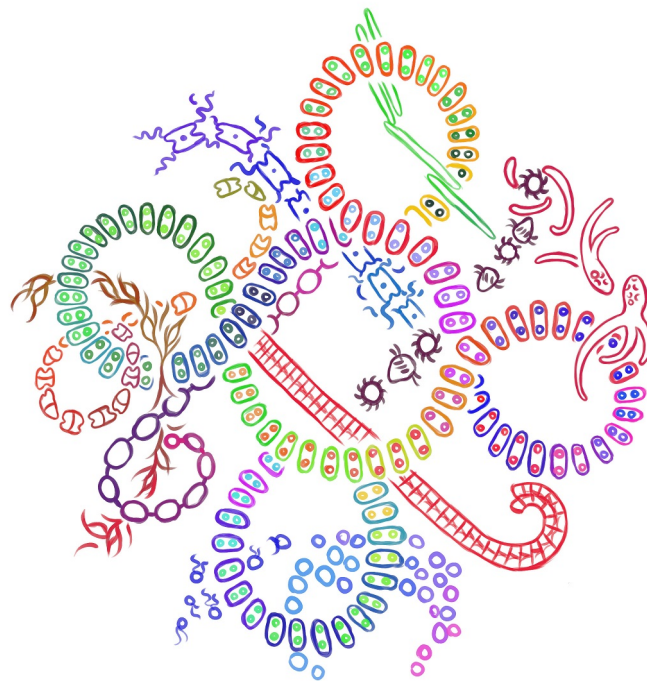


Discernment of phytoplankton groups from optical properties.



Andrew Orkney

October 11, 2021

Primary supervisor: Heather Bouman
Co-supervisors: David McKee, Helen Johnson

Department of Earth Sciences
St. Hugh's College, University of Oxford,
Doctor of Philosophy in Earth Sciences

Short Abstract

‘Discernment of phytoplankton groups from optical properties,’

The Barents Sea is a highly seasonal and productive shelf ecosystem, representing a gateway between the Atlantic and Arctic Oceans. Anthropogenic climate change is causing rapid changes in temperature, ocean chemistry, water column structure and sea-ice persistence. Sea-ice extents are declining twice as fast in the Barents Sea as other Arctic regions, establishing a mandate to investigate the impact of these environmental changes on productivity in the Seasonal Ice Zone. This is the remit of Arctic **PRIZE** (**PR**oductivity in the Seasonal **I**ce **Z**on**E**), and this thesis was conducted under the auspices of work package 3, studying phytoplankton compositional and phenological variability.

Remote-sensing is a key means by which ocean ecosystems are monitored and there have been recent efforts in the literature to employ ocean-colour to recognise different assemblages of phytoplankton. It is routine practice in the remote-sensing literature to distinguish functional assemblages from cell-counts and then seek bio-optical signatures to be targeted in remotely-sensed imagery. The analysis herein suggests that switching between different data types in this way introduces a translation error that undermines remote-detection efforts. We therefore recommend that functional assemblages are defined and identified directly from their bio-optical properties.

Thereafter, a scheme of bio-optical functional types is defined, based on Barents Sea phytoplankton communities, and a bio-optical algorithm is developed with the aim of retrieving this structure in remotely-sensed ocean-colour. This algorithm is applied to study community compositional change, facilitating the integration of previous *in-situ* studies towards a more complete understanding of incipient ecosystem change.

In addition, the ascendancy of temperate-mode autumnal phytoplankton blooms is investigated and relationships to evolving environmental conditions are explored, producing strong evidence that Atlantic inflow currents regulate the intensity of autumnal blooms in the Barents Sea.

In sum, growing Atlantic influence in the Barents Sea is found both to cause community compositional changes and concomitant phenological shifts that may change Barents Sea ecosystem function.

Long Abstract

‘Discernment of phytoplankton groups from optical properties,’

Marine Arctic ecosystems are among the most rapidly changing of Earth’s biological communities, owing to disproportionately rapid warming of high northern latitudes. Arctic ecosystems must contend with a variety of ancillary environmental changes in addition to rising temperatures. Anthropogenic warming is causing changes to Meridional Overturning Circulation, changing patterns of sea-ice persistence and broad shifts in prevailing physical and chemical conditions.

The majority of marine Arctic primary production is hosted on rich shelf systems, where relatively high nutrient provision and perennial spring sunlight coincide to produce suitable conditions for massive phytoplankton blooms. Of these shelf systems, the Barents Sea often hosts more marine primary production than any other Arctic shelf in a given year. The Barents Sea is a region typified by strong spatial gradients and temporal variability in hydrographic conditions. Its close proximity to intense Atlantic inflows render this shelf particularly sensitive to any changes to Global Ocean circulation. The Barents Sea is hence an area of utmost interest for the study of changing Arctic marine ecosystem function in the face of anthropogenic climate change – especially given the ecosystem services it provides as a fishery of great economic importance.

Yet, there is no single position in the scientific literature anticipating how this productive shelf ecosystem will change. Some researchers expect that a longer growing season and reduced sea-ice cover – hence greater illumination – will foster increased primary productivity, boosting sustenance to higher trophic levels. However, this view is not shared universally, because declining stocks of nutrients in the Nordic Seas may result in a shift towards recycling in the microbial loop rather than further new production.

The character of primary production and its likelihood to be exported depends on phytoplankton community composition and phenology, hence Arctic **PRIZE** (**PR**oductivity in the Seasonal **I**ce **Z**on**E**) has devoted a work-package to understanding the succession of phytoplankton assemblages and their dynamics in the Seasonal Ice Zone of the Barents Sea, with a view to improving understanding of this community’s baseline phenological and compositional variance. This should inform appraisals of the response of Arctic phytoplankton to evolving climatological conditions and the significance of this change for higher trophic levels and overall ecosystem function.

My Doctorate has concentrated on investigation of Phytoplankton Functional Types and the influence of methodological choice on their detection; development of bio-optical

algorithms for the discrimination of Phytoplankton Functional Types; and remote-sensing studies of secular change in the predominance and distribution of Phytoplankton Functional Types in the Barents Sea and shifts in their phenology.

My study of phytoplankton functional variety brings together measures of taxon-indicative pigments, cell-counts, flow-cytometry and bio-optical properties, exploring the extent to which an environmental dataset of hydrographic properties and nutrient availability constrains this variance. Multivariate statistics are applied to reveal functional assemblages. My analysis finds cell-counts and flow cytometry reveal a larger number of significant environmental relationships than any other community compositional dataset. This results from a succession of highly-abundant diatom species with distinct environmental niches. These different diatom species are, however, similar in their optical properties and pigment complements. Functional assemblages based on cell-count datasets may therefore prove difficult to detect as targets in remotely-sensed ocean-colour. By contrast, I discover a distinctive clustering of phytoplankton samples in bio-optical properties: diatoms, *Phaeocystis* colonies, nanophytoplankton and dinoflagellates. This form of variation corresponds closely to the concepts laid out by Margalef's mandala – a succession of phytoplankton functional assemblages proposed to covary with declining ocean turbulence following the onset of stratification. I find there may exist a realistic capacity to reliably detect these functional assemblages in remotely-sensed ocean-colour, as they possess distinct bio-optical signatures.

Therefore, I recommend that researchers take care to consider that functional structure may vary depending on the dataset type that is used to describe the phytoplankton. Furthermore, when mixed types of dataset are employed in validation, there is likely to be a 'translation bias'. This bias may be avoided by prioritising schemes of functional variety defined directly from bio-optical properties.

Having developed a scheme of bio-optically-informed Phytoplankton Functional Types I explore the feasibility of their detection in remotely-sensed ocean-colour, constructing a database of concomitant level-3 ocean-colour measurements of MODIS-Aqua with *in-situ* samples. I found it necessary to identify and remove ocean-colour data influenced strongly by either coccolithophore blooms, suspended sediment or dissolved humic substances in river-plumes, because these sources of variance dominated the dataset structure. It also proved necessary to remove the variance in ocean-colour attributable to varying concentrations of phytoplankton, which far exceeded the variation in ocean-colour attributable to differences in the composition of phytoplankton communities.

After these sources of variation were removed I was left with a database of residua, that could hypothetically be attributed to differences in phytoplankton community composition. From this database I built a Linear Discriminant Analysis and was able to show that phytoplankton groups with high pigment packaging, such as diatoms or *Phaeocystis*, could be reliably distinguished from nanophytoplankton. Cross-validation demonstrated that, when compared with other existing approaches, this Linear Discriminant Analysis performed favourably within the Barents Sea. I then applied an existing experimental algorithm, originally designed to retrieve signal of chlorophyll-c3 absorption from *in-situ* radiometry, and found convincing evidence that diatoms could be distinguished from *Phaeocystis* colonies, which are rich in chlorophyll-c3. I thus assembled a bio-optical algorithm to distinguish salient Phytoplankton Functional Types in the Barents Sea from remotely-sensed ocean-

colour.

I applied the ocean-colour algorithm to categorise Barents Sea functional types, synthesising a synoptic-scale record of the predominance and geographic distribution of these types in the Barents Sea during July, between 2002 and 2018. I chose the month of July because reports of increasing remotely-sensed primary production in the Barents Sea often cite a lengthened growing season and enhanced mid-summer productivity.

I was able to reproduce findings that reported an ongoing redistribution of coccolithophores further North-east in Atlantic inflow currents, and extended these by also finding a North-western redistribution in the West Spitsbergen Current, which had been predicted by previous studies. No significant change to the geographic distribution of diatom or colonial *Phaeocystis* blooms was found. However, inter-annual variation in *Phaeocystis* predominance was found and recapitulated the pattern reported by sporadic *in-situ* studies. In addition to this, the synoptic and temporally continuous nature of the ocean-colour data set allowed me to show that there has been a significant base-line increase in the dominance of *Phaeocystis*, in agreement with some predictions.

As *Phaeocystis* is routinely associated with low export production in contrast to diatoms, and higher regeneration of nutrients within the microbial loop, this finding has implications for how we envisage changing ecosystem structure and overall function in the Barents Sea.

Finally, I employed OC-CCI ocean-colour datasets and reanalysis products of Sea Surface Temperature, current velocities and wind fields, to investigate the relationship between energetic winds and the ascendancy of autumnal phytoplankton blooms in the Barents Sea.

Existing studies indicate that high winds should precipitate autumnal blooms by engendering vertical mixing that brings deeply sequestered nutrients to the surface ocean. It has been suggested that decreased sea-ice extents and a more energetic atmosphere have resulted in increasing autumnal blooms in the Arctic. This mode of change is of great interest because it recapitulates temperate modes of autumnal primary production, and therefore investigating autumnal blooms will improve understanding of phenological change in the Arctic in response to anthropogenic perturbation.

Contrary to expectations, I did not find a clear positive relationship between high winds and autumnal bloom intensity in the Barents Sea. Instead, the speed of Atlantic inflow underpins variability in autumnal bloom intensity. This finding emphasises the unique hydrographic set-up of the Barents Sea and the role of ascendant Atlantic influences in incipient phenological change.

Declaration

ANDREW ORKNEY

Student number: 662336 Candidate number: 554288

ST. HUGH'S COLLEGE

Supervisor: Dr. HEATHER BOUMAN

Title of thesis: DISCERNMENT OF PHYTOPLANKTON TYPES FROM OPTICAL PROPERTIES

DPhil Earth Sciences (Full-time)

Word count: 236 pages (Department of Earth Sciences regulations state thesis body must be under 250 pages.)

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism.

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism.

The thesis I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, either for this Honour School or qualification or for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis.

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted. I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Andrew Orkney

06-04-02021

Acknowledgements

I acknowledge the invaluable assistance of my primary supervisor, Heather Bouman, and my co-supervisors David McKee and Helen Johnson, for facilitating all aspects of the work that is represented in this thesis.

I acknowledge the guidance and technical input of Trevor Platt and Shubha Sathyendranath.

I acknowledge Jo Hopkins, Finlo Cottier and Jørgen Berge for overseeing the organisation of field sampling missions.

I acknowledge Elaine Mitchell for teaching optical microscopy techniques to me.

I acknowledge Sharon Mcneill for cross-checking instrument calibrations with me.

I acknowledge Duncan Purdie and John Gittins for training me in HPLC processing techniques.

I acknowledge Sian Henley for organising Arctic **PRIZE** research meetings.

I acknowledge Bhavani Narayanaswamy for providing insightful discussion that was pivotal in establishing some of my research directions.

I acknowledge my twin, Matthew Orkney, for proof-reading parts of my thesis.

I acknowledge Erin Saupe for acting as my college supervisor.

I acknowledge Roger Benson, Alex Bjarnason, Neil Brocklehurst, Isabel Fenton and Brooke Johnson for providing insightful discussion of statistical methods and research approaches. Research institutions, facilities and funding bodies are acknowledged in sections at the end of each appropriate chapter.

Abbreviations

Abbreviation	Expansion	Definition
AIC	Akaike Information Criterion	Statistical representation of model prediction error employed in model selection.
a_{c3}	Chlorophyll-c3 absorption	Absorption attributable to phytoplankton pigment Chlorophyll-c3 [m^{-1}]
$a_{ph}(\lambda)$	Wavelength-dependent (λ) phytoplankton (ph) absorption (a)	Absorption spectrum attributable to absorption by phytoplankton [m^{-1}]
$a_{ph}^*(\lambda)$	Chlorophyll-specific absorption	Absorption spectrum attributable to absorption by phytoplankton at specified [<i>chl-a</i>]. [$\text{m}^2 \text{mg}^{-1} \text{chl-a}$]
AUC	Area Under Curve	Diagnostic capability statistic derived from ROC curve (see ROC)
b_b	Particulate backscattering	Portion of light scattered-back by suspended particles such as phytoplankton cells. [m^{-1}]
CDOM	Coloured Dissolved Organic Matter	A complex mix of many organic chemical species, such as decaying matter or humic substances, that absorb strongly at blue wavelengths.
CGF	Component Gaussian Function	Gaussian curves derived from spectral deconstruction of complex function.
[<i>chl-a</i>]	Chlorophyll-a concentration	[mg m^{-3}]
CTD	Conductivity Temperature Depth rosette	Sampling device to perform hydrographic profiling and collect seawater.
DMS	dimethyl-sulphide	An aerosol species produced by some phytoplankton species, that plays a key role nucleating cloud formation.
f -ratio		Ratio of primary production supported by inorganic nitrogen sources. (dimensionless)

GLS	Generalised Least Squares	Procedure for performing linear regressions.
HNAN	Heterotrophic Nanoflagellates	
HPLC	High Performance Liquid Chromatography	An analytical method here used to measure concentrations of taxon-indicative phytoplankton pigments.
Lab-OC4L diatom	Labrador-Sea diatom algorithm	Algorithm for inference of $[chl-a]$ from ocean-colour, assuming diatom-dominated water column.
Lab-OC4L prymn	Labrador-Sea prymnesiophyte algorithm	Algorithm for inference of $[chl-a]$ from ocean-colour, assuming prymnesiophyte-dominated water column.
LDA	Linear Discriminant Analysis	Statistical approach to find linear combinations of variables that divide classes.
MODIS	Moderate Resolution Imaging Spectroradiometer	Remote-sensing satellite system.
OC-CCI	Ocean Colour Climate Change Initiative	Merged remote-sensing ocean-colour product.
E_{PAR}	Irradiance (E) from Photosynthetically Active Radiation (400-700 nm)	Spectrally-adjusted measure of energy available to drive photosynthesis. $[\mathbf{W\ m^{-2}}]$
PCA	Principal Component Analysis	Statistical procedure to explore variance structure of datasets.
PFT	Phytoplankton Functional Type	An assemblage of phytoplankton defined by a common shared function in ecosystem or biogeochemical cycling.
PNAN	Phototrophic Nanoflagellates	
PPC	Photoprotective carotenoid	A family of carotenoid pigments employed to shield photosystems from intense light. $[\mathbf{mg\ m^{-3}}]$
PSU	Practical Salinity Units	Mass proportion of seawater occupied by salts. (dimensionless)
σ_θ	Potential density anomaly	Density of fluid parcel if brought to surface pressure adiabatically, compared with surface value. $[\mathbf{kg\ m^{-3}}]$
RDA	Redundancy Analysis	Statistical method to perform regressions between multiple dependent and independent variables.

ROC	Receiver Operator Characteristics	Statistical convention to represent diagnostic capability of a binary classifier.
R_{rs}	Remote-sensing reflectance	Ratio of water-leaving to down-welling irradiance perceived at the top of the atmosphere. (dimensionless)
R_{rsm}	Magnitude-normalised remote-sensing reflectance	wavelength-specific R_{rs} values normalised to sum of R_{rs} at all wavelengths. (dimensionless)
$R_{rsm}^{e,(d,p)}$	Expected spectral shape	R_{rs} spectral shape expected (e) under assumption of diatom (d) or prymnesiophyte (p)-dominated water columns, respectively. (dimensionless)
R_{rsRe}	Residual spectral shape	The residual (Re) difference between measured and expected R_{rs} spectral shape. (dimensionless)
SeaWiFS	Sea-viewing Wide Field of view Sensor	Remote-sensing satellite system.
SCM	Sub-surface Chlorophyll Maximum	$[chl-a]$ concentrations at depth exceeding surface values.
SST	Sea Surface Temperature	$[^{\circ}\text{C}]$ when absolute values are stated, and presented in Kelvin in mathematical analyses.

Contents

1	Introduction	1
1.1	The Barents Sea and its context in an evolving Arctic hydrography	1
1.1.1	Plain Language Summary	1
1.1.2	Gross hydrography	2
1.1.3	Intricate environmental heterogeneity	6
1.1.4	Atlantification	10
1.2	Phytoplankton of the Barents Sea and the wider ecosystem	11
1.2.1	Pelagic modes of production	14
1.2.2	Sympagic modes of production and ‘hidden’ blooms	20
1.2.3	Trophic cascades: energy transfer through the Barents Sea ecosystem	24
1.3	Ecosystem change in the Barents Sea: unresolved questions	33
1.3.1	<i>In-situ</i> studies	35
1.3.2	Remote-sensing and modelling studies	38
1.3.3	The need to represent functional variety of phytoplankton in remote- sensing studies	43
1.3.4	Arctic PRIZE	44
1.3.5	Arctic PRIZE in a historic and contemporary context	44
1.4	Thesis overview	46
2	Describing an ecosystem: Phytoplankton Functional Types	60
2.1	Prelude	61
2.2	Significance statement	61
2.3	Abstract	63
2.4	Plain Language Summary	64
2.5	Introduction	65
2.6	Materials	69
2.6.1	Data collection	69
2.6.2	Environmental data	71
2.6.3	Phytoplankton community structure	72
2.6.4	Data Analysis	79
2.7	Results	86
2.8	Discussion	91
2.8.1	Influence of data type on perception of functional relationships	91

2.8.2	Capacity of community composition data sets to validate remote-sensing algorithms based on bio-optical variation	94
2.9	Conclusion	95
2.10	Data availability statement	96
2.11	Author contributions	98
3	Development of a Bio-Optical algorithm	105
3.1	Prelude	106
3.2	Abstract	107
3.3	Plain Language Summary	108
3.4	Introduction	109
3.5	Fundamentals of remote-sensing	112
3.6	Taxonomic algorithms	115
3.7	Collection of phytoplankton data	118
3.8	Identification of distinct bio-optical phytoplankton communities in <i>in-situ</i> samples	121
3.8.1	Ocean-colour data	131
3.8.2	Masking unsuitable ocean-colour observations	132
3.9	Relation of variability in ocean-colour to <i>in-situ</i> bio-optical communities	135
3.9.1	Discrimination of diatoms and <i>Phaeocystis</i>	145
3.10	Summary of resulting bio-optical Algorithm structure	151
3.11	Correction for drift in MODIS-Aqua ocean-colour data	154
3.12	Review of bio-optical algorithm classifications and concluding remarks	158
4	Bio-Optical evidence for increasing <i>Phaeocystis</i> dominance in the Barents Sea	169
4.1	Prelude	170
4.2	Abstract	171
4.3	Plain Language Summary	172
4.4	Introduction	172
4.5	Methods	180
4.6	Results	184
4.7	Discussion	192
4.7.1	Phytoplankton community shift in the context of “Atlantification”	192
4.7.2	Implications	194
4.7.3	Limitations to our experimental bio-optical algorithm	196
4.8	Conclusions	197
5	Atlantic inflow the primary driver of fall blooms in European Arctic	205
5.1	Prelude	206
5.2	Abstract	208
5.3	Plain Language Summary	208
5.4	Introduction	209
5.5	Dataset and methods	212
5.5.1	Satellite data	212
5.5.2	Environmental data	213
5.5.3	Analysis	214

5.6	Results	217
5.7	Discussion	222
5.7.1	Limitations	226
5.8	Summary	228
5.9	Author contributions	228
5.10	Supplement	229
5.10.1	Coccolithophore Mask	229
5.10.2	Preliminary analysis	230
6	Conclusion and future work	238
6.1	Conclusion	238
6.2	Future work directions	243
6.2.1	Harmful Algal Bloom detection in Norwegian waters	243
6.2.2	Community composition of autumnal blooms	250
6.2.3	The ascendancy of machine learning and the integration of measurement platforms	254
7	Appendix: Arctic PRIZE	264
7.1	Project outline	264
7.2	Transects	268
8	Appendix: Mathematical methods	278
8.1	Opening	278
8.2	1 Principal Component Analysis	279
8.3	2 Two Block Partial Least Squares	289
8.4	3 Redundancy Analysis	294
8.5	4 Linear Discriminant Analysis	303
8.6	5 Fuzzy Clusters	312
8.7	6 Temporal Autocorrelation	318

List of Figures

1.1	Map of Arctic Ocean Circulation	3
1.2	Typical watercolumn structure	4
1.3	Map of Barents Sea and associated circulation	8
1.4	Pan-Arctic Primary Production	12
1.5	Global Primary Production	13
1.6	Plankton illustration	15
1.7	Cartoon progression of blooms	17
1.8	Photographs of sympagic algae	21
1.9	Ecosystem interactions network	31
2.1	Sampling map with hydrographic properties	70
2.2	Algal classes HPLC	74
2.3	Bio-optical assemblages	80
2.4	Redundancy analyses	84
2.5	Comparison of redundancy analyses	85
2.6	ROC plots	90
3.1	Schematic of remote-sensing physics	114
3.2	Sampling map	119
3.3	Component Gaussian Functions	123
3.4	2 Block Partial Least Squares analysis	125
3.5	Phytoplankton Bio-optical assemblages	128
3.6	Hyperspectral absorption normalised to chlorophyll-a	130
3.7	Masking of invalid ocean-colour data	134
3.8	Validation of coccolithophore detection	135
3.9	Regression of ocean-colour with evolving chlorophyll	138
3.10	Example algorithm application	142
3.11	Contrast to existing algorithms	143
3.12	Retrieval of chlorophyll-c3 signal	149
3.13	Example application of algorithm	150
3.14	Bio-optical algorithm flow chart	153
3.15	Temporal evolution of Chlorophyll-c3 signal	155
3.16	Temporal drift in r2018.1 reprocessing	157
3.17	Labrador Sea algorithm application	160
3.18	Barents Sea algorithm application	161
4.1	Map of the Barents Sea	173

4.2	<i>Phaeocystis</i> algorithm validation	183
4.3	Patterns of geographic chance in Phytoplankton Functional Types	185
4.4	Temporal evolution of Phytoplankton Functional Types	187
4.5	Geographic patterns of changing phytoplankton assemblages	190
4.6	Synthesis cartoon	191
5.1	Map of Barents Sea in Global context	212
5.2	Map of salient variables for study	215
5.3	Linear trends in conditions and autumnal bloom intensity	218
5.4	Linear relationships to autumnal bloom intensity	219
5.5	Relationships between autumnal bloom intensity and Atlantic Waters	221
5.6	Coccolithophore mask	230
5.7	All variables	232
6.1	Chlorophyll-c3 detection	244
6.2	Harmful Algal Bloom	245
6.3	Ichthyotoxic bloom cartoon	249
6.4	Autumn sampling campaign map	252
6.5	Autumn sampling variability	253
6.6	Integration of diverse data streams	259
7.1	Map of sampling locations	266
7.2	CTD Temperature	268
7.3	CTD Salinity	269
7.4	CTD TON	270
7.5	CTD Si	271
7.6	CTD PAR	272
7.7	CTD in-vivo Chlorophyll fluorescence	273
7.8	CTD Turner-derived Chlorophyll fluorescence	274
7.9	CTD Oxygen	275
7.10	CTD Ammonium	276
8.1	Bread buns	281
8.2	Bread buns PCA	284
8.3	Ocean Colour PCA	286
8.4	Ocean Colour PCA v Chl	287
8.5	Ocean Colour PCA v True Colour	288
8.6	coccolithophore mask v True Colour	289
8.7	Weather Barents Sea 2002	292
8.8	Salience 1 of 2BPLS	293
8.9	Salience 2 of 2BPLS	294
8.10	Redundancy Analysis example	298
8.11	Redundancy Analysis Dendrogram	300
8.12	Redundancy Analysis Tanglegram	301
8.13	Redundancy Analysis contrasts	302
8.14	Northern and Southern Barents map	307
8.15	LDA confusion matrix example	308
8.16	LDA posterior probabilities	309

8.17	LDA coefficients	310
8.18	LDA confusion matrix example 2	311
8.19	Fuzzy clustering of phytoplankton absorption spectra	315
8.20	Example mean absorption spectra	316
8.21	Example of fuzzy cluster failure	317
8.22	Autocorrelated time series example	321
8.23	Time series of salient variables	323
8.24	Geographic relationship sst and wind	324
8.25	Geographic relationship sst and open water area	325
8.26	Geographic relationship sst and open water area	326
8.27	Geographic relationship sst and open water area	327

List of Tables

3.1	PC loadings, bio-optical assemblages	129
3.2	Confusion matrix	140
3.3	Pigment packaging discriminant coefficients	145
3.4	Inherent Optical Properties	146
5.1	GLS model results	217
5.2	Coccolithophore mask coefficients	230
7.1	Cruise durations	265
8.1	Bread bun rotation matrix	283
8.2	Covariance matrix	290
8.3	Left PLS vectors	293
8.4	Right PLS vectors	293
8.5	Unconstrained variance RDA	299

Chapter 1

Introduction

1.1 The Barents Sea and its context in an evolving Arctic hydrography

1.1.1 Plain Language Summary

The Barents Sea is a transitional region, representing the gateway between the interior Arctic Ocean and the North-east Atlantic.

Climatological and hydrographic conditions in the Barents Sea are dominated by the influence of water-masses from these adjacent regions; the South-western Barents Sea does not experience seasonal ice cover, and the physical structure of the water column is determined by temperature, while the North-east experiences recurrent growth of a thin seasonal ice cover and water column structure is dominated by vertical changes in salinity.

This physical regime means that overall stratification on the Barents shelf is weak, compared to the rest of the Arctic. This means that nutrient-rich Atlantic waters, which enter the shelf in the South-west, and which are generally overlain by fresher water-masses, are readily mixed up to the surface over winter. As a consequence, when ice-melt and sunshine induce physical stratification in spring, phytoplankton are supplied with plentiful light and

nutrients; supporting transient blooms. These blooms are so rich that they place the Barents Sea among the most productive shelf seas of the Arctic Ocean, as well as being a productive sea by global standards.

Environmental conditions in the Barents Sea may vary considerably from year to year, resulting in different assemblages of phytoplankton or different intensities of phytoplankton blooms between years; some assemblages rely on sea-ice, while others prefer open-ocean, for example. Some phytoplankton species may require specific nutrients, be preferentially eaten by different consumers or sink at different rates. The style of phytoplankton primary production therefore has significant consequences for the way the Barents Sea ecosystem functions, but understanding of these relationships is incomplete.

The hydrography of the Barents Sea is rapidly evolving in response to climate change. There is increasing Atlantic influence and waning sea-ice formation, which may result in overall changes to the types of phytoplankton assemblage that predominate, amounts of primary production that are supported and its transfer through the Barents Sea food-web.

There is therefore a mandate to increase knowledge of baseline variability in Barents Sea phytoplankton assemblages and its relationship to environmental change, so that wider change to the Barents Sea ecosystem can be better anticipated. This will require the integration of *in-situ* sampling campaigns and synoptic-scale remote-sensing to monitor the occurrence of different phytoplankton assemblages in the Barents Sea.

1.1.2 Gross hydrography

The Arctic Ocean is a Mediterranean-type sea typified by extensive shallows on the continental shelves of North America and Eurasia, which enclose deep central basins (Carmack and Wassmann 2006) (Figure 1.1).

The entire region functions as a double estuary, wherein waters from the North Atlantic and Pacific enter over the shelves and are freshened by continental run-off and sea-ice melt

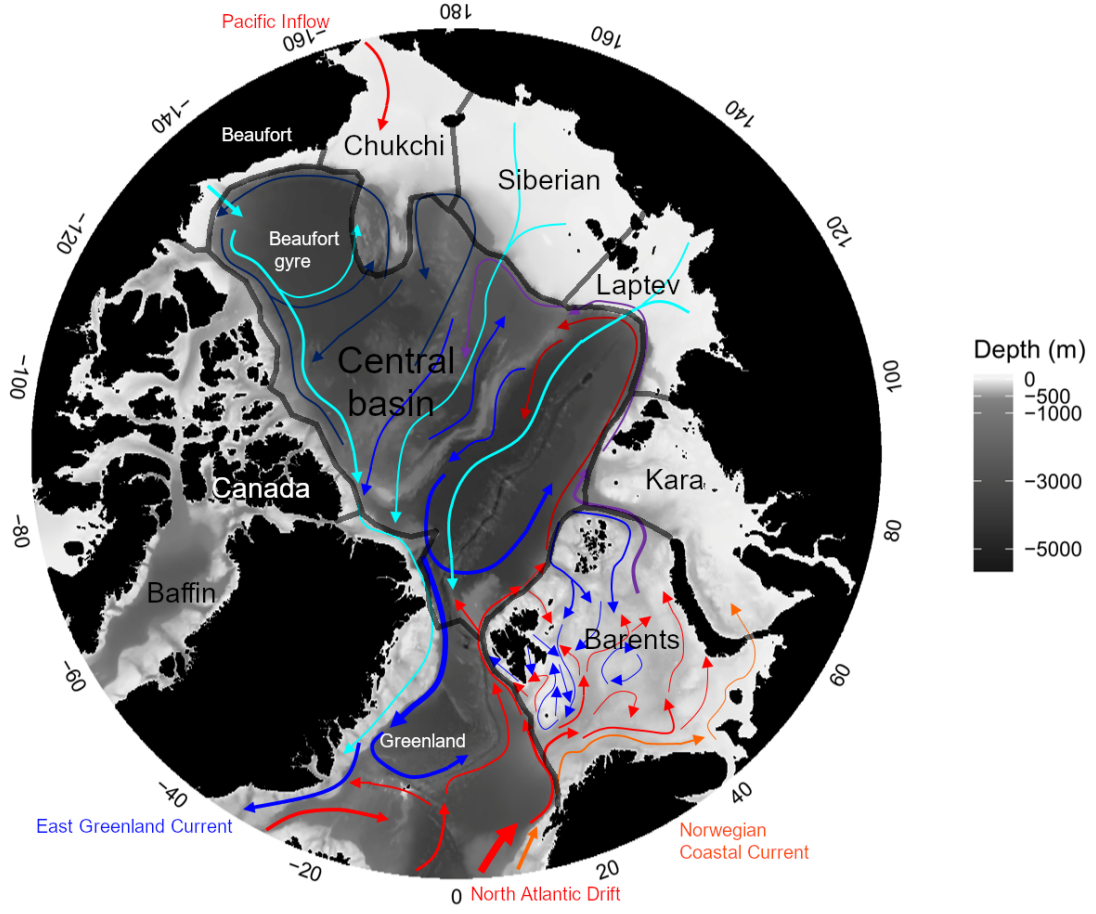


Figure 1.1: Arctic Ocean Circulation, with higher resolution in the Barents Sea. Red arrows represent warm saline inflows, orange represents the Norwegian Coastal Current, purple represents modified Barents Sea Water, blue represents Arctic water-masses and cyan represents fresh water contributed from continental run-off from Siberian and North American rivers. Bathymetry is represented at 4km^2 resolution, constructed by querying the ETOPO1 bathymetric database, hosted by NOAA, with the R *marmap* package (Pante and Simon-Bouhet 2013). The major circulation patterns of the Arctic Ocean were reproduced from Björk et al. 2018. Circulation currents of the Barents Sea reproduced and simplified with reference to Grøsvik et al. 2018. Regional boundaries reproduced from Lewis et al. 2020.

in summer; and lose buoyancy through brine-rejection and radiative cooling in winter (Carmack 2007). This confluence of water-masses and complex physical processes results in the

Arctic Ocean hosting a diverse array of environments, sometimes separated by intense spatial gradients in conditions. The Barents Sea, placed at the juncture between the interior Eurasian shelves, central Arctic Ocean and Norwegian Sea, is one such transitional region. The Barents Sea functions as the gateway for inflowing Atlantic Water delivery to the Arctic. Wind patterns and variability in the Barents Sea control the volume and route of warm, saline Atlantic Water admitted into the Arctic Ocean (Lien et al. 2013), establishing the Barents Sea as a region of great interest in the regulation of Arctic hydrography.

The comparatively high salinity of inflowing Atlantic Waters ($>34.8\text{‰}$ Våge et al. 2016) results in their subduction beneath a ‘polar-lid’ of Arctic-derived water-masses and sea-ice melt (see Figure 1.2).

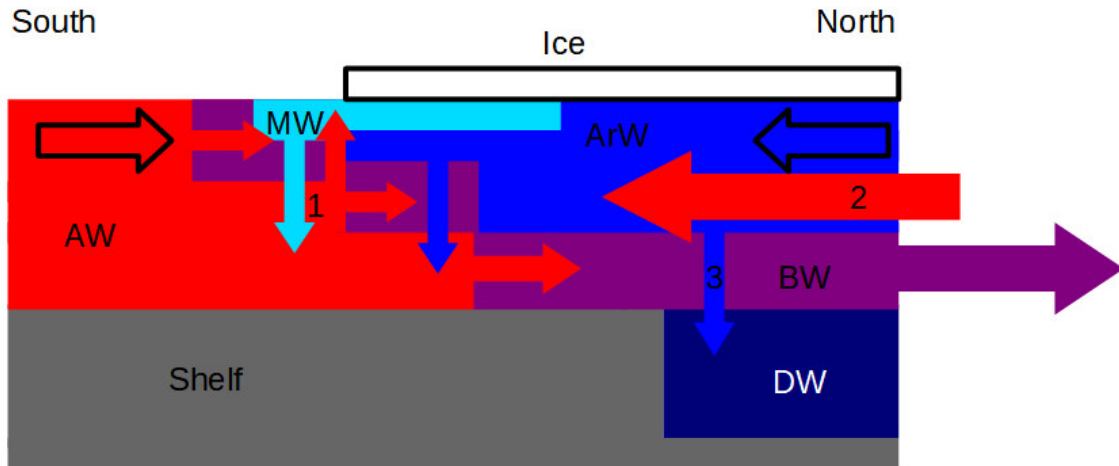


Figure 1.2: Cartoon meridional section of water-mass distribution over the Barents Shelf. AW: Atlantic Water, ArW: Arctic Water, BW: Barents Water, MW: Melt Water, DW: Deep Water. 1 Transient double diffusive convection and salt-fingering can engender vertical mixing in spring. 2 Ephemeral retrograde flow of Atlantic Waters following shelf-break. 3 Episodic Deep Water formation in autumn and winter.

This results in a salinity-stratified water-column dominating over the Northern Barents

Sea and much of the interior Arctic Ocean (Schauer et al. 2002) (see Figure 7.3). High-latitude seas that are stratified by salinity are known as β^1 oceans, in opposition to the thermally-stratified α^2 oceans (Carmack 2007). Inflowing Atlantic Water is modified by mixing with Arctic Waters and radiative cooling to form Barents Water, which exits the shelf into the Kara Sea and continues in an anticyclonic circulation across the Eurasian shelf seas, eventually even bringing Atlantic influence to the Canada Basin (Rudels et al. 1994). The delivery of Atlantic Water and its subduction underneath surface Arctic Waters is crucial to the hydrographic and ecosystem set-up of the Barents Sea because Atlantic Water is replete in macronutrients³ essential to support primary production⁴ (Oziel et al. 2016; Slagstad and Støle-Hansen 1991).

The Barents Sea’s unique situation as a transitional shelf sea underpins its capacity to support high levels of primary production. While the Barents Sea is located entirely North of the Arctic circle, its proximal position to warm Atlantic inflows (Figure 1.1) brings a moderating influence. As a result, the Barents Sea represents a high-latitude boundary between α and β type ocean, (Carmack 2007) with temperature-stratified waters pervading the Southern and western portions of the shelf and salinity-stratified waters dominant in the far North (see Figure 7.2). While, in the central Arctic strong haline stratification inhibits the supply of exogenous nutrients to the surface, the relatively shallow bathymetry of the Barents Sea (averaging only 230 m) and its weaker stratification mean that deep nutrient-replete water-masses can be resupplied to the surface layer more easily. Typically, surface waters are recharged with nutrients as a result of winter mixing that penetrates to hundreds of metres when the seasonal thermocline⁵ is relaxed over the polar night (Loeng 1991) (see Figure 7.4). High salinity ($>34.8\text{‰}$, Våge et al. (2016)), combined with low

1. Where β refers to the density-dependence on salinity.
2. Where α refers to the thermal expansion coefficient.
3. Nutrients that are usually present at micromolar concentrations.
4. ‘Primary production’ refers to the synthesis of organic molecules, such as phytoplankton tissues, from CO_2 .
5. An abrupt change in temperature with depth, separating the surface and interior ocean.

amounts of fresh riverine input⁶, low autumnal ice-cover⁷, and the relaxation of physical stratification in winter, result in the extensive densification of Barents Water and the formation of Deep Waters. This deep dense water forms a cyclonic overflow exiting South out of the Nordic Seas. This deep return flow combines with other sources of Deep Water, completing the Atlantic Meridional Overturning Circulation (AMOC) (Buckley and Marshall 2016). Hence, to a first order, the Barents is a highly-seasonal epeiric Arctic Sea with a South-west North-east gradient in hydrographic conditions resulting from the confluence of Atlantic and Arctic-derived water-masses, constituting an important component in broader meridional patterns of ocean circulation. The unique hydrographic conditions that the Barents Sea experiences, as a result of its physical situation, result in comparatively high nutrient availability, which results in high primary production. Changes to the the physical and hydrographic set-up of the Barents Sea are therefore of paramount interest because of their capacity to cause cascades in physical conditions through the rest of the Arctic and because of the potential changes to Arctic primary production that might occur as a result of an altered Barents Sea.

1.1.3 Intricate environmental heterogeneity

Spatial variability

An intricate pattern of recirculations is super-imposed on the gross hydrography of the Barents Sea. These currents are steered by bathymetric features (Figure 1.1) and their effects can result in latitudinally-reversed gradients in temperature, antithetical to the broader patterns of hydrography across the shelf. For example, the West Spitsbergen Current – the pre-eminent source of salt and heat into the central Arctic (Mosby 1962; Schlichtholz and Houssais 1999) – separates from its parent, the North Atlantic Drift, before following the

6. Total Arctic run-off totals ~ 0.1 Sv, and is linearly proportional to catchment area. The Barents Sea is surrounded by relatively little land, resulting in only minor run-off (Prowse and Flegg 2000).

7. Facilitating high levels of fresh sea-ice formation in North-eastern polynas (Martin and Cavalieri 1989).

edge of the Norwegian shelf through the Fram Strait (Lien et al. 2013). The West Spitsbergen Current then branches into a series of smaller currents, flowing North over the Yermak Plateau on into the Nansen Basin, East towards the Svalbard Archipelago, or deviating South-west in an Atlantic return flow (see Figure 1.3 and Figure 7.2).

As a consequence of recirculating ocean currents, the North-western coast of Spitsbergen is almost permanently ice-free, whilst Sørkapp – the Southern most point of the Svalbard Archipelago – may experience perennial sea-ice cover and the ice-edge historically extended as far South as Bjørnøya (Shapiro et al. 2003). In general, longitudinal gradients are stronger than latitudinal variation in Sea Surface Temperature across the Svalbard Archipelago.

Spatial changes in hydrographic conditions can be sharp at the juncture of Atlantic and Arctic water-masses. Bathymetry, again, plays an important role in this aspect of the Barents Sea’s hydrographic regime. Topographic features, such as the Spitsbergen Bank, act to constrain the position of the Polar Front that forms between Atlantic and Arctic water-masses (Oziel et al. 2016). Shallow topographic features and frontal hydrographic regions also interact to create regions of enhanced vertical upwelling of deeper waters by prompting diapycnal⁸ mixing over rough bathymetry (Wassmann et al. 1999).

8. To occur across potential density gradients.

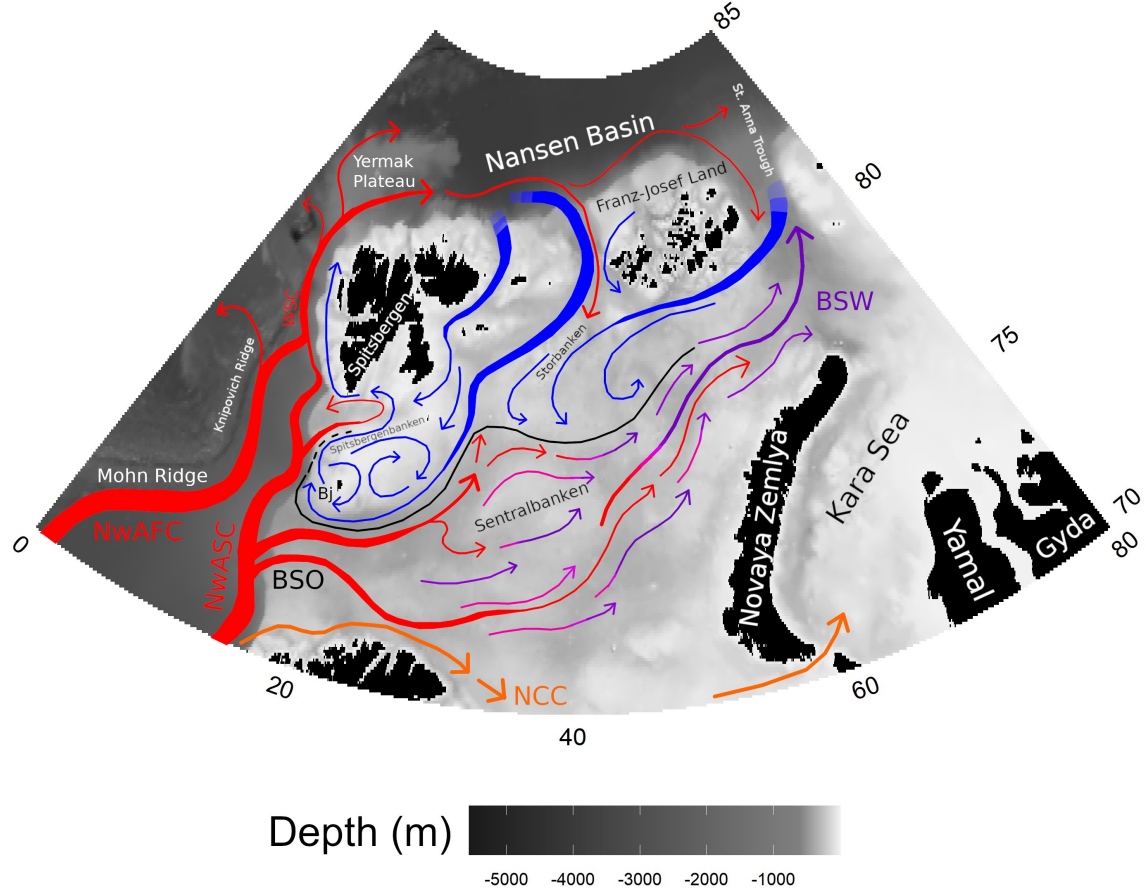


Figure 1.3: Map of the hydrographic set-up of the Barents Sea: Bj– Bjørnøya, BSO– Barents Sea Opening, BSW– Barents Sea Water, NCC– Norwegian Coastal Current, NwAFC– Norwegian-Atlantic Frontal Current, NwASC– Norwegian-Atlantic Slope Current, WSC– West Spitsbergen Current. Red currents represent warm, saline Atlantic-derived inflows. Blue currents represent colder, fresher Arctic-derived Waters. The Polar Front, represented by a black line, divides these water-masses. The Norwegian Coastal Current (orange) carries a signal of riverine input. The mixing of Atlantic and Arctic-derived water-masses, and the alteration of Atlantic Waters through radiative cooling and the contribution of Melt Waters, results in the genesis of Barents Sea Waters (purple) which flow on into the Kara Sea and down-stream Eurasian shelf.

Temporal variability

In addition to intense seasonality and geographic shifts in conditions, hydrography in the Barents Sea has a significant component of temporal variability. On short time-scales, water-mass exchange across sharp spatial gradients in hydrography can be achieved by eddies (Porter et al. 2020) or by episodic mixing caused by double-diffusive processes and salt-fingering⁹ in the Marginal Ice Zone¹⁰ (May and Kelley 2001; Sundfjord et al. 2007). On an inter-annual time-scale the Atlantic inflow's strength, and the routes it takes onto the Barents Shelf, respond to variation in the state of the North Atlantic and the Atlantic Multidecadal Oscillations (Lien et al. 2013; Oziel et al. 2016). Indeed, anthropogenic radionuclide pollution in North-west Europe is an effective tracer of Atlantic inflow to the Arctic circulation. Measurements and modelling studies of radionuclide distribution converge to show that shifts in the state of the Arctic Oscillation¹¹ drive changes in the Nordic Seas, which propagate Eastwards over the Eurasian Basin. These cascading changes ultimately affect circulation structure in the whole Arctic (Karcher et al. 2012).

Studies of historical sea-ice records demonstrate that maximum annual sea-ice extent has greater variability close to Atlantic inflows (Shapiro et al. 2003), likely reflecting the variability of the heat and salt fluxes associated with these features. The proximal position of the Barents Sea to the warm Atlantic inflow, which flows onto the shelf through the Barents Sea Opening between Northern Norway and Bjørnøya, means that the Barents Sea hydrography is particularly sensitive to changes in Atlantic inflow, compared with other Arctic shelves.

9. A process of turbulent mixing that may occur across the halocline in the Marginal Ice Zone.

10. The region peripheral to and influenced by the sea-ice edge.

11. An annular mode of sea level pressure anomalies surrounding the Northern polar region.

1.1.4 Atlantification

Bathymetry acts to stabilise the position of hydrographic features in the Barents Sea, such as the Polar Front, on inter-annual time scales. However, anthropogenic changes to the global climate system have doubled the volume of Atlantic-derived waters in the Barents Sea over recent decades – a process known as ‘Atlantification’. This substantial change is forcing the poleward retreat of the Polar Front (Francis and Hunter 2007). The migration of hydrographic fronts is causing rapid changes to the physical properties of the water-column across the Barents Sea. The increased admission of warm Atlantic-derived waters is undermining the persistence of sea-ice. Sea-ice itself controls water-column stratification, by encouraging the formation of fresh Melt Water lenses. A reduction in the persistence of sea-ice has therefore undermined the capacity of the Barents Sea to form stratified water-columns permissive to ice formation in the first place – resulting in a positive reinforcement. Indeed, the Polar Front has recently started delimiting the maximum extent of Barents sea-ice (Barton et al. 2020). In the coming decades, the evolution a positive feedback loop between reduced sea-ice survival, particularly multi-year ice, and haline stratification may undermine sea-ice formation, even in the Northern Barents Sea (Polyakov et al. 2017). Thus, the hydrographic set-up of the Barents Sea is uniquely sensitive to anthropogenic climate change, and faces the possibility of a rapid future reorganisation. Changes in the Barents Sea may herald future shifts that will propagate across the entire Eurasian Basin. The heightened sensitivity of the hydrography of the Barents Sea to climate change therefore makes it a model system to investigate the anticipated effects of future climate change more broadly over the Arctic shelves of Eurasia and perhaps the Arctic system as a whole. The possibility of rapid future change means that research documenting baseline ecosystem properties and incipient change now has special value as a reference point to place future observations in context.

1.2 Phytoplankton of the Barents Sea and the wider ecosystem

The complex hydrography of the Barents Sea establishes physical conditions that permit high levels of primary production; up to $200 \text{ g C m}^{-2} \text{ yr}^{-1}$, contributing between a quarter and half of the total Arctic shelf sea primary production (Carmack et al. 2006; Sakshaug 2004). Putting this value in a pan-Arctic context, it becomes evident that this degree of primary production is comparatively high; exceeded unambiguously only in the Chukchi Sea (Figure 1.4).

Arctic primary production can be placed in a global context, by comparison with estimates derived from global annually-integrated estimates of primary production, derived from ocean-colour (Figure 1.5).

The astute reader will note that there is a marked discrepancy between Arctic primary production values reported by Carmack et al. (2006) and Behrenfeld and Falkowski (1997), and that the Arctic region is generally typified by the absence of ocean-colour observations in Figure 1.5. The exact reasons for the difficulty quantifying marine Arctic primary production from remotely-sensed ocean colour will be discussed in more detail through this thesis. Suffice to say, the persistent presence of sea-ice and cloud cover at high latitudes limit observational capacities of remote-sensors in this region, and that global primary production algorithms routinely contain parameterisations that may not describe Arctic Seas well. For example, the retrieval of $[chl-a]$ from remotely-sensed ocean-colour is often optimised for lower latitudes (O'Reilly et al. 1998). Primary production algorithms, furthermore, often use ancillary data such as SST to inform estimates of phytoplankton physiology (with an Arrhenius dependency on temperature often assumed), which may describe phytoplankton well at the global scale, but which fail to adequately represent the specific physiological adaptations of polar phytoplankton species, that are most productive at the coldest

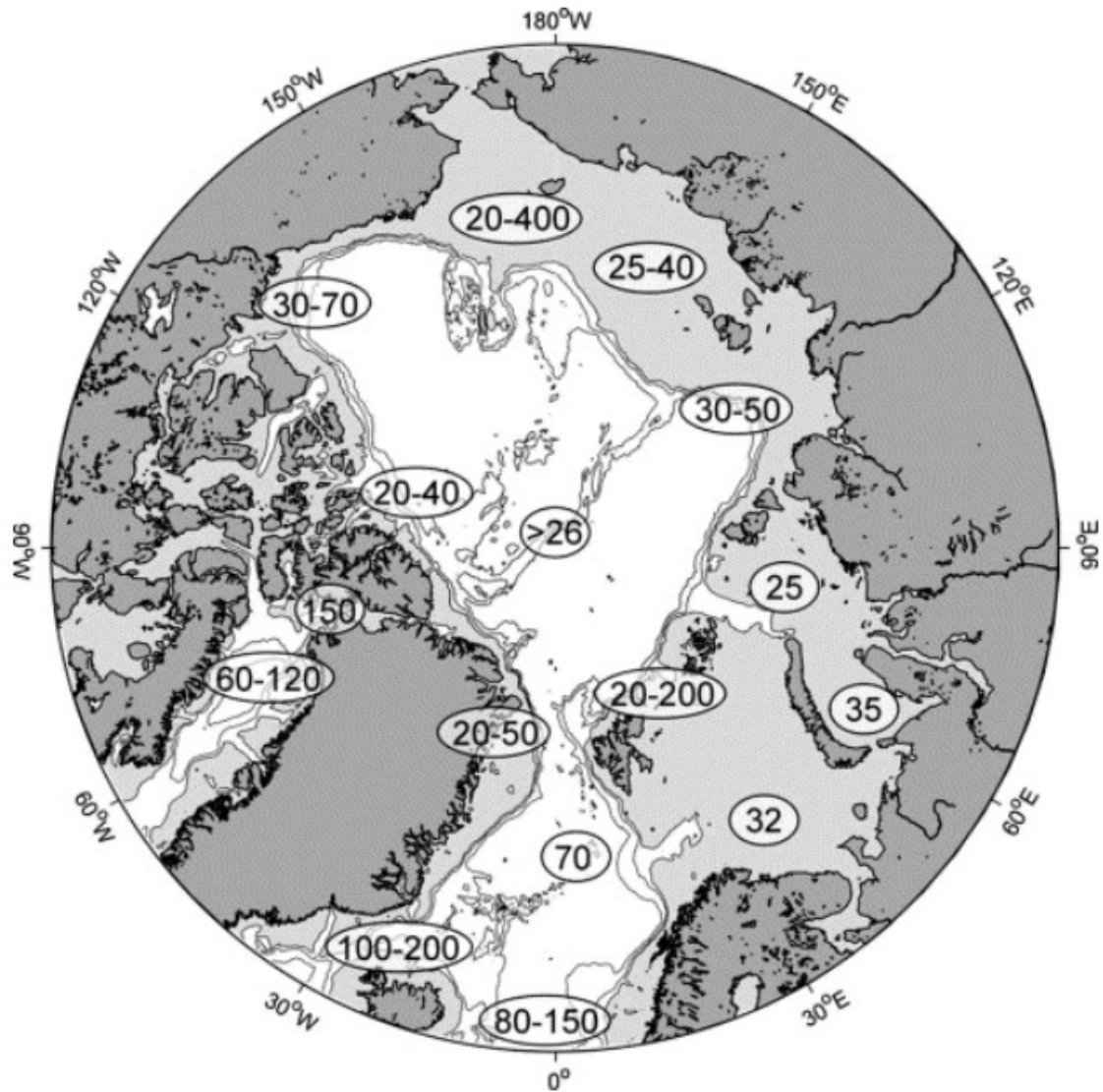


Figure 1.4: Map of representative values of annual primary production ($\text{g C m}^{-2} \text{yr}^{-1}$) from across the Arctic, taken from Carmack et al. 2006. It is evident that the Arctic shelf seas support greater primary production than the central basin, and that inflow shelves in particular are characterised by greater primary production. Reprinted from Progress in Oceanography, 71 (2-4), Carmack et al., "Climate variability and physical forcing of the food webs and the carbon budget on panarctic shelves" 145-181., Copyright (2006), with permission from Elsevier. <https://doi.org/10.1016/j.pocean.2006.10.005>

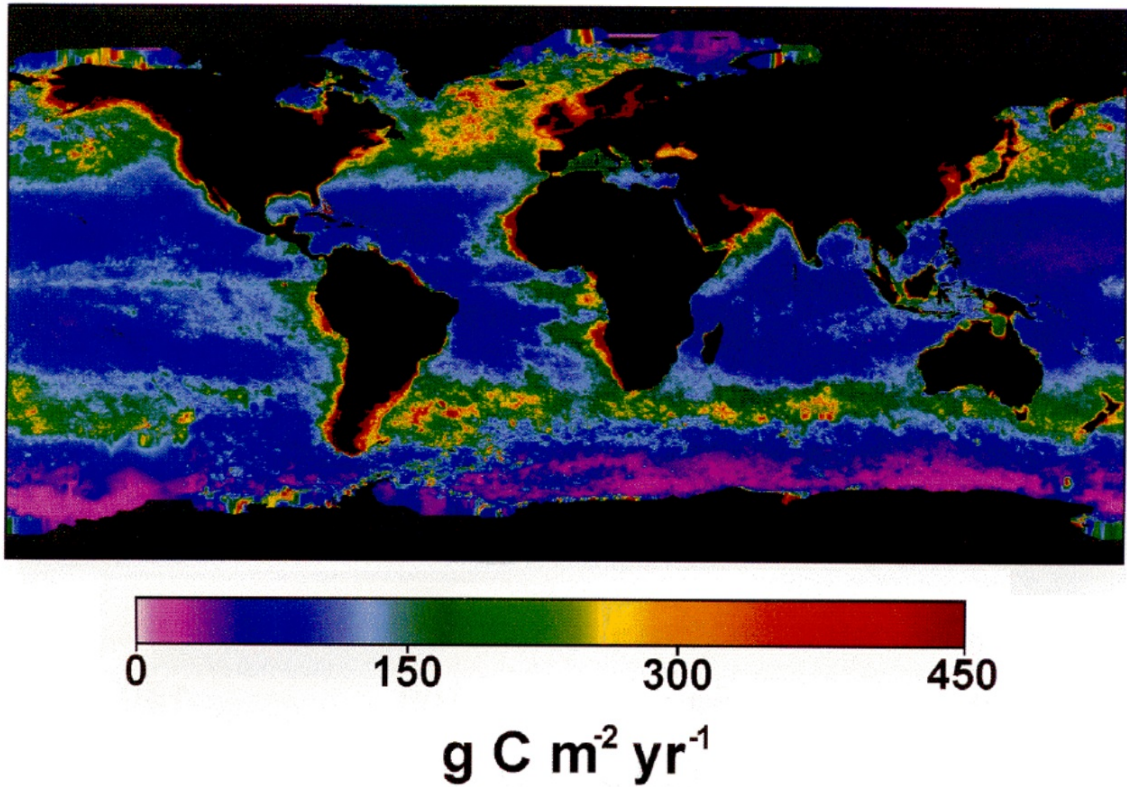


Figure 1.5: Map of estimated annual primary production ($\text{g C m}^{-2} \text{yr}^{-1}$) from across the globe, produced by Behrenfeld and Falkowski (1997), using the Coastal Zone Colour Scanner instrument to collect ocean colour, and estimating primary production through Behrenfeld et al.’s Vertically Generalised Production Model. Reprinted from Limnology and Oceanography, 42 (1), Behrenfeld & Falkowski., “Photosynthetic rates derived from satellite-based chlorophyll concentration” 1-20., Copyright (1997), with permission from John Wiley and Sons. <https://doi.org/10.4319/lo.1997.42.1.0001>

point of the year (Berge et al. 2020). Despite these deficiencies, iterations of the Vertically Generalised Production Model continue to be used with modern remote-sensing platforms (Behrenfeld et al. 2006; Brown et al. 2014) to inform assessments of global marine primary production and phytoplankton dynamics. The unique challenges of using remote-sensors to adequately represent ecosystem dynamics and contributions to biogeochemical cycles will be a recurrent subject of this thesis, but for now, let us return to our discussion of Barents

Sea marine primary production.

Intense phytoplankton blooms typically occur in the Barents Sea in spring, when the alleviation of nutrient limitation, evolving as a consequence of deep winter mixing (Olsen et al. 2002), acts in unison with the stabilisation of the water-column by surface heating and sea-ice melt to create a shallow nutrient-replete surface mixed layer, ideal for photoautotrophy¹² (see Figures 7.7 and 7.8). High primary production, in conjunction with physical processes and biological mechanisms that mediate transport of organic matter beneath the surface mixed layer, results in intense carbon export (Andreassen and Wassmann 1998). Indeed, the Barents Sea exhibits some of the highest biological carbon pump efficiencies¹³ in the Global Ocean (Buesseler et al. 2020), and supports significant economically important fisheries (Drinkwater 2006). Rapid Arctic climate change, and the effect it may have on ecosystem function in the Barents Sea, is therefore of great concern. There is a mandate to study the intersection of physical and ecosystem processes in order to better understand how biogeochemical contributions and ecosystems services of this region are liable to change over coming decades.

1.2.1 Pelagic modes of production

Melting sea-ice has an important role in bloom initiation in the Barents Sea. The release of fresh water, and resulting haline stratification, leads to a rapid stabilisation of the water column compared to the open-ocean. The recession of sea-ice therefore results in ‘ice-edge’ blooms (Wassmann et al. 1999), which are often dominated by opportunistic fast-growing diatom genera, such as *Chaetoceros*, *Thalassiosira* and *Fragilariopsis*, reaching [*chl-a*] in

12. Photosynthesis of biomolecules from inorganic carbon.

13. A measure of the capacity for biological processes to fix and export atmospheric carbon below the surface mixed layer.

excess of 10 mg m^{-3} (Båmstedt et al. 1991). An illustrative figure of abundant Barents plankton species is provided in Figure 1.6.

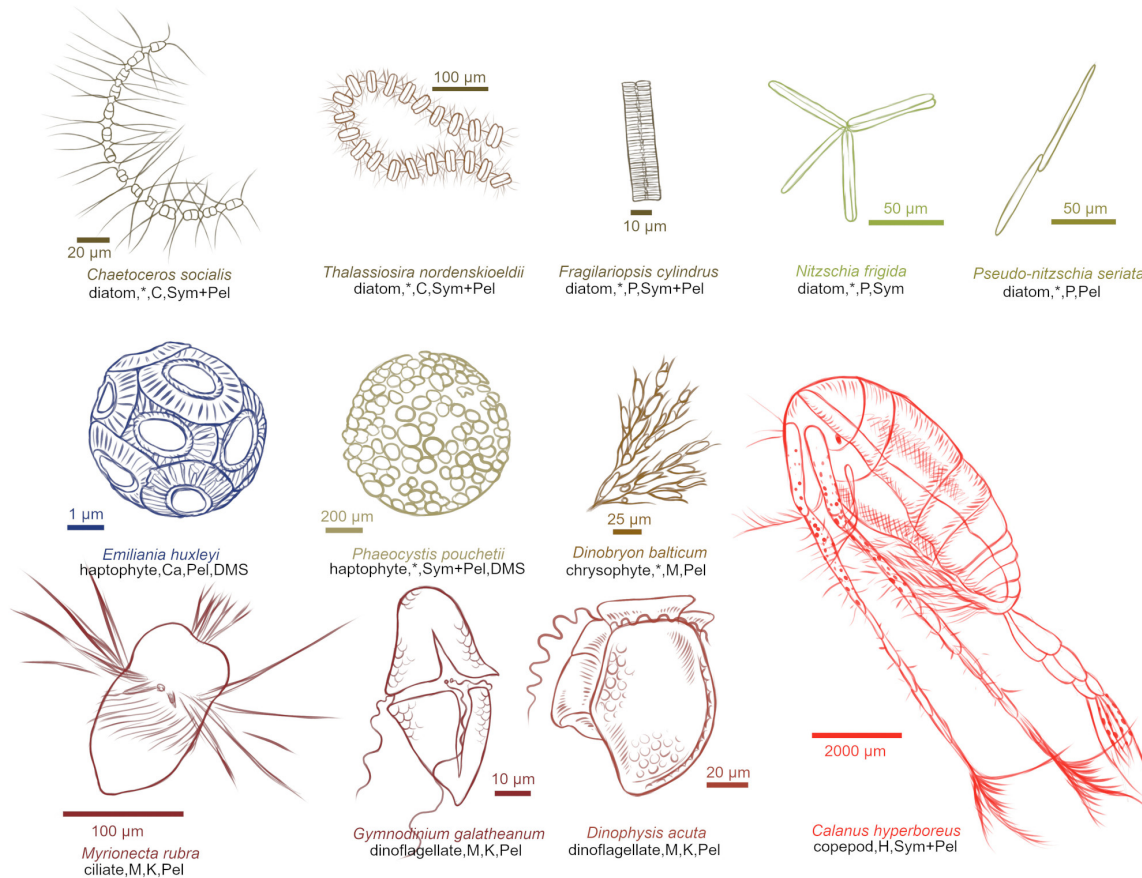


Figure 1.6: Illustrations of abundant plankton species in the Barents Sea and indications of functional associations. Colonial (*), calcifying (Ca), mixotrophic (M), kleptoplastid (K), heterotrophic (H), centric diatom (c), pennate diatom (p), sympagic (Sym), pelagic (Pel), DMS-producer (DMS). Colours indicate trophic status; green and blue (photoautotrophy), purple (mixotrophy), red (heterotrophy). Illustrations produced by Andrew Orkney, based on optical microscopy training with Elaine Mitchell at the Scottish Association for Marine Science. The abundance of species such as *Emiliana huxleyi*, *Phaeocystis pouchetii*, *Dinobryon balticum* and species of *Chaetoceros*, *Thalassiosira* and *Fragilariopsis* is supported by Sakshaug et al. (2009), pages 167-208. The specific dinoflagellate species illustrated here were conspicuous in the optical microscopy, but *Dinophysis acuta* is attested as an abundant species in Rat'kova and Wassmann (2002).

Diatom-dominated ice-edge blooms are, however, not the only possible scenario; intense blooms of *Phaeocystis pouchetii* have been reported developing under the sea-ice North of Svalbard (Assmy et al. 2017; Pavlov et al. 2017; Kauko et al. 2019). Additionally, blooms of pelagic diatoms, originating in the open ocean, may be advected underneath the sea-ice (Johnsen et al. 2018).

Half of the annual primary production in the Barents Sea can occur in these short-lived pulses of production that may only persist less than a month before nutrient stocks in the surface mixed layer are depleted. Despite their short duration, the intense biological activity they host places the Barents Sea among the most productive seas on Earth (Wassmann et al. 1999; Carmack and Wassmann 2006). Standing stocks of phytoplankton on the Arctic shelves reach such concentrations that they colour the ocean surface – and centennial scale quantitative records of ocean-colour (Wernand and Woerd 2010) show that the Barents Sea has historically been the greenest open sea of all (Wernand et al. 2013).

By contrast, the warmer Atlantic-influenced waters South of the Seasonal Ice Zone¹⁴, where sea-ice melt cannot induce rapid stratification, begin to bloom later. This is because sufficient time must pass for a thermal gradient in the water-column to become established and the Mixed Layer Depth to shoal above the euphotic depth¹⁵ (see Figure 1.7 assemblage 4).

14. The region which may experience seasonal sea-ice cover.

15. The depth at which Photosynthetically Active Radiation declines to 1% of the surface value.

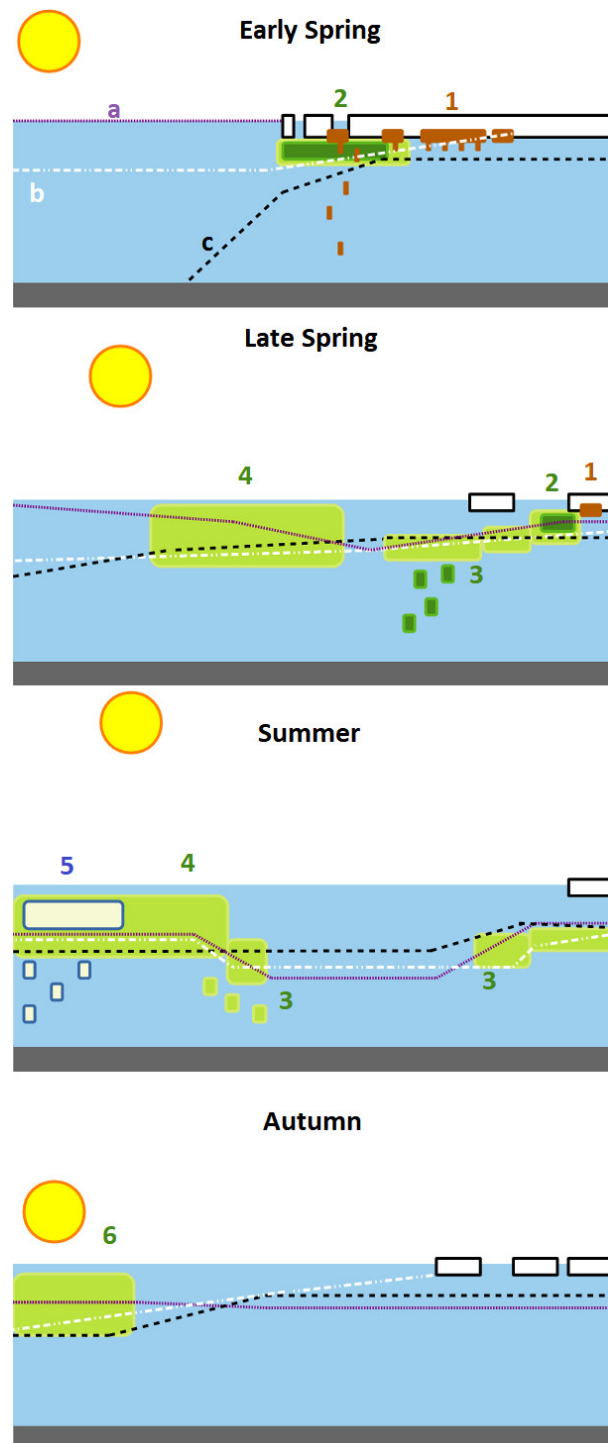


Figure 1.7: Seasonal progression of phytoplankton blooms in the Barents Sea. **a:** nutricline, **b:** euphotic depth, **c:** Mixed Layer Depth, **1:** ice-algae, **2:** ice-edge bloom, **3:** sub-surface bloom, **4:** pelagic bloom, **5:** coccolithophore bloom, **6:** secondary autumnal bloom. This generalised progression is not comprehensive. Specific bloom dynamics may vary between years, for example ice-edge and pelagic blooms may occur discretely in warm years, but be super-imposed in cold years (Oziel et al. 2017).

These waters tend to be dominated by smaller prymnesiophyte species such as *Emiliana huxleyi* and *Phaeocystis pouchetii* (Figure 1.6). While blooms South of the ice-edge do not tend to reach such high biomass concentrations, Atlantic-influenced phytoplankton communities are widely argued in the literature to attain higher cumulative levels of net¹⁶ primary production and new production¹⁷ over an annual time scale. This is believed to result from their more protracted occurrence and the deeper vertical mixing of the sea in these more Southern locations, which makes accessible a greater portion of the deep nutrient inventory (Drinkwater 2006; Wassmann et al. 2010; Degen et al. 2016).

Atlantic-influenced waters of the Barents Sea may even exhibit a smaller secondary autumnal bloom in September as thermal stratification is relaxed and a deepening mixed layer provides a second charge of nutrients to surface waters – a mechanism of bloom formation more associated with temperate waters (Ardyna et al. 2014; Dong et al. 2020). A cartoon of the general progression of phytoplankton blooms over the productive season in the Barents Sea is illustrated in Figure 1.7.

As blooms exhaust available nutrients, mixotrophic¹⁸ or heterotrophic¹⁹ taxa such as ciliates and dinoflagellates, or specialist diatom species such as *Pseudo-nitzschia*, become more dominant (Wassmann et al. 2006a) (see Figure 1.6 and Figure 7.8). While meridional differences in phytoplankton community composition are to some extent determined by distinct ecological preferences among some highly abundant phytoplankton species (Wassmann et al. 2006a; Rat’kova and Wassmann 2002), the advection of source populations also plays a critical role in determining the community composition of ensuing blooms (Syvertsen 1991; Boetius et al. 2013; Johnsen et al. 2018; Oziel et al. 2020).

16. The difference between energy fixed by primary production and that used by respiration.

17. The fraction of production supported by newly supplied nutrients, as opposed to recycled forms of nutrients.

18. Organisms that are capable of both auto- and heterotrophy.

19. Organisms that rely on organic carbon for sustenance.

High levels of inter-annual variability in climatology and the delivery of different water-masses and nutrient stocks onto the shelf (Rey 2012; Hátún et al. 2017) are associated with inter-annual variability in primary production in the Barents Sea – on the order of 20% (Carmack and Wassmann 2006). Concomitant with this variability in ocean productivity, the relative dominance of diatom and prymnesiophyte genera are also subject to a high degree of spatio-temporal variation (Reigstad et al. 2002). This is a consequence of both changing environmental conditions favouring the growth of different species – and fluctuations in the contributions of different inflowing water-masses, which deliver different phytoplankton stocks. Typically, the composition of phytoplankton blooms in the spring tends to be more homogeneous across the shelf (Degerlund and Eilertsen 2010), and spatial gradients in composition become more pronounced as blooms mature into summer²⁰.

As the season progresses to high and late summer, blooms of the calcite-producing coccolithophore *Emiliana huxleyi* can emerge in the Atlantic-derived water-masses of the Southern Barents Sea (Smyth et al. 2004; Hovland et al. 2013), exploiting warm ($>6^{\circ}\text{C}$), well-stratified and oligotrophic²¹ waters (Downes et al. 2020). Such coccolithophore ‘whitings’ have historically occurred well South of the Seasonal Ice Zone, but have begun to disperse further onto the Eurasian shelves in strengthening Atlantic inflow currents (Oziel et al. 2016; Oziel et al. 2017; Neukermans et al. 2018; Oziel et al. 2020; Orkney et al. 2020). The presence of *Emiliana huxleyi* has been indicated in August as far north as Atlantic-derived waters around the Svalbard Archipelago by the taxon-indicative pigment gyroxanthin-diester (Pettersen et al. 2011). In addition, Pettersen et al. (2011) routinely detected the prasinophyte-indicative pigments prasinoxanthin and chlorophyll-b (Pettersen et al. 2011), supporting earlier reports of abundant *Micromonas pusilla* picophytoplankton in the Nordic Seas (Lovejoy et al. 2007). This clearly indicates that the phytoplankton diversity, and potentially functional variety and significance, of the Barents Sea extends

20. See meridional gradient in community composition evident in July 2018; (Downes et al. 2020).

21. Conditions of nutrient scarcity.

beyond nano- and microplankton. It must be expressed, however, that forms such as *Emiliana huxleyi* and *Micromonas pusilla* are united with diatoms and *Phaeocystis* in all being eukaryotes.

Blooms of *Emiliana huxleyi* rarely occur North of the Polar Front, even when conditions there are suitable, because hydrographic fronts impose barriers to their distribution to these areas (Oziel et al. 2020). Coccolithophore fossils, preserved in Pleistocene sediments of the central Arctic Ocean (Backman et al. 2009) do reveal that coccolithophores have expanded into ice-influenced environments in previous interglacials. The periodic occurrence of coccolithophore whittings is significant because of their contributions to the benthic flux of calcite and production of the climatologically important gas species dimethyl-sulphide (DMS) (Galí et al. 2019). Coccolithophore production of calcite also liberates carbon dioxide to the water-column and hence potentially modulates the dissolution of atmospheric carbon into the surface ocean (Shutler et al. 2013).

1.2.2 Sympagic modes of production and ‘hidden’ blooms

In addition to pelagic modes of production, there are ‘hidden’ forms of production in the Barents Sea that are not routinely quantified because they are not observable at the ocean surface or occur in inaccessible environments. Such clandestine forms of production include phytoplankton blooms occurring in or underneath sea-ice, or in deep ‘Sub-surface Chlorophyll Maxima’ (SCMs) that become especially pronounced in the summer season (see Figure 7.8). Downes et al. (2020) noted the widespread occurrence of productive SCMs, in the Arctic-derived water-masses of the Barents Sea, in July 2018. Remote-sensing approaches, such as the observation of ocean-colour, can only observe water-column properties within 1 optical depth²² of the ocean surface. The thickness of the optical depth can decline below 1-2 m in productive settings (Johnsen et al. 2018), meaning that remote-sensing approaches

22. A measure proxying the depth interval over which incident light is attenuated, conventionally defined at the depth from which >90% of the upwelling surface radiance originates (Gordon and McCluney 1975)

can only be reliably applied in ice-free contexts, where phytoplankton biomass is concentrated near the ocean surface. There is hence a hidden world of phytoplankton activity, beyond observation by remote-sensors.

Ice-algal production is largely achieved by specialist shade-adapted colonial diatoms living directly on the ice subsurface or forming languid spaghetti-like colonies that dangle into the water-column beneath. These include the pennate *Nitzschia frigida*, which tends to form colonies of millimetre thickness under first-year ice (see Figure 1.8), and the centric diatom *Melosira arctica* that can form metre-long filaments – typically under multi-year ice (Syvertsen 1991; Boetius et al. 2013).



Figure 1.8: **Left:** Photograph of collapsed sympagic algae filaments on over-turned ice, RRS James Clark Ross, summer 2017, JR16006. **Right:** Photograph of sympagic algae colonies revealed as the wake of the RRS James Clark Ross disrupts the sea-ice field in summer 2018, cruise JR17006. The dark colouration likely reflects densely packaged concentrations of pigment, including diatom-indicative pigments such as fucoxanthin, which confer an orange appearance.

While we might imagine that shaded conditions should depress photoautotrophy, fixture to the ice subsurface ensures sympagic²³ communities are not mixed beneath the euphotic depth. In addition, a pan-Arctic conveyor belt that imports ice from the Siberian Shelf

23. Ice-associated.

– where it is enriched by terrigenous input (Krumpen et al. 2019) can mean that sea-ice contains its own pool of nutrients. These nutrients are gradually released as the ice melts, which makes sea-ice a lucrative environmental niche.

