity can be addressed by training satellite algorithms with phytoplankton composition data derived from other in situ methods, such as light-microscopy, flow cytometry (e.g., Sosik et al., 2010; Graff et al., 2012), quantitative cell imaging (IFCB; Olson and Sosik, 2007); however, the availability of this dataset is limited in space and time, not widely distributed as HPLC pigment data and the taxonomic size range offered by these techniques are limited to certain PSCs or group.

Our satellite-derived PFT products lack per-pixel uncertainty, a significant gap in utilizing ocean color radiometry to estimate PFTs (Bracher et al., 2017). To quantitatively estimate per-pixel uncertainty from ocean color imagery, it is essential to have quantitative estimates of the uncertainties associated with input parameters. In our case, uncertainties related to POLYMER-derived Rrs are currently unavailable, and, as a result, per-pixel uncertainty estimates from Sentinel 3 OLCI imagery remain undetermined. Efforts have been initiated to assess the uncertainties associated with POLYMER-derived Rrs. Consequently, future studies focusing on characterizing PFTs from ocean color imagery processed using the POLYMER AC scheme will be able to provide the per-pixel uncertainty of the output PFT product.

Another limitation of our work is the lack of utilization of the various physical and chemical environmental driver data. Here, we utilized several environmental driver data in the open ocean study but were limited to SST and SSS in the SoG. Analysis with various physical (i.e., SST, SSS, Light, MLD, Brunt–Väisälä frequency, Cloud cover, Wind and Freshwater discharge), chemical (i.e., Nitrate, Nitrite, Ammonia, Urea, Silicate, Phosphate, Dissolved iron), and biological (i.e., Zooplankton grazing) variables would deepen the understanding of the factors controlling the complex dynamics of phytoplankton community within the system that remained unexplained by the drivers included in this study.

6.3 Concluding Remarks

All existing research has predominantly utilized HPLC pigment data, CHEMTAX software, microscopy, and physical-biological models to investigate phytoplankton community composition in both Case-1 and Case-2 waters of the SNEP. While these studies offer a fundamental understanding of phytoplankton community composition, seasonality, and the factors influencing their regional environmental drivers, they are often limited to small geographical areas, failing to utilize the potential integration of both in situ and Earth Observation satellites, such as Sentinel 3A and 3B, and the upcoming PACE. Additionally, the utilization of high-resolution radiometry data from autonomous platforms has not been adequately explored. A comprehensive approach that combines these methods could enhance our ability to characterize phytoplankton composition and understand their spatial-temporal dynamics over a larger spatial-temporal scale and resolve with high-resolution capability, which motivates this doctoral thesis.

The testing of multiple objectives through various chapters in this doctoral thesis has provided a baseline for estimating phytoplankton community composition from in situ,

satellites and above-water radiometry across this highly dynamic region. Additionally, our study contributes to the broader literature on remote sensing of phytoplankton community composition across this region. One of the critical contributions of this study is the quantitative retrieval of dominant phytoplankton groups using Sentinel 3 OLCI imagery, high-resolved hyperspectral Rrs data within the Case-1 open ocean and optically complex Case-2 waters of SNEP. Our satellite-based retrievals from the west coast of Canada showed distinct spring and fall bloom peaks with diatoms dominating the community, and they are highly localized within the central, north, and off of Vancouver Island. Additionally, our algorithm captured the largest Raphidophyte bloom, predominantly composed of harmful algae, *Heterosigma akashiwo*. However, phytoplankton groups retrieved from the iron-poor open ocean waters of the GoA showed weak spatial-temporal variability, revealing a discernible gradient from the open ocean to the continental/coastal region. In addition, distinct physicochemical variables shape their dynamics within this region's iron-poor open ocean and continental shelf/coastal waters. Finally, phytoplankton groups obtained along the ferry track using highly-resolved hyperspectral data revealed diatom dominated spring bloom, notably localized in the western section of the transect that had higher SSS and that were least influenced by the Fraser River. A community succession from diatom-dominated spring bloom to flagellate-dominated summer conditions was observed in this region, with flagellates' biomass being highly variable within the transect, showing a tendency towards the low salinity Fraser River plume waters. Furthermore, the results of this study captured large harmful raphidophyte blooms (*Heterosigma akashiwo*), which showed expected trends showing association with the Fraser River plume. Beyond these trends, our transect analysis demonstrated the value of having highly resolved phytoplankton community composition data, with distinct features in phytoplankton group concentrations, likely driven by nutrient-rich waters, that would not have been captured using discrete samples. These features would not have been easily captured via in situ sampling and may provide valuable information on the drivers of phytoplankton community composition.

References

Aguilar-Islas, A.M., Séguret, M.J.M., Rember, R., Buck, K.N., Proctor, P., Mordy, C.W., Kachel, N.B., 2016. Temporal variability of reactive iron over the Gulf of Alaska shelf. Deep. Res. Part II Top. Stud. Oceanogr. 132, 90–106. https://doi.org/10. 1016/j.dsr2.2015.05.004.

Allen, S., Latornell, D., Olson, E., 2018. "Spring phytoplankton bloom timing, interannual summer productivity in the Strait of Georgia", in State of the Physical, Biological and Selected Fishery Resources of Pacific Canadian Marine Ecosystems in 2018, eds. J. L. Boldt, J. Leonard, and P. C. Chandler (Victoria: In Canadian Technical Report of Fisheries and Aquaculture Sciences. Fisheries and Oceans Canada Institute Ocean Sciences), 136–140.

Alexander, H., Rouco, M., Haley, S. T., Wilson, S. T., Karl, D. M., & Dyhrman, S. T. (2015). Functional group-specific traits drive phytoplankton dynamics in the oligotrophic ocean. Proceedings of the National Academy of Sciences, (12), E5972–E5979. https://doi.org/10.1073/pnas.1518165112.

Alikas, K., Ansko, I., Vabson, V., Ansper, A., Kangro, K., Uudeberg, K., & Ligi, M. (2020). Consistency of radiometric satellite data over lakes and coastal waters with local field measurements. Remote Sensing, 12(2), 1–34. https://doi.org/ 10.3390/rs12040616.

Alikas, K., Vabson, V., Ansko, I., Tilstone, G.H., Olmo, G.D., Nencioli, F., Vendt, R., Donlon, C., Casal, T., 2020. Comparison of Above-Water Seabird and TriOS Radiometers along an Atlantic Meridional Transect. Remote Sens. 12, 1–29.

Allen, S. E., and Wolfe, M. A. (2013). Hindcast of the timing of the spring phytoplankton bloom in the Strait of Georgia, 1968-2010. Prog. Oceanogr. 115, 6–13. doi: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. Res. Part I Oceanogr. Res. Pap. 52, 1989–2004. https://doi.org/10.1016/j.dsr.2005.06.015.

Amaya, D.J., Miller, A.J., Xie, S.P., Kosaka, Y., 2020. Physical drivers of the summer 2019 North Pacific marine heatwave. Nat. Commun. 11, 1–9. https://doi.org/ 10.1038/s41467-020-15820-w.

Antoine, D., Morel, A., 1999. A multiple scattering algorithm for atmospheric correction of remotely sensed ocean color (MERIS instrument): Principle and implementation for atmospheres carrying various aerosols including absorbing ones, International Journal of Remote Sensing. https://doi.org/10.1080/014311699212533.

Armbrecht, L. H., Wright, S. W., Petocz, P., and Armand, L. K. (2015). A new approach to testing the agreement of two phytoplankton quantification techniques: Microscopy and CHEMTAX. Limnol. Oceanogr. Methods 13, 425–437. doi:10.1002/lom3.10037.

Aurin, D., Mannino, A., Lary, D.J., 2018. Remote sensing of CDOM, CDOM spectral slope, and dissolved organic carbon in the global ocean. Appl. Sci. 8, 1–37. https://doi.org/10.3390/app8122687.

Bailey, S.W., Werdell, P.J., 2006. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. Remote Sens. Environ. 102, 12–23. https://doi.org/10.1016/j.rse.2006.01.015.

Banse, K., English, D.C., 1999. Comparing phytoplankton seasonality in the eastern and western subarctic Pacific and the western Bering Sea. Prog. Oceanogr. 43, 235–288. https://doi.org/10.1016/S0079-6611(99)00010-5.

Barnes, B. B., Cannizzaro, J. P., English, D. C., and Hu, C. (2019). Validation of VIIRS and MODIS reflectance data in coastal and oceanic waters: An assessment of methods. Remote Sens. Environ. 220, 110–123. doi:10.1016/j.rse.2018.10.034.

Batten, S. D., & Crawford, W. R. (2005). The influence of coastal origin eddies on oceanic plankton distributions in the eastern Gulf of Alaska. Deep Sea Research Part II: Topical Studies in Oceanography, 52, 991–1009. https://doi.org/10.1016/j.dsr 2.2005.02.009.

Beamish, R. J., Neville, C. E., Thomson, B. L., Harrison, P. J., and St John, M. (1994). A relationship between Fraser River discharge and interannual production of Pacific salmon (Oncorhynchus spp.) and Pacific herring (Clupea pallasi) in the Strait of Georgia. Can. J. Fish. Aquat. Sci. 51, 2843–2855. doi:10.1139/f94-283.

Beamish, R. J., Neville, C., Sweeting, R., and Lange, K. (2012). The synchronous failure of juvenile Pacific salmon and herring production in the Strait of Georgia in 2007 and the poor return of sockeye salmon to the Fraser River in 2009. Mar. Coast. Fish. 4, 403–414. doi:10.1080/19425120.2012. 676607.

Beamish, R.J., 2017. What the past tells us about the future of Pacific salmon research. Fish Fish. 18, 1161–1175. https://doi.org/10.1111/faf.12231.

Beardall, J., & Raven, J. A. (2004). The potential effects of global climate change on microalgal photosynthesis, growth and ecology. Phycologia, 43(1), 26–40. https://doi.org/10.2216/i0031-8884-43-1-26.1.

Beardall, J., & Stojkovic, S. (2006). Microalgae under global environmental change: Implications for growth and productivity, populations and trophic flow. Scienceasia, 1(s1), 1–10. https://doi.org/10.2306/scienceasia1513-1874.2006.32(s1).001.

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., et al. (2006). Climate-driven trends in contemporary ocean productivity. Nat. Lett. 444, 752–755. doi:10.1038/nature05317.

Ben Mustapha, Z., Alvain, S., Jamet, C., Loisel, H., Dessailly, D., 2014. Automatic classification of water-leaving radiance anomalies from global SeaWiFS imagery: Application to the detection of phytoplankton groups in open ocean waters. Remote Sens. Environ. 146, 97–112. https://doi.org/10.1016/j.rse.2013.08.046.

Bendriem, N., Roman, R., Gibson, D., & Sumaila, U. R. (2019). A review of the fate of southern British Columbia coho salmon over time. Fisheries Research, 218, 10–21. https://doi.org/10.1016/j.fishres.2019.04.002.

Bi, S., Li, Y., Wang, Q., Lyu, H., Liu, G., Zheng, Z., et al. (2018). Inland water Atmospheric Correction based on Turbidity Classification using OLCI and SLSTR synergistic observations. Remote Sensing, 10(6), 1–29. https://doi.org/ 10.3390/rs10071002.

Bishop, J.K.B., Davis, R.E., Sherman, J.T., 2002. Robotic observations of dust storm enhancement of carbon biomass in the North Pacific. Science (80). 298, 817–820. https://doi.org/10.1126/science.1074961.

Blanchet, F. G., Legendre, P., & Borcard, D. (2008). Forward selection of explanatory variables. Ecology, 89(9), 2623–2632. https://doi.org/10.1890/07-0986.1.

Booth, B.C., Lewin, J., Postel, J.R., 1993. Temporal variation in the structure of autotrophic and heterotrophic communities in the subarctic Pacific. Prog. Oceanogr. 32, 57–99. https://doi.org/10.1016/0079-6611(93)90009-3.

Bopp, L., Aumont, O., Cadule, P., Alvain, S., Gehlen, M., 2005. Response of diatoms distribution to global warming and potential implications: A global model study. Geophys. Res. Lett. 32, 1–4. https://doi.org/10.1029/2005GL023653.

Bopp, L., Resplandy, L., Orr, J.C., Doney, S.C., Dunne, J.P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J., Vichi, M., 2013. Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models. Biogeosciences 10, 6225–6245. https://doi.org/10.5194/bg-10-6225-2013.

Bouman, H.A., Ulloa, O., Scanlan, D.J., Zwirglmaier, K., Li, W.K.W., Platt, T., Stuart, V., Barlow, R., Leth, O., Clementson, L., Lutz, V., Fukasawa, M., Watanabe, S., Sathyendranath, S., 2006. Oceanographic basis of the global surface distribution of Prochlorococcus ecotypes. Science (80). 312, 918–921. https://doi.org/ 10.1126/science.1122692.

Boyce, D.G., Frank, K.T., Leggett, W.C., 2015. From mice to elephants: Overturning the "one size fits all" paradigm in marine plankton food chains. Ecol. Lett. 18, 504–515. https://doi.org/10.1111/ele.12434.

Boyd, P. W., Wong, C. S., Merrill, J., Whitney, F., Snow, J., Harrison, P. J., & Gower, J. (1998). Atmospheric iron supply and enhanced vertical carbon flux in the NE subarctic Pacific: Is there a connection? Global Biogeochemical Cycles, 12(3), 429–441. https://doi.org/10.1029/98GB00745.

Boyd, P., & Harrison, P. J. (1999). Phytoplankton dynamics in the NE subarctic Pacific. Deep-Sea Research Part II: Topical Studies in Oceanography, 46, 2405–2432. https://doi.org/10.1016/S0967-0645(99)00069-7.

Bracher, A., Bouman, H.A., Brewin, R.J.W., Bricaud, A., Brotas, V., Ciotti, A.M., Clementson, L., Devred, E., Di Cicco, A., Dutkiewicz, S., Hardman-Mountford, N.J., Hickman, A.E., Hieronymi, M., Hirata, T., Losa, S.N., Mouw, C.B., Organelli, E., Raitsos, D.E., Uitz, J., Vogt, M., Wolanin, A., 2017. Obtaining Phytoplankton Diversity from Ocean Color: A Scientific Roadmap for Future Development. Front. Mar. Sci. 4, 1–15. https://doi.org/10.3389/fmars.2017.00055.

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., R¨ottgers, R., Peeken, I., 2009. Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. Biogeosciences 6, 751–764. https://doi.org/https:// doi.org/ 10.5194/bg-6-751-2009.

Bracher, A., Xi, H., Dinter, T., Mangin, A., Strass, V., Appen, W. Von, Wiegmann, S., 2020. High Resolution Water Column Phytoplankton Composition Across the Atlantic Ocean From Ship-Towed Vertical Undulating Radiometry. Front. Mar. Sci. 7, 1–22. https://doi.org/10.3389/fmars.2020.00235.

Brando, V.E., Lovell, J.L., King, E.A., Boadle, D., Scott, R., Schroeder, T., 2016. The Potential of Autonomous Ship-Borne Hyperspectral Radiometers for the Validation of Ocean Color Radiometry Data. Remote Sens. 8, 1–18. https://doi.org/ 10.3390/rs8020150.

Brewin, R. J. W., Ciavatta, S., Sathyendranath, S., Jackson, T., Tilstone, G., Curran, K., et al. (2017). Uncertainty in ocean-color estimates of chlorophyll for phytoplankton groups. Frontiers in Marine Science, 4(104), 1–22. https://doi.org/10.3389/ fmars.2017.00104.

Brewin, R. J. W., Sathyendranath, S., Jackson, T., Barlow, R., Brotas, V., Airs, R., & Lamont, T. (2015). Influence of light in the mixed-layer on the parameters of a three-component model of phytoplankton size class. Remote Sensing of Environment, 168, 437–450. https://doi.org/10.1016/j.rse.2015.07.004.

Brewin, R.J.W., Dall'Olmo, G., Pardo, S., van Dongen-Vogels, V., Boss, E.S., 2016. Underway spectrophotometry along the Atlantic Meridional Transect reveals high performance in satellite chlorophyll retrievals. Remote Sens. Environ. 183, 82–97. https://doi.org/10.1016/j.rse.2016.05.005.

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. Ecol. Modell. 221, 1472–1483. https://doi.org/10.1016/ j.ecolmodel.2010.02.014.

Bricaud, A., Claustre, H., Ras, J., Oubelkheir, K., 2004. Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations. J. Geophys. Res. Ocean. 109, 1–12. https://doi.org/ 10.1029/2004JC002419.

Brickley, P.J., Thomas, A.C., 2004. Satellite-measured seasonal and inter-annual chlorophyll variability in the Northeast Pacific and Coastal Gulf of Alaska. Deep. Res. Part II Top. Stud. Oceanogr. 51, 229–245. https://doi.org/10.1016/j.dsr2.2003.06.003.

Brito, A. C., Sá, C., Mendes, C. R., Brand, T., Dias, A. M., Brotas, V., & Davidson, K. (2015). Structure of late summer phytoplankton community in the Firth of Lorn (Scotland) using microscopy and HPLC-CHEMTAX. Estuarine, Coastal and Shelf Science, 167, 86–101. https://doi.org/10.1016/j.ecss.2015.07.006.

Brotas, V., Brewin, R. J. W., Sá, C., Brito, A. C., Silva, A., Mendes, C. R., et al. (2013). Deriving phytoplankton size classes from satellite data: Validation along a trophic gradient in the eastern Atlantic Ocean. Remote Sens. Environ. 134, 66–77. doi:10.1016/j.rse.2013.02.013.

Budge, S.M., Wang, S.W., Ormseth, O.A., Rand, K.M., 2022. Foraging ecology of nearshore fishes in the Gulf of Alaska. Deep. Res. Part II Top. Stud. Oceanogr. 195, 1–13. https://doi.org/10.1016/j.dsr2.2021.105013.

Burger, J., Gochfeld, M., Niles, L., Powers, C., Brown, K., Clarke, J., et al. (2015). Complexity of bioindicator selection for ecological, human, and cultural health: Chinook salmon and red knot as case studies. Environmental Monitoring and Assessment, 187(3), 102. https://doi.org/10.1007/s10661-014-4233-4.

Cabré, A., Marinov, I., Leung, S., 2015. Consistent global responses of marine ecosystems to future climate change across the IPCC AR5 earth system models. Clim. Dyn. 45, 1253–1280. https://doi.org/10.1007/s00382-014-2374-3.

Cael, B.B., Chase, A., Boss, E., 2020. Information content of absorption spectra and implications for ocean color inversion. Appl. Opt. 59, 3971–3984. https://doi.org/ 10.1364/ao.389189.

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 Sens. 9, 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, 246–264. https://doi.org/10.1002/2017JC013195.

Chase, A. P., Boss, E. S., Haëntjens, N., Culhane, E., Roesler, C., & Karp-Boss, L. (2022). Plankton imagery data inform satellite-based estimates of diatom carbon. Geophysical Research Letters, 49(13), 1–11. https://doi.org/10.1029/2022GL098076.

Chase, A. P., Boss, E., Cetinic, I., and Slade, W. (2017). Estimation of phytoplankton accessory pigments from hyperspectral reflectance spectra: Toward a global algorithm. J. Geophys. Res. Ocean. Res. 122, 9725–9743. doi:10.1002/2017JC012859.

Chase, A. P., Kramer, S. J., Haëntjens, N., Boss, E. S., Karp-Boss, L., Edmondson, M., et al. (2020). Evaluation of diagnostic pigments to estimate phytoplankton size classes. Limnol. Oceanogr. Methods 18, 570–584. doi:10.1002/lom3.10385.

Chase, A., Boss, E., Zaneveld, R., Bricaud, A., Claustre, H., Ras, J., Dall, G., Westberry, T.K., 2013. Decomposition of in situ particulate absorption spectra. Methods Oceanogr. 7, 110–124. https://doi.org/10.1016/j.mio.2014.02.002.

Chassot, E., Bonhommeau, S., Dulvy, N. K., Melin, F., Watson, R., Gascuel, D., et al. (2010). Global marine primary production constrains fisheries catches. Ecol. Lett. 13, 495–505. doi:10.1111/j.1461-0248.2010.01443.x.

Chavez, F. P., Michisaki, R. P., Friederich, G. E., Pennington, J. T., Schlining, B., Fayos, C., et al. (2000). A ten-year time series from Monterey Bay, California: Seasonal, interannual and long-term patterns. Poster, USA.

Chavez, F. P., Ryan, J., Lluch-Cota, S. E., & Ñiquen, C. M. (2003). From anchovies to sardines and back: Multidecadal change in the Pacific Ocean. Science, 299(5604), 217–221. https://doi.org/10.1126/science.1075880.

Childers, A.R., Whitledge, T.E., Stockwell, D. A., 2005. Seasonal and interannual variability in the distribution of nutrients and chlorophyll a across the Gulf of Alaska shelf: 1998-2000. Deep. Res. Part II Top. Stud. Oceanogr. 52, 193–216. https://doi.org/10.1016/j.dsr2.2004.09.018.

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

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. Limnol. Oceanogr. Methods 4, 237–253. https://doi.org/10.4319/lom.2006.4.237.

Claustre, H., Hooker, S. B., Van Heukelem, L., Berthon, J.-F., Barlow, R., Ras, J., et al. (2004). An intercomparison of HPLC phytoplankton pigment methods using in situ samples: application to remote sensing and database activities. Marine Chemistry, 85, 41–61. https://doi.org/10.1016/j.marchem.2003.09.002.

Collins, A. K., Allen, S. E., and Pawlowicz, R. (2009). The role of wind in determining the timing of the spring bloom in the Strait of Georgia. Can. J. Fish. Aquat. Sci. 66, 1597–1616. doi:10.1139/F09-071.

Correa-Ramirez, M., Morales, C.E., Letelier, R., Anabalón, V., Hormazabal, S., 2018. Improving the remote sensing retrieval of phytoplankton functional types (PFT) using empirical orthogonal functions: A case study in a coastal upwelling region. Remote Sens. 10, 1–25. https://doi.org/10.3390/rs10040498.

Costalago, D., Forster, I., Nemcek, N., Neville, C., Perry, R. I., Young, K., et al. (2020). Seasonal and spatial dynamics of the planktonic trophic biomarkers in the Strait of Georgia (northeast Pacific) and implications for fish. Sci. Rep. 10, 1–12. doi:10.1038/s41598-020-65557-1.

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. Appl. Opt. 45, 5414–5425. https://doi.org/10.1364/AO.45.005414.

Crawford, W. R. (2005). Heat and fresh water transport by eddies into the Gulf of Alaska. Deep-Sea Research Part II:, 52, 893–908. https://doi.org/10.1016/ j.dsr2.2005.02.003.

Cullen, J.T., Chong, M., Ianson, D., 2009. British Columbia continental shelf as a source of dissolved iron to the subarctic northeast Pacific Ocean. Global Biogeochem. Cycles 23, 1–12.

https://doi.org/10.1029/2008GB003326.

Cury, P., Shin, Y., Planque, B., Durant, J., Fromentin, J., Kramerschadt, S., et al. (2008). Ecosystem oceanography for global change in fisheries. Trends in Ecology & Evolution, 23(6), 338–346. https://doi.org/10.1016/j.tree.2008.02.005.

Cushing, D. H. (1990). Plankton production and year-class strength in fish populations: An update of the match/mismatch hypothesis. Adv. Mar. Biol. 26, 249–293. doi:10.1016/S0065-2881(08)60202-3.

Dagg, M., Strom, S., Liu, H., 2009. High feeding rates on large particles by Neocalanus flemingeri and N. plumchrus, and consequences for phytoplankton community structure in the subarctic Pacific Ocean. Deep. Res. Part I 56, 716–726. https://doi.org/10.1016/j.dsr.2008.12.012.

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. Estuar. Coasts 44, 1419–1439. https://doi.org/https://doi.org/10.1007/ s12237-020-00858-2 Phytoplankton.

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. J. Geophys. Res. Ocean. 111, 1–11. https://doi.org/10.1029/ 2005JC002880.

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

Di Lorenzo, E., Mantua, N., 2016. Multi-year persistence of the 2014/15 North Pacific marine heatwave. Nat. Clim. Chang. 6, 1042–1047. https://doi.org/ 10.1038/ nclimate3082.

Di Lorenzo, E., Schneider, N., Cobb, K. M., Franks, P. J. S., Chhak, K., Miller, A. J., et al. (2008). North Pacific Gyre Oscillation links ocean climate and ecosystem change. Geophysical Research Letters, 35(8), 1–6. https://doi.org/10.1029/2007GL032838.

Dierssen, H., Bracher, A., Brando, V., Loisel, H., Ruddick, K., 2020. Data needs for hyperspectral detection of algal diversity across the globe.

Dierssen, H.M., Ackleson, S.G., Joyce, K.E., Hestir, E.L., Castagna, A., Lavender, S., Mcmanus, MA, 2021. Living up to the Hype of Hyperspectral Aquatic Remote Sensing : Science , Resources and Outlook. Front. Environ. Sci. 9, 1–26. https://doi.org/10.3389/fenvs.2021.649528.

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M., Féménias, P., Frerick, J., et al. (2012). The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. Remote Sensing of Environment, 120, 37–57. https://doi.org/10.1016/ j.rse.2011.07.024.

Dosser, H. V., Waterman, S., Jackson, J.M., Hannah, C.G., Evans, W., Hunt, B.P.V., 2021. Stark Physical and Biogeochemical Differences and Implications for Ecosystem Stressors in the Northeast Pacific Coastal Ocean. J. Geophys. Res. Ocean. 126, 1–18. https://doi.org/10.1029/2020JC017033.

Dragoo, D. E., Renner, H. M., & Kaler, R. S. A. (2017). Breeding status, population trends and diets of seabirds in Alaska, 2016. U.S. Fish and Wildlife Service Report AMNWR 2017/06. Homer, Alaska. https://doi.org/10.2307/1521664.

Dutkiewicz, S., Cermeno, P., Jahn, O., Follows, M.J., Hickman, A.A., Taniguchi, D.A.A., Ward, B.A., 2020. Dimensions of marine phytoplankton diversity. Biogeosciences 17, 609–634. https://doi.org/10.5194/bg-17-609-2020.

Dutkiewicz, S., Scott, J.R., Follows, M.J., 2013. Winners and losers: Ecological and biogeochemical changes in a warming ocean. Global Biogeochem. Cycles 27, 463–477. https://doi.org/10.1002/gbc.20042.

Dutkiewicz, S., Ward, B.A., Scott, J.R., Follows, M.J., 2014. Understanding predicted shifts in diazotroph biogeography using resource competition theory. Biogeosciences 11, 5445–5461. https://doi.org/10.5194/bg-11-5445-2014.

El Hourany, R., Abboud-Abi Saab, M., Faour, G., Aumont, O., Crépon, M., and Thiria, S. (2019). Estimation of secondary phytoplankton pigments from satellite observations using Self-Organizing Maps (SOMs). J. Geophys. Res. Ocean. 124, 1357–1378. doi:10.1029/2018JC014450.

Endo, H., Ogata, H., & Suzuki, K. (2018). Contrasting biogeography and diversity patterns between diatoms and haptophytes in the central Pacific Ocean. Scientific Reports, 8(1), 1–13. https://doi.org/10.1038/s41598-018-29039-9.

Esenkulova, S., Suchy, K. D., Pawlowicz, R., and Costa, M. (2021). Harmful algae and oceanographic conditions in the Strait of Georgia, Canada based on citizen science monitoring. Front. Mar. Sci. 8: 725092. 1–23. doi:10.3389/fmars.2021.725092.

EUMETSAT, 2017. Sentinel-3 OLCI Marine User Handbook. In: REF: SOLVO/EUM/16/VCA/D8. ISSUE: 1.3. July 2017. Darmstadt, Germany.

EUMETSAT, 2018. Sentinel-3 OLCI Marine User Handbook 2018, Avaliable onlie at https://earth.esa.int/eogateway/documents/20142/1564943/Sentinel-3-OLCI-Marine-User-Handbook.pdf.

Evans, C., Brussaard, C.P.D., 2012. Viral lysis and microzooplankton grazing of phytoplankton throughout the Southern Ocean. Limnol. Oceanogr. 57, 1826–1837. https://doi.org/10.4319/lo.2012.57.6.1826.

Evans, G.T., Parslow, J.S., 1985. A model of annual plankton cycles. Biol. Oceanogr. 3, 327–347. https://doi.org/10.1016/0198-0254(85)92902-4.

Evans, W., Pocock, K., Hare, A., Weekes, C., Hales, B., Jackson, J., et al. (2019). Marine CO2 patterns in the Northern Salish Sea. Front. Mar. Sci. 5, 1–18. doi:10.3389/fmars.2018.00536.

Falkowski, P. G., Barber, R. T., and Smetacek, V. (1998). Biogeochemical controls and feedbacks on ocean primary production. Science 281, 200–206. doi:10.1126/science. 281.5374.200.

Falkowski, P., 2012. The power of plankton. Nature 483, S17–S20. https://doi.org/https:// doi.org /10.1038/483S17a.

Falkowski, P.G., Oliver, M.J., 2007. Mix and match: How climate selects phytoplankton. Nat. Rev. Microbiol. 5, 813–819. https://doi.org/10.1038/nrmicro1751.

Favorite, F., Dodimead, A.J., Nasu, K., 1976. Oceanography of the Subarctic Pacific Region. Bull. Int. North Pac. Fish. Comm 33, 1–87.

Ferrari, G. M., Dowell, M. D., Grossi, S., and 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. Mar. Chem. 55, 299–316. doi:10.1016/S0304-4203(96)00061-8.

Ferreira, A., Sá, C., Silva, N., Beltrán, C., Dias, A. M., and Brito, A. C. (2020). Phytoplankton community dynamics in a coastal bay under upwelling influence (Central Chile). Estuar. Coast. Shelf Sci. 245, 1–15. doi:10.1016/j.ecss.2020.106968.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., Falkowski, P., 1998. Primary Production of the Biosphere : Integrating Terrestrial and Oceanic Components. Science (80). 281, 237–240. https://doi.org/DOI: 10.1126/science.281.5374.237.

Finkel, Z. V., Beardall, J., Flynn, K.J., Quigg, A., Rees, T.A. V., Raven, J.A., 2010. Phytoplankton in a changing world: Cell size and elemental stoichiometry. J. Plankton Res. 32, 119–137. https://doi.org/10.1093/plankt/fbp098.

Foreman, M. G. G., Callendar, W., MacFadyen, A., Hickey, B. M., Thomson, R. E., and Di Lorenzo, E. (2008). Modeling the generation of the Juan de Fuca eddy. J. Geophys. Res. Ocean. 113, 1–18. doi:10.1029/2006JC004082.

Friedland, K.D., Moisan, J.R., Maureaud, A.A., Brady, D.C., Davies, A.J., Bograd, S.J., Watson, R.A., Rousseau, Y., 2021. Trends in phytoplankton communities within large marine ecosystems diverge from the global ocean. Can. J. Fish. Aquat. Sci. 00, 1–12. https://doi.org/10.1139/cjfas-2020-0423.

Frost, B.W., 1991. The role of grazing in nutrient-rich areas of the open sea. Limnol. Oceanogr. 36, 1616–1630. https://doi.org/10.4319/lo.1991.36.8.1616.

Fujiki, T., Matsumoto, K., Honda, M.C., Kawakami, H., Watanabe, S., 2009. Phytoplankton composition in the subarctic North Pacific during autumn 2005. J. Plankton Res. 31, 179–191. https://doi.org/10.1093/plankt/fbn108.

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

Gilerson, A., Herrera-Estrella, E., Foster, R., Agagliate, J., Hu, C., Ibrahim, A., Franz, B., 2022. Determining the Primary Sources of Uncertainty in Retrieval of Marine Remote

Sensing Reflectance From Satellite Ocean Color Sensors. Front. Remote Sens. 3, 1–19. https://doi.org/10.3389/frsen.2022.857530.

Gittings, J.A., Brewin, R.J.W., Raitsos, D.E., Kheireddine, M., Ouhssain, M., Jones, B.H., Hoteit, I., 2019. Remotely sensing phytoplankton size structure in the Red Sea. Remote Sens. Environ. 234, 1–13. https://doi.org/10.1016/j.rse.2019.111387.

Goela, P. C., Danchenko, S., Icely, J. D., Lubian, L. M., Cristina, S., & Newton, A. (2014). Using CHEMTAX to evaluate seasonal and interannual dynamics of the phytoplankton community off the south-west coast of Portugal. Estuarine, Coastal and Shelf Science, 151, 112–123. https://doi.org/10.1016/j.ecss.2014.10.001.

Gossn, J. I., Ruddick, K. G., and Dogliotti, A. I. (2019). Atmospheric correction of OLCI imagery over extremely turbid waters based on the red, NIR and 1016 nm bands and a new baseline residual technique. Remote Sens. 11, 1–24. doi:10.3390/rs11030220.

Gower, J., and King, S. (2018). "Satellite observations of B.C. waters," in State of the physical, biological and selected fishery resources of Pacific Canadian marine ecosystems in 2018., eds. J. L. Boldt, J. Leonard, and P. C. Chandler (Victoria: In Canadian Technical Report of Fisheries and Aquaculture Sciences. Fisheries and Oceans Canada Institute Ocean Sciences), 50–53.

Grasshoff, K., Kremling, K., & Ehrhardt, M. (1999). Methods of Seawater Analysis. (K. Grasshoff, K. Kremling, & M. Ehrhardt, Eds.) (Third). New York: WILEY-VCH.

Gregg, W. W., & Casey, N. W. (2007). Modeling coccolithophores in the global oceans. Deep Sea Research Part II: Topical Studies in Oceanography, 54, 447–477. https://doi.org/10.1016/j.dsr2.2006.12.007.

Guidi, L., Stemmann, L., Jackson, G.A., Ibanez, F., Claustre, H., Legendre, L., Picheral, M., Gorsky, G., 2009. Effects of phytoplankton community on production, size and export of large aggregates : A world-ocean analysis. Limnol. Oceanogr. 54, 1951–1963. https://doi.org/https://doi.org/10.4319/lo.2009.54.6.1951.

Haigh, N., and Esenkulova, S. (2014). "Economic losses to the British Columbia salmon aquaculture industry due to harmful algal blooms 2009–2012," in Proceedings of the workshop on economic impacts of harmful algal blooms on fisheries and aquaculture, vol. 47

.Eds. V. L. Trainer and T. Yoshida (Sidney, British Columbia, Canada: PIECES Scientific Report) pp. 2–6.

Haigh, R., and Taylor, E. J. R. (1991). Mosaicism of microplankton communities in the northern Strait of Georgia, British Columbia. Mar. Biol. 110, 301–314. https://doi.org/10.1007/BF01313717.