As a consequence of the unique environmental conditions associated with ice, sympagic algae can begin growing as early as February, before the onset of physical stratification in the water-column (Syvertsen 1991). Such ice-algae assemblages possess distinct $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ signatures and unique fatty acid species. The chemical signatures of benthic²⁴ sediments or animal tissues can therefore be analysed to estimate contributions from sympagic algae. The stable isotopic compositions of key copepod (*Calanus glacialis* and *C. hyperboreus* – Figure 1.6) and amphipod species, that play important roles transferring production to higher trophic levels (Vernet et al. 2017), indicate substantial sustenance is derived from ice-algae (Kohlbach et al. 2016). Standard $\delta^{14}\text{C}$ uptake incubations in sympagic and pelagic settings in the Northern Barents Sea indicate 16-22% of total primary production is ice-associated (Hegseth 1998). Kauko et al. (2018) showed that there is a succession, independent of environmental forcing, from young ice-algal communities dominated by ciliates and dinoflagellates, towards mature communities dominated by pennate diatoms. Importantly, the establishment of mature communities was fostered by adjacent areas of older sea-ice seeding new territory in first-year ice (Kauko et al. 2018), potentially because multi-year ice acts as a repository for seeding ice-algal populations (Olsen et al. 2017). The capacity for mature sympagic communities to develop may therefore be fundamentally altered as older sea-ice becomes less prevalent (Kwok 2018) and conveyor-like circulations that re-distribute such ice less efficacious (Krumpen et al. 2019).

As the sea-ice melts, ice-algal colonies may be liberated into the water-column and act either as seeding populations for pelagic blooms, or sink directly to the benthos (see Figure 1.7 early spring assemblage 1) (Syvertsen 1991; Boetius et al. 2013). Ice-algal assemblages

24. Ecosystems at the sea bottom.

hence play a role ‘inoculating’ ice-edge blooms and therefore, while they are rarely observed without ice-breaking vessels, ice-algal blooms are inescapably convolved with pelagic modes of production.

It is evident from the discussion thus far that some sympagic diatom species live double lives, practising both sessile and planktic ecologies. Sessile modes are reminiscent of their heterokont sisters, the brown-algae, some of which develop differentiated multicellular tissues. Pluricellularity is widespread among eukaryotic phytoplankton of the Arctic shelf seas, occurring not only in diatoms but also the heterokont golden alga *Dinobryon* (Hansen et al. 1990) and even in haptophytes such as *Phaeocystis* species. Indeed, there may be a differentiation between vegetative and sexual tissues in *Phaeocystis* (Peperzak et al. 2000a). Both of these phytoplankton genera occur routinely in fresher, ice-associated settings (Degerlund and Eilertsen 2010). Colony formation in *Phaeocystis pouchetii*, mediated by the secretion of gelatinous extracellular polymers, can result in the flocculation of cells into rapidly-sinking aggregates that result in mass-sedimentation and export of organic matter beneath the mixed layer (Wassmann et al. 1990).

An additional layer of complexity is super-imposed by interspecific interactions, such as the nucleation of *Phaeocystis* colonies around senescent diatom frustules (Weisse et al. 1986), or the colonisation of existing *Phaeocystis* colonies by specialist pennate diatoms (Sazhin et al. 2007). These processes, and the secretion of toxic compounds – such as the climatologically relevant aerosol species DMS – may act to modulate the efficacy with which *Phaeocystis* blooms are metabolised by heterotrophs or exported beneath the mixed layer.

As nutrients are depleted in the well-stratified surface waters of the Marginal Ice Zone, phytoplankton blooms are unable to continue producing new biomass and are forced to descend towards the nutricline²⁵ (Kristiansen and Farbrot 1991; Cullen 2015), forming SCMs. The significance of production in the SCM as a source of sustenance for the wider ecosys-

25. A rapid increase in nutrient stocks with depth.

tem and as a means to sequester atmospheric carbon has proven contentious. Some authors have argued that SCM phytoplankton communities are routinely moribund, representing phytoplankton ‘graveyards’ responsible for only a fraction ($< 1\%$) of total net primary production (Arrigo et al. 2011), while others have suggested that their contributions may be significant (Hill et al. 2013). A recent study indicated that when SCMs occur as narrow horizons within the euphotic zone that they may receive sufficient illumination to support equal production to surface blooms (Bouman et al. 2020).

1.2.3 Trophic cascades: energy transfer through the Barents Sea ecosystem

80% of harvestable production in the Barents Sea is transferred through secondary production in bathypelagic and benthic communities (Wassmann et al. 2006a), with the remainder largely grazed by copepod species. Processes that transfer pelagic and sympagic production to the benthos (pelagic-benthic coupling) are therefore of paramount importance to the efficient function of the Barents Sea ecosystem. They determine the transfer of energy to higher trophic levels and sequestration of carbon in shelf sediments.

On the shelf a large proportion of benthic respiration, as indicated by oxygen consumption, is mediated by epifaunal²⁶ megabenthic animals (> 1 cm length) such as abundant brittle stars – particularly in rocky areas, while trough or shelf-break environment respiration is dominated by smaller infaunal²⁷ animals such as polychaete worms. Sedimentary respiration is mostly achieved by microbial communities (Piepenburg et al. 1995). Up to 75% of oxygen consumption in the benthic community can be attributed to epifauna in shallow settings (Renaud et al. 2008), and a majority of sinking carbon and pigments fuelling secondary production are inferred to be diatoms and copepod faecal pellets.

Paradoxically, spatial gradients in pelagic primary production in the Barents Sea are

26. To live above the sediment-water interface.

27. To live below the sediment-water interface.

anti-correlated with gradients in benthic secondary production (Jørgensen et al. 2014; Degen et al. 2016). Enhanced trawling pressure in the South-west Barents Sea might be implicated, because it removes benthic biomass. However, estimated net primary production in the Northern Barents Sea is routinely lower than the measured benthic production and therefore trawling alone is not a full explanation (Degen et al. 2016). Degen et al. (2016) also reported a positive relationship between trawling and benthic production, possibly caused by the removal of larger, older and less productive benthic organisms in trawls, dismissing trawling as an explanation for antithetical gradients in surface and benthic production.

Spatial variation of the f -ratio²⁸, and efficacy of pelagic-benthic coupling, offer potential resolutions; a smaller amount of net primary production might support greater benthic production if it is transferred very efficiently to the benthos, while a region with high net primary production may contribute relatively little sustenance to higher trophic levels if a large fraction of net primary production is supported by regenerated nutrients being recycled in the microbial loop. Clandestine forms of production, such as under-ice or SCM blooms, may need to be invoked to fully explain scenarios in which benthic production exceeds estimated surface primary production. If these hidden forms of production occur preferentially in Northern waters regarded as less productive, they offer a simple explanation for increasing secondary benthic production with latitude. Alternatively, the mass-advection of pelagic organisms (Eriksen et al. 2017) might redistribute food to less productive regions, meaning that local pelagic and benthic forms of production may not be in direct equilibrium. The possible explanations for reversed gradients in detectable pelagic and benthic production may therefore include a complex network of interacting physical and biological processes – which confounds the task of inferring their potential response to anthropogenic climate change. I hence consider each category here in turn and contrive a scheme of their

28. A measure of the utilisation of inorganic compared to total nitrogen sources – representing the ratio of new to total (new and regenerated) production. Conventionally, nitrate is taken to proxy inorganic sources and ammonium to proxy organic sources.

interactions.

Physical causes for enhanced Northern export

Degen et al. 2016 suggest that ice-edge upwelling could justify higher new production in the Marginal Ice Zone and that other ice-edge physical processes might be instrumental in the vertical export of carbon to depth. Indeed, transient occurrence of salt-fingering in this environment may penetrate beneath a relatively shallow mixed layer caused by haline stratification (Porter et al, in prep), facilitating enhanced export. By contrast, the mixed layer takes longer to shoal in the thermally-stratified Atlantic Waters of the Southern Barents Sea. Phytoplankton biomass evolving in these regions may therefore be less liable to sudden export.

Esoteric physical processes, such as ballasting of phytoplankton blooms by gypsum crystals shed from sea-ice, may also in some circumstances contribute to efficient export production (Wollenburg et al. 2018). Some of these physical processes, such as the ballasting of colonial phytoplankton, are implicitly convolved with the quality and quantity of primary production (see 1.2.3). For example, the suggestion that enhanced wind stress at the ice-edge enhances export (Degen et al. 2016) is related to the quantity of primary production, because upwelled nutrients may fuel additional primary production.

Biological quality of production

As the Barents Sea transitions into summer, communities of nanophytoplankton²⁹ tend to predominate in Atlantic Waters. The larger microphytoplankton³⁰ in the Atlantic water-masses tend to be mixotrophic dinoflagellates (Downes et al. 2020), rather than obligate autotrophs. Smaller autotrophic phytoplankton cells do not sink as rapidly as large frustule-

29. Phytoplankton that are between 2 and 20 μm diameter.

30. Phytoplankton that are over 20 μm diameter.

clad diatom colonies (Falkowski et al. 1998; Jiang et al. 2005) and their less refractory coatings may render them less protected from consumption by heterotrophs, possibly eroding the downward carbon flux. Indeed, studies of *Phaeocystis* occurrence in the Southern Barents Sea imply that their biomass is more rapidly remineralised as they sink, in contrast to diatoms (Reigstad and Wassmann 2007).

As the Barents Sea’s production evolves from spring through summer it is increasingly supported by regenerated nutrients. The delayed evolution of stratification in the Southern Barents Sea may therefore mean f -ratios are perennially lower and that higher net primary production in the surface ocean may result in only modest contributions to export (Erga et al. 2014). Such an interpretation is compatible with the evolution of high ammonium concentrations in the Southern Barents Sea during summer (see Figure 7.10).

Thus, the biological quality of production, and the role it plays in moderating export, is not only structured by bottom-up controls such as the physical environment (see 1.2.3), but also top-down controls from higher trophic levels (see 1.2.3).

As variation in the biological quality of production is a pre-eminent means by which pelagic-benthic coupling is mediated, its study is a central subject of this thesis. The biological quality of production is directly amenable to study under work package 3 of Arctic **PRIZE**, as a consequence of the detailed collection of taxon-indicative pigment ratios and seawater samples stained with Lugol’s solution for the enumeration of microphytoplankton cells under optical microscope. There are currently sustained efforts in the literature to recognise and monitor phytoplankton functional assemblages in remotely-sensed ocean-colour (Quere et al. 2005; Sathyendranath et al. 2004; Jackson et al. 2010; Sathyendranath et al. 2014; Liu et al. 2018), but studies specific to the Barents Sea have, thus far, have disproportionately focussed on discrimination of coccolithophores (Smyth et al. 2004; Hovland et al. 2013; Oziel et al. 2016; Oziel et al. 2017; Neukermans et al. 2018; Oziel et al. 2020), owing to the strong influence that scattering by their biomineral liths has on

upwelling radiance (Hovland et al. 2014). However, there have been previous efforts to distinguish diatoms from mixed phytoplankton populations in the Labrador Sea (e.g. Stuart et al. (2000)). However, such algorithms have only a regional scope of application, owing to region-specific patterns in phytoplankton hyperspectral absorption (Jackson et al. 2010). As hundreds of measurements of *in-vivo* phytoplankton hyperspectral absorption (hereafter simply referred to as bio-optical properties) were collected under Arctic **PRIZE**, there is a potential for similar efforts to be extended to the Barents Sea. This offers the prospect for Functional assemblages – which might be deduced from relationships between community compositional and environmental datasets – to be recognised and related to bio-optical properties of the ocean, such as hyperspectral absorption ($a_{ph}(\lambda)$), and thereby remotely-sensed ocean-colour. As different community compositions of phytoplankton may have significantly different functional significances, such as their contributions to export production, this could provide additional insights into ecosystem state in the Barents Sea. Although alacrity to detect functional assemblages and infer biogeochemical dynamics must be tempered with the knowledge that biogeochemical relationships can be highly variable and influenced by environment-specific circumstances (Le Moigne et al. 2015; Wollenburg et al. 2018). Keeping these nuances in mind, a key ambition of this thesis is to explore the extent to which optically distinct phytoplankton assemblages might be recognised in the Barents Sea from space, in aid of improving insights into the state of the ecosystem and probing its response to anthropogenic perturbation. We must temper our enthusiasm, however, with the knowledge that bio-optical properties of distinct phytoplankton assemblages can be dynamic. Phytoplankton may modulate their production of different intracellular pigments, or change the way in which pigment-protein complexes are bonded (Johnsen and Sakshaug 2007), which may significantly alter the wavelength and amplitude of pigment absorption spectra. Evidence of photoacclimation can also be found in the ratio of key protective or photosynthetic carotenoid pigments (e.g. Johnsen et al. (2018) and Kauko

et al. (2019). In combination with self-shading (e.g. Barros et al. (2003)) pigment-specific absorption ($a_{ph}^*(\lambda)$) may be altered by a factor of 2-4 by photoacclimation, which could seriously disrupt researcher efforts to infer pigment concentration or spectral signatures of distinct phytoplankton assemblages from ocean colour.

Heterotrophic controls on export production

Pelagic heterotrophs, such as grazing copepod species, are usually credited with enhancing export production through much of the Global Ocean. This results from Diel Vertical Migration, which allows them to excrete fast-sinking faecal pellets beneath the surface mixed layer. However, copepod grazers in the Southern Barents Sea can reduce export because of coprophagous behaviour in response to food scarcity (e.g. Riser et al. 2002, Riser et al. 2007). The specific life-histories of grazers in the Barents Sea, caused by variation in the phenology³¹ of their autotrophic food-supply, therefore has great influence on their moderations of export production.

In addition to indigenous heterotrophs, supported by autochthonous production, the widely variable flux of pelagic heterotrophs onto the shelf that is mediated by ocean currents must be considered. Atlantic currents may introduce juvenile fish and krill into the South-west Barents Sea. Massive increases in the abundance of these organisms, likely driven by increasing Atlantic inflows, drove a doubling in the Barents Sea pelagic compartment biomass after 2003 (Eriksen et al. 2017). This means that there is a significant degree of allochthonous influence in the Barents System on top-down controls of ecosystem function.

Export production is therefore the result of the superimposition of multiple fields; a field of the initial new production and phytoplankton community composition, physical proclivity for export and the action of heterotrophs moderating the survival of sinking particles.

31. Cyclic, seasonal biological phenomena.

It is evident, from this discussion, that there is no consistent globally-applicable relationship between remotely-sensed [*chl-a*] and net primary production, and between net primary production and resultant export production. The efficiency of export in the global ocean varies widely (Buesseler et al. 2020) and therefore, converting net primary production into export estimates requires knowledge of ancillary proxies of this variability. SST is routinely invoked to this end (Laws et al. 2000), as it covaries with a host of biological and physical processes that sum to influence export. However, we should remind ourselves that even a single species, such as a calanoid copepod, can significantly alter export efficiency (e.g. Riser et al. (2002) and Riser et al. (2007)), and that we must therefore accept that there exists some degree of uncertainty that cannot be easily represented in algorithmic representations of ocean processes, but which must be quantified empirically and inferred intuitively by the skilled biologist.

It might be expected that a change in the overall productivity should have straightforward implications for export. However, it must be considered that there may be accompanying changes in phytoplankton community composition and the survival, composition and phenology of heterotrophs. It must also be acknowledged that benthic heterotrophs, that rely on a vernal pulse of productivity to sustain them, are constrained to be in equilibrium with the autochthonous production. By comparison, pelagic heterotrophs are often highly migratory or liable to rapid dispersals as a result of mechanical mixing.

The network of interactions between environmental conditions, autotrophic primary production, heterotrophs and resulting export production, is illustrated in Figure 1.9.

Environmental conditions are instrumental in facilitating autotrophy, for example by creating conditions permissive to growth or by advecting phytoplankton stocks onto the shelf. Resulting production might be directly exported to the benthos (1), or form a flux to the benthos that is moderated by heterotrophy (2). The composition of the autotrophic community might influence the structure of the heterotrophic community, resulting in a

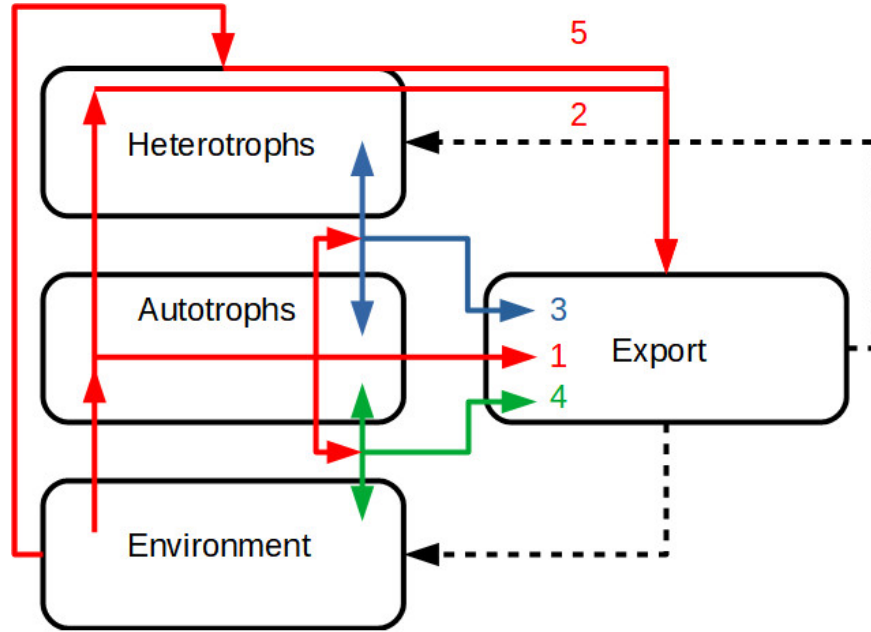


Figure 1.9: Diagram of relationships between bottom-up and top-down controls on primary production and their relationships to export. Red: pathways of interaction caused by environmental change. Green: pathways associated with change in the autotroph community’s interactions with the environment. Blue: pathways associated with change in the heterotroph community. Dashed: feedbacks associated with export production.

feedback between autotrophy and heterotrophic moderation of export (3). Changes in the physical environment may also moderate the flux of exported production (4). There is some capacity for autotrophs to themselves influence the evolution of the physical environment, for example by engendering surface heating of the water-column by altering the albedo of sea-ice or the absorption of light in pelagic settings (Park et al. 2015), or by releasing climatologically relevant gasses (Charlson et al. 1987). The physical environment may also directly determine the survival of heterotrophs, or their dispersal onto the shelf, modifying the capacity of the heterotrophic community to moderate export production (5).

Ultimately, export production has a role in determining the nourishment of bathy-

pelagic and benthic heterotrophs. Export therefore represents a feedback to the pelagic heterotrophic community because some pelagic heterotrophs are meroplankton³² with benthic or bathypelagic adult stages. This means their long term reproductive success may be influenced by the quantity, quality and style of export production. Some heterotrophs may themselves be consumed by other heterotrophs that have benthic stages; Arctic cod (*Boreogadus saida*) and capelin (*Mallotus villosus*) both practice demersal³³ ecologies and are prolific consumers of calanoid copepods (Hop and Gjøsaeter 2013). Export production also influences the environment and its capacity to support autotrophy, because it results in nutrient draw-down.

This discussion of species-specific interactions in the heterotroph community is a reminder that each compartment modifying export production possesses complex internal structure – and that anthropogenic influence on ecosystem function is not merely limited to perturbation of long-term environmental conditions, but also by human demands for ecosystem services such as commercial fishing. It should be considered, for instance, that the dominant role of epifaunal megabenthic animals in transferring production to higher trophic levels is fundamentally altered by bottom trawling. Bottom trawling is one of the major factors controlling the community composition of benthic ecosystems in the Barents Sea (Denisenko 2001). The sensitivity of the Barents Sea to climate change is therefore shaped by human activities. For instance, observing that bottom-trawling significantly decreases the benthic faunal biomass (Jørgensen et al. 2016), there is good reason to believe that their capacity to fulfil the functional roles they play in the ecosystem will be undermined. For example, if animals that recycle benthic deposits are extirpated by trawling, then this resource of food is made less accessible to demersal-foraging fish species. For example, we can envisage a scenario in which greater Atlantic influence displaces the keystone species polar cod (*Boreogadus saida*) towards their North-east refuge (Hop and Gjøsaeter

32. Organisms that are only planktic for part of their life-cycle.

33. Living close to the benthos.

2013) and reduced sea-ice extents permit greater trawling within this refuge – potentially resulting in a combination of additional stresses on this species.

While there is tentative evidence that it may be possible to detect zooplankton swarms in remotely-sensed ocean-colour (see Basedow et al. (2019)) the development of such algorithms is currently in its infancy. This means that the documentation of variation in zooplankton populations is reliant on *in-situ* sampling (e.g. Eriksen et al. (2017)). The role grazers play regulating the occurrence or durations of phytoplankton blooms, and pelagic-benthic coupling, is therefore beyond the thematic scope of this thesis – which is focussed pre-eminently on the monitoring of phytoplankton stocks from space. Upon reviewing Figure 1.9, it can be seen that processes (2), (3) and (5) will remain the subject of future research, unless they are so tightly coupled to phytoplankton functional assemblages that they can be considered deterministic extensions of the biological quality of primary production. However, rigorous study of the pelagic phytoplankton community, relation of environmental associations to distinct functional assemblages, and the recognition of such functional types from remotely-sensed ocean-colour, will offer significant capacity to improve understanding of baseline ecosystem function and its response to environmental forcing. Furthermore, improvements in understanding and ecosystem monitoring, that will result from pursuing these questions, can build a foundation essential to future holistic studies that include higher trophic levels.

1.3 Ecosystem change in the Barents Sea: unresolved questions

While understanding of baseline ecosystem function in the Barents Sea is clearly incomplete, scientists are faced with the additional challenge of accommodating rapid ecosystem change

into their conceptualisation of Barents Sea ecosystem processes. These rapid changes are attested by a range of sampling campaigns, retrieving *in-situ* measurements that proxy ecosystem state, and remote-sensing surveys that facilitate synoptic-scale inferences about ecosystem function. Ecosystem modelling may also be used to further understanding of ecosystem variability and anticipated modes of change. In recent years, accelerating progress in the miniaturisation of electronics has facilitated the development and deployment of a wide range of sensor platforms to Arctic environments (Berge et al. 2020) (chapter 9). *In-situ* sampling by ships, records by autonomous moorings, and remotely-sensed ocean-colour from nationally-funded space-borne platforms, are therefore now accompanied by observations from remotely-operated, semi-autonomous or fully automated robotic submersibles (e.g. (Porter et al. 2020)), drones and miniature satellites. This heterogeneous sum of data-streams, including measurements as diverse as photography by subaqueous robotic landers, hydrography, fluorometric quantification of $[chl-a]$, backscattering by particulate matter and nutrient assays, heralds the dawn of a new data-rich era of Arctic science. The integration of these diverse datasets – representing spatial and temporal scales that may vary between individual seconds at fixed locations (e.g. landers) or multi-day integrations over thousands of square kilometres (remote sensing), presents a unique challenge for a new generation of researchers (see e.g. Kostakis et al. (2020) for the integration of *in-situ* absorption and backscattering measurements acquired by robotic platforms with remotely-sensed ocean-colour, and Degen et al. (2016) for the use of derived biological variables, inferred by a neural network). If they can be effectively integrated, these additional data-streams provide researchers with the opportunity to gain an unparalleled holistic insight into Arctic ecosystem function. In particular, the long-term deployment of autonomous robotics, is allowing researchers to obviate the strong bias towards transient vernal and summer sampling, and permitting greater access to historically under-sampled environments, such as underneath the sea ice (e.g. Lee et al. (2017)) .

For now, however, I concentrate on discussing the legacy of traditional *in-situ* sampling, remote-sensing methods and modelling approaches, that have come to shape the established schools of thought concerning the biology of the Barents Sea, identify their deficiencies and where those data collected under Arctic **PRIZE** may prove helpful.

1.3.1 *In-situ* studies

Sampling of the phytoplankton community and sediment traps North-west of Svalbard have revealed that the phytoplankton community composition is increasingly dominated by small haptophyte nanophytoplankton, such as *Emiliana huxleyi* and *Phaeocystis pouchetii* (Sakshaug et al. 2009; Pettersen et al. 2011; Nöthig et al. 2015). This implies that increased Atlantic inflow into the Barents Sea is importing temperate water species, recapitulating changes inferred from sedimentary cores of previous interglacials (Backman et al. 2009). Alternatively, enhanced physical stratification from increased surface warming may favour Phytoplankton Functional Types that are more suited to oligotrophy (e.g. stratification and resultant nutrient limitation causing a succession in Functional Types in the Norwegian Sea Frago et al. (2021). Indeed Frago et al. (2021) were able to show that bio-optical properties, such as the ratio of fluorescence to particular scattering, covaried with functional assemblages.). A signal of such a shift has already been reported in the Canada Basin (Li et al. 2009). Other than stratification dynamics, reduced nutrient content of Atlantic inflows (Hátún et al. 2017) might also explain declining end-winter nutrient inventories and resultant secular shifts in Phytoplankton Functional Types.

In addition to these pelagic changes, a variety of potential new styles of production have been attested under the sea-ice, such as diatom colonies forming under thinner sea-ice in the Nansen basin (Boetius et al. 2013), enhanced primary production north of the MIZ in the Fram Strait (Dybwad et al. 2020), and some authors have speculated that an increasingly thin and dynamic sea-ice cover will result in reduced snow-cover and more refrozen

leads, potentially transmitting more light and hence supporting greater ice-algal production (Kauko et al. 2017).

It can, however, routinely prove difficult to fully reconcile studies in differing locations or environmental contexts in the Barents Sea. Recent research at the shelf break North of Svalbard implies that nutrients are charged to higher concentrations over winter, and draw-down is more protracted in spring, when ice-cover is low (Henley et al. 2020), potentially contradicting our expectation of ongoing secular decline in end-winter nutrient inventories and a transition towards phytoplankton functional assemblages dominated by smaller cells and greater recycling of production. The authors suggest that increased Atlantic influence will boost access to nutrients, and that this may lead to enhanced nitrate uptake and possibly higher export production – a conclusion supported by other researchers working in the area (Randelhoff et al. 2018). These apparent contradictions may result from unique ecosystem function and environmental conditions in the shelf-break region (Personal correspondence, S.Henley 2020) which is unlikely to be representative of the broader shelf ecosystem. It is also worth considering that protracted nitrate draw-down is not necessarily indicative of greater integrated export production over the productive season, as the slower draw-down may instead result from ongoing remineralisation. Henley et al. (2020) argue that their findings must be integrated with changing stratification regimes and shifts in phytoplankton community composition. Indeed, a shift in community composition away from diatoms and towards smaller haptophyte phytoplankton species should be expected to result in increasingly regenerated production (Reigstad et al. 2002). Along with haptophyte genera, such as *Emiliania* and *Phaeocystis*, we might also expect picoeukaryotic prasinophytes, such as *Micromonas pusilla*, to be ascendant in a new Arctic Ocean. Indeed, recent discoveries of mixotrophy in this species (McKie-Krisberg and Sanders 2014) point to piceukaryotic prasinophytes potentially finding lucrative niches in a more oligotrophic Arctic Ocean recycling production.

Recent studies of bio-perturbation at the sedimentary interface imply that the benthic community's role in the fate of sinking carbon can be rapidly modified by changing environmental conditions, such as reduced ice-cover, even when the long term structure of benthic communities is relatively robust (Solan et al. 2020). A greater biomass and functional diversity among Northern benthic animal assemblages may make predicting these communities' responses to environmental change – and the significance they have for benthic production and carbon sequestration – particularly difficult, especially since it has proven challenging to relate the high degree of variability in Northern benthic communities to any salient environmental conditions (Souster et al. 2020). These studies, more than anything, reveal the unknown environmental dependencies and functional versatility in the benthic animal community. Indeed, distinct studies of benthic faunal assemblages in the same times and locations, using the same statistical methods³⁴ still disagree on the nature of variation in faunal composition. Solan et al. (2020) identify Northern and Southern clusters, while Souster et al. (2020) describes variation as clinal.

Faust et al. (2020) suggests, based on a study of sediment chemistry, that ongoing change in the Barents Sea may impede the long term sequestration of carbon in sediments. Sedimentary organic carbon content was found to be higher in historically ice-covered regions. Curiously, elevated Northern organics anti-correlate with calcareous sediment content³⁵ in the Southern Barents Sea. The authors attribute this to variation in the strength of pelagic-benthic coupling, or to a larger degree of recycling of sinking organics in the Southern Barents Sea – recapitulating suggestions by other authors of greater regeneration in this region (Degen et al. 2016)

In sum, *in-situ* studies provide an uncertain impression of the functional significance of ongoing environmental change in the Barents Sea. This results from an incomplete understanding of how food-web interactions determine the fate of fixed-carbon and how those

34. Non-Metric Multi-dimensional Scaling analysis.

35. Considered a proxy of primary productivity.

interactions are themselves being re-structured by anthropogenic perturbation.

1.3.2 Remote-sensing and modelling studies

Researchers might ask whether exploring different parameterisation schemes of ecological interactions in ecosystem models could increase understanding of how sensitive predictions of ecosystem change are to different hypotheses. They might also wonder whether seeming contradictions between *in-situ* studies might actually simply be the result of spatial heterogeneity in the Barents Sea (frontal, shelf-break or bank ecosystems might all fundamentally differ), and as such may seek means of monitoring synoptic-scale change in the Barents Sea. Model simulations offer the potential for synoptic studies, and remote-sensing approaches – though limited to the surface ocean – could also provide valuable means of measuring changes in phytoplankton distribution, concentration and phenology. Models and remote-sensing approaches coincide in some regards, as remote-sensing approaches may be used to validate models or may provide data that is assimilated into them directly.

Remote-sensing studies, which infer chlorophyll-a concentration ($[chl-a]$) from the colour of surface waters, have repeatedly indicated growing Arctic net primary production (e.g. Arrigo and Dijken (2011), Arrigo and Dijken (2015)). The astute reader should take note, however, that satellite-derived $[chl-a]$ can possess systematic bias (see discussion surrounding Figure 1.5), and that algorithms designed to infer primary production often make use of ancillary environmental variables to infer physiology and therefore the efficacy with which a standing stock of phytoplankton biomass converts downwelling irradiance into primary production. In Arrigo and Dijken (2015), for example, the daily primary production algorithm of Pabi et al. (2008) is employed:

$$PP = \int_{100}^{z=0} \int_{24}^{t=0} Chl(z) \frac{C}{Chl} G(z, t) dt dz, \quad (1.1)$$

where primary production (PP) is integrated between $z = 0$ and 100 m, over $t = 24$ hours, the carbon (C) to $[chl-a]$ (Chl) ratio is assumed to be 90 g g^{-1} , and physiology is parameterised by a growth term, G , varying over depth and time, that is itself an exponential function of SST and downwelling irradiance (vis-a-vis Arrhenius). It is therefore evident that, along with $[chl-a]$, variation in SST will be assumed to control overall primary production, and that assumptions researchers make about the vertical depth-distribution of $[chl-a]$ might significantly impact primary production estimates. It should therefore be emphasised that overall net primary production does not follow directly from surface $[chl-a]$. In particular, I petition the reader to consider that an assumption of a constant carbon to $[chl-a]$ ratio is a dramatic simplification (Johnsen and Sakshaug 2007; Brunet et al. 2011), and that physiological adaptations may vary widely among the phytoplankton. For example, a temperature dependence may have limited value in Arctic-influenced water-masses, where phytoplankton blooms are most intense at the coldest point of the year – showing remarkable agnosticism to SST (Hodal et al. 2012).

The findings of Arrigo and Dijken (2015) are, furthermore, partly driven by receding sea-ice revealing more ocean surface to sensors. Whether this represents enhanced production is therefore debatable, as it must be acknowledged under-ice production, which can be prolific (Arrigo et al. 2012; Dybwad et al. 2020; Ardyna et al. 2020), is neglected from consideration. In addition, there is reason to be sceptical that a longer period of ice-free conditions necessarily represents a longer period of new production rather than a longer period of regenerated cycling in the microbial loop. Of note, decreasing sea-ice results in greater annually-integrated remotely-sensed $[chl-a]$ in Barents Sea, but lower values in spring (Oziel et al. 2017). Researchers should be cautioned to interpret secular change in remotely-sensed $[chl-a]$ with care, as summer primary production is more likely to be regenerated.

Lewis et al. (2020) and Siswanto (2020) both provide remote-sensing evidence of a secular increase in $[chl-a]$ in the Southern and central Barents Sea. They argue that net primary production must be increasing, even if increasing open-water area is not considered. Lewis et al. (2020) develop upon this argument, making the case that increases in inferred $[chl-a]$ are limited to the Arctic inflow shelves and that this must necessarily mean that greater inflows of temperate nutrient-rich waters are supporting greater new production in these regions. However, the entirety of the Barents Sea is considered an inflow shelf in this analysis, and furthermore *in-situ* studies of the concentrations of crucial nutrients, such as silicic acid, indicate a secular decline in end-winter inventories (Rey 2012), potentially driven by ongoing anthropogenic changes to patterns of ocean circulation (Hátún et al. 2017).

It must also be considered that Arctic specialist phytoplankton and temperate phytoplankton species differ in their bio-optical properties (Stuart et al. 2000), and that previous authors have demonstrated that taxon-specific variation in bio-optical properties is a major driver of error in remotely-inferred $[chl-a]$ (Dutkiewicz et al. 2019). Global algorithms for the inference of $[chl-a]$, from remotely-sensed ocean-colour, are usually trained on empirical calibration match-ups between low-latitude $[chl-a]$ measurements and matching ocean-colour. Arctic phytoplankton assemblages have much higher pigment packaging³⁶ than low-latitude communities. Global algorithms may therefore under-estimate Arctic pigment concentrations by a factor of two, as a result of their lower chlorophyll-specific³⁷ absorption³⁸ (Cota et al. 2003). The dispersal of bio-optically aberrant temperate phytoplankton assemblages into the Barents Sea may therefore drive a spurious secular increase in perceived $[chl-a]$ concentration.

36. Absorption by a phytoplankton community *in-vivo*, in which pigments are packaged inside cells, compared to absorption of those pigments suspended in solution (Duyens 1956).

37. Absorption, normalised to $[chl-a]$. Commonly denoted $a_{ph}^*(\lambda)$ in the literature, measured in units of $m^2 mg^{-1} chl-a$.

38. Wavelength-dependent attenuation of light per unit distance (m^{-1}).

Ecosystems modelling approaches to simulate Barents Sea processes have yielded conclusions that are, in some cases, difficult to resolve with observational studies. Wassmann et al. (2006b) succeed in reproducing temporal and spatial modes of variability of suspended biomass, indicated by *in-situ* measurements, but their 3-dimensional model underestimates export production³⁹. Perhaps of crucial relevance to this underestimation, the fluid dynamical aspects of the simulation set-up fail to resolve sudden downwelling events, retention by grazing may be overestimated and May [*chl-a*] is underestimated. Authors, such as Jørgensen et al. (2014) and Degen et al. (2016), have suggested that export efficiency can be amplified during spring blooms by ice-edge physical interactions, and that this represents a leading explanation for relatively high benthic production in regions with high sea-ice persistence. Inadequate representation of ice-associated processes in ecosystem models may therefore contribute to a positive modelled relationship between temperature and export production (Wassmann et al. 2006b). There is hence a potential that incomplete representation of sub-mesoscale physics or parameterisations of trophic interactions could result in models providing conclusions antithetical to reality. However, even in so doing, modelling studies provide researchers with useful indications of which physical or ecosystem processes are not fully represented and provide a qualitative assessment of their relative importance.

Modelling approaches possess some advantages over remote-sensing inferences, as sub-surface processes, inaccessible to satellite-based observations, can be probed. Slagstad and Støle-Hansen (1991) employed 1 and 3-dimensional approaches to study the evolution of phytoplankton biomass and sedimentation in the context of the Barents Sea hydrography and were able to simulate the development of summer SCM blooms. Slagstad and Støle-Hansen (1991) commented that their modelling study was partially limited by insufficient knowledge of differing trophic interactions of diatom and *Phaeocystis* assemblages⁴⁰, but

39. The underestimate varies seasonally, but is as large as $1000 \text{ mg C m}^{-2} \text{ day}^{-1}$ in May – an underestimate of 80%.

40. In particular, mass sedimentation (Wassmann et al. 1990) may play a salient role in controlling pelagic-benthic coupling.

recent observational campaigns may have obviated these knowledge deficiencies (Dybwad et al. 2020).

Vernet et al. (2017) explicitly modelled changes in trophic interactions resulting from increasing temperatures favouring a transition towards *Phaeocystis* dominance over diatoms, in order to assess their potential significance. The approach of Vernet et al. (2017) required a combination of linear inverse modelling and also network analysis representing trophic interactions and therefore pathways of carbon transfer. Networks were based on unpublished empirical biological datasets gathered in the Fram Strait under different climatological conditions. These compartmentalised networks contained different phytoplankton size classes, zooplankton, fishes and heterotrophic bacteria. The flow of carbon through the networks was assumed to be in equilibrium and an ‘inverse analysis’ was then used to find solutions to the systems of equations describing these networks, vis-a-vis Vézina and Platt (1988). Vernet et al. (2017) found that, while trophic pathways of carbon transfer were substantially restructured, plasticity in zooplankton dietary preferences meant there was little overall change in carbon export. The study did, in line with predictions based on *in-situ* studies (Reigstad et al. 2002), predict that increased *Phaeocystis* dominance would result in enhanced microbial loop cycling. Hence, whether or not zooplankton buffer effects on overall carbon export, this modelling study may still provide a means by which remote-sensing studies reporting increasing primary production (e.g. Arrigo and Dijken 2015; Lewis et al. 2020) can be unified with *in-situ* studies that provide a much more ambiguous picture (e.g. Nöthig et al. 2015; Dybwad et al. 2020). A greater preponderance of microbial loop cycling would not contribute significantly to elevating export production. However, it would result in increases to remotely-sensed [*chl-a*] and therefore inference of higher net primary production – which may be incorrectly assumed to suggest increased export.

The recognition of different phytoplankton functional assemblages at scale in synoptic-scale remote-sensing observations would offer a capacity to test such hypotheses. Ap-

proaches by authors such as Galí et al. (2019) have indicated a pervasive shift towards *Phaeocystis* dominance, and therefore presumably greater microbial loop cycling, but Galí et al. (2019) use Mixed Layer Depth to proxy variation in community composition and phytoplankton functional variety is therefore not being retrieved directly.

1.3.3 The need to represent functional variety of phytoplankton in remote-sensing studies

It can be seen, from this previous discussion, that there is a requirement for functional variety in the phytoplankton to be adequately represented in remote-sensing studies, to ensure that they do not provide us with uncertain or even potentially misleading appraisals of ecosystem change. Researchers may, for example, incorrectly interpret enhanced cycling in the microbial loop as a signal of increasing new production.

Community composition and the influence it has on productivity, export and ecosystem function, can be highly multivariate and complex. It is not reasonable to expect this complete sum of information to be retrievable from ocean-colour datasets with only 10 degrees of freedom. Indeed, this information may not be retrievable even with hyper-spectral data (Cael et al. 2020). However, it may still be possible to contrive condensed descriptions of ecosystem state related to functional variety in the phytoplankton. While these classifications may prove simple, they would provide a much greater opportunity to develop insights into ecosystem change than having no available proxies of functional variety.

Discovering a means to define a representative scheme of such functional types, and to monitor them remotely, thus has an important value in providing information in order to create a synthesis in pursuit of resolving outstanding questions about ecosystem change in the Barents Sea.

1.3.4 Arctic PRIZE

A detailed review of the Arctic **PRIZE** (**P**roductivity in the Seasonal **I**ce **Z**one) project is provided in the Appendix.

The research chapters in this thesis make use of a myriad of environmental and biological datasets, collected under the auspices of Arctic **PRIZE**, with a view to better describing and understand the community compositional and phenological variability of the pelagic phytoplankton communities of the Barents Sea.

Sampling missions were undertaken between 2017 and 2018, with a focus on the productive spring, early and later summer phytoplankton communities, so that a full seasonal succession could be captured. Hydrographic properties, under water light environments and nutrient assays were all systematically recorded in order to fully characterise environmental conditions. Phytoplankton bio-optical properties, taxon-indicative pigments, optical microscopic and flow-cytometric cell-counts were all recorded to comprehensively document community variability.

The integration of this multitude of data streams with remotely-sensed ocean-colour data, and decadal-scale re-analyses of meteorological and oceanographic conditions, permit a unique perspective on the processes facilitating Arctic change. This perspective provides us with the opportunity to engage some of the unanswered questions and knowledge-deficiencies that we have identified.

1.3.5 Arctic PRIZE in a historic and contemporary context

Arctic **PRIZE** is among the latest of a range of multi-disciplinary observational campaigns aimed at increasing understanding of the hydrography and baseline ecosystem function of the Barents Sea. For example, the Pro Mare project (see Sakshaug et al. (1991) for a review), was conducted between 1984 and 1990, and at the time was one of the most sophisticated and wide-ranging scientific studies into pelagic ecology in the Barents Sea. As

well as attempting to describe baseline ecosystem function in several pelagic compartments (phytoplankton, microbes, zooplankton (Dalpadado and Skjoldal 1991), fish (Mehl 1991), pinnipeds), there was also an emphasis on probing the effects of anthropogenic perturbation, such as oil-slicks.

Key findings of the Pro Mare project included detailed physical descriptions of the hydrographic set-up and seasonal variation in the Barents Sea (Loeng 1991) and their consequence for the phenology of phytoplankton blooms.

Secular change in the physical hydrography of the Barents Sea has meant that there is a considerable risk of historic studies becoming obviated. The most salient changes have been the expansion of Atlantic Waters across a greater region of the shelf and a decline in sea-ice extents and durations, especially multi-year sea-ice. More recent scientific campaigns have therefore sought to investigate novel sea-ice dynamics and the influence these have for ecosystem processes in a Barents Sea whose Northern reaches are increasingly dominated by younger, thinner ice. These were the objectives of the Norwegian young sea ICE project. The insights of this campaign, which integrated a diverse set of data-streams, including atmospheric, pelagic, cryospheric and biological, have revealed a Northern Barents Sea where increased transfer of momentum from storms into the surface ocean has resulted in an increased deformational and dynamic nature to a thinner sea-ice cover (Itkin et al. 2017). The formation of ultra-transparent refrozen cracks and leads potentially provides ‘windows’ for the transfer of downwelling irradiance to ice-algae (Assmy et al. 2017), and there is an increasing focus in the literature on pelagic phytoplankton blooms being supported underneath a thinner more transparent ice-pack (Ardyna et al. 2020).

Ongoing research efforts include the largest scientific investigation into the polar regions in history; MOSAIC (Multidisciplinary Drifting Observatory for Study of Arctic Climate), which integrates researchers from 20 nations and has a more wide-reaching focus that aims to quantify the significance of Arctic change for global climate processes by collecting hy-

drographic and cryospheric measurements in sparsely-sampled regions near the North pole. MOSAIC began in 2019 and entailed maintaining an observational research platform (RV Polarstern) near the North pole for an entire year (Shupe et al. 2018). The long-duration of the mission also provided opportunity for researchers to study sympagic algal and animal communities and their over-wintering strategies.

In addition, the ongoing Nansen Legacy project (2018-2023) brings together hundreds of researchers from 10 Norwegian institutions, extending to include geologists in its multi-disciplinary milieu. This research programme, that takes advantage of the new ‘Kronprins Haakon’ research vessel and sophisticated submersible robotics, has already delivered wide-reaching insights into Arctic ecosystem and hydrographic dynamics in a changing climate, revealed by palaeontological and biogeochemical analysis of fossilised faunal assemblages indicative of bottom water conditions and sea-ice extents – demonstrating that sea-ice persisted even in previous thermal maxima between glacials (Pieńkowski et al. 2021).

While these sampling campaigns may differ individually in their focus, they are all necessarily multidisciplinary and integrate a diverse range of observational platforms, including semi- and autonomous robotics, with sophisticated laboratory, statistical analyses and mathematical modelling. This establishes a blueprint for a successful application of the diverse datasets collected under Arctic **PRIZE**.

1.4 Thesis overview

It can be seen, from the state of the literature discussed, that the study of phytoplankton phenology and community compositional variation from space offers immense value to monitor the base of Arctic marine ecosystems and will be an indispensable tool in the ongoing study of Arctic change. However, remote-sensing studies in the Barents Sea have rarely extended beyond the detection of coccolithophore blooms, even though it has been convention to represent functional assemblages such as diatoms and *Phaeocystis* in ecosystem models

for at least 3 decades. Developing means to distinguish diatoms and *Phaeocystis*-dominated functional assemblages offers particular value for ecosystem monitoring and model validation.

In chapter 2, a comprehensive collection of phytoplankton cell-counts, flow-cytometric measurements, taxon-indicative pigments and bio-optical properties are investigated and related to a concurrent environmental dataset of hydrographic conditions and nutrient concentrations. This facilitates the identification of functional associations in the phytoplankton. A total of 4 bio-optically recognisable and functionally-distinct phytoplankton assemblages are identified, including diatom and *Phaeocystis*-dominated assemblages. It is noteworthy that the retrieved assemblages can be conceptually related to Margalef’s mandala⁴¹. However, it is also found that different community compositional datasets possess differing capacities to retrieve environmental associations, or indeed to be successfully related to bio-optical properties. This raises the possibility that previous efforts in the literature to distinguish phytoplankton functional assemblages from space have been mired by error introduced by the conventional practice of using pigment ratios or cell-counts, rather than bio-optical properties, to validate the presence phytoplankton assemblages.

In chapter 3, the scheme of bio-optical Phytoplankton Functional Types that was developed in chapter 2 is used to compile a dataset of ocean-colour observations made with the MODIS-Aqua remote-sensing platform. Geo-spatial information and the date of sampling are used to label the ocean-colour observations with putative Phytoplankton Functional Type identities. The variance structure of this ocean-colour data set is explored with Linear Discriminant Analysis. While this proves to be a highly successful method to distinguish diatom and nanophytoplankton-dominated assemblages, difficulty remains discriminating diatom and *Phaeocystis*-dominated assemblages. This difficulty is obviated by the use of radiative-transfer modelling, based on the Inherent Optical Properties of these

41. A succession of phytoplankton functional assemblages related to declining ocean turbulence with the onset of stratification (Margalef 1978).

two assemblages. This demonstrates that they might be discerned from one another by *Phaeocystis*'s preferential absorption in the blue-green wavelengths, as a result of its investment in pigments such as chlorophyll-c3. Indeed, absorption of blue-green wavelengths has been observed to be a key mechanism underlying photoacclimation to shade in the genus *Phaeocystis* (Center 1987). An existing algorithm in the literature was successfully employed to demonstrate that this feature distinguished two unique phytoplankton assemblages. A multi-annual *in-situ* dataset of *Phaeocystis* abundance in the Labrador Sea was well-described at a basin scale by concomitant assessments of *Phaeocystis* dominance derived from ocean-colour. However, this effort required the careful correction of drift in the concurrent MODIS-Aqua ocean-colour processing. The absorption signal attributable to chlorophyll-c3 is so subtle that it can easily be eclipsed by relatively minor changes in sensor output attributable to degradation or stochastic error.

Chapter 4 takes advantage of the bio-optical algorithm developed in chapter 3, in order to produce a record of dominant Phytoplankton Functional Types in the Barents Sea in July, between 2002 and 2018. With the exception of coccolithophores, no sustained patterns in the geographic occurrence Phytoplankton Functional Types is found. A North-easterly progression of coccolithophores with enhanced Atlantic-inflow currents reproduces the consensus already identified in the literature. The state of the literature – a variety of *in-situ* measurements attesting a possible transition towards greater *Phaeocystis*-dominance in regions once more associated with diatom assemblages – is improved upon by retrieving evidence of a systemic secular increase in *Phaeocystis* prevalence in the West and Southern Barents Sea. These changes are compatible with predictions made by previous authors and have implications for the capacity of remote-sensing approaches that are agnostic to community composition to make faithful assessments of variables, such as new production. The potential implications baseline shifts in phytoplankton community composition have for pelagic-benthic coupling and over-arching ecosystem function are discussed in detail.

Finally, chapter 5 extends to a remote-sensing study of changing phytoplankton phenologies, following reports in the literature of the ascendancy of autumnal blooms on the Arctic shelves. It is currently hypothesised that these are driven by increasing action of wind upon an ever more ice-free Arctic Ocean, recapitulating temperate phenologies. However, it is found that the Barents Sea represents an exceptional case, whereby wind fields possess no significant or slight negative relationships to the intensity of autumnal blooms. Rather, it was found that the major driver of autumnal blooms in the Barents Sea is the average speed of inflowing currents – the main contributor to the Barents Sea nutrient inventory – in the Atlantic gateway onto the shelf. The significance of this mechanism is considered with careful reference to the established literature on the development of autumnal blooms and provides an increased understanding of incipient phenological change in the Barents Sea. The bio-optical algorithm that was applied to study bloom intensity was agnostic to phytoplankton community composition⁴². I did not, moreover, possess confidence that phytoplankton functional assemblages occurring in the autumn in the Barents Sea would be directly comparable to those that had been encountered in sampling missions concentrated in the spring and summer. The importance of autumnal sampling missions in a future Arctic, in which a greater share of annual primary production may occur in secondary autumnal blooms, is emphasised. Such studies are required to investigate whether there are unique phytoplankton functional assemblages occurring in this context and to discern their contributions to export production and any significance for the wider ecosystem.

42. On account of the decision to use a merged ocean-colour product that could not be interrogated by our bio-optical algorithm to retrieve functional types.

References

- Andreassen, Inger J, and Paul Wassmann. 1998. "Vertical flux of phytoplankton and particulate biogenic matter in the marginal ice zone of the Barents Sea in May 1993." *Marine Ecology Progress Series* 170:1–14.
- Ardyna, Mathieu, Marcel Babin, Michel Gosselin, Emmanuel Devred, Luc Rainville, and Jean-Éric Tremblay. 2014. "Recent Arctic Ocean sea ice loss triggers novel fall phytoplankton blooms." *Geophysical Research Letters* 41 (17): 6207–6212.
- Ardyna, Mathieu, CJ Mundy, Nicolas Mayot, Lisa C Matthes, Laurent Oziel, Christopher Horvat, Eva Leu, Philipp Assmy, Victoria Hill, Patricia A Matrai, et al. 2020. "Under-ice phytoplankton blooms: Shedding light on the "invisible" part of Arctic primary production." *Frontiers in Marine Science* 7:985.
- Arrigo, Kevin R, and Gert L van Dijken. 2015. "Continued increases in Arctic Ocean primary production." *Progress in Oceanography* 136:60–70.
- . 2011. "Secular trends in Arctic Ocean net primary production." *Journal of Geophysical Research: Oceans* 116 (C9).
- Arrigo, Kevin R, Patricia A Matrai, and Gert L Van Dijken. 2011. "Primary productivity in the Arctic Ocean: Impacts of complex optical properties and subsurface chlorophyll maxima on large-scale estimates." *Journal of Geophysical Research: Oceans* 116 (C11).
- Arrigo, Kevin R, Donald K Perovich, Robert S Pickart, Zachary W Brown, Gert L Van Dijken, Kate E Lowry, Matthew M Mills, Molly A Palmer, William M Balch, Frank Bahr, et al. 2012. "Massive phytoplankton blooms under Arctic sea ice." *Science* 336 (6087): 1408–1408.
- Assmy, Philipp, Mar Fernández-Méndez, Pedro Duarte, Amelie Meyer, Achim Randelhoff, Christopher J Mundy, Lasse M Olsen, Hanna M Kauko, Allison Bailey, Melissa Chierici, et al. 2017. "Leads in Arctic pack ice enable early phytoplankton blooms below snow-covered sea ice." *Scientific reports* 7 (1): 1–9.
- Backman, Jan, Eliana Fornaciari, and Domenico Rio. 2009. "Biochronology and paleoceanography of late Pleistocene and Holocene calcareous nannofossil abundances across the Arctic Basin." *Marine Micropaleontology* 72 (1-2): 86–98.
- Båmstedt, Ulf, Hans Christian Eilertsen, Kurt S Tande, Dag Slagstad, and Hein Rune Skjoldal. 1991. "Copepod grazing and its potential impact on the phytoplankton development in the Barents Sea." *Polar Research* 10 (2): 339–354.
- Barros, MP, M Pedersén, P Colepiccolo, and P Snoeijs. 2003. "Self-shading protects phytoplankton communities against H₂O₂-induced oxidative damage." *Aquatic microbial ecology* 30 (3): 275–282.
- Barton, Benjamin I, Yueng-Djern Lenn, and Camille Lique. 2020. "Observed atlantification of the Barents Sea causes the Polar Front to limit the expansion of winter sea ice." *JPO* 500:18–0003.
- Basedow, Sünne L, David McKee, Ina Lefering, Astthor Gislason, Malin Daase, Emilia Trudnowska, Einar Skarstad Egeland, Marvin Choquet, and Stig Falk-Petersen. 2019. "Remote sensing of zooplankton swarms." *Scientific Reports* 9 (1): 686.
- Behrenfeld, Michael J, and Paul G Falkowski. 1997. "Photosynthetic rates derived from satellite-based chlorophyll concentration." *Limnology and oceanography* 42 (1): 1–20.
- Behrenfeld, Michael J, Robert T O'Malley, David A Siegel, Charles R McClain, Jorge L Sarmiento, Gene C Feldman, Allen J Milligan, Paul G Falkowski, Ricardo M Letelier, and Emmanuel S Boss. 2006. "Climate-driven trends in contemporary ocean productivity." *Nature* 444 (7120): 752–755.
- Berge, Jørgen, Geir Johnsen, and Jonathan H Cohen. 2020. *POLAR NIGHT marine ecology: life and light in the dead of night*. Vol. 4. Springer Nature.

- Björk, Göran, Martin Jakobsson, Karen Assmann, Leif G Andersson, Johan Nilsson, Christian Stranne, and Larry Mayer. 2018. "Bathymetry and oceanic flow structure at two deep passages crossing the Lomonosov Ridge." *Ocean Science* 14 (1): 1–13.
- Boetius, Antje, Sebastian Albrecht, Karel Bakker, Christina Bienhold, Janine Felden, Mar Fernández-Méndez, Stefan Hendricks, Christian Katlein, Catherine Lalande, Thomas Krumpen, et al. 2013. "Export of algal biomass from the melting Arctic sea ice." *Science* 339 (6126): 1430–1432.
- Bouman, Heather A, Thomas Jackson, Shubha Sathyendranath, and Trevor Platt. 2020. "Vertical structure in chlorophyll profiles: influence on primary production in the Arctic Ocean." *Philosophical Transactions of the Royal Society A* 378 (2181): 20190351.
- Brown, Christopher W, Stephanie Schollaert Uz, and Bruce H Corliss. 2014. "Seasonality of oceanic primary production and its interannual variability from 1998 to 2007." *Deep Sea Research Part I: Oceanographic Research Papers* 90:166–175.
- Brunet, Christophe, Geir Johnsen, Johann Lavaud, and Suzanne Roy. 2011. "Pigments and photoacclimation processes." *Phytoplankton pigments: characterization, chemotaxonomy and applications in oceanography*.
- Buckley, Martha W, and John Marshall. 2016. "Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review." *Reviews of Geophysics* 54 (1): 5–63.
- Buesseler, Ken O, Philip W Boyd, Erin E Black, and David A Siegel. 2020. "Metrics that matter for assessing the ocean biological carbon pump." *Proceedings of the National Academy of Sciences* 117 (18): 9679–9687.
- Cael, B.B., Alison Chase, and Emmanuel Boss. 2020. "Information content of absorption spectra and implications for ocean color inversion." *Applied optics* 59 (13): 3971–3984.
- Carmack, Eddy. 2007. "The alpha/beta ocean distinction: A perspective on freshwater fluxes, convection, nutrients and productivity in high-latitude seas." *Deep Sea Research Part II: Topical Studies in Oceanography* 54 (23-26): 2578–2598.
- Carmack, Eddy, David Barber, Jens Christensen, Robie Macdonald, Bert Rudels, and Egil Sakshaug. 2006. "Climate variability and physical forcing of the food webs and the carbon budget on panarctic shelves." *Progress in Oceanography* 71 (2-4): 145–181.
- Carmack, Eddy, and Paul Wassmann. 2006. "Food webs and physical-biological coupling on pan-Arctic shelves: unifying concepts and comprehensive perspectives." *Progress in Oceanography* 71 (2-4): 446–477.
- Center, Scientific Data. 1987. "Spectral light absorption and quantum yield of photosynthesis in sea ice microalgae and a bloom of *Phaeocystis pouchetii* from McMurdo Sound, Antarctica." *Marine Ecology-Progress Series* 39:175–189.
- Charlson, Robert J, James E Lovelock, Meinrat O Andreae, and Stephen G Warren. 1987. "Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate." *Nature* 326 (6114): 655–661.
- Cota, Glenn F, W Glen Harrison, Trevor Platt, Shubha Sathyendranath, and Venetia Stuart. 2003. "Bio-optical properties of the Labrador Sea." *Journal of Geophysical Research: Oceans* 108 (C7).
- Cullen, John J. 2015. "Subsurface chlorophyll maximum layers: enduring enigma or mystery solved?"
- Dalpadado, Padmini, and Hein Rune Skjoldal. 1991. "Distribution and life history of krill from the Barents Sea." *Polar Research* 10 (2): 443–460.
- Degen, Renate, Lis Lindal Jørgensen, Pavel Ljubin, Ingrid H Ellingsen, Hendrik Pehlke, and Thomas Brey. 2016. "Patterns and drivers of megabenthic secondary production on the Barents Sea shelf." *Marine Ecology Progress Series* 546:1–16.
- Degerlund, Maria, and Hans Christian Eilertsen. 2010. "Main species characteristics of phytoplankton spring blooms in NE Atlantic and Arctic waters (68–80 N)." *Estuaries and coasts* 33 (2): 242–269.

- Denisenko, SG. 2001. "Long-term changes of zoobenthos biomass in the Barents Sea." *Proceedings of the Zoological Institute of the Russian Academy of Sciences* 289:59–66.
- Dong, Kaixing, Kristina Øie Kvile, Nils Chr Stenseth, and Leif Chr Stige. 2020. "Associations among temperature, sea ice and phytoplankton bloom dynamics in the Barents Sea." *Marine Ecology Progress Series* 635:25–36.
- Downes, Patrick P, Stephen J Goult, E Malcolm S Woodward, Claire E Widdicombe, Karen Tait, and Joanna L Dixon. 2020. "Phosphorus dynamics in the Barents Sea." *Limnology and Oceanography*.
- Drinkwater, Kenneth F. 2006. "The regime shift of the 1920s and 1930s in the North Atlantic." *Progress in Oceanography* 68 (2-4): 134–151.
- Dutkiewicz, Stephanie, Anna E Hickman, Oliver Jahn, Stephanie Henson, Claudie Beaulieu, and Erwan Monier. 2019. "Ocean colour signature of climate change." *Nature communications* 10 (1): 578.
- Duysens, LNM. 1956. "The flattening of the absorption spectrum of suspensions, as compared to that of solutions." *Biochimica et Biophysica Acta* 19:1–12.
- Dybwad, Christine, Philipp Assmy, Lasse M Olsen, Ilka Peeken, Anna Nikolopoulos, Thomas Krumpen, Achim Randelhoff, Agnieszka Tatarek, Józef M Wiktor, and Marit Reigstad. 2020. "Carbon Export in the Seasonal Sea Ice Zone North of Svalbard From Winter to Late-Summer." *Frontiers in Marine Science* 7:1137.
- Erga, Svein Rune, Nicolausi Ssebiyonga, Børge Hamre, Øyvind Frette, Francisco Rey, and Ken Drinkwater. 2014. "Nutrients and phytoplankton biomass distribution and activity at the Barents Sea Polar Front during summer near Hopen and Storbanken." *Journal of Marine Systems* 130:181–192.
- Eriksen, Elena, Hein Rune Skjoldal, Harald Gjøsæter, and Raul Primicerio. 2017. "Spatial and temporal changes in the Barents Sea pelagic compartment during the recent warming." *Progress in Oceanography* 151:206–226.
- Falkowski, Paul G, Richard T Barber, and Victor Smetacek. 1998. "Biogeochemical controls and feedbacks on ocean primary production." *Science* 281 (5374): 200–206.
- Faust, Johan C, Mark A Stevenson, Geoffrey D Abbott, Jochen Knies, Allyson Tessin, Isobel Mannion, Ailbe Ford, Robert Hilton, Jeffrey Peakall, and Christian März. 2020. "Does Arctic warming reduce preservation of organic matter in Barents Sea sediments?" *Philosophical Transactions of the Royal Society A* 378 (2181): 20190364.
- Fragoso, Glaucia M, Geir Johnsen, Matilde S Chauton, Finlo Cottier, and Ingrid Ellingsen. 2021. "Phytoplankton community succession and dynamics using optical approaches." *Continental Shelf Research* 213:104322.
- Francis, Jennifer A, and Elias Hunter. 2007. "Drivers of declining sea ice in the Arctic winter: A tale of two seas." *Geophysical Research Letters* 34 (17).
- Galí, Martí, Emmanuel Devred, Marcel Babin, and Maurice Levasseur. 2019. "Decadal increase in Arctic dimethylsulfide emission." *Proceedings of the National Academy of Sciences* 116 (39): 19311–19317.
- Gordon, Howard R, and WR McCluney. 1975. "Estimation of the depth of sunlight penetration in the sea for remote sensing." *Applied optics* 14 (2): 413–416.
- Grøsvik, Bjørn E, Tatiana Prokhorova, Elena Eriksen, Pavel Krivosheya, Per A Horneland, and Dmitry Prozorkevich. 2018. "Assessment of marine litter in the Barents Sea, a part of the Joint Norwegian–Russian Ecosystem Survey." *Frontiers in Marine Science* 5:72.
- Hansen, B, UC Berggreen, KS Tande, and H Chr Eilertsen. 1990. "Post-bloom grazing by *Calanus glacialis*, *C. finmarchicus* and *C. hyperboreus* in the region of the Polar Front, Barents Sea." *Marine Biology* 104 (1): 5–14.
- Hátún, Hjálmar, K Azetsu-Scott, R Somavilla, Francisco Rey, C Johnson, Moritz Mathis, Uwe Mikolajewicz, P Coupel, J-É Tremblay, S Hartman, et al. 2017. "The subpolar gyre regulates silicate concentrations in the North Atlantic." *Scientific reports* 7 (1): 1–9.

- Hegseth, Else Nøst. 1998. "Primary production of the northern Barents Sea." *Polar Research* 17 (2): 113–123.
- Henley, Sian F, Marie Porter, Laura Hobbs, Judith Braun, Robin Guillaume-Castel, Emily J Venables, Estelle Dumont, and Finlo Cottier. 2020. "Nitrate supply and uptake in the Atlantic Arctic sea ice zone: seasonal cycle, mechanisms and drivers." *Philosophical Transactions of the Royal Society A* 378 (2181): 20190361.
- Hill, Victoria J, Patricia A Matrai, Elise Olson, S Suttles, Mike Steele, LA Codispoti, and Richard C Zimmerman. 2013. "Synthesis of integrated primary production in the Arctic Ocean: II. In situ and remotely sensed estimates." *Progress in Oceanography* 110:107–125.
- Hodal, Helene, Stig Falk-Petersen, Haakon Hop, Svein Kristiansen, and Marit Reigstad. 2012. "Spring bloom dynamics in Kongsfjorden, Svalbard: nutrients, phytoplankton, protozoans and primary production." *Polar Biology* 35 (2): 191–203.
- Hop, Haakon, and Harald Gjøsæter. 2013. "Polar cod (*Boreogadus saida*) and capelin (*Malotus villosus*) as key species in marine food webs of the Arctic and the Barents Sea." *Marine Biology Research* 9 (9): 878–894.
- Hovland, Erlend Kjeldsberg, Heidi M Dierssen, Ana Sofia Ferreira, and Geir Johnsen. 2013. "Dynamics regulating major trends in Barents Sea temperatures and subsequent effect on remotely sensed particulate inorganic carbon." *Marine Ecology Progress Series* 484:17–32.
- Hovland, Erlend Kjeldsberg, Kasper Hancke, Morten Omholt Alver, Ken Drinkwater, Jo Høkedal, Geir Johnsen, Mark Moline, and Egil Sakshaug. 2014. "Optical impact of an *Emiliania huxleyi* bloom in the frontal region of the Barents Sea." *Journal of Marine Systems* 130:228–240.
- Itkin, Polona, Gunnar Spreen, Bin Cheng, Martin Doble, Fanny Girard-Ardhuin, Jari Haapala, Nick Hughes, Lars Kaleschke, Marcel Nicolaus, and Jeremy Wilkinson. 2017. "Thin ice and storms: Sea ice deformation from buoy arrays deployed during N-ICE 2015." *Journal of Geophysical Research: Oceans* 122 (6): 4661–4674.
- Jackson, Thomas, Heather A Bouman, Shubha Sathyendranath, and Emmanuel Devred. 2010. "Regional-scale changes in diatom distribution in the Humboldt upwelling system as revealed by remote sensing: implications for fisheries." *ICES Journal of Marine Science* 68 (4): 729–736.
- Jiang, Lin, Oscar ME Schofield, and Paul G Falkowski. 2005. "Adaptive evolution of phytoplankton cell size." *The American Naturalist* 166 (4): 496–505.
- Johnsen, Geir, Marit Norli, Mark Moline, Ian Robbins, Cecilie von Quillfeldt, Kai Sørensen, Finlo Cottier, and Jørgen Berge. 2018. "The advective origin of an under-ice spring bloom in the Arctic Ocean using multiple observational platforms." *Polar Biology* 41 (6): 1197–1216.
- Johnsen, Geir, and Egil Sakshaug. 2007. "Biooptical characteristics of PSII and PSI in 33 species (13 pigment groups) of marine phytoplankton, and the relevance for pulse-amplitude-modulated and fast-repetition-rate fluorometry 1." *Journal of Phycology* 43 (6): 1236–1251.
- Jørgensen, Lis Lindal, Pavel Ljubin, Hein Rune Skjoldal, Randi B Ingvaldsen, Natalia Anisimova, and Igor Manushin. 2014. "Distribution of benthic megafauna in the Barents Sea: baseline for an ecosystem approach to management." *ICES Journal of Marine Science* 72 (2): 595–613.
- Jørgensen, Lis Lindal, Benjamin Planque, Trude Hauge Thangstad, and Grégoire Certain. 2016. "Vulnerability of megabenthic species to trawling in the Barents Sea." *ICES Journal of Marine Science* 73 (suppl.1): i84–i97.
- Karcher, Michael, John N Smith, Frank Kauker, Rüdiger Gerdes, and William M Smethie. 2012. "Recent changes in Arctic Ocean circulation revealed by iodine-129 observations and modeling." *Journal of Geophysical Research: Oceans* 117 (C8).

- Kauko, Hanna M, Lasse M Olsen, Pedro Duarte, Ilka Peeken, Mats A Granskog, Geir Johnsen, Mar Fernández-Méndez, Alexey K Pavlov, Christopher J Mundy, and Philipp Assmy. 2018. "Algal colonization of young Arctic sea ice in spring." *Frontiers in Marine Science* 5:199.
- Kauko, Hanna M, Alexey K Pavlov, Geir Johnsen, Mats A Granskog, Ilka Peeken, and Philipp Assmy. 2019. "Photoacclimation state of an Arctic underice phytoplankton bloom." *Journal of Geophysical Research: Oceans* 124 (3): 1750–1762.
- Kauko, Hanna M, Torbjørn Taskjelle, Philipp Assmy, Alexey K Pavlov, CJ Mundy, Pedro Duarte, Mar Fernández-Méndez, Lasse M Olsen, Stephen R Hudson, Geir Johnsen, et al. 2017. "Windows in Arctic sea ice: Light transmission and ice algae in a refrozen lead." *Journal of Geophysical Research: Biogeosciences* 122 (6): 1486–1505.
- Kohlbach, Doreen, Martin Graeve, Benjamin A. Lange, Carmen David, Ilka Peeken, and Hauke Flores. 2016. "The importance of ice algae-produced carbon in the central Arctic Ocean ecosystem: Food web relationships revealed by lipid and stable isotope analyses." *Limnology and Oceanography* 61 (6): 2027–2044.
- Kostakis, Ina, Rüdiger Röttgers, Andrew Orkney, Heather A Bouman, Marie Porter, Finlo Cottier, Jørgen Berge, and David McKee. 2020. "Development of a bio-optical model for the Barents Sea to quantitatively link glider and satellite observations." *Philosophical Transactions of the Royal Society A* 378 (2181): 20190367.
- Kristiansen, Svein, and Tove Farbrot. 1991. "Nitrogen uptake rates in phytoplankton and ice algae in the Barents Sea." *Polar research* 10 (1): 187–192.
- Kruppen, Thomas, H Jakob Belter, Antje Boetius, Ellen Damm, Christian Haas, Stefan Hendricks, Marcel Nicolaus, Eva-Maria Nöthig, Stephan Paul, Ilka Peeken, et al. 2019. "Arctic warming interrupts the Transpolar Drift and affects long-range transport of sea ice and ice-rafted matter." *Scientific Reports* 9 (1): 5459.
- Kwok, Ron. 2018. "Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018)." *Environmental Research Letters* 13 (10): 105005.
- Laws, Edward A, Paul G Falkowski, Walker O Smith Jr, Hugh Ducklow, and James J McCarthy. 2000. "Temperature effects on export production in the open ocean." *Global Biogeochemical Cycles* 14 (4): 1231–1246.
- Le Moigne, Frederic AC, Alex J Poulton, Stephanie A Henson, Chris J Daniels, Glaucia M Fragoso, Elaine Mitchell, Sophie Richier, Benjamin C Russell, Helen EK Smith, Geraint A Tarling, et al. 2015. "Carbon export efficiency and phytoplankton community composition in the Atlantic sector of the Arctic Ocean." *Journal of Geophysical Research: Oceans* 120 (6): 3896–3912.
- Lee, Craig M, Jim Thomson, Marginal Ice Zone, and Arctic Sea State Teams. 2017. "An autonomous approach to observing the seasonal ice zone in the western Arctic." *Oceanography* 30 (2): 56–68.
- Lewis, KM, GL van Dijken, and KR Arrigo. 2020. "Changes in phytoplankton concentration now drive increased Arctic Ocean primary production." *Science* 369 (6500): 198–202.
- Li, William KW, Fiona A McLaughlin, Connie Lovejoy, and Eddy C Carmack. 2009. "Smallest algae thrive as the Arctic Ocean freshens." *Science* 326 (5952): 539–539.
- Lien, Vidar S, Frode B Vikebø, and Øystein Skagseth. 2013. "One mechanism contributing to co-variability of the Atlantic inflow branches to the Arctic." *Nature Communications* 4 (1): 1–6.
- Liu, Xiaohan, Emmanuel Devred, and Catherine Johnson. 2018. "Remote sensing of phytoplankton size class in Northwest Atlantic from 1998 to 2016: bio-optical algorithms comparison and application." *Remote Sensing* 10 (7): 1028.
- Loeng, Harald. 1991. "Features of the physical oceanographic conditions of the Barents Sea." *Polar Research* 10 (1): 5–18.
- Lovejoy, Connie, Warwick F Vincent, Sylvia Bonilla, Suzanne Roy, Marie-Josée Martineau, Ramon Terrado, Marianne Potvin, Ramon Massana, and Carlos Pedrós-Alió. 2007.

- “Distribution, phylogeny, and growth of cold-adapted picoprasinophytes in Arctic seas 1.” *Journal of Phycology* 43 (1): 78–89.
- Margalef, R. 1978. “Life forms of phytoplankton as survival alternatives in an unstable environment.” *Oceanologica Acta* 1:4–493.
- Martin, Seelye, and Donald J Cavalieri. 1989. “Contributions of the Siberian shelf polynyas to the Arctic Ocean intermediate and deep water.” *Journal of Geophysical Research: Oceans* 94 (C9): 12725–12738.
- May, Brian D, and Dan E Kelley. 2001. “Growth and steady state stages of thermohaline intrusions in the Arctic Ocean.” *Journal of Geophysical Research: Oceans* 106 (C8): 16783–16794.
- McKie-Krisberg, Zaid M, and Robert W Sanders. 2014. “Phagotrophy by the picoeukaryotic green alga *Micromonas*: implications for Arctic Oceans.” *The ISME journal* 8 (10): 1953–1961.
- Mehl, Sigbjørn. 1991. “The Northeast Arctic cod stock’s place in the Barents Sea ecosystem in the 1980s: an overview.” *Polar research* 10 (2): 525–534.
- Mosby, Håkon. 1962. “Water, salt and heat balance of the North Polar Sea and of the Norwegian Sea.” *Geophysica Norvegica* 24 (11): 289–313.
- Neukermans, Griet, Laurent Oziel, and Marcel Babin. 2018. “Increased intrusion of warming Atlantic water leads to rapid expansion of temperate phytoplankton in the Arctic.” *Global change biology* 24 (6): 2545–2553.
- Nöthig, Eva-Maria, Astrid Bracher, Anja Engel, Katja Metfies, Barbara Niehoff, Ilka Peeken, Eduard Bauerfeind, Alexandra Cherkasheva, Steffi Gäbler-Schwarz, Kristin Hardge, et al. 2015. “Summertime plankton ecology in Fram Strait- a compilation of long and short-term observations.” *Polar Research* 34 (1): 23349.
- Olsen, Are, Leif G Anderson, and Truls Johannessen. 2002. “The impact of climate variations on fluxes of oxygen in the Barents Sea.” *Continental shelf research* 22 (8): 1117–1128.
- Olsen, Lasse M, Samuel R Laney, Pedro Duarte, Hanna M Kauko, Mar Fernández-Méndez, Christopher J Mundy, Anja Rösel, Amelie Meyer, Polona Itkin, Lana Cohen, et al. 2017. “The seeding of ice algal blooms in Arctic pack ice: the multiyear ice seed repository hypothesis.” *Journal of Geophysical Research: Biogeosciences* 122 (7): 1529–1548.
- O’Reilly, John E, Stephane Maritorena, B Greg Mitchell, David A Siegel, Kendall L Carder, Sara A Garver, Mati Kahru, and Charles McClain. 1998. “Ocean color chlorophyll algorithms for SeaWiFS.” *Journal of Geophysical Research: Oceans* 103 (C11): 24937–24953.
- Orkney, Andrew, Trevor Platt, Bhavani E Narayanaswamy, Ina Kostakis, and Heather A Bouman. 2020. “Bio-Optical evidence for increasing *Phaeocystis* dominance in the Barents Sea.” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*.
- Oziel, Laurent, Alberto Baudena, Mathieu Ardyna, Philippe Massicotte, A Randelhoff, J-B Sallée, Randi Brunvær Ingvaldsen, Emmanuel Devred, and Marcel Babin. 2020. “Faster Atlantic currents drive poleward expansion of temperate phytoplankton in the Arctic Ocean.” *Nature communications* 11 (1): 1–8.
- Oziel, Laurent, G Neukermans, M Ardyna, C Lancelot, J-L Tison, Paul Wassmann, Jérôme Sirven, Diana Ruiz-Pino, and J-C Gascard. 2017. “Role for Atlantic inflows and sea ice loss on shifting phytoplankton blooms in the Barents Sea.” *Journal of Geophysical Research: Oceans* 122 (6): 5121–5139.
- Oziel, Laurent, Jérôme Sirven, and Jean-Claude Gascard. 2016. “The Barents Sea frontal zones and water masses variability (1980–2011).” *Ocean Science* 12(1):169–184.
- Pabi, Sudeshna, Gert L van Dijken, and Kevin R Arrigo. 2008. “Primary production in the Arctic Ocean, 1998–2006.” *Journal of Geophysical Research: Oceans* 113 (C8).