Hallegraeff, G.M., Anderson, D.M., Cembella, A.D., 2004. Manual on HarmfulMarineMicroalgae, Monographs on Oceanographic Methodology. Paris: UNESCO.

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

Halverson, M., and Pawlowicz, R. (2016). Tide, wind, and river forcing of the surface currents in the Fraser River plume. Atmos. - Ocean 54, 131–152. doi:10.1080/07055900.2016.1138927.

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. Cont. Shelf Res. 59, 52–64. https://doi.org/10.1016/j.csr.2013.04.010.

Hamme, R.C., Webley, P.W., Crawford, W.R., Whitney, F.A., Degrandpre, M.D., Emerson, S.R., Eriksen, C.C., Giesbrecht, K.E., Gower, J.F.R., Kavanaugh, M.T., Pea, M.A., Sabine, C.L., Batten, S.D., Coogan, L.A., Grundle, D.S., Lockwood, D., 2010. Volcanic ash fuels anomalous plankton bloom in subarctic northeast Pacific. Geophys. Res. Lett. 37, 1–5. https://doi.org/10.1029/2010GL044629.

Harris, S.L., 2001. Size-Fractionated Chlorophyll and Primary Productivity and Nutrient Distributions off The West Coast Of Vancouver Island. The University of British Columbia.

Harris, S.L., Varela, D.E., Whitney, F.W., Harrison, P.J., 2009. Nutrient and phytoplankton dynamics off the west coast of Vancouver Island during the 1997/98 ENSO event. Deep. Res. Part II Top. Stud. Oceanogr. 56, 2487–2502. https://doi.org/10.1016/j.dsr2.2009.02.009.

Harrison, P.J., Boyd, P.W., Varela, D.E., Takeda, S., Shiomoto, A., Odate, T., 1999. Comparison of factors controlling phytoplankton productivity in the NE and NW subarctic Pacific gyres. Prog. Oceanogr. 43, 205–234. https://doi.org/10.1016/S0079-6611(99)00015-4.

Harrison, P.J., Fulton, J.D., Taylor, F.J.R., Parssons, T.R., 1983. Review of the Biological Oceanography of the Strait of Georgia: Pelagic Environment. Can. J. Fish. Aquat. Sci. 40, 1064–1094. https://doi.org/https://doi.org/10.1139/f83-129.

Harrison, P.J., Whitney, F.A., Tsuda, A., Saito, H., Tadokoro, K., 2004. Nutrient and plankton dynamics in the NE and NW Gyres of the subarctic Pacific Ocean. J. Oceanogr. 60, 93–117. https://doi.org/10.1023/B:JOCE.0000038321.57391.2a.

Henson, S.A., 2007. Water column stability and spring bloom dynamics in the Gulf of Alaska. J. Mar. Res. 65, 715–736. https://doi.org/10.1016/b978-0-408-70700-8.50018-5.

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

Hickey, B. M., and Banas, N. S. (2008). Why is the northern end of the California Current System so productive? Oceanography 21, 90–107. doi:10.5670/oceanog.2008.07.

Hieronymi, M. (2019). Spectral band adaptation of ocean color sensors for applicability of the multi-water biogeo-optical algorithm ONNS. Opt. Express 27, A707–A724. doi:10.1364/oe.27.00a707.

Higgins, H. W., Wright, S. W., & Schlüter, L. (2011). Quantitative interpretation of chemotaxonomic pigment data. In S. Roy, C. A. Llewellyn, E. S. Egeland, & G. Johnsen (Eds.), Phytoplankton Pigments: Characterization, Chemotaxonomy and Applications in Oceanography (pp. 257–313). Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511732263.010.

Hijmans, R.J., 2020. raster: Geographic Data Analysis and Modeling. Retrieved from https://CRAN.R-project.org/package=raster.

Hirata, T., Hardman-Mountford, N.J., Brewin, R.J.W., Aiken, J., Barlow, R., Suzuki, K., Isada, T., Howell, E., Hashioka, T., Noguchi-Aita, M., Yamanaka, Y., 2011. Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. Biogeosciences 8, 311–327. https://doi.org/10.5194/ bg-8-311-2011.

Hobday, A.J., Pecl, G.T., 2014. Identification of global marine hotspots: Sentinels for change and vanguards for adaptation action. Rev. Fish Biol. Fish. 24, 415–425. https://doi.org/10.1007/s11160-013-9326-6.

Hood, R. R., Laws, E. A., Armstrong, R. A., Bates, N. R., Brown, C. W., Carlson, C. A., et al. (2006). Pelagic functional group modeling: Progress, challenges and prospects. Deep-Sea Research Part II: Topical Studies in Oceanography, 53, 459–512. https://doi.org/10.1016/j.dsr2.2006.01.025.

Hooker, S.B., Morel, A., 2003. Platform and Environmental Effects on Above-Water Determinations of Water-Leaving Radiances. Am. Meteorol. Soc. 20, 187–205. https://doi.org/https://doi.org/10.1175/1520-0426(2003)020<0187:PAEEOA>2.0.CO;2.

Hu, C., Feng, L., Lee, Z., 2013. Uncertainties of SeaWiFS and MODIS remote sensing reflectance: Implications from clear water measurements. Remote Sens. Environ. 133, 168–182. https://doi.org/10.1016/j.rse.2013.02.012.

IOCCG (2000). Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters. In S. Sathyendranath (Eds.), Eds., Reports of the international ocean color coordinating group (No. 3). IOCCG.

IOCCG (2014). Phytoplankton functional types from space. In S. Sathyendranath (Eds.), Eds., Reports of the international ocean color coordinating group (No. 15). IOCCG. http://dx.doi.org/10.25607/OBP-106.

IOCCG. (2019). Uncertainties in ocean color remote sensing, In F. Melin (Eds.), Eds., Reports of the International Ocean Color Coordinating Group (No. 18). IOCCG.

Isada, T., Hirawake, T., Kobayashi, T., Nosaka, Y., Natsuike, M., Imani, I., Suzuki, K., Saitoch, S.-I., 2015. Hyperspectral optical discrimination of phytoplankton community structure in Funka Bay and its implications for ocean color remote sensing of diatoms. Remote Sens. Environ. 159, 134–151. https://doi.org/10.1016/j.rse.2014.12.006.

Jackson, J.M., Thomson, R.E., Brown, L.N., Willis, P.G., Borstad, G.A., 2015. Satellite chlorophyll off the British Columbia Coast, 1997–2010. J. Geophys. Res. Ocean. 4709–4728. https://doi.org/10.1002/jgrc.20224.

Johannessen, S. C., Macdonald, R. W., and Paton, D. W. (2003). A sediment and organic carbon budget for the greater Strait of Georgia. Estuar. Coast. Shelf Sci. 56, 845–860. doi:10.1016/S0272-7714(02)00303-7.

Johannessen, S. C., Macdonald, R. W., and Strivens, J. E. (2021). Has primary production declined in the Salish sea? Can. J. Fish. Aquat. Sci. 78, 312–321. doi:10.1139/cjfas-2020-0115.

Johannessen, S.C., Masson, D., Macdonald, R.W., 2006. Distribution and cycling of suspended particles inferred from transmissivity in the Strait of Georgia, Haro Strait and Juan de Fuca Strait. Atmos. - Ocean 44, 17–27. https://doi.org/10.3137/ao.440102.

Johnson, S.W., Neff, A.D., Thedinga, J.F., Lindeberg, M.R., Maselko, J.M., 2012. Atlas of Nearshore Fishes of Alaska : A Synthesis of Marine Surveys from 1998 to 2011. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-AFSC-239, 261 p.

Johnson, W. K., Miller, L. A., Sutherland, N. E., & Wong, C. S. (2005). Iron transport by mesoscale Haida eddies in the Gulf of Alaska. Deep-Sea Research Part II, 52, 933–953. https://doi.org/10.1016/j.dsr2.2004.08.017.

Juranek, L. W., White, A. E., Dugenne, M., Henderikx Freitas, F., Dutkiewicz, S., Ribalet, F., et al. (2020). The importance of the phytoplankton "middle class" to ocean net community production. Global Biogeochemical Cycles, 34(12). https://doi.org/ 10.1029/2020GB006702.

Kelley, D., & Richards, C. (2021). gsw: Gibbs Sea Water Functions. Retrieved from https://cran.r-project.org/web/packages/gsw/index.html.

Ketchen, K.S., Bourne, N., Butler, T.H., 1983. History and present status of fisheries for marine fishes and invertebrates in the Strait of Georgia, British Columbia. Can. J. Fish. Aquat. Sci. 40, 1095–1119. https://doi.org/10.1139/f83-130.

Komick, N. M., Costa, M. P. F., and Gower, J. (2009). Bio-optical algorithm evaluation for MODIS for western Canada coastal waters: An exploratory approach using in situ reflectance. Remote Sens. Environ. 113, 794–804. doi:10.1016/j.rse.2008.12.005.

Konik, M., Peña, M. A., Hirawake, T., Hunt, B. P. V., Suseelan, P., Eisner, L. B., et al. (n.d.). Bioregionalization of the subarctic Pacific based on phytoplankton phenology and composition. Submitted to Progress in Oceanography.

Kostadinov, T.S., Siegel, D.A., Maritorena, S., 2009. Retrieval of the particle size distribution from satellite ocean color observations. J. Geophys. Res. Ocean. 114, 1–22. https://doi.org/10.1029/2009JC005303.

Kostadinov, T.S., Siegel, D.A., Maritorena, S., 2010. Global variability of phytoplankton functional types from space: Assessment via the particle size distribution. Biogeosciences 7, 3239–3257. https://doi.org/10.5194/bg-7-3239-2010.

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, 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(215), 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 Sens. Environ. 270, 1.14. https://doi.org/10.1016/j.rse.2021.112879.

Kratzer, S., Kyryliuk, D., Brockmann, C., 2020. Inorganic suspended matter as an indicator of terrestrial influence in Baltic Sea coastal areas — Algorithm development and validation, and ecological relevance. Remote Sens. Environ. 237, 1–17. https://doi.org/10.1016/j.rse.2019.111609.

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., R. Christian, J., P. Dunne, J., Gehlen, M., Ilyina, T., G. John, J., Lenton, A., Li, H., S. Lovenduski, N., C. Orr, J., Palmieri, J., Santana-Falcón, Y., Schwinger, J., Séférian, R., A. Stock, C., Tagliabue, A., Takano, Y., Tjiputra, J., Toyama, K., Tsujino, H., Watanabe, M., Yamamoto, A., Yool, A., Ziehn, T., 2020. Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. Biogeosciences 17, 3439–3470. https://doi.org/ 10.5194/bg-17-3439-2020.

Lamont, T., Brewin, R. J. W., and Barlow, R. G. (2018). Seasonal variation in remotely-sensed phytoplankton size structure around southern Africa. Remote Sens. Environ. 204, 617–631. doi:10.1016/j.rse.2017.09.038.

Landry, M.R., Monger, B.C., Selph, K.E., 1993. Time-dependency of microzooplankton grazing and phytoplankton growth in the subarctic Pacific. Prog. Oceanogr. 32, 205–222. https://doi.org/10.1016/0079-6611(93)90014-5.

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.

Latasa, M., Scharek, R., Vidal, M., Vila-Reixach, G., Gutiérrez-Rodríguez, A., Emelianov, M., et al. (2010). Preferences of phytoplankton groups for waters of different trophic status in the northwestern Mediterranean sea. Mar. Ecol. Prog. Ser. 407, 27–42. doi:10.3354/meps08559.

Laurie Van Heukelem Stanford B. Hooker, 2011. The importance of a quality assurance plan for method validation and minimizing uncertainties in the HPLC analysis of phytoplankton pigments, in: Suzanne, R., Llewellyn, C.A., Egeland, E.S., Ohnsen, G. (Eds.), Phytoplankton Pigments: Characterization, Chemotaxonomy, and Applications in Oceanography. Cambridge University Press., Cambridge, UK, pp. 195–242.

Laza-Martinez, A., Seoane, S., Zapata, M., and Orive, E. (2007). Phytoplankton pigment patterns in a temperate estuary: From unialgal cultures to natural assemblages. J. Plankton Res. 29, 913–929. doi:10.1093/plankt/fbm069.

Le Quéré, C., Harrison, S.P., Prentice, I.C., Buitenhuis, E.T., Aumounts, O., Bopp, L., Claustre, H., Cunha, L.C. Da, Geinder, R., Giraud, X., Christine, K., Kochfeld, K.E., Legendre, L., Manizza, M., Platt, T., Rivkin, R.B., Sathyendranath, S., Uitz, J., Watson, A.J., Wolf-Gladrow, D., 2005. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. Glob. Chang. Biol. 11, 2016–2040. https://doi.org/10.1111/j.1365-2486.2005.01004.x.

Leaf, R. T., & Friedland, K. D. (2014). Autumn bloom phenology and magnitude influence haddock recruitment on Georges Bank. ICES Journal of Marine Science, 71(8), 2017–2025. https://doi.org/10.1093/icesjms/fsu076.

Lee, Z.P., Carder, K., Arnone, R., He, M.X., 2007. Determination of primary spectral bands for remote sensing of aquatic environments. Sensors 7, 3428–3441. https://doi.org/10.3390/s7123428.

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. Appl. Opt. 50, 3155–3167. https://doi.org/10.136 4/AO.50.003155.

Legendre, L. (1990). The significance of microalgal blooms for fisheries and for the export of particulate organic carbon in oceans. Journal of Plankton Research, 12(4), 681–699. https://doi.org/10.1093/plankt/12.4.681.

Legendre, L., & Legendre, P. (1998). Numerical Ecology. (Second English Edition ed. Amsterdam, The Netherlands: Elsevier Science, Vol. 132). Elsevier Scientific Pub. Co. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S030438000000291X.

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. Oecologia, 129, 271–280. https://doi.org/10.1007/s004420100716.

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

Li, J., Jamet, C., Zhu, J., Han, B., Li, T., Yang, A., et al. (2019). Error budget in the validation of radiometric products derived from OLCI around the China Sea from open ocean to coastal waters compared with MODIS and VIIRS. Remote Sensing, 11(2400), 1–22. https://doi.org/10.3390/rs11202400.

Li, M., Gargett, A., and Denman, K. (2000). What determines seasonal and interannual variability of phytoplankton and zooplankton in strongly estuarine systems? Application to the semi-enclosed estuary of Strait of Georgia and Juan de Fuca Strait. Estuar. Coast. Shelf Sci. 50, 467–488. doi:10.1006/ecss.2000.0593.

Li, Y.-Y., Chen, X.-H., Xie, Z.-X., Li, D.-X., Wu, P.-F., Kong, L.-F., et al. (2018). Bacterial diversity and nitrogen utilization strategies in the upper layer of the northwestern Pacific Ocean. Frontiers in Microbiology, 9(797), 1–14. https://doi.org/10.3389/fmicb.2018.00797.

Lin, J., OLMO, G., Tilstone, G., Brewin, R., Vabson, V., Ansko, I., EVERS-KING, H., Casal, T., Donlon, C., 2022. Derivation of uncertainty budgets for continuous above-water radiometric measurements along an Atlantic Meridional Transect. Opt. Express 30, 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. Appl. Environ. Microbiol. 85, 1–14.

Liu, B., D'Sa, E. J., Maiti, K., Rivera-Monroy, V. H., and Xue, Z. (2021). Biogeographical trends in phytoplankton community size structure using adaptive sentinel 3-OLCI chlorophyll a and spectral empirical orthogonal functions in the estuarine-shelf waters of the northern Gulf of Mexico. Remote Sens. Environ. 252, 1–20. doi:10.1016/j.rse.2020.112154.

Lohrenz, S. E., Carroll, C. L., Weidemann, A. D., and Tuel, M. (2003). Variations in phytoplankton pigments, size structure and community composition related to wind forcing and water mass properties on the North Carolina inner shelf. Cont. Shelf Res. 23, 1447–1464. doi:10.1016/S0278-4343(03)00131-6.

Lombard, F., Boss, E., Waite, A.M., Uitz, J., Stemmann, L., Sosik, H.M., Schulz, J., Romagnan, J.B., Picheral, M., Pearlman, J., Ohman, M.D., Niehoff, B., Möller, K.O., Miloslavich, P., Lara-Lopez, A., Kudela, R.M., Lopes, R.M., Karp-Boss, L., Kiko, R., Jaffe, J.S., Iversen, M.H., Irisson, J.O., Hauss, H., Guidi, L., Gorsky, G., Giering, S.L.C., Gaube, P., Gallager, S., Dubelaar, G., Cowen, R.K., Carlotti, F., Briseño-Avena, C., Berline, L., Benoit-Bird, K.J., Bax, N.J., Batten, S.D., Ayata, S.D., Appeltans, W., 2019. Globally consistent quantitative observations of planktonic ecosystems. Front. Mar. Sci. 6, 1–21. https://doi.org/10.3389/fmars.2019.00196.

Loos, E. A., and Costa, M. (2010). Inherent optical properties and optical mass classification of the waters of the Strait of Georgia, British Columbia, Canada. Prog. Oceanogr. 87, 144–156. doi:10.1016/j.pocean.2010.09.004.

Loos, E., Costa, M., and Johannessen, S. (2017). Underwater optical environment in the coastal waters of British Columbia, Canada. Facets 2, 872–891. doi:10.1139/facets-2017-0074.