- Pante, Eric, and Benoit Simon-Bouhet. 2013. "marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R." *PLoS ONE* 8 (9): e73051. doi:10.1371/journal.pone.0073051.
- Park, Jong-Yeon, Jong-Seong Kug, Jürgen Bader, Rebecca Rolph, and Minho Kwon. 2015. "Amplified Arctic warming by phytoplankton under greenhouse warming." *Proceedings of the National Academy of Sciences* 112 (19): 5921–5926.
- Pavlov, Alexey K, Torbjørn Taskjelle, Hanna M Kauko, Børge Hamre, Stephen R Hudson, Philipp Assmy, Pedro Duarte, Mar Fernández-Méndez, CJ Mundy, and Mats A Granskog. 2017. "Altered inherent optical properties and estimates of the underwater light field during an Arctic under-ice bloom of *Phaeocystis pouchetii*." *Journal of Geophysical Research: Oceans* 122 (6): 4939–4961.
- Peperzak, L, F Colijn, EG Vrieling, WWC Gieskes, and JCH Peeters. 2000a. "Observations of flagellates in colonies of *Phaeocystis globosa* (Prymnesiophyceae); a hypothesis for their position in the life cycle." *Journal of Plankton Research* 22 (12): 2181–2203.
- Pettersen, Ragnhild, Geir Johnsen, Jørgen Berge, and Erlend Kjeldsberg Hovland. 2011. "Phytoplankton chemotaxonomy in waters around the Svalbard archipelago reveals high amounts of Chl b and presence of gyroxanthin-diester." *Polar biology* 34 (5): 627–635.
- Pieńkowski, Anna J, Katrine Husum, Simon T Belt, Ulysses Ninnemann, Denizcan Köseoğlu, Dmitry V Divine, Lukas Smik, Jochen Knies, Kelly Hogan, and Riko Noormets. 2021. "Seasonal sea ice persisted through the Holocene Thermal Maximum at 80 N." *Communications Earth & Environment* 2 (1): 1–10.
- Piepenburg, D, T.H. Blackburn, C.F. von Dorrien, J Gutt, O.J. Hall, S Hulth, M.A. Kendall, K.W. Opalinski, E. Rachor, and K.M. Schmid. 1995. "Partitioning of benthic community respiration in the Arctic (northwestern Barents Sea)." *Marine Ecology Progress Series* 118:199–213.
- Polyakov, Igor V, Andrey V Pnyushkov, Matthew B Alkire, Igor M Ashik, Till M Baumann, Eddy C Carmack, Ilona Goszczko, John Guthrie, Vladimir V Ivanov, Torsten Kanzow, et al. 2017. "Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean." *Science* 356 (6335): 285–291.
- Porter, Marie, SF Henley, Andrew Orkney, HA Bouman, B Hwang, E Dumont, EJ Venables, and F Cottier. 2020. "A polar surface eddy obscured by thermal stratification." *Geophysical Research Letters* 47 (6): e2019GL086281.
- Prowse, TD, and PO Flegg. 2000. "The magnitude of river flow to the Arctic Ocean: dependence on contributing area." *Hydrological Processes* 14 (16-17): 3185–3188.
- Quere, Corinne Le, Sandy P Harrison, I Colin Prentice, Erik T Buitenhuis, Olivier Aumont, Laurent Bopp, Hervé Claustre, Leticia Cotrim Da Cunha, Richard Geider, Xavier Giraud, et al. 2005. "Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models." *Global Change Biology* 11 (11): 2016–2040.
- Randelhoff, Achim, Marit Reigstad, Melissa Chierici, Arild Sundfjord, Vladimir Ivanov, Mattias Rolf Cape, Maria Vernet, Jean-Éric Tremblay, Gunnar Bratbak, and Svein Kristiansen. 2018. "Seasonality of the physical and biogeochemical hydrography in the inflow to the Arctic Ocean through Fram Strait." *Frontiers in Marine Science* 5:224.
- Rat'kova, Tatjana N, and Paul Wassmann. 2002. "Seasonal variation and spatial distribution of phyto- and protozooplankton in the central Barents Sea." *Journal of Marine Systems* 38 (1-2): 47–75.
- Reigstad, Marit, and Paul Wassmann. 2007. "Does *Phaeocystis* spp. contribute significantly to vertical export of organic carbon?" In *Phaeocystis, major link in the biogeochemical cycling of climate-relevant elements*, 217–234. Springer.
- Reigstad, Marit, Paul Wassmann, Christian Wexels Riser, Sigrid Øygarden, and Francisco Rey. 2002. "Variations in hydrography, nutrients and chlorophyll a in the marginal ice-zone and the central Barents Sea." *Journal of Marine Systems* 38 (1-2): 9–29.

- Renaud, Paul E, Nathalie Morata, Michael L Carroll, Stanislav G Denisenko, and Marit Reigstad. 2008. "Pelagic–benthic coupling in the western Barents Sea: processes and time scales." *Deep Sea Research Part II: Topical Studies in Oceanography* 55 (20-21): 2372–2380.
- Rey, Francisco. 2012. "Declining silicate concentrations in the Norwegian and Barents Seas." *ICES Journal of Marine Science* 69 (2): 208–212.
- Riser, Christian Wexels, Marit Reigstad, Paul Wassmann, Elena Arashkevich, and Stig Falk-Petersen. 2007. "Export or retention? Copepod abundance, faecal pellet production and vertical flux in the marginal ice zone through snap shots from the northern Barents Sea." *Polar Biology* 30 (6): 719–730.
- Riser, Christian Wexels, Paul Wassmann, Kalle Olli, Anna Pasternak, and Elena Arashkevich. 2002. "Seasonal variation in production, retention and export of zooplankton faecal pellets in the marginal ice zone and central Barents Sea." *Journal of Marine Systems* 38 (1-2): 175–188.
- Rudels, Bert, EP Jones, LG Anderson, and Gerhard Kattner. 1994. "On the intermediate depth waters of the Arctic Ocean." *The polar oceans and their role in shaping the global environment* 85:33–46.
- Sakshaug, E. 2004. "Primary and secondary production in the Arctic Seas." In *The organic carbon cycle in the Arctic Ocean*, 57–81. Springer.
- Sakshaug, Egil, Christopher CE Hopkins, and Nils Are Øritsland. 1991. "Proceedings of the Pro Mare Symposium on Polar Marine Ecology, Trondheim, Norway, 12-16 May 1990." *Polar Research* 10 (1): 1–4.
- Sakshaug, Egil, Geir Helge Johnsen, and Kit M Kovacs. 2009. *Ecosystem Barents Sea*. Tapir Academic Press.
- Sathyendranath, Shubha, Jim Aiken, S Alvain, R Barlow, H Bouman, Astrid Bracher, R Brewin, Annick Bricaud, CW Brown, AM Ciotti, et al. 2014. "Phytoplankton functional types from Space." In *(Reports of the International Ocean-Colour Coordinating Group (IOCCG); 15)*, 1–156. International Ocean-Colour Coordinating Group.
- Sathyendranath, Shubha, Louisa Watts, Emmanuel Devred, Trevor Platt, Carla Caverhill, and Heidi Maass. 2004. "Discrimination of diatoms from other phytoplankton using ocean-colour data." *Marine Ecology Progress Series* 272:59–68.
- Sazhin, Andrey F, L Felipe Artigas, Jens C Nejstgaard, and Marc E Frischer. 2007. "The colonization of two *Phaeocystis* species (Prymnesiophyceae) by pennate diatoms and other protists: a significant contribution to colony biomass." In *Phaeocystis, major link in the biogeochemical cycling of climate-relevant elements*, 137–145. Springer.
- Schauer, Ursula, Harald Loeng, Bert Rudels, Vladimir K Ozhigin, and Wolfgang Dieck. 2002. "Atlantic water flow through the Barents and Kara Seas." *Deep Sea Research Part I: Oceanographic Research Papers* 49 (12): 2281–2298.
- Schlichtholz, Pawel, and Marie-Noëlle Houssais. 1999. "An inverse modeling study in Fram Strait. Part I: dynamics and circulation." *Deep Sea Research Part II: Topical Studies in Oceanography* 46 (6-7): 1083–1135.
- Shapiro, Inna, Roger Colony, and Torgny Vinje. 2003. "April sea ice extent in the Barents Sea, 1850–2001." *Polar Research* 22 (1): 5–10.
- Shupe, Matthew, Gijs de Boer, Klaus Dethloff, Elizabeth Hunke, Wieslaw Maslowski, Allison McComiskey, Ola Perrson, David Randall, Michael Tjernstrom, David Turner, et al. 2018. *The Multidisciplinary Drifting Observatory for the Study of Arctic Climate (MOSAIC) Atmosphere Science Plan*. Technical report. DOE Office of Science Atmospheric Radiation Measurement (ARM) Program . . .
- Shutler, JD, PE Land, CW Brown, HS Findlay, CJ Donlon, M Medland, R Snooke, and JC Blackford. 2013. "Coccolithophore surface distributions in the North Atlantic and their modulation of the air-sea flux of CO₂ from 10 years of satellite Earth observation data." *Biogeosciences* 10 (4): 2699–2709.

- Siswanto, Eko. 2020. "Temporal variability of satellite-retrieved chlorophyll-a data in Arctic and subarctic ocean regions within the past two decades." *International Journal of Remote Sensing* 41 (19): 7427–7445.
- Slagstad, Dag, and Kjell Støle-Hansen. 1991. "Dynamics of plankton growth in the Barents Sea: model studies." *Polar Research* 10 (1): 173–186.
- Smyth, TJ, T Tyrrell, and B Tarrant. 2004. "Time series of coccolithophore activity in the Barents Sea, from twenty years of satellite imagery." *Geophysical Research Letters* 31 (11).
- Solan, Martin, Ellie R Ward, Christina L Wood, Adam J Reed, Laura J Grange, and Jasmin A Godbold. 2020. "Climate-driven benthic invertebrate activity and biogeochemical functioning across the Barents Sea polar front." *Philosophical Transactions of the Royal Society A* 378 (2181): 20190365.
- Souster, TA, DKA Barnes, and J Hopkins. 2020. "Variation in zoobenthic blue carbon in the Arctic's Barents Sea shelf sediments." *Philosophical Transactions of the Royal Society A* 378 (2181): 20190362.
- Stuart, V, S Sathyendranath, EJH Head, T Platt, B Irwin, and H Maass. 2000. "Bio-optical characteristics of diatom and prymnesiophyte populations in the Labrador Sea." *Marine Ecology Progress Series* 201:91–106.
- Sundfjord, Arild, Ilker Fer, Yoshie Kasajima, and Harald Svendsen. 2007. "Observations of turbulent mixing and hydrography in the marginal ice zone of the Barents Sea." *Journal of Geophysical Research: Oceans* 112 (C5).
- Syvertsen, Erik E. 1991. "Ice algae in the Barents Sea: types of assemblages, origin, fate and role in the ice-edge phytoplankton bloom." *Polar Research* 10 (1): 277–288.
- Våge, Kjetil, Robert S Pickart, Vladimir Pavlov, Peigen Lin, Daniel J Torres, Randi Ingvaldsen, Arild Sundfjord, and Andrey Proshutinsky. 2016. "The Atlantic Water boundary current in the Nansen Basin: Transport and mechanisms of lateral exchange." *Journal of Geophysical Research: Oceans* 121 (9): 6946–6960.
- Vernet, Maria, Tammi L Richardson, Katja Metfies, Eva-Maria Nöthig, and Ilka Peeken. 2017. "Models of plankton community changes during a warm water anomaly in Arctic waters show altered trophic pathways with minimal changes in carbon export." *Frontiers in Marine Science* 4:160.
- Vézina, Alain F, and Trevor Platt. 1988. "Food web dynamics in the ocean. 1. Best-estimates of flow networks using inverse methods." *Marine ecology progress series. Oldendorf* 42 (3): 269–287.
- Wassmann, Paul, Tatjana Ratkova, Inger Andreassen, Maria Vernet, Gunnar Pedersen, and Francisco Rey. 1999. "Spring bloom development in the marginal ice zone and the central Barents Sea." *Marine Ecology* 20 (3-4): 321–346.
- Wassmann, Paul, Marit Reigstad, Tore Haug, Bert Rudels, Michael L Carroll, Haakon Hop, Geir Wing Gabrielsen, Stig Falk-Petersen, Stanislav G Denisenko, Elena Arashkevich, et al. 2006a. "Food webs and carbon flux in the Barents Sea." *Progress in Oceanography* 71 (2-4): 232–287.
- Wassmann, Paul, Dag Slagstad, and Ingrid Ellingsen. 2010. "Primary production and climatic variability in the European sector of the Arctic Ocean prior to 2007: preliminary results." *Polar Biology* 33 (12): 1641–1650.
- Wassmann, Paul, Dag Slagstad, Christian Wexels Riser, and Marit Reigstad. 2006b. "Modelling the ecosystem dynamics of the Barents Sea including the marginal ice zone: II. Carbon flux and interannual variability." *Journal of Marine Systems* 59 (1-2): 1–24.
- Wassmann, Paul, Maria Vernet, B Greg Mitchell, and Francisco Rey. 1990. "Mass sedimentation of *Phaeocystis pouchetii* in the Barents Sea." *Marine Ecology Progress Series* 66:183–195.
- Weisse, T, N Grimm, W Hickel, and P Martens. 1986. "Dynamics of *Phaeocystis pouchetii* blooms in the Wadden Sea of Sylt (German Bight, North Sea)." *Estuarine, Coastal and Shelf Science* 23 (2): 171–182.

- Wernand, Marcel R, Hendrik J van der Woerd, and Winfried WC Gieskes. 2013. "Trends in ocean colour and chlorophyll concentration from 1889 to 2000, worldwide." *PLOS one* 8 (6).
- Wernand, MR, and HJ Van der Woerd. 2010. "Spectral analysis of the Forel-Ule Ocean colour comparator scale." *Journal of the European Optical Society-Rapid Publications* 5.
- Wollenburg, JE, C Katlein, Gernot Nehrke, E-M Nöthig, Jens Matthiessen, Dieter A Wolf-Gladrow, A Nikolopoulos, F Gázquez-Sanchez, L Rossmann, Philipp Assmy, et al. 2018. "Ballasting by cryogenic gypsum enhances carbon export in a *Phaeocystis* under-ice bloom." *Scientific Reports* 8 (1): 7703.

Chapter 2

Describing an ecosystem: Phytoplankton Functional Types

Andrew Orkney^{1,a}, Keith Davidson^{2,b}, Elaine Mitchell^{2,c},
Sian Henley^{3,d} & Heather Bouman^{1,e}

^a: Corresponding author, andrew.orkney@earth.ox.ac.uk

<https://orcid.org/0000-0003-4972-2541>

^b: <https://orcid.org/0000-0001-9269-3227>

^c: <https://orcid.org/0000-0001-7274-1009>

^d: <https://orcid.org/0000-0003-1221-1983>

^e: <https://orcid.org/0000-0002-7407-9431>

¹: Department of Earth Sciences, University of Oxford, 3 South Parks Road, OX1 3AN

²: The Scottish Association for Marine Science, Oban, PA37 1QA

³: School of Geosciences, University of Edinburgh, Drummond Street, EH8 9XP

2.1 Prelude

This chapter represents research that we anticipate submitting to *Limnology and Oceanography* in the coming months.

This chapter takes a foundational role in the structure of my thesis, exploring the variety of assemblages and functional associations in Barents Sea phytoplankton communities encountered in sampling under the auspices of Arctic **PRIZE**. A scheme of bio-optical functional types (clusters with distinct absorption spectra – $a_{ph}(\lambda)$) is identified that conforms well to the conceptual cascade of phytoplankton assemblages proposed under Margalef’s mandala, raising the possibility that the identification of these bio-optically distinct types in remotely-sensed ocean-colour would provide exceptional insight into ecosystem state in the Barents Sea. Moreover, the identification of diatom and *Phaeocystis*-dominated assemblages recapitulates directly the choice of community compositional variation that is conventionally selected in modelling approaches – which means that there is a potential for approaches based on these bio-optical types to be employed in the validation of or assimilation into ecosystem models. The scheme of functional types identified here is used as a foundation in subsequent chapters to develop bio-optical algorithms (a set of computations that can be undertaken to discern, from remotely-detected ocean-colour, derived variables about ecosystem state such as $[chl-a]$ or dominant phytoplankton size class) to retrieve Phytoplankton Functional Types from remotely-sensed ocean-colour, and thereafter to explore secular change in ecosystem state in the Barents Sea over a decadal time-scale.

2.2 Significance statement

Phytoplankton community structure is essential to understanding primary production, nutrient utilisation, export and the sustenance of higher trophic levels and hence overall

ecosystem structure, ecosystems services and biogeochemical processes. Owing to its complexity, phytoplankton community structure is usually represented by simplified schemes of categories or ‘Functional Types’ that each perform distinct and important roles in an ecosystem by way of their differing utilisation of nutrients, sinking rates, or food-web interactions (Quere et al. 2005; McGill et al. 2006). A commonly-employed suite of Functional Types distinguishes different assemblages of phytoplankton by their size for example (Liu et al. 2018), because this property has high relevance to the sinking rate of fixed carbon and its transfer through ecosystems (Armstrong et al. 2009). Such representations are key to the implementation of ecosystem models and ecosystem monitoring.

The identification of distinct Phytoplankton Functional Types in remotely-sensed ocean-colour has therefore gained attention as a goal in biological oceanography (e.g. Sathyendranath et al. (2004) Hovland et al. (2013) Liu et al. (2018)), because of the capacity this would offer to facilitate synoptic real-time monitoring of ocean ecosystem processes and the validation of ecosystem models. However, attempts to use remotely-sensed ocean-colour to distinguish Phytoplankton Functional Types have had mixed success (Brewin et al. 2011). Hence the identification of the reasons for difficulties achieving this goal have special significance for overcoming obstacles to an improved understanding of our oceans.

Understanding boreal shelf seas is of particular importance because of the wide variety of functional roles the phytoplankton of these seas perform (e.g. dimethyl-sulphide (DMS) production, calcification that manifests as ocean whittings) and the great Northern fisheries they support owing to their high productivity (Drinkwater 2006; Wassmann et al. 2006a). The occasional occurrence of ichthyotoxic blooms in Northern coastal waters (Johnsen et al. 1999) may also make the detection of Harmful Algal Blooms desirable, although this is an especially challenging task in fjordic environments where ocean-colour is influenced substantially by non-algal sources of variance. These Northern shelves are undergoing rapid climatological and hydrographic change. The difficulty that remote-sensing approaches face

distinguishing Phytoplankton Functional Types in productive shelf-settings, and the lack of direct bio-optical algorithms to identify DMS producers, are of particular relevance as current challenges for biological oceanographers. Some researchers have attempted to infer DMS production from profiling ARGO float hydrography – assuming a deeper Mixed Layer Depth favours *Phaeocystis* and therefore higher DMS (Galí et al. 2019).

We explore the possibility that, as suggested previously in the literature, the choice of observational methods used to describe phytoplankton community structure may have importance for the capacity to perceive distinct Phytoplankton Functional Types and recognise their presence in *in-situ* samples for the purpose of validating remote-sensing algorithms.

2.3 Abstract

Phytoplankton community structure and its relationship to the environment are key to understanding primary production in the oceans, provision of food to higher trophic levels and contributions to global biogeochemical cycles. For this reason, the remote detection of Phytoplankton Functional Types has received increasing attention as a means to monitor ocean ecosystems. However, the consistent definition of Functional Types and their identification from *in-situ* samples for validation of remote-sensing algorithms remains challenging, especially given the use of a variety of different methods to describe phytoplankton community structure. Here we compare efforts to describe community structure from cell-counts, taxon-indicative pigment ratios and bio-optical properties ($a_{ph}(\lambda)$ / spectrally-dependent absorption) across multiple seasons in the Barents Sea – an ecosystem experiencing rapid climatological change – to explore how different representations of community structure influence the perception of Phytoplankton Functional Types. We find that retrieved relationships of phytoplankton community structure to environmental conditions differ slightly between these data types, suggesting that the choice of representation of phytoplankton

community structure will influence the definition and detection of Phytoplankton Functional Types. We find cell-counts reveal the greatest number of significant environmental relationships. However, the relatively weak capacity of these data to predict concomitant bio-optical structure suggests a ‘translation’ error will be introduced if cell-counts are used to validate remote-sensing algorithms, which must necessarily exploit bio-optical variation. We suggest this error may be avoided by using bio-optical properties to construct schemes of Phytoplankton Functional Types and validate remote-sensing algorithms. Refined definitions of Phytoplankton Functional Types offer the opportunity to better monitor regime shifts in a changing Barents Sea ecosystem.

2.4 Plain Language Summary

Hundreds of different phytoplankton species are known to occur in the Barents Sea. Different phytoplankton species may have different environmental preferences, or be delivered into the Barents Sea by different inflowing currents. Therefore, as a consequence of variability in environmental conditions and the exchange of different water-masses, the exact mixture of species that occurs from one bloom to the next is variable.

Different phytoplankton species may have different nutrient requirements, physiologies, sink at different rates, produce different climatologically-relevant gasses, or be more or less likely to be eaten by other organisms. Therefore variation in phytoplankton assemblages is an important component of overall ecosystem function and we can state that there is ‘functional variety’ among the phytoplankton.

Researchers often simplify this variety to schemes of ‘Phytoplankton Functional Types’. Quantifying this variety in phytoplankton assemblages, in order to identify PFTs, is a challenge for researchers, because counting phytoplankton cells individually is laborious and can rarely be achieved exhaustively. Researchers may therefore exploit species-specific variation in phytoplankton pigments, or absorptive spectral properties (determined by the combina-

tion of intracellular pigments, pigment-protein complexes and morphological properties of the cell or colonies of cells), to gain insights into the compositional variety of phytoplankton communities.

This chapter explores the influence of the different methods researchers use to describe the composition of phytoplankton communities on the perception of functional relationships. This is achieved by studying datasets of Barents Sea phytoplankton absorption, cell-counts and pigments; and discerning functional relationships by relating them to a dataset of explanatory variables, composed of hydrography, nutrient concentrations, light and oxygen concentration.

While the broadest functional preferences for different water-masses remain similar across the different phytoplankton datasets, it is found that cell-counts offer the capacity to identify the highest number of significant environmental relationships. This is probably because some individual species of phytoplankton have distinctive environmental niches, but do not possess distinctive pigment inventories or absorptive properties. It must be cautioned that variation in phytoplankton community composition inferred from cell-counts is not always strongly related to variation in phytoplankton absorption.

Therefore, should researchers seek to identify different phytoplankton assemblages from remotely-sensed ocean-colour (a growing ambition in the literature), they must carefully consider which forms of functional variety have perceptible influence on the bio-optical properties of phytoplankton. A scheme of PFTs defined from variation in phytoplankton absorption may be more clearly related to variety in ocean-colour than a scheme of PFTs defined from cell-counts, for instance.

2.5 Introduction

The task of relating variation in phytoplankton community composition to environmental conditions and biogeochemical processes is complicated by the highly multivariate nature

of community composition. While environmental variation can be measured in highly standardised frameworks (González Taboada and Anadón 2014; Fragoso et al. 2017), there is no single method to quantify phytoplankton community composition. Common approaches include optical microscopy, either counting taxa comprehensively (Choudhury and Pal 2011) or enumerating select key species or size-classes (Widdicombe et al. 2010; Degerlund and Eilertsen 2010), or measurement of ratios of taxon-indicative pigments (Roy et al. 2011; Fragoso et al. 2017). Other, more esoteric approaches, have included genetic sequencing (Onda et al. 2017), and scanning electron microscopy. Scanning electron microscopy is particularly useful for studying nano- and picophytoplankton (Gran-Stadniczeňko et al. 2017), but the identification of species depends on researcher expertise and volumes of processed cells are low. Genetic sequencing offers the advantage of much higher throughputs, but capacity to identify species depends on the completeness of genetic reference libraries – indeed Gran-Stadniczeňko et al. (2017) found a majority of species they encountered, in a comparative analysis of genetic sequencing and microscopy, had not been recorded in existing libraries. Furthermore, the biomass of genetic molecules in a cell varies as a function of phylogeny and environment, and the same genetic mix of phytoplankton may have different functional significances depending on whether they are organised as free-cells or as colonies (e.g. *Phaeocystis*). Recent developments in the imaging of cells processed with flow-cytometry (a continuous stream of samples cells passed through a processor; a technique developed from medical applications) has also made it possible to rapidly assess and enumerate the morphological and optical properties (e.g. fluorescence) of phytoplankton cells in real time (Poulton 2016). Variables such as size structure can be retrieved accurately, but taxonomic classification depends on artificial intelligence such as neural network classifiers, and is thus contingent on access to a comprehensive library of classified sample images.

All of these data are multivariate, and it is necessary to collapse them onto a reduced

number of axes to facilitate analysis. Variance in phytoplankton community composition is routinely summarised into distinct clusters dominated by well-defined environmental associations or nutrient requirements; these functionally-differentiated clusters are known as ‘Phytoplankton Functional Types’ (PFTs) (McGill et al. 2006).

The concept of Functional Types is useful because it allows complex multivariate change in phytoplankton community composition to be related to nutrient cycling, sinking rates and contributions to export production (Fuhrman and Capone 2001; Quere et al. 2005; Nair et al. 2008). Depending on research goals, different schemes of PFTs may be contrived to emphasise trophic variety, nutrient demands, size-structure (Sieburth et al. 1978) or a combination of these factors (Quere et al. 2005). Such simplified representations of phytoplankton functional diversity are routinely employed in ecosystem models (e.g. Wassmann et al. 2010; Zhang et al. 2015; Yool et al. 2011).

Detection of PFTs in remotely-sensed ocean-colour imagery has gained increasing attention as a goal in the literature (Platt et al. 2006; Bracher et al. 2017), to monitor ecosystems in real time and provide synoptic-scale validation of ecosystem models. Nowhere is this goal more urgent than the productive Arctic shelf seas, where climate and hydrography are evolving rapidly (Overland et al. 2018), and commensurate changes in community structure occurring (Li et al. 2009; Lalande et al. 2013; Nöthig et al. 2015; Oziel et al. 2017), that have the capacity to significantly alter sinking rates and export production (Armstrong et al. 2009; Wassmann et al. 1990; Wollenburg et al. 2018).

However, efforts to detect PFTs from ocean-colour have had mixed success, especially in productive shelf regions, where no existing algorithm has been assessed to reliably retrieve phytoplankton size-structure (Liu et al. 2018). Difficulty developing reliable algorithms may stem from challenges validating the occurrence of different Functional Types from *in-situ* samples and relating this information to variance in ocean-colour, as it is possible different means of describing phytoplankton community composition are more or less disposed to re-

trieve different forms of functional diversity (Sosik et al. 2014). Common approaches, such as defining schemes of PFTs on cell-counts or pigment ratios, before seeking bio-optical signatures to recognise in remote-sensing (e.g. Astoreca et al. (2008), Jackson et al. (2010), and Oziel et al. (2017)), may falsely assume that variance in phytoplankton bio-optical properties is mirrored by variance in community structure perceived in cell-counting and pigments. However, chromatic acclimation may distort these relationships and bio-optical variety may exist in phytoplankton communities that is not apparent from microscopy or pigment ratios (Goericke and Montoya 1998; Astoreca et al. 2008). In particular, colonial phytoplankton aggregates, that may possess unique bio-optical properties and functional roles (Peperzak et al. 2000b; Astoreca et al. 2008; Fragoso et al. 2017), are often not represented in cell-counts at all, due to difficulties finding standardised approaches to counting them (Degerlund and Eilertsen 2010).

There is hence a mandate to investigate the influence of different approaches to describe phytoplankton community composition on the perception of functional diversity in productive Arctic shelf seas and to explore the capacity of different data types to predict concomitant bio-optical variation – a key assumption underlying their use for validation in remote-sensing applications.

In this study, we explore the consistency of relationships between phytoplankton community structure and environmental conditions across multiple different methods for describing phytoplankton community composition. Our samples originate from the Barents Sea, an ecosystem frequently dominated by *Phaeocystis* and a diversity of diatoms of a range of cell sizes. This will allow us to assess the influence of the choice of method for quantifying phytoplankton community structure has on the determination of schemes of PFTs. We also explore the capacity of cell-counts, pigment ratios and bio-optics to identify bio-optical groups from *in-situ* samples, to probe the effects of method choice on the accuracy of validation for bio-optical remote-sensing algorithms. We also explore a recent revival in

bio-optical approaches; Gaussian decomposition methods (Liu et al. 2019). Gaussian decomposition breaks down phytoplankton absorption spectra into a series of Gaussian curves, whose heights proxy the contribution of taxon-indicative suites of phytoplankton pigments (Hoepffner and Sathyendranath 1993; Lohrenz et al. 2003; Chase et al. 2013). A full review of Gaussian decomposition, multivariate analysis of results and other bio-optical methods (e.g. fluorescence spectra) can be found in chapter 13 of Roy et al. (2011). We explore whether Gaussian decompositions offer any advantage over raw hyperspectral properties in the identification of functional relationships.

2.6 Materials

2.6.1 Data collection

Phytoplankton samples were collected aboard four research cruises conducted as part of the UK-led Changing Arctic Ocean research programme, throughout the Western Barents Sea, with a special focus on sampling the Seasonal Ice Zone along the 30th Eastern Meridian (Cottier 2018, Cottier and Berge 2018a, Cottier and Berge 2018b, Hopkins 2018).

The locations of hydrographic and biological sampling are presented in Figure 2.1.

The four cruises were distributed across the winter (HH180101; 4–18th January 2018), spring (HH180423; 23rd April– 5th May, 2018), early (JR17006; 11th June– 6th July 2018) and late (JR16006; 30th June– 8th August 2017) summer in order to capture the seasonal evolution of hydrographic conditions in the Barents Sea and commensurate succession of phytoplankton communities. A component of inter-annual variability in environmental conditions is also convolved, as a consequence of the distribution of the cruises across two years, and this is especially evident in the variable ice conditions between the two summer cruises. The Northern limit of sampling was restricted on winter and spring cruises by heavy ice

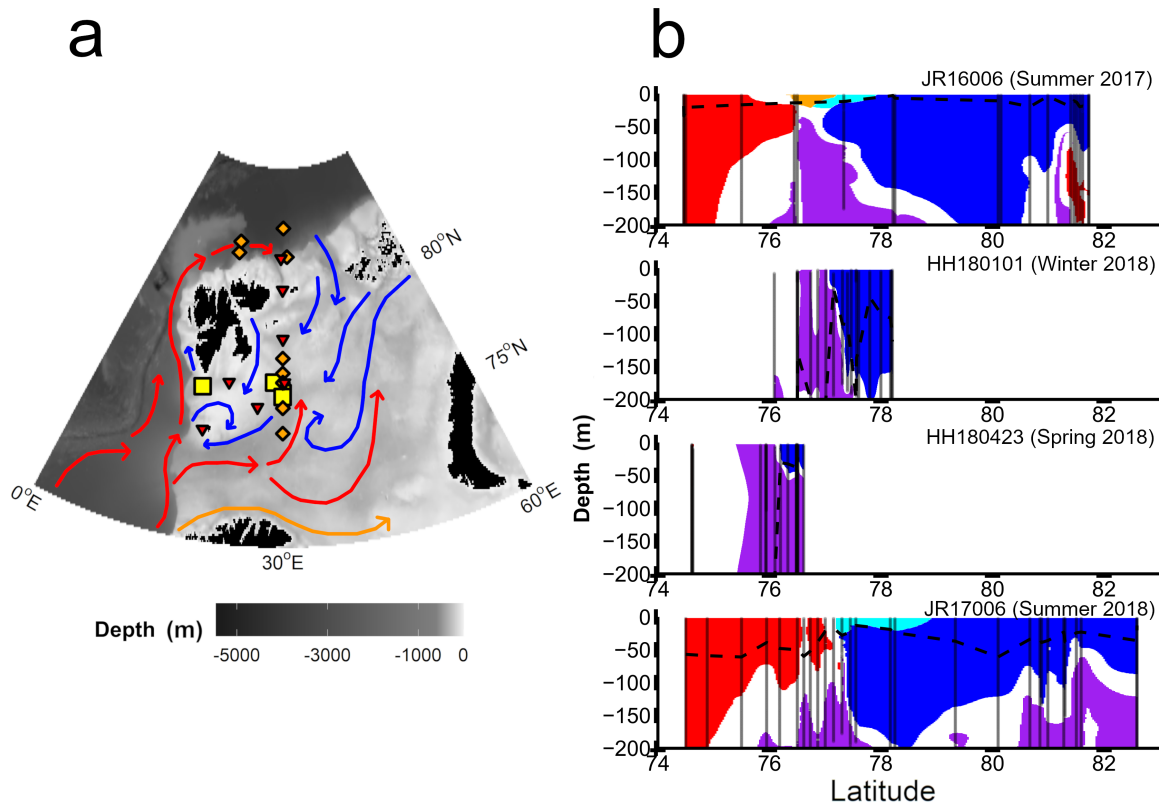


Figure 2.1: **a)** Map of phytoplankton samples considered in this study; HH180423 (large yellow squares), JR17006 (medium orange diamonds), JR16006 (small red triangles). **b)** From top to bottom: JR16006 (summer 2017), HH180101 (winter 2018), HH180423 (spring 2018), JR17006 (summer 2018). Water-masses: Arctic (blue), Atlantic (red), Barents (purple), Coastal (orange) and Melt (cyan). The Mixed Layer Depth is represented by a dashed black line. Vertical lines represent CTD profiles.

conditions, and the Southern limit of ice extent evolved with the Arctic water-mass (Barton et al. 2020).

The seasonality of hydrographic conditions is illustrated in Figure 2.1 (also see 7.2,7.3), and water-masses are defined in accordance with Oziel et al. (2016) and Våge et al. (2016). Mixed Layer Depth, which is shown to reflect the density structure of the upper water column, is defined as the depth at which potential density anomaly (σ_θ) increases to 0.1 kg m^{-3} above the surface reference (Peralta-Ferriz and Woodgate 2015).

A Conductivity-Temperature-Depth (CTD) rosette equipped with a Chelsea Seabird Aqua fluorometer was used to produce downward profiles of *in-vivo* [*chl-a*] fluorescence (7.7) and hydrography, which informed the selection of sampling depths for the collection of waters for biological sampling on the upward cast. Typically the surface, chlorophyll-maximum, and 2-3 depths bracketing these features, were targeted for sampling. Niskin bottles were then sub-sampled for measurements of phytoplankton bio-optical properties, HPLC pigment and nutrient concentrations, and algal cell abundances.

2.6.2 Environmental data

Hydrography

The sensory array on the CTD package recorded profiles of temperature (7.2), practical salinity (7.3), depth, oxygen saturation (7.9) and downwelling Photosynthetically Available Radiation (E_{PAR} ; W m^{-2}) (7.6) at a 1 m vertical resolution on all deployments. oxygen saturation was determined using a Seabird polarographic oxygen probe, calibrated with Winkler titrations (Carpenter 1965). Temperature was measured with SBE3plus sensors, salinity inferred from conductivity readings derived from SBE 4C sensors. Further detail on the preparation of hydrographic data is available in the cruise reports: (Cottier 2018, Cottier and Berge 2018a, Cottier and Berge 2018b, Hopkins 2018). The depth of the maximum Brunt-Väisälä frequency at each station was computed from temperature, pressure and salinity in the context of sampling latitude, with the *gsw_Nsquared()* function of the *oce* R package version 1.2-0 (Kelley and Richards 2020).

Nutrients

Triplicate 50 mL subsamples were taken from Niskin bottles for assays of nitrate plus nitrite (hereafter referred to as ‘nitrate’ (7.4) as contributions of nitrite are minor), silicic acid (7.5) and ammonium (7.10), with acid-washed polythene vials. Samples were filtered through

clean 200 μm nylon mesh to remove zooplankton. Samples were brought to room temperature before measurement with a Lachat *Quikchem 8500* flow injection auto-analyser. The analyser was standardised using international certified reference materials for nutrients in seawater (KANSO Limited, Japan), and samples were analysed against standard solutions made up in a low-nutrient seawater (OSIL batch 26) standard matrix. Five calibration standards and blank seawater were analysed before each batch process and a linear correction was applied to correct for drift over each analytical run. The end results are represented as micromolar volumetric concentrations ($\mu\text{Mol L}^{-1}$). Different phytoplankton groups may exploit environments dominated by different nutrient concentrations, or exhibit differential use of key nutrients (e.g. silicification in diatoms). Therefore, concomitant gradients in nutrient concentration across a dataset of phytoplankton samples can be used as an explanatory variable to constrain variance in the community composition and identify functional associations. See the mathematical methods appendix ‘Redundancy Analysis’ for further mathematical details.

2.6.3 Phytoplankton community structure

Bio-Optical properties

Bio-optical properties are herein represented by phytoplankton absorption ($a_{ph}(\lambda)$). Sub-samples of up to 1 L were collected from Niskins and filtered through 0.7 μm pore GF/F filters under pressure not exceeding one half atmosphere. Samples were frozen in liquid nitrogen and transferred to -80°C storage. In the laboratory, samples were thawed to -20°C in a dark environment and moistened with a drop of filtered deep Atlantic Water. The absorption spectra of the moistened filter pads were recorded with a Shimadzu UV-Vis spectrophotometer equipped with an integrating sphere at a 1 nm resolution, following the method of Kishino et al. (1985), including the extraction of pigments with methanol to

determine absorption by detritus. Filter pads were rotated by a quarter revolution 4 times, collecting replicates that were subsequently averaged, to control for uneven distribution of cells on the filter pads. We used the binomial path length correction equation of Hoepffner and Sathyendranath (1992) to convert the optical densities into absorption coefficients. We subtracted exponential fits to the detrital absorption spectra from the total particulate absorption spectra to attain the absorption spectra of the phytoplankton. No valid absorption spectra could be generated for winter samples, as concentrations of phytoplankton were too low. Therefore no winter-derived samples are included in further analysis.

Pigment analysis

Subsamples, typically between 1 and 2 L, were collected from Niskins and filtered through 0.7 μm pore GF/F filter pads. Samples were frozen in liquid nitrogen and transferred to -80°C storage. Samples were shipped to the Dansk Hydrualisk Institut (DHI; Hørsholm, Danmark) for HPLC analysis. The samples were prepared in a vortex, sonicated and extracted in 90% acetone solution at 4°C before further mixing. The samples were filtered through a 0.2 μm Teflon syringe filter (to remove debris from the digested filter pads) and the filtrate transferred to vials for analysis. Buffer solution and filtrate were mixed in a 5:2 ratio, and analysed following the HPLC method of Van Heukelem & Thomas (2005), detailed in Hooker (2005).

HPLC (High Performance Liquid Chromatography) is an analytical method that permits the measurement of taxon-indicative phytoplankton pigments, from which phytoplankton community composition can be inferred. An example repertoire of indicative pigment species and the algal classes that possess them is presented in Figure 2.2.

It can be seen from Figure 2.2 that few pigment species can be treated as indicative of a single taxon (e.g. Peridinin). Therefore, converting inventories of HPLC-derived pigments into estimates of the contribution of different algal classes to a phytoplankton community

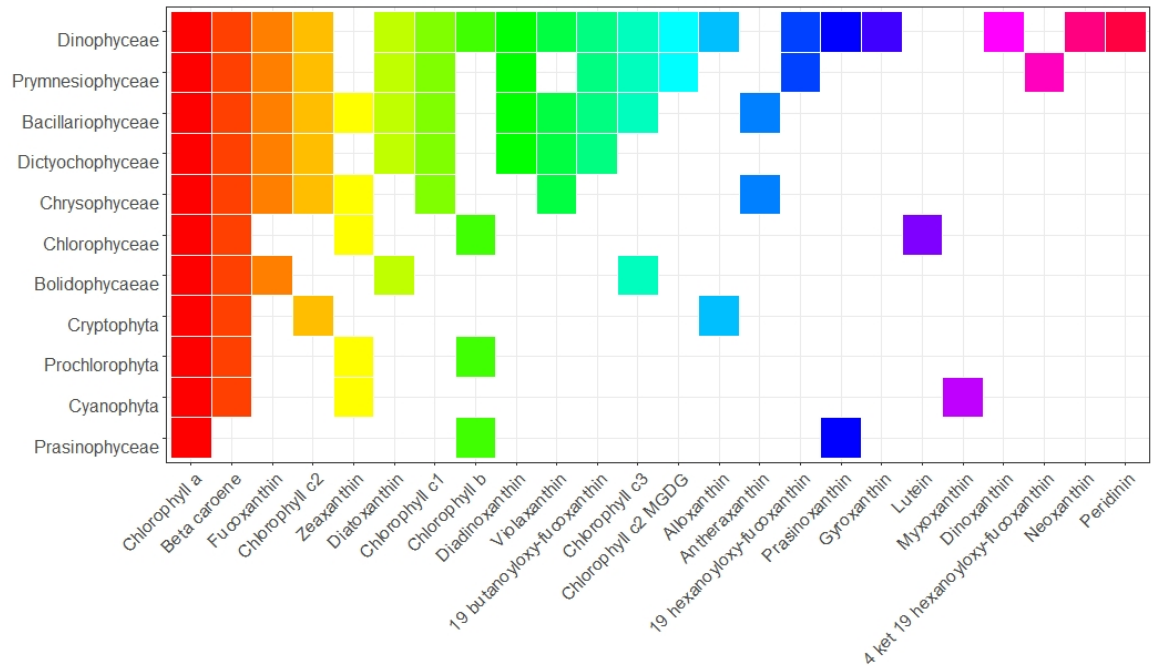


Figure 2.2: Taxon-indicative algal pigments that can be retrieved through HPLC, and the algal classes that possess them. Data aggregated from Roy et al. (2011) and processed in R. Algal classes are sorted by number of pigments species possessed. Pigment species are sorted by frequency across algal classes. Some pigment species names are simplified.

necessitates solving a linear un-mixing problem. This is not trivial, because this is a ‘many to many’ type statistical problem and there may be multiple equally valid solutions. To provide an analogy, we can compare this to a painter mixing different brands of paint, that themselves contain different balances of pigment. A final mix might contain 1 part cadmium yellow and 2 parts ultramarine. If the painter has, in their possession, 3 brands of paint (‘sunflower’; 1 part cadmium yellow, ‘grass green’; 1 part cadmium yellow, 1 part ultramarine, ‘sea green’; 1 part cadmium yellow, 3 parts ultramarine), then we can see that we could arrive at the final mix in multiple ways. Perhaps we could mix 2 parts sea green with 1 part sunflower, but we could equally well mix equal parts grass and sea green.

As interrogation of HPLC data may typically involve >20 pigment species, we can imagine that the solutions to this many-to-many problem form a landscape, where the topography represents a measure of ‘goodness of fit’ that we need to optimise as we propose different combinations of algal classes to satisfy the constraints from the final pigment mix. Different mixes will entail different final errors, and perfect solutions will contain zero error. The landscape may not slope down uniformly to the perfect solution, and there may be multiple local minima in this landscape, and multiple (or zero) perfect solutions. The ‘goodness of fit’ measure we propose is, furthermore, a subjective choice. Perhaps we simply choose residual error between our proposed solution and the final mix, or perhaps we incorporate some measure of the complexity of our guess, and choose to penalise more complex guesses, for example.

A commonly-employed algorithmic approach to arrive at solutions is the CHEMTAX (Chemo-taxonomy) method (Mackey et al. 1996). Providing a loose outline, CHEMTAX views this multivariate landscape and follows a steepest-descent from an initial guessed solution like a skier down alpine slopes. This analysis is re-run multiple times with random variations to the initial statement of the weighting coefficients that relate algal classes to pigments. This means that, should those initial statements be slightly wrong, or should there be some small amount of variance across the original dataset, that the skiers’ various routes will result in numerous minima. The lowest of that sum of minima can thereafter be inferred to be the optimal solution. R packages to solve linear inverse problems exist that loosely reflect the process of CHEMTAX (Meersche et al. 2009). These methods require the user to state the starting assumptions of linear coefficients that relate the stated algal classes to the studied pigments. Different descent algorithms are typically employed (e.g. random walk). As CHEMTAX uses linear coefficients, constant across the dataset, to relate algal classes to pigments, processes such as photo-acclimation (e.g. Johnsen and Sakshaug (1993)), or internal phylogenetic variation in pigment ratios within an algal class, will be poorly

represented by this method and result in erroneous estimates of class contributions. As a consequence, the initial matrix of weighting coefficients that relate algal classes and pigments must be informed carefully by ancillary methods for describing community composition, such as optical microscopy, and it is necessary that CHEMTAX is run on subsets of the larger dataset (e.g. Fragoso et al. (2017), Table 3) to ensure that the linear coefficients are consistently supported across the dataset.

As the final proposed groupings of CHEMTAX are weighted linear combinations of the original pigment inventory, we can reason that our recipe of algal classes estimated for a group of samples is simply a rotation of the original pigment inventory data through multivariate space (see my mathematical methods chapter appendix ‘PCA’ for another example of ordination by rotation). This rotation differs from PCA, however, in the respect that if the number of retrieved algal classes is small and fitting errors are non-negligible, then some portion of the original data structure has been sacrificed in the conversion. Therefore rotation by coefficients, derived from CHEMTAX, is unlikely to improve the capacity of other rotation-based exploratory methods to detect relationships between pigment inventory and functional associations. For example, the ratio of diatoms to dinoflagellates inferred from CHEMTAX may well proxy gradients in silicic acid concentration, but so too would a linear combination of fucoxanthin, peridinin and dinoxanthin. Hence, in this analysis, I do not elect to process the inventory of sample pigments with methods such as CHEMTAX.

The HPLC analysis quantified concentrations of pigments presented in Table 2.1 below.

Pigment species	Phytoplankton groups associated therewith	Function
Chlorophyll-a	Ubiquitous	Photosynthetic
Chlorophyllide-a	Degradation product	
Chlorophyll-b	Prasinophytes, chlorophytes, diatoms	Photosynthetic

Chlorophyll-c3	Prymnesiophytes, chrysophytes and some dinoflagellates	Photosynthetic
Chlorophyll-c2 & c1	Diatoms, prymnesiophytes, dinoflagellates, cryptophytes and chrysophytes	Photosynthetic
Chlorophyll-c3 MDGD	Prymnesiophytes and dinoflagellates	derivative of c2
Phaeopigment	Degradation product	
Phaeophytin-a	Degradation product	
Peridinin	Dinoflagellates	Photosynthetic carotenoid
19'-butanoyloxyfucoxanthin	Prymnesiophytes, chrysophytes and some dinoflagellates	Photosynthetic carotenoid
Fucoxanthin	Diatoms, prymnesiophytes, chrysophytes and some dinoflagellates	Photosynthetic carotenoid
Neoxanthin	Dinoflagellates	Photoprotective carotenoid
Prasincoxanthin	Prasinophytes	Photoprotective carotenoid
Keto-19'-hexanoyloxyfucoxanthin	Metabolic product of 19'-hexanoyloxyfucoxanthin	
Violaxanthin	Chlorophytes and prasinophytes	Photoprotective carotenoid
19'-hexanoyloxyfucoxanthin	Prymnesiophytes and some dinoflagellates	Photosynthetic carotenoid

Diadinoxanthin	Diatoms, prymnesiophytes, dinoflagellates and chrysophytes	Photoprotective carotenoid
Dinoxanthin	Dinoflagellates	Photoprotective carotenoid
Myxoxanthin	Cyanobacteria (we do not anticipate encountering these)	Photoprotective carotenoid
Antheraxanthin	Diatoms and chrysophytes	Photoprotective carotenoid
Alloxanthin	Mostly dinoflagellates	Photoprotective carotenoid
Diatoxanthin	Diatoms, prymnesiophytes, dinoflagellates and chrysophytes	Photoprotective carotenoid
Zeaxanthin and Lutein	Chlorophytes and some prasinophytes	Photoprotective carotenoid
Gyroxanthin	Dinoflagellates and prymnesiophytes	Possibly photosynthetic carotenoid
α -carotene		Photoprotective carotenoid
β -carotene	Diatoms, prasinophytes and chlorophytes	Photoprotective carotenoid

Table 2.1 is based on information aggregated from Pettersen et al. (2011) and Roy et al. (2011) and Fragoso et al. (2017).

Phaeopigment concentrations were below the detection limit in all samples, indicating

a low degree of degradation.

Dinoxanthin, myxoxanthin, lutein, gyroxanthin and α -carotene never constituted >1% pigment biomass in any sample.

Cell-counts

50 mL subsamples, preserved with 1% Lugol's solution, were prepared for optical microscopy following the Utermöhl technique (Utermöhl 1958). Chamber counts were made with a Zeiss Axio S100 inverted microscope with x20 magnification and phase-2 contrast conditions, following established practice (Davidson et al. 2013). Counts of recognisable genera and species were produced in consultation with Tomas (1997), Throndsen et al. (2007), and Hoppenrath et al. (2009). Colonial forms, such as *Phaeocystis* and *Dinobryon* were excluded, reflecting routine practice in the literature (Degerlund and Eilertsen 2010). We supplemented the microphytoplankton counts with flow-cytometric measurements of nanophytoplankton, following the example of Ardyna et al. (2011), and the method of Zubkov and Burkill (2006) to stain and process samples. We divided samples into small and large PNAN (Phototrophic nanoflagellates) and HNAN (Heterotrophic nanoflagellates), distinguishing them by differential red fluorescence (Zubkov et al. 2007). As cell-counts did not always originate from the same Niskin bottles as pigment and bio-optical samples, we matched them to one another under the criterion they must originate from the same location, within 2 metres of depth.

2.6.4 Data Analysis

Bio-optical community structure in the phytoplankton was summarised by dividing samples into four clusters with distinct bio-optical properties, with a fuzzy clustering algorithm (Cebeci et al. 2019) following Orkney et al. (2020) (see the fuzzy clustering section of the mathematical methods appendix). These clusters were validated by comparison with microscopic observations of plankton collected from net-tows. Nanophytoplankton (insignificant

number of large cells observed), diatoms, *Phaeocystis* colonies and dinoflagellates were recognised. An ordination space, based on a projection of the data along the leading principal components identified in samples from cruise JR16006, is presented (Figure 2.3).

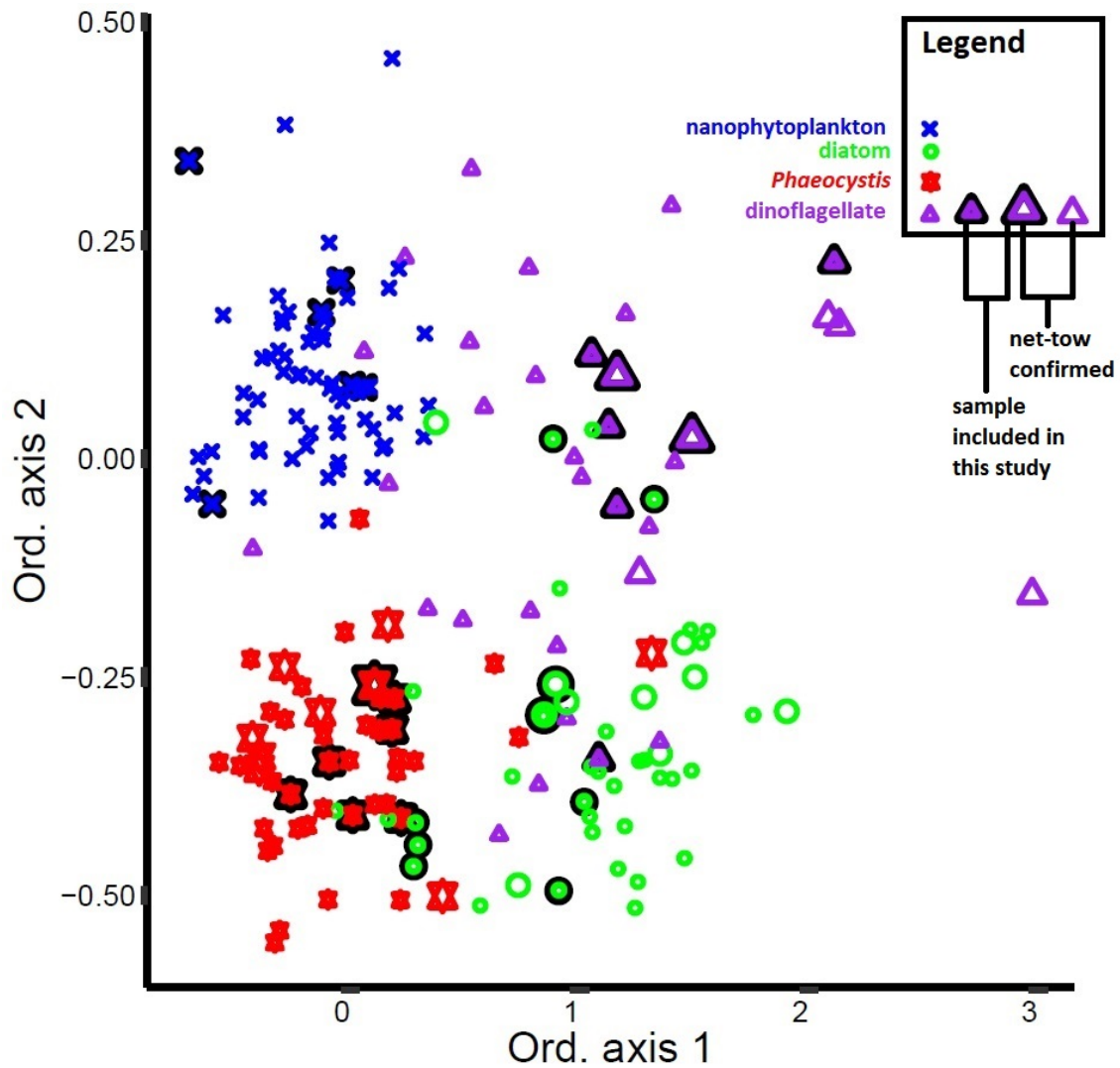


Figure 2.3: Ordination space representing salient variance structure in Barents phytoplankton bio-optics (191 points). A dark outline indicates sample inclusion in this study (n=29), large point size indicates samples whose taxonomic composition was confirmed by ship-board microscopic analysis of net-tows. Crosses: nanophytoplankton, Circles: diatoms, Stars: *Phaeocystis* colonies, Triangles: dinoflagellates.

In total, our analysis includes 6 samples with bio-optical properties consistent with dominance by nanophytoplankton, 9 consistent with diatoms, 7 consistent with dinoflagellates and 7 consistent with colonial *Phaeocystis*.

Further detail on the ordination axes presented in Figure 2.3 can be found in Table 3.1, where the linear coefficients of different spectral regions are presented in full. To summarise, the first ordination axis may represent a partial proxy of pigment packaging, while the second ordination axis likely represents the relative abundance of photoprotective spectral adaptations.

A detailed view of the ordination space (based on PCA; see mathematical methods section), and the average spectra of the different bio-optical clusters is presented in detail in Figure 3.5. Diatom-labelled samples included in this study are characterised by high pigment packaging and lots of photosynthetic carotenoids. *Phaeocystis*-labelled samples are similar to diatoms, but have a stronger contribution of chl-c3, and a concomitant higher absorption at 469 nm. Nanophytoplankton-labelled samples have low pigment packaging, a lower ratio of *chl-a* to other pigments and a greater proportion of photoprotective carotenoids. Dinoflagellate-labelled samples do not appear to have a distinct pigmentation (lacking dinoflagellate-specific pigments), and have a greater dispersion in the ordination space. This cluster is predominantly recognised by confirmation in net-tow observations, rather than its possession of a single distinct spectral signature or pigment ratio. However, dinoflagellate-labelled samples do appear to occupy a region of the ordination space that is not highly populated by other clusters. See the results section for a detailed discussion of the pigment complements associated with each cluster.

Redundancy Analysis (RDA), implemented in R version 3.6.3 (R Core Team 2020) using the *vegan* package version 2.5-6 (Oksanen et al. 2019), was used to explore functional relationships in each community compositional dataset. This resembles methods employed in previous studies (e.g. Janišová et al. (2007) and Ardyna et al. (2011)) and Fragoso et

al. (2017) – including the calculation of marginal, termwise and pure effects to quantify the significance of different explanatory variables’ constraint of the dependent variables. See the mathematical methods appendix ‘Redundancy Analysis’ for a worked example and publicly available code.

Bio-optical absorption spectra were decomposed into Gaussian functions following Chase et al. (2013) and their amplitudes normalised to their sum. Raw absorption spectra were similarly prepared by normalisation to the integral between 400 and 750 nm. HPLC pigment ratios were normalised to the sum of pigments in each sample. Cell-counts were log-transformed, as in (Peperzak 2010), and normalised to the sum of all log-transformed counts in each sample.

Biplots, representing the leading axes of the Redundancy Analyses, are presented in Figure 2.4.

It can be seen from inspection that the structures of some of these biplots are more similar than others (compare Gaussian and raw optics, which we should expect to be similar), but an objective and quantitative measure extending beyond the first two axes is necessary for rigorous comparison.

The results of any individual Redundancy Analysis can be represented by a matrix n samples deep and p columns wide. In this matrix, the p^{th} value on the n^{th} row represents the angle between the vector, from the origin to the n^{th} sample, and the vector describing the p^{th} environmental variable. These two vectors exist in a p -dimensional space, defined by the RDA axes (Legendre and Legendre 2012).

Thus, as this is a Euclidean space, both sample scores and environmental vectors were scaled by the square-root of the p RDA eigenvalues, and hierarchical clustering was used to summarise the structure of each Redundancy Analysis as a dendrogram. The dendrogram describes the similarity of the environmental associations among the n phytoplankton

samples. The similarity of pair-wise combinations of dendrograms was then derived by computing the average cophenetic distance of samples between dendrograms, using the *cor.dendlist()* function of R package *dendextend* (Galili 2015). The data were resampled 1000 times with replacement, to arrive at a robust appraisal of the correlation structure of the phytoplankton functional relationships uncovered by each of these analyses (Figure 2.5). A worked example of this process is available in the Redundancy Analysis section of the mathematical methods appendix.

The *anova.cca()* function of the *vegan* R package (Oksanen et al. 2019) was employed to compute termwise, marginal and pure effects (presented in Figure 2.5) to assess relative significance of different environmental variables' contribution to explained variance in the phytoplankton datasets.

Lastly, the capacity of different data types to identify distinct bio-optical phytoplankton groups was assessed. Linear Discriminant Analyses were trained, provided with 3 randomly chosen samples of 4 bio-optical clusters – diatoms, nanophytoplankton, dinoflagellates and *Phaeocystis* (presented in Figure 2.3). They were then used to predict the bio-optical identities of the remaining samples. Posterior probabilities were used to fit logistic models with the *glm()* function, and ROC (Receiver Operating Characteristic) curves computed, using the *pROC* R package (Robin et al. 2011). This process was repeated 100 times for each of the data types, for each of the 4 phytoplankton bio-optical clusters. This allows us to assess the rate of true and false positive classifications and hence the capacity for each data type to be used for validation of remote-sensing algorithms that are based on bio-optical structure. See the Linear Discriminant Analysis section of the mathematical methods appendix for a fuller discussion of this method.

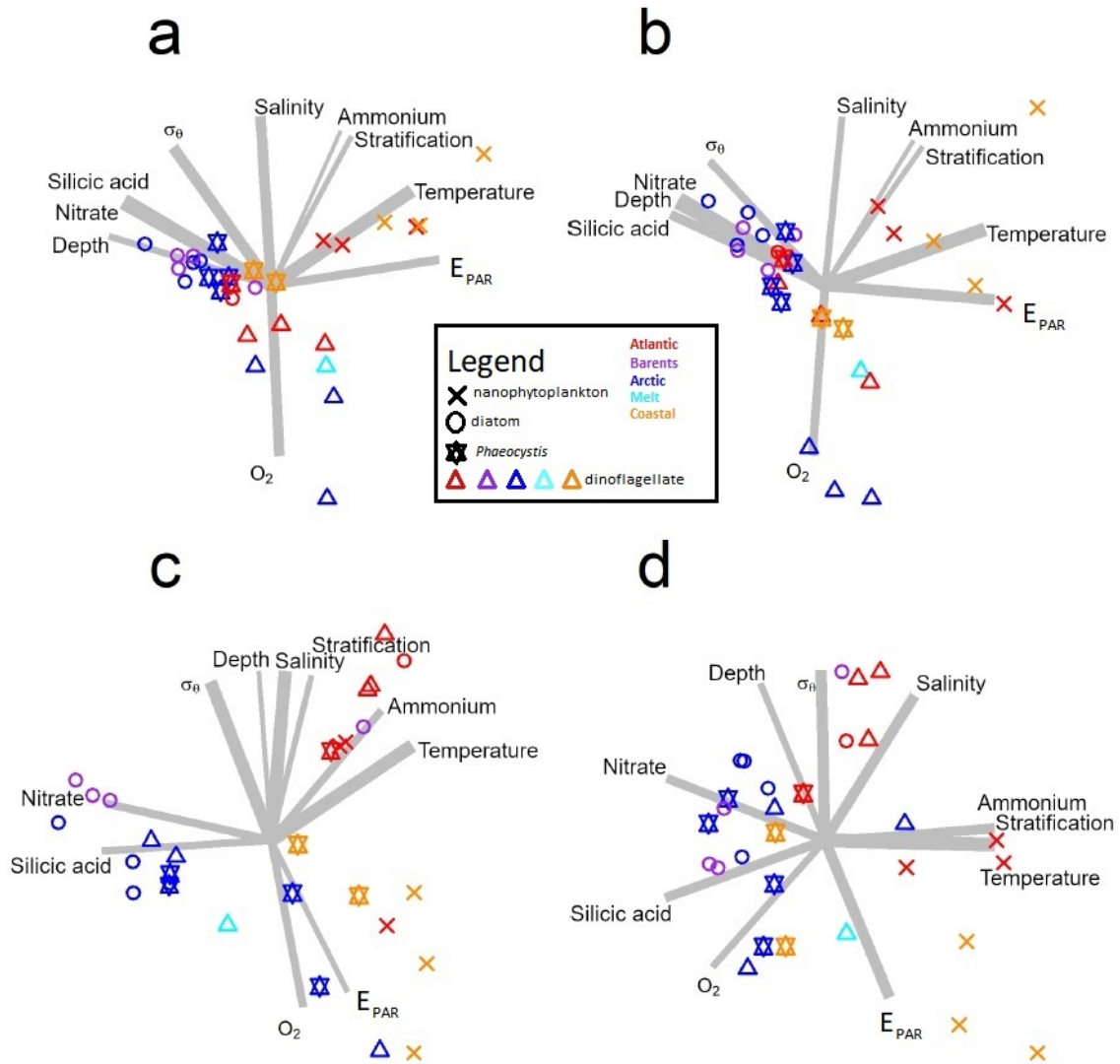


Figure 2.4: Biplots of redundancy analyses relating **a)** Gaussian optical properties, **b)** raw spectra, **c)** cell-counts and **d)** pigment ratios to environmental variables. The magnitude of the axes and environmental vectors for each analysis has been scaled to unity. Points are discriminated by shape and colour, indicating their affinity to different bio-optical phytoplankton groups identified in Orkney et al., (2020): *Phaeocystis* (stars), diatoms (circles), nanophytoplankton (crosses) and dinoflagellates (triangles). Border-colours indicate water-masses; red (Atlantic Water), blue (Arctic Water), purple (Barents Water), cyan (Melt Water), orange (Coastal Water). The thickness of the environmental vectors indicates their relative importance in explaining variance.

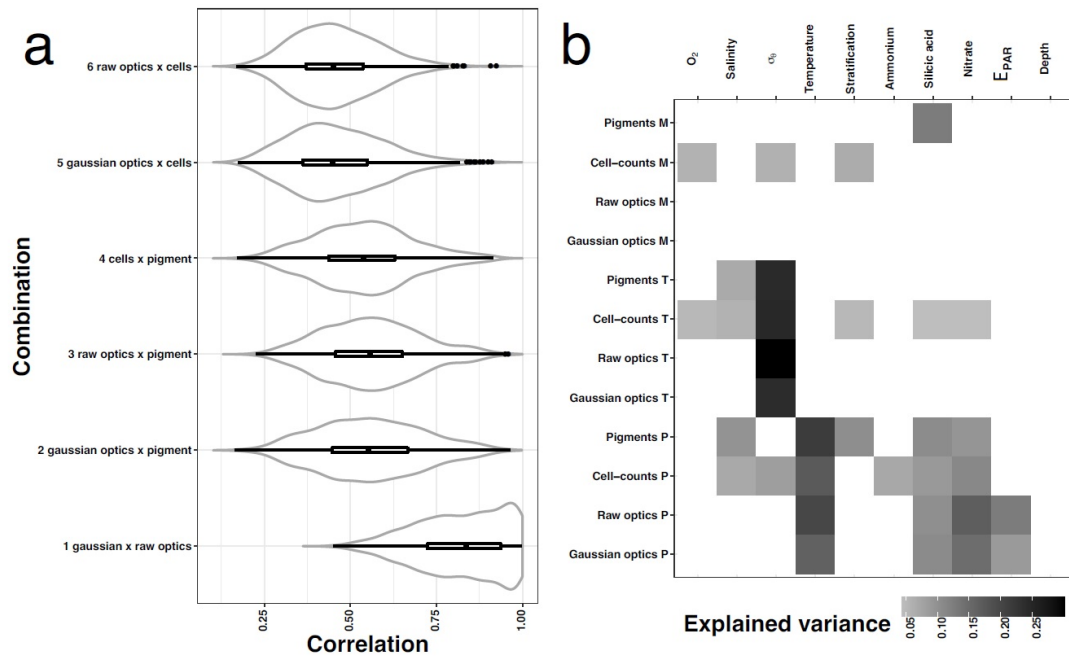


Figure 2.5: a) Violin plot representing the correlation of redundancy analyses based on cell-counts, Gaussian optics and pigments to redundancy analyses based on the other data types. b) Punnett plot representing significant marginal (M), termwise (T) and pure (P) effects associated with with *anova.cca()* function of the *vegan* R package (Oksanen et al. 2019).

2.7 Results

Among samples with diatom-like optical properties, 18% of counted cells were diatoms, compared to 1.6% among other samples. The most abundant microphytoplankton in the diatom cluster were the diatoms *Chaetoceros socialis*, species of *Fragillariopsis*, *Fossula*, *Pseudo-nitzschia delicatissima* and *Thalassiosira nordenskiöldii*. Among samples with nanophytoplankton-like optical properties, 82% of counted cells were phototrophic nanophytoplankton (PNAN), compared with 67% among other samples. The most abundant microphytoplankton species within the nanophytoplankton cluster were the diatom *Pseudo-nitzschia delicatissima*, species of the ciliate genera *Strombidium*, diatoms of the genus *Navicula*, dinoflagellates of *Scrippsiella* and diatoms of *Cylindrotheca*. *Phaeocystis*-affiliated samples were also dominated by PNAN (76% of counts). Among samples of the *Phaeocystis* cluster, the most common microphytoplankton were the diatoms *Pseudo-nitzschia delicatissima*, *Chaetoceros socialis*, species of *Thalassiosira*, *Chaetoceros furcellatus* and coastal species of *Chaetoceros*. Dinoflagellates tended to make up only a minor component of counted cells in the dataset (0.11%). Samples with dinoflagellate-like optical properties were however distinguished by a relatively high contribution of heterotrophic nanophytoplankton (HNAN) of 33%, compared with an average of 20% among other samples. Given that relatively large dinoflagellate cells are flexible, and their DNA can be recovered even when sample water is filtered through a pore size of 3 μm (Lovejoy et al. 2006), we suspect that hetero- or mixotrophic dinoflagellates in our dataset are largely subsumed within the HNAN flow cytometry counts. The most common microphytoplankton counted in the dinoflagellate cluster were species of the diatom *Attheya*, *Chaetoceros furcellatus*, *Chaetoceros socialis*, *Pseudo-nitzschia delicatissima* and *Pseudo-nitzschia seriata*.

Among samples with diatom-like optical properties, 57% of pigment biomass was *chl-a*, 21% was fucoxanthin, 8% was *chl-c2* and 3% was *chl-c3*, according well with expectations. Among samples with nanophytoplankton-like optical properties, 48% of pigment biomass

was *chl-a*, 11% was fucoxanthin, 6% was chl-c2, 5% was chl-b, 4% was 19-hex, 4% was diadinoxanthin, 3% was chl-c3, 2% was peridinin, alloxanthin, 19-but, phaeophytin-a, and the rest were trace. Among samples with *Phaeocystis*-like optical properties, 57% of pigment biomass was *chl-a*, 18% was fucoxanthin, 8% was chl-c2, 6% was chl-c3, 2% was chl-b and diadinoxanthin, and the rest were trace. Richer chl-c3 in this cohort accords with expectations that this pigment is indicative of prymnesiophyte-dominated assemblages. In samples with optical properties suggestive of dinoflagellates, 53% of pigment biomass was *chl-a*, 20% was fucoxanthin, 7% was chl-c2, 4% was chl-c1, 3% was chl-b and diadinoxanthin, 2% was chl-c3 and phaeophytin, and there were traces of β -carotene, alloxanthin and violaxanthin. The paucity of dinoflagellate-indicative pigments such as peridinin (0.4%) and gyroxanthin (not detected) was surprising.

Patterns of functional diversity are similar across the different community compositional datasets (Figure 2.4), but there are some nuanced differences. Leading axes discriminate samples with optical properties suggestive of diatoms and nanophytoplankton. Diatom-labelled samples coincide with higher potential density anomalies (σ_θ) and nutrient inventories, while nanophytoplankton-labelled samples are associated with higher temperatures, E_{PAR} and ammonium. There is a broad distinction of gradients associating community composition and hydrography, between Arctic and Atlantic influenced water-masses; extending between microphytoplankton and nanophytoplankton dominance in Atlantic-influenced water-masses, and between diatom and dinoflagellate-labelled samples in Arctic Waters. Review, for instance, subplot a of Figure 2.4 that shows increasing SST is strongly associated with nanophytoplankton-labelled samples. A facsimile of this pattern exists in subplots c and d, except that the distribution of nanophytoplankton-labelled samples is more coherent with E_{PAR} . It should also be observed that concentrations of silicic acid tend to increase towards regions of the ordination space dominated by diatoms, which is as expected given that

diatoms depend on this nutrient to build their frustules. In analyses based on bio-optical properties and pigments (subplots a, b and d) explanatory vectors describing hydrography establish two gradients in phytoplankton community composition that coincide in diatom-rich Barents Waters (denoted as purple circles), associated with diatom-labelled samples. However, this group of samples are divergent in the cell-count analysis (Figure 2.4c), being more widely dispersed in the plot space. This is an indication that the leading functional relationships between explanatory variables and phytoplankton community composition can change depending on how phytoplankton community composition is quantified. This is reflected in the correlation structure of functional relationships derived from the different phytoplankton datasets (Figure 2.5). Strong correlation between functional relationships, derived from raw absorption spectra and Gaussian decompositions, represents the baseline for agreement between essentially identical data structures. Correlations between all of the other pairwise combinations are significantly lower than this, emphasising that the perception of functional relationships is significantly influenced by the choice of community compositional data type. This conclusion is also supported by the perception of a greater number of significant environmental relationships when cell-counts data are investigated, compared to other data types (Figure 2.5b). Gaussian component functions do not appear to offer any major advantage over raw absorption spectra in achieving a greater coherence to the functional relationships derived from other data types, or the perception of any greater number significant environmental relationships. Overall, functional relationships derived from bio-optical analyses appear to be more concordant with those derived from pigments, and less with those derived from cell-counts. In turn, functional relationships derived from cell-counts analyses are more concordant with those derived from pigments than they are with those derived from bio-optical properties.

ROC curves, representing the capacity for different phytoplankton community datasets to distinguish bio-optical groups, are presented in Figure 2.6.

The straight diagonal lines represent an equal rate of true and false positive classifications – equivalent to a prediction without value. Greater Area Under the Curve (‘AUC’) indicates a greater predictive power.

The highest AUCs are associated with ROC curves using raw optical data as their input (values range from 0.89-0.98), followed by Gaussian optics (0.80-0.89). This compares to AUCs ranging between 0.76-0.91 for pigment ratios and 0.61-0.76 for cell-counts. The wider range of values for cell-counts and pigments reflect bias in their capacity to identify certain different bio-optical groups (Figure 2.6). The capacity to distinguish diatom-labelled samples from cell-counts data is clearly higher than the capacity to recognise *Phaeocystis* colonies for example. The higher AUC values for bio-optical datasets is effectively tautological, as the groups we are aiming to distinguish were originally defined using bio-optical data. Nevertheless, we view this as a useful result, given that it is routine practice to use different data types to validate remote-sensing algorithms (e.g. Liu et al. (2018), using pigment ratios to identify different phytoplankton size-classes with a simple linear additive analysis of key pigment species, known as ‘Diagnostic Pigment Analysis’ (Vidussi et al. 2001)). We emphasise here that switching between different data types in this process is liable to introduce error and potential bias.

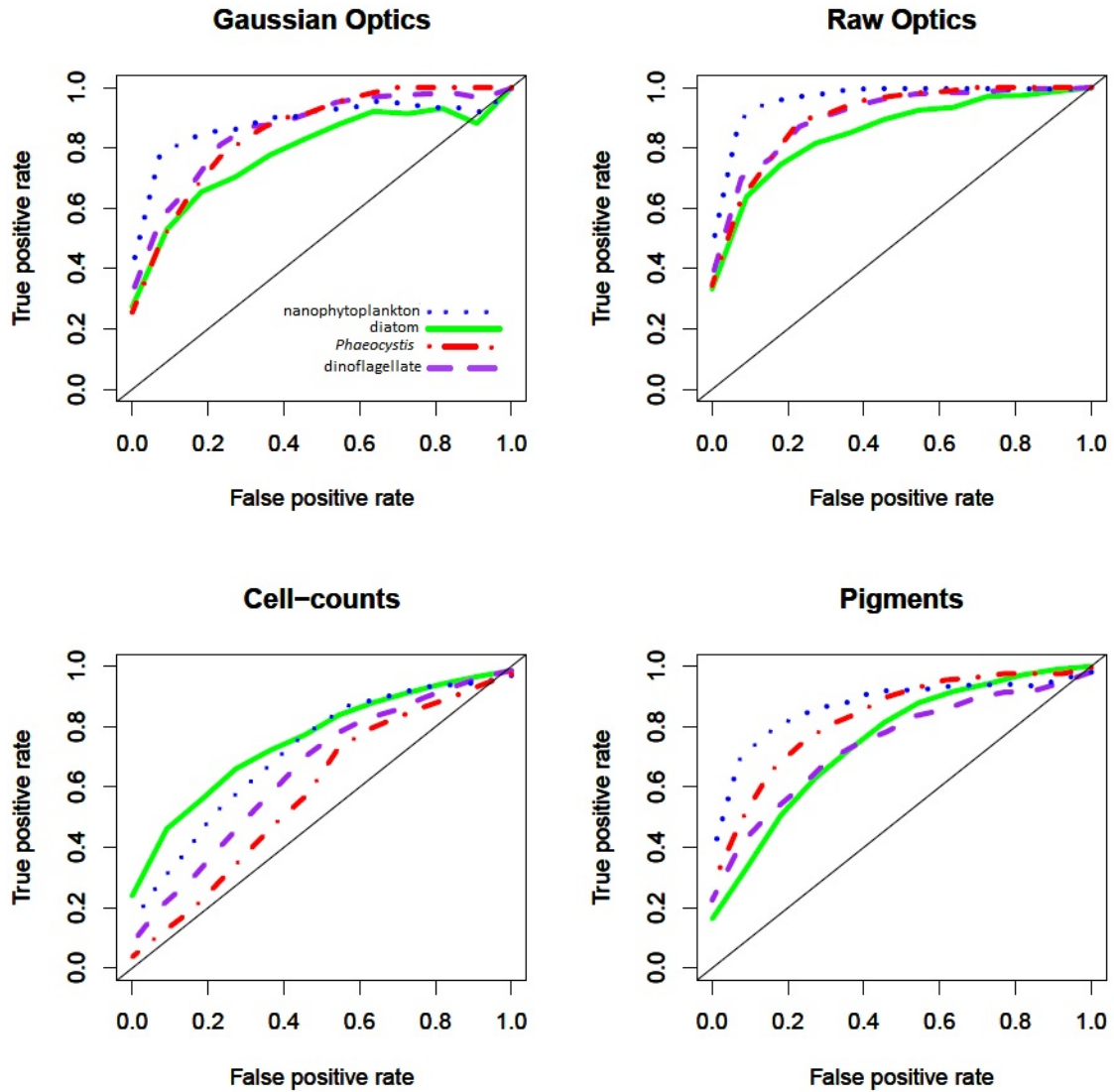


Figure 2.6: ROC plots for the capacity of different phytoplankton community structure data sets to correctly classify phytoplankton samples to distinct bio-optical groups. **Top-left)** Gaussian optics **Top-right)** Raw optics **Bottom-left)** Cell-counts **Bottom-right)** Pigment ratios. Green solid line (diatom-dominated), blue dotted line (nanophytoplankton-dominated), purple dashed line (dinoflagellate-dominated), red dot-dashed line (*Phaeocystis*-dominated). In each case 100 random re-samplings have been used to generate average ROC curves. The black solid line represents a line of equal probability that a sample is correctly or incorrectly classified.

2.8 Discussion

2.8.1 Influence of data type on perception of functional relationships

The intersection of Atlantic and Arctic clines in diatom-dominated spring samples conforms to reports of a diatom-dominated spring bloom, relatively homogeneous in composition across the Barents Sea (Degerlund and Eilertsen 2010). The large number of samples at this intersection belonging to the Barents Water is as expected, given that most sampling in spring was confined to this water-mass (Figure 2.1). The deviation of cell-counts from this pattern, in which diatom-labelled samples are divergent and clustering is more strongly associated with distinct water-masses, is compatible with previous studies showing a small number of highly abundant diatom ‘indicator’ species, with specific environmental preferences (Rat’kova and Wassmann 2002), dominate seasonal succession. The leading variables in RDA derived from cell-counts are all highly-abundant diatom species (*Chaetoceros socialis*, *Fragilariopsis*, *Chaetoceros furcellatus*, *Thalassiosira nordenskioeldii*, *Nitzschia frigida* and *Navicula pelagica*) previously noted in the literature for their strong environmental preferences and seasonal succession (Wassmann et al. 2006a; Sakshaug et al. 2009). For example, *Chaetoceros socialis* and some species of *Thalassiosira* tend to be restricted to colder waters. *Navicula* *Fragilariopsis* and *Nitzschia* often occur in sympagic contexts, and *Pseudo-Nitzschia* tends to be confined to post-bloom settings later in the season (Rat’kova and Wassmann 2002; Sakshaug et al. 2009; Degerlund and Eilertsen 2010; Berge et al. 2020).