Lu, L., Jiang, T., Xu, Y., Zheng, Y., Chen, B., Cui, Z., et al. (2018). Succession of phytoplankton functional groups from spring to early summer in the central Bohai Sea using HPLC–CHEMTAX approaches. J. Oceanogr. 74, 381–392. doi:10.1007/s10872-018-0469-x.

Lubac, B., & Loisel, H. (2007). Variability and classification of remote sensing reflectance spectra in the eastern English Channel and southern North Sea. Remote Sensing of Coastal Aquatic Environments. Technologies, Techniques and Applications, 110(2), 45–58. https://doi.org/10.1016/j.rse.2007.02.012.

MacFadyen, A., and Hickey, B. M. (2010). Generation and evolution of a topographically linked, mesoscale eddy under steady and variable wind-forcing. Cont. Shelf Res. 30, 1387–1402. doi:10.1016/j.csr.2010.04.001.

Mackas, D.L., Harrison, P.J., 1997. Nitrogenous nutrient sources and sinks in the Juan de Fuca Strait/Strait of Georgia/Puget Sound estuarine system: Assessing the potential for eutrophication. Estuar. Coast. Shelf Sci. 44, 1–21. https://doi.org/10.1006 /ecss.1996.0110.

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(1–3), 265–283. https://doi.org/10.3354/meps144265.

Mahadevan, A., and Campbell, J. W. (2002). Biogeochemical patchiness at the sea surface. Geophys. Res. Lett. 29, 1–4. doi:10.1029/2001GL014116.

Maldonado, M. T., Boyd, P. W., Harrison, P. J., & Price, N. M. (1999). Co-limitation of phytoplankton growth by light and Fe during winter in the NE subarctic Pacific Ocean. Deep Sea Research Part II: Topical Studies in Oceanography, 46(11–12), 2475–2485.

Malick, M.J., Cox, S.P., Mueter, F.J., Peterman, R.M., 2015. Linking phytoplankton phenology to salmon productivity along a north-south gradient in the Northeast Pacific Ocean. Can. J. Fish. Aquat. Sci. 72, 697–708. https://doi.org/10.1139/cjfas-2014-0298.

Marañón, E., 2015. Cell Size as a Key Determinant of Phytoplankton Metabolism and Community Structure. Ann. Rev. Mar. Sci. 7, 241–264. https://doi.org/10.1146/ annurev-marine-010814-015955.

Marañón, E., Cermeño, P., Latasa, M., & Tadonléké, R. D. (2012). Temperature, resources, and phytoplankton size structure in the ocean. Limnology and Oceanography, 57(5), 1266–1278. https://doi.org/10.4319/lo.2012.57.5.1266.

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. Front. Mar. Sci. 9, 1–22. https://doi.org/10.3389/fmars.2022.968470.

Marchetti, A., Juneau, P., Whitney, F.A., Wong, C.S., Harrison, P.J., 2006. Phytoplankton processes during a mesoscale iron enrichment in the NE subarctic Pacific: Part II-Nutrient utilization. Deep. Res. Part II 53, 2114–2130. https://doi.org/ 10.1016 j.dsr2.2006.05.031.

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

Marinov, I., Doney, S.C., Lima, I.D., 2010. Response of ocean phytoplankton community structure to climate change over the 21st century: Partitioning the effects of nutrients, temperature and light. Biogeosciences 7, 3941–3959. https://doi.org/ 10.5194/bg-7-3941-2010.

Maritorena, S., Siegel, D.A., Peterson, A.R., 2002. Optimization of a semianalytical ocean color model for global-scale applications. Opt. Soc. Am. 41, 2705–2714.

Masson, D. (2002). Deep water renewal in the strait of Georgia. Estuar. Coast. Shelf Sci. 54, 115–126. doi:10.1006/ecss.2001.0833.

Masson, D., and Cummins, P. F. (2000). Fortnightly modulation of the estuarine circulation in Juan de Fuca Strait. J. Mar. Res. 58, 439–463. doi:10.1357/002224000321511106.

Masson, D., and Cummins, P. F. (2004). Observations and modeling of seasonal variability in the Straits of Georgia and Juan de Fuca. Journal of Marine Research 62, 491–516. doi: 10.1357/0022240041850075.

Masson, D., and Peña, A. (2009). Chlorophyll distribution in a temperate estuary: The Strait of Georgia and Juan de Fuca Strait. Estuar. Coast. Shelf Sci. 82, 19–28. doi:10.1016/j.ecss.2008.12.022.

Matua, N. J., & Hare, S. R. (2002). The Pacific Decadal Oscillation. Journal of Oceanography, 58, 35–44. https://doi.org/10.1023/A:1015820616384.

McClain, C.R., 2009. A Decade of Satellite Ocean Color Observations. Ann. Rev. Mar. Sci. 1, 19–42. https://doi.org/10.1146/annurev.marine.010908.163650.

McEwan, N., Pawlowicz, R., Pakhomov, E., Maldonado, M.T., 2023. Seasonality of modelled planktonic food web structure in the Strait of Georgia, Canada. Ecol. Modell. 482, 1–14. https://doi.org/10.1016/j.ecolmodel.2023.110402.

McGowan, D.W., Horne, J.K., Parker-Stetter, S.L., 2019. Variability in species composition and distribution of forage fish in the Gulf of Alaska. Deep. Res. Part II Top. Stud. Oceanogr. 165, 221–237. https://doi.org/10.1016/j.dsr2.2016.11.019.

McLaskey, A. K., Forster, I., Del Bel Belluz, J., and Hunt, B. P. V. (2022). A high-resolution time series of particulate matter fatty acids reveals temporal dynamics of the composition and quality available to zooplankton in a temperate coastal ocean. Prog. Oceanogr. 206, 1–19. doi:10.1016/j.pocean.2022.102843.

Meyer, M.G., Gong, W., Kafrissen, S.M., Torano, O., Varela, D.E., Santoro, A.E., Cassar, N., Gifford, S., Niebergall, A.K., Sharpe, G., Marchetti, A., 2022. Phytoplankton size-class contributions to new and regenerated production during the EXPORTS Northeast Pacific Ocean field deployment. Elementa 10, 1–23. https://doi.org/10.1525/elementa.2021.00068.

Miller, C. B. (1993). Pelagic production processes in the Subarctic Pacific. Progress in Oceanography, 32(1), 1–15.

Miller, C.B., Frost, B.W., Wheeler, P.A., Landry, M.R., Welschmeyer, N., Powell, T.M., Limnology, S., Phytoplankton, W.C., Miller, C.B., Frost, B.W., Wheeler, P.A., Powell, T.M., 1991. Ecological dynamics in the subarctic Pacific, a possibly iron-limited ecosystem. Limnol. Oceanogr. 36, 1600–1615. https://doi.org/https://doi.org/ 10.4319/lo.1991.36.8.1600.

Mobley, C.D., 1999. Estimation of the remote-sensing reflectance from above-surface measurements. Appl. Opt. 38, 7442–7455.

Mograne, M.A., Jamet, C., Loisel, H., Vantrepotte, V., Mériaux, X., Cauvin, A., 2019. Evaluation of Five Atmospheric Correction Algorithms over French Optically-Complex Waters for the Sentinel-3A OLCI Ocean Color Sensor. Remote Sens. 11, 1–25. https://doi.org/10.3390/rs11060668.

Moisan, T. A., Rufty, K. M., Moisan, J. R., and Linkswiler, M. A. (2017). Satellite observations of phytoplankton functional type spatial distributions, phenology, diversity, and ecotones. Front. Mar. Sci. 4, 1–24. doi:10.3389/fmars.2017.00189.

Mojica, K. D. A., Van De Poll, W. H., Kehoe, M., Huisman, J., Timmermans, K. R., Buma, A. G. J., et al. (2015). Phytoplankton community structure in relation to vertical stratification along a north-south gradient in the Northeast Atlantic Ocean: Phytoplankton and vertical

stratification. Limnology and Oceanography, 60(5), 1498–1521. https://doi.org/10.1002/lno.10113.

Moore, G.F., Aiken, J., Lavender, S.J., 1999. The atmospheric correction of water colour and the quantitative retrieval of suspended particulate matter in Case II waters: Application to MERIS. Int. J. Remote Sens. 20, 1713–1733. https://doi.org/10.1080/ 014311699212434.

Moore, G.F., Mazeran, C., Huot, J.P., 2017. Case II.S Bright Pixel Atmospheric Correction, Algorithm Theoretical Basis Document.

Moore, T. S., & Brown, C. W. (2020). Incorporating environmental data in abundance-based algorithms for deriving phytoplankton size classes in the Atlantic Ocean. Remote Sensing of Environment, 240, 1–16. https://doi.org/10.1016/j.rse.2020.111689.

Moore, T. S., Campbell, J. W., and Feng, H. (2015). Characterizing the uncertainties in spectral remote sensing reflectance for SeaWiFS and MODIS-Aqua based on global in situ matchup data sets. Remote Sens. Environ. 159, 14–27. doi:10.1016/j .rse.2014.11.025.

Morel, A., Gentili, B., 1996. Diffuse reflectance of oceanic waters III Implication of bidirectionality for the remote-sensing problem. Appl. Opt. 35, 4850–4862. https://doi.org/10.1364/ao.35.004850.

Moses, W. J., Ackleson, S. G., Hair, J. W., Hostetler, C. A., and Miller, W. D. (2016). Spatial scales of optical variability in the coastal ocean: Implications for remote sensing and in situ sampling. J. Geophys. Res. Ocean. 121, 4194–4208. doi:10.1002/ 2016JC011767.

Mouw, C.B., 2019. A Satellite Assessment of Environmental Controls of Phytoplankton Community Size Structure Global Biogeochemical Cycles 540–558. https://doi.org/ 10.1029/2018GB006118.

Mouw, C.B., Barnett, A., Mckinley, G.A., Gloege, L., Pilcher, D., 2016. Phytoplankton size impact on export flux in the global ocean. Global Biogeochem. Cycles 30, 1542–1562. https://doi.org/10.1002/2015GB005355.Received.

Mouw, C.B., Hardman-Mountford, N.J., Alvain, S., Bracher, A., Brewin, R.J.W., Bricaud, A., Ciotti, A.M., Devred, E., Fujiwara, A., Hirata, T., Hirawake, T., Kostadinov, T.S., Roy, S., Uitz, J., 2017. A Consumer's Guide to Satellite Remote Sensing of Multiple Phytoplankton Groups in the Global Ocean. Front. Mar. Sci. 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. (J. L. Mueller, G. S. Fargion, & C. R. McClain, Eds.) (Vol. 5). Greenbelt, US. Retrieved from http://oceancolor.gsfc.nasa.gov/cms/techdocs.

Murphy, J., & Riley, J. P. (1962). A modified single solution method for the determination of phosphate in the natural waters. Analytica Chimica Acta, 27, 31–36.

Myers, K. W., Irvine, J. R., Longerwell, E. A., Urawa, S., Naydenko, S. V., Zavolokin, A. V., & Davis, N. D. (2016). Pacific salmon and steelhead: Life in a changing winter ocean. In S. Urawa, M. Trudel, R. J. Beamish, E. V. Farley, J. R. Guyon, J. R. Irvine, et al. (Eds.), Pacific Salmon and Steelhead Production in a Changing Climate: Past, Present, and Future (pp. 113–138). Victoria, BC: North Pacific Anandromous Fish Commission.

Nair, A., Sathyendranath, S., Platt, T., Morales, J., Stuart, V., Forget, M.H., Devred, E., Bouman, H., 2008. Remote sensing of phytoplankton functional types. Remote Sens. Environ. 112, 3366–3375. https://doi.org/10.1016/j.rse.2008.01.021.

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. Front. Remote Sens. 3, 1–17. https://doi.org/10.3389/frsen.2022.876748.

Navarro, G., Alvain, S., Vantrepotte, V., and Huertas, I. E. (2014). Identification of dominant phytoplankton functional types in the Mediterranean Sea based on a regionalized remote sensing approach. Remote Sens. Environ. 152, 557–575. doi:10.1016/j.rse.2014.06.029.

Nayar, S., and Chou, L. M. (2003). Relative efficiencies of different filters in retaining phytoplankton for pigment and productivity studies. Estuar. Coast. Shelf Sci. 58, 241–248. doi:10.1016/S0272-7714(03)00075-1.

Nemcek, N., & Peña, A. (2014). Institute of Ocean Sciences Protocols for Phytoplankton Pigment Analysis by HPLC (Technical Report) (pp. 1–79). Sidney, BC: Fisheries and Oceans Canada.

Nemcek, N., Hennekes, M., and Perry, I. (2018). “Seasonal dynamics of the phytoplankton community in the Salish Sea from HPLC measurements 2015-2018,” in State of the Physical, Biological and Selected Fishery Resources of Pacific Canadian Marine Ecosystems in 2018, eds. J. L. Boldt, J. Leonard, and P. C. Chandler (Victoria: In Canadian Technical Report of Fisheries and Aquaculture Sciences. Fisheries and Oceans Canada Institute Ocean Sciences), 141–145.

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

Nesbitt, H. K., and Moore, J. W. (2016). Species and population diversity in Pacific salmon fisheries underpin indigenous food security. J. Appl. Ecol. 53, 1489–1499. doi:10.1111/1365-2664.12717.

Nishioka, J., Obata, H., Hirawake, T., Kondo, Y., Yamashita, Y., Misumi, K., & Yasuda, I. (2021). A review: Iron and nutrient supply in the subarctic Pacific and its impact on phytoplankton production. Journal of Oceanography, 77(4), 561–587. https://doi.org/10.1007/s10872-021-00606-5.

Obayashi, Y., Tanoue, E., Suzuki, K., Handa, N., Nojiri, Y., & Wong, C. S. (2001). Spatial and temporal variabilities of phytoplankton community structure in the northern North Pacific as determined by phytoplankton pigments. Deep-Sea Research Part I: Oceanographic Research Papers, 48, 439–469. https://doi.org/10.1016/S0967-0637(00)00036-4.

Oksanen, A.J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P.R., Hara, R.B.O., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2020. Vegan: Community Ecology Package. Retrieved from https://cran.r-project.org/web/packages/vegan/index.html. https://doi.org/10.1007/978-94-024-1179-9_301576.

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

Owens, D., Abeysirigunawardena, D., Biffard, B., Chen, Y., Conley, P., Jenkyns, R., Kerschtien, S., Lavallee, T., MacArthur, M., Mousseau, J., Old, K., Paulson, M., Pirenne, B.,

Scherwath, M., Thorne, M., 2022. The Oceans 2.0/3.0 Data Management and Archival System. Front. Mar. Sci. 9, 1–33. https://doi.org/10.3389/ fmars.2022.806452.

Pahlevan, N., Mangin, A., Balasubramanian, S. V., Smith, B., Alikas, K., Arai, K., et al. (2021). ACIX-Aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters. Remote Sensing of Environment, 258(112366), 1–22. https://doi.org/10.1016/j.rse.2021.112366.

Pakhomov, E. A., Figurkin, A., Hunt, B. P. V., Somov, A., & Mahara, N. (2022). Oceanographic Conditions During the Gulf of Alaska 2019 and 2020 Expeditions (Technical Report No. 18) (pp. 10–14). Vancouver, Canada: North Pacific Anadromous Fish Commission: NPAFC. Retrieved from https://www.researchgate.net/ publication/360376054.

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, 7553–7574. https://doi.org/10.5194/bg-10-7553-2013.

Palevsky, H. I., Ribalet, F., Swalwell, J. E., Cosca, C. E., Cokelet, E. D., Feely, R. A., et al. (2013). The influence of net community production and phytoplankton community structure on CO2 uptake in the Gulf of Alaska. Global Biogeochemical Cycles, 27(3), 664–676. https://doi.org/10.1002/gbc.20058.

Pan, X., Mannino, A., Russ, M. E., Hooker, S. B., and Harding, L. W. (2010). Remote sensing of phytoplankton pigment distribution in the United States northeast coast. Remote Sens. Environ. 114, 2403–2416. doi:10.1016/j.rse.2010.05.015.

Pančić, M., Torres, R.R., Almeda, R., Kiørboe, T., 2019. Silicified cell walls as a defensive trait in diatoms. Proc. R. Soc. B Biol. Sci. 286, 1–9. https://doi.org/10.1098/rspb.2019.0184.

Park, Y. J., & Ruddick, K. (2005). Model of remote-sensing reflectance including bidirectional effects for case 1 and case 2 waters. Applied Optics, 44(7), 1236–1249. https://doi.org/10.1364/AO.44.001236.

Pawlowicz, R., 2002. Observations and linear analysis of sill-generated internal tides and estuarine flow in Haro Strait. J. Geophys. Res. 107, 1–13. https://doi.org/10.1029/2000jc000504.

Pawlowicz, R., Riche, O., and Halverson, M. (2007). The circulation and residence time of the Strait of Georgia using a simple mixing-box approach. Atmos. - Ocean 45, 173–193. doi:10.3137/ao.450401.

Pawlowicz, R., Suzuki, T., Chappell, R., Ta, A., and Esenkulova, S. (2020). Atlas of oceanographic conditions in the Strait of Georgia (2015-2019) based on the Pacific Salmon Foundation's Citizen Science Dataset. Can. Tech. Rep. Fish. Aquat. Sci. 3374, 1–116.

Peña, A., & Nemcek, N. (2020). Nutrient and Phytoplankton Along the Line P in State of the Physical, Biological and Selected Fishery Resources of Pacific Canadian Marine Ecosystems in 2019. In Canadian Technical Report of Fisheries and Aquac. (J. L. Boldt, A. Javorski, & P. C. Chandler, Eds.). Victoria, BC: Fisheries & Oceans Canada Institute Ocean Sciences.

Peña, A., & Nemcek, N. (2021). Nutrient and Phytoplankton Along the Line P in State of the Physical, Biological and Selected Fishery Resources of Pacific Canadian Marine Ecosystems in 2020. In Canadian Technical Report of Fisheries and Aquac. (J. L. Boldt, A. Javorski, & P. C. Chandler, Eds.). Victoria, BC: Fisheries & Oceans Canada Institute Ocean Sciences.

Peña, A., and Nemcek, N. (2018). "Phytoplankton in surface waters along line P and off the west coast of Vancouver Island," in State of the Physical, Biological and Selected Fishery Resources of Pacific Canadian Marine Ecosystems in 2018., eds. J. L. Boldt, J. Leonard, and P. C. Chandler (Victoria: In Canadian Technical Report of Fisheries and Aquaculture Sciences. isheries and Oceans Canada Institute Ocean Sciences), 54–58.

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

Peña, M.A., Fine, I., Callendar, W., 2019. Interannual variability in primary production and shelf-offshore transport of nutrients along the northeast Pacific Ocean margin. Deep. Res. Part II Top. Stud. Oceanogr. 169–170, 104637. https://doi.org/10.1016/j.dsr2.2019.104637.

Peña, M.A., Nemcek, N., Robert, M., 2019. Phytoplankton responses to the 2014–2016 warming anomaly in the northeast subarctic Pacific Ocean. Limnol. Oceanogr. 64, 515–525. https://doi.org/10.1002/lno.11056.

Peña, M.A., Varela, D.E., 2007. Seasonal and interannual variability in phytoplankton and nutrient dynamics along Line P in the NE subarctic Pacific. Prog. Oceanogr. 75, 200–222. https://doi.org/10.1016/j.pocean.2007.08.009.

Perry, R. I., Young, K., Galbraith, M., Chandler, P., Velez-Espino, A., and Baillie, S. (2021). Zooplankton variability in the Strait of Georgia, Canada, and relationships with the marine survivals of Chinook and Coho salmon. PLoS One 16, 1–35. doi:10.1371/journal.pone.0245941.

Peterson, T.D., Crawford, D.W., Harrison, P.J., 2011. Evolution of the phytoplankton assemblage in a long-lived mesoscale eddy in the eastern Gulf of Alaska. Mar. Ecol. Prog. Ser. 424, 53–73. https://doi.org/10.3354/meps08943.

Peterson, T.D., Harrison, P.J., 2012. Diatom dynamics in a long-lived mesoscale eddy in the northeast subarctic Pacific Ocean. Deep. Res. Part I Oceanogr. Res. Pap. 65, 157–170. https://doi.org/10.1016/j.dsr.2012.03.007.

Peterson, T.D., Toews, H.N.J., Robinson, C.L.K., Harrison, P.J., 2007. Nutrient and phytoplankton dynamics in the Queen Charlotte Islands (Canada) during the summer upwelling seasons of 2001-2002. J. Plankton Res. 29, 219–239. https://doi.org/10.1093/plankt/fbm010.

Petrou, K., Kranz, S. A., Trimborn, S., Hassler, C. S., Ameijeiras, S. B., Sackett, O., et al. (2016). Southern Ocean phytoplankton physiology in a changing climate. Journal of Plant Physiology, 203, 135–150. https://doi.org/10.1016/j.jplph.2016.05.004.

Phillips, S. R., and Costa, M. (2017). Spatial-temporal bio-optical classification of dynamic semi-estuarine waters in western North America. Estuar. Coast. Shelf Sci. 199, 35–48. doi:10.1016/j.ecss.2017.09.029.

Pinckney, J. L. (2010) “The fourth SeaWiFS HPLC analysis round-robin experiment (SeaHARRE-4),” in NASA Tech. memo, the fourth SeaWiFS HPLC analysis round-robin experiment (SeaHARRE-4) EdsS. B. Hooker, C. S. Thomas, L.V. Heukelem, L. Schluter, M. E. Russ, J. Ras and J. L. Pinckney. (Greenbelt, MD,USA: NASA).

Plant, J. N., Johnson, K. S., Sakamoto, C. M., Jannasch, H. W., Coletti, L. J., Riser, S. C., & Swift, D. D. (2016). Net community production at Ocean Station Papa observed with nitrate

and oxygen sensors on profiling float. Global Biogeochemical Cycles, 30, 859–879. https://doi.org/10.1002/2015GB005349.

Pramlall, S., Jackson, J.M., Konik, M., Costa, M., 2023. Merged Multi-Sensor Ocean Colour Chlorophyll Product Evaluation for the British Columbia Coast. Remote Sens. 15, 1–17. https://doi.org/10.3390/rs15030687.

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. Limnol. Oceanogr. 53, 605–613. https://doi.org/10.4319/lo.2008.53.2.0605.

Ramette, A. (2007). Multivariate analyses in microbial ecology. FEMS Microbiology Ecology, 62, 142–160. https://doi.org/10.1111/j.1574-6941.2007.00375.x.

Raven, J.A., 1998. The twelfth Tansley Lecture. Small is beautiful: The picophytoplankton. Funct. Ecol. 12, 503–513. https://doi.org/10.1046/j.1365-2435.1998.00233.x.

Redfield, A. C., Ketchum, B. H., & Richards, F. A. (1963). The influence of organisms on the composition of sea water. In M. N. Hill (Ed.), The sea (Vol. 2). New York: John Wiley & Sons.

Reilly, J.E.O., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., Mcclain, C., 1998. Ocean color chlorophyll algorithms for SeaWiFS. J. Geophys. Res. 103, 24,937-24,953.

Rensel, J. E. J., Haigh, N., and Tynan, T. J. (2010). Fraser River sockeye salmon marine survival decline and harmful blooms of Heterosigma akashiwo. Harmful Algae 10, 98–115. doi:10.1016/ j.hal.2010.07.005.

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

Rêve-Lamarche, A.H., Alvain, S., Racault, M.F., Dessailly, D., Guiselin, N., Jamet, C., Vantrepotte, V., Beaugrand, G., 2017. Estimation of the potential detection of diatom assemblages based on ocean color radiance anomalies in the North Sea. Front. Mar. Sci. 4, 1–20. https://doi.org/10.3389/fmars.2017.00408.

Ribalet, F., Marchetti, A., Hubbard, K.A., Brown, K., Durkin, C.A., Morales, R., Robert, M., Swalwell, J.E., Tortell, P.D., Armbrust, E.V., 2010. Unveiling a phytoplankton hotspot at a narrow boundary between coastal and offshore waters. Proc. Natl. Acad. Sci. U. S. A. 107, 16571–16576. https://doi.org/10.1073/pnas.1005638107.

Roesler, C., Uitz, J., Claustre, H., Boss, E., Xing, X., Organelli, E., Briggs, N., Bricaud, A., Schmechtig, C., Poteau, A., D'Ortenzio, F., Ras, J., Drapeau, S., Haëntjens, N., Barbieux, M., 2017. Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: A global analysis of WET Labs ECO sensors. Limnol. Oceanogr. Methods 15, 572–585. https://doi.org/10.1002/lom3.10185.

Ross, T., & Peña, A. (2022). Trends in the Observations at Ocean Station P and Relevance to the Two Expeditions (Technical Report No. 18) (pp. 70–80). Vancouver, Canada: North Pacific Anadromous Fish Commission.

Röttgers, R., Heymann, K., and Krasemann, H. (2014). Suspended matter concentrations in coastal waters: Methodological improvements to quantify individual measurement uncertainty. Estuar. Coast. Shelf Sci. 151, 575–582. doi:10.1016/j.ecss.2014.10.010.

Roy, S., Sathyendranath, S., Bouman, H., and Platt, T. (2013). The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. Remote Sens. Environ. 139, 185–197. doi:10.1016/j.rse.2013.08.004.

Royer, S. J., Levasseur, M., Lizotte, M., Arychuk, M., Scarratt, M. G., Wong, C. S., et al. (2010). Microbial dimethylsulfoniopropionate (DMSP) dynamics along a natural iron gradient in the northeast subarctic Pacific. Limnology and Oceanography, 55(4), 1614–1626. https://doi.org/10.4319/lo.2010.55.4.1614.

Royer, T. C. (1983). Observations of the Alaska Coastal Current. In H. G. Gade, A. Edwards, & H. Svendsen (Eds.), Coastal Oceanography (1st ed., pp. 9–30). Boston, MA: Springer. https://doi.org/10.1007/978-1-4615-6648-9.

Ruddick, K., De Cauwer, V., Van Mol, B., 2005. Use of the near infrared similarity reflectance spectrum for the quality control of remote sensing data, in: Proceedings of SPIE. pp. 1–12. https://doi.org/10.1117/12.615152.

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

Ruddick, K.G., Voss, K., Banks, A.C., Boss, E., Castagna, A., Frouin, R., Hieronymi, M., Jamet, C., Johnson, B.C., Kuusk, J., Lee, Z., Ondrusek, M., Vabson, V., 2019. A Review of Protocols for Fiducial Reference Measurements of Downwelling Irradiance for the Validation of Satellite Remote Sensing Data over Water.

Sá, C., D'Alimonte, D., Brito, A.C., Kajiyama, T., Mendes, C.R., Vitorino, J., Oliveira, P.B., da Silva, J.C.B., Brotas, V., 2015. Validation of standard and alternative satellite ocean-color chlorophyll products off Western Iberia. Remote Sens. Environ. 168, 403–419. https://doi.org/10.1016/j.rse.2015.07.018.

Sadeghi, A., Dinter, T., Vountas, M., Taylor, B. B., Altenburg-Soppa, M., Peeken, I., et al. (2012). Improvement to the PhytoDOAS method for identification of coccolithophores using hyper-spectral satellite data. Ocean Sci. 8, 1055–1070. doi:10.5194/os-8-1055-2012.

Sambrotto, R.N., Lorenzen, C.J., 1987. The Gulf of Alaska: Physical Environment and Biological Resources, in: Hood, D.W., Zimmerman, S.T. (Eds.), The Gulf of Alaska: Physical Environment and Biological Resources. Ocean Assessments Division, NOAA. U.S. Department of Commerce, Washington DC, pp. 249–282.

Sastri, A.R., Dower, J.F., 2009. Interannual variability in chitobiase-based production rates of the crustacean zooplankton community in the Strait of Georgia. Mar. Ecol. Prog. Ser. 388, 147–157. https://doi.org/10.3354/meps08111.

Sathyendranath, S., Brewin, R.J.W., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., Cipollini, P., Couto, A.B., Dingle, J., Doerffer, R., Donlon, C., Dowell, M., Farman, A., Grant, M., Groom, S., Horseman, A., Jackson, T., Krasemann, H., Lavender, S., Martinez-vicente, V., Mazeran, C., Mélin, F., Moore, T.S., Frouin, R., Taberner, M., Müller, D., Regner, P., Roy, S., Steele, C.J., Steinmetz, F., Swinton, J., Thompson, A., Valente, A., Zühlke, M., Brando, V.E., Feng, H., Feldman, G., Franz, B.A., Jr, RWG, 17, S.B.H., Kahru, M., Kratzer, S., Mitchell, B.G., Muller-Karger, F.E., Sosik, H.M., Voss, K.J., Werdell, J., Platt, T., 2019. An Ocean-Colour Time Series for Use in Climate Studies : The Experience of the Ocean-Colour Climate Change Initiative (OC-CCI). Sensors 19, 1–33. https://doi.org/10.3390/s19194285.

Schwing, F. B., Murphree, T., DeWitt, L., & Green, P. M. (2002). The evolution of oceanic and atmospheric anomalies in the northeast Pacific during the El Niño and La Niña events of 1995–2001. Progress in Oceanography, 54(1–4), 459–491. https://doi.org/10.1016/S0079-6611(02)00064-2.

Scott, J.P., Werdell, P.J., 2019. Comparing level-2 and level-3 satellite ocean color retrieval validation methodologies. Opt. Express 27, 30140–30157. https://doi.org/10.1364/oe.27.030140.

Shen, M., Duan, H., Cao, Z., Xue, K., Qi, T., Ma, J., et al. (2020). Sentinel-3 OLCI observations of water clarity in large lakes in eastern China: Implications for SDG 6.3.2 evaluation. Remote Sens. Environ. 247, 1–17. doi:10.1016/j.rse.2020.111950.

Shen, M., Duan, H., Cao, Z., Xue, K., Qi, T., Ma, J., et al. (2020). Sentinel-3 OLCI observations of water clarity in large lakes in eastern China: Implications for SDG 6.3.2 evaluation. Remote Sensing of Environment, 247(6), 1–17. https://doi.org/10.1016 /j.rse.2020.111950.

Siddon, E.C., De Forest, L.G., Blood, D.M., Doyle, M.J., Matarese, A.C., 2019. Early life history ecology for five commercially and ecologically important fish species in the eastern and western Gulf of Alaska. Deep. Res. Part II Top. Stud. Oceanogr. 165, 7–25. https://doi.org/10.1016/j.dsr2.2016.06.022.

Soja-Woźniak, M., Craig, S. E., Kratzer, S., Wojtasiewicz, B., Darecki, M., & Jones, C. T. (2017). A novel statistical approach for ocean colour estimation of inherent optical properties and cyanobacteria abundance in optically complex waters. Remote Sensing, 9, 1–22. https://doi.org/10.3390/rs9040343.

Stabeno, P. J., Bell, S., Cheng, W., Danielson, S., Kachel, N. B., & Mordy, C. W. (2016). Long-term observations of Alaska Coastal Current in the northern Gulf of Alaska. Deep-Sea Research Part II: Topical Studies in Oceanography, 132, 24–40. https://doi.org/10.1016/j.dsr2.2015.12.016.

Stabeno, P. J., Bond, N. A., Hermann, A. J., Kachel, N. B., Mordy, C. W., & Overland, J. E. (2004). Meteorology and oceanography of the Northern Gulf of Alaska. Continental Shelf Research, 24(7–8), 859–897. https://doi.org/10.1016/ j.csr.2004.02.007.

Stabeno, P. J., Reed, R. K., & Schumacher, J. D. (1995). The Alaska Coastal Current: Continuity of transport and forcing. Journal of Geophysical Research, 100(C2), 2477–2485.

Stavn, R. H., Rick, H. J., and 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. Estuar. Coast. Shelf Sci. 81, 575–582. https://doi.org/10.1016/j.ecss.2008.12.017.

Steinberg, D.K., Landry, M.R., 2017. Zooplankton and the Ocean Carbon Cycle. Ann. Rev. Mar. Sci. 9, 413–444. https://doi.org/10.1146/annurev-marine-010814-015924.

Steinmetz, F., Deschamps, P.-Y., Ramon, D., 2011. Atmospheric correction in presence of sun glint: application to MERIS. Opt. Express 19, 9783–9800. https://doi.org/10.1364/oe.19.009783.

Steinmetz, F., Ramon, D., Deschamps, P.-Y., 2016. ATBD v1 - Polymer Atmospheric Correction Algorithm ref: D2.3 Date: 23/12/2016 Issue: 2.1. PML, United Kingdom.

Stock, A., and Subramaniam, A. (2020). Accuracy of empirical satellite algorithms for mapping phytoplankton diagnostic pigments in the open ocean: A supervised learning perspective. Front. Mar. Sci. 7, 1–19. doi:10.3389/fmars.2020.00599.

Stock, C. A., Dunne, J. P., & John, J. G. (2014). Global-scale carbon and energy flows through the marine planktonic food web: An analysis with a coupled physical-biological model. Progress in Oceanography, 120(7), 1–28. https://doi.org/10.1016/j.pocean.2013.07.001.

Stockner, J.G., Cliff, D.D., Shortreed, K.R.S., 1979. Phytoplankton Ecology of the Strait of Georgia, British Columbia. J. Fish. Res. Board Canada 36, 657–666. https://doi.org/10.1139/f79-095.

Stoecker, D. K., & Lavrentyev, P. J. (2018). Mixotrophic plankton in the polar seas: A Pan-Arctic review. Frontiers in Marine Science, 5(292), 1–12. https://doi.org/10.3389/fmars.2018.00292.

Strom, S. L., & Welschmeyer, N. A. (1991). Pigment-specific rates of phytoplankton growth and microzooplankton grazing in the open subarctic Pacific Ocean. Limnology and Oceanography, 36(1), 50–63. https://doi.org/10.4319/lo.1991.36.1.0050.

Strom, S. L., Macri, E. L., & Fredrickson, K. A. (2010). Light limitation of summer primary production in the coastal Gulf of Alaska: Physiological and environmental causes. Marine Ecology Progress Series, 402, 45–57. https://doi.org/10.3354/meps08456.

Strom, S.L., Fredrickson, K.A., Bright, K.J., 2016. Spring phytoplankton in the eastern coastal Gulf of Alaska: Photosynthesis and production during high and low bloom years. Deep. Res. Part II Top. Stud. Oceanogr. 132, 107–121. https://doi.org/10.1016/j.dsr2.2015.05.003.

Strom, S.L., Olson, M.B., Macri, E.L., Mordy, CW, 2006. Cross-shelf gradients in phytoplankton community structure, nutrient utilization, and growth rate in the coastal Gulf of Alaska. Mar. Ecol. Prog. Ser. 328, 75–92. https://doi.org/10.3354/meps328075.