The strong environmental preferences of these indicator species, which are not clearly discriminated by their bio-optical properties or pigment contents, may explain the perception of a greater number of significant environmental relationships in the cell-counts dataset (Figure 2.5b), compared to other descriptions of phytoplankton community composition.

The weak clustering of dinoflagellate-labelled samples in cell-counts and pigment ratio

analyses, compared to bio-optics, may result from their subsumption within the HNAN counts, and with their lack of possession of taxon-specific photosynthetic pigments, consistent with a heterotrophic lifestyle. The lack of explicit enumeration of *Phaeocystis* in the cell-counts dataset may explain why *Phaeocystis* is more weakly clustered in the cell-counts analysis than the pigment or bio-optical analyses.

In spite of these nuanced differences between the analyses based on the different community compositional datasets, the patterns of seasonal succession in Arctic and Atlantic water-masses are broadly similar across the analyses. The Arctic succession from autotrophic diatom-labelled samples and colonial *Phaeocystis*, towards mixotrophic communities with higher proportions of dinoflagellates, ciliates and heterotrophic nanoflagellates, recapitulates the sequence of phytoplankton lifeforms suggested by Margalef (1978). Phytoplankton assemblages resorting to mixotrophic or heterotrophic lifestyles can be interpreted as a response to prevailing oligotrophy that results from nutrient draw-down by autotrophs and the development of a more stratified water column as the season progressed (see Sakshaug et al. (2009) pages 167-208 for a generalised scheme of the the phenology of primary production in the Barents Sea). Observe, for example, how ammonium and stratification covary across the Redundancy Analysis biplots in Figure 2.4, and how they are opposed to the vectors describing nutrient concentration. We also remark that a prevalence of mixotrophic organisms in high summer reminds us of reports of a mixotrophic and heterotrophic community assemblages dominating sparse over-wintering communities in the Barents Sea (see Berge et al. (2020), pages 67-112). Post-bloom summer communities differ from the Polar Night, where autotrophy is prohibited by darkness, rather than by oligotrophy.

In contrast to the phenological cascade in Arctic Waters, the Atlantic succession grades from microphytoplankton into nanophytoplankton-dominated communities, which may result from weaker physical inhibition of vertical mixing in the thermally-stratified waters of the Southern Barents Sea. This may enable regenerated production via the microbial loop

to continue supporting autotrophic lifestyles late into the summer season. The diverging correlation of nanophytoplankton-labelled samples with ammonium, and diatom-labelled samples with nitrate, is robust across all datasets (Figure 2.4) and is compatible with this interpretation. Additionally, the common logarithm of the depth of maximum Brunt-Väisälä frequency is significantly associated with salinity (p -value = 0.0071, R^2 = 0.083), supporting our assessment that access to deep nutrient stocks differs between haline- and thermally-stratified waters.

A seasonal succession in the Southern Barents Sea, culminating in regenerated production by autotrophic nanophytoplankton, is compatible with reports of greater annually-integrated primary production compared to more Arctic-influenced sites (e.g. Wassmann et al. (2006a), Wassmann et al. (2006b)). Particulate organic carbon flux measurements, furthermore, demonstrate weaker surface retention in Arctic water-masses (Reigstad et al. 2008). A part of this variation in tendency to retention and regenerated production is explain by faecal pellet export efficacy, which is lower in Atlantic-derived Barents Sea water-masses, resulting from mechanisms such as coprophagy (e.g. Riser et al. (2002) Riser et al. (2007)), with over 90% of faecal pellet production retained in the upper 50 m in post-bloom Atlantic settings. Lower export ratios in Southern Barents Waters are commensurate with reports in the literature that megafaunal benthic biomass is lower in regions of the Barents benthos nourished by Atlantic-dominated water-masses (Jørgensen et al. 2014; Degen et al. 2016).

Overall, the different data types converge to show similar patterns of seasonal succession and hydrographic associations in phytoplankton community structure, but we caution that a scheme of PFTs based on cell-counts may result in Functional Types being defined that cannot easily be discriminated by pigment ratios and which may not represent feasible targets for detection in remotely-sensed ocean-colour.

2.8.2 Capacity of community composition data sets to validate remote-sensing algorithms based on bio-optical variation

It is possible that it is difficult to predict *Phaeocystis*-labelled samples from cell-counts data, as the colonies may have been too large to pass through the aperture of the cytometer. Alternatively, low predictive power may result because PNAN cells are usually a majority of counted cells and hence their presence is not a unique indicator of *Phaeocystis* colonies.

There are several possible explanations for difficulty predicting dinoflagellate-labelled samples from cell-counts data. The low representation of dinoflagellate cells in our dataset (0.11% on average) may be responsible. At this scale of sampling, signal to noise ratios may be too low to facilitate meaningful inferences. Alternatively, if many hetero- or mixotrophic dinoflagellate cells have been enumerated as HNAN by flow cytometry, it is possible that dinoflagellate counts have been pooled with a variety of other taxa and hence that HNAN counts are, by themselves, not a unique indicator of dinoflagellate presence. This may be especially likely for dinoflagellates that lack armour.

As HPLC samples were typically over 1 L in volume, we do not believe low sampling intensity can explain low capacity to predict dinoflagellate-labelled samples from pigment ratios. Kleptoplastidy, the appropriation of plastids from prey, is practised in dinoflagellate species belonging to genera such as *Gymnodinium* (Skovgaard 1998) and *Dinophysis* (Nishitani et al. 2012), both of which feature in our floristic descriptions of Barents Sea samples, perhaps meaning that pigment ratios associated with dinoflagellate presence may be especially plastic, or simply reflect the ratios present in the wider community of prey species.

We also remark that the dispersion of dinoflagellate optical properties is relatively large, compared to other bio-optical clusters (Figure 2.3). We interpret this as compatible with reports that there are up to 6 different dinoflagellates clusters with distinct pigment complements (Jeffrey et al. 2011). This complexity may obfuscate clear relationships between

dinoflagellate-labelled samples as a whole and matrices of their pigment complements.

The relatively high capacity to predict diatom-labelled samples from the cell-counts data is probably a result of highly abundant diatom cells being well-represented in counts. However, it should be noted that the AUC values for cell-counts tend to be pervasively lower, compared to other data types. The low relative AUCs for diatoms, compared to other bio-optical clusters, in Figure 2.6, may be a result from the selection of diatom-labelled samples in our study not being representative. It can be seen in Figure 2.3 that diatom-labelled samples available in our study were diffusely distributed and few came from the central region of the diatom-cluster identified in a broader range of Barents phytoplankton samples.

2.9 Conclusion

In conclusion, our investigations across multiple data types, describing phytoplankton community composition, all retrieved similar functional relationships to a first order between the phytoplankton and hydrography. However, there are differences in the subtler environmental associations revealed by different data types. We find cell-counts have the greatest capacity to identify significant environmental variables explaining variation in phytoplankton community structure, because of a small number of highly abundant indicator species.

Overall, the gross relationships between phytoplankton community structure and environment derived from pigment-ratios and bio-optical properties are more coherent with one another than either of them are with cell-counts. We believe this means that, for the purpose of identifying Phytoplankton Functional Types as targets for remote-sensing, where bio-optical variation is the only property amenable to analysis, that bio-optical (here $a_{ph}(\lambda)$) properties are the most appropriate means of quantifying phytoplankton community structure. The study of fluorescence-excitation spectra, which may also impart distinct spectral signatures (Johnsen and Sakshaug 2007) that could influence remotely-sensed ocean-colour,

is beyond the remit of the current work but identified as a potential future avenue of exploration.

If $a_{ph}(\lambda)$ measurements are unavailable, pigment ratios should be chosen ahead of cell-counts. We stress that these conclusions are specific to the Barents Sea, because the relationships between cell-counts, bio-optics and pigment ratios may be structured differently in other parts of the Global Ocean.

When we probed the capacity of cell-counts, pigment-ratios and bio-optical properties to identify distinct bio-optical clusters of phytoplankton, a key step in the validation of remote-sensing algorithms, we found that raw hyperspectral bio-optical properties possessed the greatest predictive skill. Pigment ratios and cell-counts both had substantially less predictive skill and exhibited considerable bias.

In sum, we suggest that the difficulty previous authors have encountered in the validation of remote-sensing algorithms of Phytoplankton Functional Types may result from the use of pigment ratios or cell-counts to represent variation in phytoplankton community structure, because these representations of community structure do not map directly onto bio-optical variation.

2.10 Data availability statement

Bio-optical datasets are available via the British Oceanographic Data Centre:

Orkney A.; Bouman H.(2019). Phytoplankton absorption spectra, derived from measurements of absorption from phytoplankton samples taken in the Barents Sea during cruise JR16006. British Oceanographic Data Centre, National Oceanography Centre, NERC, UK: [doi:10.5285/97daa7ea-8792-6cff-e053-6c86abc0dd46]

Orkney A.; Bouman H.(2019). Phytoplankton absorption spectra, derived from measurements of absorption from phytoplankton samples taken in the Barents Sea during cruise HH180423. British Oceanographic Data Centre, National Oceanography Centre, NERC,

UK:

[doi:10.5285/97daa7ea-8792-6cff-e053-6c86abc0dd46]

Orkney A.; Bouman H.(2019). Phytoplankton absorption spectra, derived from measurements of absorption from phytoplankton samples taken in the Barents Sea during cruise JR17006. British Oceanographic Data Centre, National Oceanography Centre, NERC, UK: [doi:10.5285/982b6da2-7e12-060a-e053-6c86abc09389]

Cell-count datasets are available at:

Mitchell, E., McNeil, S., Whyte, C., Cottier, F., Hopkins, J., & Davidson, K. (2021). Phytoplankton enumeration and biomass calculations from samples collected in the Barents Sea during 2017-2018 (Version 1.0) [Data set]. UK Polar Data Centre, Natural Environment Research Council, UK Research & Innovation.

[https://doi.org/10.5285/69ae1fa2-cf1d-44da-8415-d9cc512c0256]

Flow-cytometry datasets are available at:

Mitchell, E., McNeil, S., Whyte, C., Cottier, F., Hopkins, J., & Davidson, K. (2021). Flow cytometry analysis of water samples for bacterial and Pico-plankton enumeration, samples collected in the Barents Sea during 2017-2018 (Version 1.0) [Data set]. UK Polar Data Centre, Natural Environment Research Council, UK Research & Innovation.

[https://doi.org/10.5285/2182fd8d-0a0d-4f56-9bf9-4dffaa67ff4b]

Hydrographic data are available via the British Oceanographic Data Centre:

Dumont E., Brand T., Hopkins J.(2019). CTD data from NERC Changing Arctic Ocean Cruise JR16006 on the RRS James Clark Ross, June-August 2017. British oceanographic Data Centre, National Oceanography Centre, NERC, UK:

[doi:10.5285/89a3a6b8-7223-0b9c-e053-6c86abc0f15d]

Nutrient assay data are available via the British Oceanographic Data Centre:

Brand T.; Norman L.; Henley S.F.; Mahaffey C.; Tuerena R.; Crocket K.C.(2020). Dissolved nutrient samples collected in the Barents Sea as part of the Changing Arctic Ocean

programme for the Arctic **PRIZE** and ARISE projects during RRS James Clark Ross cruise JR16006. British Oceanographic Data Centre, National Oceanography Centre, NERC, UK: [doi:10.5285/b4c1537e-c729-6463-e053-6c86abc0c7de]

Brand T.; Henley S.F.; Mahaffey C.; Crockett K.C.; Norman L.; Tuerena R.(2020). Dissolved nutrient samples collected in the Barents Sea as part of the Changing Arctic Ocean programme during cruise JR17006, June-July 2018. British Oceanographic Data Centre, National Oceanography Centre, NERC, UK: [doi:10.5285/b62f2d5d-1f3f-0c2c-e053-6c86abc0265d]

2.11 Author contributions

Elaine Mitchell acquired cell-counts and flow cytometric data.

Keith Davidson provided expert conceptual input on cell-counts and flow cytometric datasets and methods used in the existing literature to relate phytoplankton community structure to environmental variation.

Heather Bouman supervised, providing key conceptual input in analysis, and assisted in acquisition of bio-optical and pigment datasets.

Sian Henley provided critical input on nutrient data and biogeochemical processes.

Andrew Orkney acquired bio-optical and pigment samples, drafted the manuscript and performed analysis.

All authors contributed to the text of the manuscript. All authors have read and approved the final manuscript.

Elaine Mitchell and Keith Davidson's work was funded by Arctic **PRIZE** grant NE/P006302/1.

Heather Bouman and Andrew Orkney's work was funded by the Natural Environment Research Council (NERC) NE/P006507/1.

References

- Ardyna, Mathieu, Michel Gosselin, Christine Michel, Michel Poulin, and Jean-Éric Tremblay. 2011. "Environmental forcing of phytoplankton community structure and function in the Canadian High Arctic: contrasting oligotrophic and eutrophic regions." *Marine Ecology Progress Series* 442:37–57.
- Armstrong, Robert A, Michael L Peterson, Cindy Lee, and Stuart G Wakeham. 2009. "Settling velocity spectra and the ballast ratio hypothesis." *Deep Sea Research Part II: Topical Studies in Oceanography* 56 (18): 1470–1478.
- Astoreca, Rosa, Véronique Rousseau, Kevin Ruddick, Cécile Knechciak, Barbara Van Mol, Jean-Yves Parent, and Christiane Lancelot. 2008. "Development and application of an algorithm for detecting *Phaeocystis globosa* blooms in the Case 2 Southern North Sea waters." *Journal of Plankton Research* 31 (3): 287–300.
- Barton, Benjamin I, Yueng-Djern Lenn, and Camille Lique. 2020. "Observed atlantification of the Barents Sea causes the Polar Front to limit the expansion of winter sea ice." *JPO* 500:18–0003.
- Berge, Jørgen, Geir Johnsen, and Jonathan H Cohen. 2020. *POLAR NIGHT marine ecology: life and light in the dead of night*. Vol. 4. Springer Nature.
- Bracher, Astrid, Heather A Bouman, Robert JW Brewin, Annick Bricaud, Vanda Brotas, Aurea M Ciotti, Lesley Clementson, Emmanuel Devred, Annalisa Di Cicco, Stephanie Dutkiewicz, et al. 2017. "Obtaining phytoplankton diversity from ocean color: a scientific roadmap for future development." *Frontiers in Marine Science* 4:55.
- Brewin, Robert JW, Nick J Hardman-Mountford, Samantha J Lavender, Dionysios E Raitos, Takafumi Hirata, Julia Uitz, Emmanuel Devred, Annick Bricaud, Aurea Ciotti, and Bernard Gentili. 2011. "An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing." *Remote Sensing of Environment* 115 (2): 325–339.
- Carpenter, James H. 1965. "The Chesapeake Bay Institute technique for the Winkler dissolved oxygen method." *Limnology and Oceanography* 10 (1): 141–143.
- Cebeci, Zeynel, Figen Yildiz, A. Tuna Kavlak, Cagatay Cebeci, and Hasan Onder. 2019. "ppclust: Probabilistic and Possibilistic Cluster Analysis." R package version 0.1.2, <https://CRAN.R-project.org/package=ppclust>.
- Chase, Alison, Emmanuel Boss, Ronald Zaneveld, Annick Bricaud, Herve Claustre, Josephine Ras, Giorgio Dall’Olmo, and Toby K Westberry. 2013. "Decomposition of in situ particulate absorption spectra." *Methods in Oceanography* 7:110–124.
- Choudhury, Avik Kumar, and Ruma Pal. 2011. "Variations in seasonal phytoplankton assemblages as a response to environmental changes in the surface waters of a hypo saline coastal station along the Bhagirathi–Hooghly estuary." *Environmental monitoring and assessment* 179 (1-4): 531–553.
- Cottier, Finlo R. 2018. "RRS James Clark Ross JR17006," https://www.bodc.ac.uk/resources/inventories/cruise_inventory/report/16981/.
- Cottier, Finlo R, and Jørgen Berge. 2018a. "Helmer Hanssen HH180101," https://www.bodc.ac.uk/resources/inventories/cruise_inventory/report/16917/.
- . 2018b. "Helmer Hanssen HH180423 (Arctic PRIZE, Arctic SIZE)," https://www.bodc.ac.uk/resources/inventories/cruise_inventory/report/17059/.
- Davidson, Keith, Linda C Gilpin, Romain Pete, Debra Brennan, Sharon McNeill, Grigorios Moschonas, and Jonathon Sharples. 2013. "Phytoplankton and bacterial distribution and productivity on and around Jones Bank in the Celtic Sea." *Progress in Oceanography* 117:48–63.
- Degen, Renate, Lis Lindal Jørgensen, Pavel Ljubin, Ingrid H Ellingsen, Hendrik Pehlke, and Thomas Brey. 2016. "Patterns and drivers of megabenthic secondary production on the Barents Sea shelf." *Marine Ecology Progress Series* 546:1–16.

- Degerlund, Maria, and Hans Christian Eilertsen. 2010. "Main species characteristics of phytoplankton spring blooms in NE Atlantic and Arctic waters (68–80 N)." *Estuaries and coasts* 33 (2): 242–269.
- Drinkwater, Kenneth F. 2006. "The regime shift of the 1920s and 1930s in the North Atlantic." *Progress in Oceanography* 68 (2-4): 134–151.
- Fragoso, Glaucia M, Alex J Poulton, Igor M Yashayaev, Erica JH Head, and Duncan A Purdie. 2017. "Spring phytoplankton communities of the Labrador Sea (2005–2014): pigment signatures, photophysiology and elemental ratios." *Biogeosciences* 14:1235–1259.
- Fuhrman, Jed A, and Douglas G Capone. 2001. "Nifty nanoplankton." *Nature* 412 (6847): 593–594.
- Galí, Martí, Emmanuel Devred, Marcel Babin, and Maurice Levasseur. 2019. "Decadal increase in Arctic dimethylsulfide emission." *Proceedings of the National Academy of Sciences* 116 (39): 19311–19317.
- Galili, Tal. 2015. "dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering." *Bioinformatics*, <https://doi.org/10.1093/bioinformatics/btv428>. eprint: <https://academic.oup.com/bioinformatics/article-pdf/31/22/3718/17122682/btv428.pdf>. <https://academic.oup.com/bioinformatics/article/31/22/3718/240978/dendextend-an-R-package-for-visualizing-adjusting>.
- Goericke, Ralf, and Joseph P Montoya. 1998. "Estimating the contribution of microalgal taxa to chlorophyll a in the field— variations of pigment ratios under nutrient-and light-limited growth." *Marine Ecology Progress Series* 169:97–112.
- González Taboada, Fernando, and Ricardo Anadón. 2014. "Seasonality of North Atlantic phytoplankton from space: impact of environmental forcing on a changing phenology (1998–2012)." *Global change biology* 20 (3): 698–712.
- Gran-Stadniczenko, Sandra, Luka Šupraha, Elianne D Egge, and Bente Edvardsen. 2017. "Haptophyte diversity and vertical distribution explored by 18S and 28S ribosomal RNA gene metabarcoding and scanning electron microscopy." *Journal of Eukaryotic Microbiology* 64 (4): 514–532.
- Hoepffner, Nicolas, and Shubha Sathyendranath. 1992. "Bio-optical characteristics of coastal waters: Absorption spectra of phytoplankton and pigment distribution in the western North Atlantic." *Limnology and Oceanography* 37 (8): 1660–1679.
- . 1993. "Determination of the major groups of phytoplankton pigments from the absorption spectra of total particulate matter." *Journal of Geophysical Research: Oceans* 98 (C12): 22789–22803.
- Hooker, Stanford Baird. 2005. *Second SeaWiFS HPLC analysis round-robin experiment (SeaHARRE-2)*. Vol. 212785. National Aeronautics / Space Administration, Goddard Space Flight Center.
- Hopkins, Jo. 2018. "RRS James Clark Ross JR16006," <https://www.bodc.ac.uk/resources/inventories/cruise.inventory/report/16774/>.
- Hoppenrath, Mona, Malte Elbrächter, and Gerhard Drebes. 2009. "Marine phytoplankton."
- Hovland, Erlend Kjeldsberg, Heidi M Dierssen, Ana Sofia Ferreira, and Geir Johnsen. 2013. "Dynamics regulating major trends in Barents Sea temperatures and subsequent effect on remotely sensed particulate inorganic carbon." *Marine Ecology Progress Series* 484:17–32.
- Jackson, Thomas, Heather A Bouman, Shubha Sathyendranath, and Emmanuel Devred. 2010. "Regional-scale changes in diatom distribution in the Humboldt upwelling system as revealed by remote sensing: implications for fisheries." *ICES Journal of Marine Science* 68 (4): 729–736.
- Janišová, Monika, Richard Hrivnák, Dušan Gömöry, Karol Ujházy, Milan Valachovič, Erika Gömöryová, Katarína Hegedúšová, and Iveta Škodová. 2007. "Changes in understorey

- vegetation after Norway spruce colonization of an abandoned grassland.” In *Annales Botanici Fennici*, 256–266. JSTOR.
- Jeffrey, SW, SW Wright, and M Zapata. 2011. “Microalgal classes and their signature pigments.”
- Johnsen, Geir, Runar Dalløkken, Wenche Eikrem, Catherine Legrand, Jan Aure, and Hein Rune Skjoldal. 1999. “Eco-physiology, bio-optics and toxicity of the ichthyotoxic *Chrysochromulina leadbeateri* (Prymnesiophyceae).” *Journal of Phycology* 35 (6): 1465–1476.
- Johnsen, Geir, and Egil Sakshaug. 1993. “Bio-optical characteristics and photoadaptive responses in the toxic and bloom-forming dinoflagellates *Gyrodinium aureolum*, *Gymnodinium galatheanum* and two strains of *Prorocentrum minimum*.” *Journal of Phycology* 29 (5): 627–642.
- . 2007. “Biooptical characteristics of PSII and PSI in 33 species (13 pigment groups) of marine phytoplankton, and the relevance for pulse-amplitude-modulated and fast-repetition-rate fluorometry 1.” *Journal of Phycology* 43 (6): 1236–1251.
- Jørgensen, Lis Lindal, Pavel Ljubin, Hein Rune Skjoldal, Randi B Ingvaldsen, Natalia Anisimova, and Igor Manushin. 2014. “Distribution of benthic megafauna in the Barents Sea: baseline for an ecosystem approach to management.” *ICES Journal of Marine Science* 72 (2): 595–613.
- Kelley, Dan, and Clark Richards. 2020. *oce: Analysis of Oceanographic Data*. R package version 1.2-0. <https://CRAN.R-project.org/package=oce>.
- Kishino, M, M Takahashi, N Okami, and S Ichimura. 1985. “Estimation of the spectral absorption coefficients of phytoplankton in the sea.” *Bulletin of Marine Science* 37 (2): 634–642.
- Lalande, Catherine, Eduard Bauerfeind, Eva-Maria Nöthig, and Agnieszka Beszczynska-Möller. 2013. “Impact of a warm anomaly on export fluxes of biogenic matter in the eastern Fram Strait.” *Progress in Oceanography* 109:70–77.
- Legendre, Pierre, and Loic FJ Legendre. 2012. *Numerical ecology*. Elsevier.
- Li, William KW, Fiona A McLaughlin, Connie Lovejoy, and Eddy C Carmack. 2009. “Smallest algae thrive as the Arctic Ocean freshens.” *Science* 326 (5952): 539–539.
- Liu, Xiaohan, Emmanuel Devred, and Catherine Johnson. 2018. “Remote sensing of phytoplankton size class in Northwest Atlantic from 1998 to 2016: bio-optical algorithms comparison and application.” *Remote Sensing* 10 (7): 1028.
- Liu, Yangyang, Emmanuel Boss, Alison Chase, Hongyan Xi, Xiaodong Zhang, Rüdiger Röttgers, Yanqun Pan, and Astrid Bracher. 2019. “Retrieval of phytoplankton pigments from underway spectrophotometry in the Fram Strait.” *Remote Sensing* 11 (3): 318.
- Lohrenz, Steven E, Alan D Weidemann, and Merritt Tuel. 2003. “Phytoplankton spectral absorption as influenced by community size structure and pigment composition.” *Journal of Plankton Research* 25 (1): 35–61.
- Lovejoy, Connie, Ramon Massana, and Carlos Pedrós-Alió. 2006. “Diversity and distribution of marine microbial eukaryotes in the Arctic Ocean and adjacent seas.” *Applied and environmental microbiology* 72 (5): 3085–3095.
- Mackey, MD, DJ Mackey, HW Higgins, and SW Wright. 1996. “CHEMTAX-a program for estimating class abundances from chemical markers: application to HPLC measurements of phytoplankton.” *Marine Ecology Progress Series* 144:265–283.
- Margalef, R. 1978. “Life forms of phytoplankton as survival alternatives in an unstable environment.” *Oceanologica Acta* 1:4–493.
- McGill, Brian J, Brian J Enquist, Evan Weiher, and Mark Westoby. 2006. “Rebuilding community ecology from functional traits.” *Trends in ecology & evolution* 21 (4): 178–185.
- Meersche, Karel Van den, Karline Soetaert, and Dick Van Oevelen. 2009. “xsample (): an R function for sampling linear inverse problems.” *Journal of Statistical Software* 30 (Code Snippet 1).

- Nair, Anitha, Shubha Sathyendranath, Trevor Platt, Jesus Morales, Venetia Stuart, Marie-Hélène Forget, Emmanuel Devred, and Heather Bouman. 2008. "Remote sensing of phytoplankton functional types." *Remote Sensing of Environment* 112 (8): 3366–3375.
- Nishitani, Goh, Satoshi Nagai, Shiho Hayakawa, Yuki Kosaka, Kiyonari Sakurada, Takashi Kamiyama, and Takashi Gojobori. 2012. "Multiple plastids collected by the dinoflagellate *Dinophysis mitra* through kleptoplastidy." *Appl. Environ. Microbiol.* 78 (3): 813–821.
- Nöthig, Eva-Maria, Astrid Bracher, Anja Engel, Katja Metfies, Barbara Niehoff, Ilka Peeken, Eduard Bauerfeind, Alexandra Cherkasheva, Steffi Gäbler-Schwarz, Kristin Hardge, et al. 2015. "Summertime plankton ecology in Fram Strait- a compilation of long and short-term observations." *Polar Research* 34 (1): 233–49.
- Oksanen, Jari, F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, et al. 2019. *vegan: Community Ecology Package*. R package version 2.5-6. <https://CRAN.R-project.org/package=vegan>.
- Onda, Deo FL, Emmanuelle Medrinal, André M Comeau, Mary Thaler, Marcel Babin, and Connie Lovejoy. 2017. "Seasonal and interannual changes in ciliate and dinoflagellate species assemblages in the Arctic Ocean (Amundsen Gulf, Beaufort Sea, Canada)." *Frontiers in Marine Science* 4:16.
- Orkney, Andrew, Trevor Platt, Bhavani E Narayanaswamy, Ina Kostakis, and Heather A Bouman. 2020. "Bio-Optical evidence for increasing *Phaeocystis* dominance in the Barents Sea." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*.
- Overland, James, Edward Dunlea, Jason E Box, Robert Corell, Martin Forsius, Vladimir Kattsov, Morten Skovgård Olsen, Janet Pawlak, Lars-Otto Reiersen, and Muyin Wang. 2018. "The urgency of Arctic change." *Polar Science*.
- Oziel, Laurent, G Neukermans, M Ardyna, C Lancelot, J-L Tison, Paul Wassmann, Jérôme Sirven, Diana Ruiz-Pino, and J-C Gascard. 2017. "Role for Atlantic inflows and sea ice loss on shifting phytoplankton blooms in the Barents Sea." *Journal of Geophysical Research: Oceans* 122 (6): 5121–5139.
- Oziel, Laurent, Jérôme Sirven, and Jean-Claude Gascard. 2016. "The Barents Sea frontal zones and water masses variability (1980–2011)." *Ocean Science* 12(1):169–184.
- Peperzak, L, WWC Gieskes, R Duin, and F Colijn. 2000b. "The vitamin B requirement of *Phaeocystis globosa* (Prymnesiophyceae)." *Journal of Plankton Research* 22 (8): 1529–1537.
- Peperzak, Louis. 2010. "An objective procedure to remove observer-bias from phytoplankton time-series." *Journal of Sea Research* 63 (2): 152–156.
- Peralta-Ferriz, Cecilia, and Rebecca A Woodgate. 2015. "Seasonal and interannual variability of pan-Arctic surface mixed layer properties from 1979 to 2012 from hydrographic data, and the dominance of stratification for multiyear mixed layer depth shoaling." *Progress in Oceanography* 134:19–53.
- Pettersen, Ragnhild, Geir Johnsen, Jørgen Berge, and Erlend Kjeldsberg Hovland. 2011. "Phytoplankton chemotaxonomy in waters around the Svalbard archipelago reveals high amounts of Chl b and presence of gyroxanthin-diester." *Polar biology* 34 (5): 627–635.
- Platt, T, S Sathyendranath, and V Stuart. 2006. "Why study biological oceanography?" *Aquabiology* 28 (5): 542.
- Poulton, Nicole J. 2016. "FlowCam: Quantification and classification of phytoplankton by imaging flow cytometry." In *Imaging Flow Cytometry*, 237–247. Springer.
- Quere, Corinne Le, Sandy P Harrison, I Colin Prentice, Erik T Buitenhuis, Olivier Aumont, Laurent Bopp, Hervé Claustre, Leticia Cotrim Da Cunha, Richard Geider, Xavier Giraud, et al. 2005. "Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models." *Global Change Biology* 11 (11): 2016–2040.

- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rat'kova, Tatjana N, and Paul Wassmann. 2002. "Seasonal variation and spatial distribution of phyto-and protozooplankton in the central Barents Sea." *Journal of Marine Systems* 38 (1-2): 47–75.
- Reigstad, Marit, Christian Wexels Riser, Paul Wassmann, and Tatjana Ratkova. 2008. "Vertical export of particulate organic carbon: attenuation, composition and loss rates in the northern Barents Sea." *Deep Sea Research Part II: Topical Studies in Oceanography* 55 (20-21): 2308–2319.
- Riser, Christian Wexels, Marit Reigstad, Paul Wassmann, Elena Arashkevich, and Stig Falk-Petersen. 2007. "Export or retention? Copepod abundance, faecal pellet production and vertical flux in the marginal ice zone through snap shots from the northern Barents Sea." *Polar Biology* 30 (6): 719–730.
- Riser, Christian Wexels, Paul Wassmann, Kalle Olli, Anna Pasternak, and Elena Arashkevich. 2002. "Seasonal variation in production, retention and export of zooplankton faecal pellets in the marginal ice zone and central Barents Sea." *Journal of Marine Systems* 38 (1-2): 175–188.
- Robin, Xavier, Natacha Turck, Alexandre Hainard, Natalia Tiberti, Frédérique Lisacek, Jean-Charles Sanchez, and Markus Müller. 2011. "pROC: an open-source package for R and S+ to analyze and compare ROC curves." *BMC Bioinformatics* 12:77.
- Roy, Suzanne, Carole A Llewellyn, Einar Skarstad Egeland, and Geir Johnsen. 2011. *Phytoplankton pigments: characterization, chemotaxonomy and applications in oceanography*. Cambridge University Press.
- Sakshaug, Egil, Geir Helge Johnsen, and Kit M Kovacs. 2009. *Ecosystem Barents Sea*. Tapir Academic Press.
- Sathyendranath, Shubha, Louisa Watts, Emmanuel Devred, Trevor Platt, Carla Caverhill, and Heidi Maass. 2004. "Discrimination of diatoms from other phytoplankton using ocean-colour data." *Marine Ecology Progress Series* 272:59–68.
- Sieburth, John McN, Victor Smetacek, and Jürgen Lenz. 1978. "Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions 1." *Limnology and oceanography* 23 (6): 1256–1263.
- Skovgaard, Alf. 1998. "Role of chloroplast retention in a marine dinoflagellate." *Aquatic Microbial Ecology* 15 (3): 293–301.
- Sosik, HM, S Sathyendranath, J Uitz, H Bouman, and A Nair. 2014. "In situ methods of measuring phytoplankton functional types." *Phytoplankton Functional Types from Space. Reports of the International Ocean-Colour Coordinating Group (IOCCG)*, no. 15, 21–38.
- Thrandsen, Jahn, Grethe Rytter Hasle, and Karl Tangen. 2007. *Phytoplankton of Norwegian coastal waters*. Almatel Forlag AS.
- Tomas, Carmelo R. 1997. *Identifying marine phytoplankton*. Elsevier.
- Utermöhl, Hans. 1958. "Zur vervollkommnung der quantitativen phytoplankton-methodik: mit 1 Tabelle und 15 abbildungen im Text und auf 1 Tafel." *Internationale Vereinigung für theoretische und angewandte Limnologie: Mitteilungen* 9 (1): 1–38.
- Våge, Kjetil, Robert S Pickart, Vladimir Pavlov, Peigen Lin, Daniel J Torres, Randi Ingvaldsen, Arild Sundfjord, and Andrey Proshutinsky. 2016. "The Atlantic Water boundary current in the Nansen Basin: Transport and mechanisms of lateral exchange." *Journal of Geophysical Research: Oceans* 121 (9): 6946–6960.
- Vidussi, Francesca, Hervé Claustre, Beniamino B Manca, Anna Luchetta, and Jean-Claude Marty. 2001. "Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter." *Journal of Geophysical Research: Oceans* 106 (C9): 19939–19956.
- Wassmann, Paul, Marit Reigstad, Tore Haug, Bert Rudels, Michael L Carroll, Haakon Hop, Geir Wing Gabrielsen, Stig Falk-Petersen, Stanislav G Denisenko, Elena Arashkevich,

- et al. 2006a. "Food webs and carbon flux in the Barents Sea." *Progress in Oceanography* 71 (2-4): 232–287.
- Wassmann, Paul, Dag Slagstad, and Ingrid Ellingsen. 2010. "Primary production and climatic variability in the European sector of the Arctic Ocean prior to 2007: preliminary results." *Polar Biology* 33 (12): 1641–1650.
- Wassmann, Paul, Dag Slagstad, Christian Wexels Riser, and Marit Reigstad. 2006b. "Modelling the ecosystem dynamics of the Barents Sea including the marginal ice zone: II. Carbon flux and interannual variability." *Journal of Marine Systems* 59 (1-2): 1–24.
- Wassmann, Paul, Maria Vernet, B Greg Mitchell, and Francisco Rey. 1990. "Mass sedimentation of *Phaeocystis pouchetii* in the Barents Sea." *Marine Ecology Progress Series* 66:183–195.
- Widdicombe, CE, D Eloire, D Harbour, RP Harris, and PJ Somerfield. 2010. "Long-term phytoplankton community dynamics in the Western English Channel." *Journal of plankton research* 32 (5): 643–655.
- Wollenburg, JE, C Katlein, Gernot Nehrke, E-M Nöthig, Jens Matthiessen, Dieter A Wolf-Gladrow, A Nikolopoulos, F Gázquez-Sanchez, L Rossmann, Philipp Assmy, et al. 2018. "Ballasting by cryogenic gypsum enhances carbon export in a *Phaeocystis* under-ice bloom." *Scientific Reports* 8 (1): 7703.
- Yool, A, EE Popova, and TR Anderson. 2011. "Medusa-1.0: a new intermediate complexity plankton ecosystem model for the global domain." *Geoscientific Model Development* 4 (2): 381–417.
- Zhang, Jinlun, Carin Ashjian, Robert Campbell, Yvette H Spitz, Michael Steele, and Victoria Hill. 2015. "The influence of sea ice and snow cover and nutrient availability on the formation of massive under-ice phytoplankton blooms in the Chukchi Sea." *Deep Sea Research Part II: Topical Studies in Oceanography* 118:122–135.
- Zubkov, Mikhail V, and Peter H Burkill. 2006. "Syringe pumped high speed flow cytometry of oceanic phytoplankton." *Cytometry Part A: The Journal of the International Society for Analytical Cytology* 69 (9): 1010–1019.
- Zubkov, Mikhail V, Peter H Burkill, and Juliette N Topping. 2007. "Flow cytometric enumeration of DNA-stained oceanic planktonic protists." *Journal of Plankton Research* 29 (1): 79–86.

Chapter 3

Development of a Bio-Optical algorithm

Andrew Orkney ^{1,a}, Ina Kostakis ^{2,b} David McKee ^{3,c} Trevor Platt ⁴, Shubha Sathyendranath ^{4,d},
and Heather Bouman ^{1,e}

^a <https://orcid.org/0000-0003-4972-2541>

^b <https://orcid.org/0000-0003-1167-2227>

^c <https://orcid.org/0000-0001-8023-5923>

^d <https://orcid.org/0000-0003-3586-192X>

^e <https://orcid.org/0000-0002-7407-9431>

¹Department of Earth Sciences, University of Oxford, Oxford, United Kingdom

²School of Computing, University of Portsmouth, Portsmouth, United Kingdom

³Department of Physics, University of Strathclyde, Glasgow, United Kingdom

⁴Plymouth Marine Laboratory, Plymouth, United Kingdom

3.1 Prelude

This chapter forms the prelude and supplement to a submission “ Bio-optical evidence for increasing *Phaeocystis* dominance in the Barents Sea ”, published in Philosophical Transactions of the Royal Society A. Further detail is provided in the prelude to chapter 4.

This chapter serves to explain how bio-optically distinct (distinctive $a_{ph}(\lambda)$ signatures) phytoplankton functional assemblages (occupying different environmental or phenological niches) were identified, in samples taken from the Barents Sea under Arctic **PRIZE**, and related to patterns of variance in remotely-sensed ocean-colour data from the MODIS-Aqua sensor. This effort culminates in the development of a bio-optical algorithm to distinguish salient community compositional variation in the Barents Sea. Cross-validation approaches and comparisons to a multi-annual *in-situ* phytoplankton sampling record in the Labrador Sea are employed to demonstrate efficacy. However, the requirement to correct for temporal drift in the concurrent MODIS-Aqua reprocessing emphasises that the discrimination of functional assemblages exploits fine variation between ocean-colour wavebands. Hence, great care must be taken in the use of such ocean-colour algorithms and, as such, any new application of our bio-optical algorithm would require a bespoke validation scheme. We emphasise that the greatest value is extracted from bio-optical ocean-colour algorithms when they are explicitly considered in conjunction with *in-situ* sampling campaigns. Space-borne ocean-colour sensors experience accumulating damage and consequent decay in function or drift in output values, which requires constant correction through re-processings. Re-processing is usually focussed on correction of commonly-used derived variables such as $[chl-a]$, rather than ensuring the fidelity of individual waveband ratios. Researchers must therefore be keenly aware that algorithms developed to infer derived variables, such as Phytoplankton Functional Types, may experience loss of function should damage to remote-sensors accumulate or re-processing methods change. A salient example is reported by Gregg et al. (2017), who demonstrated that, following a new re-processing of SEAW-

iFS, MODIS-Aqua and VIIRS data, apparent secular declines in remotely-sensed [*chl-a*] (Behrenfeld et al. 2006) were no longer significant. Hence, researchers retrieving Functional Types from ocean-colour should be careful to restrict themselves to studying only temporal periods and geographic regions where there are records of *in-situ* phytoplankton community composition. When conducted in this context, such research has the capacity to integrate *in-situ* records and provide a broader synoptic assessment of ecosystem state and evolution.

3.2 Abstract

In the ocean, satellite-based retrieval of biologically-informative variables, such as [*chl-a*], offers the opportunity to study evolution of biomass and ecosystem state on synoptic scales – provided that variables such as photoacclimation, depth-distribution of phytoplankton and carbon-to-chlorophyll ratios can be approximated as constant. However, regional variations in the optical properties (e.g. $a_{ph}(\lambda)$) of phytoplankton – often associated with changes in phytoplankton community structure, photoacclimation, depth or intracellular pigment concentration – are known to bias the retrieval of [*chl-a*], especially in high latitudes. Thus, there is potential to improve [*chl-a*] retrieval from satellite, if phytoplankton community structure can be inferred from remotely-sensed ocean-colour. Furthermore, variations in community composition are frequently accompanied by significant variety in physiology and ecological function. The retrieval of phytoplankton community structure from remotely-sensed ocean-colour is therefore an important goal in the study of ocean ecology. Retrieval of [*chl-a*] in boreal regions may also be biased by variable CDOM. Therefore, researchers aiming to retrieve derived variables, such as [*chl-a*], from remotely-sensed ocean-colour must develop regionally-tuned ocean-colour algorithms that accommodate both variation in community structure and incorporate knowledge of CDOM.

Here we describe the construction of a regionally-tuned algorithm to diagnose major phytoplankton assemblages in the Barents Sea. We use exploratory multivariate analyses

of the spectral shape of phytoplankton absorption coefficients ($a_{ph}(\lambda)$) to identify optically-distinct phytoplankton groups in natural seawater samples and elucidate their taxonomy. The *in-vivo* absorption spectra are then used to compile an empirical training dataset of ocean-colour observations from level 3 daily MODIS-Aqua data. This training dataset is used to train a Linear Discriminant Analysis in R to distinguish blooms of pluricellular aggregates, such as diatoms and *Phaeocystis*, from single-celled nanophytoplankton. We use Principal Component Analysis to identify coccolithophore blooms, and adapt an existing algorithm for the detection of *Phaeocystis*, by way of the detection of chlorophyll-c3, to discriminate *Phaeocystis* from diatoms.

3.3 Plain Language Summary

In biological oceanography, the term ‘bio-optical algorithms’ refers to a family of approaches that aim to interpret remotely-sensed ocean-colour to infer properties of the phytoplankton; for example the concentration of photosynthetic pigments such as chlorophyll-a, the average size of cells in a phytoplankton community or the presence of calcifying coccolithophores (these release calcite plates that imbue the ocean with a milky colouration). There has been a particular emphasis on the development of bio-optical algorithms to retrieve appraisals of phytoplankton community structure (or ‘Phytoplankton Functional Types’) in recent years, because this would provide researchers with increased insight into ocean ecological processes.

A dataset of phytoplankton absorption spectra in the Western Barents Sea was collected under Arctic PRIZE. Variation in phytoplankton absorption should influence remotely-sensed ocean-colour. Therefore, if variety in phytoplankton absorption corresponds meaningfully to different functional assemblages of phytoplankton, this dataset could be applied to develop a bio-optical algorithm in the Barents Sea to distinguish PFTs.

In this chapter, I demonstrate that variety in phytoplankton absorption corresponds well

with a small number of ship-board optical microscopy descriptions. Samples typified by abundant diatom cells, *Phaeocystis* colonies, dinoflagellates or a paucity of large phytoplankton cells (and therefore dominated by smaller cells; nanophytoplankton), can all be distinguished by their absorption spectra. I therefore divided phytoplankton absorption measurements into clusters that represented each of these groups, and compiled a record of coincident remotely-sensed ocean-colour observations. Variety in the ocean-colour observations was clearly useful for distinguishing some of these clusters; observations associated with diatoms and *Phaeocystis* colonies could be distinguished from observations associated with nanophytoplankton. However, it was difficult to distinguish ocean-colour observations associated with diatoms and *Phaeocystis* colonies from one another.

I used the original phytoplankton absorption measurements to perform simulations, to explore whether subtle differences in remotely-sensed ocean-colour should exist. I found that *Phaeocystis* colonies should exhibit slightly greater absorption near 469 nm. An algorithm had previously been developed in the literature to detect this spectral feature, but had not been applied to remotely-sensed ocean-colour. I discovered that, when careful corrections are made for drift in the output of remotely-sensed ocean-colour records, application of this algorithm successfully discriminates *Phaeocystis* colonies from diatoms. This chapter therefore concludes with the development of a bio-optical algorithm to infer dominant PFTs in the Barents Sea, that could permit a detailed study of change in phytoplankton community composition and ecology at a synoptic scale in this rapidly changing Arctic shelf.

3.4 Introduction

The Arctic Ocean is undergoing rapid climate change, having experienced an 11.5% decline in summer minimum sea-ice cover per decade between 1979 and 2012 (Stocker et al. 2013) and a 60% loss in the lateral extent of multi-year sea-ice compared with the 1980s (Kwok 2018). The Arctic Ocean may be completely transformed over the next decades even if the

goal of limiting global warming to within 2°C of pre-industrial temperatures is achieved, given the disproportional warming of the polar regions compared with lower latitudes (Overland et al. 2018). Some subregions of the Arctic are proving especially sensitive to climate change; the seas of the Eurasian Basin are experiencing an increased input of warmer, more saline Atlantic Water as the Global Ocean circulation responds to anthropogenic forcing (Francis and Hunter 2007). This process of ‘Atlantification’ is causing exceptionally rapid loss of sea-ice cover in the Barents Sea, the frontier between the Atlantic Waters of the Nordic Seas and the Arctic Waters of the Eurasian Basin (Cavalieri and Parkinson 2012). Ultimately, the increased input of Atlantic Water and decreasing sea-ice cover in the Arctic threaten to interact in a positive feedback that may undermine salinity-based water-column stratification, exposing warmer waters to the atmosphere and potentially precluding perennial sea-ice formation in the Barents Sea altogether (Polyakov et al. 2017).

In this context of rapid climatological and hydrographic change, evidence is emerging that the Arctic marine ecosystem is already changing. Studies in the Fram Strait region have demonstrated that the summer phytoplankton community composition has responded to an increasing inflow of Atlantic Waters, progressively shifting from dominance by diatom species towards haptophyte species, such as *Phaeocystis pouchetii* and *Emiliania huxleyi* (Nöthig et al. 2015).

Ocean-colour is a useful means to monitor ecosystem change on synoptic scales; remote-sensing studies have revealed that a poleward shift in double-bloom phytoplankton phenologies has emerged in response to decreased sea-ice cover. Lower sea-ice cover permits wind-induced vertical mixing to recharge the autumn ocean surface with nutrients. This, in turn, leads to a second autumnal bloom, following the main spring bloom– a paradigm more commonly associated with temperate waters (Ardyna et al. 2014). Average annual remotely-sensed [*chl-a*] in the region is increasing; Arrigo and Dijken (2015) have suggested that pelagic net primary production in the region has increased by 30% between 1998 and

2012). However, inferences of increased Arctic productivity in the satellite era may be difficult to reconcile with centennial-scale declines in sub-Arctic productivity indicated by biochemical proxies in Greenland ice cores (Osman et al. 2019).

Ocean-colour has also been used to study changes in phytoplankton assemblages with distinct optical properties, for example the detection of a statistically-significant Eastward shift in the occurrence of late-summer coccolithophore blooms in the Barents Sea (Smyth et al. 2004; Hovland et al. 2013; Oziel et al. 2017; Neukermans et al. 2018; Oziel et al. 2020). Coccolithophore blooms can be recognised easily in remotely-sensed ocean-colour because the strong scattering properties of their calcite liths impart a bright colouration to seawater, quite distinct from other phytoplankton groups (Hovland et al. 2013; Hovland et al. 2014). Distinguishing other phytoplankton groups from one another is more challenging. Retrieval of a greater range of phytoplankton assemblages from ocean-colour would provide the opportunity to monitor boreal ecosystems in greater detail and to account for taxonomic biases in $[chl-a]$ estimation (Sathyendranath et al. 2004).

Indeed, Dierssen (2010) found that estimates of $[chl-a]$ concentration from ocean-colour are accurate only within a factor of 5, owing to suspended minerals, Coloured Dissolved Organic Matter (CDOM) and taxonomic variation in the optical properties of phytoplankton. More recently, work by Sathyendranath et al. (2019) has shown that global $[chl-a]$ algorithms produced to interpret merged time series produced in the OC-CCI (Ocean-Colour Climate Change Initiative) are associated with root-mean-squared errors on the order of 0.3 $\log_{10}$ units of $[chl-a]$ (measured in mg m^{-3}). Jackson et al. (2017) has also published root-mean-squared errors of this magnitude; varying between 0.2 and 0.4 $\log_{10}$ units. Recent modelling approaches also predict that a fundamental change in the relationships underpinning empirical global algorithms may occur in the future (Dutkiewicz et al. 2019), such that the development of ocean-colour algorithms that incorporate taxonomic variations will have special value over the next decades.

We also invite the reader to consider that photoacclimation can cause the carbon-to-chlorophyll ratio of phytoplankton tissues to vary by a factor of 2-5 (Rodriguez et al. 2006). This means that, even if researchers can faithfully reconstruct $[chl-a]$ estimates from ocean-colour, that the inference of primary, new and export production represent additional and incomplete problems that require additional ancillary information to solve.

We have therefore established that there is a mandate in the wider literature for researchers to develop regional bio-optical algorithms to distinguish phytoplankton eco-physiological and functional variety.

3.5 Fundamentals of remote-sensing

As we have seen in the introduction section, numerous researchers have made insightful comments about ecosystem state and dynamics in Arctic seas by studying remotely-sensed ocean-colour. It is therefore worthwhile building a foundational understanding of the nature of this rich resource of data.

Ocean-colour apparent to remote-sensors is, effectively, the sum of light refracted by the atmosphere, reflected from the ocean surface, or modified by interactions within the surface ocean. It is hence clear that variation in ocean-colour will reflect the cumulative sum of a many different processes.

We can simplify our interrogation of ocean-colour by making a practical assumption. We will assume that the ratio of water-leaving radiance to down-welling irradiance (R_{rs}), is determined principally by elastic scattering (R^E) (the result of a flux of light penetrating and interacting with a field of particles – in this case water molecules and suspended constituents – in events that can be approximated as elastic collisions).

Elastic scattering is itself a wavelength-dependent (λ) function of absorption (a) and backscattering (b_b):

$$R^E(\lambda, 0) = r \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}, \quad (3.1)$$

where r is a coefficient of proportionality that varies with factors such as angle of incident light and the viewing angle (0 indicates the equation describes ocean-colour at sea-level).

The overall contribution of absorption by pure water, phytoplankton and CDOM to total absorption is additive;

$$a(\lambda) = a_w(\lambda) + a_{ph}(\lambda) + a_{cdom}(\lambda). \quad (3.2)$$

By assessing spectral shape – the wavelength dependent variation in R_{rs} , we can avoid the requirement to consider variation in r (assuming it does not vary with wavelength).

$a_w(\lambda)$ can be assumed to be constant. Furthermore, previous research has shown that, in the open waters of the Barents Sea, $a_{cdom}(\lambda)$ (absorption due to CDOM) is invariant across major hydrographic features and can therefore be treated as constant (Hancke et al. 2014), with an absorption coefficient of 0.04 m^{-1} at 400 nm, decaying according to the relationship found in Bricaud et al. (1981);

$$a_{cdom}(\lambda) = 0.04e^{(-0.0014(\lambda-440))}. \quad (3.3)$$

We have therefore greatly simplified our understanding of ocean-colour with a simple representation of the underlying physics. This simple representation has allowed us to conclude that many potential sources of variance in ocean-colour can be approximated as constant and therefore need not occupy our minds.

We can therefore assume, with good justification, that a and b_b in the Barents Sea are driven principally by variation in phytoplankton community composition and pigment biomass.

An illustrative schematic of these physical processes is presented in Figure 3.1.

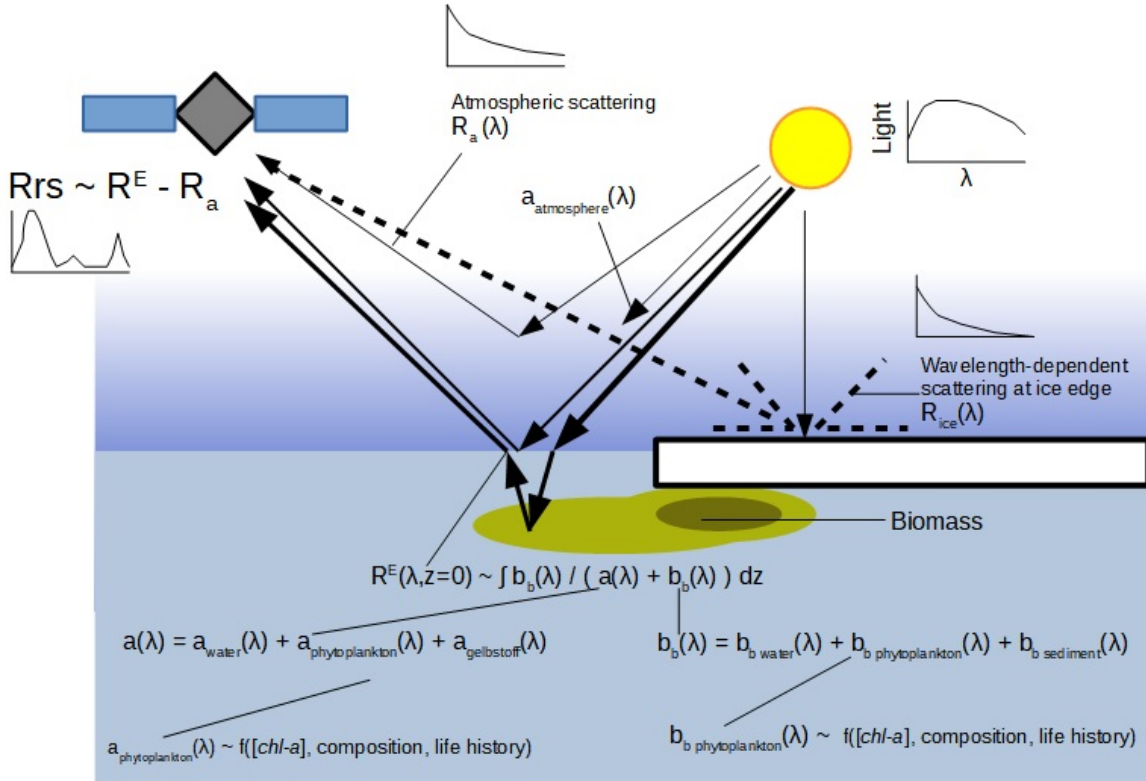


Figure 3.1: Schematic illustration of incident sunlight interacting with the atmosphere, ice-field, ocean-surface and sub-surface. Note that valid retrieval of R_{rs} relies on an accurate atmospheric correction and the assumption that ice-associated scattering can be neglected. Note that R_{rs} is a complicated wavelength-dependent function with a multivariate dependence on CDOM, phytoplankton concentrations, compositions and life-histories. ‘Gelbstoff’, German for ‘Yellow stuff’, is the combination of CDOM and non-algal particulate matter: $a_{gelbstoff} = a_{CDOM} + a_{NAP}$.

General reviews of the physics and interpretation of remotely-sensed ocean-colour can be found in Lee et al. (2006), and a special consideration of the unique challenges of polar regions is found in Babin et al. (2015), including a discussion of wavelength-dependent scattering from sea-ice (Bélanger et al. 2007).

It can be seen from Figure 3.1 that we can intuitively expect variation in the standing stock of phytoplankton biomass, pigment composition, depth-distribution or packaging of

pigments to result in changes to R_{rs} . Should individual taxon-indicative pigments have distinctive spectral signatures, this could result in specific wavelengths of R_{rs} varying in indicative ways, that could possibly be used to proxy the presence of different unique assemblages of phytoplankton. Where variation in community composition causes complex simultaneous variations in pigment packaging and complement, we might expect variation in R_{rs} to also be a complex function. Differences may be subtle, and variation in other factors, such as overall $[chl-a]$, may be superimposed upon them.

We can therefore see that, while we have been able to make some simplifying assumptions, and sketch a physical basis to justify an expectation that remotely-sensed ocean-colour should contain useful information about phytoplankton community composition, that extracting this information will require a detailed understanding of multivariate interactions and will not be easy.

Let us, then, review previous attempts in the literature to infer information about phytoplankton community composition from remotely-sensed ocean-colour; ‘taxonomic algorithms’.

3.6 Taxonomic algorithms

Various approaches have been applied to discern phytoplankton community structure from remotely-sensed ocean-colour.

Alvain et al. (2005) explored the empirical relationships between dominant phytoplankton groups (discerned by HPLC/ High Performance Liquid Chromatography) and coincident SeaWiFS measurements of ocean-colour in 5 wavebands, to develop the PHYSAT taxonomic algorithm. Alvain et al. (2005) computed the $[chl-a]$ concentration with the OC4v4 algorithm and then used an empirical relationship to extract the ocean-colour signal attributable to pigments other than $[chl-a]$, which are diagnostic for different phytoplankton

groups. The PHYSAT algorithm is informed by *in-situ* samples gathered largely in the low and mid-latitudes. Additionally, the PHYSAT algorithm was trained exclusively on pixels with a $[chl-a]$ concentration under 3 mg m^{-3} , to exclude waters influenced by sediment. These aspects of the PHYSAT algorithm make it unsuitable for applications in boreal seas, where $[chl-a]$ concentrations can exceed 10 mg m^{-3} as a matter of course (see Wassmann et al. 1990; Reigstad et al. 2002; Hancke et al. 2014). Application of the PHYSAT algorithm to the Global Ocean implies that the Barents Sea is dominated by nanoeukaryotes (Alvain et al. 2005, Figure 6 and Alvain et al. 2008, Figure 4) across all seasons.

Other authors (e.g., Zhang et al. (2018)) have proposed regionally-tuned taxonomic algorithms. They collected phytoplankton samples in the Bering and Chukchi Seas and performed HPLC analyses, identifying nine distinct phytoplankton assemblages. The authors determined the phytoplankton absorption spectra ($a_{ph}(\lambda)$) associated with each of these assemblages. They were then able to apply linear-unmixing models to $a_{ph}(\lambda)$ to estimate contributions from 5 optically-distinct phytoplankton assemblages to the overall $a_{ph}(\lambda)$. These approaches have, however, not yet been successfully extended to remotely-sensed ocean-colour.

Semi-empirical approaches, such as those of Sathyendranath et al. (2004) and Jackson et al. (2010), are examples of another strategy, in which phytoplankton community composition is inferred from HPLC data, then corresponding $a_{ph}(\lambda)$ representative of different communities are identified, and radiative transfer models are employed to establish remotely-sensed ocean-colour waveband ratios associated with different communities for a range of $[chl-a]$ concentrations. When remotely-sensed ocean-colour is used as input, at least two waveband ratios are employed to calculate $[chl-a]$ concentrations for different presumed community structures. The community that produces the most consistent result for multiple waveband ratios is then chosen as the most probable community composition. The approach offers the advantage of necessarily producing a taxonomically-informed estimate

of $[chl-a]$ concentration.

Devred et al. (2011) developed a system of equations relating $[chl-a]$ associated with three different phytoplankton size-classes in the North-west Atlantic, to the total absorption coefficient at several wavelengths. The size-classes were derived from HPLC analysis. The model could then be used to solve for the $[chl-a]$ associated with each size-class from remotely-sensed ocean-colour data at multiple wavelengths. This method offers the advantage of describing community composition according to the $[chl-a]$ concentration associated with each component of the community, and gives $[chl-a]$ concentration for each size-class, rather than just identifying the dominant class. However, size-classes are not necessarily synonymous with taxonomic assemblages of phytoplankton.

Clearly, there are many strategies available to retrieve information on phytoplankton community structure from ocean-colour data, but all of the methods discussed above rely on labour-intensive HPLC analysis in order to distinguish different taxonomic assemblages or size-classes of phytoplankton in natural samples. These assemblages are then effectively treated as targets, for which distinct bio-optical properties or ocean-colour signals must be identified. This approach requires the freezing and shipping of phytoplankton samples, which represents a source of delay that encumbers the ability of researchers to develop bio-optical algorithms rapidly. Furthermore, chapter 2 demonstrated that defining target functional assemblages directly from bio-optical properties may eliminate some sources of error that confound the development of bio-optical algorithms.

Thus, in this chapter, we discuss the construction of a remote-sensing algorithm for community structure in the Barents Sea, which did not require HPLC analysis of new samples in its development (samples for HPLC analysis had been collected for the Arctic **PRIZE** phytoplankton dataset, but had not reached the laboratory at the time of writing). We employ methods similar to those used by Hoepffner and Sathyendranath (1991), Chase et al. (2013), and Liu et al. (2019) to infer the presence of taxonomically-informative pigments

from $a_{ph}(\lambda)$, which can be measured using benchtop or profiling spectrophotometers, and show that they are sufficient to identify major phytoplankton assemblages. We compare this means of categorisation of assemblages with an existing data set from the Labrador Sea (collected by and previously discussed in Stuart et al. (2000), Cota et al. (2003), and Sathyendranath et al. (2004)) and demonstrate that optical properties have a similar utility in the recognition of broad taxonomic assemblages, such as ‘diatom-dominated’ or ‘*Phaeocystis*-dominated’, to HPLC analysis. We use a dataset of 191 phytoplankton samples from the Barents Sea to demonstrate that empirical, regionally-tuned algorithms for the remote-sensing of phytoplankton community composition can be constructed rapidly using this method. We assess the limitations of solely empirically-informed algorithms and obviate deficiencies with the use of semi-empirical radiative transfer modelling, as in Sathyendranath et al. (2004) and Jackson et al. (2010).

3.7 Collection of phytoplankton data

Let us begin our efforts to develop a ‘taxonomic’, or ‘bio-optical’, algorithm for the Barents Sea by identifying gross community structure in the Barents Sea phytoplankton samples collected under Arctic **PRIZE**.

We collected natural assemblages of phytoplankton from the Western Barents Sea on 3 research cruises aboard British and Norwegian vessels, as part of the Arctic **PRIZE** (**P**roductivity in the Seasonal **I**ce **Z**one) research project, with the aim of capturing seasonal variation in the ecosystem of the Seasonal Ice Zone. The geographic distribution of the sampling sites is presented in Figure 3.2.

The sampling locations are concentrated along the 30th Eastern Meridian, such that they form a transect across the Polar Front, which divides warmer Atlantic-influenced water-masses from Arctic Waters of a North-easterly origin in the Eurasian Basin (Reigstad et

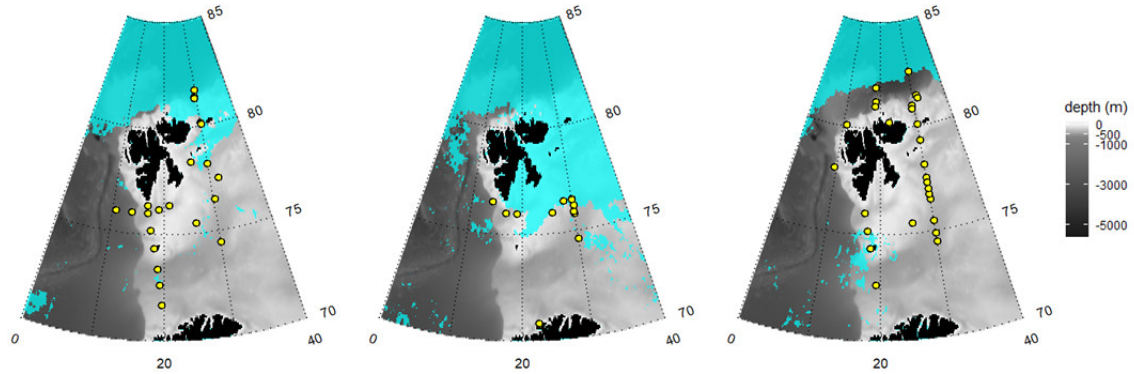


Figure 3.2: (Left) Cruise JR16006 Julian days 189-217. (Centre) Cruise HH23042018 Julian days 114-122. (Right) Cruise JR17006 Julian days 163-185. Regions with no MODIS-Aqua satellite observations during the cruise window are covered by a cyan mask, which indicates the extent of the ice edge. Yellow markers indicate sampling sites.

al. 2002).

A Conductivity-Temperature-Depth (CTD) rosette, equipped with an *in-vivo* Chelsea Aqua 3 fluorometer, was used to collect seawater samples. Real-time readings of $[chl-a]$ fluorescence were used to select sampling depths, typically targeting the surface, Sub-surface Chlorophyll Maximum (SCM) and 1-3 intermediary depths. Seawater, collected using the Niskin bottles, was subsampled for analysis of bio-optical properties ($a_{ph}(\lambda)$) and $[chl-a]$.

We made initial ship-board optical microscopic observations on a subset of the samples we collected, so that we could gain an overview of the salient variation in phytoplankton community composition that we had encountered. These microscopic observations simply recorded the presence or absence of major phytoplankton groups, such as diatoms, dinoflagellates or *Phaeocystis* colonies. Where we recorded no microphytoplankton we concluded the sample must represent a nanophytoplankton-dominated community.

Samples were filtered through 25 mm GF/F filters (0.7 μm pore diameter) to collect phytoplankton cells. We collected triplicates of 0.2 L samples for fluorometric $[chl-a]$ measurement.

We immersed $[chl-a]$ samples in 90% acetone solution and stored them in the dark at -20°C overnight to facilitate pigment extraction. A Turner Designs Trilogy fluorometer was used to measure $[chl-a]$. Typical sample volumes for $a_{ph}(\lambda)$ ranged between 0.25 and 1 L; we aimed to filter a sufficient volume such that colour was visible to the naked eye. Filters for phytoplankton $a_{ph}(\lambda)$ were flash-frozen in liquid nitrogen and transferred to a -80°C freezer for long term storage. Later, in the laboratory, $a_{ph}(\lambda)$ was measured by a Shimadzu UV-Vis spectrophotometer equipped with an integrating sphere at a 1 nm resolution, following the method of Kishino et al. (1985), including the extraction of pigments with methanol to determine absorption by detrital matter. The filter pads were rotated by a quarter revolution 4 times, collecting replicate measurements that were subsequently averaged, to control for uneven distribution of cells on the disk of the filter pads. We used the binomial pathlength correction equation of Hoepffner and Sathyendranath (1992) to convert the optical densities into absorption coefficients;

$$a_T(\lambda) = \frac{0.057575\pi d^2}{V} (0.31X_f(\lambda) + 0.57X_f^2(\lambda)), \quad (3.4)$$

where a_T is total particulate absorption measured in m^{-1} , d is the diameter of the filtration area measured in cm, X_f is the absorbance measured from the filter-pad by a Shimadzu spectrophotometer, V is the volume of seawater that was filtered through that filter-pad measured in Litres (L), and λ specifies wavelength (rearranged from equations 2 and 3 in Hoepffner and Sathyendranath (1992)). We chose to match our method to Stuart et al. (2000) so that we could pool our Barents dataset with a similar dataset of $a_{ph}(\lambda)$ collected in the Labrador Sea (Stuart et al. 2000; Cota et al. 2003; Sathyendranath et al. 2004), in which diatom and *Phaeocystis*-dominated blooms were sampled – similar to the compositions known to be frequent in the Barents Sea (Degerlund and Eilertsen 2010). The Labrador Sea dataset was collected between 1996 and 2001 in multiple seasons.

Comparing $a_{ph}(\lambda)$ in the Barents and Labrador Seas therefore allowed us to gain greater confidence that any recognition of distinct bio-optical assemblages common to both Seas is robust to seasonal and inter-annual variation.

We fitted exponential curves to the detrital absorption spectra and subtracted them from the total absorption spectra to attain the absorption spectra of the phytoplankton;

$$a_{ph}(\lambda) = a_T(\lambda) - a_d(\lambda), \quad (3.5)$$

where $a_{ph}(\lambda)$, $a_T(\lambda)$ and $a_d(\lambda)$ denote phytoplanktic, total and detrital absorption, respectively, at wavelength λ . In total, 191 spectra are presented in this study.

Variability in the absorption spectrum of phytoplankton samples, and any relationship it might have to our ship-board optical microscopy, will provide us with an impression of bio-optically distinct phytoplankton communities in the Barents Sea that we might feasibly be able to identify in remotely-sensed ocean-colour observations.

3.8 Identification of distinct bio-optical phytoplankton communities in *in-situ* samples

We have gathered a resource of hyperspectral $a_{ph}(\lambda)$ data for phytoplankton communities in the Barents Sea, but how do we begin exploring this dataset for structure that could reflect salient variation in phytoplankton community composition?

Measurements of $a_{ph}(\lambda)$ have many dimensions. We therefore sought a means to reduce the dimensionality of our Barents Sea dataset before analysis, so that any patterns of variation that exist might prove more tractable.

Methods for decomposition of $a_{ph}(\lambda)$ were developed by Hoepffner and Sathyendranath

(1991), Chase et al. (2013), and Liu et al. (2019). We make use of the version presented in Chase et al. (2013), which takes total absorption readings from an AC9 spectrophotometer, removes absorption attributable to detritus to isolate $a_{ph}(\lambda)$ and then breaks this down into component Gaussian functions that correspond to the concentrations of different combinations of taxonomically-indicative phytoplankton pigments. The Gaussian function decomposition method has been used successfully by Liu et al. (2019) to develop algorithms that describe the contribution from taxonomically-indicative pigments to $a_{ph}(\lambda)$ measured in the Fram Strait. Thus, there is a precedent for this method’s successful application in the European Arctic.

Alternative dimensionality reduction approaches include stepwise-discriminant analysis, which was employed by Johnsen et al. (1994) to discover 5 taxonomically-sensitive wavelengths of $a_{ph}(\lambda)$ that were shown to discriminate 10 different phytoplankton assemblages.

Having performed Gaussian decomposition on the Barents Sea $a_{ph}(\lambda)$ dataset, we also decomposed the $a_{ph}(\lambda)$ dataset, gathered by Stuart et al. (2000), so that we had a resource of data for comparison.

We chose this Labrador Sea $a_{ph}(\lambda)$ dataset, to contrast with our Barents Sea $a_{ph}(\lambda)$ dataset, because the climatology and phytoplankton community composition is similar to the Barents Sea (a high latitude sea experiencing perennial ice cover) and because its phytoplankton community is known to comprise a similar group of assemblages that follow similar phenologies and probably have similar adaptations to life in the boreal seas as Barents Sea phytoplankton (Wassmann 1994; Wassmann et al. 1999; Cota et al. 2003; Carmack and Wassmann 2006).

The Labrador Sea phytoplankton samples had already been prescribed likely identities as diatom or nanophytoplankton-dominated, based on HPLC (Stuart et al. 2000), but we added a third category, classifying those blooms in which literature reports noted strong DMS (dimethyl-sulphide) odours or gelatinous mucilage as *Phaeocystis*, which commonly

occur off the Greenland Shelf (Cota et al. 2003; Fragoso et al. 2017). This parallel dataset of $a_{ph}(\lambda)$ will, in conjunction with our ship-board optical microscopic observations, help us explore how variation in $a_{ph}(\lambda)$ maps onto variation in phytoplankton community composition.

Gaussian decomposition was performed by a translation of Chase et al.’s decomposition script, which made use of the *lsqnonneg()* function of the *pracma* package (Borchers 2018). The script was modified to remove the step that accounts for detrital absorption, as this had already been achieved by subtraction of detrital absorption spectra.

Original Barents Sea $a_{ph}(\lambda)$, and the resulting decomposition into component Gaussian functions, are presented in Figure 3.3.

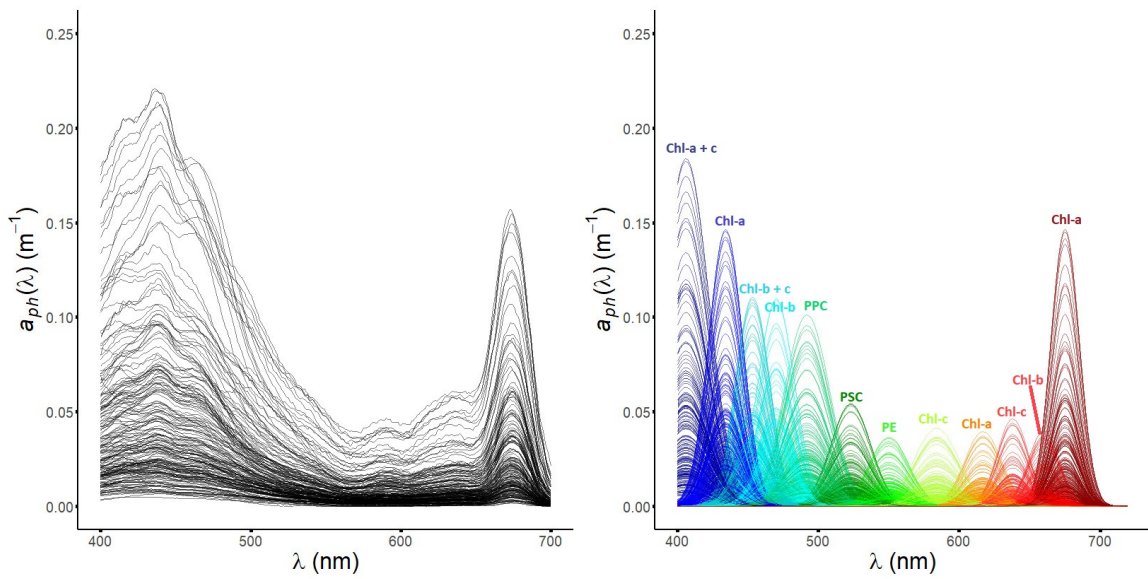


Figure 3.3: (Left) Original hyperspectral absorption spectra for 191 Barents Sea phytoplankton samples. (Right) Component Gaussian functions, coloured to foster clarity and labelled with the pigments they indicate. ‘PE’, which has not been introduced yet, stands for ‘phycoerythrin’.

We established a strong relationship between taxon-indicative pigments and the magnitude of Gaussian components, by performing a Two-Block Partial Least Squares analysis

(2BPLS) (see Rohlf and Corti (2000) for details) on the Labrador Sea dataset, the results of which are presented in Figure 3.4. A discussion of 2BPLS and an illustrative example is provided in the mathematical methods appendix. The analysis is a multivariate exploratory method, initially adopted in sociology (e.g. Sampson et al. (1989)), but now widely used in comparative anatomy and increasingly employed in ecological studies, (e.g. Contreras et al. (2017)). The approach has also been applied to study covariance between $a_{ph}(\lambda)$ and pigment complement across phytoplankton samples and cultures (see Roy et al. (2011), pages 351-352 for a review). This is hence an established approach for the comparison of spectral and pigment data in biological oceanography.

2BPLS is analogous to PCA (also see our mathematical methods appendix), but rather than finding a system of orthogonal linear combinations of variables that explain variance in a single block of data, it finds a system of orthogonal pairs of linear combinations of variables that explain covariance between two or more blocks of data (Rohlf and Corti 2000). The weightings of linear variable pair combinations are referred to as ‘salience’ and each battery of saliences makes up a ‘latent variable’, analogous to a PC axis.

We prepared the left block of the analysis by normalising HPLC pigment concentrations to the sum of all pigments (accessory + $[chl-a]$) in each sample and performed PCA, truncating the block of PC-scores when the axial explained variance dropped below 5%.

We prepared the right block of the analysis by normalising the magnitudes of the component Gaussian functions (which we shall refer to as ‘CGF-scores’) to the sum of all CGF-scores in each sample and performed PCA, truncating the block of PC-scores when the axial explained variance dropped below 5%. We then performed 2BPLS, and randomly re-sampled the input blocks, with replacement, 999 times to compute a p -value for assessing significance of the first latent variable.

Finally, we computed the effect size, as described in Adams and Collyer (2016).

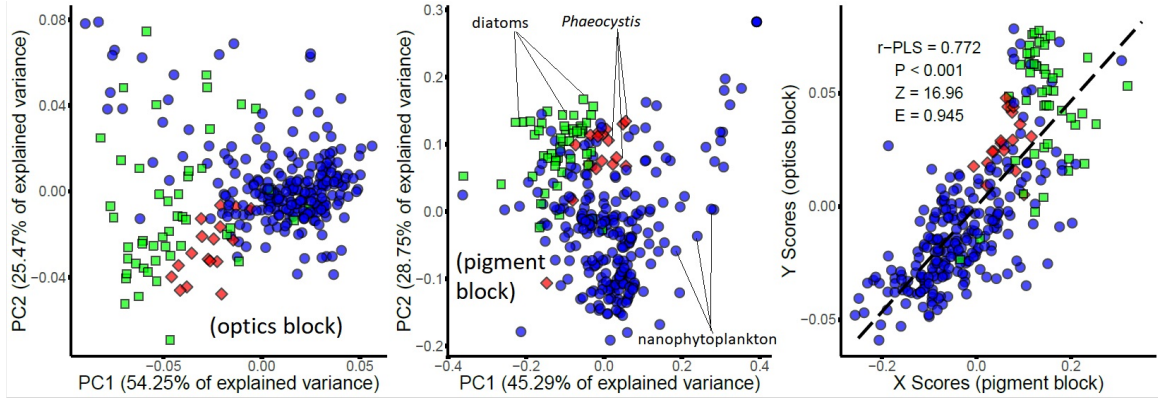


Figure 3.4: **Left:** Biplot of the leading PC axes in the optical properties of the Labrador Sea samples. **Centre:** Biplot of the leading PC axes in the pigment composition of the Labrador Sea samples. **Right:** The leading latent variable retrieved by 2BPLS. Green squares (diatoms), red diamonds (*Phaeocystis*), blue circles (nanophytoplankton).

We retrieved a maximum correlation coefficient between input blocks of 0.772, with significance testing indicating a p -value < 0.001 , showing that this relationship is highly significant. The retrieved effect size indicated that the relationship was 16.96 standard deviations above the association that would be expected from a randomly permuted dataset with no real association between input blocks.

The first latent variable explained 94.49% of covariance between the input blocks. The stratification of distinct Labrador Sea phytoplankton communities along the leading saliences in this analysis confirm a strong relationship between pigments and Gaussian decompositions, and establish that they offer a credible means by which to distinguish prevailing phytoplankton communities. We can therefore proceed to explore their variance structure with confidence.

Given that $a_{ph}(\lambda)$ is determined not only by pigment complement, but also by the packaging of those pigments and the morphology of phytoplankton cells (eg. Kirk 1994; Cota et al. 2003), we should expect that these Gaussian components are proxies not only of pigment complement, but also of cell morphology. Let us explore the variance structure in our

Barents Sea $a_{ph}(\lambda)$ dataset by performing Principal Component Analysis on the matrix of Gaussian decomposition magnitudes (peak heights).

We found, upon normalising each of the Gaussian peaks to the fourth Gaussian peak (which represents the *in-vivo* absorption by *chl-b*) and then applying PCA (Principal Component Analysis; see the mathematical methods appendix section on PCA for an involved discussion) to the absorption data from JR16006, that scores on the second Principal Component have a bimodal distribution, as seen in Figure 3.5.

We projected the remaining Barents and Labrador Sea $a_{ph}(\lambda)$ data along these axes and found that different taxonomic groups of samples, identified by ship-board optical microscopy or identified in (Stuart et al. 2000; Cota et al. 2003; Sathyendranath et al. 2004), tended to cluster in different regions of the space.

Diatoms tend to possess high PC1 scores and low PC2 scores, for example, while dinoflagellates possess both high PC1 and PC2 scores. This suggests that clustering in the $a_{ph}(\lambda)$ of the Barents Sea phytoplankton represents variability in community composition.

Only a subset of our Barents Sea $a_{ph}(\lambda)$ samples possessed corresponding ship-board microscopy to record their phytoplankton community composition. However, we can see that the clustered distribution of $a_{ph}(\lambda)$ corresponds well to salient variation in phytoplankton community composition. Hence, fuzzy-clustering was performed using the R package *ppclust*'s *fcm()* function (Cebeci et al. 2019) (a description of fuzzy-clustering and a worked example is provided in the mathematical methods appendix). We normalised the Gaussian peak heights to the sum of all such peak heights for each sample, to control for variance in magnitude, set fuzziness to 1.8 and elected to solve for 4 clusters. We chose to solve for 4 clusters on the basis that there are 3 Labrador Sea clusters with clear Barents Sea analogues, and our ship-board optical-microscopy indicated that dinoflagellate-dominated samples likely formed a fourth cluster. We have therefore crisply partitioned our Barents Sea $a_{ph}(\lambda)$ dataset into four clusters. We have a justified confidence that these clusters map

onto salient variation in phytoplankton community composition and that they are dominated by nanophytoplankton, diatoms, colonial *Phaeocystis* and dinoflagellates.

We could have used a variety of different clustering techniques to partition our data, but we chose fuzzy clustering because it permits the calculation of partial memberships. We used these partial memberships as weights when we calculated the mean $a_{ph}(\lambda)$ for each of the 4 clusters. The mean $a_{ph}(\lambda)$ for the i^{th} cluster was calculated as a weighted average of the memberships of all samples most closely aligned to the i^{th} cluster. This is the means by which representative spectra presented in Figure 3.5 were produced.

A comparison of the Barents Sea spectra with the mean spectra for the Labrador Sea phytoplankton clusters shows a high degree of similarity, with the exception of the fourth cluster, that our ship-board microscopy indicated was rich in dinoflagellates.

The complete inventory of $a_{ph}(\lambda)$ for each cluster, normalised to $[chl-a]$, is presented in Figure 3.6.

It can be seen in Figure 3.5 that the *Phaeocystis* cluster is optically distinct from nanophytoplankton – exhibiting a much greater degree of pigment packaging and having a greater resemblance to the diatoms; a shoulder in $a_{ph}(\lambda)$ around 469 nm distinguishes the *Phaeocystis* and diatom clusters. This feature probably results from absorption due to chlorophyll-c3 (Astoreca et al. 2008). There is a possible ancillary absorption peak around 585 nm (see Figure 3.6 diatom and *Phaeocystis* subplots), which is also indicative of chlorophyll-c3 (Johnsen et al. 1994), but this does not align with any spectral waveband commonly retrieved by remote-sensors, so we do not place considerable emphasis on this feature (as our ultimate ambition is to relate variety in $a_{ph}(\lambda)$ to remotely-sensed ocean-colour).

Chlorophyll-c3 was treated by Astoreca et al. (2008) as a characteristic marker pigment indicating the presence of *Phaeocystis globosa* in the Southern North Sea. It must be noted

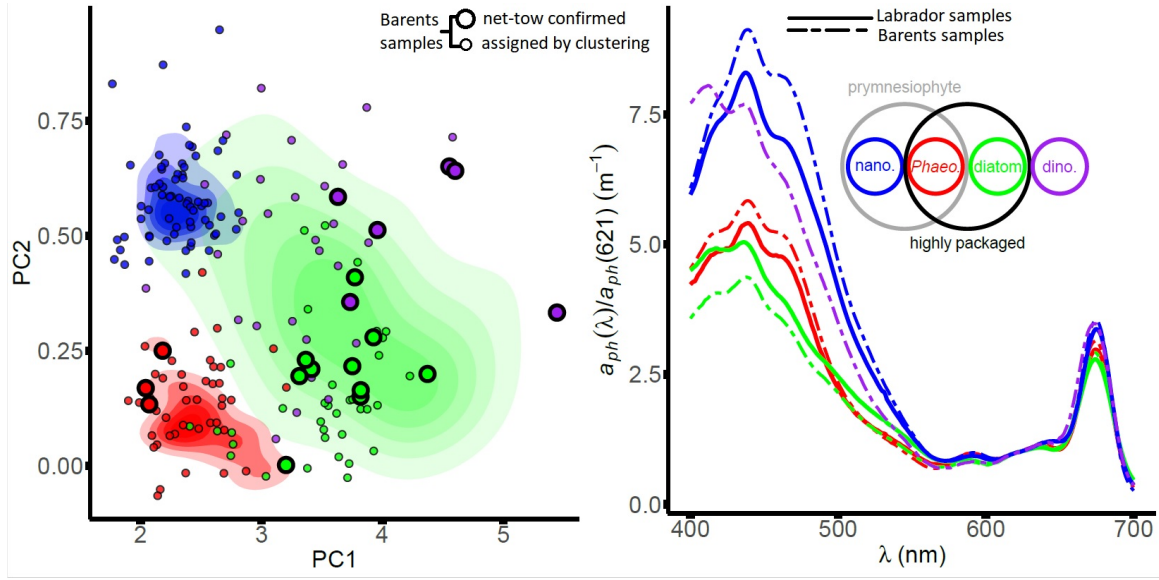


Figure 3.5: **Left:** A PCA space representing variation in the optical properties of the Barents phytoplankton samples. The density gradients represent the regions of this space occupied by nanophytoplankton, diatom and *Phaeocystis* samples identified in the Labrador Sea (blue, green, red). The bold points represent a sub-set of Barents Sea phytoplankton which optical microscopy showed to be dominated exclusively by *Phaeocystis*, diatoms or dinoflagellates (red, green, purple), while the smaller points represent the rest of the Barents phytoplankton samples after a fuzzy clustering algorithm was instructed to divide them into four groups. **Right:** Representative average absorption spectra for nanophytoplankton, dinoflagellates, *Phaeocystis* and diatoms (blue, purple, green, red) for the Labrador (bold) and Barents (dashed) Sea datasets, showing high similarity.

that chlorophyll-c3 is not exclusive to this genus or even more broadly to prymnesiophytes (e.g. in *Emiliania huxleyi* (Garrido and Zapata 1998)); its occurrence has been noted by Roy et al. (2011) in the heterokont picophytoplankton class bolidophyceae, silicoflagellates and their sisters the pelagophytes, as well as some diatoms and dinoflagellates. Indeed, Johnsen and Sakshaug (2000) describe chlorophyll-c3 as indicative of prymnesiophytes (including *Chrysochromulina* and *Phaeocystis*) and dinoflagellates in the Norwegian Sea. Nevertheless, the pronounced shoulder in *Phaeocystis* $a_{ph}(\lambda)$ at 469 nm, indicates that chlorophyll-c3 functions as an effective marker pigment of colonial *Phaeocystis* dominance in the Arctic PRIZE Barents Sea $a_{ph}(\lambda)$ dataset 3.5). The astute reader should, however, note that the

perception of chlorophyll-c3 should not be taken as unambiguous evidence of any specific genus of phytoplankton without ancillary analyses or justification.

A review of the PC loadings of the axes, that define the space depicted in Figure 3.5, provides insight into the variation in optical ecology that distinguishes clusters (Table 3.1).

Table 3.1: *PC loadings*

Axis	<i>Chla + c</i>	<i>Chla</i>	<i>Chlb + c</i>	<i>Chlb</i>	<i>PPC</i>	<i>PSC</i>	<i>PE</i>	<i>Chlc</i>	<i>Chla</i>	<i>Chlc</i>	<i>Chlb</i>	<i>Chla</i>
PC1	0.876	0.189	0.233	0	0.176	0.151	0.134	0.067	0.080	0.056	0.069	0.229
PC2	-0.069	0.181	0.163	0	0.687	0.225	0.0296	-0.204	-0.186	-0.171	-0.157	-0.532

Note. All optical characters normalised to *Chlb*, all values rounded to 3 significant figures. *PPC* stands for photoprotective carotenoids, *PSC* for photosynthetic carotenoids and *PE* for Phycoerythrin.

Table 3.1 shows that all of the loadings for PC1 are positive. Different CGF-scores that serve as proxies for the same pigments, such as [*chl-a*], have different loadings, and there is a trend for those CGF-scores that tend to have higher magnitudes in Figure 3.3 to have higher loadings as well, suggesting that PC1 may be a proxy for the degree of pigment packaging in phytoplankton samples. This interpretation is compatible with the observation that dinoflagellate and diatom-replete samples, which are more highly packaged, are discriminated from nanophytoplanktic prymnesiophytes along this axis (Figure 3.5).

The loadings for PC2 vary between negative and positive. The dominating pattern is a large positive loading in CGF-scores corresponding to PPC (Photoprotective Carotenoids) and a negative loading in [*chl-a*]. This suggests that PC2 is discriminating organisms with optical adaptations to high-intensity light, an interpretation that could be consistent with the separation of nanophytoplankton and dinoflagellates, which both tend to inhabit stratified waters and hence are more exposed to sunlight (Margalef 1978), from diatoms and *Phaeocystis*, which occur in concentrated blooms that provide sufficient self-shading to render investment in copious amounts of PPC unnecessary. Indeed, we refer the reader to the ‘Results’ section of chapter 2, where the pigment complements associated with these clusters are reported and can be observed to generally be commensurate with ecological

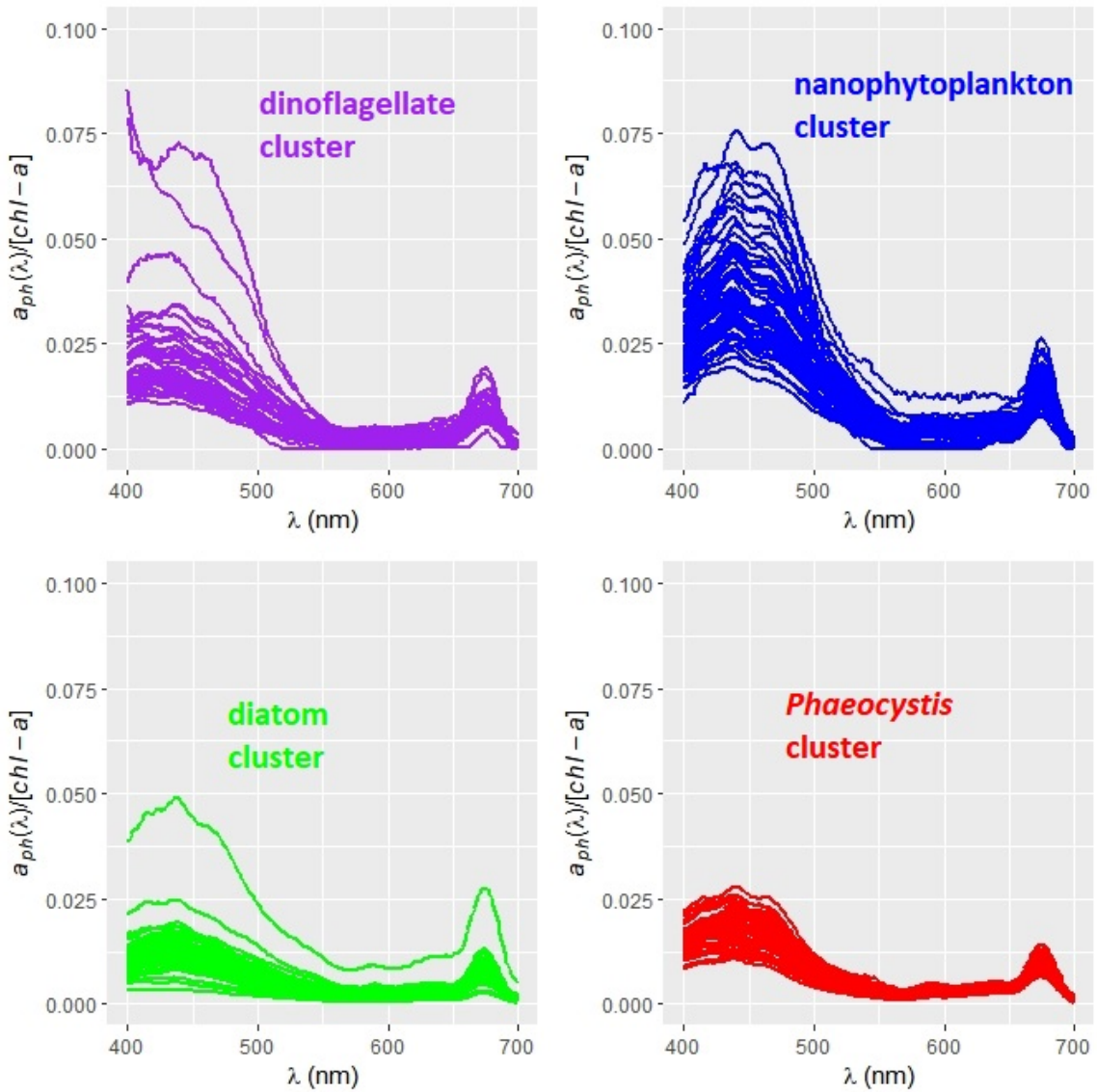


Figure 3.6: Hyperspectral absorption spectra for samples belonging to each of the four clusters identified in the Arctic **PRIZE** $a_{ph}(\lambda)$ dataset. Note that diatoms, *Phaeocystis* and dinoflagellates all exhibit greater pigment packaging than nanophytoplankton. A shoulder in absorption at 469 nm suggests high relative concentrations of chlorophyll-c3 in the *Phaeocystis* cluster. There is potentially a greater variety in spectral shapes in dinoflagellates, consistent with their greater dispersion in the ordination plot in Figure 3.5.

interpretations based on $a_{ph}(\lambda)$.

3.8.1 Ocean-colour data

Our ambition to construct an algorithm, that can predict bio-optical variation in the phytoplankton from ocean-colour, requires a resource of ocean-colour observations.

We selected the MODIS-Aqua platform, owing to its provision of data at a decadal scale with a relatively high spectral resolution. Other remote-sensing platforms concurrent with our sampling have either shorter flight histories or fewer available spectral bands in the visible range. For example, the Visible Infrared Imaging Radiometer Suite (VIIRS) (OBPG 2020) has a 2011 launch-date and 5 spectral bands, while the Ocean-Colour-CCI dataset, which merges data from MERIS (MEdium Resolution Imaging Spectrometer), MODIS-Aqua, SeaWiFS (Sea-View Wide Field-of-View Sensor) and VIIRS, has 6 spectral bands and begins in 1998 (Sathyendranath et al. 2012). By comparison, MODIS-Aqua has a launch date of 2002 and 10 spectral bands.

We decided that the extra spectral resolution of the MODIS-Aqua sensor outweighed the importance of the 4 extra years available in the OC-CCI dataset, especially since MODIS-Aqua offers bands at blue wavelengths (443, 469, 488 nm) that should be useful for distinguishing *Phaeocystis* (Astoreca et al. 2008), a particularly important and optically distinct phytoplankton assemblage in the Barents Sea (Reigstad et al. 2002; Rat’kova and Wassmann 2002; Sakshaug et al. 2009).

The preferential absorption of chlorophyll-c3 at 469 nm (Johnsen et al. 1994), should mean that prymnesiophyte species rich in this pigment (such as the *Phaeocystis* cluster we have identified) should have a decreased reflectance/ R_{rs} value at 469 nm relative to adjacent wavebands.

We downloaded Level-3 daily binned ocean-colour observations spanning the duration of our research cruises (Table 7.1) from the MODIS-Aqua website (OBPG 2019). These data represent ocean-colour observations centred at wavebands $\lambda = 412, 443, 469, 488, 531, 547, 555, 645, 667$ and 678 nm, within 1° latitude or longitude of *in-situ* samples, on the

same Julian day as sampling. The atmospheric correction and quality control of these data belongs to the r2018.1 reprocessing. We collected 18,000 ocean-colour observations in total.

It should be noted that the pigment-specific $a_{ph}(\lambda)$ of some pigment species is a function of phytoplankton photoacclimation state and the way in which pigments are chemically bound (Johnsen and Sakshaug 2007) to the thylakoid membrane – a surface within the chloroplast of photoautotrophic cells upon which the light-dependent reactions of photosynthesis are mediated (Kirk 1994).

Hence, the ratios of reflected light across different ocean-colour wavebands may be influenced by the life-history of phytoplankton communities, in addition to their pigment complement, community composition, depth-distribution and overall concentration in the water-column.

3.8.2 Masking unsuitable ocean-colour observations

In our assessment, of the fundamental physics underlying remote-sensing, we made the simplifying assumption that change in $a(\lambda)$ is driven by change in $a_{ph}(\lambda)$. We must also neglect consideration of variance in $b_b(\lambda)$ associated with phytoplankton community composition or life history, even though this may be significant (Pavlov et al. 2017). This is a simplifying assumption we are forced to make, because a comprehensive set of phytoplankton backscattering properties was not collected under Arctic **PRIZE**.

Our aim, to construct an algorithm to relate variation in remotely-sensed ocean-colour to bio-optical variation in our *in-situ* samples, could therefore be undermined by sources of variance in remotely-sensed ocean-colour that are unrelated to the absorption of phytoplankton.

Violations of our assumptions may occur when non-algal particles, which have their own wavelength-dependent absorption (Babin et al. 2003) and backscattering properties (Mes-

lard et al. 2018), are present in substantial concentrations. River plumes can also carry terrigenous organic matter, increasing $a_{cdom}(\lambda)$ (Drozdova et al. 2017), and coccolithophore blooms will violate our assumptions too, because of the influence of backscattering of light off of their biomineral liths (Hovland et al. 2013).

We therefore need a means by which to identify ocean-colour observations in which our assumption that variation in spectral shape is driven by phytoplankton absorption is violated, so that we can mask them from consideration.

We performed a Principal Component Analysis (PCA) to explore variation in ocean-colour R_{rs} measurements taken from the Barents Sea between 0 and 80° East and 0 and 85° North, in July 2014. See the mathematical methods appendix on PCA for an involved discussion of this approach.

We selected this region and time because coccolithophore blooms have been reported (Carlowicz et al. 2014). The geographic range of our sample also included the Ob River plume (Drozdova et al. 2017) and sediment-dominated fjords of the Svalbard archipelago (Meslard et al. 2018). The sampled region therefore includes all of the sources of variance that are causing us concern.

We identified the geographic regions corresponding to coccolithophore blooms, river plumes and coastal sediments and devised a scheme of linear dividers that distinguish these R_{rs} measurements from those of the open-ocean, where our simplifying assumptions can reasonably be expected to hold true (Figure 3.7).

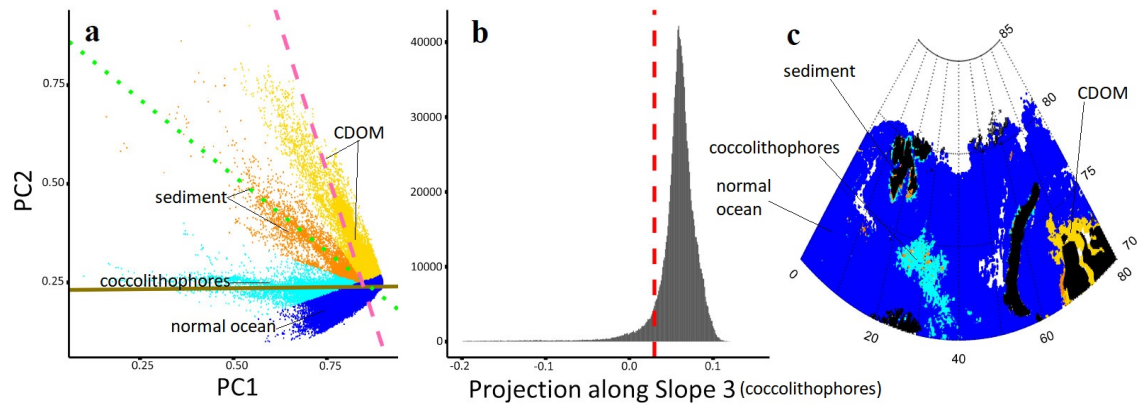


Figure 3.7: a) A Principal Component Analysis of ocean-colour in the Barents Sea, July 2014; blue, yellow, orange and cyan points correspond to normal ocean, river plumes, sediment and coccolithophores. The pink, dashed; green, dotted and brown solid lines correspond to the linear trends associated with river plumes, sediment and coccolithophores; linear dividers were defined to bisect the angles between them. b) The red, dashed line represents the division between open-ocean ocean-colour and those masked as river plumes, sediment or coccolithophores. c) A map of the classifications, showing coherent geographic domains.

We used the linear dividers we had contrived to distinguish these optically distinct clines related to coccolithophores, river plumes and fjordic sediment to develop masking criteria that we could employ to dismiss ocean-colour observations unfit for analysis. The capacity for our masking criteria to identify coccolithophores is demonstrated in Figure 3.8, in which several coccolithophore blooms identified by the NASA Earth Observatory (GSFC 2003; Hansen et al. 2016; Carlowicz et al. 2018) are correctly recognised.

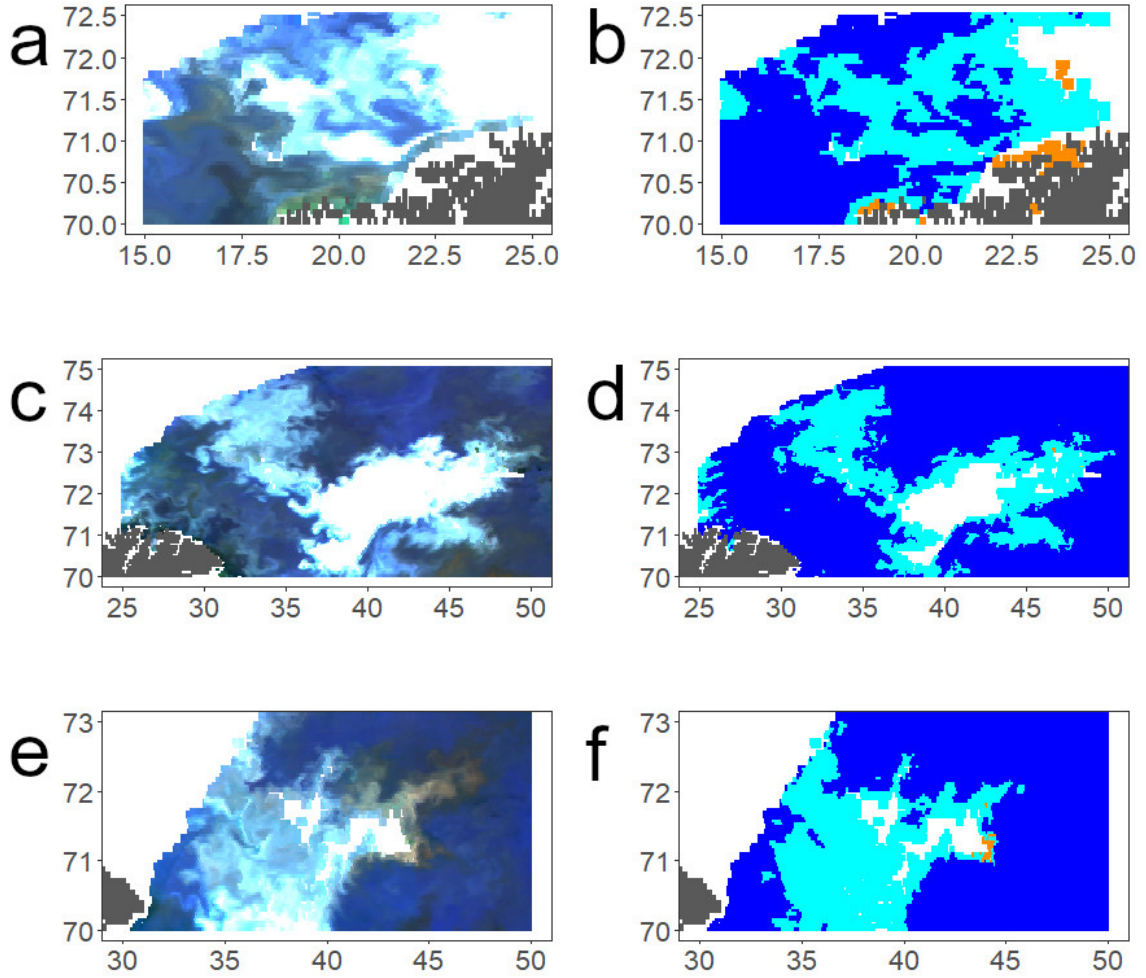


Figure 3.8: Examples of true colour images of coccolithophore blooms on the 24th of July 2003, 6th of July 2016 and 18th of July 2018 (**a,c,e**) and corresponding coccolithophore masks (**b,d,f**). Normal ocean is masked blue, coccolithophores are masked cyan and sediment is masked orange.

3.9 Relation of variability in ocean-colour to *in-situ* bio-optical communities

Now that we have established clear bio-optical variation exists in the Barents Sea $a_{ph}(\lambda)$ samples, and that it conforms to meaningful variation in community structure, we are

motivated to ask whether variance in the absorptive properties of the phytoplankton is recapitulated in the ocean surface reflectance observations in the MODIS-Aqua ocean-colour dataset we have assembled.

Given that we have labelled distinct bio-optical clusters in the *in-situ* samples, we can relate them to corresponding ocean-colour measurements. We can follow a similar approach to Kurekin et al. (2014), by using subsets of our total dataset to train Linear Discriminant Analyses (LDA). See our mathematical methods appendix section on Linear Discriminant Analysis for a discussion of this analytical technique and a worked example.

We began by removing from our ocean-colour dataset those measurements with optical properties suggestive of coccolithophores, river plumes or fjordic sediments, as in Figure 3.7.

Following the isolation and removal of these ocean-colour measurements, we subsequently sought to reduce variance in the training dataset originating from sources other than community composition. For example, variability in the elevation of the sun, due to latitude and season, could significantly vary brightness and surface scattering properties (see Sathyendranath et al. (2004), equation 2). We therefore normalised each ocean-colour retrieval by the magnitude of the sum of the contributions from all wavebands;

$$R_{rsm}(\lambda) = R_{rs}(\lambda) / \sum_1^n R_{rs}(\lambda), \quad (3.6)$$

where R_{rsm} is the magnitude-normalised remote-sensing reflectance, R_{rs} is remote-sensing reflectance, λ represents the MODIS-Aqua waveband, and n is the total number of wavebands considered (10).

Finally, we endeavour to control for variability in $[chl-a]$ content of the ocean-colour measurements in the training dataset. Variation in the $[chl-a]$ concentration is probably

one of the leading sources of variance in $R_{rsm}(\lambda)$ and it is likely that it covaries with community composition (Brewin et al. 2011). Indeed, in the Barents Sea study area, $[chl-a]$ exceeding 3 mg m^{-3} were the exclusive domain of diatom-rich samples. An attempt should therefore be made to deconvolve these two sources of variance in $R_{rsm}(\lambda)$, so that algorithms constructed on the training dataset are not biased by $[chl-a]$.

To achieve this, we treated $R_{rsm}(\lambda)$ as shape data. We selected all of the cleaned training dataset $R_{rsm}(\lambda)$ retrievals considered to be diatoms and performed a Procrustes-Anova regression (see Goodall 1991 for details) against the $[chl-a]$ calculated for each retrieval with the diatom power-law of Cota et al. (2003). We chose to use the power-law published by Cota et al. (2003), rather than seek to develop our own independent $[chl-a]$ proxy for the Barents Sea, because Cota’s review of the literature showed that most high-latitude $[chl-a]$ algorithms are similar (Cota et al., 2003, Figure 9).

The blue-green ratio (R) can be defined using the closest available MODIS-Aqua wavebands (Cota et al. 2003).

$$R = R_{rs}(443 > 488 > 531)/R_{rs}(555). \quad (3.7)$$

Calculation of inferred ocean-surface $[chl-a]$ for the i_{th} retrieval can be expressed as

$$chl - a_i = 10^{(n_0 + n_1 R_i)}, \quad (3.8)$$

where $n_0 = 0.7291$ and $n_1 = -2.258$, based on the power law for retrieving $[chl-a]$ in diatom-dominated waters (**Lab-OC4L diatom**), presented in (Cota et al. 2003), whereas prymnesiophytic waters (**Lab-OC4L prymn**) follow a trend in which $n_0 = 0.5313$ and

$$n_1 = -2.4$$

The relationship retrieved by Procrustes-Anova regression is robust, with an effect size score indicating the degree of association is 8.7 standard deviations higher than that expected from a randomly scrambled dataset. The constants derived from Procrustes-Anova regression can then be used to predict the values of $R_{rsm}(\lambda)$ for each retrieval, under the assumption that the pixel is dominated by diatoms; examples are provided in Figure 3.9.

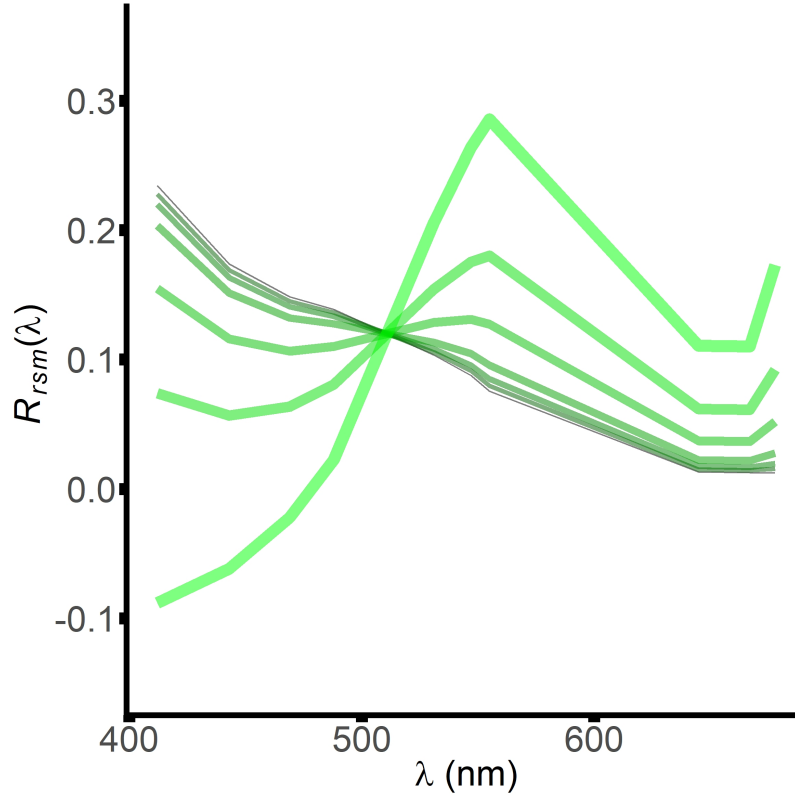


Figure 3.9: Predicted $R_{rsm}(\lambda)$ based on a Procrustes-Anova regression between remotely-sensed $[chl-a]$ and $R_{rsm}(\lambda)$ in diatom-dominated pixels from the cleaned training dataset. Increasingly thick/green lines correspond to $[chl-a]$ of 0.1, 0.5, 1, 2, 5, 10 and 20 mg m^{-3} . Note that at high $[chl-a]$ the bluest wavebands reach negative $R_{rsm}(\lambda)$ values, which are not physically tenable.

We used the diatom power-law of Cota et al. (2003) (**Lab-OC4L diatom**) to calculate a nominal $[chl-a]$ for each R_{rsm} in the training dataset. We then computed the ocean-colour

retrieval that we would expect from such a $[chl-a]$ in diatom-dominated waters ($R_{rsm}^{e,d}(\lambda)$), where the d superscript indicates that the nominal $[chl-a]$ is calculated with the diatom power-law. Thereafter, we found the difference between the measured ($R_{rsm}(\lambda)$) and expected ($R_{rsm}^{e,d}(\lambda)$) retrievals to arrive at ‘residual shape’ ($R_{rsRe}^d(\lambda)$);

$$R_{rsRe}^d(\lambda) = R_{rsm}(\lambda) - R_{rsm}^{e,d}(\lambda). \quad (3.9)$$

This dataset of residual shapes is used to train a Linear Discriminant Analysis (LDA).

We then created an analogous set of residual shapes from the training dataset, but using the prymnesiophyte power law (**Lab-OC4L prymn**) of Cota et al. (2003) to calculate the nominal $[chl-a]$, before the subtraction of the $R_{rsm}^{e,p}(\lambda)$ that would be associated with this $[chl-a]$ under the assumption of diatom taxonomy (where the p superscript here indicates that the nominal $[chl-a]$ is calculated with the prymnesiophyte power-law).

We decided to estimate the $[chl-a]$ of all ocean-colour observations that we hope to classify in accordance with the **Lab-OC4L prymn** $[chl-a]$ algorithm published by Cota et al. (2003), in order to be especially conservative in our removal of the signal of $[chl-a]$ concentration from ocean-colour, because **Lab-OC4L prymn** generally predicts higher values than **Lab-OC4L diatom**. This has the effect of decreasing bias to conflate change in $[chl-a]$ with change in phytoplankton groups. Hence the predicted $R_{rsm}(\lambda)$ that is subtracted has a higher ratio of blue-to-green wavelengths (Figure 3.9) and the $R_{rsRe}^p(\lambda)$ residual appears greener – as though it has a higher concentration of pigment. For diatom-dominated retrievals that already have high concentrations of pigment, this should not engender misclassification, but there is an increased risk that prymnesiophyte-dominated retrievals might be misclassified as diatoms.

Inspection of Figure 3.9 shows that high $[chl-a]$ produce some negative $R_{rsm}^e(\lambda)$ values at blue wavelengths, which is physically untenable, resulting from extrapolation of the lin-

ear Procrustes-distance based fit between $R_{rsm}(\lambda)$ and inferred $[chl-a]$. This will result in higher $R_{rsRe}(\lambda)$ values at blue wavelengths for ocean-colour retrievals with high inferred $[chl-a]$.

To assess the utility of the empirical algorithm we have constructed, we applied hold-p-out cross-validation, to predict the taxonomic identity of random selections of 500 retrievals from the training dataset. The results are presented as a confusion matrix in Table 3.2.

Table 3.2: *A confusion matrix representing the results of 1000 permutations of hold-p-out cross validation of the training dataset.*

		Predicted			
		Diatoms	Dinoflagellates	Nanophytoplankton	<i>Phaeocystis</i>
Actual	Diatoms	95.87	0.3	3.08	0.75
	Dinoflagellates	18.08	57.45	15.26	9.21
	Nanophytoplankton	16.31	13.15	68.46	2.08
	<i>Phaeocystis</i>	53.11	14.39	11.94	20.57

Confusion matrices are a simple means of presenting the mean success of a classifier, and how it varies across multiple groups.

A target set of data to be classified can be considered, for which the classes are known. The classifier can then be used to generate predicted classes, and the pairwise relationships between predicted and actual classes can be represented as a matrix.

In this matrix, the diagonal will represent successful classifications for which the predicted classes match the actual classes. The off-diagonals will represent incorrect classifications. If the actual classes are organised along the rows of the confusion matrix, and the predicted classes as columns, then each row i represents the predicted classifications for the i^{th} class. This matrix therefore succinctly provides the reader with information about whether a classifier's predictions are more or less likely to confuse some classes than others. A worked example and code, generating a confusion matrix, is provided in the mathematical methods appendix section on Linear Discriminant Analysis.

The results in the confusion matrix show that diatom and nanophytoplankton-dominated samples can be recognised with high levels of reliability, whereas our ability to discriminate dinoflagellate and *Phaeocystis*-dominated ocean-colour observations is less strong. Dinoflagellates are predicted with a higher accuracy than random, while *Phaeocystis* retrievals are usually mistaken for diatoms.

The difficulty recognising dinoflagellates and *Phaeocystis* could be attributable to their forming blooms that have smaller spatial extents than the 1° box used to assemble the training dataset. For example, a *Phaeocystis* bloom might occur mixed with a diatom bloom on a spatial scale much finer than this, if *Phaeocystis* colonies are nucleated on senescent diatom frustules (Peperzak et al. 2000b). Additionally, d’Ovidio et al. (2010) showed that turbulent mixing of the ocean can result in phytoplankton blooms being drawn out into filaments on the order of 100 km long, but only 10 km wide – a scale of spatial heterogeneity too small to be resolved by the coarse 1° sampling we used to construct our training dataset.

Phaeocystis blooms may also be confused with diatom blooms because of similar optical properties, such as high pigment packaging (see Figure 3.6). (*Phaeocystis* assemblages sourced from Stuart et al. (2000) have $[chl-a]$ normalised absorption ($a_{ph}^*(440)$) values around $0.041 \text{ m}^2 \text{ mg}^{-1}$, compared with diatom values of $0.045 \text{ m}^2 \text{ mg}^{-1}$ and nanophytoplankton values of $0.084 \text{ m}^2 \text{ mg}^{-1}$. These values differ slightly from those presented in appendix 1 of Stuart et al. (2000))

Many of the dinoflagellate-dominated pixels have very low $[chl-a]$, so there may frequently be an insufficient number of dinoflagellate cells suspended in the surface waters to give these retrievals a distinctive spectral signature. In addition, the mixotrophic nutrition of dinoflagellates may mean that the relationship between pigment biomass and cell biomass is complex; previous authors have shown that mixotrophic dinoflagellates may retain the chloroplasts of ingested prey, altering their pigment properties (Li et al. 1999).

We used the LDA residual shape algorithm to classify phytoplankton assemblages in a

satellite image of the Southern Barents Sea that was presented by Carlowicz et al. (2014), with the tentative suggestion that it depicted the succession from diatom to coccolithophore blooms. Our linear divider that masks coccolithophores indeed identifies a bloom. In addition, the results of our LDA classification identify a spatially coherent swath of diatoms (Figure 3.10) along with the identification of dinoflagellates dominating in what appear to be deep blue oligotrophic waters.

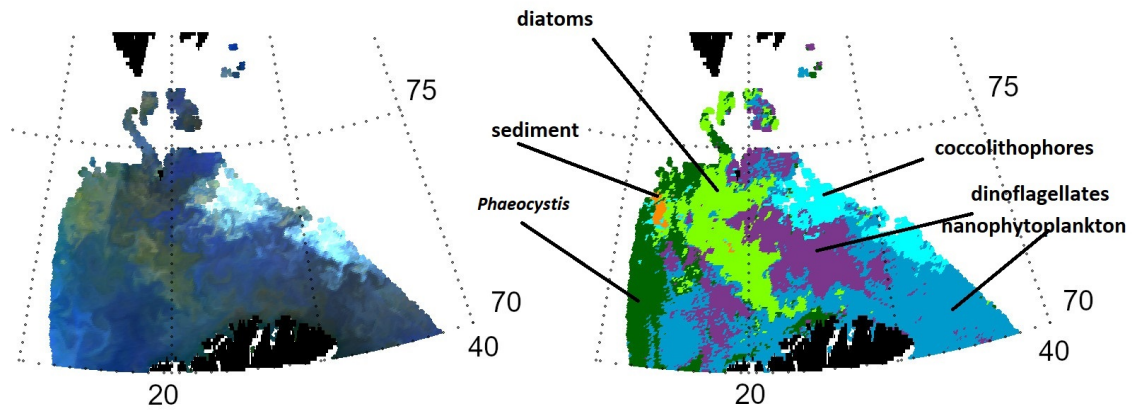


Figure 3.10: Satellite image of the Southern Barents Sea on Julian day 191, 2014. **Left:** Pseudo true-colour representation. A coccolithophore bloom is clearly visible in the North-east. **Right:** Classification by application of masks followed by LDA: cyan (coccolithophores), orange (sediment), light-green (diatoms), dark-green (*Phaeocystis*), purple (dinoflagellates), blue (nanoplanktic prymnesiophytes).

We compared the classifications of our LDA-based algorithm on the satellite image presented in Carlowicz et al. (2014), with the dichotomous semi-empirical algorithms of Sathyendranath et al. (2004) and Jackson et al. (2010) (Figure 3.11), which are tuned to the Labrador and Humboldt Seas respectively. These algorithms were constructed for use with the SeaWiFS platform; we have used the closest available MODIS-Aqua wavebands as input.

Figure 3.11 shows that the Labrador Sea and Humboldt Sea algorithms produce similar

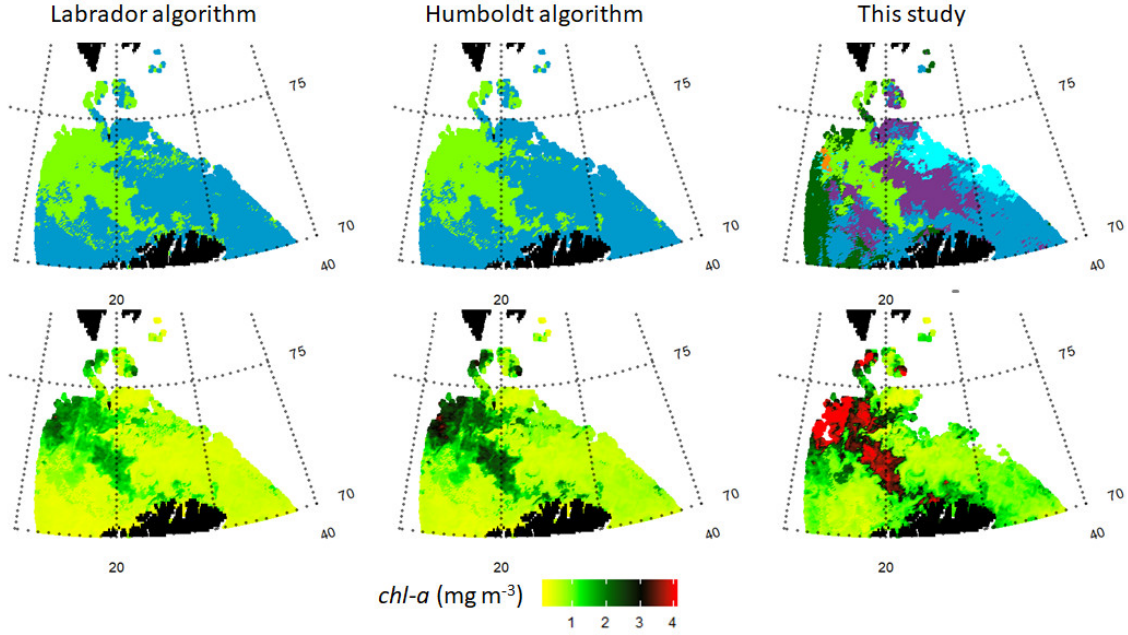


Figure 3.11: **Top-Left:** division between diatoms (light-green) and ‘mixed’ populations (blue) resulting from the Labrador-tuned semi-empirical algorithm of Sathyendranath et al. (2004). This algorithm uses spectral look-up-tables to generate guesses for $[chl-a]$ from two different waveband pairs for each proposed community composition, and then selects the community composition for which the difference between the $[chl-a]$ guesses is smallest. **Top-Centre:** division between diatoms and ‘mixed’ populations resulting from the Humboldt-tuned semi-empirical algorithm of Jackson et al. (2010). This algorithm functions identically to that of Sathyendranath et al. (2004), except that the look-up-tables are parameterised from phytoplankton samples originating in the Humboldt, rather than Labrador Sea. **Top-Right:** division between diatoms (light-green), *Phaeocystis* (dark-green), nanoplanktic prymnesiophytes (blue), dinoflagellates (purple), with sediment (orange) and coccolithophores (cyan) masked out, based on empirical algorithm developed in this study. **Bottom-Left, -Centre, -Right:** corresponding $[chl-a]$ heat-maps, scale in mg m^{-3} .

spatial patterns of classification – especially a swath of diatom-dominated retrievals that extends from the Southern tip of Spitsbergen towards Nordkapp, Norway (Figure 3.11).

Hence, although the two semi-empirical algorithms are based on simulated $R_{rs}(\lambda)$ for populations of diatoms and prymnesiophytes from different parts of the World Ocean, and possess different optical properties (see Jackson et al. (2010) Figure 2) it appears that the

differences between the optical properties of the two groups are persistent enough through the Global Ocean for algorithms parameterised in different regions to produce similar classifications in the Barents Sea.

However, although the taxonomic classifications are similar, the corresponding $[chl-a]$ predicted by the semi-empirical algorithms differ considerably (Figure 3.11). The algorithm tuned to the Humboldt Sea retrieves $[chl-a]$ on the order of half those of the Labrador algorithm.

This results because Humboldt phytoplankton have pervasively higher absorption coefficients per unit $[chl-a]$ than their Labrador Sea equivalents, hence a lower concentration of pigment is required to match the $R_{rs}(\lambda)$ band ratios of a given retrieval. Jackson et al. (2010) hypothesise that smaller phytoplankton cell sizes in the equatorial Pacific, compared with the Labrador Sea, may be associated with a diminished packaging effect, explaining the difference in chlorophyll-specific absorption (Morel and Bricaud 1981).

The empirical Barents Sea algorithm does not generate a best estimate of $[chl-a]$ for each ocean-colour retrieval as a necessity of its construction, which means that selecting appropriate $[chl-a]$ is a subjective exercise.

Here, we chose to state that $[chl-a]$ should be calculated according to the diatom power-law when a retrieval is classified as diatom or *Phaeocystis*-dominated, while the prymnesiophyte power-law is used when a retrieval is dominated by dinoflagellates or nanoplanktic prymnesiophytes (Figure 3.11). Blended approaches are also possible – for instance the posterior probability of diatom or nanoplanktic prymnesiophyte membership of a retrieval could be used to generate a weighted average of the two power-laws’ predictions.

3.9.1 Discrimination of diatoms and *Phaeocystis*

Given our weaker confidence in the detection of dinoflagellate blooms, we made the decision to exclude dinoflagellate-labelled entries from our training dataset and to concentrate on the utility of the Linear Discriminant we had found for distinguishing blooms of highly-packaged phytoplankton (diatoms and *Phaeocystis*) from ocean waters dominated by nanophytoplankton. The discriminant coefficients are presented in Table 3.3:

Table 3.3: *Discriminant coefficients, rounded to 2 decimal places.*

λ (nm)	412	443	469	488	531	547	555	645	667	678
test	-6.71	-46.67	22.62	46.64	128.03	-161.17	-43.39	76.70	-226.56	51.84

It can be seen in Table 3.3 that there is a strong variation in the coefficients of different spectral wavebands associated with absorption by *chl-a*, such as 443 and 667 nm. Variation in pigment packaging distorts the relationship between the ratio of absorption at these wavelengths and [*chl-a*] (review Figure 3.6), hence we do not consider it surprising that a facsimile of this pattern is manifest in the discriminant coefficients.

Owing to the optical similarity of diatom and *Phaeocystis*-labelled ocean-colour residual spectral shape, we considered the possibility that the differences between diatom and *Phaeocystis* ocean-colour observations are subtle and that LDA had not found them. We therefore opted to develop an additional stage in our algorithm, with the intent to take ocean-colour residual spectra labelled as highly-packaged phytoplankton and divide them into diatom and *Phaeocystis*-dominated measurements.

We made use of HydroLight-Ecolight software (Mobley and Sundman 2012) to simulate the expected ocean-colour under conditions of diatom and *Phaeocystis* dominance, for [*chl-a*] ranging between 0.1 and 25 mg m⁻³ (measured values of *in-situ* samples varied between 0.11 and 24.23 mg m⁻³).

We implemented these simulations with phytoplankton absorption spectra for our di-

atom and *Phaeocystis* groups (as in Figure 3.5), a chlorophyll-specific scattering model based on a combination of Barents and Ligurian Sea measurements and a chlorophyll-specific back-scattering model based on a combination of Barents Sea and NOMAD measurements (Valente et al. 2019).

The spectral absorption of phytoplankton was constrained to evolve according to the relationship identified in Bricaud et al. (1995) and absorption due to CDOM was treated as constant, with a value of 0.04 m^{-1} at 440 nm (Hancke et al. 2014), decaying with wavelength according to the function identified by Bricaud et al. (1981).

Scattering and back-scattering model coefficients are presented in Table 3.4, wherein scattering and back-scattering evolve according to equation 1.4:

$$b_i(\lambda) = A_i(\lambda)(chl-a)^{B_i(\lambda)}, \quad (3.10)$$

where b_1 represents total scattering of particles and seawater (m^{-1}), b_2 represents backscattering (m^{-1}); $chl-a$ is the $[chl-a]$ concentration (mg m^{-3}), and A_i and B_i represent fit parameters.

Table 3.4: *Scattering and Backscattering model coefficients and goodness of fit.*

λ (nm)	A_1	B_1	R_1^2	A_2	B_2	R_2^2
412	0.3893	0.0842	0.0177	0.0027	0.4116	0.5727
440	0.4235	0.4089	0.2606	0.0026	0.4967	0.7643
488	0.4271	0.4282	0.2814	0.0023	0.5017	0.7614
510	0.4324	0.4403	0.2961	0.0022	0.5038	0.7581
532	0.4418	0.4852	0.4803	0.0021	0.4898	0.7503
555	0.4459	0.5105	0.5071	0.0020	0.5033	0.7453
650	0.3972	0.4554	0.2833	0.0016	0.5169	0.7132
676	0.3777	0.5215	0.4662	0.0015	0.5048	0.7016
712	0.3595	0.1879	0.0884	no NOMAD data		

We recognise that, within the Arctic **PRIZE** dataset, *Phaeocystis* is a chlorophyll-c3-rich plankton, compared with diatoms. (Although we recommend that other researchers take care to interpret the presence of chlorophyll-c3 carefully, as it can occur in other

phytoplankton, including some species of diatoms, prymnesiophytes and in fucoxanthin-containing dinoflagellates (Millie et al. 1997), therefore researchers require specific justification if they mean to associate this pigment with a single genus within the scope of their studies.)

We therefore surmised that we might be able exploit this optical difference, evident in Figure 3.5, in order to distinguish our simulated spectra. If this can be achieved, it may be possible to distinguish diatom and *Phaeocystis* blooms in our training dataset of remotely-sensed ocean-colour observations.

We adopted the chlorophyll-c3 absorption algorithm, developed by (Astoreca et al. 2008), to determine whether a signal attributable to chlorophyll-c3 could be detected in the diatom and *Phaeocystis* HydroLight simulations. We made use of the **Lab-OC4L diatom** [*chl-a*] algorithm published by (Cota et al. 2003) to infer [*chl-a*] content, reasoning that *Phaeocystis* would be typified by a high chlorophyll-c3 signal per unit [*chl-a*], compared to diatoms.

The chlorophyll-c3 algorithm of Astoreca et al. (2008) relies upon use of the MODIS land band centred at 469 nm to detect a shoulder in $a_{ph}(\lambda)$ associated with chlorophyll-c3 (consult Figure 3.5 to observe this feature). The use of the MODIS land bands, including 469 nm, has already been applied to the identification of bio-optically distinct phytoplankton blooms by Hu et al. (2010), who successfully identified *Trichodesmium* blooms in optically complex coastal waters.

We found [*chl-a*] estimates were broadly lower than expected (Figure 3.12). However, given that the magnitude of simulated absorption spectra used as input to HydroLight are computed as a function of [*chl-a*], we should still expect the relationship between absorption signal attributable to chlorophyll-c3 and [*chl-a*] to be informative.

We found, upon performing a PCA on ocean-colour residual spectra in our training database labelled as diatom or *Phaeocystis*, that there was a bimodal normal distribution of the data in Principal Component 5-6 space (Figure 3.12). We suspected that this likely rep-

resented two dominating phytoplankton assemblages, but we note that the entries labelled as diatom or *Phaeocystis* are distributed across both of the Gaussian features, suggesting that the reason that LDA could not retrieve an effective discriminant was due to labelling in the training dataset confusing these groups. This confusion is probably as a result of their close spatial co-occurrence during cruises.

It is clear that the relationship between $[chl-a]$ and the common logarithm of absorption signal attributable to chlorophyll-c3 ($\log_{10}(a_{c3})$) separates diatom and *Phaeocystis* from one another in HydroLight simulations (Figure 3.12 c). Although we note that PPC (Photo-protective Carotenoids) may interfere with absorption by chlorophyll-c3 in the blue-green portion of the spectrum (Johnsen and Sakshaug 2007).

We devised a linear divider, relating $\log_{10}(a_{c3})$ and inferred $[chl-a]$, to coarsely divide diatom and *Phaeocystis* simulations. When this was applied to the training dataset, the results broadly conformed to the two Gaussian features we had previously identified (Figure 3.12).

An example application of $\log_{10}(a_{c3}) / [chl-a]$ distinguishing diatom and *Phaeocystis* assemblages is demonstrated in Figure 3.13, where it can be seen that two adjacent phytoplankton blooms – that are clearly demarcated by a front in ocean-colour – are classified as diatom and *Phaeocystis*-dominated.

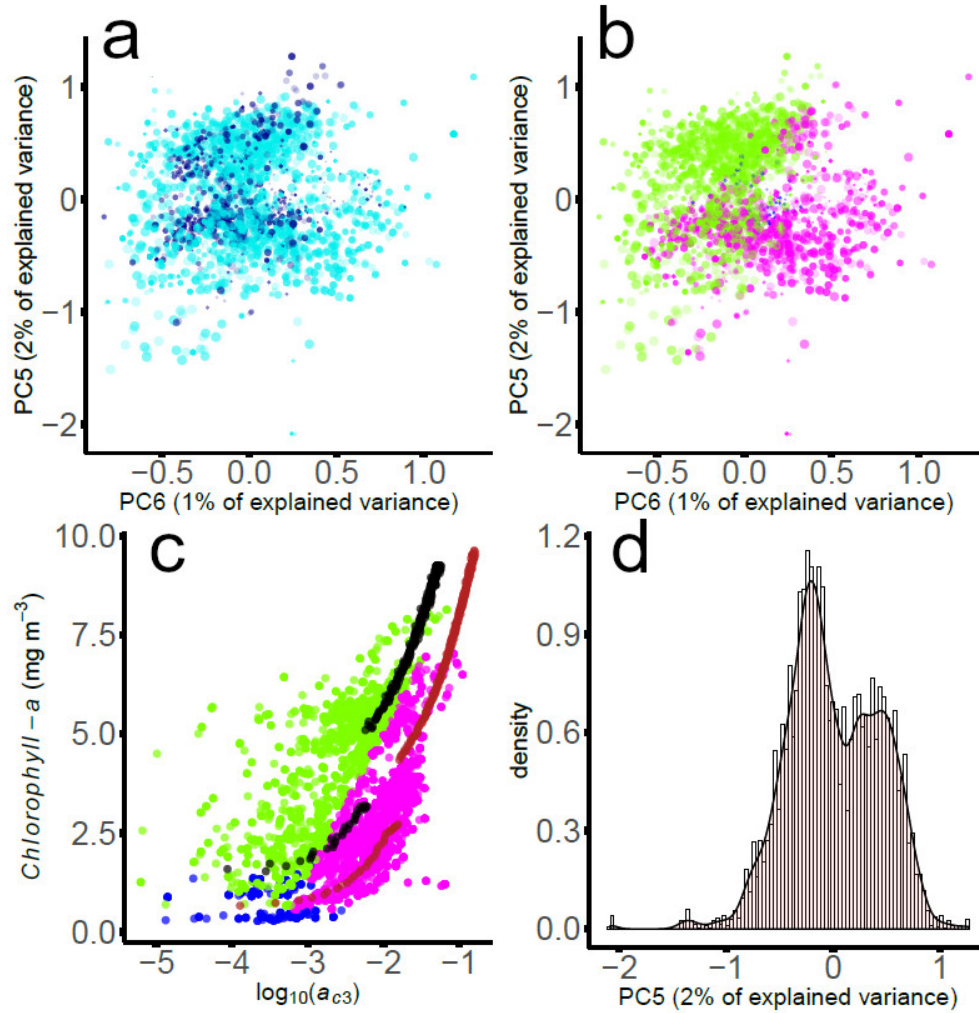


Figure 3.12: a) PC5-6 rotation of the training entries for diatoms (light blue) and *Phaeocystis* (dark blue) – point size is scaled to inferred $[chl-a]$. b) The same space as subplot a, except entries have been re-classified in accordance with our bio-optical algorithm; diatoms (green), *Phaeocystis* (magenta), nanophytoplankton (blue). c) a linear divider is developed that distinguishes simulated diatoms (black) from simulated *Phaeocystis* (dark red), with the training dataset entries from b projected onto the space. d) A histogram, illustrating bimodal distribution of training dataset entries along PC5.

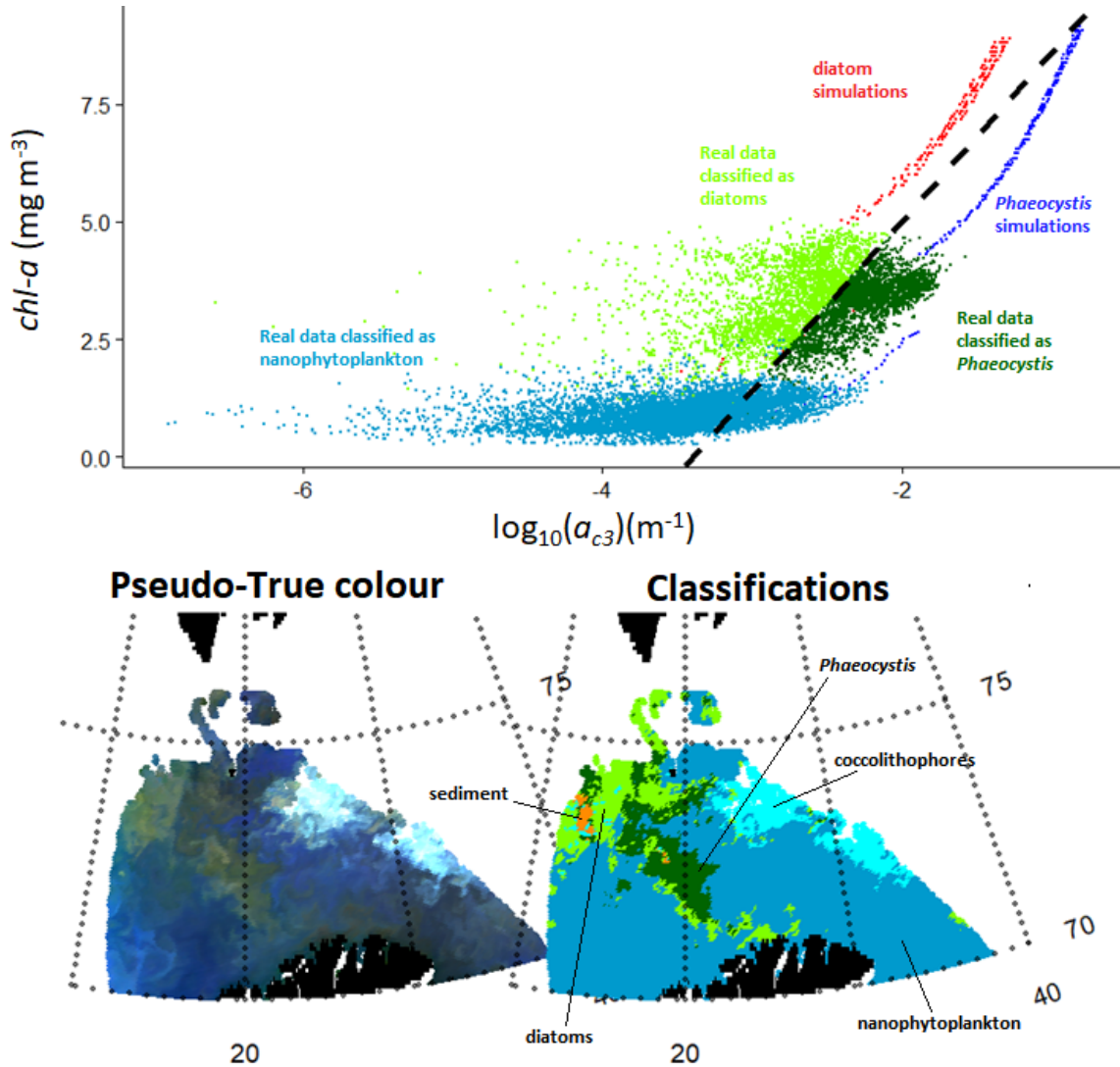


Figure 3.13: Remotely-sensed ocean-colour from the Southern Barents Sea on Julian day 191, 2014. **Top:** plot of chlorophyll-a and absorption signal attributable to chlorophyll-c3 inferred from ocean-colour. Red: trend based on HydroLight simulations of ocean-colour under diatom-bloom scenario (Mobley and Sundman 2012). Blue: trend under simulated colonial *Phaeocystis* scenario. Light green: suspected diatom-dominated community in real image. Dark green: suspected *Phaeocystis*. Teal: suspected unicellular nanophytoplankton. Note that some nanophytoplankton have relatively high chlorophyll-c3 signals. This could be indicative of other prymnesiophyte genera known to possess chlorophyll-c3, such as *Chrysochromulina* (Johnsen and Sakshaug 2000). **Bottom:** pseudo-true colour image (left) and inferred phytoplankton functional types (right), with coccolithophores masked in cyan and waters possibly influenced by sediment masked in orange.

3.10 Summary of resulting bio-optical Algorithm structure

A modular flow chart of the structure of the Barents Sea bio-optical algorithm we have developed is presented in Figure 3.14.

In module 1, R_{rs} input data are quality-checked in order to remove pixels whose bio-optical properties are strongly influenced by factors other than absorption by phytoplankton, such as those hosting significant loads of sediment, CDOM or coccolithophores.

Only those waters assessed to be unaffected by these substances are passed to module 2, in which an attempt is made to disentangle variance in R_{rs} ratios caused by the phytoplankton community composition and $[chl-a]$. When this is achieved a judgement is made about whether the R_{rs} signal attributable to community composition is more compatible with highly packaged assemblages of phytoplankton or with single-celled nanophytoplankton. After this judgement is made, appropriate algorithms are applied to estimate $[chl-a]$ concentration.

Data that are judged to represent highly packaged assemblages of phytoplankton are passed to the third module. It is presumed that these data must represent either diatom or *Phaeocystis* communities – the two abundant highly-packaged phytoplankton assemblages that regularly recur in the Barents Sea. The chlorophyll-c3 concentration of these data are inferred and compared with estimates of $[chl-a]$ concentration in order to establish whether the data are indicative of *Phaeocystis* blooms, which we believe to be richer in chlorophyll-c3. We acknowledge that other phytoplankton taxa may also be rich in chlorophyll-c3, such as some dinoflagellates (Millie et al. 1997), prymnesiophytes such as *Emiliania huxleyi* (Garrido and Zapata 1998; Johnsen and Sakshaug 2000) and silicoflagellates (Roy et al. 2011).

The modular algorithm therefore concludes having identified 4 phytoplankton community compositions; coccolithophores, nanophytoplankton, diatoms and *Phaeocystis*.

$[chl-a]$ estimates are provided for nanophytoplankton, diatoms and *Phaeocystis* but cannot be estimated for coccolithophores, as a result of the strong disturbance of R_{rs} ratios

in these waters by calcite liths, obscuring the relationship between R_{rs} and absorption by phytoplankton pigments.

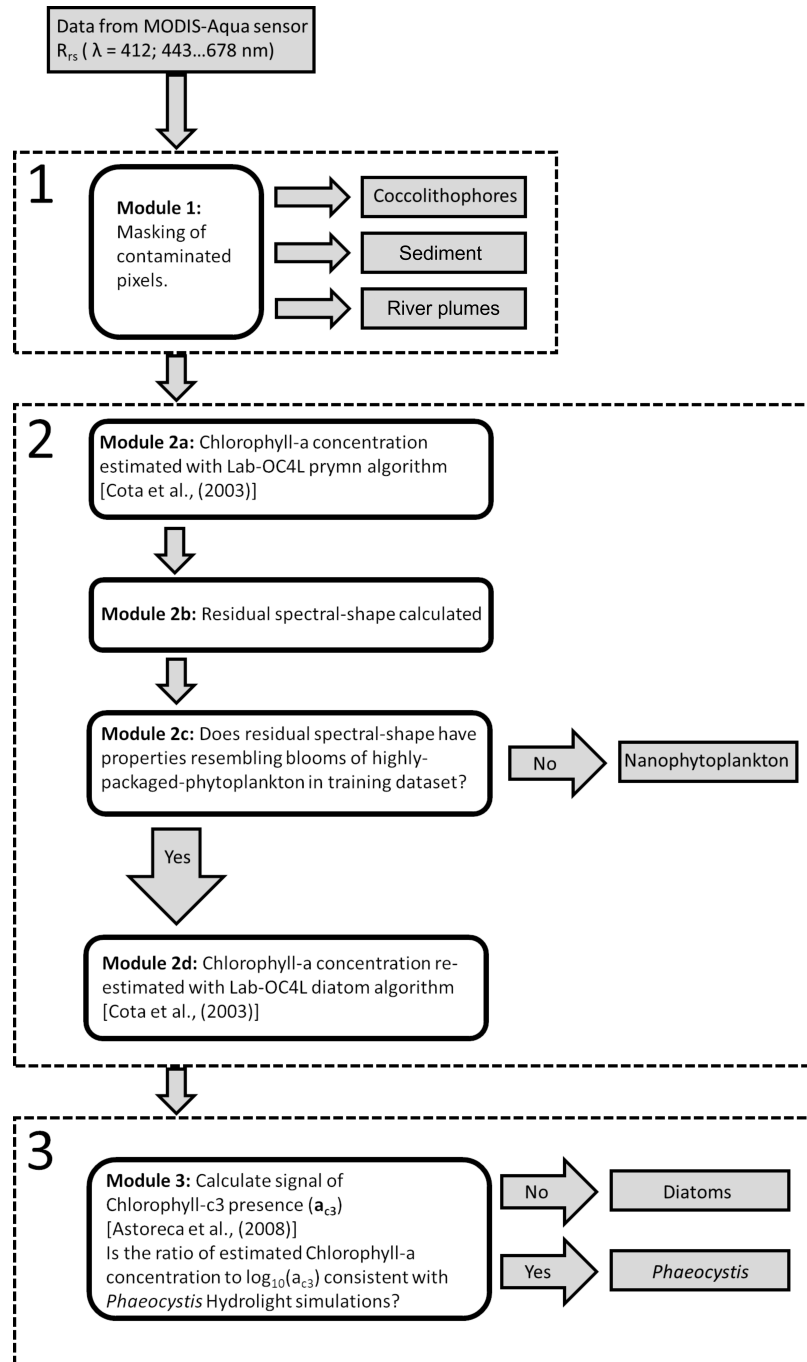


Figure 3.14: Modular flow chart of the structure of the Barents Sea bio-optical algorithm. Note that blooms of *Emiliana huxleyi*, which possess chlorophyll-c3 (Johnsen and Sakshaug 2000), are masked well before chlorophyll-c3 is employed as a marker pigment to distinguish diatom and *Phaeocystis*-dominated assemblages.

3.11 Correction for drift in MODIS-Aqua ocean-colour data

We now move on to discuss the stability of the derived variable, chlorophyll-c3 absorption signal (a_{c3}), in time series of MODIS-Aqua ocean-colour.

We assessed this by compiling time series of mean a_{c3} estimated from ocean-colour across several years in different parts of the Global Ocean that we would not expect to behave in unison. If these disparate regions show common changes it may provide evidence of systematic bias in the inference of a_{c3} .

We chose to investigate the North Atlantic Sub-Tropical Gyre in particular because we should expect the ocean-colour in the region to be relatively constant over time, as it is oligotrophic ($[chl-a] < 0.1 \text{ mg m}^{-3}$), and to a first-order may be presumed to represent clear ocean waters when contrasted with the eutrophic Barents or Labrador Seas (which we define as the regions between 70-85° North, 0-50° East and 50-65° North, 40 to 60° West).

In Figure 3.15, the North Atlantic Sub-Tropical Gyre region is assumed to occupy the region between 30-50° West and 25-35° North. We downloaded MODIS-Aqua ocean-colour observations in this region for Julian days 130-135 between 2003-2018. The fact that a_{c3} in this region is clearly related to that in the Barents and Labrador Seas (Figure 3.15) evidences drift in the representation of ocean-colour from the MODIS-Aqua r2018.1 reprocessing; a clear step-change is visible in all time series at 2010, for example. It is hence necessary to correct for this systemic form of bias.

Our correction scheme is based on that of Taylor et al. (2018), who observed a shift in the spectral output of the IASI (Infrared Atmospheric Sounding Interferometer) remote-sensing platform in 2010, which biased results in an algorithm they had developed to infer Sulphur dioxide.

To correct for this drift, Taylor et al. (2018) took the lowest concentration of inferred

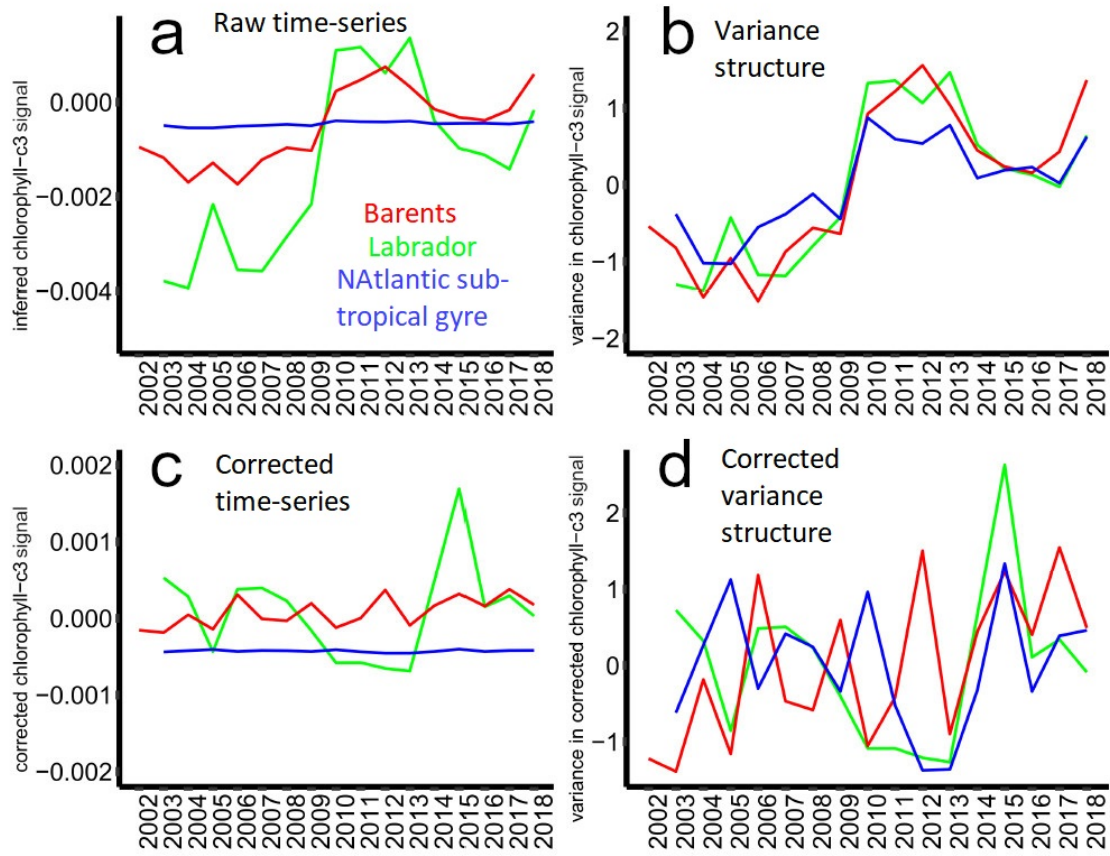


Figure 3.15: a) Annually-binned mean signal due to chlorophyll-c3 for the Labrador Sea (green), Barents Sea (red) and North-Atlantic Sub-Tropical Gyre (blue). b) Trends in inferred chlorophyll-c3 signal normalised to variance, with Sub-Tropical Gyre variance multiplied by 10 to make it visible. c) Annually-binned mean signal due to chlorophyll-c3 after a correction for drift is applied. d) Trends in corrected chlorophyll-c3 signal normalised to variance, with Sub-Tropical Gyre variance multiplied by 10 to make it perceptible.

Sulphur dioxide in any image, that was within 1 standard deviation of the mean, and added its magnitude to all remaining pixels.

In our approach we first constrained all a_{c3} signal values to within -0.01 and 0.005 (see Figure 3.15 a), so that aberrant values would not significantly disturb calculation of the standard deviation. The lower negative limit of -0.01 may be surprising, but Astoreca et al. (2008) themselves found that inferred a_{c3} signal became negative in diatom-dominated

conditions.

We then divided the values into an annually-binned time series and calculated the difference between the mean annual a_{c3} values and the lowest a_{c3} value in each annual bin that fell within 1 standard deviation of the mean a_{c3} in each bin. We added the magnitude of the lowest annually-binned a_{c3} values to the remaining a_{c3} values in each bin, and then shifted the resulting absorption values so that the mean of all a_{c3} values between 2010-2018 matched the mean of the original a_{c3} values (since our bio-optical algorithm was constructed around *in-situ* samples collected after the 2010 shift in inferred chlorophyll-c3 absorption).

In our efforts to illustrate drift in sensor output we computed pair-wise t-tests comparing mean R_{rs} values between 2003-2009 and 2010-2018 in the North Atlantic Sub-Tropical Gyre (Figure 3.16). We found that a significant shift has occurred among the most blue bands and that this was associated with a decrease in the R_{rs} of bands with the shortest wavelengths.

Astoreca et al. (2008) had speculated on the question of whether minute differences in R_{rs} near 469 nm would be perceptible by remote-sensing platforms. We believe we have found evidence that these differences are detectable (Figure 3.12), but we caution that for the particular application of distinguishing *Phaeocystis* from diatom blooms, it appears necessary to account for anomalies resulting from temporal drift. We caution that different correction schemes may be necessary for future ocean-colour reprocessings, because they may subtly alter the ratios between different ocean-colour wavebands in a systematic fashion. After a correction scheme is applied, it can be seen that temporal drift in mean annually-binned a_{c3} is largely removed and patterns of variance in the different seas can be considered decoupled (Figure 3.15).