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. Prog. Oceanogr. 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. Front. Mar. Sci. 9, 1–22. https://doi.org/10.3389/fmars.2022.832684.

Sutton, J. N., Johannessen, S. C., and 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, 7179–7194. doi:10.5194/bg-10-7179-2013.

Suzuki, K., Minami, C., Liu, H., Saino, T., 2002. Temporal and spatial patterns of chemotaxonomic algal pigments in the subarctic Pacific and the Bering Sea during the early summer of 1999. Deep. Res. Part II Top. Stud. Oceanogr. 49, 5685–5704. https://doi.org/10.1016/S0967-0645(02)00218-7.

Swan, C. M., Vogt, M., Gruber, N., & Laufkoetter, C. (2016). A global seasonal surface ocean climatology of phytoplankton types based on CHEMTAX analysis of HPLC pigments. Deep Sea Research Part I: Oceanographic Research Papers, 109, 137–156. https://doi.org/10.1016/j.dsr.2015.12.002.

Taves, R. C., Janssen, D. J., Peña, M. A., Ross, A. R. S., Simpson, K. G., Crawford, W. R., & Cullen, J. T. (2022). Relationship between surface dissolved iron inventories and net community production during a marine heatwave in the subarctic northeast Pacific.

Environmental Science: Processes & Impacts, 24(9), 1460–1473. https://doi.org/10.1039/D2EM00021K.

Taylor, B. B., Taylor, M. H., Dinter, T., & Bracher, A. (2013). Estimation of relative phycoerythrin concentrations from hyperspectral underwater radiance measurements — A statistical approach. Journal of Geophysical Research : Oceans, 118, 2948–2960. https://doi.org/10.1002/jgrc.20201.

Thomson, R. E., & Krassovski, M. V. (2010). Poleward reach of the California Undercurrent extension. Journal of Geophysical Research: Oceans, 115(9), 1–9. https://doi.org/10.1029/2010JC006280.

Thomson, R. E., Beamish, R. J., Beacham, T. D., Trudel, M., Whitfield, P. H., and Hourston, R. A. S. (2012). Anomalous ocean conditions may explain the recent extreme variability in Fraser River sockeye salmon production. Mar. Coast. Fish. 4, 415–437. doi:10.1080/19425120.2012.675985.

Thomson, R. E., Heesemann, M., Davis, E. E., and Hourston, R. A. S. (2014). Continental microseismic intensity delineates oceanic upwelling timing along the west coast of North America. Geophys. Res. Lett. 41, 6872–6880. doi:10.1002/2014GL061241.

Tilstone, G. H., Pardo, S., Simis, S. G. H., Qin, P., Selmes, N., Dessailly, D., et al. (2022). Consistency between satellite ocean colour products under high coloured dissolved organic matter absorption in the Baltic Sea. Remote Sens. 14, 1–19. doi:10.3390/rs14010089.

Tilstone, G., Olmo, G.D., Hieronymi, M., Ruddick, K., Beck, M., Ligi, M., Costa, M., Alimonte, D.D., Vellucci, V., Vansteenwegen, D., Bracher, A., Wiegmann, S., Kuusk, J., Vabson, V., Ansko, I., Vendt, R., Donlon, C., Casal, T., 2020. Field Intercomparison of Radiometer Measurements for Ocean Colour Validation. Remote Sens. 12, 1–48. https://doi.org/10.3390/rs12101587.

Tilstone, G.H., Pardo, S., Dall'Olmo, G., Brewin, R.J.W., Nencioli, F., Dessailly, D., Kwiatkowska, E., Casal, T., Donlon, C., 2021. Performance of Ocean Colour Chlorophyll a algorithms for Sentinel-3 OLCI, MODIS-Aqua and Suomi-VIIRS in open-ocean waters of the Atlantic. Remote Sens. Environ. 260, 1–19. https://doi.org/10.1016/j.rse.2021.112444.

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 Sens. Environ. 115, 2578–2593. https://doi.org/10.1016/j.rse.2011.05.014.

Travers-Smith, H., Giannini, F., Sastri, A. R., and Costa, M. (2021). Validation of non-photochemical quenching corrections for chlorophyll-a measurements aboard ships of opportunity. Front. Mar. Sci. 8, 1–15. doi:10.3389/fmars.2021.686750.

Turner, R. E., Qureshi, N., Rabalais, N. N., Dortch, Q., Shaw, R. F., & Cope, J. (1998). Fluctuating silicate: nitrate ratios and coastal plankton food webs. Proceedings of the National Academy of Sciences, 95(22), 13048–13051.

Ueno, H., Bracco, A., Barth, J. A., Budyansky, M. V., Hasegawa, D., Itoh, S., et al. (2023). Review of oceanic mesoscale processes in the North Pacific: Physical and biogeochemical impacts. Progress in Oceanography, 212(102955), 1–37. https://doi.org/10.1016/j.pocean.2022.102955.

Uitz, J., Claustre, H., Morel, A., Hooker, S.B., 2006. Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. J. Geophys. Res. Ocean. 111. https://doi.org/10.1029/2005JC003207.

Uitz, J., Huot, Y., Bruyant, F., Babin, M., Claustre, H., 2008. Relating phytoplankton photophysiological properties to community structure on large scales. Limnol. Oceanogr. 53, 614–630. https://doi.org/10.4319/lo.2008.53.2.0614.

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 Sens. Environ. 171, 58–74. https://doi.org/10.1016/j.rse.2015.09.027.

Utermöhl, H. (1958). Zur Vervollkommnung der quantitativen Phytoplankton-Methodik: Mitteilungen Internationale Vereinigung für Theoretische und Angewandte. Limnologie 9, 1–38. doi: 10.1080/05384680.1958.11904091.

Vabson, V., Kuusk, J., Ansko, I., Vendt, R., Alikas, K., Ruddick, K., Ansper, A., Bresciani, M., Burmester, H., Costa, M., D'Alimonte, D., Dall'Olmo, G., Damiri, B., Dinter, T., Giardino, C., Kangro, K., Ligi, M., Paavel, B., Tilstone, G., Van Dommelen, R., Wiegmann, S., Bracher, A., Donlon, C., Casal, T., 2019a. Laboratory intercomparison of radiometers

used for satellite validation in the 400-900 nm range. Remote Sens. 11, 1–24. https://doi.org/10.3390/rs11091129.

Vabson, V., Kuusk, J., Ansko, I., Vendt, R., Alikas, K., Ruddick, K., Ansper, A., Bresciani, M., Burmester, H., Costa, M., D'Alimonte, D., Dall'Olmo, G., Damiri, B., Dinter, T., Giardino, C., Kangro, K., Ligi, M., Paavel, B., Tilstone, G., Van Dommelen, R., Wiegmann, S., Bracher, A., Donlon, C., Casal, T., 2019b. Field intercomparison of radiometers used for satellite validation in the 400-900 nm range. Remote Sens. 11, 1–22. https://doi.org/10.3390/rs11091129.

Vaillancourt, R. D., Lance, V. P., & Marra, J. F. (2018). Phytoplankton chemotaxonomy within contiguous optical layers across the western North Atlantic Ocean and its relationship to environmental parameters. Deep-Sea Research Part I: Oceanographic Research Papers, 139, 14–26. https://doi.org/10.1016/j.dsr.2018.05.007.

Vandenberg, N., Costa, M., Coady, Y., Agbaje, T., 2017. PySciDON : A Python Scientific Framework for Development of Ocean Network Applications. IEEE.

Vanhellemont, Q., & Ruddick, K. (2021). Atmospheric correction of Sentinel-3/OLCI data for mapping of suspended particulate matter and chlorophyll-a concentration in Belgian turbid coastal waters. Remote Sensing of Environment, 256(9), 1–18. https://doi.org/10.1016/j.rse.2021.112284.

Vanni, M. J., & Findlay, D. L. (1990). Trophic cascades and phytoplankton community structure. Ecology, 71(3), 921–937. https://doi.org/10.2307/1937363.

Varela, D. E., & Harrison, P. J. (1999). Seasonal variability in nitrogenous nutrition of phytoplankton assemblages in the northeastern subarctic Pacific Ocean. Deep-Sea Research Part II: Topical Studies in Oceanography, 46, 2505–2538. https://doi.org/10.1016/S0967-0645(99)00074-0.

Vinogradov, M.E., Shushkina, E.A., Vedernikov, V.I., Nezlin, N.P., Gagarin, V.I., 1997. Primary production and plankton stocks in the Pacific Ocean and their seasonal variation according to remote sensing and field observations. Deep. Res. Part II 44, 1979–2001. https://doi.org/10.1016/S0967-0645(97)00086-6.

Vishnu, P. S., & Costa, M. (2023). Evaluating the performance of Sentinel-3A OLCI products in the subarctic northeast Pacific. Remote Sensing, 15(13), 3244. https://doi.org/10.3390/rs15133244.

Vishnu, P. S., Xi, H., Belluz, J. D. 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, 9, 1018510. https://doi.org/10.3389/fmars.2022.1018510.

Waite, J.N., Mueter, F.J., 2013. Spatial and temporal variability of chlorophyll-a concentrations in the coastal Gulf of Alaska, 1998-2011, using cloud-free reconstructions of SeaWiFS and MODIS-Aqua data. Prog. Oceanogr. 116, 179–192. https://doi.org/10.1016/j.pocean.2013.07.006.

Wang, C., Pawlowicz, R., and Sastri, A. R. (2019). Diurnal and seasonal variability of near-surface oxygen in the Strait of Georgia. J. Geophys. Res. Ocean. 124, 2418–2439. doi:10.1029/2018JC014766.

Wang, F., Xie, Y., Wu, W., Sun, P., Wang, L., & Huang, B. (2019). Picoeukaryotic diversity and activity in the northwestern Pacific Ocean based on rDNA and rRNA high-throughput sequencing. Frontiers in Microbiology, 9(3259), 1–13. https://doi.org/10.3389/fmicb.2018.03259.

Wang, G., Lee, Z., and Mouw, C. B. (2018). Concentrations of multiple phytoplankton pigments in the global oceans obtained from satellite ocean color measurements with MERIS. Appl. Sci. 8, 1–20. doi:10.3390/app8122678.

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, Croatia.

Wang, Z., Costa, M., 2022. Autonomous shipborne in Situ reflectance data in optically complex coastal waters: A case study of the salish Sea, Canada. Front. Remote Sens. 3, 1–17. https://doi.org/10.3389/frsen.2022.867570.

Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y., Zemla, J., et al. (2021). Corrplot: Visualization of a correlation matrix. Retrieved from https://github.com/taiyun/corrplot.

Weingartner, T., 2006. The Physical Environment of the Gulf of Alaska, First edit. ed, Long-term ecological change in the northern gulf of alaska. Elsevier. https://doi.org/doi.org/10.1016/B978-044452960-2/50003-5.

Weingartner, T., Coyle, K., Finney, B., Hopcroft, R., Whitledge, T., Brodeur, R., et al. (2002). The Northeast Pacific GLOBEC program: Coastal Gulf of Alaska. Oceanography, 15(2), 48–63. https://doi.org/10.5670/oceanog.2002.21.

Weingartner, T., Eisner, L., Eckert, G.L., Danielson, S., 2009. Southeast Alaska: Oceanographic habitats and linkages. J. Biogeogr. 36, 387–400. https://doi.org/10.1111/j.1365-2699.2008.01994.x.

Weingartner, T.J., Danielson, S.L., Royer, T.C., 2005. Freshwater variability and predictability in the Alaska Coastal Current. Deep. Res. Part II Top. Stud. Oceanogr. 52, 169–191. https://doi.org/10.1016/j.dsr2.2004.09.030.

Welschmeyer, N.A., Strom, S., Goericke, R., DiTullio, G., Belvin, M., Petersen, W., 1993. Primary production in the subarctic Pacific Ocean: Project SUPER. Prog. Oceanogr. 32, 101–135. https://doi.org/10.1016/0079-6611(93)90010-B.

Werdell, P. J., & Bailey, S. W. (2005). An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation. Remote Sensing of Environment, 98, 122–140. https://doi.org/10.1016/j.rse.2005.07.001.

Werdell, P.J., Behrenfeld, michael J., BontemPi, P.S., BoSS, E., CairnS, B., DaviS, G.T., Franz, B.A., GlieSe, U.B., Gorman, E.T., HaSekamP, O., Knobelspiesse, K. d., Mannino, A., MartinS, J. vanderlei, Mcclain, C.R., MeiSter, G., Remer, L.A., 2019. The plankton, aerosol, cloud, ocean ecosystem mission Status, Science, Advances. Bull. Am. Meteorol. Soc. 1775–1794. https://doi.org/10.1175/BAMS-D-18-0056.1.

Werdell, P.J., Mckinna, L.I.W., Boss, E., Ackleson, S.G., Craig, S.E., Gregg, W.W., Lee, Z., Maritorena, S., Roesler, C.S., Rousseaux, C.S., Stramski, D., Sullivan, J.M., Twardowski, M.S., Tzortziou, M., Zhang, X., 2018. An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing. Prog. Oceanogr. 160, 186–212. https://doi.org/10.1016/j.pocean.2018.01.001.

Werdell, P.J., Roesler, C.S., Goes, J.I., 2014. Discrimination of phytoplankton functional groups using an ocean reflectance inversion model. Appl. Opt. 53, 4833–4849. https://doi.org/10.1364/AO.53.004833.

Westberry, T.K., Schultz, P., Behrenfeld, M.J., Dunne, J.P., Hiscock, M.R., Maritorena, S., Sarmiento, J.L., Siegel, D.A., 2016. Annual cycles of phytoplankton biomass in the subarctic Atlantic and Pacific Ocean. Global Biogeochem. Cycles 30, 175–190. https://doi.org/https://doi.org/10.1002/2015GB005276.

Whitney, F.A., Wong, C.S., Boyd, P.W., 1998. Interannual variability in nitrate supply to surface waters of the Northeast Pacific Ocean. Mar. Ecol. Prog. Ser. 170, 15–23. https://doi.org/10.3354/meps170015.

Whitney, F. A., & Freeland, H. J. (1999). Variability in upper-ocean water properties in the NE Pacific Ocean. Deep-Sea Research Part II: Topical Studies in Oceanography, 46, 2351–2370. https://doi.org/10.1016/S0967-0645(99)00067-3.

Whitney, F. A., Crawford, W. R., & Harrison, P. J. (2005). Physical processes that enhance nutrient transport and primary productivity in the coastal and open ocean of the subarctic NE Pacific. Deep-Sea Research Part II: Topical Studies in Oceanography, 52(5–6), 681–706. https://doi.org/10.1016/j.dsr2.2004.12.023.

Whitney, F. A., Freeland, H. J., & Robert, M. (2007). Persistently declining oxygen levels in the interior waters of the eastern subarctic Pacific. Progress in Oceanography, 75(2), 179–199. https://doi.org/10.1016/j.pocean.2007.08.007.

Wickham, H. (2016). Ggplot2: Elegant Graphics for Data Analysis. (R. Gentleman, G. Parmigiani, & K. Hornik, Eds.). New York: Springer. https://doi.org/10.1007/978-0-387-78171-6.

Wolanin, A., Soppa, M.A., Bracher, A., 2016. Investigation of spectral band requirements for improving retrievals of phytoplankton functional types. Remote Sens. 8, 1–21. https://doi.org/10.3390/rs8100871.

Womble, J.N., Sigler, M.F., 2006. Seasonal availability of abundant, energy-rich prey influences the abundance and diet of a marine predator, the Steller sea lion Eumetopias jubatus. Mar. Ecol. Prog. Ser. 325, 281–293. https://doi.org/10.3354/meps325281.

Womble, J.N., Sigler, M.F., Willson, M.F., 2009. Linking seasonal distribution patterns with prey availability in a central-place forager, the Steller sea lion. J. Biogeogr. 36, 439–451. https://doi.org/10.1111/j.1365-2699.2007.01873.x.

Wong, C. S., & Matear, R. J. (1999). Sporadic silicate limitation of phytoplankton productivity in the subarctic NE Pacific. Deep-Sea Research Part II: Topical Studies in Oceanography, 46, 2539–2555. https://doi.org/10.1016/S0967-0645(99)00075-2.

Wood, E. D., Armstrong, F. A. J., & Richards, F. A. (1967). Determination of nitrate in sea water by cadmium-copper reduction to nitrite. Journal of the Marine Biological Association of the United Kingdom, 47(1), 23–31. https://doi.org/10.1017/S002531540003352X.

Wu, P.-F., Li, D.-X., Kong, L.-F., Li, Y.-Y., Zhang, H., Xie, Z.-X., et al. (2020). The diversity and biogeography of microeukaryotes in the euphotic zone of the northwestern Pacific Ocean. Science of The Total Environment, 698, 1–12. https://doi.org/10.1016/j.scitotenv.2019.134289.

Xi, H., Hieronymi, M., Krasemann, H., Röttgers, R., 2017. Phytoplankton Group Identification Using Simulated and In situ Hyperspectral Remote Sensing Reflectance 4, 1–13. 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 Sens. 7, 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), 1–27. https://doi.org/10.1029/2020jc017127.

Xi, H., Losa, S.N., Mangin, A., Soppa, M.A., Garnesson, P., Demaria, J., Liu, Y., D'Andon, O.H.F., Bracher, A., 2020. Global retrieval of phytoplankton functional types based on empirical orthogonal functions using CMEMS GlobColour merged products and further extension to OLCI data. Remote Sens. Environ. 240, 1–21. https://doi.org/10.1016/j.rse.2020.111704.