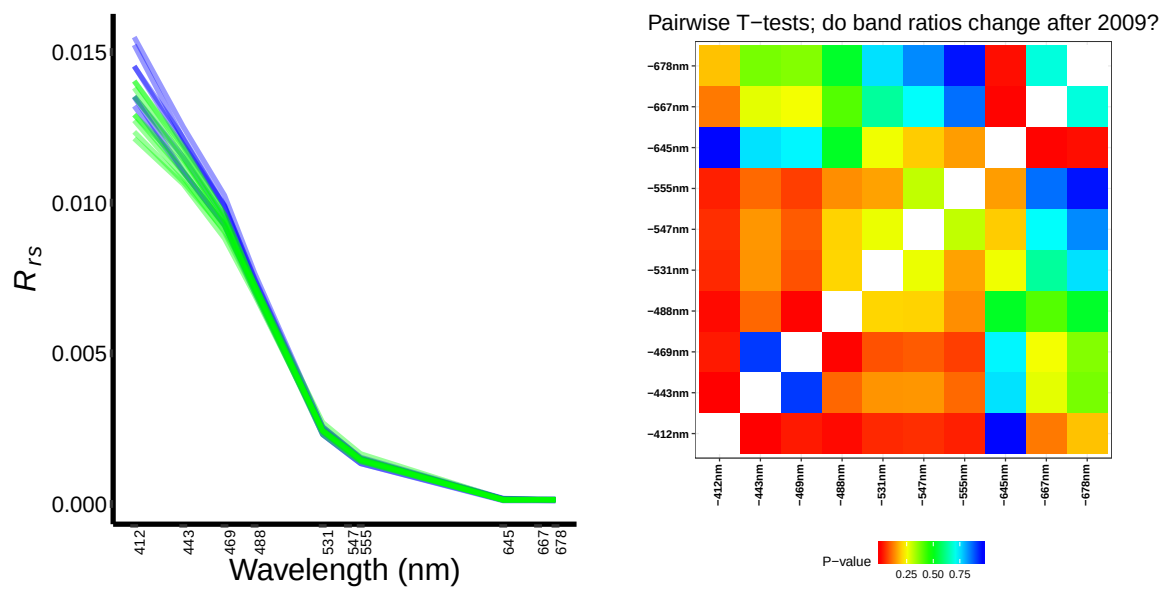


Figure 3.16: **Left:** Ocean-colour reflectance spectra for the North Atlantic Sub-Tropical Gyre before 2010 (blue) and after 2010 (green), a shift in spectral shape is evident. **Right:** Pair-wise tests for separation of means show that band ratios among the bands 443, 469 and 488 nm have shifted significantly after 2010.

3.12 Review of bio-optical algorithm classifications and concluding remarks

Figures 3.17 and 3.18 represent time series for the occurrence of different phytoplankton blooms, identified with our experimental bio-optical algorithm, in the Labrador and Barents Seas, once the correction for drift is applied. The subplots in Figure 3.17 represent integrations of ocean-colour observations from the MODIS-Aqua sensor across May in the years 2005-2014. The RGB values are set to reflect linear contributions of nanophytoplankton (blue), *Phaeocystis* (red) and diatom (green) classifications, according to the ocean-colour algorithm developed in this chapter. In addition to this, whiteness is set as a linear function of the frequency of coccolithophore classifications. Black values indicate that a location was not observed by MODIS-Aqua within the monthly integration. This therefore establishes an intuitive colour-space that the reader can interrogate to observe geographic and temporal variation in the broad phytoplankton community assemblages that are inferred from remote-sensing.

Furthermore, filled circles are overlaid upon the subplots in Figure 3.17, corresponding to *in-situ* phytoplankton samples. A discussion of these transects is presented in Fragoso et al. (2017) and the original dataset is accessible at

<https://doi.pangaea.de/10.1594/PANGAEA.871872>

The colour of the samples is arrived at through CHEMTAX analysis of pigments derived from HPLC analysis. Green samples are dominated by diatoms, red samples by *Phaeocystis*, and blue samples are dominated by neither and most likely represent mixed populations dominated by nanophytoplankton. It can be observed that the *in-situ* samples broadly show a pattern of recurrent *Phaeocystis* blooms off the Greenland coast, while diatoms dominated towards Newfoundland and Labrador. The central Labrador Sea usually supports mixed populations. The bio-optical algorithm classifications produce a similar overall geographic

pattern, providing confidence that the bio-optical algorithm developed in this chapter can provide meaningful synoptic assessments of ecosystem state and dynamics.

An analogous Figure for monthly integrations of ocean-colour in the Barents Sea in July, between 2002 and 2018, is presented in Figure 3.18. Comparison of these Figures illustrates intriguing differences in the ecosystem dynamics of these two regions; while there are clear geographic regions in the Labrador Sea that support distinct phytoplankton assemblages, the distribution of assemblages across the Barents Sea is much more homogeneous. This observation accords well with comments by Degerlund and Eilertsen (2010) that spring blooms in the Barents Sea are homogeneous in their community composition, and that the dominant mode of variation is inter-annual rather than geographical. We can perhaps persuade ourselves that coccolithophore blooms, indicated by white values, are restricted to the southern Barents Sea. This observation is compatible with comments by numerous authors (e.g. (Hovland et al. 2013; Neukermans et al. 2018; Oziel et al. 2020)). We may be able to form hypotheses, interrogating these images, about secular change. The overall distribution of coccolithophores appears to expand eastwards towards the end of the time series, for example, and there appears to be a greater contribution of *Phaeocystis*. Confirmation of these observations requires rigorous statistical analysis and will form the focus of the next chapter.

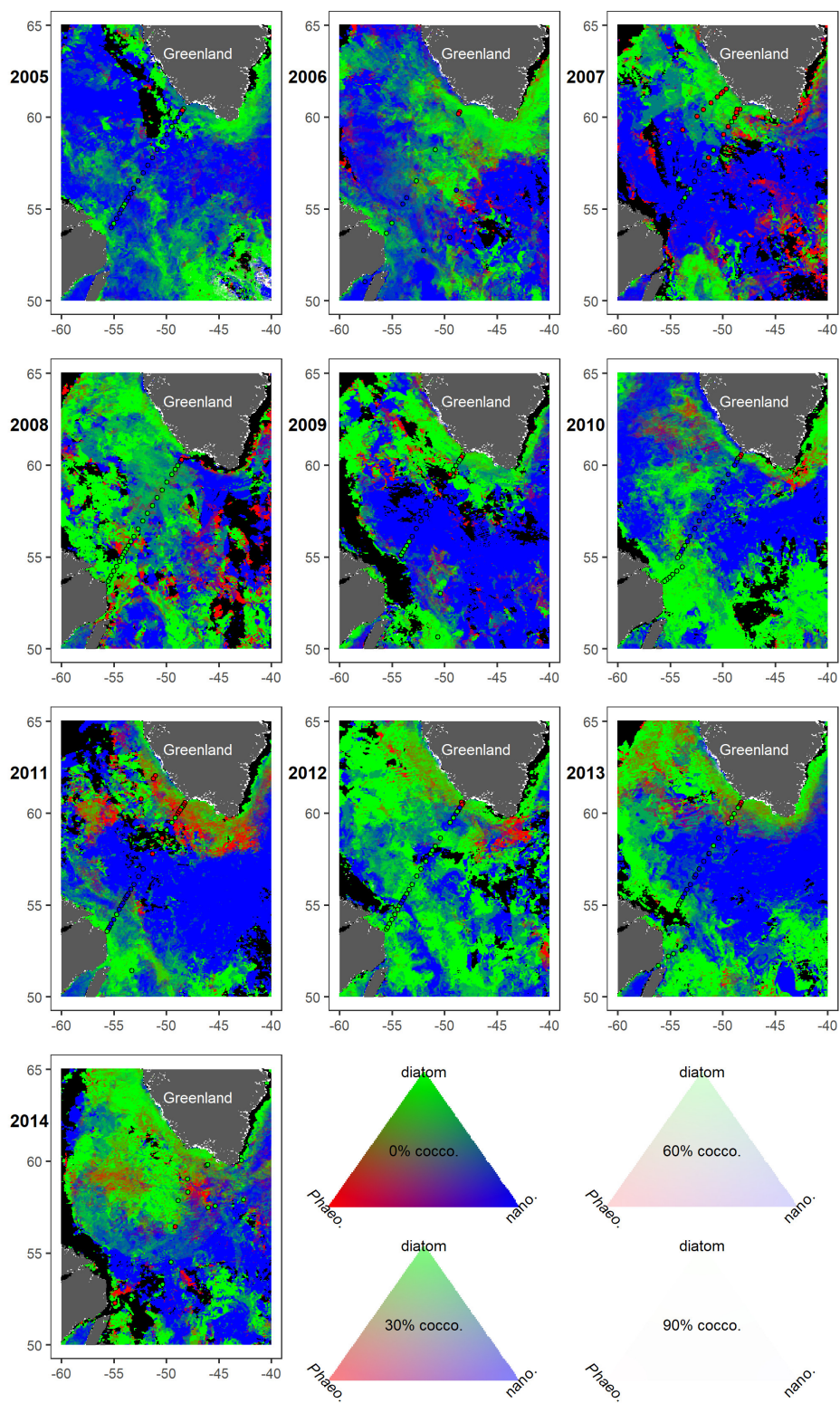


Figure 3.17
160

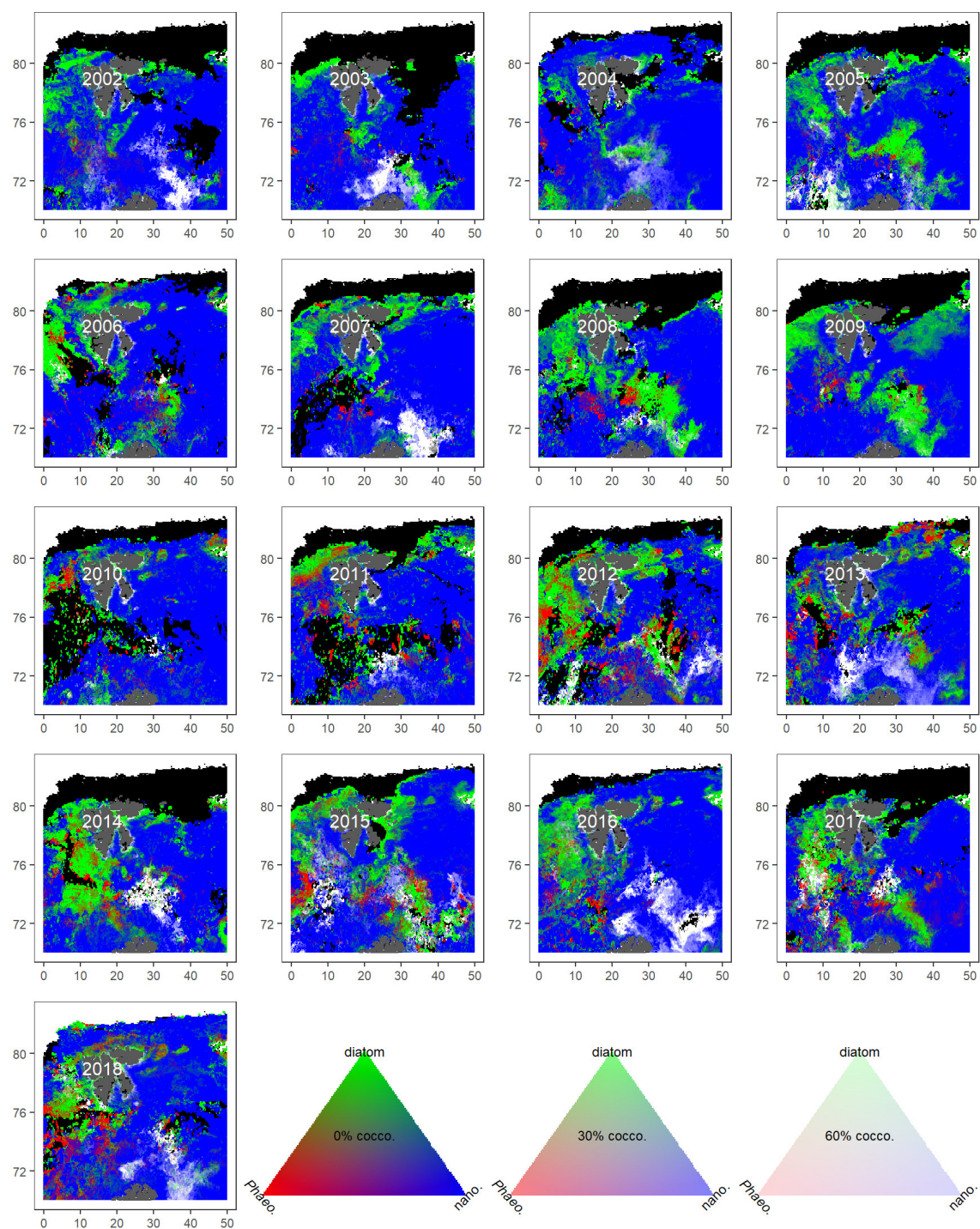


Figure 3.18

The datasets of optical properties used in the construction of the phytoplankton community composition algorithm used in this study are available at

[doi:10.5285/97daa7ea-8792-6cff-e053-6c86abc0dd46]

[doi:10.5285/982b6da2-7e11-060a-e053-6c86abc09389]

[doi:10.5285/982b6da2-7e12-060a-e053-6c86abc09389]

accompanying datasets for [*chl-a*] are available at

[doi:10.5285/97daa7ea-8793-6cff-e053-6c86abc0dd46]

[doi:10.5285/982b6da2-7e13-060a-e053-6c86abc09389]

[doi:10.5285/982b6da2-7e14-060a-e053-6c86abc09389]

MODIS-A data is available from NASA's website for ocean-colour.

[doi:10.5067/AQUA/MODIS/L3SMI/]

A. Orkney carried out analyses and drafted the manuscript.

H.A. Bouman supervised analysis and the drafting of the manuscript.

T. Platt edited the manuscript.

I. Kostakis set up HydroLight simulations.

D. McKee set up HydroLight simulations.

S. Sathyendranath assessed the validity of our correction scheme for temporal drift.

All authors read and approved the manuscript.

This research was funded by the Natural Environment Research Council (NERC) Grant number: NE/P006507/1

References

- Adams, Dean C, and Michael L Collyer. 2016. "On the comparison of the strength of morphological integration across morphometric datasets." *Evolution* 70 (11): 2623–2631.
- Alvain, Séverine, Cyril Moulin, Yves Dandonneau, and François-Marie Bréon. 2005. "Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery." *Deep Sea Research Part I: Oceanographic Research Papers* 52 (11): 1989–2004.

- Alvain, Séverine, Cyril Moulin, Yves Dandonneau, and Hubert Loisel. 2008. "Seasonal distribution and succession of dominant phytoplankton groups in the Global Ocean: A satellite view." *Global Biogeochemical Cycles* 22 (3).
- Ardyna, Mathieu, Marcel Babin, Michel Gosselin, Emmanuel Devred, Luc Rainville, and Jean-Éric Tremblay. 2014. "Recent Arctic Ocean sea ice loss triggers novel fall phytoplankton blooms." *Geophysical Research Letters* 41 (17): 6207–6212.
- Arrigo, Kevin R, and Gert L van Dijken. 2015. "Continued increases in Arctic Ocean primary production." *Progress in Oceanography* 136:60–70.
- Astoreca, Rosa, Véronique Rousseau, Kevin Ruddick, Cécile Knechciak, Barbara Van Mol, Jean-Yves Parent, and Christiane Lancelot. 2008. "Development and application of an algorithm for detecting *Phaeocystis globosa* blooms in the Case 2 Southern North Sea waters." *Journal of Plankton Research* 31 (3): 287–300.
- Babin, Marcel, Kevin Arrigo, Simon Bélanger, Marie-Helene Forget, et al. 2015. "Ocean Colour Remote Sensing in Polar Seas."
- Babin, Marcel, Dariusz Stramski, Giovanni M Ferrari, Herve Claustre, Annick Bricaud, Grigor Obolensky, and Nicolas Hoepffner. 2003. "Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe." *Journal of Geophysical Research: Oceans* 108 (C7).
- Behrenfeld, Michael J, Robert T O'Malley, David A Siegel, Charles R McClain, Jorge L Sarmiento, Gene C Feldman, Allen J Milligan, Paul G Falkowski, Ricardo M Letelier, and Emmanuel S Boss. 2006. "Climate-driven trends in contemporary ocean productivity." *Nature* 444 (7120): 752–755.
- Bélanger, Simon, Jens K Ehn, and Marcel Babin. 2007. "Impact of sea ice on the retrieval of water-leaving reflectance, chlorophyll a concentration and inherent optical properties from satellite ocean color data." *Remote Sensing of Environment* 111 (1): 51–68.
- Borchers, Hans W. 2018. "pracma: Practical Numerical Math Functions." R package version 2.2.2, <https://CRAN.R-project.org/package=pracma>.
- Brewin, Robert JW, Nick J Hardman-Mountford, Samantha J Lavender, Dionysios E Raitos, Takafumi Hirata, Julia Uitz, Emmanuel Devred, Annick Bricaud, Aurea Ciotti, and Bernard Gentili. 2011. "An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing." *Remote Sensing of Environment* 115 (2): 325–339.
- Bricaud, Annick, Marcel Babin, André Morel, and Hervé Claustre. 1995. "Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization." *Journal of Geophysical Research: Oceans* 100 (C7): 13321–13332.
- Bricaud, Annick, Andre Morel, and Louis Prieur. 1981. "Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains1." *Limnology and Oceanography* 26 (1): 43–53.
- Carlowicz, Mike, Kuring Norman, and Sergio Signorini. 2014. "Dueling Blooms NASA Earth Observatory." Accessed: 2019-04-17, <https://earthobservatory.nasa.gov/images/84135/dueling-blooms>.
- Carlowicz, Mike, Joshua Stevens, and Lauren Dauphin. 2018. "Summer Blooms in the Baltic and Barents." Accessed: 2020-03-03, <https://earthobservatory.nasa.gov/images/92462/summer-blooms-in-the-baltic-and-barents>.
- Carmack, Eddy, and Paul Wassmann. 2006. "Food webs and physical-biological coupling on pan-Arctic shelves: unifying concepts and comprehensive perspectives." *Progress in Oceanography* 71 (2-4): 446–477.
- Cavalieri, Donald J, and Claire L Parkinson. 2012. "Arctic sea ice variability and trends, 1979-2010." *The Cryosphere* 6 (4): 881.
- Cebeci, Zeynel, Figen Yildiz, A. Tuna Kavlak, Cagatay Cebeci, and Hasan Onder. 2019. "ppclust: Probabilistic and Possibilistic Cluster Analysis." R package version 0.1.2, <https://CRAN.R-project.org/package=ppclust>.

- Chase, Alison, Emmanuel Boss, Ronald Zaneveld, Annick Bricaud, Herve Claustre, Josephine Ras, Giorgio Dall’Olmo, and Toby K Westberry. 2013. “Decomposition of in situ particulate absorption spectra.” *Methods in Oceanography* 7:110–124.
- Contreras, Jorge E, Camilo Rodriguez-Valentino, Mauricio F Landaeta, Guido Plaza, Manuel I Castillo, and Mónica Alvarado-Niño. 2017. “Growth and mortality of larval anchoveta *Engraulis ringens*, in northern Chile during winter and their relationship with coastal hydrographic conditions.” *Fisheries Oceanography* 26 (6): 603–614.
- Cota, Glenn F, W Glen Harrison, Trevor Platt, Shubha Sathyendranath, and Venetia Stuart. 2003. “Bio-optical properties of the Labrador Sea.” *Journal of Geophysical Research: Oceans* 108 (C7).
- Degerlund, Maria, and Hans Christian Eilertsen. 2010. “Main species characteristics of phytoplankton spring blooms in NE Atlantic and Arctic waters (68–80 N).” *Estuaries and coasts* 33 (2): 242–269.
- Devred, Emmanuel, Shubha Sathyendranath, Venetia Stuart, and Trevor Platt. 2011. “A three component classification of phytoplankton absorption spectra: Application to ocean-color data.” *Remote Sensing of Environment* 115 (9): 2255–2266.
- Dierssen, Heidi M. 2010. “Perspectives on empirical approaches for ocean color remote sensing of chlorophyll in a changing climate.” *Proceedings of the National Academy of Sciences* 107 (40): 17073–17078.
- d’Ovidio, Francesco, Silvia De Monte, Séverine Alvain, Yves Dandonneau, and Marina Lévy. 2010. “Fluid dynamical niches of phytoplankton types.” *Proceedings of the National Academy of Sciences* 107 (43): 18366–18370.
- Drozdova, AN, SV Patsaeva, and DA Khundzhua. 2017. “Fluorescence of dissolved organic matter as a marker for distribution of desalinated waters in the Kara Sea and bays of Novaya Zemlya archipelago.” *Oceanology* 57 (1): 41–47.
- Dutkiewicz, Stephanie, Anna E Hickman, Oliver Jahn, Stephanie Henson, Claudie Beaulieu, and Erwan Monier. 2019. “Ocean colour signature of climate change.” *Nature communications* 10 (1): 578.
- Fragoso, Glaucia M, Alex J Poulton, Igor M Yashayaev, Erica JH Head, and Duncan A Purdie. 2017. “Spring phytoplankton communities of the Labrador Sea (2005–2014): pigment signatures, photophysiology and elemental ratios.” *Biogeosciences* 14:1235–1259.
- Francis, Jennifer A, and Elias Hunter. 2007. “Drivers of declining sea ice in the Arctic winter: A tale of two seas.” *Geophysical Research Letters* 34 (17).
- Garrido, José L, and Manuel Zapata. 1998. “Detection of new pigments from *Emiliana Huxleyi* (prymnesiophyceae) by high-performance liquid chromatography, liquid chromatography–mass spectrometry, visible spectroscopy, and fast atom bombardment mass spectrometry.” *Journal of phycology* 34 (1): 70–78.
- Goodall, Colin. 1991. “Procrustes methods in the statistical analysis of shape.” *Journal of the Royal Statistical Society: Series B (Methodological)* 53 (2): 285–321.
- Gregg, Watson W, Cécile S Rousseaux, and Bryan A Franz. 2017. “Global trends in ocean phytoplankton: a new assessment using revised ocean colour data.” *Remote Sensing Letters* 8 (12): 1102–1111.
- GSFC, MODIS Rapid Response Team NASA. 2003. “Spectacular Bloom in the Barents Sea.” Accessed: 2020-03-03, <https://earthobservatory.nasa.gov/images/3649/spectacular-bloom-in-the-barents-sea>.
- Hancke, Kasper, Erlend K Hovland, Zsolt Volent, Ragnhild Pettersen, Geir Johnsen, Mark Moline, and Egil Sakshaug. 2014. “Optical properties of CDOM across the Polar Front in the Barents Sea: Origin, distribution and significance.” *Journal of Marine Systems* 130:219–227.
- Hansen, Kathryn, Jeff Schmaltz, and Joshua Stevens. 2016. “The Barents Sea Abloom.” Accessed: 2020-03-03, <https://earthobservatory.nasa.gov/images/88316/the-barents-sea-abloom>.

- Hoepffner, Nicolas, and Shubha Sathyendranath. 1992. "Bio-optical characteristics of coastal waters: Absorption spectra of phytoplankton and pigment distribution in the western North Atlantic." *Limnology and Oceanography* 37 (8): 1660–1679.
- . 1991. "Effect of pigment composition on absorption properties of phytoplankton." *Marine Ecology Progress Series* 73 (1): 1–23.
- Hovland, Erlend Kjeldsberg, Heidi M Dierssen, Ana Sofia Ferreira, and Geir Johnsen. 2013. "Dynamics regulating major trends in Barents Sea temperatures and subsequent effect on remotely sensed particulate inorganic carbon." *Marine Ecology Progress Series* 484:17–32.
- Hovland, Erlend Kjeldsberg, Kasper Hancke, Morten Omholt Alver, Ken Drinkwater, Jo Høkedal, Geir Johnsen, Mark Moline, and Egil Sakshaug. 2014. "Optical impact of an *Emiliana huxleyi* bloom in the frontal region of the Barents Sea." *Journal of Marine Systems* 130:228–240.
- Hu, Chuanmin, Jennifer Cannizzaro, Kendall L Carder, Frank E Muller-Karger, and Robert Hardy. 2010. "Remote detection of *Trichodesmium* blooms in optically complex coastal waters: Examples with MODIS full-spectral data." *Remote Sensing of Environment* 114 (9): 2048–2058.
- Jackson, Thomas, Heather A Bouman, Shubha Sathyendranath, and Emmanuel Devred. 2010. "Regional-scale changes in diatom distribution in the Humboldt upwelling system as revealed by remote sensing: implications for fisheries." *ICES Journal of Marine Science* 68 (4): 729–736.
- Jackson, Thomas, Shubha Sathyendranath, and Frédéric Mélin. 2017. "An improved optical classification scheme for the Ocean Colour Essential Climate Variable and its applications." *Remote Sensing of Environment* 203:152–161.
- Johnsen, G, and E Sakshaug. 2000. "Monitoring of harmful algal blooms along the Norwegian coast using bio-optical methods." *South African Journal of Marine Science* 22 (1): 309–321.
- Johnsen, Geir, and Egil Sakshaug. 2007. "Biooptical characteristics of PSII and PSI in 33 species (13 pigment groups) of marine phytoplankton, and the relevance for pulse-amplitude-modulated and fast-repetition-rate fluorometry 1." *Journal of Phycology* 43 (6): 1236–1251.
- Johnsen, Geir, Oddgeir Samset, Lois Granskog, and Egil Sakshaug. 1994. "In vivo absorption characteristics in 10 classes of bloom-forming phytoplankton: taxonomic characteristics and responses to photoadaptation by means of discriminant and HPLC analysis." *Marine Ecology Progress Series*, 149–157.
- Kirk, John TO. 1994. *Light and photosynthesis in aquatic ecosystems*. Cambridge University Press.
- Kishino, M, M Takahashi, N Okami, and S Ichimura. 1985. "Estimation of the spectral absorption coefficients of phytoplankton in the sea." *Bulletin of Marine Science* 37 (2): 634–642.
- Kurekin, AA, PI Miller, and HJ Van der Woerd. 2014. "Satellite discrimination of *Karenia mikimotoi* and *Phaeocystis* harmful algal blooms in European coastal waters: Merged classification of ocean colour data." *Harmful Algae* 31:163–176.
- Kwok, Ron. 2018. "Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018)." *Environmental Research Letters* 13 (10): 105005.
- Lee, Zhong-Ping, et al. 2006. *Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications*. International Ocean Colour Coordinating Group (IOCCG).
- Li, Aishao, Diane K Stoecker, and Jason E Adolf. 1999. "Feeding, pigmentation, photosynthesis and growth of the mixotrophic dinoflagellate *Gyrodinium galatheanum*." *Aquatic Microbial Ecology* 19 (2): 163–176.

- Liu, Yangyang, Emmanuel Boss, Alison Chase, Hongyan Xi, Xiaodong Zhang, Rüdiger Röttgers, Yanqun Pan, and Astrid Bracher. 2019. "Retrieval of phytoplankton pigments from underway spectrophotometry in the Fram Strait." *Remote Sensing* 11 (3): 318.
- Margalef, R. 1978. "Life forms of phytoplankton as survival alternatives in an unstable environment." *Oceanologica Acta* 1:4–493.
- Meslard, Florian, François Bourrin, Gaël Many, and Philippe Kerhervé. 2018. "Suspended particle dynamics and fluxes in an Arctic fjord (Kongsfjorden, Svalbard)." *Estuarine, Coastal and Shelf Science* 204:212–224.
- Millie, DF, O Schofield, GJ Kirkpatrick, G Johnsen, PA Tester, and BT Vinyard. 1997. "Phytoplankton pigments and absorption spectra as potential 'biomarkers' for harmful algal blooms: A case study of the Florida red-tide dinoflagellate, *Gymnodinium breve*." *Limnol. Oceanogr* 42:1240–1251.
- Mobley, CD, and LK Sundman. 2012. "HydroLight 5.1 EcoLight 5.1 Users' Guide, Sequoia Scientific." *Inc., Bellevue*.
- Morel, André, and Annick Bricaud. 1981. "Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton." *Deep Sea Research Part A. Oceanographic Research Papers* 28 (11): 1375–1393.
- Neukermans, Griet, Laurent Oziel, and Marcel Babin. 2018. "Increased intrusion of warming Atlantic water leads to rapid expansion of temperate phytoplankton in the Arctic." *Global change biology* 24 (6): 2545–2553.
- Nöthig, Eva-Maria, Astrid Bracher, Anja Engel, Katja Metfies, Barbara Niehoff, Ilka Peeken, Eduard Bauerfeind, Alexandra Cherkasheva, Steffi Gäbler-Schwarz, Kristin Hardge, et al. 2015. "Summertime plankton ecology in Fram Strait- a compilation of long and short-term observations." *Polar Research* 34 (1): 23349.
- OBPG. 2019. "Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Daily L3SMI Ocean Colour Data." *NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Daily L3SMI Ocean Colour Data Data; NASA OB.DAAC, Greenbelt, MD, USA., <https://doi.org/doi:10.5067/AQUA/MODIS/L3SMI/accessedon11/04/2019..>*
- . 2020. "NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Visible and Infrared Imager/Radiometer Suite (VIIRS) Daily L3SMI Ocean Colour Data." *NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Visible and Infrared Imager/Radiometer Suite (VIIRS) Daily L3SMI Ocean Colour Data Data; NASA OB.DAAC, Greenbelt, MD, USA., <https://doi.org/doi:10.5067/NPP/VIIRS/L3M/RRS/2018accessedon19/08/2020..>*
- Osman, Matthew B, Sarah B Das, Luke D Trusel, Matthew J Evans, Hubertus Fischer, Mackenzie M Grieman, Sepp Kipfstuhl, Joseph R McConnell, and Eric S Saltzman. 2019. "Industrial-era decline in subarctic Atlantic productivity." *Nature* 569:551–555.
- Overland, James, Edward Dunlea, Jason E Box, Robert Corell, Martin Forsius, Vladimir Kattsov, Morten Skovgård Olsen, Janet Pawlak, Lars-Otto Reiersen, and Muyin Wang. 2018. "The urgency of Arctic change." *Polar Science*.
- Oziel, Laurent, Alberto Baudena, Mathieu Ardyna, Philippe Massicotte, A Randelhoff, J-B Sallée, Randi Brunvør Ingvaldsen, Emmanuel Devred, and Marcel Babin. 2020. "Faster Atlantic currents drive poleward expansion of temperate phytoplankton in the Arctic Ocean." *Nature communications* 11 (1): 1–8.
- Oziel, Laurent, G Neukermans, M Ardyna, C Lancelot, J-L Tison, Paul Wassmann, Jérôme Sirven, Diana Ruiz-Pino, and J-C Gascard. 2017. "Role for Atlantic inflows and sea ice loss on shifting phytoplankton blooms in the Barents Sea." *Journal of Geophysical Research: Oceans* 122 (6): 5121–5139.

- Pavlov, Alexey K, Torbjørn Taskjelle, Hanna M Kauko, Børge Hamre, Stephen R Hudson, Philipp Assmy, Pedro Duarte, Mar Fernández-Méndez, CJ Mundy, and Mats A Granskog. 2017. "Altered inherent optical properties and estimates of the underwater light field during an Arctic under-ice bloom of *Phaeocystis pouchetii*." *Journal of Geophysical Research: Oceans* 122 (6): 4939–4961.
- Peperzak, L, WWC Gieskes, R Duin, and F Colijn. 2000b. "The vitamin B requirement of *Phaeocystis globosa* (Prymnesiophyceae)." *Journal of Plankton Research* 22 (8): 1529–1537.
- Polyakov, Igor V, Andrey V Pnyushkov, Matthew B Alkire, Igor M Ashik, Till M Baumann, Eddy C Carmack, Ilona Goszczko, John Guthrie, Vladimir V Ivanov, Torsten Kanzow, et al. 2017. "Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean." *Science* 356 (6335): 285–291.
- Rat'kova, Tatjana N, and Paul Wassmann. 2002. "Seasonal variation and spatial distribution of phyto-and protozooplankton in the central Barents Sea." *Journal of Marine Systems* 38 (1-2): 47–75.
- Reigstad, Marit, Paul Wassmann, Christian Wexels Riser, Sigrid Øygarden, and Francisco Rey. 2002. "Variations in hydrography, nutrients and chlorophyll a in the marginal ice-zone and the central Barents Sea." *Journal of Marine Systems* 38 (1-2): 9–29.
- Rodriguez, F, M Chauton, G Johnsen, K Andresen, LM Olsen, and M Zapata. 2006. "Photoacclimation in phytoplankton: implications for biomass estimates, pigment functionality and chemotaxonomy." *Marine Biology* 148 (5): 963–971.
- Rohlf, F James, and Marco Corti. 2000. "Use of two-block partial least-squares to study covariation in shape." *Systematic Biology* 49 (4): 740–753.
- Roy, Suzanne, Carole A Llewellyn, Einar Skarstad Egeland, and Geir Johnsen. 2011. *Phytoplankton pigments: characterization, chemotaxonomy and applications in oceanography*. Cambridge University Press.
- Sakshaug, Egil, Geir Helge Johnsen, and Kit M Kovacs. 2009. *Ecosystem Barents Sea*. Tapir Academic Press.
- Sampson, Paul D, Ann P Streissguth, Helen M Barr, and Fred L Bookstein. 1989. "Neurobehavioral effects of prenatal alcohol: Part II. Partial least squares analysis." *Neurotoxicology and Teratology* 11 (5): 477–491.
- Sathyendranath, Shubha, Bob Brewin, Dagmar Mueller, Roland Doerffer, Hajo Krasemann, Frédéric Mélin, Carsten Brockmann, Norman Fomferra, Marco Peters, Mike Grant, et al. 2012. "Ocean colour climate change initiative— Approach and initial results." In *2012 IEEE International Geoscience and Remote Sensing Symposium*, 2024–2027. IEEE.
- Sathyendranath, Shubha, Robert JW Brewin, Carsten Brockmann, Vanda Brotas, Ben Calton, Andrei Chuprin, Paolo Cipollini, André B Couto, James Dingle, Roland Doerffer, et al. 2019. "An Ocean-Colour Time Series for Use in Climate Studies: The Experience of the Ocean-Colour Climate Change Initiative (OC-CCI)." *Sensors* 19 (19): 4285.
- Sathyendranath, Shubha, Louisa Watts, Emmanuel Devred, Trevor Platt, Carla Caverhill, and Heidi Maass. 2004. "Discrimination of diatoms from other phytoplankton using ocean-colour data." *Marine Ecology Progress Series* 272:59–68.
- Smyth, TJ, T Tyrrell, and B Tarrant. 2004. "Time series of coccolithophore activity in the Barents Sea, from twenty years of satellite imagery." *Geophysical Research Letters* 31 (11).
- Stocker, Thomas F, Dahe Qin, Gian-Kasper Plattner, LV Alexander, Simon K Allen, NL Bindoff, FM Bréon, JA Church, U Cubasch, S Emori, et al. 2013. "Technical summary Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change." *Computational Geometry* 18 (2): 95–123.

- Stuart, V, S Sathyendranath, EJH Head, T Platt, B Irwin, and H Maass. 2000. "Bio-optical characteristics of diatom and prymnesiophyte populations in the Labrador Sea." *Marine Ecology Progress Series* 201:91–106.
- Taylor, Isabelle A, James Preston, Elisa Carboni, Tamsin A Mather, Roy G Grainger, Nicolas Theys, Silvana Hidalgo, and Brendan McCormick Kilbride. 2018. "Exploring the Utility of IASI for Monitoring Volcanic SO₂ Emissions." *Journal of Geophysical Research: Atmospheres* 123 (10): 5588–5606.
- Valente, André, Shubha Sathyendranath, Vanda Brotas, Steve Groom, Michael Grant, Malcolm Taberner, David Antoine, Robert Arnone, William M Balch, Kathryn Barker, et al. 2019. "A compilation of global bio-optical in situ data for ocean-colour satellite applications—version two." *Earth System Science Data* 11 (3): 1037–1068.
- Wassmann, P. 1994. "Significance of sedimentation for the termination of *Phaeocystis* blooms." *Journal of Marine Systems* 5 (1): 81–100.
- Wassmann, Paul, Tatjana Ratkova, Inger Andreassen, Maria Vernet, Gunnar Pedersen, and Francisco Rey. 1999. "Spring bloom development in the marginal ice zone and the central Barents Sea." *Marine Ecology* 20 (3-4): 321–346.
- Wassmann, Paul, Maria Vernet, B Greg Mitchell, and Francisco Rey. 1990. "Mass sedimentation of *Phaeocystis pouchetii* in the Barents Sea." *Marine Ecology Progress Series* 66:183–195.
- Zhang, Hailong, Emmanuel Devred, Amane Fujiwara, Zhongfeng Qiu, and Xiaohan Liu. 2018. "Estimation of phytoplankton taxonomic groups in the Arctic Ocean using phytoplankton absorption properties: implication for ocean-color remote sensing." *Optics Express* 26 (24): 32280–32301.

Chapter 4

Bio-Optical evidence for increasing *Phaeocystis* dominance in the Barents Sea

A. Orkney^{1,a}, T. Platt², B.E. Narayanaswamy^{3,b}, I. Kostakis^{4,5,c}, H.A. Bouman^{1,d}

^a <https://orcid.org/0000-0003-4972-2541>

^b <https://orcid.org/0000-0002-5810-9127>

^c <https://orcid.org/0000-0003-1167-2227>

^d <https://orcid.org/0000-0002-7407-9431>

¹ Department of Earth Sciences, University of Oxford,

3 South Parks Road, Oxford OX1 3AN

² Plymouth Marine Laboratory, Prospect Place, Plymouth, PL1 3DH

³ Scottish Association for Marine Science, Scottish Marine Institute, Oban, PA37 1QA

⁴ School of Computing, University of Portsmouth, Portsmouth PO1 3HE, UK

⁵ Physics Department, University of Strathclyde, Glasgow, G4 ONG, UK

4.1 Prelude

This chapter was submitted as a manuscript to Philosophical Transactions of the Royal Society A and published in the ‘Changing Arctic Ocean: Consequences for biological communities, biogeochemical processes and ecosystem functioning’ special issue on the 31st of August 2020:

<https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2019.0357>

The supplementary material can be also be accessed from the above hyperlink.

This chapter makes use of the bio-optical algorithm, developed in chapter 3, to interrogate a decadal-scale time series of ocean-colour data in July in the Barents Sea – a period of the year in which a secular increase in net primary production has been attested. The frequencies of occurrence of different Phytoplankton Functional Groups, and their geographic distributions, are monitored and compared over this period. Convincing evidence of ongoing further dispersal of coccolithophores in Atlantic inflow currents is found, conforming to previous studies and demonstrating the efficacy of our choice of statistical methods. Evidence is found of an increasing predominance of *Phaeocystis* throughout the South and Western Barents Sea. *Phaeocystis* is associated with enhanced cycling within the microbial loop and decreased export efficiency. This finding therefore contributes to increased understanding of the ecological and biological relevance of reported increases in remotely-sensed net primary production. While ocean-colour is effectively the only available means to generate synoptic-scale assessments of baseline ecosystem function in marine systems, it must be remembered that changes in ocean-colour result from a network of many-to-one relationship between ecosystem processes and the Inherent Optical Properties of the water-column.

4.2 Abstract

Increasing contributions of prymnesiophytes such as *Phaeocystis pouchetii* and *Emiliania huxleyi*, both of which possess chlorophyll-c3, to Barents Sea phytoplankton production have been suggested based on *in-situ* observations of phytoplankton community composition, but the scattered and discontinuous nature of these records confounds simple inference of community change or its relationship to salient environmental variables. However, provided that meaningful assessments of phytoplankton community composition can be inferred based on their optical characteristics, ocean-colour records offer a potential means to develop a synthesis between sporadic *in-situ* observations. Existing remote-sensing algorithms to retrieve Phytoplankton Functional Types based on [*chl-a*] or indices of pigment packaging may however fail to distinguish *Phaeocystis* from other blooms of phytoplankton with high pigment packaging, such as diatoms. Moreover, variation in the photo-acclimation state of phytoplankton may result in variation in pigment packaging that is not indicative of community compositional change, meaning that ancillary measures are needed to constrain inferred community composition. We develop a novel algorithm to distinguish major Phytoplankton Functional Types in the Barents Sea and apply it to the MODIS-Aqua ocean-colour record, to study changes in the composition of Barents Sea phytoplankton blooms in July, between 2002 and 2018, creating time series of the spatial distribution and intensity of coccolithophore, diatom and *Phaeocystis* blooms. We confirm a North-eastward expansion in coccolithophore bloom distribution, identified in previous studies, and suggest an inferred increase in [*chl-a*], reported by previous researchers, may be partly explained by increasing frequencies of *Phaeocystis* blooms.

4.3 Plain Language Summary

Previous research has suggested that phytoplankton assemblages in the Barents Sea have been shifting from large diatoms to smaller prymnesiophytes often more commonly associated with temperate waters, such as *Phaeocystis pouchetii* and the coccolithophore *Emiliana huxleyi*. There have also been reports in the literature of increasing concentrations of chlorophyll-a in the Barents Sea in mid-summer, and some researchers believe this indicates that overall photosynthesis is increasing and therefore provision of food to the ecosystem. These two positions may be difficult to resolve, however, because smaller phytoplankton species such as *Phaeocystis pouchetii* are often associated with smaller contributions of new biomass to the ecosystem and may not sink as rapidly as large heavy diatoms, limiting their ability to nourish organisms dwelling on the sea bottom.

I demonstrated, in chapter 3, that it should be possible to detect the presence of these different assemblages (diatoms, *Phaeocystis* colonies, coccolithophores) in remotely-sensed ocean-colour. Thus, I constructed a record of these phytoplankton groups, inferred from remotely-sensed ocean-colour, for each July between 2002 and 2018, to investigate further. I was able to show that increases in remotely-sensed chlorophyll-a concentrations in the Barents Sea are partly driven by an increasing frequency of *Phaeocystis* blooms. This conclusion therefore emphasises that changes in chlorophyll-a concentration, inferred from ocean-colour, must be interpreted carefully. Ongoing changes in phytoplankton community composition in the Barents Sea may mean that detected increases in chlorophyll-a cannot be assumed to indicate greater nourishment of this rapidly changing Arctic shelf.

4.4 Introduction

The Barents Sea extends North of mainland Arctic Norway, spanning the continental shelf region between the Svalbard Archipelago, Franz Josef land and Novaya Zemlya, occupying

some 1.4 million km², with a complex hydrography that results from the convergence of warmer more saline Atlantic and fresher, cooler Arctic Waters (Reigstad et al. 2002). The Barents Sea is predominantly shallow, averaging only ~230 m depth. This shallow topography means that the movements of water-masses are guided by bathymetry such as the Barents Sea Opening (~470 m deep), which serves as a gateway for the delivery of warmer more saline and nutrient-rich Atlantic Water onto the shelf (Oziel et al. 2016). These ocean dynamics are illustrated in Figure 4.1.

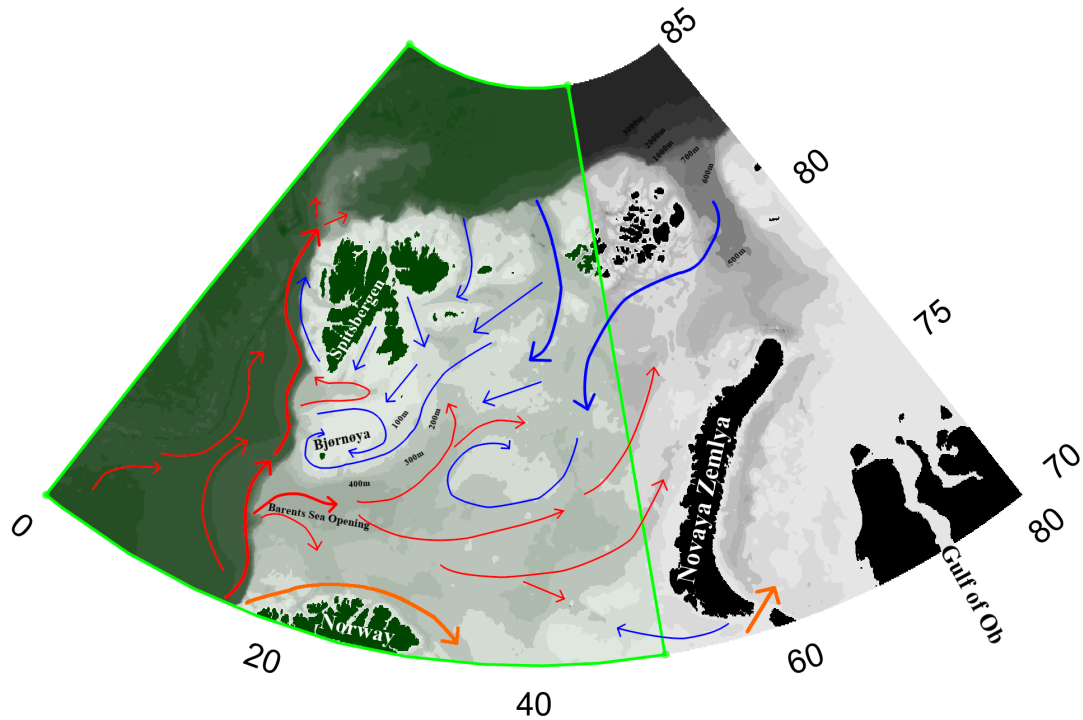


Figure 4.1: A bathymetric map of the Barents Sea, with inflowing Atlantic (red), Norwegian Coastal Current (orange) and Arctic (blue) water-masses. A front forms between the warm and cold water-masses. The area of interest in this study is outlined in green. Bathymetric data was sourced using the ETOPO1 NOAA global relief database (<https://www.ngdc.noaa.gov/mgg/global/>) with the R *marmap* package (Pante and Simon-Bouhet 2013).

Bathymetric features stabilise the position of the boundary between Atlantic and Arctic

water-masses, known as the Polar Front, which means that the front's position remains relatively constant year-on-year, insensitive to inter-annual variation in the flux of Atlantic Water caused by climatological variability, such as the Arctic Oscillation (Oziel et al. 2016). Over the last 30 years, however, a sustained increase in the baseline rate of Atlantic Water flux into the Arctic has doubled the volume occupied by Atlantic Water in the Barents Sea, causing the Northward displacement of the Polar Front (Francis and Hunter 2007; Oziel et al. 2016). As the Polar Front delimits the maximum extent of sea-ice cover (Barton et al. 2020), its retreat has resulted in the recession of average sea-ice extent in the Barents Sea, which has occurred at a disproportionately rapid rate compared with the rest of the Arctic (Cavalieri and Parkinson 2012).

The hydrography and climatology of the Barents Sea create an environment conducive to massive phytoplankton blooms, and the inter-annual variation in these environmental properties determine which phytoplankton groups will be most successful from year-to-year.

The surface ocean is charged with nutrients in the winter when dense Atlantic Waters are brought to the surface by vertical mixing (Loeng 1991) and this nutrient complement is further supplemented in spring by mineral-rich ice-rafted debris that is released by melting sea-ice imported through the Fram Strait region (Krumpen et al. 2019). Thermal stratification of the water-column in spring, and salinity-based stratification associated with melting sea-ice, create a nutrient-replete sunlit surface layer that can support phytoplankton blooms in which $[chl-a]$ exceed 10 mg m^{-3} (Båmstedt et al. 1991).

Reviews of the physical set-up of the Barents Sea, and its consequences for phytoplankton phenology, can be found in Sakshaug et al. (2009) chapters 2 (ocean physical setup) and 7 (phytoplankton and primary production), and in products of the Nansen Legacy project, such as Dong et al. (2020).

Historically, these blooms have usually been dominated by large-celled diatom gen-

era such as *Chaetoceros*, *Fragilariopsis* and *Thalassiosira* that can reproduce very rapidly when resources are abundant (Wassmann et al. 1999) and have been observed to form ice-edge blooms that follow the receding Marginal Ice Zone through summer (Carmack and Wassmann 2006). Outside of the Seasonal Ice Zone, in waters where perennial freezing does not occur, summer mixing tends to be deeper and periodic recharging of the surface nutrient reservoir may occur, favouring smaller phytoplankton species with larger proportional surface areas, such as *Emiliana huxleyi* and *Phaeocystis pouchetii* (Wassmann et al. 1999). These species have a long established presence in the Barents Sea (see Sakshaug et al. (2009) chapter 7), and may become increasingly dominant in future. Indeed there is already remote-sensing evidence of increasing abundances of *Emiliana huxleyi* in historically Arctic-influenced waters (Oziel et al. 2020). *Phaeocystis* may also have the capacity to adapt to variable light environments, and appears to be prospering in the under-ice environment beneath a thinner and more dynamic ice-pack (Assmy et al. 2017; Kauko et al. 2017; Pavlov et al. 2017). In addition, to haptophytes, there is chemotaxonomic evidence of an over-looked biomass of picoeukaryotic prasinophytes in the Barents Sea (Pettersen et al. 2011), such as *Micromonas pusilla*. Their small cell-sizes, and possible mixotrophy, may leave them well-placed to prosper in a Barents Sea with an increasingly Atlantic-influenced physical regime.

We should therefore expect that environmental changes in the Barents Sea, caused by human-induced warming, will influence the community composition of phytoplankton blooms, with ramifications for overall new production, phenology and hence the supply of food to higher trophic levels in this great Northern fishery, and influencing the ecosystem services that are rendered.

Indeed, *in-situ* investigations in the Eurasian Basin, during the summer 2012 sea-ice minimum, revealed a rare mass-sedimentation of diatomaceous ice-algae filaments, consis-

tent with the expectation of authors such as Krumpen et al. (2019) that thinner sea-ice, carrying lower levels of exogenous sediment, should transmit more sunlight and hence permit greater under-ice primary production. This carries potential implications for nutrient cycling mechanisms and carbon sequestration in a future Arctic Ocean that may be more conducive to perennial ice-algal sedimentation (Boetius et al. 2013).

In the pelagic realm, some authors have reported that coccolithophore blooms are migrating Eastward, further into historically more Arctic-influenced areas (Oziel et al. 2017), and sampling in the Fram Strait has revealed an increasing predominance of *Phaeocystis pouchetii* and *Emiliana huxleyi* (Nöthig et al. 2015). Average annual primary production, inferred from remotely-sensed [*chl-a*], has increased by 30% between 1998 and 2012 (Arrigo and Dijken 2015), but this picture of increasing primary production is difficult to reconcile with observations of declining end-winter silicic acid inventories in the Barents Sea (Rey 2012), critical for diatoms to build their frustules, and with studies of biochemical proxies in Greenlandic ice-cores that imply a centennial scale decline in the productivity of subarctic seas (Osman et al. 2019).

In addition, there is an increasing incidence of reports of under-ice phytoplankton blooms, both pelagic in origin (and being advected underneath the sea-ice (Johnsen et al. 2018)) and indigenous to the under-ice environment, where they are possibly encouraged by decreasing ice thickness and higher transparency (Assmy et al. 2017).

Remote-sensing inference of a decadal increase in Arctic dimethyl-sulphide (DMS) production (Galí et al. 2019) – often associated with prymnesiophyte species that do not require silicic acid, such as *Phaeocystis pouchetii* or *Emiliana huxleyi*, suggests that change in community composition may help explain these seemingly conflicting findings – although the algorithm of Galí et al. (2019) is based on comparisons of estimated productivity and inferred Mixed Layer Depth, rather than the recognition of optical signatures of community composition.

Time series of phytoplankton samples in the Barents Sea necessary to test the hypothesis of community compositional change are, however, sporadic and usually do not exceed the required decadal scales (Degerlund and Eilertsen 2010). We therefore may ask whether a taxonomically-informed study of changes in ocean-colour over the 21st century, might help elucidate the matter.

There already exist several approaches towards this end. Remote-sensing of ocean-colour has successfully detected coccolithophore blooms (Smyth et al. 2004; Hovland et al. 2013; Oziel et al. 2017), owing to the strong scattering properties of calcite liths. However, the detection of other Phytoplankton Functional Types has proved more challenging, although numerous attempts have been made; such as semi-analytical approaches based on forward-modelling from the optical properties of different phytoplankton groups to arrive at ocean-colour signatures that may distinguish them (Sathyendranath et al. 2004; Jackson et al. 2010) or the inversion of systems of equations relating ocean-colour and phytoplankton size classes retrieved with HPLC (High Performance Liquid Chromatography) (Devred et al. 2011). Some authors have explored linear unmixing models – another form of inversion – of phytoplankton light absorption (Zhang et al. 2018) to infer community composition. However, the capacity to infer community composition from Inherent Optical Properties measured *in-situ* is a distinct problem from investigating patterns in remotely-sensed ocean-colour, which these methods have not yet been applied to.

There have also been attempts to detect optical signatures of specific phytoplankton pigments (Cullen and Lewis 1995; Johnsen and Sakshaug 2000; Astoreca et al. 2008), which have been demonstrated to be successful with *in-situ* radiometry, but have not yet been extended to remotely-sensed observations.

Some existing remote-sensing algorithms for the inference of Phytoplankton Functional Types initially assume some [*chl-a*] dependences in their construction (Devred et al. 2011), which confounds the task of recognising shifts in Barents Sea bloom composition, given that

the inferred $[chl-a]$ in this region has been increasing, and $[chl-a]$ and community compositional change may therefore be convolved (Oziel et al. 2017; Arrigo and Dijken 2015).

We also recognised that other algorithms to distinguish diatoms in high-latitude waters are predicated on differences in specific absorption coefficient between diatoms and mixed populations (Sathyendranath et al. 2004). Pigment packaging is the phenomenon of the flattening of absorption spectra of suspensions of pigment as distributions change from dispersed to tightly packaged aggregates (Duysens 1956; Kirk 1994). *Phaeocystis* forms aggregated colonies and therefore exhibits a normalised absorption spectrum that is indicative of levels of pigment packaging similar to diatoms (Astoreca et al. 2008) (although photoacclimation (Johnsen and Sakshaug 2007) may result in a wide variation in intra-cellular pigment concentrations and therefore packaging (Rodriguez et al. 2006)), and we therefore may expect its optical properties to resemble those of diatoms. Indeed, this is indicated in Figure 2 of Alvain et al. (2008), who recognised a *Phaeocystis*-like group in ocean-colour measurements in their development upon the PHYSAT algorithm for the inference of different Phytoplankton Functional Types.

Given diatoms and *Phaeocystis* blooms are typically associated with different styles of production (with *Phaeocystis* associated with greater regeneration (Reigstad and Wassmann 2007)), remote-sensing approaches, which treat these two groups as synonymous, may produce results that provide misleading conclusions about ecosystem change. This is a salient observation that bears special consideration in the Barents Sea where gradients in remotely-sensed production anti-correlate with the mass of benthic megafauna (Degen et al. 2016), suggesting that the direct interpretation of remotely-sensed production as an indicator for export production is fraught.

We reviewed the the global PHYSAT algorithm and found that, while extensive *Phaeocystis*-like blooms are identified in the Southern Ocean, PHYSAT describes the Barents as nanophytoplankton-dominated (Alvain et al. 2008), suggesting that the global PHYSAT al-

gorithm struggles to detect different Barents Sea phytoplankton communities. The ‘*Phaeocystis*-like’ group identified in PHYSAT is based on ocean-colour observations from the Southern Ocean, where *Phaeocystis antarctica* predominates, as opposed to *Phaeocystis pouchetii* in the Arctic, so differences in ocean-colour between these two species may explain more pervasive identification of *Phaeocystis* in the Southern Ocean.

We hence opted to construct our own experimental bio-optical algorithm tuned to the bespoke Phytoplankton Functional Types (coccolithophores, nanophytoplankton, diatoms and *Phaeocystis*) and conditions prevailing in the Barents Sea, to classify ocean-colour observations (see chapter 3, ‘Development of a Bio-Optical algorithm,’ for a detailed description).

In the near future, NASA’s PACE (Plankton, Aerosol, Cloud, Ecosystem) mission aims to measure ocean-colour at a 5 nm spectral resolution (Petro et al. 2020), which may offer researchers the capacity to detect subtle hyper-spectral features in ocean-colour attributable to key taxon-specific phytoplankton pigments, and may therefore facilitate increased confidence in the identification of distinct phytoplankton groups from ocean-colour. However, the legacy of ocean-colour measurements already collected will still represent a rich historic dataset and we should therefore retain an interest in exploring the detection of distinct phytoplankton groups in existing satellite records (e.g. Sathyendranath et al. 2004; Alvain et al. 2005; Devred et al. 2011) and assessing the extent to which bio-optical algorithms can be marshalled to describe ecosystem baselines or detect incipient signals of ecosystem change in our warming climate (Quere et al. 2005; Nair et al. 2008; Sathyendranath et al. 2014; Dutkiewicz et al. 2019), especially as it is this historical record that will place future observations in context.

Here we develop a bio-optical algorithm in the mould of Alvain et al. (2008). We constructed a training dataset of ocean-colour, in order to inform a Linear Discriminant Analysis to identify blooms of phytoplankton with a high degree of pigment packaging (hereafter

referred to as ‘highly-packaged phytoplankton’), such as diatom or *Phaeocystis* blooms. We employ an adaptation of the experimental chlorophyll-c3 algorithm of Astoreca et al. (2008), which compares the inferred absorption by phytoplankton at 469 nm (indicative of the bio-marker chlorophyll-c3) to adjacent wavelengths, to test whether it might also be possible to distinguish *Phaeocystis* from diatom blooms in the Barents Sea and assess whether it is possible to confirm isolated reports of changing phytoplankton community composition and extend our understanding by using ocean-colour to construct a regional synthesis. Previous authors have shown that detection of bio-marker pigments, such as chlorophyll-c3, from *in-situ* absorption spectra can indicate the presence of a range of species that form harmful algal blooms – including *Phaeocystis* (Johnsen and Sakshaug 2000).

4.5 Methods

Level-3 daily binned ocean-colour observations during the month of July 2002–2018 were downloaded from the MODIS-Aqua website, for $\lambda = 412, 443, 469, 488, 531, 547, 555, 645, 667$ and 678 nm (OBPG 2019) between $0\text{--}50^\circ$ East and $70\text{--}85^\circ$ North. We chose to investigate only ocean-colour observations from July, because while annual net primary production has been increasing (Arrigo and Dijken 2015) lower ice extents are associated with decreased $[chl-a]$ biomass in spring (Oziel et al. 2017). We hence chose to study summer blooms alone to avoid the convolution of these antithetical changes.

Our experimental algorithm is a stack of 3 individual algorithm modules, which can be visualised as a flow-chart (see Figure 3.14). The first module masks ocean-colour observations that are classified as being influenced by sediment, Coloured Dissolved Organic Matter (CDOM)-rich river plumes or coccolithophore blooms. These masks were based on MODIS-Aqua ocean-colour observations from July 2014 between $0\text{--}80^\circ$ East and $70\text{--}85^\circ$ North. Modes of ocean-colour variance associated with the Ob river plume – a feature of high CDOM output, the Svalbard coast – high sediment load, and coccolithophore blooms

in the central Barents Sea were identified and used to inform criteria to establish masks (see Figures 3.7, 3.8). Only un-masked ocean-colour observations are passed on to the second module.

The second module applies a Linear Discriminant – trained on a collection of Level-3 daily binned MODIS-Aqua ocean-colour observations, measured on the same day and within 1° latitude or longitude of as a series of natural phytoplankton samples collected in 2017-2018. This LDA is used to distinguish whether ocean-colour properties are compatible with blooms of highly-packaged phytoplankton (either diatom or colonial *Phaeocystis* cells). The use of Level-3 daily data for the purpose of identifying distinct phytoplankton groups follows from the example set by the development of the PHYSAT algorithm, and our 1° criterion follows from that used in PHYSAT validation (Alvain et al. 2005; Alvain et al. 2008). Hold-p-out cross-validation on our training dataset demonstrated that highly-packaged phytoplankton could be distinguished from nanophytoplankton effectively (97 and 80% of diatom and *Phaeocystis*-affiliated ocean-colour observations were assessed to be highly-packaged, and some 68% of nanophytoplankton were correctly identified).

After ocean-colour observations are classified as highly-packaged phytoplankton or as unicellular nanophytoplankton populations, $[chl-a]$ estimates are recalculated, using the **Lab-OC4L diatom** algorithm for all highly-packaged phytoplankton and the **Lab-OC4L prymn** relationship for all nanophytoplankton populations. We justified our choice to use the **Lab-OC4L diatom** algorithm to estimate $[chl-a]$ for all highly-packaged phytoplankton (whether they would subsequently go on to be classified as diatoms or as *Phaeocystis*) because Cota et al. (2003) had described remote-sensing algorithms developed by Arrigo et al. (1998) for the inference of diatom, *Phaeocystis* and (diatom+*Phaeocystis*) communities as ‘very similar, perhaps indistinguishable’ (Cota et al. 2003; Arrigo et al. 1998).

The final module takes all observations classified as blooms of highly-packaged phytoplankton and uses the algorithm developed by Astoreca et al. (2008) to distinguish between

Phaeocystis and diatom blooms. This is done on the basis that Stuart et al. (2000) had used chlorophyll-c3 as a marker pigment to distinguish diatom and prymnesiophyte-dominated samples. Discrimination is informed by a suite of HydroLight simulated ocean-colour reflectances based on optical properties of *in-situ* samples of diatom and *Phaeocystis* communities. It was found the ratio of retrieved [*chl-a*] to inferred chlorophyll-c3 absorption separated these simulations. We recognised that MODIS-Aqua ocean-colour reflectance may not be amenable to the detection of subtle hyperspectral features, so we performed a sensitivity analysis to evaluate the effects of noise and variable band-width on the resolution of this feature, and diatom and *Phaeocystis* HydroLight simulations remained distinct. We found that our discriminant for *Phaeocystis* from diatoms, when applied to our training dataset, corresponded to a bimodal feature in the optical properties of highly-packaged phytoplankton (see Figure 3.13).

We found that a significant shift in mean retrieved chlorophyll-c3 signal (we use the term ‘signal’ rather than ‘absorption’ here because this metric sometimes takes negative values. For example, Astoreca et al. (2008) recorded a diatom bloom with an apparent negative chlorophyll-c3 absorption) occurred after 2009 in disparate parts of the global ocean (see Figures 3.15, 3.16) and adapted the method of Taylor et al. (2018), who corrected for drift in the output of the IASI (Infrared Atmospheric Sounding Interferometer) platform. Corrected chlorophyll-c3 absorption signal was then used in module 3 of our algorithm to infer the presence or absence of *Phaeocystis* blooms. The bio-optical algorithm was validated against a yearly time series of spring phytoplankton samples in the Labrador Sea, reported by Fragoso et al. (2017)

(dataset accessible at <https://doi.pangaea.de/10.1594/PANGAEA.871872>). A linear regression between the yearly *Phaeocystis* proportion of *in-situ* samples, collected by Fragoso et al. (2017) (defined as the portion of [*chl-a*] associated with *Phaeocystis* divided by the

sum of $[chl-a]$ associated with either *Phaeocystis* or diatoms) shows a highly significant correlation with the yearly *Phaeocystis* proportion in ocean-colour classifications (defined as the ratio of *Phaeocystis* to the sum of *Phaeocystis* and diatom classifications) (Figure 4.2). The relationship between the inter-annual differences of these two curves is also highly significant. None of the curves show a significant linear trend over the time period 2005-2014.

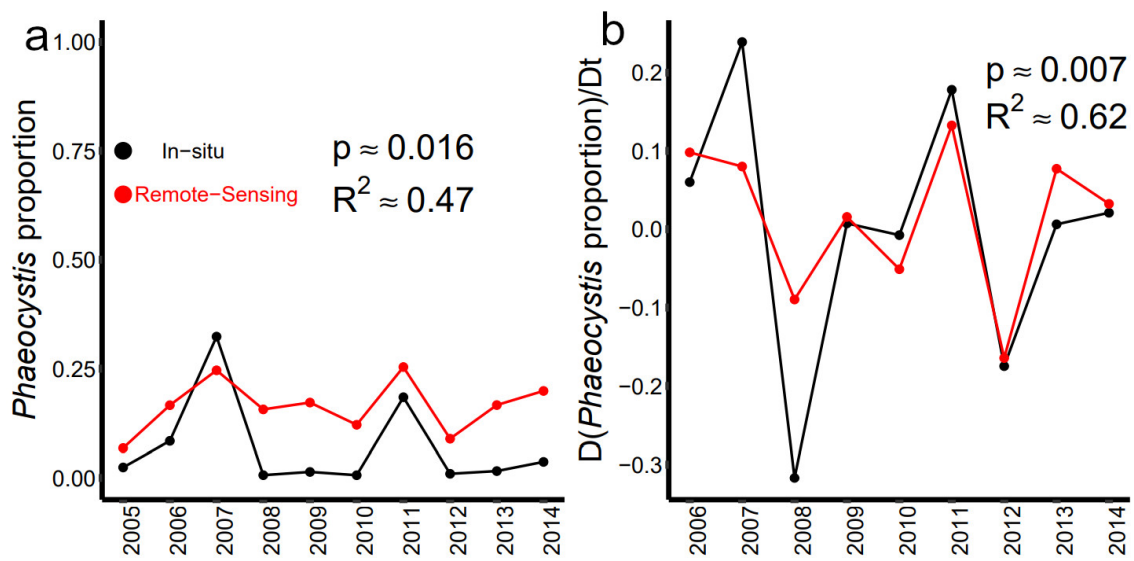


Figure 4.2: a) Year-to-year variation in the dominance of *Phaeocystis* blooms in the *in-situ* observations of Fragoso et al. (2017) (black line) and from ocean-colour classifications across the Labrador Sea (red line). *in-situ* estimates of the *Phaeocystis* proportion were sought from chemotaxonomic analysis of pigments performed by Fragoso et al. (2017). b) Year-to-year variation in the inter-annual difference of *Phaeocystis* dominance.

We noted that the *Phaeocystis* dominance estimated across the Labrador Sea from remote-sensing was often biased high, compared to the *in-situ* samples (Figure 4.2a). This may be because the dataset of Fragoso et al. (2017) does not contain observations along the East-Greenland coast, which we found was often pervaded by *Phaeocystis* classifications (see Figure 3.17 years 2007-8, 2010-12). We believe we have demonstrated that our bio-optical

algorithm facilitates the construction of time series of diatom and *Phaeocystis* blooms that show biogeographical patterns consistent with *in-situ* observations in another Subarctic Atlantic system.

4.6 Results

We first tested for changes in the geographic distribution of phytoplankton blooms with time. We discuss the example of coccolithophores to clarify our approach. We controlled for changes in the geography of the sampled area, as a result of the inter-annual recession in sea-ice extents for example, by excluding pixels poleward of 80° North – the approximate limit of ice at the start of our time series. We also excluded observations in the box between $10\text{--}30^\circ$ East and $76\text{--}80^\circ$ North, because the coastal regions of Svalbard are often identified as coccolithophore-dominated. It is not clear whether these classifications are spurious, but even if they are genuine their association with the coast means that they do not reflect any biogeographic changes in open-ocean coccolithophore blooms. We took all the unique pixels in our time series and computed the relative frequency of coccolithophores in each pixel for a given year. A pixel observed once as coccolithophore-dominated is 100% coccolithophore-dominated and a pixel observed four times, but classified as coccolithophore only twice, is 50% coccolithophore-dominated. The accuracy of the frequency estimate increases with greater observations. We then calculated the longitudes and latitudes of the quantiles of the pixel distribution, weighted by the coccolithophore frequencies in each pixel, for 10 equally-spaced values of τ (the quantile being used for quantile regression; e.g. $\tau = 0.5$ if a researcher is interested in estimating the conditional median of the dependent variable) between 0.1 and 0.9, corresponding to the 10th and 90th quantiles of the distribution. We performed this exercise for each year in the time series. We then took our time series of latitudes and longitudes for each quantile value and performed a linear regression with time, weighted

by the number of coccolithophore observations in each year, so that a year with a strange geographic distribution but few observations would not exert undue influence. Finally, we inverted the linear models to predict changes over time in latitude and longitude for each value of τ . We built multiple linear regressions for all latitude-longitude combinations of varying τ against time, and assessed their significance by calculating p -values. The result is a field of vectors, which visually describe the change in coccolithophore distribution with time (Figure 4.3). The most North-eastern vector corresponds to the model for change in geography of the 90th quantiles for both latitude and longitude, while the most South-western vector corresponds to the model for change in the geography of the 10th quantiles for both latitude and longitude. Coccolithophores were the only phytoplankton community which demonstrated significant changes in biogeography between 2002 and 2018.

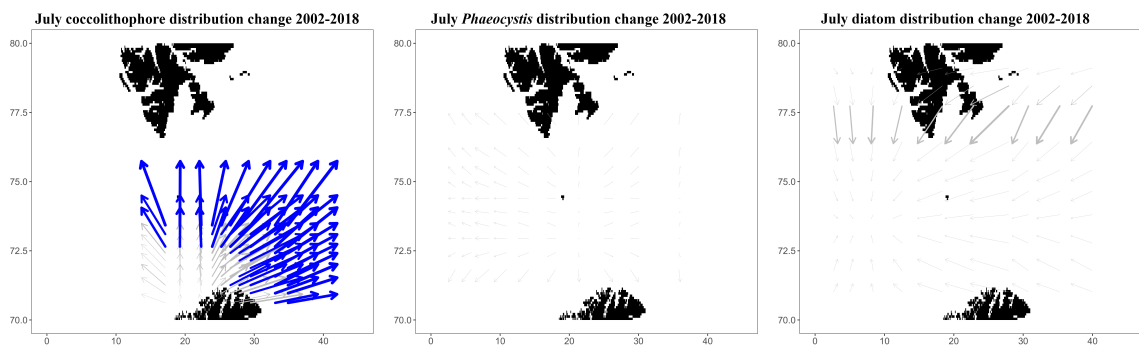


Figure 4.3: **Left:** changing distributions of coccolithophore ocean-colour observations between 2002 and 2018. Significant (p -value < 0.05) vectors are coloured blue. **Centre:** vector field for *Phaeocystis* distribution change, no significant vectors. **Right:** vector field for diatom distribution change, no significant vectors.

The Eastern limit of coccolithophore blooms has shifted some 7° Eastward over the study window, while the Northern limit has expanded about 1.5° poleward, in accord with the findings of previous studies (Oziel et al. 2017; Hovland et al. 2013). Interestingly, Figure 4.3 also shows that the Western limit of coccolithophore blooms has expanded Westward by 4°. The vector field of coccolithophore distribution change in Figure 4.3 implies that

there are distinct Eastern and Western coccolithophore blooms. It appears that the Eastern bloom dominates the distribution and is migrating to the North-east, while a minor branch of Western blooms is migrating towards the Fram Strait region. These two pelagic blooms may be diverging as they migrate North, because the Svalbard archipelago separates the Western and Eastern portions of the Barents Sea, and consequently warm ocean currents bifurcate as they head North, steered by bathymetry (Figure 4.1). Coccolithophore blooms identified by Smyth et al. (2004), between 1987 and 2002, were typically centred between 30 and 40° East and rarely extended beyond 40° East. By contrast, this study shows that it has become routine in recent years for coccolithophore blooms to extend past 40°.

We also constructed time series of the proportion of all pixels in the study area dominated by different phytoplankton communities (Figure 4.4). Hovland et al. (2013) commented that a large mass of Atlantic Water penetrated far into the East of the Barents Sea during 2007, compared with 2010, which was a year with relatively less Atlantic influence. Figure 4.4 shows that 2007 has a greater proportion of coccolithophore detections than 2010, conforming to the results of these earlier studies. Although inter-annual variations in diatom frequency are large, varying between 10% and 25% of observations, no discernible pattern of biogeographic change prevails over the study window and there is no significant linear trend in their occurrence between 2002-2018 (Figure 4.4). Sub-significant patterns in the vector field of Figure 4.3 imply a slight gravitation of diatoms towards the Atlantic inflow region and this may reflect a local increase in diatom identifications here (Figure 4.5). By contrast, *Phaeocystis* frequency is low at the start of the study window, but climbs rapidly in 2010 (Figure 4.4). We find this particularly interesting, given that previous authors found a massive increase in the occurrence of *Phaeocystis* and associated marine gels in the Fram Strait in 2010 (Engel et al. 2017). We also note a positive anomaly in *Phaeocystis* detections in 2006, coinciding with reports of an increase in *Phaeocystis* in the Fram Strait by Nöthig et al. (2015). We consider increasing *Phaeocystis* predominance compatible with the

predictions that Atlantification of the Barents should be expected to increase *Phaeocystis* frequency and result in a Northward expansion of more Atlantic-associated species such as *E. huxleyi* (Rat'kova and Wassmann 2002).

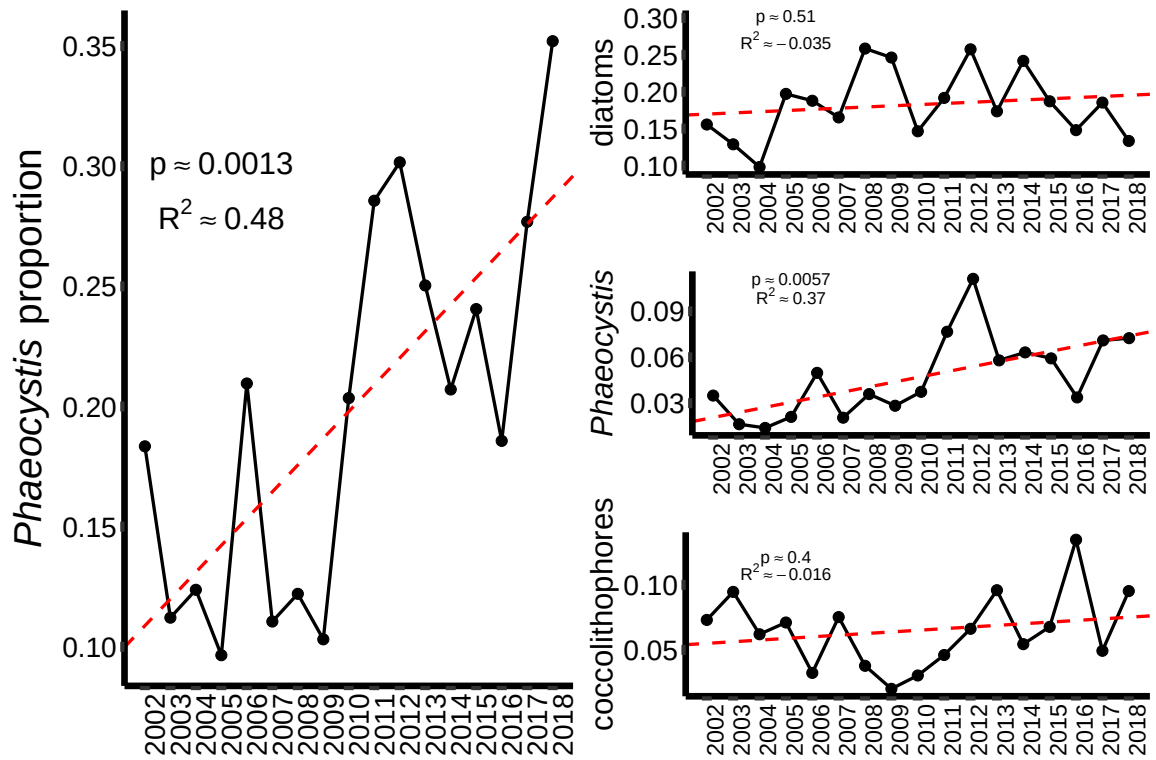


Figure 4.4: Evolution of the dominance of different phytoplankton groups in the Barents Sea in July. *Phaeocystis* proportion is defined as the number of *Phaeocystis* classifications divided by the sum of diatom and *Phaeocystis* classifications. All other subplots record the number of classifications of a specific phytoplankton group as a proportion of the total number of observations. Linear regressions with time are represented by red dashed lines. Coccolithophores were identified by performing a rotation on ocean-colour data and masking measurements above a threshold value. Diatoms were identified by multiplying ocean-colour data by a set of discriminant coefficients designed to represent pigment packaging. *Phaeocystis* was identified by the combination of possessing an ocean-colour signature indicative of high pigment packaging, and a high ratio of inferred chlorophyll-c3 to [*chl-a*].

Given that a pronounced change in phytoplankton community structure was found over the

study period, difference maps showing the change in the frequency of occurrence of different phytoplankton groups were created, so that we could identify the geographic regions driving these patterns (Figure 4.5). We limited the temporal extent to within 2005-2014 to match the period of our validation scheme in the Labrador Sea (Figure 4.2), so that we could have confidence in the results. The difference maps were derived by computing linear models of the variable of interest, such as *[chl-a]* concentration or relative frequency of coccolithophores, in each unique pixel location as a function of time. We calculated the average change in each pixel that the linear models predicted and inferred significance of plankton community shifts from their *p*-values (expressed in Figure 4.5 as ‘Certainty’, defined as $1 - p\text{-value}$).

Figure 4.5 shows that there have been increases in the mean July *[chl-a]* concentration in the Western Barents Sea, Southern Fram Strait and far South-eastern Barents Sea. The largest increases exceed 1 mg m^{-3} , West of Bjørnøya. There is an equivocal pattern of change in the frequency of diatom occurrence; while there are strong increases in the Atlantic inflow region West of Bjørnøya there are decreases further down-stream in the Northern Fram Strait or over the Sentralbanken. *Phaeocystis* increases are also achieved throughout much of the Western and Southern Barents Sea, but are absent beyond the typical position of the Polar Front in the Arctic-derived waters of the North-east. Changes in patterns of coccolithophore occurrence are in agreement with Figure 4.3. The frequency of coccolithophores has declined in the South-west of their geographic range and increased in the North-east. The migration of coccolithophores into the Sentralbanken region may offer an explanation for declines in the frequency of diatom detections here. Curiously, patterns of change in SST data from the ESA SST CCI reprocessed SST analyses, (temperature maps for the 15th of July in each year downloaded from CMEMS; Copernicus Marine Environment Monitoring Service: <http://marine.copernicus.eu/>) did not correspond clearly to changes in phytoplankton community composition. The pattern of change reproduces the regional change

described for 2005-2016 in Figure 2C, Barton et al. (2020), with the exception that there is slight cooling in the inflow region, where Barton et al. (2020) detected only insignificant changes of low magnitude. This suggests that SST evolution in the Barents Sea exhibits seasonal structure. This hypothesis is supported by assessments of SST evolution in chapter 5, which show that the inflow region exhibits either an ambiguous or warming signal in September (see Figure 5.3).

We might tentatively suggest that a broad region of declining SST in the Atlantic inflow region aligns with increasing [*chl-a*] concentrations, while warming regions of the Northern Fram Strait align with declines in diatoms and increases in *Phaeocystis*. Interestingly patterns of temperature change in the Fugløya Atlantic inflow section (corresponding to the Barents Sea Opening: <https://ocean.ices.dk/iroc/>) show that there has been a broad increase in sea temperatures between 50 and 200m between 2002-2018, whereas surface temperature data from CMEMS implies a decline over this same period. This emphasises that consideration of SST in isolation may provide misleading explanations for variation in the state of ecosystems.

The sum of all the changes to phytoplankton community composition that have been found in this study are presented in a synthesis as a single cartoon in Figure 4.6.

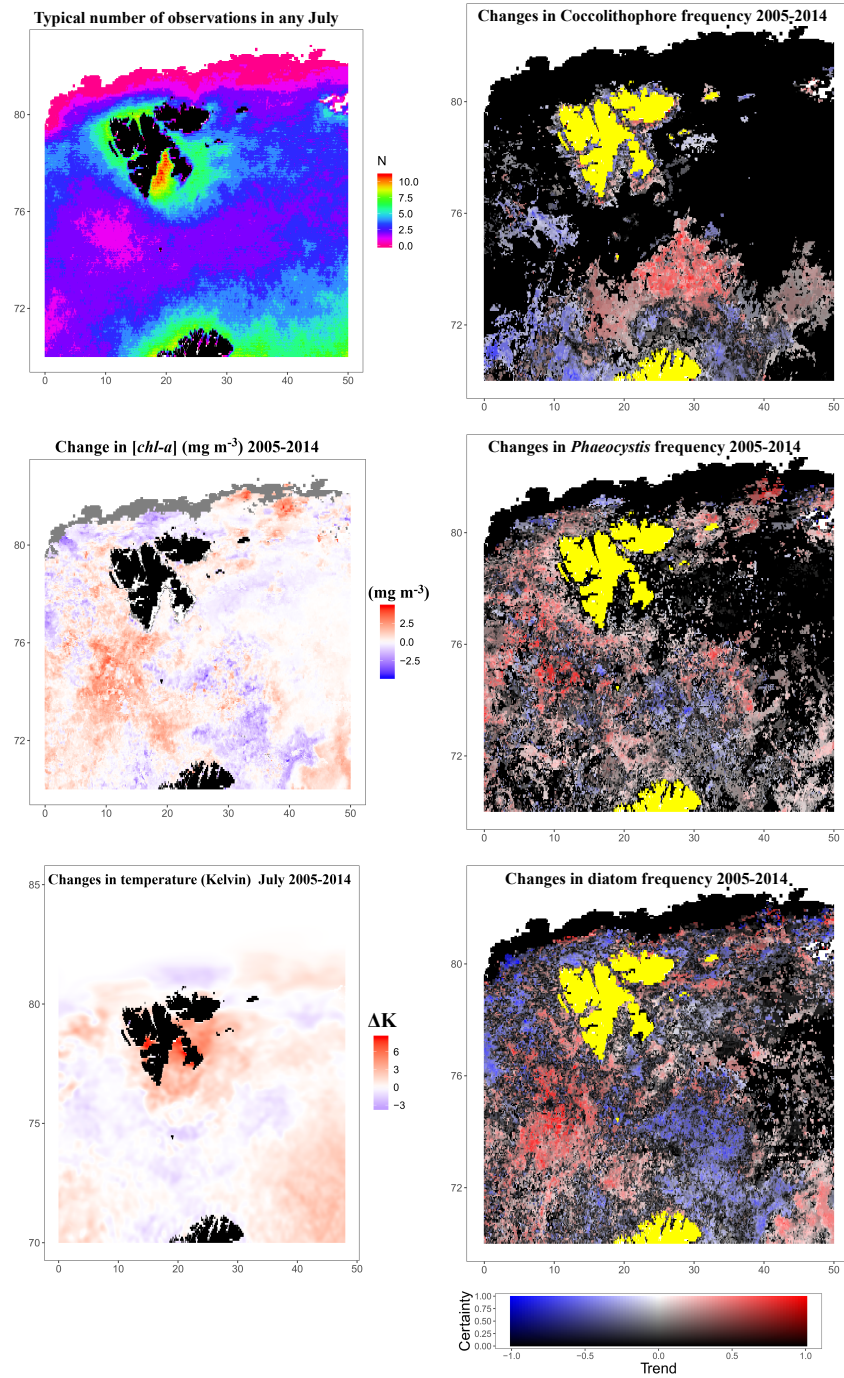


Figure 4.5: Maps of the changes of phytoplankton [chl-a], temperature and community composition 2005-2014. The top-left sub-plot is a map of the average number of ocean-colour observations in a typical July. Colour-bars in the right-hand subplots display trend (hue) and certainty (saturation) of the occurrence frequency of coccolithophores, *Phaeocystis* and diatoms through this interval.

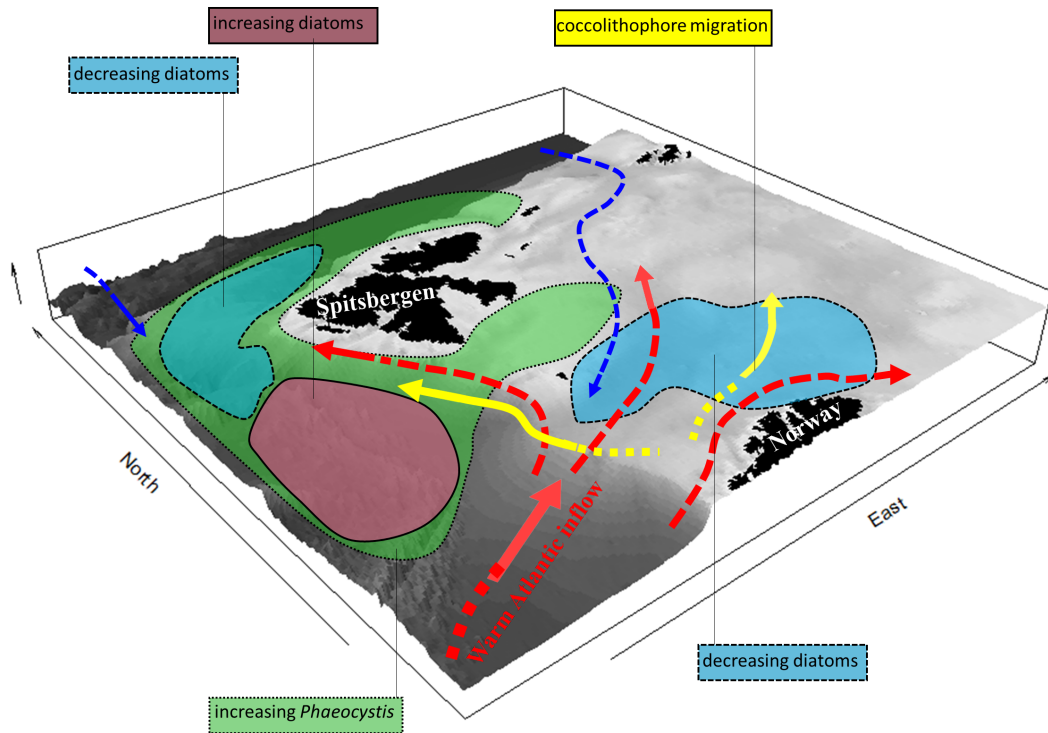


Figure 4.6: A synthesis of all the changes to phytoplankton community composition found in this study. Warm inflowing currents are represented in red, whilst cold waters flowing from the Nansen Basin are represented in blue. There is a general pattern of dispersal of coccolithophore blooms further into historically more Arctic-influenced areas as Atlantic inflow currents intensify. There is an ambiguous change in the predominance of diatom blooms; while they may be increasing close to the Atlantic inflow region, their incidence is in decline more broadly. The incidence of ocean-colour measurements indicative of *Phaeocystis* is broadly increasing throughout the Western Barents Sea, suggesting that ascendant Atlantic influences are benefiting this group.

4.7 Discussion

4.7.1 Phytoplankton community shift in the context of “Atlantification”

The most significant trend we have uncovered is an inferred increasing predominance of *Phaeocystis* in midsummer in the Barents Sea. The distribution of change illustrated in Figure 4.6 suggests that the findings of ship-board sampling by Nöthig et al. (2015) apply generally to the more Atlantic-influenced portions of the Barents Sea, consistent with expectations expressed in Reigstad et al. (2002). This may well be consistent with predictions by Randelhoff et al. (2018) that enhanced Atlantic Water input onto the Western shelves of the Barents Sea would alleviate late summer nutrient limitation and hence support increased pelagic primary production. It is also compatible with the report of Oziel et al. (2017) of increasing [*chl-a*] in the Barents Sea in late Summer, and the observation of Arrigo and Dijken (2015) of increasing annual pelagic primary production in the region. The Fram Strait region has also experienced a significant increase in *Phaeocystis* occurrence, consonant with the existing literature (Nöthig et al. 2015). Given the association of *Phaeocystis* blooms with the influx of warmer Atlantic derived waters into the Fram Strait (Nöthig et al. 2015) it is perhaps no surprise that the least Atlantified waters of the Barents Sea prove the most resistant to increases in *Phaeocystis* predominance.

These findings are complementary to our investigations of coccolithophore biogeography, that extend upon the work of authors such as Oziel et al. (2017), Smyth et al. (2004) and Hovland et al. (2013), showing that coccolithophore blooms have dispersed downstream in Atlantic inflow currents. The co-association of these changes is strongly suggestive of a signal of ‘Atlantification’ in midsummer Barents phytoplankton communities – especially as patterns in change appear to be delimited by the position of major fronts – implying a bottom-up control of inter-annual variation in phytoplankton community structure. We think it is curious that no significant change in biogeography for *Phaeocystis* was recov-

ered, but this may be a result of this genus already having a cosmopolitan distribution in the Barents Sea (Degerlund and Eilertsen 2010). We note that increases in *Phaeocystis* incidence illustrated in Figure 4.6 generally occur downstream of the sites of intense coccolithophore blooms and beyond the ‘Southern Front’ identified in Oziel et al. (2016) but not beyond the Polar Front, perhaps implying that phytoplankton communities of different Barents water-masses exhibit subtly different responses to environmental change. Diatom detections exhibited a complex change between 2005-2014, increasing their incidence in inflow regions but declining further downstream. This might result from the dual changes of increasing Atlantic inflow and decreasing Atlantic inflow silicic acid inventories (Rey 2012). Proximal to the inflow, silicic acid may not be limiting, so an increased flux may justify higher diatom growth, but further downstream, the declining silicic acid content of Atlantic Water may mean that phytoplankton groups less reliant on this nutrient are favoured.

We stress that simple studies of [*chl-a*], or the use of bio-optical algorithms that do not sub-divide highly-packaged phytoplankton into diatom and *Phaeocystis*-dominated assemblages, would provide a much coarser impression of ecosystem change (for example see Figure 4.5 [*chl-a*] change). Coarser products such as these might be interpreted as being concomitant with suggestions of increasing new production (Arrigo and Dijken 2015). However, our finding of increasing *Phaeocystis* predominance suggests that the style of production is likely becoming increasingly regenerated in character. This was explicitly considered by Arrigo and Dijken (2015) as a possibility, and might be inferred from isolated *in-situ* records (Nöthig et al. 2015; Engel et al. 2017), but we have here provided the first bio-optical evidence of this. Other authors have also pointed out that deepening mixed layers in midsummer could explain greater regenerated production (Oziel et al. 2017), a hypothesis compatible with increasing *Phaeocystis*, which is a shade-tolerant genus that prospers under conditions of deep mixing (Fragoso et al. 2017). These interpretations are consistent with observations that remotely-sensed [*chl-a*] are concentrated in mid and late

summer, when episodic mixing prevails, and that $[chl-a]$ detected in spring – when production depends on the nutrient inventory stocked over winter – have declined in years with lower sea-ice (Wassmann et al. 1999; Oziel et al. 2017). This shows that the retrieval of dominant phytoplankton groups from ocean-colour can inform a more nuanced understanding of incipient ecosystem change in the Arctic in the face of human-induced climate change. We anticipate that the launch of hyper-spectral sensors (PACE) into space will further invigorate taxonomically-informed approaches in the study of ocean-colour.

4.7.2 Implications

The ecological significance of increasing *Phaeocystis* predominance is difficult to ascertain. Previous authors have argued that zooplankton may modify their diets, buffering any substantial effects on higher trophic levels (Vernet et al. 2017). By contrast, other authors have suggested that the efficiency of the biological pump should decline under *Phaeocystis* predominance (Wassmann 1994), because descending *Phaeocystis* aggregates are remineralised faster than diatoms (Reigstad and Wassmann 2007) – an observation consistent with the faster sinking rates reported for more massive diatom cells (Falkowski et al. 1998; Jiang et al. 2005). Suggestions of declining biological pump efficiency can be contrasted in turn with rare reports of mass-sedimentation of *Phaeocystis* (Wassmann et al. 1990), when colonies are ballasted by gypsum crystals rejected from sea-ice (Wollenburg et al. 2018), or the frustules of senescent diatom blooms (Peperzak et al. 2000a; Cadée and Hegeman 1991). *Phaeocystis* blooms frequently succeed diatoms in the Barents Sea and the co-occurrence of diatoms ballasting *Phaeocystis* colonies has been frequently recorded (Sazhin et al. 2007; Verity et al. 2007). Hence, there are credible reasons in the literature to believe that increasing *Phaeocystis* should decrease carbon export, perhaps increase it or have little effect at all.

Empirical studies of benthic carbon flux and trawls of benthic organisms also yield con-

flicting results. Lalande et al. (2013) reports greater biogenic and faecal pellet export under high sea-ice conditions – suggesting a warmer less diatom-dominated Barents Sea should have a weaker biological pump, while Cochrane et al. (2009) report lower macrobenthic animal biomass and sediment carbon content North of the oceanic Polar Front, implying that Atlantification and the Northward recession of the Polar Front ought to increase nourishment to the benthos. These conflicting reviews can perhaps be resolved by Jørgensen et al. (2014), who notes that the benthic expression of the Polar Front extends much further South than the ocean surface outcrop. Jørgensen et al. (2014) found that megafaunal (> 1 cm body length) biomass distribution differed from macrofaunal (> 1 mm body length, but < 1 cm), with greater biomass North of the benthic Polar Front. Jørgensen et al. (2014) hypothesised that a longer equilibration time-scale for larger benthic animals may mean changes in the benthos lag environmental change and subsequent research showed that gradients in benthic megafaunal biomass are decoupled from surface production and therefore that carbon flux must be controlled to a first order by biological pump efficiency – perhaps mediated by bloom composition, dynamics or ice-edge physics (Degen et al. 2016).

We may therefore synthesise these findings to suggest that a history of frequent fast-sinking diatom blooms justify greater megafaunal biomass in Northern sites, and that the Atlantification of these waters, and a switch to *Phaeocystis* predominance, will result in a poorer nourishment of the benthos and a shift in the long-term to lower megafaunal biomass. We must also consider that, during mid and late summer, phytoplankton blooms in the Barents Sea descend toward the nutricline to form Sub-surface Chlorophyll Maxima (SCMs) (Slagstad and Støle-Hansen 1991), meaning some production is missed by remote-sensing approaches. In this study, this may have the effect of giving us the false impression that phytoplankton communities North of the Polar Front are resistant to environmental change (Figure 4.5), when communities residing at depth here may in fact be changing. As there is currently much debate in the literature as to the effects of neglecting the SCM on

production estimates

(e.g. Hill et al. (2013) and Arrigo et al. (2011) and Bouman et al. (2020)), Arctic shelf SCMs clearly warrant further investigation.

4.7.3 Limitations to our experimental bio-optical algorithm

Overall, much greater confidence can be invested in the detection of coccolithophore blooms, which are clearly defined by the characteristic light-scattering properties of their shed coccoliths. We should, however, consider that, while coccolithophores may shed a large number of liths, and therefore dominate variance in ocean-colour signals, their blooms have previously been reported to be concomitant with diatoms, chrysophytes and dinoflagellates. Coccolithophore may only make relatively minor contributions to inventories of biomass and $[chl-a]$ in such settings (Volent et al. 2011).

Our approach is also necessarily limited by the available spectral resolution of the MODIS-Aqua sensor and by the underlying assumption in our bio-optical algorithm's construction that CDOM is invariant. This assumption is fair in the Barents Sea, where CDOM absorption remains constant even across major hydrographic features (Hancke et al. 2014), but is not true outside of the Barents Sea, limiting the wider application of our algorithm to other Arctic regions. We must also acknowledge that Astoreca et al. (2008) caution that their reflectance algorithm for the retrieval of chlorophyll-c3 may require modification in turbid waters and that accurate quantification of chlorophyll-c3 concentrations may not be possible even under ideal conditions – only the recognition of high concentrations indicative of *Phaeocystis* blooms.

We believe that we have taken care to use the algorithm of Astoreca et al. (2008) within these limitations. We might also regard the requirement to use the MODIS landbands (such as 469 nm) as a limitation, but we note that previous authors, such as Hu et al. (2010), have made successful use of these bands for the inference of Phytoplankton Functional Types (Hu

et al. 2010). We caution that our approach to correct for temporal drift in the inferred signal of chlorophyll-c3 absorption is specific to the MODIS-Aqua r2018.1 ocean-colour reprocessing, and that this approach may need to be modified for future reprocessings, which may subtly vary the values of different ocean-colour bands. The results from our bio-optical algorithm suggest that it is possible to construct regional algorithms for the retrieval of major Phytoplankton Functional Types in the Arctic, and that the output from such algorithms can form a useful supplement to incomplete *in-situ* records of phytoplankton community composition.

We must also be aware that sea-ice often precludes detection of ice-edge or under-ice phytoplankton blooms. Where the ice-field or ice-edge currents are dynamic, phytoplankton blooms may be advected beneath the ice (Johnsen et al. 2018), or low-light acclimated blooms may be exposed should the ice recede (Assmy et al. 2017). Photo-acclimation can influence overall intracellular pigment concentrations and complements, and therefore may bias remote-sensing ocean-colour algorithms that seek indices of pigment packaging or the ratios of taxon-indicative pigments.

4.8 Conclusions

We have developed a bio-optical algorithm to retrieve dominant phytoplankton groups in the Barents Sea and probed its applicability to investigate incipient changes in community composition. We have presented evidence that an expansion of coccolithophore range further down-stream in Atlantic inflow currents is ongoing, consistent with the expectations of previous authors. We have suggested that *Phaeocystis* predominance in the Barents Sea in midsummer is likely increasing in more Atlantic-influenced regions, which is consistent with a signal of Atlantification in the Barents phytoplankton community and the expectations of several previous authors. The combination of these changes may mean that reported increases in remotely-sensed production in the Barents Sea are driven in part by

higher standing stocks of phytoplankton associated with regenerated styles of production increasingly mediated by the microbial loop – a change with potential significance for higher trophic levels. The retrieval of these signals of Atlantification by our bio-optical algorithm emphasises that existing records of ocean-colour will remain useful for interpretation alongside newer ocean-colour products from up-coming missions such as PACE.

The datasets of optical properties used in the construction of the phytoplankton community composition algorithm used in this study are available at

[doi:10.5285/97daa7ea-8792-6cff-e053-6c86abc0dd46]

[doi:10.5285/982b6da2-7e11-060a-e053-6c86abc09389]

[doi:10.5285/982b6da2-7e12-060a-e053-6c86abc09389]

accompanying datasets for [*chl-a*] are available at

[doi:10.5285/97daa7ea-8793-6cff-e053-6c86abc0dd46]

[doi:10.5285/982b6da2-7e13-060a-e053-6c86abc09389]

[doi:10.5285/982b6da2-7e14-060a-e053-6c86abc09389]

MODIS-A data is available from NASA's website for ocean-colour.

[doi:10.5067/AQUA/MODIS/L3SMI/

A. Orkney carried out analyses and drafted the manuscript.

H.A. Bouman supervised analysis and the drafting of the manuscript.

T. Platt edited the manuscript.

B. Narayanaswamy, assisted in the interpretation of the significance of our results for the structure of benthic communities.

I. Kostakis set up HydroLight simulations.

All authors read and approved the manuscript.

This research was funded by the Natural Environment Research Council (NERC) Grant

number: NE/P006507/1

We acknowledge D. McKee of the University of Strathclyde 16 Richmond St, Glasgow G1 1XQ, for assisting in the configuration of HydroLight. We acknowledge S. Sathyendranath of the Plymouth Marine Laboratory, Prospect Place, Plymouth, PL1 3DH, for assessing the validity of our correction scheme for temporal drift.

References

- Alvain, Séverine, Cyril Moulin, Yves Dandonneau, and François-Marie Bréon. 2005. "Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery." *Deep Sea Research Part I: Oceanographic Research Papers* 52 (11): 1989–2004.
- Alvain, Séverine, Cyril Moulin, Yves Dandonneau, and Hubert Loisel. 2008. "Seasonal distribution and succession of dominant phytoplankton groups in the Global Ocean: A satellite view." *Global Biogeochemical Cycles* 22 (3).
- Arrigo, Kevin R, and Gert L van Dijken. 2015. "Continued increases in Arctic Ocean primary production." *Progress in Oceanography* 136:60–70.
- Arrigo, Kevin R, Patricia A Matrai, and Gert L Van Dijken. 2011. "Primary productivity in the Arctic Ocean: Impacts of complex optical properties and subsurface chlorophyll maxima on large-scale estimates." *Journal of Geophysical Research: Oceans* 116 (C11).
- Arrigo, Kevin R, Dale H Robinson, Denise L Worthen, Brian Schieber, and Michael P Lizotte. 1998. "Bio-optical properties of the southwestern Ross Sea." *Journal of Geophysical Research: Oceans* 103 (C10): 21683–21695.
- Assmy, Philipp, Mar Fernández-Méndez, Pedro Duarte, Amelie Meyer, Achim Randelhoff, Christopher J Mundy, Lasse M Olsen, Hanna M Kauko, Allison Bailey, Melissa Chierici, et al. 2017. "Leads in Arctic pack ice enable early phytoplankton blooms below snow-covered sea ice." *Scientific reports* 7 (1): 1–9.
- Astoreca, Rosa, Véronique Rousseau, Kevin Ruddick, Cécile Knechciak, Barbara Van Mol, Jean-Yves Parent, and Christiane Lancelot. 2008. "Development and application of an algorithm for detecting *Phaeocystis globosa* blooms in the Case 2 Southern North Sea waters." *Journal of Plankton Research* 31 (3): 287–300.
- Båmstedt, Ulf, Hans Christian Eilertsen, Kurt S Tande, Dag Slagstad, and Hein Rune Skjoldal. 1991. "Copepod grazing and its potential impact on the phytoplankton development in the Barents Sea." *Polar Research* 10 (2): 339–354.
- Barton, Benjamin I, Yueng-Djern Lenn, and Camille Lique. 2020. "Observed atlantification of the Barents Sea causes the Polar Front to limit the expansion of winter sea ice." *JPO* 500:18–0003.
- Boetius, Antje, Sebastian Albrecht, Karel Bakker, Christina Bienhold, Janine Felden, Mar Fernández-Méndez, Stefan Hendricks, Christian Katlein, Catherine Lalande, Thomas Krumpen, et al. 2013. "Export of algal biomass from the melting Arctic sea ice." *Science* 339 (6126): 1430–1432.
- Bouman, Heather A, Thomas Jackson, Shubha Sathyendranath, and Trevor Platt. 2020. "Vertical structure in chlorophyll profiles: influence on primary production in the Arctic Ocean." *Philosophical Transactions of the Royal Society A* 378 (2181): 20190351.
- Cadée, Gerhard C, and Jan Hegeman. 1991. "Phytoplankton primary production, chlorophyll and species composition, organic carbon and turbidity in the Marsdiep in 1990, compared with foregoing years." *Hydrobiological Bulletin* 25 (1): 29–35.

- Carmack, Eddy, and Paul Wassmann. 2006. "Food webs and physical-biological coupling on pan-Arctic shelves: unifying concepts and comprehensive perspectives." *Progress in Oceanography* 71 (2-4): 446–477.
- Cavalieri, Donald J, and Claire L Parkinson. 2012. "Arctic sea ice variability and trends, 1979-2010." *The Cryosphere* 6 (4): 881.
- Cochrane, Sabine KJ, Stanislav G Denisenko, Paul E Renaud, Christopher S Emblow, William G Ambrose Jr, Ingrid H Ellingsen, and Jofrid Skardhamar. 2009. "Benthic macrofauna and productivity regimes in the Barents Sea- ecological implications in a changing Arctic." *Journal of Sea Research* 61 (4): 222–233.
- Cota, Glenn F, W Glen Harrison, Trevor Platt, Shubha Sathyendranath, and Venetia Stuart. 2003. "Bio-optical properties of the Labrador Sea." *Journal of Geophysical Research: Oceans* 108 (C7).
- Cullen, John J, and Marlon R Lewis. 1995. "Biological processes and optical measurements near the sea surface: Some issues relevant to remote sensing." *Journal of Geophysical Research: Oceans* 100 (C7): 13255–13266.
- Degen, Renate, Lis Lindal Jørgensen, Pavel Ljubin, Ingrid H Ellingsen, Hendrik Pehlke, and Thomas Brey. 2016. "Patterns and drivers of megabenthic secondary production on the Barents Sea shelf." *Marine Ecology Progress Series* 546:1–16.
- Degerlund, Maria, and Hans Christian Eilertsen. 2010. "Main species characteristics of phytoplankton spring blooms in NE Atlantic and Arctic waters (68–80 N)." *Estuaries and coasts* 33 (2): 242–269.
- Devred, Emmanuel, Shubha Sathyendranath, Venetia Stuart, and Trevor Platt. 2011. "A three component classification of phytoplankton absorption spectra: Application to ocean-color data." *Remote Sensing of Environment* 115 (9): 2255–2266.
- Dong, Kaixing, Kristina Øie Kvile, Nils Chr Stenseth, and Leif Chr Stige. 2020. "Associations among temperature, sea ice and phytoplankton bloom dynamics in the Barents Sea." *Marine Ecology Progress Series* 635:25–36.
- Dutkiewicz, Stephanie, Anna E Hickman, Oliver Jahn, Stephanie Henson, Claudie Beaulieu, and Erwan Monier. 2019. "Ocean colour signature of climate change." *Nature communications* 10 (1): 578.
- Duysens, LNM. 1956. "The flattening of the absorption spectrum of suspensions, as compared to that of solutions." *Biochimica et Biophysica Acta* 19:1–12.
- Engel, Anja, Judith Piontek, Katja Metfies, Sonja Endres, Pim Sprong, Ilka Peeken, Steffi Gäbler-Schwarz, and Eva-Maria Nöthig. 2017. "Inter-annual variability of transparent exopolymer particles in the Arctic Ocean reveals high sensitivity to ecosystem changes." *Scientific Reports* 7 (1): 4129.
- Falkowski, Paul G, Richard T Barber, and Victor Smetacek. 1998. "Biogeochemical controls and feedbacks on ocean primary production." *Science* 281 (5374): 200–206.
- Fragoso, Glaucia M, Alex J Poulton, Igor M Yashayaev, Erica JH Head, and Duncan A Purdie. 2017. "Spring phytoplankton communities of the Labrador Sea (2005–2014): pigment signatures, photophysiology and elemental ratios." *Biogeosciences* 14:1235–1259.
- Francis, Jennifer A, and Elias Hunter. 2007. "Drivers of declining sea ice in the Arctic winter: A tale of two seas." *Geophysical Research Letters* 34 (17).
- Galí, Martí, Emmanuel Devred, Marcel Babin, and Maurice Levasseur. 2019. "Decadal increase in Arctic dimethylsulfide emission." *Proceedings of the National Academy of Sciences* 116 (39): 19311–19317.
- Hancke, Kasper, Erlend K Hovland, Zsolt Volent, Ragnhild Pettersen, Geir Johnsen, Mark Moline, and Egil Sakshaug. 2014. "Optical properties of CDOM across the Polar Front in the Barents Sea: Origin, distribution and significance." *Journal of Marine Systems* 130:219–227.
- Hill, Victoria J, Patricia A Matrai, Elise Olson, S Suttles, Mike Steele, LA Codispoti, and Richard C Zimmerman. 2013. "Synthesis of integrated primary production in the Arctic

- Ocean: II. In situ and remotely sensed estimates." *Progress in Oceanography* 110:107–125.
- Hovland, Erlend Kjeldsberg, Heidi M Dierssen, Ana Sofia Ferreira, and Geir Johnsen. 2013. "Dynamics regulating major trends in Barents Sea temperatures and subsequent effect on remotely sensed particulate inorganic carbon." *Marine Ecology Progress Series* 484:17–32.
- Hu, Chuanmin, Jennifer Cannizzaro, Kendall L Carder, Frank E Muller-Karger, and Robert Hardy. 2010. "Remote detection of *Trichodesmium* blooms in optically complex coastal waters: Examples with MODIS full-spectral data." *Remote Sensing of Environment* 114 (9): 2048–2058.
- Jackson, Thomas, Heather A Bouman, Shubha Sathyendranath, and Emmanuel Devred. 2010. "Regional-scale changes in diatom distribution in the Humboldt upwelling system as revealed by remote sensing: implications for fisheries." *ICES Journal of Marine Science* 68 (4): 729–736.
- Jiang, Lin, Oscar ME Schofield, and Paul G Falkowski. 2005. "Adaptive evolution of phytoplankton cell size." *The American Naturalist* 166 (4): 496–505.
- Johnsen, G, and E Sakshaug. 2000. "Monitoring of harmful algal blooms along the Norwegian coast using bio-optical methods." *South African Journal of Marine Science* 22 (1): 309–321.
- Johnsen, Geir, Marit Norli, Mark Moline, Ian Robbins, Cecilie von Quillfeldt, Kai Sørensen, Finlo Cottier, and Jørgen Berge. 2018. "The advective origin of an under-ice spring bloom in the Arctic Ocean using multiple observational platforms." *Polar Biology* 41 (6): 1197–1216.
- Johnsen, Geir, and Egil Sakshaug. 2007. "Biooptical characteristics of PSII and PSI in 33 species (13 pigment groups) of marine phytoplankton, and the relevance for pulse-amplitude-modulated and fast-repetition-rate fluorometry 1." *Journal of Phycology* 43 (6): 1236–1251.
- Jørgensen, Lis Lindal, Pavel Ljubin, Hein Rune Skjoldal, Randi B Ingvaldsen, Natalia Anisimova, and Igor Manushin. 2014. "Distribution of benthic megafauna in the Barents Sea: baseline for an ecosystem approach to management." *ICES Journal of Marine Science* 72 (2): 595–613.
- Kauko, Hanna M, Torbjørn Taskjelle, Philipp Assmy, Alexey K Pavlov, CJ Mundy, Pedro Duarte, Mar Fernández-Méndez, Lasse M Olsen, Stephen R Hudson, Geir Johnsen, et al. 2017. "Windows in Arctic sea ice: Light transmission and ice algae in a refrozen lead." *Journal of Geophysical Research: Biogeosciences* 122 (6): 1486–1505.
- Kirk, John TO. 1994. *Light and photosynthesis in aquatic ecosystems*. Cambridge University Press.
- Kruppen, Thomas, H Jakob Belter, Antje Boetius, Ellen Damm, Christian Haas, Stefan Hendricks, Marcel Nicolaus, Eva-Maria Nöthig, Stephan Paul, Ilka Peeken, et al. 2019. "Arctic warming interrupts the Transpolar Drift and affects long-range transport of sea ice and ice-rafted matter." *Scientific Reports* 9 (1): 5459.
- Lalande, Catherine, Eduard Bauerfeind, Eva-Maria Nöthig, and Agnieszka Beszczynska-Möller. 2013. "Impact of a warm anomaly on export fluxes of biogenic matter in the eastern Fram Strait." *Progress in Oceanography* 109:70–77.
- Loeng, Harald. 1991. "Features of the physical oceanographic conditions of the Barents Sea." *Polar Research* 10 (1): 5–18.
- Nair, Anitha, Shubha Sathyendranath, Trevor Platt, Jesus Morales, Venetia Stuart, Marie-Hélène Forget, Emmanuel Devred, and Heather Bouman. 2008. "Remote sensing of phytoplankton functional types." *Remote Sensing of Environment* 112 (8): 3366–3375.
- Nöthig, Eva-Maria, Astrid Bracher, Anja Engel, Katja Metfies, Barbara Niehoff, Ilka Peeken, Eduard Bauerfeind, Alexandra Cherkasheva, Steffi Gäbler-Schwarz, Kristin Hardge, et al. 2015. "Summertime plankton ecology in Fram Strait- a compilation of long and short-term observations." *Polar Research* 34 (1): 23349.

- OBPG. 2019. "Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Daily L3SMI Ocean Colour Data." *NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Daily L3SMI Ocean Colour Data Data; NASA OB.DAAC, Greenbelt, MD, USA.*, <https://doi.org/doi:10.5067/AQUA/MODIS/L3SMI/accessedon11/04/2019..>
- Osman, Matthew B, Sarah B Das, Luke D Trusel, Matthew J Evans, Hubertus Fischer, Mackenzie M Grieman, Sepp Kipfstuhl, Joseph R McConnell, and Eric S Saltzman. 2019. "Industrial-era decline in subarctic Atlantic productivity." *Nature* 569:551–555.
- Oziel, Laurent, Alberto Baudena, Mathieu Ardyna, Philippe Massicotte, A Randelhoff, J-B Sallée, Randi Brunvær Ingvaldsen, Emmanuel Devred, and Marcel Babin. 2020. "Faster Atlantic currents drive poleward expansion of temperate phytoplankton in the Arctic Ocean." *Nature communications* 11 (1): 1–8.
- Oziel, Laurent, G Neukermans, M Ardyna, C Lancelot, J-L Tison, Paul Wassmann, Jérôme Sirven, Diana Ruiz-Pino, and J-C Gascard. 2017. "Role for Atlantic inflows and sea ice loss on shifting phytoplankton blooms in the Barents Sea." *Journal of Geophysical Research: Oceans* 122 (6): 5121–5139.
- Oziel, Laurent, Jérôme Sirven, and Jean-Claude Gascard. 2016. "The Barents Sea frontal zones and water masses variability (1980–2011)." *Ocean Science* 12(1):169–184.
- Pante, Eric, and Benoit Simon-Bouhet. 2013. "marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R." *PLoS ONE* 8 (9): e73051. [doi:10.1371/journal.pone.0073051](https://doi.org/doi:10.1371/journal.pone.0073051).
- Pavlov, Alexey K, Torbjørn Taskjelle, Hanna M Kauko, Børge Hamre, Stephen R Hudson, Philipp Assmy, Pedro Duarte, Mar Fernández-Méndez, CJ Mundy, and Mats A Granskog. 2017. "Altered inherent optical properties and estimates of the underwater light field during an Arctic under-ice bloom of *Phaeocystis pouchetii*." *Journal of Geophysical Research: Oceans* 122 (6): 4939–4961.
- Peperzak, L, F Colijn, EG Vrieling, WWC Gieskes, and JCH Peeters. 2000a. "Observations of flagellates in colonies of *Phaeocystis globosa* (Prymnesiophyceae); a hypothesis for their position in the life cycle." *Journal of Plankton Research* 22 (12): 2181–2203.
- Petro, Susanna, Karen Pham, and George Hilton. 2020. "Plankton, Aerosol, Cloud, Ocean Ecosystem (PACE) Mission Integration and Testing."
- Pettersen, Ragnhild, Geir Johnsen, Jørgen Berge, and Erlend Kjeldsberg Hovland. 2011. "Phytoplankton chemotaxonomy in waters around the Svalbard archipelago reveals high amounts of Chl b and presence of gyroxanthin-diester." *Polar biology* 34 (5): 627–635.
- Quere, Corinne Le, Sandy P Harrison, I Colin Prentice, Erik T Buitenhuis, Olivier Aumont, Laurent Bopp, Hervé Claustre, Leticia Cotrim Da Cunha, Richard Geider, Xavier Giraud, et al. 2005. "Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models." *Global Change Biology* 11 (11): 2016–2040.
- Randelhoff, Achim, Marit Reigstad, Melissa Chierici, Arild Sundfjord, Vladimir Ivanov, Mattias Rolf Cape, Maria Vernet, Jean-Éric Tremblay, Gunnar Bratbak, and Svein Kristiansen. 2018. "Seasonality of the physical and biogeochemical hydrography in the inflow to the Arctic Ocean through Fram Strait." *Frontiers in Marine Science* 5:224.
- Rat'kova, Tatjana N, and Paul Wassmann. 2002. "Seasonal variation and spatial distribution of phyto-and protozooplankton in the central Barents Sea." *Journal of Marine Systems* 38 (1-2): 47–75.
- Reigstad, Marit, and Paul Wassmann. 2007. "Does *Phaeocystis* spp. contribute significantly to vertical export of organic carbon?" In *Phaeocystis, major link in the biogeochemical cycling of climate-relevant elements*, 217–234. Springer.

- Reigstad, Marit, Paul Wassmann, Christian Wexels Riser, Sigrid Øygarden, and Francisco Rey. 2002. "Variations in hydrography, nutrients and chlorophyll a in the marginal ice-zone and the central Barents Sea." *Journal of Marine Systems* 38 (1-2): 9–29.
- Rey, Francisco. 2012. "Declining silicate concentrations in the Norwegian and Barents Seas." *ICES Journal of Marine Science* 69 (2): 208–212.
- Rodriguez, F, M Chauton, G Johnsen, K Andresen, LM Olsen, and M Zapata. 2006. "Photoacclimation in phytoplankton: implications for biomass estimates, pigment functionality and chemotaxonomy." *Marine Biology* 148 (5): 963–971.
- Sakshaug, Egil, Geir Helge Johnsen, and Kit M Kovacs. 2009. *Ecosystem Barents Sea*. Tapir Academic Press.
- Sathyendranath, Shubha, Jim Aiken, S Alvain, R Barlow, H Bouman, Astrid Bracher, R Brewin, Annick Bricaud, CW Brown, AM Ciotti, et al. 2014. "Phytoplankton functional types from Space." In *(Reports of the International Ocean-Colour Coordinating Group (IOCCG); 15)*, 1–156. International Ocean-Colour Coordinating Group.
- Sathyendranath, Shubha, Louisa Watts, Emmanuel Devred, Trevor Platt, Carla Caverhill, and Heidi Maass. 2004. "Discrimination of diatoms from other phytoplankton using ocean-colour data." *Marine Ecology Progress Series* 272:59–68.
- Sazhin, Andrey F, L Felipe Artigas, Jens C Nejstgaard, and Marc E Frischer. 2007. "The colonization of two *Phaeocystis* species (Prymnesiophyceae) by pennate diatoms and other protists: a significant contribution to colony biomass." In *Phaeocystis, major link in the biogeochemical cycling of climate-relevant elements*, 137–145. Springer.
- Slagstad, Dag, and Kjell Støle-Hansen. 1991. "Dynamics of plankton growth in the Barents Sea: model studies." *Polar Research* 10 (1): 173–186.
- Smyth, TJ, T Tyrrell, and B Tarrant. 2004. "Time series of coccolithophore activity in the Barents Sea, from twenty years of satellite imagery." *Geophysical Research Letters* 31 (11).
- Stuart, V, S Sathyendranath, EJH Head, T Platt, B Irwin, and H Maass. 2000. "Bio-optical characteristics of diatom and prymnesiophyte populations in the Labrador Sea." *Marine Ecology Progress Series* 201:91–106.
- Taylor, Isabelle A, James Preston, Elisa Carboni, Tamsin A Mather, Roy G Grainger, Nicolas Theys, Silvana Hidalgo, and Brendan McCormick Kilbride. 2018. "Exploring the Utility of IASI for Monitoring Volcanic SO₂ Emissions." *Journal of Geophysical Research: Atmospheres* 123 (10): 5588–5606.
- Verity, Peter G, Corina P Brussaard, Jens C Nejstgaard, Maria A van Leeuwe, Christiane Lancelot, and Linda K Medlin. 2007. "Current understanding of *Phaeocystis* ecology and biogeochemistry, and perspectives for future research." *Biogeochemistry* 83 (1-3): 311–330.
- Vernet, Maria, Tammi L Richardson, Katja Metfies, Eva-Maria Nöthig, and Ilka Peeken. 2017. "Models of plankton community changes during a warm water anomaly in Arctic waters show altered trophic pathways with minimal changes in carbon export." *Frontiers in Marine Science* 4:160.
- Volent, Zsolt, Geir Johnsen, Erlend K Hovland, Are Folkestad, Lasse M Olsen, Karl Tangen, and Kai Sorensen. 2011. "Improved monitoring of phytoplankton bloom dynamics in a Norwegian fjord by integrating satellite data, pigment analysis, and Ferrybox data with a coastal observation network." *Journal of Applied Remote Sensing* 5 (1): 053561.
- Wassmann, P. 1994. "Significance of sedimentation for the termination of *Phaeocystis* blooms." *Journal of Marine Systems* 5 (1): 81–100.
- Wassmann, Paul, Tatjana Ratkova, Inger Andreassen, Maria Vernet, Gunnar Pedersen, and Francisco Rey. 1999. "Spring bloom development in the marginal ice zone and the central Barents Sea." *Marine Ecology* 20 (3-4): 321–346.
- Wassmann, Paul, Maria Vernet, B Greg Mitchell, and Francisco Rey. 1990. "Mass sedimentation of *Phaeocystis pouchetii* in the Barents Sea." *Marine Ecology Progress Series* 66:183–195.

- Wollenburg, JE, C Katlein, Gernot Nehrke, E-M Nöthig, Jens Matthiessen, Dieter A Wolf-Gladrow, A Nikolopoulos, F Gázquez-Sanchez, L Rossmann, Philipp Assmy, et al. 2018. "Ballasting by cryogenic gypsum enhances carbon export in a *Phaeocystis* under-ice bloom." *Scientific Reports* 8 (1): 7703.
- Zhang, Hailong, Emmanuel Devred, Amane Fujiwara, Zhongfeng Qiu, and Xiaohan Liu. 2018. "Estimation of phytoplankton taxonomic groups in the Arctic Ocean using phytoplankton absorption properties: implication for ocean-color remote sensing." *Optics Express* 26 (24): 32280–32301.

Chapter 5

Atlantic inflow the primary driver of fall blooms in European Arctic

Andrew Orkney ^{1,a}, Shubha Sathyendranath ^{2,b} Thomas Jackson ^{2,c} Marie Porter ^{3,d}, and
Heather Bouman ^{1,e}

^a <https://orcid.org/0000-0003-4972-2541>

^b <https://orcid.org/0000-0003-3586-192X>

^c <https://orcid.org/0000-0003-4336-1597>

^d <https://orcid.org/0000-0003-3686-353X>

^e <https://orcid.org/0000-0002-7407-9431>

¹Department of Earth Sciences, University of Oxford, Oxford, United Kingdom

²Plymouth Marine Laboratory, Plymouth, United Kingdom

³Scottish Association for Marine Science, Scottish Marine Institute, Oban, PA37 1QA

5.1 Prelude

Thus far, this thesis has investigated functional variety in Barents Sea phytoplankton assemblages and explored their consistent definition and detection from bio-optical properties; developed a bespoke bio-optical algorithm to detect Phytoplankton Functional Types in the Barents Sea; and explored recent changes in the biogeography and community composition of phytoplankton blooms in the Barents Sea in mid-summer. This work has provided a synoptic perspective that aids the integration of isolated and temporally discontinuous *in-situ* studies.

We have therefore pre-eminently concentrated on community compositional change, but must now turn our attention to an equally salient aspect of Arctic change; phenology. Historically, Arctic shelf seas were typified by a single-bloom phenology, in which an intense spring bloom occurs concomitant with sea-ice melt – resulting in a trophic cascade that provides nourishment to the ecosystem for the remainder of the year. This differs from typical temperate double-bloom phenologies, in which wind-induced mixing resulting from de-stabilisation of the autumnal atmosphere can cause vertical mixing and upwelling of nutrients. Temperate shelf seas may therefore host two equinoctial blooms each year. In recent decades, such autumnal blooms have begun to occur in Arctic shelf settings, perhaps prompted by a decline in sea-ice extent that has facilitated greater action of wind upon the sea and provided greater illumination more conducive to photosynthesis. However, while this hypothesis has broadly been validated over the interior Eurasian shelves, evidence of increasing autumnal blooms and their drivers has proven more elusive in the Barents Sea.

In order to investigate this question, we exploit a decadal-scale record of ocean-colour to investigate September [*chl-a*] in the Barents Sea and make use of model reanalyses of climatological conditions to explore potential environmental drivers of variability in autumnal bloom occurrence. We use a standard off-the-shelf [*chl-a*] product, based on OC-CCI ocean-colour, which is agnostic to Phytoplankton Functional Types. While OC-CCI lacks

spectral resolution at blue wavelengths required to distinguish Phytoplankton Functional Types, the longer period of observations is necessary in order to perceive robust changes in autumnal bloom occurrence.

We, furthermore, elected not to attempt to retrieve Phytoplankton Functional Types in Autumnal time series because the Arctic **PRIZE** communities we have described thus far are exclusive to the Spring and Summer. In the absence of knowledge of autumnal phytoplankton bloom compositions and their $a_{ph}(\lambda)$ variety, we are unable to confidently infer phytoplankton compositions in remotely-sensed ocean-colour within the autumn. We specifically identify the further investigation of autumnal bloom composition as a subject of future research. We believe our decision not to attempt to infer Phytoplankton Functional Types accords with our previous commentary that estimates of this property from ocean-colour must be integrated with *in-situ* sampling campaigns. In addition, we remind the reader that we assumed in our appraisal of the physics of remote-sensing that angle-dependent changes in R_{rs} are not a strong function of λ . This assumption will become more difficult to defend as solar zenith angles increase (as the sun becomes lower in the sky), which is the case in equinoctial periods.

Our investigation produces evidence of increasing autumnal bloom intensity – as shown by higher September $[chl-a]$ – in the Barents Sea, but the strongest increases are concentrated in perennially ice-free regions, calling the role of declining sea-ice persistence into question. Our analysis suggests that the delivery of Atlantic Waters, a key contributor to Barents Sea macronutrient inventories, is critical to the establishment of conditions permissive to autumnal blooms.

As our analysis is agnostic to Phytoplankton Functional Types, and relatively few Arctic sampling missions – including Arctic **PRIZE** – have sampled phytoplankton communities in the autumnal bloom, discerning the compositional identity and functional significance of autumnal blooms for ecosystem functions remains a fertile area of potential future research.

5.2 Abstract

Arctic seas are typified by a single spring bloom. This contrasts with temperate seas, where subsidiary autumnal blooms occur routinely. Local controls, such as nutrient limitation, are often invoked to explain the absence of autumnal Arctic blooms. Alleviation of nutrient limitation by high winds has been implicated in recent reports of rising frequencies of autumnal Arctic blooms. However, despite consistently high winds, autumnal blooms have only become routine more recently in the Barents Sea, one of the most productive regions of the Arctic Ocean.

We perform geographically-resolved regressions between remotely-sensed September [*chl-a*] and environmental conditions in the Barents Sea. We rule out a positive relationship between high winds and autumn bloom intensity, clearly distinguishing autumnal bloom controls here from other Arctic shelves. We find that Atlantic inflow currents are strongly associated with autumnal bloom intensity, highlighting the need to extend understanding of bloom controls beyond 1-dimensional conceptual models.

5.3 Plain Language Summary

A majority of annual Arctic phytoplankton productivity occurs in single spring blooms, contrasting with seas in temperate regions, where smaller secondary blooms often occur in the autumn. It is widely believed that local conditions, such as dwindling sunlight or exhaustion of nutrients, prevent autumnal Arctic blooms.

In recent years, high winds – which cause the upwelling of deep nutrient-rich waters – have accompanied an increase in autumnal blooms over much of the Arctic Ocean, but not

in the Barents Sea.

We investigated further by exploring the relationship between Barents Sea phytoplankton bloom intensity, inferred from satellite observations, and environmental conditions such as high winds, Sea Surface Temperature and current speeds. We found that there was no significant relationship between high winds and bloom intensity in the Barents Sea. We found that the speed of currents delivering nutrient-rich Atlantic Waters is strongly related to autumn bloom intensity in the Barents Sea.

This emphasises that outside influences, rather than only local forcing, need to be further considered by biological oceanographers who seek to understand how environmental conditions coincide to control the occurrence of phytoplankton blooms.

5.4 Introduction

The Arctic shelf seas are a productive Northern ecosystem that has historically been typified by a single intense period of phytoplankton blooms, associated with the recession of melting sea-ice and the consequent establishment of a shallow and nutrient-replete surface mixed layer (Båmstedt et al. 1991; Wassmann et al. 1999). This contrasts with temperate ‘double-bloom’ seasonalities of the adjacent North Atlantic region (González Taboada and Anadón 2014), whereby an initial spring bloom is curtailed by nutrient exhaustion and a secondary autumnal bloom occurs as fall storm systems promote the upwelling of nitrate-replete deep waters (Findlay et al. 2006). While autumnal blooms have been recorded in the Barents Sea, they have been comparatively less common than spring blooms and require special conditions to occur (Strass and Nöthig 1996).

Typical explanations for the absence of autumnal blooms in Arctic settings have emphasised limitation by local controls, such as grazing by zooplankton (Cushing 1959) or depressed photosynthesis due to falling light levels in high Northern latitudes (Platt et al. 2009).

In recent years, double-blooms have increased in frequency in Arctic shelf settings, and investigation has demonstrated their occurrence to be associated with an increasing incidence of high winds, sufficient to cause vertical mixing, as in temperate waters (Ardyna et al. 2014). Ardyna et al. (2014) hypothesised that decreased sea-ice extents and increasing frequencies of storms coincide to alleviate nutrient limitation. It is implicit in this hypothesis that neither low autumn light or persistent grazing prevent autumnal Arctic blooms. Indeed, Ardyna et al. (2014) comments that declining sea-ice extents may increase light availability as well as predisposing the ocean to overturning by storms. Fundamentally, this conceptual model remains similar to previous hypotheses of limitation, in that it emphasises local forcing and is inherently 1-dimensional in character.

The increasing frequency of autumnal Arctic blooms implies that, as anthropogenic perturbation of the climate system results in reduced sea-ice extents (Francis and Hunter 2007; Cavalieri and Parkinson 2012; Shapiro et al. 2003; Kwok 2018), increased Sea Surface Temperatures (SST) (Boitsov et al. 2012; Eriksen et al. 2017; Barton et al. 2020), and a more energetic atmosphere (Simmonds and Keay 2009), the hydrography and stratification of the Arctic shelf seas increasingly resemble temperate waters – with consequences for the seasonality, community composition of phytoplankton blooms and the significance of their production to higher trophic levels.

The Barents Sea, the Westernmost Arctic shelf sea of Eurasia (Figure 5.1), is the most influenced by temperate Atlantic water-masses (Smedsrud et al. 2013), and we may therefore expect an increasingly Atlantic character to first manifest here, before propagating to the interior shelves, as with other modes of climate variability such as the signal of the Arctic Oscillation (Karcher et al. 2012). Indeed, an increasing inflow of Atlantic-derived waters to this region is already undermining haline stratification (Polyakov et al. 2017), reducing physical barriers to wind-induced vertical mixing. Yet, in a time series of blooms between 1998 and 2012, the Barents Sea was not clearly characterised by a greater incidence of

autumnal blooms than the Kara or Laptev Seas, and the increase in autumnal Barents Sea blooms ($\sim 25\%$) was well below the increases sustained on the interior shelves ($\sim 48\text{--}56\%$) (Ardyna et al. 2014).

More recent time series have, however, pointed to the development of a more definitive prevalence of autumnal blooms in the Barents Sea, with fall productivity peaking in September (Oziel et al. 2017; Dong et al. 2020). We may therefore question why autumnal blooms first increased farthest from the influence of Atlantic water-masses, before progressing towards the Atlantic inflow, especially as it is those seas closest to nutrient-rich Atlantic inflow that have historically supported the greatest productivity (Carmack et al. 2006; Sakshaug 2004) and the richest fisheries (Wassmann et al. 2006a). Investigating this question may improve insight into the physical mechanisms that regulate productive Northern shelf seas and the nature and pattern of change to phytoplankton productivity and seasonality that should be anticipated under continued warming.

To this end, we have produced a time series of monthly integrations of inferred surface chlorophyll-a ($[chl-a]$) concentrations in the Barents Sea, and sought to relate them to reanalyses of SST and the frequency of energetic winds. We extend beyond considering 1-dimensional local controls on bloom evolution, by investigating the role of inflow currents that deliver macronutrients onto the Barents Shelf, essential to support primary productivity.

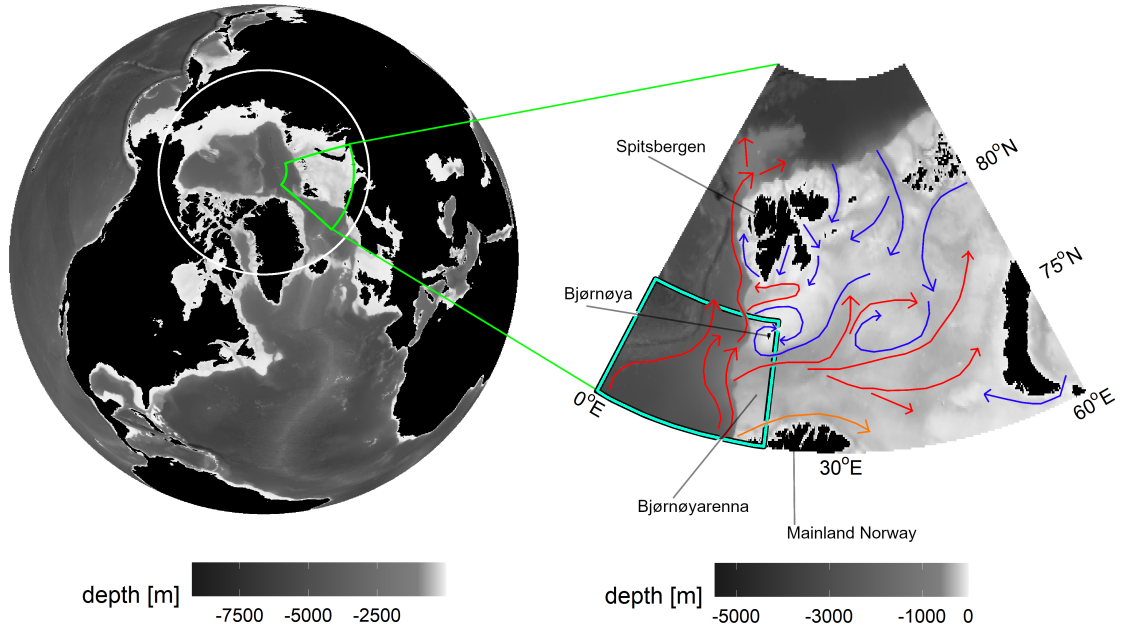


Figure 5.1: Map of the geographic context of the Barents Sea (enclosed in green) in the Global hydrography and Arctic region (encircled in white), and the pattern of recirculating pathways of Atlantic (red), Arctic (blue) and Coastal (orange) Water delivery, guided by its shelf topography. The primary region of Atlantic inflow is surrounded by a teal box. Bathymetric data downloaded and plotted using the R *marmap* package (Pante and Simon-Bouhet 2013).

5.5 Dataset and methods

5.5.1 Satellite data

Daily OC-CCIv5 (Ocean Colour Climate Change Initiative (Sathyendranath et al. 2019)) data were downloaded from the thredds server (<https://www.oceancolour.org/thredds/catalog-cci.html>), for the ocean-colour wavebands $\lambda = 412, 443, 490, 510, 560$ and 665 nm, at a nominal 4 km resolution, for September between 2002 and 2019.

The OC6_MERIS ocean-colour algorithm was employed to estimate $[chl-a]$ concentra-

tions from these ocean-colour data (O'Reilly and Werdell 2019). Coccolithophore blooms, which tend to occur in late summer in the Barents Sea (Oziel et al. 2017) were masked, employing a method based on principal component decomposition, similar to that in Orkney et al. (2020). It was necessary to mask coccolithophore blooms as their calcite liths enhance scattering, resulting in a flattened blue to green ratio that can manifest as apparently high $[chl-a]$ (Table 5.2).

Following the application of a coccolithophore mask, daily $[chl-a]$ estimates were then averaged over each September period, within 1° lat.-lon. grid cells between 0° East; 70° North; 60° East, 85° North and compiled into a stack of 18 images. Individual observations were weighted by their surface area, a function of latitude, when computing the mean value within grid cells.

A preliminary analysis of the temporal evolution of monthly $[chl-a]$ implies that sampling biases, caused by variable ice and cloud cover, do not significantly result in a direct distortion of remotely-sensed $[chl-a]$ estimates in September (Figure 5.7). A 1 year lag between variation in wind field and monthly satellite coverage is evident over the time series however Figure 5.7, showing that there is an interaction between some environmental variables and the possibility of inter-annual persistence. Analyses must therefore be designed to take temporal autocorrelation into account.

5.5.2 Environmental data

We sought monthly environmental data for hydrographic variables from the ARCTIC_REANALYSIS_PHYS_002_003 dataset, which we obtained from CMEMS (Copernicus Marine Environment Monitoring Service: <https://marine.copernicus.eu/>).

ARCTIC_REANALYSIS_PHYS_002_003 is a reanalysis product offering SST, salinity and current speed data, based on a coupled ocean and sea-ice data assimilation (TOPAZ4) and

the HYCOM ocean model, which performs well reproducing surface measurements (Xie et al. 2017).

The SST and salinity data were averaged within 1° lat.-lon. grid cells between 0° East; 70° North; 60° East, 85° North, in order to create a stack of 18 images of September conditions over the Barents Sea. Current speed was averaged in a similar fashion, computed as the Euclidean distance defined by the meridional and longitudinal vectors for each data point, and then averaged within each grid cell.

Daily September wind speed data was also sought from CMEMS, from the WIND_GLO_WIND_L3_REP_OBSERVATIONS_012_005 gridded scatterometer-derived wind speed measurements. Data at a 25 km resolution from ascending passes of the SeaWinds instrument aboard the QuikScat satellite were downloaded for the interval September 2002-2009, while data at a 25 km resolution from ascending passes of the ASCAT instrument of the MetOp-A satellite were downloaded for the interval September 2010-2019. The number of days experiencing wind speeds above 10 ms^{-1} (the metric for energetic wind events used by Ardyna et al. (2014)) was then computed for each unique location, and the mean number of such windy days within each grid cell was computed, taking care to weight observations in accordance with their latitude (as the projection was not stereographic).

A complete stack of 18 images of gridded SST, salinity, current and wind speed data was thus assembled.

The time series of inferred $[chl-a]$, SST, windiness and current speed in the Atlantic inflow region (teal box in Figure 5.1) are presented in Figure 5.2.

5.5.3 Analysis

As oceanic variability can have a year-to-year persistence, it was necessary to design an analytical approach robust to temporal autocorrelation. We decided to use a lag-time of 1 year, as the length of our time series (18 years) is relatively short and we did not expect

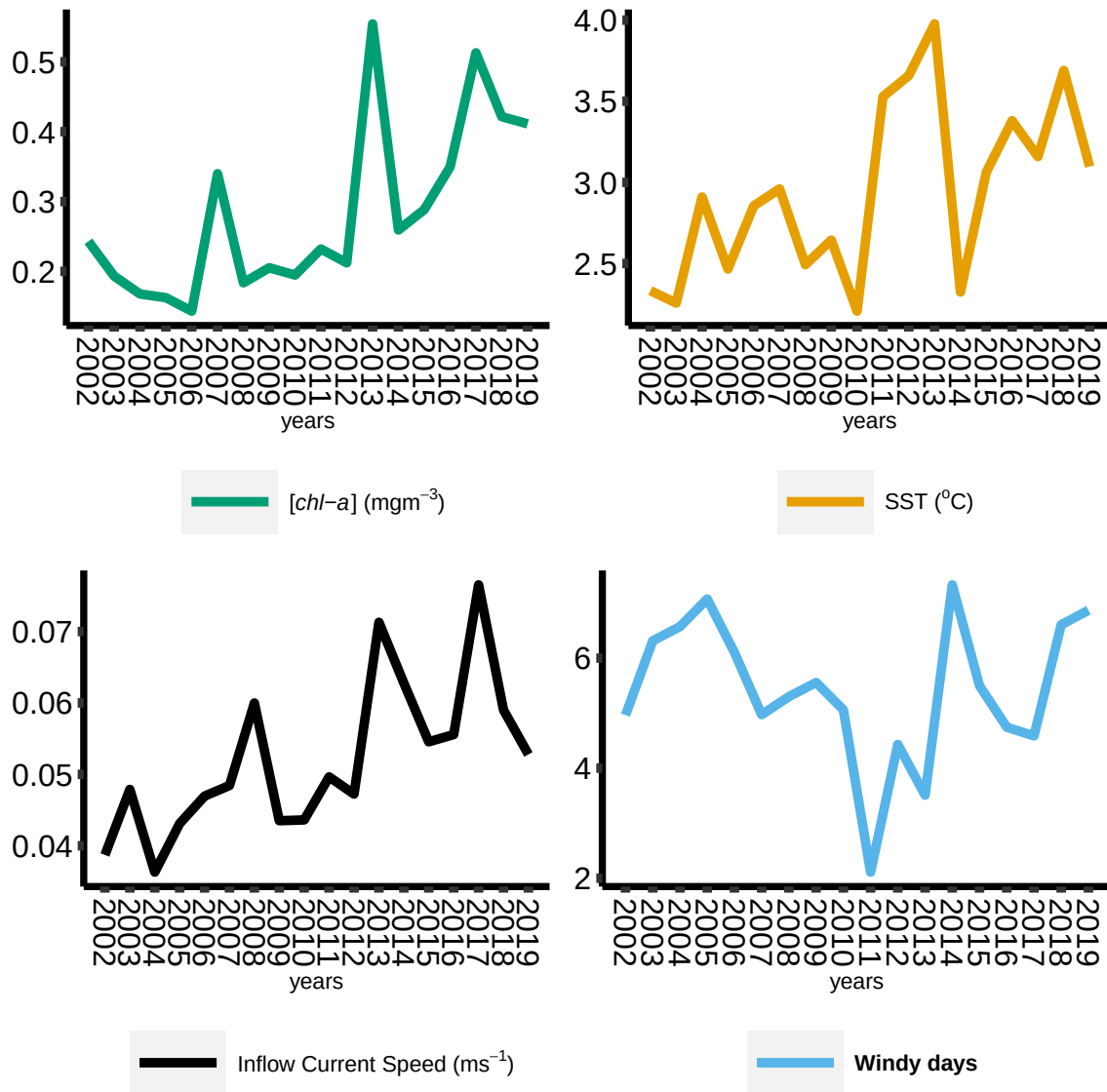


Figure 5.2: Plots of the temporal evolution of inferred $[chl-a]$, SST, windiness and current speed in the Atlantic inflow region, showing long-term trends and patterns of inter-annual variability.

longer modes of autocorrelation to be manifest in these data. Our approach conforms to practice elsewhere in the literature (e.g. (Santer et al. 2008)).

We performed Generalised Least Squares (GLS) regressions with the *gls()* function of the

nlme R package, version 3.1-144 (Pinheiro et al. 2020), both with and without 1 year autocorrelation represented. Autocorrelated fits were implemented with the *corAR1()* function of the R *nlme* package. The R *stats* analysis of variance (*anova()*) function was then used to compare fits and the autocorrelated model was selected if it was significantly favoured at a 0.05 alpha level. An example application of this statistical approach is provided in the mathematical methods section on temporal autocorrelation and hyperlinks are provided to an example code and dataset.

Our initial analysis simply computed the linear trends in the variation of $[chl-a]$, SST, windy days and current speed within grid cells. We then computed the linear relationships between gridded $[chl-a]$ and SST, windy days and current speed, in order to ascertain whether the environmental variables individually explain variation in autumnal bloom intensity. We also computed the relationship between gridded SST and windy days, as we expected that vertical overturning of the ocean, engendered by windy events, might alter SST. This therefore represents a test of the realism of the interaction of reanalyses we employed to represent environmental variability.

We investigated the relationships between $[chl-a]$ and the current speed in the inflow region of the Barents Sea (which we defined as being South of 75° North and West of 20° East-see Figure 5.1); between the inter-annual differences of these two variables; and between the inter-annual differences of $[chl-a]$ and the frequency of windy days. In these final analyses we projected maps of confidence intervals of the geographic occurrence of Atlantic-derived water-masses (defined as having SST $>3^\circ\text{C}$ and a PSU $>34.8\text{‰}$ (Oziel et al. 2016; Våge et al. 2016))

We have also plotted the variance-normalised shelf-wide trends in $[chl-a]$ and the environmental variables, weighted by surface area, so that the variation in the time series at a synoptic scale may be inspected visually.

Generalised Least Squares models were computed for pairwise combinations of $[chl-a]$

and the environmental variables, for Barents-wide time series. GLS models featuring 1 year autocorrelation are presented when preferred at an alpha level of 0.05. The Akaike Information Criterion (AIC) was computed to facilitate model comparison.

5.6 Results

Table 5.1: *Summary of Generalised Least Squares model results.*

	B	SST	C_i	W
B		$AIC = 50.27887$ $p = 0.0059$	$AIC = 43.73159$ $p = 0.0002$	$AIC_A^* = 51.9273$ $p_A^* = 0.0158$
SST			$AIC = 54.91231$ $p = 0.0787$	$AIC^* = 53.54153$ $p^* = 0.0355$
C_i				$AIC_A^* = 55.19371$ $p_A^* = 0.2792$
W				

Note. GLS regressions fitted with the *gls()* function, relating $[chl-a]$ (B), SST, inflow current speed (C_i) and the windy events index (W). An ‘A’ subscript indicates a lag-1 autocorrelated model is preferred at an alpha level of 0.05. ‘*’ superscripts indicate negative relationships. Model fits are described by their Akaike Information Criterion (AIC) and p -values.

We found strong evidence of increasing autumnal bloom intensity in the Barents Sea, (Figure 5.3a), with increases most pronounced near the Atlantic inflow region (Figure 5.1). SST increases were consistent with previous reports in the literature (Figure 5.3b) (Barton et al. 2020). Changes in wind field were uncertain, with the exception of increases in the far North-east (Figure 5.3c), while current speeds showed evidence of increasing inflow onto the shelf (Figure 5.3d) but declining speeds in the Fram Strait. We suggest this pattern of wind and current changes is consistent with the expectation that winds in the Northern Barents Sea regulate pathways of Atlantic Water delivery (Lien et al. 2013).

We found that SST is well-correlated with autumnal bloom intensity, especially close to the inflow region (Figure 5.4a). The pattern of association to windy days simply reproduces secular change (Figure 5.4b), suggesting any relationship is driven by a mutual relationship to time; we do not recover the strong positive relationship reported more widely across the Arctic by Ardyna et al. (2014). Geographically-resolved current speed shows some significant positive association with autumnal bloom intensity in the inflow region (Figure 5.4c). An assessment of wind and SST shows the anticipated physical relationships – windy days

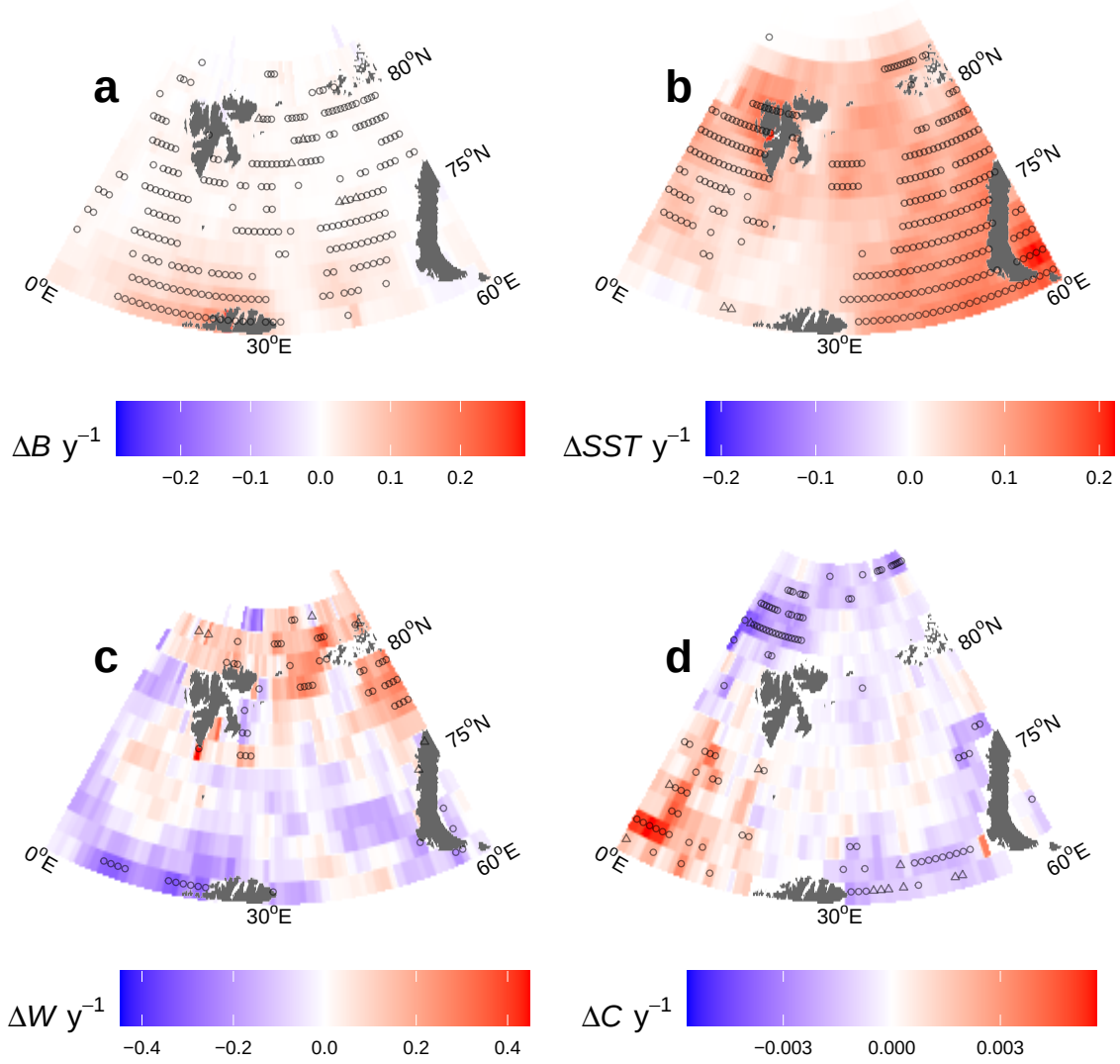
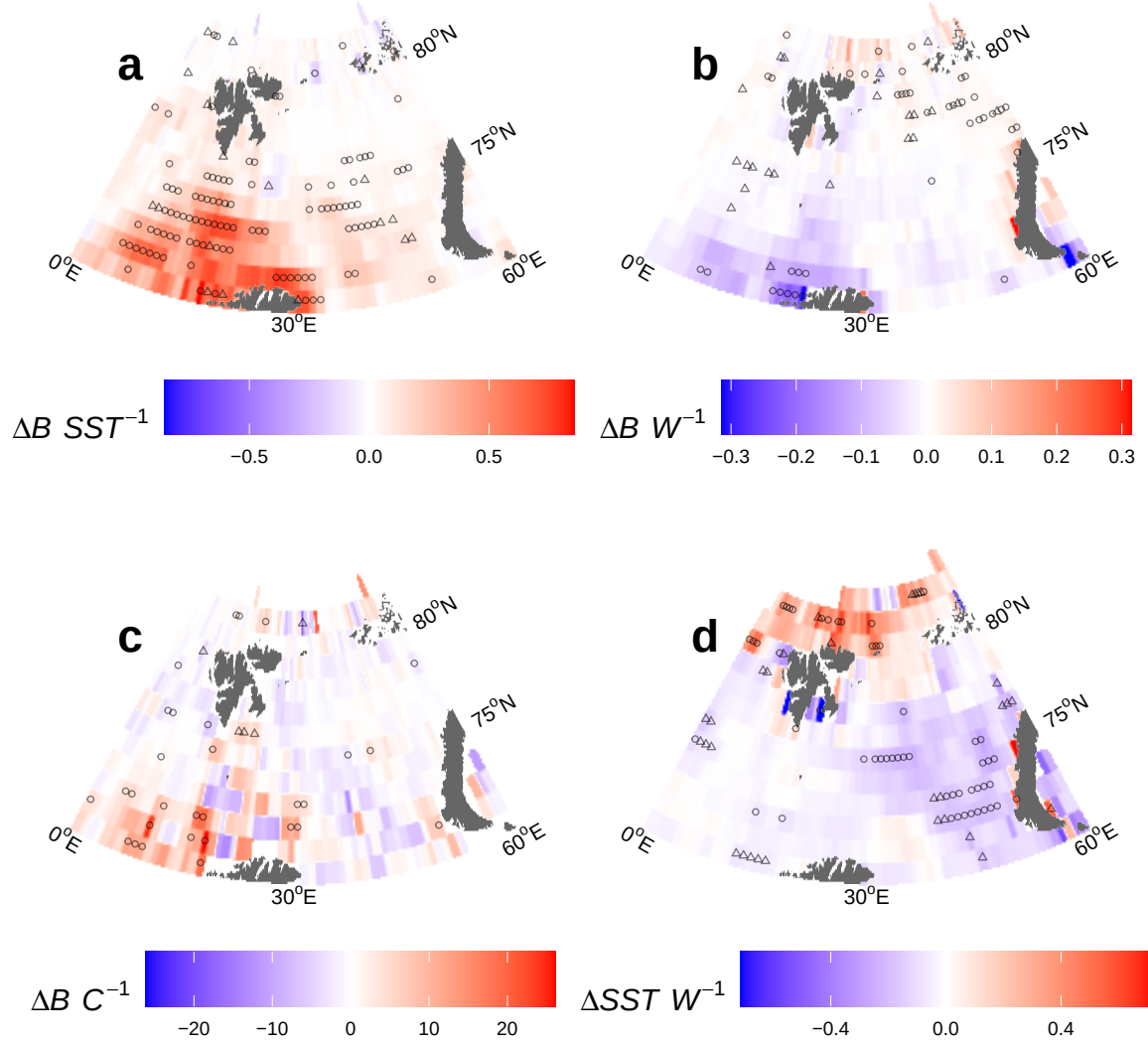


Figure 5.3: Slopes of linear models of the temporal evolution of $[chl-a]$ (B) (mg m^{-3}) (a), Sea Surface Temperature (SST) (Kelvin) (b), windy days (W) (c) and current speed (C) (ms^{-1}) (d), between 2002 and 2019. Stipples mark significant ($p < 0.05$) trends. Triangles indicate an autocorrelated fit is preferred. Secular increases in biomass and SST are evident. Trends in windy days and current speeds show regional complexity.

mix warmer waters to the surface in the far Northern haline-stratified sea, while they retrieve cooler deep waters in the more Southern temperature-stratified seas (Figure 5.4d)



(Carmack 2007).