Xin, Q., Qin, X., Wu, G., Ding, X., Wang, X., Hu, Q., et al. (2023). Phytoplankton community structure in the Western Subarctic Gyre of the Pacific Ocean during summer determined by a combined approach of HPLC-pigment CHEMTAX and metabarcoding sequencing. Frontiers in Marine Science, 10(1116050), 01–13. https://doi.org/10.3389/fmars.2023.1116050.

Yang, B., Emerson, S.R., Angelica Penã, M., 2018. The effect of the 2013-2016 high temperature anomaly in the subarctic Northeast Pacific (the "blob") on net community production. Biogeosciences 15, 6747–6759. https://doi.org/10.5194/bg-15-6747-2018.

Yin, K., Harrison, P.J., Goldblatt, R.H., St. John, M.A., Beamish, R.J., 1997. Factors controlling the timing of the spring bloom in the Strait of Georgia estuary, British Columbia, Canada. Can. J. Fish. Aquat. Sci. 54, 1985–1995. https://doi.org/ 10.1139/f97-106.

Yuan, X., & Talley, L. D. (1996). The subarctic frontal zone in the North Pacific: Characteristics of frontal structure from climatological data and synoptic surveys. Journal of Geophysical Research, 101(C7), 16491–16508.

Zhang, H., Wang, Y., Xiu, P., Chai, F., 2021. Modeling the seasonal variability of phytoplankton in the subarctic northeast Pacific Ocean. Mar. Ecol. Prog. Ser. 680, 33–50. https://doi.org/10.3354/meps13914.

Zhang, H.R., Wang, Y., Xiu, P., Qi, Y., Chai, F., 2021. Roles of Iron Limitation in Phytoplankton Dynamics in the Western and Eastern Subarctic Pacific. Front. Mar. Sci. 8, 1–18. https://doi.org/10.3389/fmars.2021.735826.

Zibordi, G., Melin, F., and Berthon, J. F. (2018). A regional assessment of OLCI data products. IEEE Geosci. Remote Sens. Lett. 15, 1490–1494. doi:10.1109/ LGRS.2018.2849329.

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 Sci. 8, 567–586. https://doi.org/10.5194/os-8-567-2012.

Appendix

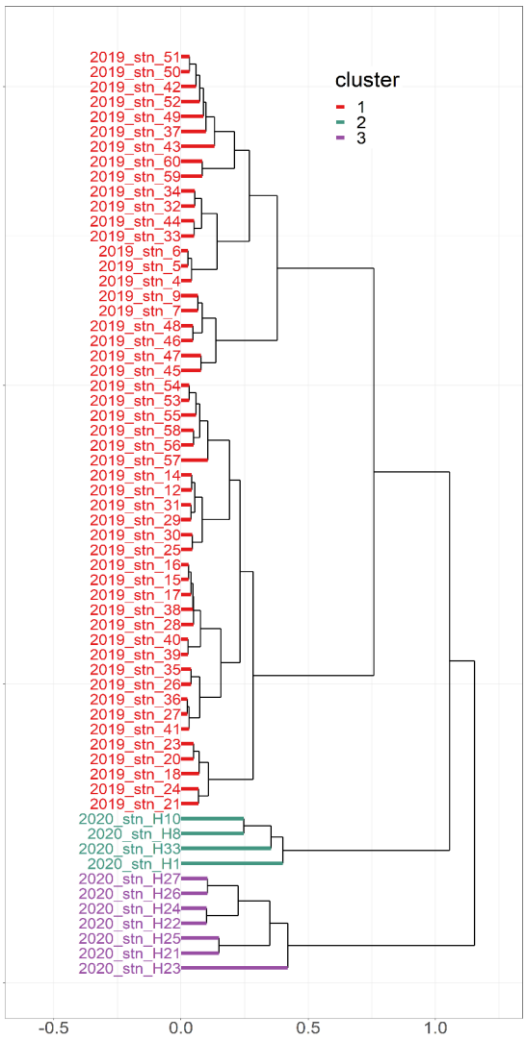

Figure S1. Dendrogram showing the three major clusters of stations with similar pigment:TChla ratios.

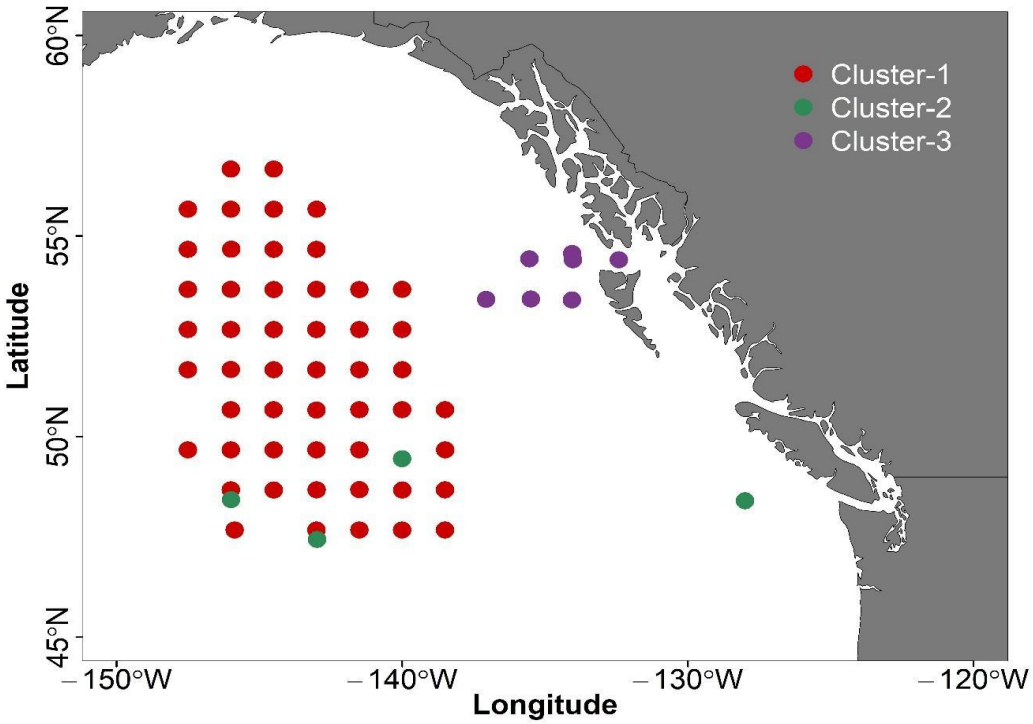

Figure S2. Spatial distribution of the three major clusters.

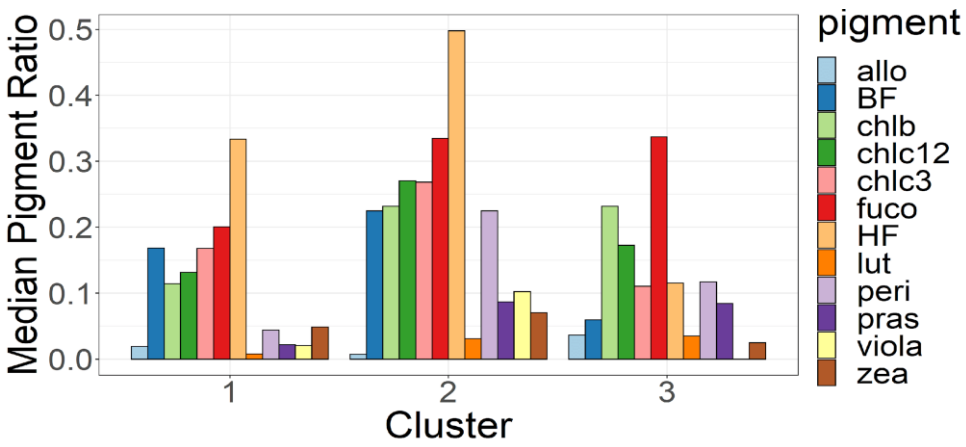

Figure S3. Median pigment:TChla ratios for the three major clusters.

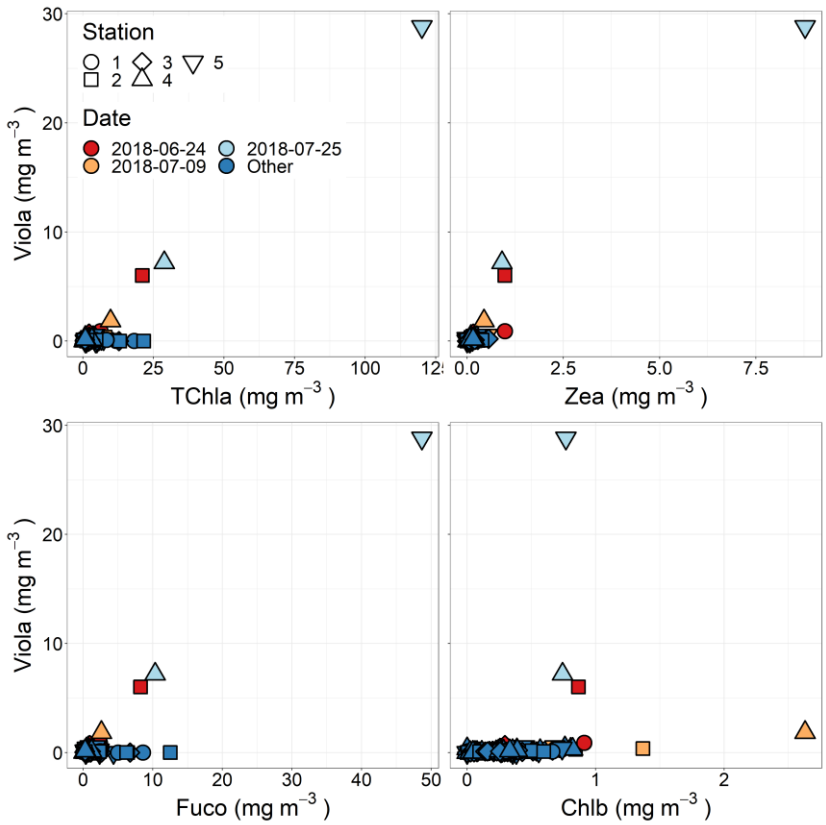

Figure S4. Scatterplots of Violaxanthin (Viola) and total Chlorophyll a (TChla, mg/m^3), Zeaxanthin (Zea), Fucoxanthin (Fuco) and Chlorophyll b (Chlb) showing distinct relationships during suspected Raphidophyte (*Heterosigma akashiwo*) blooms in the central Strait of Georgia (SoG). All concentrations are in mg/m^3.

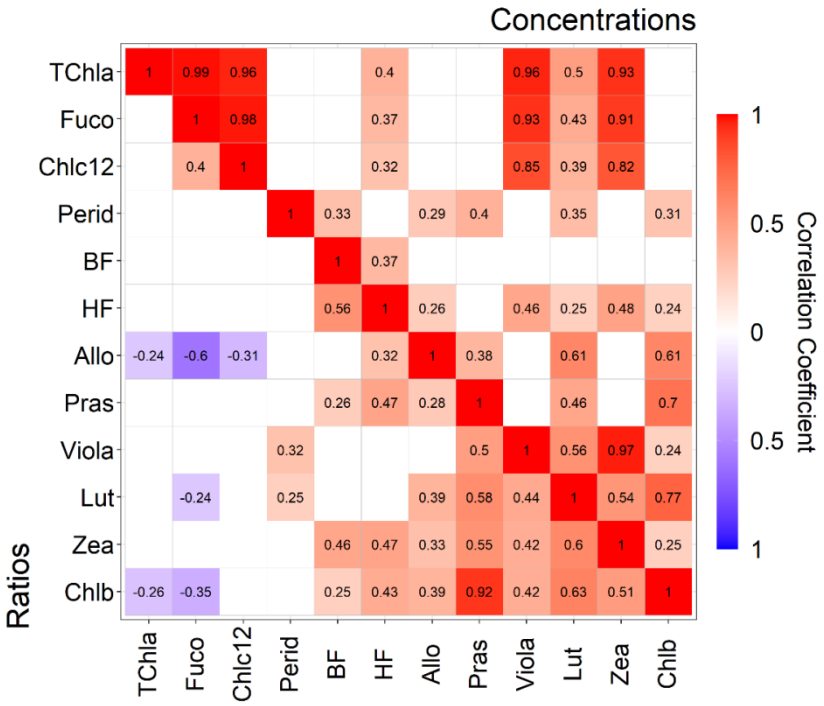

Figure S5. Correlation plot of Pearson correlations (R) between all pigments used for CHEMTAX analysis in the Strait of Georgia (SoG) dataset. Similar to Kramer and Siegal (2019), the top right corner shows correlations for absolute pigment concentrations (mg/m^3) while the bottom left corner shows correlations between pigments normalized to TChla (pigment:TChla) and between normalized pigments to TChla concentrations. Red colors represent positive correlations and blue colors represent negative correlations. Insignificant correlations ($p > 0.01$) were removed and represented by white squares.

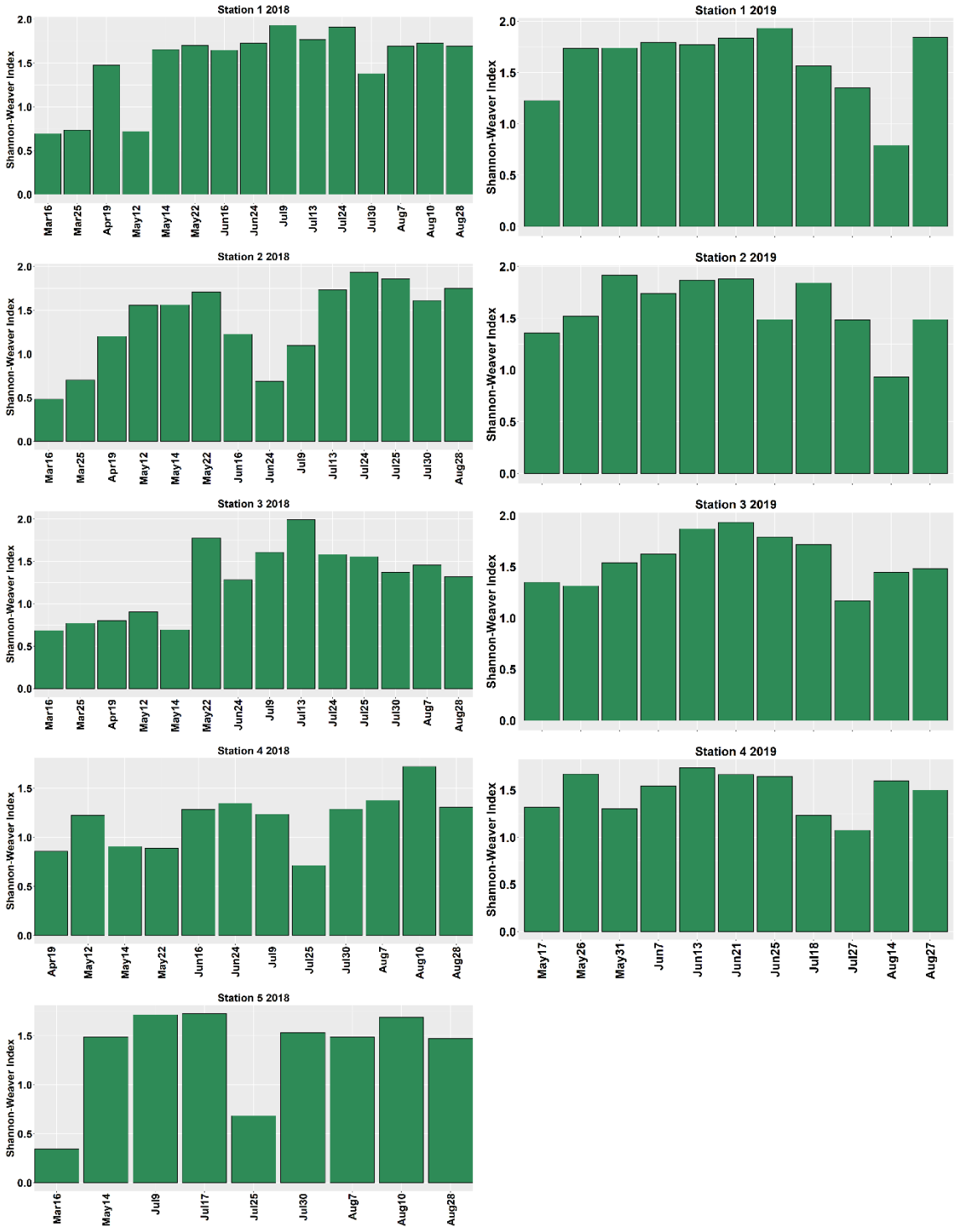

Figure S6. Shannon-Weaver Index for the stations (1 to 5) in the central Strait of Georgia during 2018 and 2019.

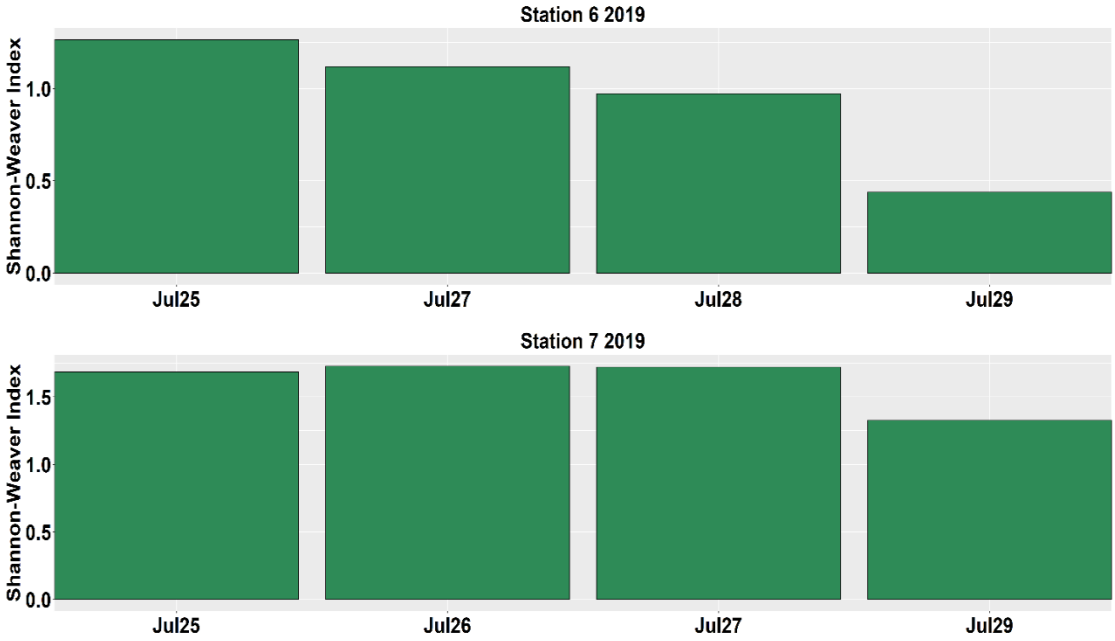

Figure S7. Shannon-Weaver Index for the stations (6 to 7) in the northern Strait of Georgia during 2019.

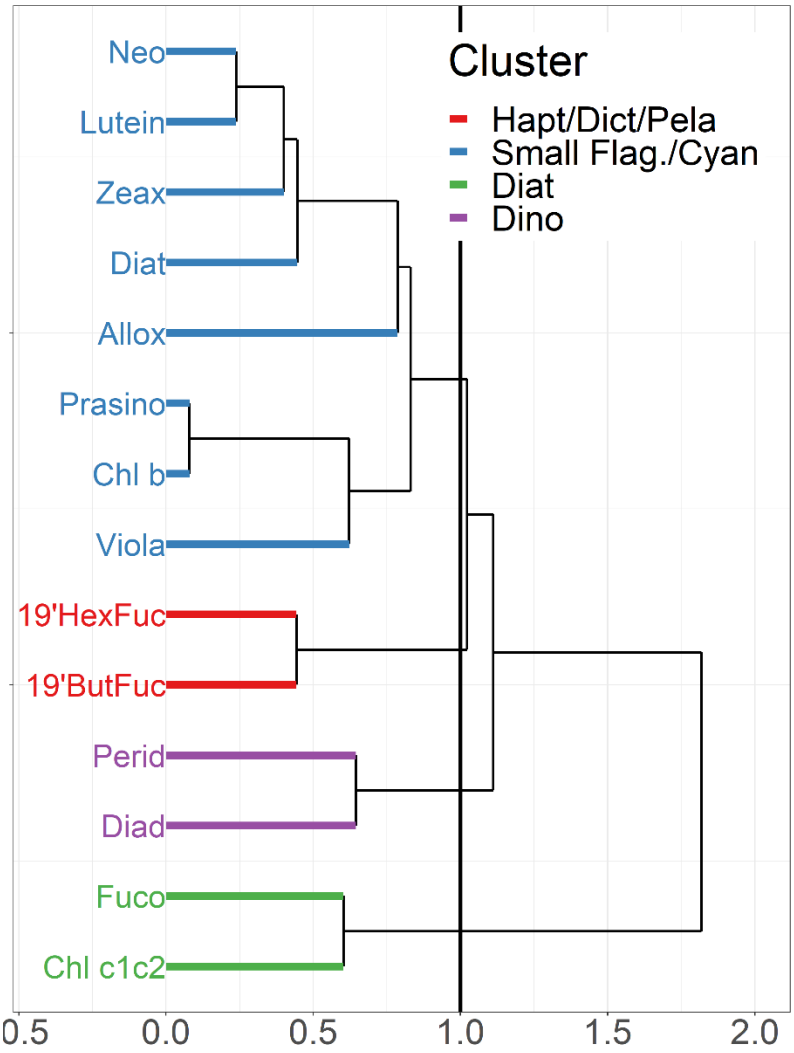

Figure S8. Hierarchical clustering of pigment:TChla ratios using the Pearson correlation distance and Ward's linkage method following Catlett and Siegal (2017); Kramer and Siegal (2019) and Kramer et al. (2020). The cophenetic correlation coefficient for the clustering was 0.84 indicating that the clustering was a good representation of the data (Kramer and Siegal (2019). To define clusters, a linkage distance of 1 was used following Kramer and Siegal (2019) and Kramer et al. (2020). Different colors are used to show the four clusters, which are made up of pigments representing different phytoplankton groups.

Ratios based on Peña et al. (2019) from Line-P.

Table S1. CHEMTAX output ratios for Cluster-1 and Cluster-2.

Group	Chlc3	$Chlc_{1+2}$	Peri	19BF	Fuco	19HF	Pras	Viol	Allo	Lut	Zea	Chlb	TChla
Cyan	0	0	0	0	0	0	0	0	0	0	0.60	0	1
Hapt	0.26	0.22	0	0.04	0.14	0.68	0	0	0	0	0	0	1
Pelago	0.23	0	0	0.77	0.62	0	0	0	0	0	0	0	1
Chlor	0	0	0	0	0	0	0	0.05	0	0.17	0.03	0.33	1
Pras	0	0	0	0	0	0	0.20	0.17	0	0.02	0.06	0.78	1
Cryp	0	0.18	0	0	0	0	0	0	0.39	0	0	0	1
Dino	0	0.21	0.68	0	0	0	0	0	0	0	0	0	1
Diat	0.07	0.28	0	0	1.18	0	0	0	0	0	0	0	1

*Ratios based on Peña et al. (2019)

Table S2. CHEMTAX output ratios for Cluster-3.

Group	Chlc3	$Chlc_{1+2}$	Peri	19BF	Fuco	19HF	Pras	Viol	Allo	Lut	Zea	Chlb	TChla
Cyan	0	0	0	0	0	0	0	0	0	0	0.60	0	1
Hapt	0.25	0.22	0	0.04	0.31	0.53	0	0	0	0	0	0	1
Pelago	0.22	0	0	0.59	0.76	0	0	0	0	0	0	0	1
Chlor	0	0	0	0	0	0	0	0.86	0	0.17	0.03	0.34	1
Pras	0	0	0	0	0	0	0.32	0.05	0	0.02	0.06	0.82	1
Cryp	0	0.21	0	0	0	0	0	0	0.37	0	0	0	1
Dino	0	0.23	0.58	0	0	0	0	0	0	0	0	0	1
Diat	0.08	0.28	0	0	1.14	0	0	0	0	0	0	0	1

*Ratios based on Peña et al. (2019)

Table S3. Full-fit versus Cross-validation statistics were obtained from the EOF analysis performed on the matchup between OLCI Rrs versus the merged CHEMTAX-derived Phytoplankton Community Composition.

	N	R^2	MAE (mg/m^3)	BIAS (%)	MDPD (%)	R^2cv	MAEcv	MDPDcv	p-value
Hapto/Pelago/Cyano	61	0.45	0.03	5.17	16.85	0.24	0.05	18.20	0.3710^{-8}
Diatom/Dino/GA	61	0.42	0.02	5.34	20.51	0.25	0.03	23.93	0.1410^{-8}
Crypto	59	0.24	0.01	26.37	43.49	0.21	0.01	44.74	0.7910^{-5}
TChla	61	0.27	0.04	2.33	12.42	0.23	0.01	13.85	0.1810^{-5}

Table S4: Chlorophyll-a concentration (mg/m^3) of Cryptophytes, Haptophytes/ Pelagophytes/Cyanobacteria, Diatoms/Dinoflagellates/Green algae, and TChla concentration obtained from weekly binned imagery of the Ocean Land Color instrument for the years 2019 and 2020. The corresponding dates are displayed on the spatial map.

Years	Phytoplankton groups	First Composite	Second Composite	Third Composite	Fourth Composite
2019	Crypto	[0 - 0.05]	[0 - 0.05]	[0 - 0.05]	[0 - 0.05]
		0.03±0.08	0.03±0.06	0.03±0.07	0.02±0.06
	Hapto/Pelago/	[0 - 0.4]	[0 - 0.4]	[0 - 0.4]	[0 - 0.4]
	Cyano	0.20±0.08	0.21±0.07	0.21±0.07	0.20±0.06
	Diatoms/Dino/GA	[0 - 0.3]	[0 - 0.3]	[0 - 0.3]	[0 - 0.3]
		0.12±0.13	0.11±0.09	0.11±0.10	0.11±0.09
	TChla	[0 - 1]	[0 - 1]	[0 - 1]	[0 - 1]
		0.35±0.09	0.36±0.07	0.35±0.07	0.34±0.06
2020	Crypto	[0 - 0.05]	[0 - 0.05]	[0 - 0.05]	[0 - 0.05]
		0.02±0.05	0.02±0.04	0.02±0.03	0.02±0.03
	Hapto/Pelago/	[0 - 0.4]	[0 - 0.4]	[0 - 0.4]	[0 - 0.4]
	Cyano	0.17±0.06	0.17±0.06	0.17±0.06	0.16±0.06
	Diatom/Dino/GA	[0 - 0.3]	[0 - 0.3]	[0 - 0.3]	[0 - 0.3]
		0.12±0.11	0.12±0.10	0.11±0.09	0.12±0.11
	TChla	[0 - 1]	[0 - 1]	[0 - 1]	[0 - 1]
		0.32±0.07	0.32±0.06	0.32±0.06	0.32±0.06

Table S5. Dates of suspected raphidophyte (Heterosigma akashiwo) blooms and accompanying pigment concentrations (mg m^{-3}) and ratios with TChla.

Date	Station	Viola (mg m^{-3})	TChla (mg m^{-3})	Fuco:Tchla	Viola:Tchla	Zea:TChla
2018-07-25	5	28.82	120.11	0.41	0.24	0.07
2018-07-25	4	7.19	28.82	0.36	0.25	0.03
2018-06-24	2	6.01	21.08	0.39	0.29	0.05
2018-07-09	4	1.84	9.81	0.27	0.19	0.05

Table S6. Input ratios for the CHEMTAX analysis in the central Strait of Georgia.

Group	$Chlc_{1+2}$	Peri	19BF	Fuco	19HF	Pras	Viol	Allo	Zea	Lut	Chlb	TChla
Cyan	0	0	0	0	0	0	0	0	0.64	0	0	1
Hapt	0.21	0	0.04	0.31	0.47	0	0	0	0	0	0	1
GA	0	0	0	0	0	0.25	0.05	0	0.06	0.02	0.70	1
Cryp[*1]	0.06	0	0	0	0	0	0	0.08	0	0	0	1
Dino	0.22	0.56	0	0	0	0	0	0	0	0	0	1
Raph[*2]	0.19	0	0	0.57	0	0	0.27	0	0.06	0	0	1
Dict	0.11	0	0.10[*1]	0.35	0	0	0	0	0	0	0	1
Diat	0.23	0	0	0.89	0	0	0	0	0	0	0	1

[*1]Derived from field ratios used by Belluz et al. (2021)
[*2]Ratios taken from Lewitus et al. (2005).

Table S7. Input ratios for the CHEMTAX analysis in the northern Strait of Georgia.

Group	$Chlc_{1+2}$	Peri	19BF	Fuco	19HF	Pras	Allo	Zea	Lut	Chlb	TChla
Cyan	0	0	0	0	0	0	0	0.64	0	0	1
Hapt	0.21	0	0.04	0.31	0.47	0	0	0	0	0	1
GA	0	0	0	0	0	0.25	0	0.06	0.01	0.70	1
Cryp	0.05	0	0	0	0	0	0.07	0	0	0	1
Dino	0.22	0.56	0	0	0	0	0	0	0	0	1
Dict	0.11	0	0.10	0.35	0	0	0	0	0	0	1
Diat	0.23	0	0	0.89	0	0	0	0	0	0	1

Table S8. ΔAIC predictor error obtained from the EOF model developed with the matchup between OLCI Rrs and CHEMTAX data.

OLCI Rrs	EOF1	EOF2	EOF3	EOF4	EOF5	EOF6	EOF7	EOF8	EOF9
TChla	16.06	17.75	6.29	36.97	11.30	NA	NA	NA	0.94
Diatom	NA	2.24	4.25	23.67	6.13	NA	5.76	NA	NA
Dictyo	NA	NA	NA	13.67	NA	NA	NA	1.68	2.56
Raphido	6.51	10.80	NA	7.75	14.97	1.93	NA	NA	0.09
Hapto	NA	15.02	NA	NA	NA	2.01	1.19	NA	NA
Prasino	7.72	0.80	0.76	4.11	NA	1.61	NA	NA	NA
Crypto	4.49	11.51	1.36	2.26	NA	NA	NA	0.44	NA
Dino	NA	NA	NA	NA	NA	NA	NA	4.80	NA
Cyano	NA	5.98	NA	1.37	NA	NA	NA	NA	NA

Table S9. Input ratios for the CHEMTAX analysis in the central Strait of Georgia.

Group	$Chlc_{1+2}$	Peri	19BF	Fuco	19HF	Pras	Viol	Allo	Zea	Lut	Chlb	TChla
Cyan	0	0	0	0	0	0	0	0	0.64	0	0	1
Hapt	0.21	0	0.04	0.31	0.47	0	0	0	0	0	0	1
GA	0	0	0	0	0	0.25	0.05	0	0.06	0.02	0.70	1
Cryp[*1]	0.06	0	0	0	0	0	0	0.08	0	0	0	1
Dino	0.22	0.56	0	0	0	0	0	0	0	0	0	1
Raph[*2]	0.19	0	0	0.57	0	0	0.27	0	0.06	0	0	1
Dict	0.11	0	0.10[*1]	0.35	0	0	0	0	0	0	0	1
Diat	0.23	0	0	0.89	0	0	0	0	0	0	0	1

[*1]Derived from field ratios used by Belluz et al. (2021)
[*2]Ratios taken from Lewitus et al. (2005).

Table S10. Input ratios for the CHEMTAX analysis in the northern Strait of Georgia.

Group	$Chlc_{1+2}$	Peri	19BF	Fuco	19HF	Pras	Allo	Zea	Lut	Chlb	TChla
Cyan	0	0	0	0	0	0	0	0.64	0	0	1
Hapt	0.21	0	0.04	0.31	0.47	0	0	0	0	0	1
GA	0	0	0	0	0	0.25	0	0.06	0.01	0.70	1
Cryp	0.05	0	0	0	0	0	0.07	0	0	0	1
Dino	0.22	0.56	0	0	0	0	0	0	0	0	1
Dict	0.11	0	0.10	0.35	0	0	0	0	0	0	1
Diat	0.23	0	0	0.89	0	0	0	0	0	0	1

Authors Contributions to Thesis Chapters

Chapter 2

In Chapter 2, I designed the research methodology, conceptualization, data curation, formal analysis, visualization and wrote the whole manuscript by myself. Maycira Costa helped to design the structure of the paper and reviewed it multiple times to improve the manuscript.

Chapter 3

In Chapter 3, I designed the research methodology, conceptualization, data curation, formal analysis, visualization, and wrote the entire manuscript by myself. Justin Del Bel Belluz helped with structuring the manuscript and writing, provided expertise in statistical analysis, and reviewed the draft multiple times to improve the manuscript. Angelica Peña reviewed the manuscript and provided comments and suggestions to improve it. Brian P.V. Hunt and Anna Vazhova provided physicochemical data, reviewed the manuscript, and provided comments and suggestions to improve it. Midhun Shah Hussain helped with developing the script for the application of the EOF algorithm to satellite imagery. Hongyan Xi and Astrid Bracher helped with the methodology, reviewed the manuscript and provided comments and suggestions to improve it. Maycira Costa helped structure the manuscript and reviewed the draft multiple times to improve it.

Chapter 4

In Chapter 4, I designed the research methodology, conceptualization, data curation, formal analysis, visualization, and wrote the entire manuscript by myself. Hongyan Xi helped with the methodology, reviewed the manuscript and provided comments and suggestions to improve it. Justin Del Bel Belluz helped with the methodology and writing, reviewed the manuscript multiple times, and provided comments and suggestions to improve it. Midhun Shah Hussain helped with developing the script for the application of the EOF algorithm to satellite imagery. Astrid Bracher reviewed the manuscript and provided comments and suggestions to improve it. Maycira Costa helped structure the manuscript and reviewed the draft multiple times to improve it.

Chapter 5

In Chapter 5, I designed the research methodology, conceptualization, data curation, formal analysis, visualization, and wrote the entire manuscript by myself. Justin Del Bel Belluz helped with structuring the manuscript and writing, provided expertise in statistical analysis, and reviewed the draft multiple times to improve the manuscript. Hongyan Xi helped with the methodology, reviewed the manuscript and provided comments and suggestions to improve it. Midhun Shah Hussain helped with developing the script for the application of the EOF algorithm to hyperspectral data. Astrid Bracher reviewed the manuscript and provided comments and suggestions to improve it. Maycira Costa helped structure the manuscript and reviewed the draft multiple times to improve it.