Inflow current speed (defined as the current speed in the inflow region of Figure 5.1) and autumnal bloom intensity show a strong positive relationship (Figure 5.5a), as do their inter-annual differences (Figure 5.5b), with the strongest signal evident in the Atlantic inflow region. We suggest that a relationship between autumnal bloom intensity and SST (Figure 5.4a) results from SST acting as a tracer of Atlantic Water delivery. The inter-annual differences of autumnal bloom intensity and wind show a slightly negative relationship in the inflow region (Figure 5.5c), indicating that windier years are associated with depressed bloom intensity in this region.

The secular variance-normalised trends in bloom intensity, SST, inflow current speed and windiness are displayed in Figure 5.5d, where it is clear inflow current speed possesses the strongest relationship to bloom intensity. These results are reinforced by our Generalised Least Squares models (Table 5.1), revealing that the lowest p -value and AIC are common to a model relating bloom intensity and inflow current speed. Wind possesses broadly negative relationships to inflow current speed, bloom intensity and SST, with the former two possessing significant auto-correlation, which may suggest a role for multi-annual feedbacks such as the wind-loop (Smedsrud et al. 2013).

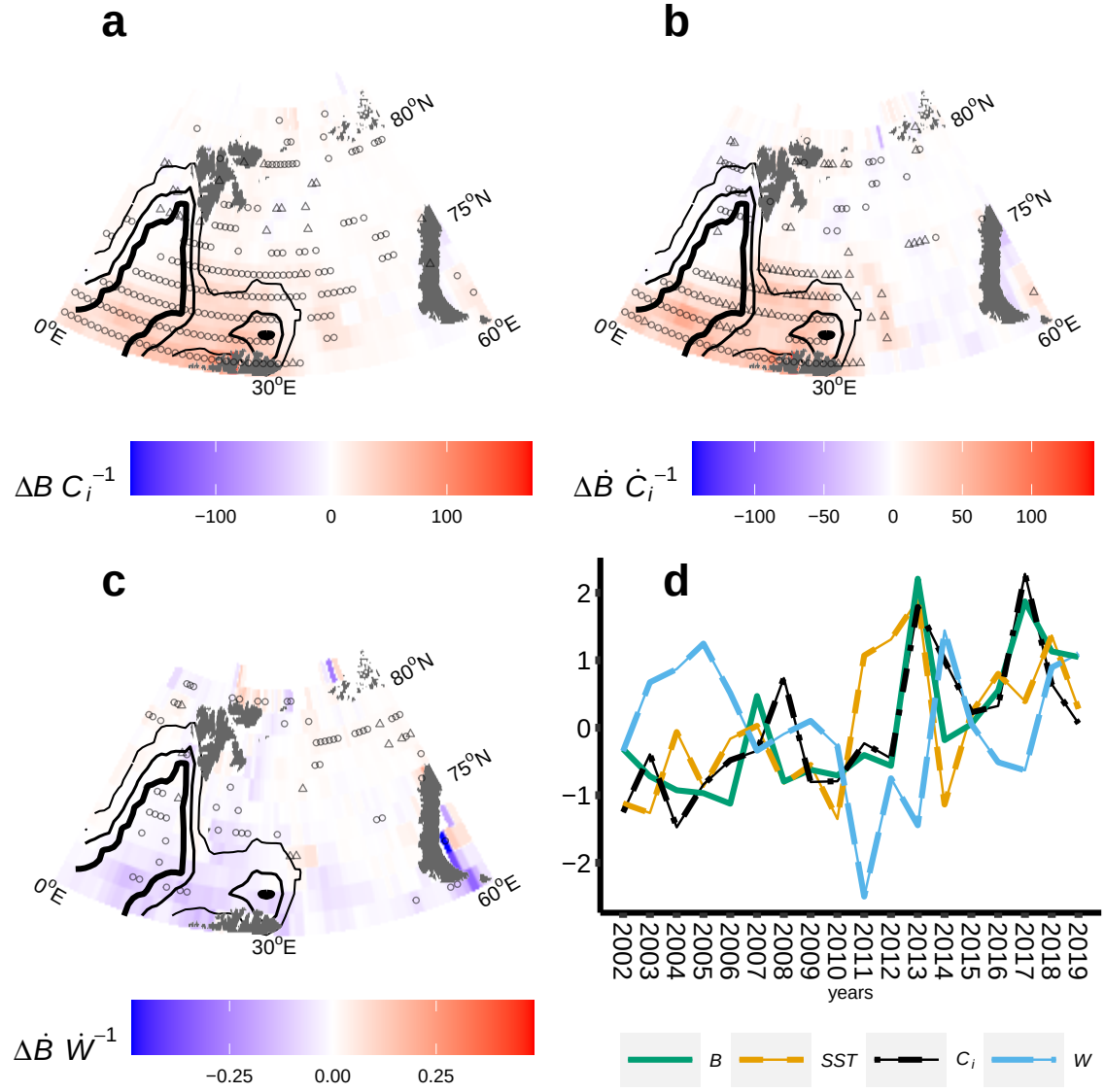


Figure 5.5: Slopes of linear relationships between $[chl-a]$ (B) (mg m^{-3}) and inflow current speed (C_i) (m s^{-1}) (a), the temporal differences of $[chl-a]$ ($\dot{B}$) ($\text{mg m}^{-3} \text{ yr}^{-1}$) and inflow current speed ($\dot{C}_i$) ($\text{m s}^{-1} \text{ yr}^{-1}$) (b), the temporal differences of $[chl-a]$ ($\text{mg m}^{-3} \text{ yr}^{-1}$) and windy days ($\dot{W}$) (c), and the temporal evolution of mean regional variance-normalised $[chl-a]$, SST, inflow current speed and windy days (d). Stipples mark significant ($p < 0.05$) trends. Triangles indicate an autocorrelated fit is preferred. Thin, medium and bold contours map the 5%, 50% and 95% occurrence intervals of the geographic distribution of Atlantic Water in the time series, based on salinity ($>34.8 \text{‰}$) and temperature ($>3^\circ\text{C}$). Strong positive relationships between biomass and Atlantic inflow currents are evident, especially within the region occupied by Atlantic Waters. Inter-annual variability in windiness anti-correlates with biomass.

5.7 Discussion

The frequency of high autumnal winds has not changed significantly across our time series over much of the Barents Sea (Figure 5.3c), meaning a different physical mechanism must be driving autumnal blooms here, compared with those identified by Ardyna et al. (2014).

In addition, alleviation of light limitation – resulting for example from reduced sea-ice extents – cannot be implicated to explain greater autumnal $[chl-a]$, as the Southern Barents Sea is perennially ice free (Shapiro et al. 2003).

We turn to Findlay et al. (2006)’s analysis of a system of three ordinary differential equations, describing development of temperate autumn blooms, to seek alternative hypotheses.

Findlay et al. (2006) states that change in the Mixed Layer Depth, M , can be represented as

$$\frac{dM}{dt} = h(t), \quad (5.1)$$

where t represents the evolution of the system with time, and h is some linear rate of change in the Mixed Layer Depth. Observe that this equation depends on no variables other than time; it is going to force the system.

Change in phytoplankton biomass P can then be represented as

$$\frac{dP}{dt} = \frac{N}{k + N}r(M)P - gP - \frac{m + h(t)}{M}P, \quad (5.2)$$

where the first term is the growth rate attributable to uptake half-saturation k of nutrients N within the context of light-dependent $r(M)$ photosynthesis that declines as the depth of mixing increases. g is simply a general linear loss term, while the third term is composed of loss due to diffusion m beneath the mixed layer, which will evolve as a function of change in its depth $h(t)$.

The inventory of nutrients is, as a consequence of phytoplankton uptake, constrained to

evolve as:

$$\frac{dP}{dt} = -\frac{N}{k+N}r(M)P + \frac{m+h(t)}{M}(N_0 - N), \quad (5.3)$$

where it can be seen that all phytoplankton uptake results in loss, and loss is replaced by diffusion of a deep source of nutrients N_0 , that will evolve as the Mixed Layer Depth changes.

Should the Mixed Layer Depth be held constant ($h(t) = 0$), then the phytoplankton and nutrient equations possess equilibrium solutions.

We can see that, by rearranging to make N the subject of equation 5.2, that we will reach a statement that

$$N \propto \frac{\frac{dP}{dt}}{r(M) - g - d}. \quad (5.4)$$

As we cannot divide by zero, we must therefore state that the equilibrium can only be achieved if light is not so limiting that loss will exceed growth.

We can also reason that N_0 must be sufficiently high that diffusion of new nutrients from the deep sea can support the loss of nutrients taken up by sinking phytoplankton.

We can then see that, at equilibrium, N will take a value determined by N_0 ; decreasing with more efficient phytoplankton uptake ($1/k$), but increasing with loss of phytoplankton through mortality or diffusion out of the mixed layer, and modulated by the difference between potential growth rate $r(M)$ and overall losses.

Phytoplankton concentration at equilibrium must necessarily be positively related to new nutrient, diffusing over a depth defined by the mixed layer, and inversely related to loss or diffusion of phytoplankton out of the mixed layer.

This would imply that a greater Mixed Layer depth would attenuate phytoplankton concentrations by spreading them out over a larger depth interval, decreasing average light availability and increasing the length-scale over which new nutrients must diffuse to reach phytoplankton. We can also see that greater mortality of the phytoplankton will depress

their concentration even further.

Therefore, we can see that – at equilibrium – a deeper mixed layer by itself cannot be expected to establish conditions for high biomass.

Findlay et al. (2006) therefore concludes that autumnal blooms are inherently disequilibrium processes, depending on sudden mixing events. We can permit phytoplankton concentrations to grow, for example, if the upwelled nutrients are injected much faster than sinking phytoplankton particles can remove them from the mixed layer. This is, however, conditional on the mixed layer deepening being shallow enough that illumination remains high enough to permit rapid growth of the phytoplankton (if M extended to infinity there would be a phytoplankton concentration in the surface ocean of 0 and the available light for phytoplankton would be 0). Furthermore, sources of mortality such as grazers, must be small enough so as not to control the evolution of phytoplankton biomass.

Findlay et al. (2006) was therefore able to conclude that autumnal blooms are transient phenomenon that will only occur while change in the Mixed Layer Depth is evolving in a specific parameter space, with a strong influence from the initial conditions established by the previous history of the water column (for example the concentration of deep nutrients, the presence of grazers, existing stocks of phytoplankton and so forth). If the Mixed Layer Depth's evolution should stop, the system will reach equilibrium and therefore phytoplankton stocks will decline to a lower value commensurate with a decreased availability of light.

We can present this in plain English; rapid deepening of the mixed layer may dilute grazer concentrations, or retrieve deep nutrient stocks – both of which may cause a phytoplankton bloom. However, should the same physical conditions develop gradually, then the delivery of deep nutrient stores will occur at a more modest rate and grazer population growth or phytoplankton loss to sinking may keep pace with phytoplankton growth, preventing a bloom.

It is in the context of this mechanistic understanding of autumnal bloom dynamics that

we can reach a fuller appreciation of the reason for the recent emergence of more intense autumnal blooms in the Barents Sea.

We might hypothesise that, in the salinity-stratified North-eastern waters of the Barents Sea, and the interior Eurasian shelves, windy days represent significant excursions from equilibrium and therefore prompt blooms by retrieving a deep nutrient stock. By contrast, in the South-western Barents Sea, windy days are routine. The historic depth of the mixed layer over the season means that the deeper nutrient stock has already been more depleted by periodic upwelling. Therefore a rapid deepening of the mixed layer may not retrieve sufficient new nutrients to trigger autumnal blooms. Instead of rapid mixed layer deepening, the occurrence of autumnal blooms then becomes controlled by the deep stock of nutrients, which is strongly influenced by Atlantic inflow deliveries. This provides an explanation for the clear relationship uncovered here between the inflow current speed – the main source of macronutrients to the Barents Shelf – and the concentration of autumnal $[chl-a]$. We remark that, contrary to the mechanisms outlined by Cushing (1959), Platt et al. (2009), and Ardyna et al. (2014), it is the influence of changes outside the Barents Sea that are controlling these aspects of seasonal cycles of phytoplankton activity, rather than local conditions.

This hypothesis also provides us with an explanation for the relationship between autumnal $[chl-a]$ and SST, especially in the South-west (Figure 5.4a). SST acts as a tracer of the delivery of warm Atlantic-derived waters to the shelf. The single region in the Atlantic inflow that does not exhibit a significant relationship between $[chl-a]$ and inflow current speed is typified by a slight increase in high winds (Figure 5.3c), and inter-annual differences of $[chl-a]$ and storminess anti-correlate significantly here (Figure 5.5c). This suggests that the action of storms diluting surface $[chl-a]$ plays a role in modifying inferred $[chl-a]$ and that it may potentially obscure relationships between $[chl-a]$ and other environmental variables. Alternatively, repeated vertical mixing of the phytoplankton may impose light

limitation, impeding their ability to bloom in spite of high nutrient supplies (Edwards et al. 2004; Platt et al. 2003)

5.7.1 Limitations

The temporal resolution of our analyses may not be fine enough to detect complex high-frequency interactions. For example, wind may cause an instantaneous decrease in $[chl-a]$ by diluting surface $[chl-a]$, but justify a greater $[chl-a]$ later as a result of enhanced nutrient upwelling.

We note that it is understood that the delivery of krill populations, that may graze phytoplankton blooms, is controlled by inflow events (Eriksen et al. 2017), and that grazer delivery may have some relationship to Barents Sea autumnal blooms. However, we consider testing this beyond the limitations of our current research.

We also note that inventories of essential macronutrients in the Barents Sea, such as nitrate (Oziel et al. 2017), and silicic acid (Rey 2012), have been declining – and that these declines have been attributed to upstream processes outside of the Barents Sea (Hátún et al. 2017). This may represent a potential challenge to hypotheses of increasing inflow currents providing enhanced nutrient supply to autumnal blooms in recent years. The significance of nutrient inventory declines depends on whether enhanced delivery of Atlantic Waters is sufficient to overcome nutrient limitation in autumnal conditions.

Between 1980 and 2011, the Atlantic Water volume in the Barents Sea doubled (Oziel et al. 2016), equivalent to a change of 6.45% per year (assuming linearity). End-winter silicate concentrations in the Western Barents Sea decreased from approximately $6\ \mu\text{Mol}$ to $4.5\ \mu\text{Mol}$ between 1990 and 2015. Again, assuming linearity, this is a decline of 1% per year. Nitrate, by contrast, showed a decrease of just 7% over 20 years (Rey 2012).

The Western Barents Sea has been dominated by Atlantic Water through this interval

(Loeng 1991), and nutrient inventories in the Barents Sea can be regarded as being dependent on Atlantic delivery (Tuerena et al. 2021). Therefore, we can observe that, even though there has been a pronounced increase in Atlantic inflow, it has been insufficient to restock silicate over-winter in the Western Barents Sea. Hence, rather than secular increases in concentrations of nutrients below the Mixed Layer Depth, Autumnal bloom occurrence must be modulated either by variation in access to this nutrient stock, or by temporal heterogeneity in nutrient delivery. A greater rate of delivery of Atlantic Waters could ‘top-up’ regions that have become depleted, or could thicken the submarine layer of Atlantic Water in stratified regions (and therefore reduce the depth of Mixing required to retrieve nutrients).

Atlantic inflow can be highly variable (Ingvaldsen et al. 2002), and therefore pulses of greater delivery possess the capacity to explain corresponding pulses in $[chl-a]$ (Figure 5.5 d) even if a secular decline in nutrient inventories is occurring. The secular increase in $[chl-a]$ would only be incompatible with declining Atlantic nutrient inventories, if inflowing concentrations should fall below the concentrations limiting for phytoplankton growth ($2 \mu\text{Mol}$ for diatoms (Egge and Aksnes 1992)). As a value of $4.5 \mu\text{Mol}$ is well above this threshold, our results do not appear incompatible with secular trends in nutrient inventories in the Atlantic inflow waters – although we must comment that end-winter inventories represent the integration of an entire season of nutrient resupply.

Our remote-sensing approach and use of a global algorithm, agnostic to phytoplankton bio-optical variation, may mean that some variation in perceived biomass is driven by variability in phytoplankton community composition (Dutkiewicz et al. 2019), or that variation in cloud and ice fields influences the likelihood of perceiving intense blooms. In addition, variation in photoacclimation (Johnsen and Sakshaug 2007) state may significantly alter intracellular pigment concentrations (Garrido and Zapata 1998), which could result in bias in the perception of remotely-sensed $[chl-a]$. We, furthermore, are unable to investigate

ice-affiliated modes of production, or sub-surface blooms, which develop over summer and may be re-invigorated by upwelling of deep nutrients in autumn.

5.8 Summary

Our findings highlight the unique hydrography and climatology of the Barents Sea, compared with the rest of the Arctic. The great Atlantic influence in this Arctic gateway means that conventional local mechanisms routinely invoked in the literature to explain phytoplankton dynamics of seasonal seas, such as rapid mixed layer deepening, have limited influence here. Instead, it is upstream processes outside of the Barents Sea that ultimately control the occurrence of autumnal blooms by delivering Atlantic Waters. We hypothesise this results from dependency on Atlantic Waters for the delivery of macronutrient reserves essential to support blooms. As the Barents Sea is one of the most productive regions in the Arctic, our findings establish a mandate to extend mechanistic understanding of bloom dynamics beyond 1-dimensional models.

5.9 Author contributions

Environmental datasets were obtained from CMEMS (Copernicus Marine Environment Monitoring Service: <https://marine.copernicus.eu/>), and hence we acknowledge the service and its European Union funding.

The datasets accessed were as follows:

ARCTIC_REANALYSIS_PHYS_002_003

(Physical reanalysis product for hydrographic conditions)

WIND_GLO_WIND_L3_REP_OBSERVATIONS_012_005

(Wind speed measurements)

Ocean-colour data were downloaded from the OC-CCIv5 project thredds server:

<https://www.oceancolour.org/thredds/catalog-cci.html>

We acknowledge research funding from the Natural Environment Research Council (NERC)

Grant number: NE/P006507/1

Andrew Orkney designed and conducted the analysis and wrote the manuscript.

Shubha Sathyendranath provided expert input on remotely-sensed ocean-colour, biological dynamics of Arctic Waters and analytical design.

Thomas Jackson provided expert input on the processing of remotely-sensed ocean-colour.

Marie Porter provided expert input on hydrography, physical setup and dynamics of the Barents Sea.

Heather Bouman supervised, provided expert input on biological dynamics and analytical design.

All authors read, made contributions to and approved the final manuscript.

5.10 Supplement

5.10.1 Coccolithophore Mask

We rotated ocean-colour data onto a set of two new axes by multiplying the 6 *Rrs* OC-CCIv5 remotely-sensed reflectance values by a series of wavelength-specific (λ) coefficients, summing, and subtracting a constant;

$$S_i = (\sum_{\lambda=1}^6 a_{i\lambda} Rrs_{\lambda}) - b_i, \quad (5.5)$$

where S_i is the result, a_i represents the slew of coefficients and b_i is a constant.

Two sets of values were used (Table 5.2), such that two series of values, S_1 and S_2 were defined. Thereafter, all ocean-colour data, for which $S_2 > -0.5 S_1 + 0.0015$ and $S_2 < 0.4 S_1 + 0.003$, were considered to represent coccolithophores and masked. An exam-

Table 5.2: Wavelengths and coefficients employed to mask coccolithophores in ocean-colour data.

λ (nm)	412	443	490	510	560	665
a_1	0.47091009	0.48800140	0.49696639	0.43672448	0.31694570	0.04404408
a_2	-0.5822403	-0.3063442	0.1096942	0.3814065	0.6141931	0.1800353

Note. $b_1 = 0.008820334$ and $b_2 = -0.0005476582$.

ple of the application of this algorithm to ocean-colour data is provided in Figure 5.6.

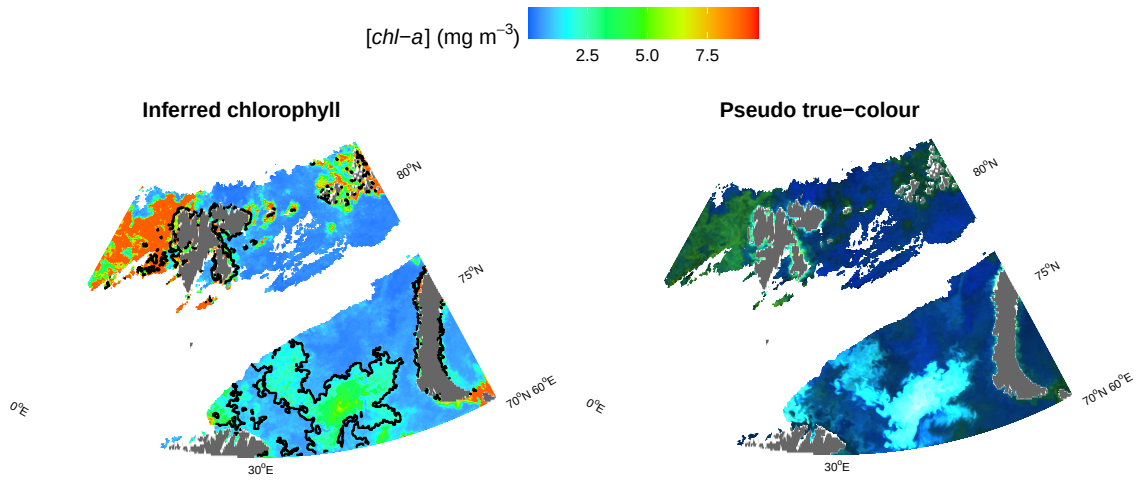


Figure 5.6: Maps of $[chl-a]$ inferred from application of OC6_meris to OC-CCIv5 data in the Barents Sea on the 6th of July 2016 (**left**) and a pseudo true-colour representation (**right**). An intense coccolithophore bloom is seen in the South-eastern Barents Sea (Hansen et al. 2016), and corresponds to a local high in inferred $[chl-a]$. Our custom coccolithophore mask is seen as a black contour in the left panel, correctly identifying this feature. $[chl-a]$ has been capped below 9 mg m^{-3} in this image, to aid visualisation.

5.10.2 Preliminary analysis

Prior to analysis we explored the relationship between retrieved $[chl-a]$ and sample completeness, in order to probe potential sources of bias in remotely-sensed $[chl-a]$ integrations,

so that we could have confidence they do not substantially affect the dataset. We contrived a sampling completeness ‘coverage metric’, by recording the total number of unique latitude-longitude combinations perceived in any monthly increment. Should a persistent ice or cloud field be present, that may prevent the observation of a strong bloom, we should expect a relationship to exist between average monthly $[chl-a]$ inferred by remote-sensing and variation in the coverage metric.

Simple linear regressions were constructed between time series of variance-normalised $[chl-a]$ and the coverage metric within each September, and the inter-annual differences.

As seen in Figure 5.7, a significant increase in average monthly $[chl-a]$ values is evident across our time series, consistent with reports of increasing biomass in the Barents Sea (e.g. (Siswanto 2020; Lewis et al. 2020)).

We note with concern however that the average September normalised coverage metric is also increasing across our time series, which implies that some increases in perceived $[chl-a]$ may not reflect an increasing incidence of blooms, but simply an increasing likelihood of remote sensors observing them as a consequence, for example, of declining ice or cloud cover.

Overall, relationships between linear change in coverage and $[chl-a]$, and their inter-annual differences, are not associated with high R^2 values or significant p -values. (secular trends: $R^2 = 0.1519$, $p = 0.06144$, interannual differences: $R^2 = 0.05088$, $p = 0.193$) We interpret this as being compatible with the low sea-ice extents that normally evolve by September, meaning that remote-sensing sampling completeness is more equitable between years, as variable ice cover is not a substantial driver.

Lastly, noting patterns of inter-annual correlation and the possibility of periodicity in $[chl-a]$ and coverage, we seek to relate variation in September coverage to SST, storminess and inflow current speed. Storminess was defined by the average number of days within

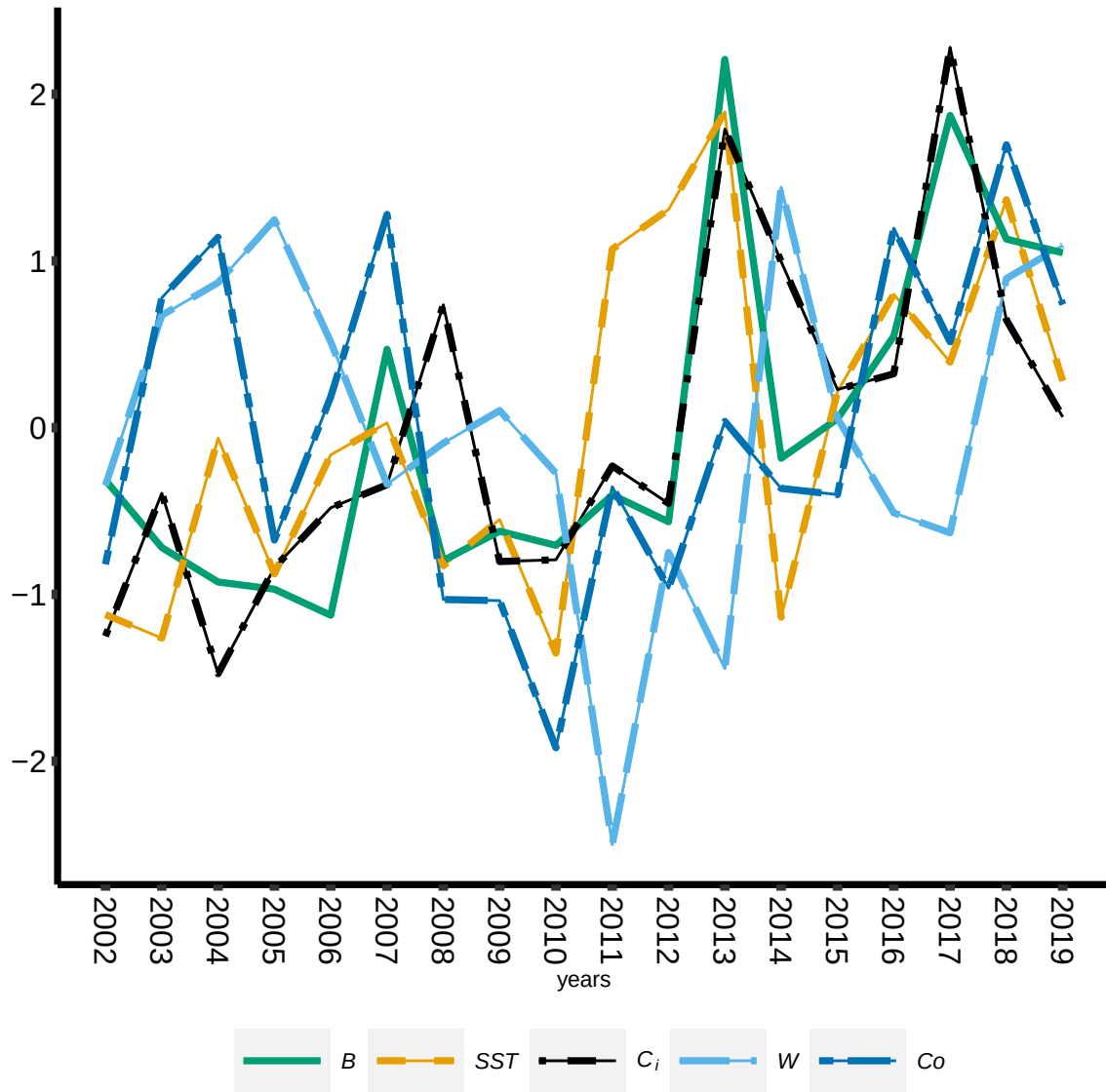


Figure 5.7: Variance-normalised monthly $[chl-a]$ (B) inferred from application of OC6_meris to OC-CCIv5 data in the Barents Sea, compared with the variance-normalised coverage metric (Co), storminess index (W), SST and inflow current speed (C_i).

1°lat.-lon. grid cells, with average wind speeds $>10\text{ ms}^{-1}$, based on

WIND_GLO_WIND_L3_REP_OBSERVATIONS_012_005 gridded scatterometer-derived wind speed measurements, sought through CMEMS (see main manuscript for acknowledgement).

SST and inflow current speed were derived from the ARCTIC_REANALYSIS_PHYS_002_003 reanalysis product, sought through CMEMS. Only the region between 70-75°North and 0-20°East was considered in the calculation of inflow current speed.

The results of time series analysis (Figure 5.7) show that there are not only relationships between secular trends among the different variables, but that there are also lagged relationships, evidencing complex multi-annual resonances and feedbacks. For example, in a simple linear regression, variation in the storminess index is significantly related to the coverage metric of the previous year ($p = 0.017$). This is consistent with hypotheses of a positive ‘wind loop’ feedback, whereby high Atlantic inflow causes high open ocean area – hence a high coverage metric – and thereby facilitates greater loss of oceanic heat to the atmosphere, resulting in low-pressure systems that draw in further strong Westerly winds (Smedsrud et al. 2013). The 1 year lag in our time series suggests that the wind loop feedback has an inter-annual persistence. It can also be seen that storminess and SST excursions anti-correlate, consistent with higher winds engendering greater open ocean area and vertical mixing and hence enhancing heat loss to the atmosphere. There is an exceptionally close match between inflow current speed and perceived [*chl-a*]. We discuss the relationships between perceived biomass and environmental variability in greater detail in our main manuscript.

Suffice to say, from the evidence apparent here, open ocean area, a significant contributor to coverage metric, is significantly related to multi-annual feedback mechanisms controlling environmental variability in the Barents Sea. The potential bias that coverage has in the perception of mean monthly [*chl-a*] cannot therefore be easily disentangled from potential mechanistic drivers of [*chl-a*] variance, without the use of repeated *in-situ* measurements. We conclude, for this reason, that it is best not to endeavour to remove signal associated with variation in coverage. However, we do note that, given multi-annual feedbacks between

environmental variables are evident, that it is necessary to account for temporal autocorrelation before analysis can decipher the mechanistic drivers of autumnal [*chl-a*] variation.

References

- Ardyna, Mathieu, Marcel Babin, Michel Gosselin, Emmanuel Devred, Luc Rainville, and Jean-Éric Tremblay. 2014. "Recent Arctic Ocean sea ice loss triggers novel fall phytoplankton blooms." *Geophysical Research Letters* 41 (17): 6207–6212.
- Båmstedt, Ulf, Hans Christian Eilertsen, Kurt S Tande, Dag Slagstad, and Hein Rune Skjoldal. 1991. "Copepod grazing and its potential impact on the phytoplankton development in the Barents Sea." *Polar Research* 10 (2): 339–354.
- Barton, Benjamin I, Yueng-Djern Lenn, and Camille Lique. 2020. "Observed atlantification of the Barents Sea causes the Polar Front to limit the expansion of winter sea ice." *JPO* 500:18–0003.
- Boitsov, Vladimir D, Alexey L Karsakov, and Alexander G Trofimov. 2012. "Atlantic water temperature and climate in the Barents Sea, 2000–2009." *ICES Journal of Marine Science* 69 (5): 833–840.
- Carmack, Eddy. 2007. "The alpha/beta ocean distinction: A perspective on freshwater fluxes, convection, nutrients and productivity in high-latitude seas." *Deep Sea Research Part II: Topical Studies in Oceanography* 54 (23–26): 2578–2598.
- Carmack, Eddy, David Barber, Jens Christensen, Robie Macdonald, Bert Rudels, and Egil Sakshaug. 2006. "Climate variability and physical forcing of the food webs and the carbon budget on panarctic shelves." *Progress in Oceanography* 71 (2–4): 145–181.
- Cavalieri, Donald J, and Claire L Parkinson. 2012. "Arctic sea ice variability and trends, 1979–2010." *The Cryosphere* 6 (4): 881.
- Cushing, DH. 1959. "The seasonal variation in oceanic production as a problem in population dynamics." *ICES Journal of Marine Science* 24 (3): 455–464.
- Dong, Kaixing, Kristina Øie Kvile, Nils Chr Stenseth, and Leif Chr Stige. 2020. "Associations among temperature, sea ice and phytoplankton bloom dynamics in the Barents Sea." *Marine Ecology Progress Series* 635:25–36.
- Dutkiewicz, Stephanie, Anna E Hickman, Oliver Jahn, Stephanie Henson, Claudie Beaulieu, and Erwan Monier. 2019. "Ocean colour signature of climate change." *Nature communications* 10 (1): 578.
- Edwards, Andrew M, Trevor Platt, and Shubha Sathyendranath. 2004. "The high-nutrient, low-chlorophyll regime of the ocean: limits on biomass and nitrate before and after iron enrichment." *Ecological modelling* 171 (1–2): 103–125.
- Egge, JK, and DL Aksnes. 1992. "Silicate as regulating nutrient in phytoplankton competition." *Marine ecology progress series. Oldendorf* 83 (2): 281–289.
- Eriksen, Elena, Hein Rune Skjoldal, Harald Gjøsæter, and Raul Primicerio. 2017. "Spatial and temporal changes in the Barents Sea pelagic compartment during the recent warming." *Progress in Oceanography* 151:206–226.
- Findlay, Helen S, Andrew Yool, Marianna Nodale, and Jonathan W Pitchford. 2006. "Modelling of autumn plankton bloom dynamics." *Journal of plankton research* 28 (2): 209–220.
- Francis, Jennifer A, and Elias Hunter. 2007. "Drivers of declining sea ice in the Arctic winter: A tale of two seas." *Geophysical Research Letters* 34 (17).
- Garrido, José L, and Manuel Zapata. 1998. "Detection of new pigments from *Emiliania Huxleyi* (prymnesiophyceae) by high-performance liquid chromatography, liquid chromatography–mass spectrometry, visible spectroscopy, and fast atom bombardment mass spectrometry." *Journal of phycology* 34 (1): 70–78.
- González Taboada, Fernando, and Ricardo Anadón. 2014. "Seasonality of North Atlantic phytoplankton from space: impact of environmental forcing on a changing phenology (1998–2012)." *Global change biology* 20 (3): 698–712.

- Hansen, Kathryn, Jeff Schmaltz, and Joshua Stevens. 2016. "The Barents Sea Abloom." Accessed: 2020-03-03, <https://earthobservatory.nasa.gov/images/88316/the-barents-sea-abloom>.
- Hátún, Hjálmar, K Azetsu-Scott, R Somavilla, Francisco Rey, C Johnson, Moritz Mathis, Uwe Mikolajewicz, P Coupel, J-É Tremblay, S Hartman, et al. 2017. "The subpolar gyre regulates silicate concentrations in the North Atlantic." *Scientific reports* 7 (1): 1–9.
- Ingvaldsen, Randi, Harald Loeng, and Lars Asplin. 2002. "Variability in the Atlantic inflow to the Barents Sea based on a one-year time series from moored current meters." *Continental Shelf Research* 22 (3): 505–519.
- Johnsen, Geir, and Egil Sakshaug. 2007. "Biooptical characteristics of PSII and PSI in 33 species (13 pigment groups) of marine phytoplankton, and the relevance for pulse-amplitude-modulated and fast-repetition-rate fluorometry 1." *Journal of Phycology* 43 (6): 1236–1251.
- Karcher, Michael, John N Smith, Frank Kauker, Rüdiger Gerdes, and William M Smethie. 2012. "Recent changes in Arctic Ocean circulation revealed by iodine-129 observations and modeling." *Journal of Geophysical Research: Oceans* 117 (C8).
- Kwok, Ron. 2018. "Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018)." *Environmental Research Letters* 13 (10): 105005.
- Lewis, KM, GL van Dijken, and KR Arrigo. 2020. "Changes in phytoplankton concentration now drive increased Arctic Ocean primary production." *Science* 369 (6500): 198–202.
- Lien, Vidar S, Frode B Vikebø, and Øystein Skagseth. 2013. "One mechanism contributing to co-variability of the Atlantic inflow branches to the Arctic." *Nature Communications* 4 (1): 1–6.
- Loeng, Harald. 1991. "Features of the physical oceanographic conditions of the Barents Sea." *Polar Research* 10 (1): 5–18.
- O'Reilly, John E, and P Jeremy Werdell. 2019. "Chlorophyll algorithms for ocean color sensors-OC4, OC5 & OC6." *Remote sensing of environment* 229:32–47.
- Orkney, Andrew, Trevor Platt, Bhavani E Narayanaswamy, Ina Kostakis, and Heather A Bouman. 2020. "Bio-Optical evidence for increasing *Phaeocystis* dominance in the Barents Sea." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*.
- Oziel, Laurent, G Neukermans, M Ardyna, C Lancelot, J-L Tison, Paul Wassmann, Jérôme Sirven, Diana Ruiz-Pino, and J-C Gascard. 2017. "Role for Atlantic inflows and sea ice loss on shifting phytoplankton blooms in the Barents Sea." *Journal of Geophysical Research: Oceans* 122 (6): 5121–5139.
- Oziel, Laurent, Jérôme Sirven, and Jean-Claude Gascard. 2016. "The Barents Sea frontal zones and water masses variability (1980–2011)." *Ocean Science* 12(1):169–184.
- Pante, Eric, and Benoit Simon-Bouhet. 2013. "marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R." *PLoS ONE* 8 (9): e73051. [doi:10.1371/journal.pone.0073051](https://doi.org/10.1371/journal.pone.0073051).
- Pinheiro, Jose, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, and R Core Team. 2020. *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-144. <https://CRAN.R-project.org/package=nlme>.
- Platt, Trevor, David S Broomhead, Shubha Sathyendranath, Andrew M Edwards, and Eugene J Murphy. 2003. "Phytoplankton biomass and residual nitrate in the pelagic ecosystem." *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 459 (2033): 1063–1073.
- Platt, Trevor, George N White III, Li Zhai, Shubha Sathyendranath, and Shovonlal Roy. 2009. "The phenology of phytoplankton blooms: Ecosystem indicators from remote sensing." *Ecological Modelling* 220 (21): 3057–3069.

- Polyakov, Igor V, Andrey V Pnyushkov, Matthew B Alkire, Igor M Ashik, Till M Baumann, Eddy C Carmack, Ilona Goszczko, John Guthrie, Vladimir V Ivanov, Torsten Kanzow, et al. 2017. "Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean." *Science* 356 (6335): 285–291.
- Rey, Francisco. 2012. "Declining silicate concentrations in the Norwegian and Barents Seas." *ICES Journal of Marine Science* 69 (2): 208–212.
- Sakshaug, E. 2004. "Primary and secondary production in the Arctic Seas." In *The organic carbon cycle in the Arctic Ocean*, 57–81. Springer.
- Santer, Benjamin D, PW Thorne, L Haimberger, Karl E Taylor, TML Wigley, JR Lanzante, Susan Solomon, M Free, Peter J Gleckler, PD Jones, et al. 2008. "Consistency of modelled and observed temperature trends in the tropical troposphere." *International Journal of Climatology: A Journal of the Royal Meteorological Society* 28 (13): 1703–1722.
- Sathyendranath, Shubha, Robert JW Brewin, Carsten Brockmann, Vanda Brotas, Ben Calton, Andrei Chuprin, Paolo Cipollini, André B Couto, James Dingle, Roland Doerffer, et al. 2019. "An Ocean-Colour Time Series for Use in Climate Studies: The Experience of the Ocean-Colour Climate Change Initiative (OC-CCI)." *Sensors* 19 (19): 4285.
- Shapiro, Inna, Roger Colony, and Torgny Vinje. 2003. "April sea ice extent in the Barents Sea, 1850–2001." *Polar Research* 22 (1): 5–10.
- Simmonds, Ian, and Kevin Keay. 2009. "Extraordinary September Arctic sea ice reductions and their relationships with storm behavior over 1979–2008." *Geophysical Research Letters* 36 (19).
- Siswanto, Eko. 2020. "Temporal variability of satellite-retrieved chlorophyll-a data in Arctic and subarctic ocean regions within the past two decades." *International Journal of Remote Sensing* 41 (19): 7427–7445.
- Smedsrud, Lars H, Igor Esau, Randi B Ingvaldsen, Tor Eldevik, Peter M Haugan, Camille Li, Vidar S Lien, Are Olsen, Abdirahman M Omar, Odd H Otterå, et al. 2013. "The role of the Barents Sea in the Arctic climate system." *Reviews of Geophysics* 51 (3): 415–449.
- Strass, Volker H, and E-M Nöthig. 1996. "Seasonal shifts in ice edge phytoplankton blooms in the Barents Sea related to the water column stability." *Polar Biology* 16 (6): 409–422.
- Tuerena, Robyn E, Joanne Hopkins, Raja S Ganeshram, Louisa Norman, Camille de la Vega, Rachel Jeffreys, and Claire Mahaffey. 2021. "Nitrate assimilation and regeneration in the Barents Sea: insights from nitrate isotopes." *Biogeosciences* 18 (2): 637–653.
- Våge, Kjetil, Robert S Pickart, Vladimir Pavlov, Peigen Lin, Daniel J Torres, Randi Ingvaldsen, Arild Sundfjord, and Andrey Proshutinsky. 2016. "The Atlantic Water boundary current in the Nansen Basin: Transport and mechanisms of lateral exchange." *Journal of Geophysical Research: Oceans* 121 (9): 6946–6960.
- Wassmann, Paul, Tatjana Ratkova, Inger Andreassen, Maria Vernet, Gunnar Pedersen, and Francisco Rey. 1999. "Spring bloom development in the marginal ice zone and the central Barents Sea." *Marine Ecology* 20 (3-4): 321–346.
- Wassmann, Paul, Marit Reigstad, Tore Haug, Bert Rudels, Michael L Carroll, Haakon Hop, Geir Wing Gabrielsen, Stig Falk-Petersen, Stanislav G Denisenko, Elena Arashkevich, et al. 2006a. "Food webs and carbon flux in the Barents Sea." *Progress in Oceanography* 71 (2-4): 232–287.
- Xie, Jiping, Laurent Bertino, François Counillon, Knut A Lisæter, and Pavel Sakov. 2017. "Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013." *Ocean Science* 13 (1).

Chapter 6

Conclusion and future work

6.1 Conclusion

A leading conclusion of this thesis is that functional relationships and differences perceived in different phytoplankton assemblages are shaped by the methodological approaches used to describe *in-situ* samples (Sosik et al. 2014). Additionally, while researchers may have predetermined schemes of Phytoplankton Functional Types they wish to seek in remotely-sensed imagery, such as size-structure (Devred et al. 2011; Liu et al. 2018), the necessity that they must exploit ocean surface reflectance to infer this structure requires that there is a robust relationship between the desired structure and bio-optical variety. However, Inherent Optical Properties of the water column have a multifactorial dependence on a great number of biological properties of phytoplankton assemblages, including their morphology and pigment complement (Liu et al. 2019), which can both co-vary with eco-physiological condition of the organisms, such as their photoacclimation state (Barros et al. 2003; Johnsen and Sakshaug 2007; Lewis et al. 2019) and the availability of nutrients. Variation in eco-physiology therefore means that relationships between size-structure, pigment ratios and $a_{ph}(\lambda)$ can vary in complex ways, that are not a straight-forward consequence of community

composition. Furthermore, the formation of colonies – which may even involve interspecific interactions (Sazhin et al. 2007) – can alter bio-optical properties such as absorption and backscattering (Assmy et al. 2017). Colony formation is not normally represented in commonly-employed indices of size-structure such as cell-counts (e.g. Degerlund and Eilertsen (2010)). Colony formation can strongly influence chlorophyll-specific absorption; chapters 2 and 3 both demonstrate that colonial *Phaeocystis* is optically distinct compared to unicellular nanophytoplankton communities.

In addition, variable concentrations of substances with wavelength-dependent bio-optical properties, such as CDOM, are often assumed to covary with $[chl-a]$ in the open ocean. However, their concentration may become decoupled from phytoplankton biomass in some productive shelf-environments, such as the Barents Sea (Hancke et al. 2014). This may act to further obfuscate relationships between phytoplankton community structure and ocean-colour. This represents an additional challenge to researchers seeking to infer aspects of community structure, such as size-classes, from remotely-sensed ocean-colour.

This thesis presents a case-study in the Barents Sea, which quantifies differences in inferred functional relationships, resulting from various representations of phytoplankton community composition, providing researchers with a clear example that demonstrates the influence of observational methods on the perception of Functional Types. While first-order functional relationships are broadly found to be similar, regardless of the dataset used to describe community composition, cell-counts yield a greater number of significant environmental associations than bio-optics or pigment ratios. This may be attributed to diatom species with specific environmental niches, that are not distinct from one another in their pigment-ratios or bio-optical properties. It is also shown that cell-counts possess a relatively low capacity to predict bio-optical functional assemblages. This may reflect the routine practice of excluding abundant colonial forms, such as *Phaeocystis*, from cell-counts. The analyses shows that colonial assemblages play an important role in overall bio-optical

variety in the Barents Sea.

Analyses identifying distinct bio-optical phytoplankton assemblages in the Barents Sea from measurements of $a_{ph}(\lambda)$ underpinned the development of a bio-optical algorithm to detect dominant Phytoplankton Functional Types in remotely-sensed ocean-colour. The pre-eminent conclusion of this work was that some phytoplankton assemblages can be distinguished much more reliably than others. The difference in bio-optical properties between diatom and nanophytoplankton assemblages was stark, for example, as a result of pervasively stronger pigment packaging in diatoms. However, diatoms and colonial *Phaeocystis* samples in the Arctic **PRIZE** $a_{ph}(\lambda)$ dataset both possessed strong pigment packaging. These two assemblages were most clearly distinguished by a stronger blue absorption in *Phaeocystis* at around 469 nm – which probably resulted in part from chlorophyll-c3, which HPLC analysis demonstrated occurred at higher concentrations when this group was present. This difference in spectral signatures, between diatoms and *Phaeocystis*, was relatively small. It is important to consider variability in pigment packaging and the balance of intracellular pigment species is observed as a photoacclimatory response across phytoplankton (Rodriguez et al. 2006; Johnsen and Sakshaug 2007; Brunet et al. 2011; Kauko et al. 2019; Lewis et al. 2019), and therefore that different life-histories may result in variation in the spectral signatures of the phytoplankton assemblages identified in this thesis. High pigment packaging can occur as a response both to low light (Lewis et al. 2019) and as a self-shading protective mechanism in high light (Barros et al. 2003). The availability of essential nutrients, such as nitrate, can control the ability of phytoplankton assemblages to photoacclimate, by inhibiting the ability of phytoplankton to manufacture key proteins in their photosynthetic machinery (Lewis et al. 2019). Therefore, the spectral signatures identified in the Arctic **PRIZE** $a_{ph}(\lambda)$ dataset are not guaranteed to be reproduced in future sampling campaigns, especially if the nature of the bio-geochemical and light environments

in the Barents Sea are subject to change.

I also note that, while chlorophyll-c3 was employed as a marker pigment for *Phaeocystis*-dominated assemblages in this thesis, there are a number of other phytoplankton groups that possess this pigment, such as some diatoms, dinoflagellates and numerous other prymnesiophytes (Johnsen et al. 1994; Johnsen and Sakshaug 2000; Roy et al. 2011).

It was clear that the ocean-colour reprocessing used (MODIS-Aqua) was specifically curated to optimise retrieval of $[chl-a]$, rather than preserving individual waveband ratios. It was therefore necessary to contrive a correction scheme, so that inter-annual drift in the MODIS-Aqua reprocessing did not exceed the fine margin between diatom and *Phaeocystis* spectral signatures. This is a valuable case study for future researchers, who may seek to develop bio-optical algorithms for the recognition of Functional Types from remote-sensing platforms that are optimised for the retrieval of derived variables, such as $[chl-a]$. Thus, while different phytoplankton assemblages may have distinct functional roles and contributions to trophic structure and biogeochemical cycling, the bio-optical structure between phytoplankton assemblages exerts a strong influence on functional relationships that are perceptible to researchers exploiting remotely-sensed ocean-colour.

It must be emphasised that the seasonal Arctic **PRIZE** bio-optical dataset was limited to measurements of $a_{ph}(\lambda)$, and that variation in scattering and backscattering among different phytoplankton assemblages can be significant (Pavlov et al. 2017). Unfortunately, it was not possible to collect a comprehensive set of scattering and backscattering measurements to match the $a_{ph}(\lambda)$ dataset, and this must necessarily therefore be identified as a subject of potential future research. Furthermore, previous authors have demonstrated that the integration of $a_{ph}(\lambda)$ data with fluorescence-emission spectra (quanta of light re-radiated as a consequence of absorption by and excitation of bio-molecules) can provide useful information about the physiological state of phytoplankton (Johnsen and Sakshaug 2007). Just as with $a_{ph}(\lambda)$, fluorescence-excitation varies as a function of the variety in intracellular pigments

and the way in which they are complexed with proteins. Light harvesting pigments transfer energy to the photosynthetic reaction centres, which are excited and consequently fluoresce. *In-vivo* fluorescence-excitation does not carry a signal of photoprotective pigments, (as a consequence of these pigment species not transferring energy to the photosystem II reaction centre where the majority of fluorescent emission is produced). Therefore comparison of $a_{ph}(\lambda)$ and fluorescence-excitation can provide insight into the functional roles of the intra-cellular pigment complements responsible for bio-optical properties. Hence, future research missions could be designed to collect a wider repertoire of bio-optical measurements. The collection of *in-vivo* absorption, backscattering and fluorescence measurements in combination could provide improved insight into the functional significance of variation in $a_{ph}(\lambda)$, and improve the accuracy of radiative transfer simulations aimed at modelling anticipated R_{rs} .

The custom bio-optical algorithm developed in this thesis, and an off-the-shelf [*chl-a*] algorithm, were employed to investigate secular change in mid-summer Barents Sea phytoplankton community composition and autumnal bloom phenology, respectively.

Overall the research herein demonstrates that existing records of remotely-sensed ocean-colour can be investigated, in conjunction with *in-situ* sampling campaigns, to reveal patterns of secular change in ecosystem function. This represents a valuable methodological approach to investigate rapidly-changing ecosystems, such as the Arctic shelf seas. It was possible to use ocean-colour records, in conjunction with isolated series of *in-situ* records (e.g. Lalande et al. (2013) and Nöthig et al. (2015)), to establish evidence of an increasing contribution from *Phaeocystis* to Barents Sea primary production. The study of autumnal bloom intensity demonstrates that increasing [*chl-a*] concentrations in the Barents Sea in September have a much stronger relationship to Atlantic inflow currents than to wind events, which trigger autumnal blooms over much of the interior Arctic Ocean (Ardyna

et al. 2014).

On reflection, the advent of new ocean-colour remote-sensing platforms, such as the PACE mission (Plankton, Aerosol, Cloud, Ecosystem) (Petro et al. 2020), may expand the capacity to apply bio-optical algorithms to study phytoplankton community composition at synoptic scales in future. Although caution must be expressed, as high levels of autocorrelation mean that ocean-colour wavebands are not statistically independent of one another. This limits the amount of information that can be extracted from them (Cael et al. 2020). Future researchers may therefore need to be prosaic not only in the selection of bio-optically distinct Functional Types as targets for remote-sensing approaches, but also the overall number of classes they hope to retrieve. Should authors attempt to solve for too many pieces of information, they may discover that there are plural equally favoured solutions.

6.2 Future work directions

6.2.1 Harmful Algal Bloom detection in Norwegian waters

Motivation

I am prompted to consider this future avenue of research, because assessment of the bio-optical algorithm developed in chapter 3 (see Figure 6.1) lead to the recognition of an unexpected spectral signature, inconsistent with any functional assemblage I had been seeking to identify. I noted that accumulations of chlorophyll-c3 sometimes occurred in waters with low perceived $[chl-a]$, such as off the coast of Troms county in June 2019 (see Figure 6.2). This satellite image coincided with a suspected bloom of *Chrysochromulina leadbeateri* that devastated farmed salmon stocks around Troms (Hansen et al. 2019). It was not possible, within the thematic scope of my thesis, to confirm whether this bio-optical signature was

consistent with *Chrysochromulina*, as there were no available *in-situ* samples for validation. Nevertheless, it is possible that these observations show that *Chrysochromulina* may occur in sufficiently dense concentrations and possess a sufficiently recognisable bio-optical signature to be the target of future remote-sensing based research.

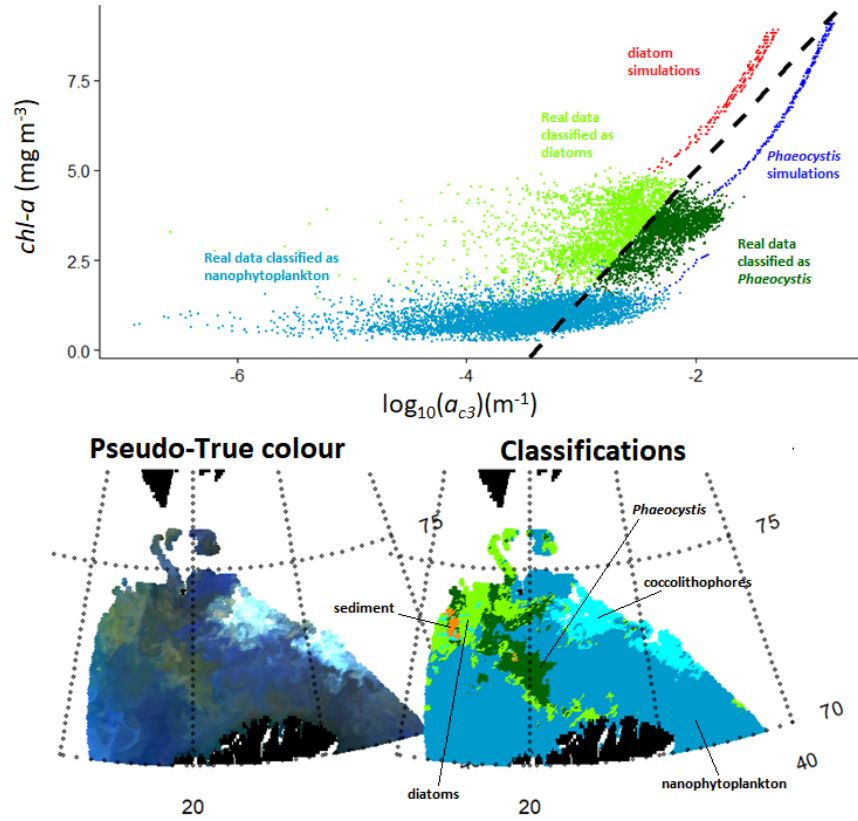


Figure 6.1: Remotely-sensed ocean-colour from the Southern Barents Sea on Julian day 191, 2014. **Top:** plot of chlorophyll-a and absorption signal attributable to chlorophyll-c3 inferred from ocean-colour. Red: trend based on HydroLight simulations of ocean-colour under diatom-bloom scenario (Mobley and Sundman 2012). Blue: trend under simulated colonial *Phaeocystis* scenario. Light green: suspected diatom-dominated community in real image. Dark green: suspected *Phaeocystis*. Teal: suspected unicellular nanophytoplankton. Note that some nanophytoplankton have relatively high chlorophyll-c3 signals. **Bottom:** pseudo-true colour image (**left**) and inferred phytoplankton functional types (**right**), with coccolithophores masked in cyan and waters possibly influenced by sediment masked in orange.

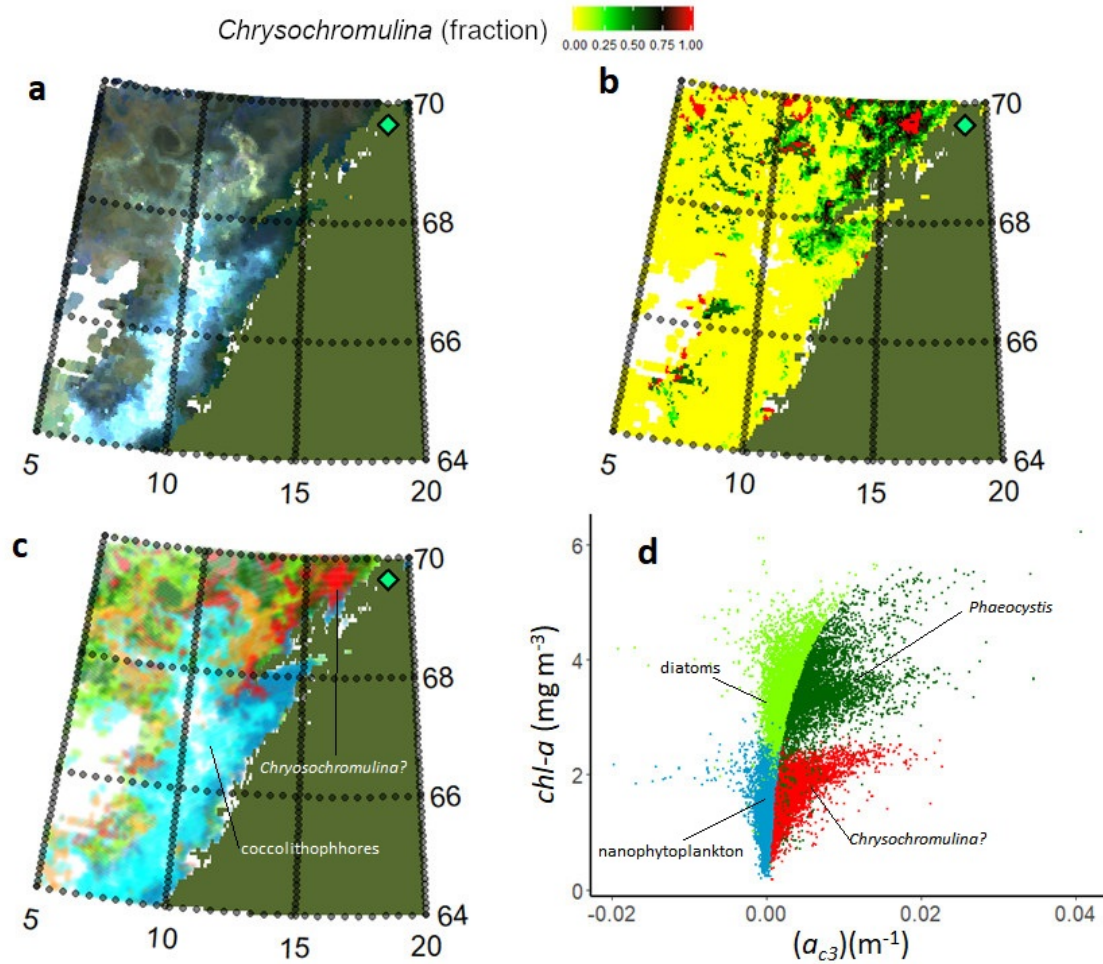


Figure 6.2: (a) Pseudo true-colour image of the Norwegian coast, averaged between Julian day 156 and 167, Tromsø is marked by a green diamond. (b) Map of the relative frequency of non-masked pixels that were putatively classified as *Chrysochromulina*. (c) Map of the different prevailing classifications of sea-water in the region (blue; normal sea-water, cyan; coccolithophores, orange; sediment-input, light-green; diatoms, dark-green; *Phaeocystis*, red; *Chrysochromulina*). (d) Classification scheme for non-coccolithophore or sediment-masked pixels- colours identical to (c). x-axis represents retrieved absorption signal attributable to chlorophyll-c3, as defined in Astoreca et al. (2008).

Chrysochromulina

Chrysochromulina leadbeateri is a nanoflagellate haptophyte, that has been associated with several mass mortalities of fish stocks in Scandinavian waters since the late 1980s (Simonsen and Moestrup 1997). It has been suggested by some authors that chemical compounds liberated by decaying fish may complex with haemolytic ichthyo-toxins produced by *Chrysochromulina*. This results in the evolution of hyper-toxic chemicals implicated in mass mortality. The growth of *Chrysochromulina* may even directly benefit from fish mortality in oligotrophic settings, as the plankton may be able to metabolise rotting fish and increase cell division (Johnsen et al. 1999). Indeed, release of ichthyo-toxins in the genus *Chrysochromulina* has been associated with oligotrophy, in particular low concentrations of phosphate (Johansson and Granéli 1999) – a bio-mineral that is found in abundance in vertebrate skeletons. This raises the possibility that *Chrysochromulina leadbeateri* behaves as a predator of fish. Other harmful algae in the Nordic Seas include species of *Pseudo-nitzschia*, which typically occur in oligotrophic contexts in summer (Wassmann et al. 2006a) and may form complex commensalisms with the colonial haptophyte genus *Phaeocystis* (Sazhin et al. 2007)

The study of Harmful Algal Blooms is increasingly relevant in extensively fished European waters, given their rising incidence in the Global Ocean (Gobler 2020). *Phaeocystis*, and *Emiliania huxleyi* can also themselves form harmful blooms, as a result of excessive DMS production; some dinoflagellates may also produce harmful blooms in Norwegian waters. This variety of species may all possess the bio-marker pigment chlorophyll-c3 (Johnsen and Sakshaug 2000). It is possible that their more frequent occurrence might be related to anthropogenic stresses acting on ecosystems.

Future research

Future research may develop bio-optical algorithms to recognise *Chrysochromulina* (or signals of chlorophyll-c3 absorption indicative of other potentially harmful phytoplankton), and apply them to study the occurrence and distribution of Harmful Algal Blooms on the Norwegian shelf.

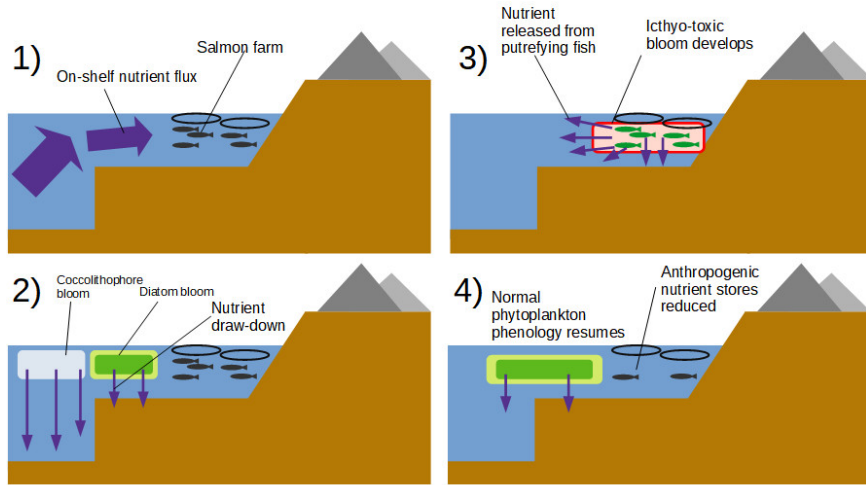
In previous work, authors have exploited absorption of chlorophyll-c3 at 469 nm to facilitate detection of this bio-marker (e.g. Astoreca et al. (2008)), but there are numerous other pigment species that absorb in this region of the spectrum, namely the photoprotective carotenoids (PPC). Work by Johnsen et al. (1994) and Johnsen and Sakshaug (2007) demonstrated that absorption of chlorophyll-c3 at 586 nm was unaffected by PPC and potentially played a role transferring energy to photosystems, perhaps facilitating the efficient capture of fluoresced light initially absorbed in the blue-green region, and resulting in the ultra-efficient capture of light energy (see discussion of photophysiology of *Chrysochromulina polylepis* in Johnsen et al. (1992)). Contemporary remote-sensing platforms lack the requisite spectral channels to monitor absorption at 586 nm, but future hyperspectral missions (e.g. Petro et al. (2020)) will make it possible to investigate this less ambiguous spectral region for signals of the bio-marker chlorophyll-c3.

Numerous potential research questions concerning the ecology and occurrence of Harmful Algal Blooms would benefit from synoptical appraisals of their occurrence; relationships to ocean-shelf nutrient transport, zooplankton, fish swarms and aquaculture could be explored.

In particular, extensive farmed salmon aquaculture is practised on the North-west coast of Europe, from the Northern United Kingdom to inside the Arctic circle in Norway. It is already understood that accumulations of farmed salmon can have deleterious influences on wild salmon and on grander ecosystem function by, for example, their contribution to the dispersal of parasite populations (Middlemas et al. 2013). We might hypothesise that an

increasing incidence of Harmful Algal Blooms in the North-east Atlantic is a manifestation of Le Chatelier's law of equilibrium. This law states that a system, whose equilibrium is disturbed, will approach a new equilibrium by processes that act to counter the initial disturbance. In this situation, declining nutrient contents in Atlantic inflows (Hátún et al. 2017) and rising populations of commercially farmed fish, represent a disturbance to ecosystem equilibrium, which is restored by agents in the ecosystem engendering mass mortality of the farmed fish – liberating their nutrients (see Figure 6.3). A geographically-resolved study, of time series of salmon aquaculture intensity and Harmful Algal Blooms in Norwegian fjords, could test this hypothesis.

a Anthropogenically perturbed system



b Unperturbed system

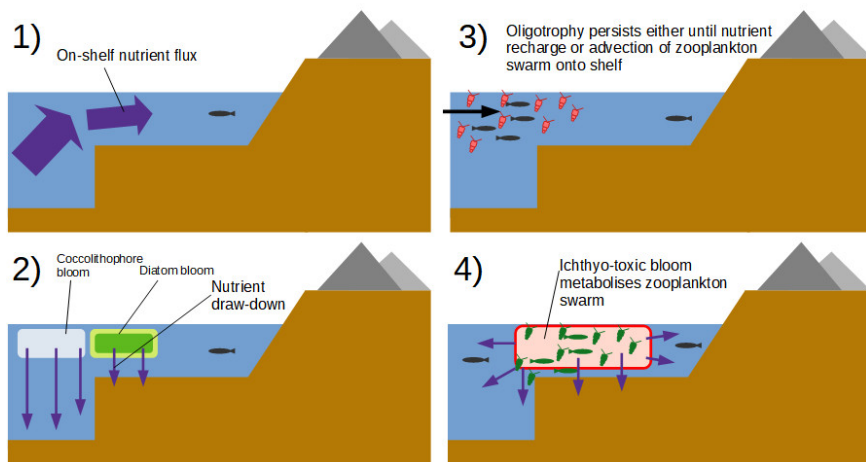


Figure 6.3: a) Hypothesised role of farmed salmon stocks in restoring nutrient imbalance through mass mortality in ichthyotoxic blooms. b) Under a scenario where fish populations are nomadic, and can actively avoid harmful algal blooms, mass mortality is more limited and occurs in contexts such as the advection of zooplankton swarms into oligotrophic enclosed fjords.

6.2.2 Community composition of autumnal blooms

Motivation

Phytoplankton sampling campaigns in the Barents Sea have usually been concentrated in the spring and summer, when climatological conditions are most amenable to research efforts and phytoplankton productivity most intense. Consequently, autumnal sampling has been eschewed. Indeed, Arctic **PRIZE** sampled extensively in the productive season, and all the *in-situ* phytoplankton samples presented in this thesis come from the spring and summer. However, accumulations of autumnal phytoplankton have been sporadically recorded in the Barents Sea for at least 4 decades; Heimdal (1983) observed diatom-dominated ice-edge communities North-west of Svalbard in September-October, 1979. *Chaetoceros*, *Leptocylindricus*, *Nitzschia* and *Thalassiosira* species predominated in biomass concentrations $>2\text{mg m}^{-3}$ [*chl-a*]. In addition, Hegseth (1997) observed autumnal ice-edge blooms on the shelf. Furthermore, recent studies, integrating *in-situ* and remote-sensing approaches, have shown that weak autumnal blooms occur in early-mid September. These blooms, with very low biomass concentrations of typically only $>1\text{mg m}^{-3}$, progress from North to South, where their concentrations and durations are largest (Dalpadado et al. 2020).

However, over this period, the evolution of climatological conditions means that autumnal hydrography has changed considerably. Hegseth (1997) encountered multi-year sea-ice 2-3 m thick, now a rarity. Today, typical ice-thickness – if ice is present at all – would be $<1\text{ m}$ (Kwok 2018). Historic autumnal ice-edge blooms probably represented a fading continuity of summer primary productivity, as diatom-rich sympagic communities continue to pursue the ice-edge until they become light-limited. This style of production cannot be reproduced in modern autumnal blooms occurring far from the ice-edge.

Coupled with declining sea-ice extents and thickness, a secular increase in autumnal bloom [*chl-a*], identified in chapter 5, establishes a mandate to further investigate autumnal styles of production and their contribution to export. The occurrence of these blooms in

perennially ice-free regions indicates they are probably underpinned by a distinct regulatory mechanism from historic sympagic blooms. They may therefore be dominated by a distinct phytoplankton community, and their significance for the function of the wider ecosystem is therefore unknown. However, it is known that these September blooms can only possess minor contributions from calcifying phytoplankton, as remotely-sensed ocean colour indicates negligible Particulate Inorganic Carbon (Oziel et al. 2017). This finding is reinforced by analyses in chapter 5, because autumnal blooms were detected despite coccolithophores being masked in the remotely-sensed ocean-colour time series (see 5.6).

Future research

Future research could address this knowledge gap in the literature by undertaking a comprehensive phytoplankton sampling programme in September conditions. Research findings, in chapter 5, indicate that autumnal bloom intensity is now underpinned by the delivery of Atlantic water-masses onto the shelf, and that [*chl-a*] is most intense in the South-west Barents Sea – an observation confirmed by Dalpadado et al. (2020). Future sampling campaigns should therefore target Bjørnøyareenna, the Atlantic inflow trough between Bjørnøya and Mainland Norway (see Figure 6.4).

Autumnal sampling for zooplankton populations is already undertaken regularly and comprehensively across the Barents Sea (see (Eriksen et al. 2017; Dalpadado et al. 2020)) therefore a future sampling mission could be designed to exploit these ongoing sampling efforts. Samples should be collected for enumeration with optical microscopy and analysis for taxon-indicative pigments with HPLC, to constrain community composition. Bio-optical measurements might, furthermore, reveal whether autumnal assemblages are comparable to the bio-optical types identified in chapters 2 and 3. The deployment of sediment traps could place constraints on autumnal pelagic-benthic coupling. Previous autumnal deployment of sediment traps has largely been limited to the Fram Strait or fjordic environments in the

Svalbard archipelago, far from the sites of interest (e.g. Hebbeln and Wefer (1991) and Weydmann-Zwolicka et al. (2021)). Zooplankton surveys in spring might also be able to explore whether autumnal bloom occurrence substantially affects the survival of zooplankton populations that enter diapause over winter.

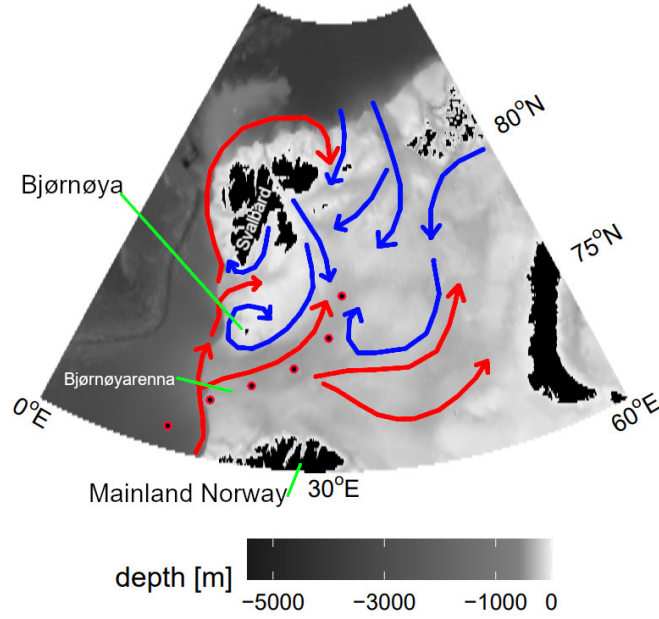


Figure 6.4: Bathymetric map of the Barents Sea, produced with the R *marmap* package (Pante and Simon-Bouhet 2013). Currents delivering Atlantic and Arctic Water are portrayed in red and blue, respectively. Suggested locations for an autumnal fieldwork campaign are illustrated as pink-bordered dots, following the Atlantic inflow onto the shelf.

Inter-annual variability in autumnal bloom intensity (see Figure 5.7) might be accompanied by variability in community composition or pelagic-benthic coupling. Therefore, a sampling campaign distributed over multiple years would be ideal. In order to estimate the influence of a greater number of yearly observations, I took the time series of variance-normalised $[chl-a]$ estimates in chapter 5 (see Figure 5.7) and divided into rolling time series of increasing length, from 2 to 17 years. Rolling subsets were chosen, because this preserves

temporal auto-correlation, which might influence variance structure. The standard deviations of the subsets were computed for each temporal length (2 years, 3 years, 4 years...) and the mean of these values was divided by the standard deviation of the whole time series. The results are presented in Figure 6.5.

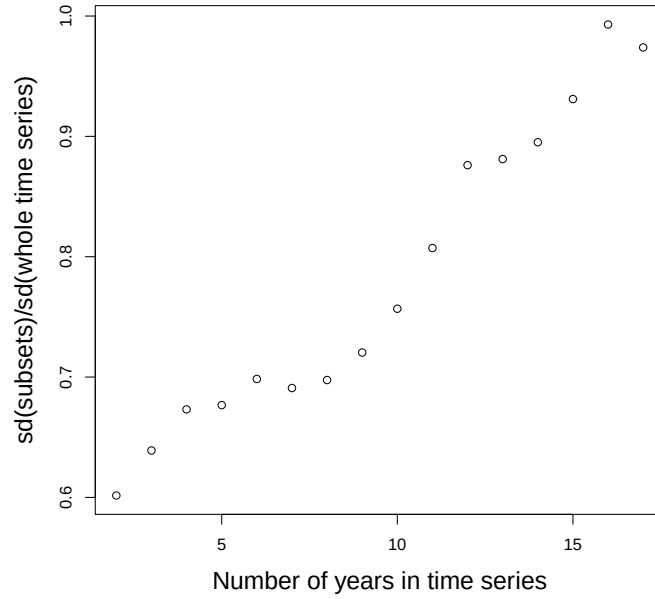


Figure 6.5: Evolution of standard deviation of $[chl-a]$ time series with greater length.

It is evident, from Figure 6.5, that there is no clear cut-off where it can be said that inter-annual variability is adequately represented. Indeed, an 18 year time series may itself not be long enough to fully represent inter-annual variability in autumnal $[chl-a]$. These time-scales are much greater than that any individual sampling campaign strategy is usually designed to meet. Therefore, the number of repeated autumnal deployments in an autumnal sampling scheme is a subjective decision. Autumnal sampling must therefore be conducted in the knowledge that variability in hydrography, climatological conditions, and the flux of zooplankton populations onto the shelf, is high (e.g. Eriksen et al. (2017)) – and that

this variability may both determine how representative datasets are as well as potentially enhancing or disrupting causative relationships between primary production, sedimentation and zooplankton populations.

6.2.3 The ascendancy of machine learning and the integration of measurement platforms

Machine learning

This thesis has applied a range of widely-employed matrix-based multivariate statistical methods. These methods have been used to engage biological questions by assessing variance structure, performing clustering and finding salient differences between pre-defined classes. The majority of these methods, such as Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA), were developed by pioneering mathematicians, statisticians and biologists of the early twentieth century (see Legendre and Legendre (2012) for a detailed discussion of these numerical methods and their history). Assumptions of linearity in the fundamental relationships underlying data structure are essential to their application. Their implementation and interpretation requires explicit supervision from the researcher. For example, the retrieval of class differences in LDA requires statements a priori of class memberships.

Thus, while these mathematical methods significantly reduce the dimensionality of datasets, therefore allowing researchers to engage multivariate problems and develop algorithms to study data at scale, the requirement for supervision in the construction of algorithms (for example in the identification of relationships underlying Principal Components, or in the assignment of class memberships and validation of Linear Discriminants) requires significant researcher attention. This represents a significant obstacle to the interrogation of large datasets. A time series of remotely-sensed ocean-colour observations could easily contain millions of measurements for example and most researchers' analysis of data

structure must therefore focus on carefully chosen subsets that are reasoned to be representative, and the researcher may be required to make informed transforms of the original dataset to ensure that salient relationships are linearised. Take, for example, the development of the PHYSAT algorithm (Alvain et al. 2005), which required a transformation of R_{rs} data to remove variation in signal attributable to $[chl-a]$, that would have otherwise represented a significant non-linear source of variance.

Mathematical methods for ‘less’-supervised exploratory analysis and classification of salient data structure began development in the mid-to-late twentieth century. This plethora of approaches is known as ‘machine learning’. Artificial neural networks are a key example of machine learning techniques, which function by imitating the architecture of brain tissue. The network is organised such that each node takes multiple inputs, weighted by a matrix, and computes an output by application of some non-linear function (Bishop et al. 1995). The network may be arranged in multiple layers, and a signal originating as input may be allowed to pass between layers multiple times before finally leaving the network as ‘output’. Layers that do not connect directly to researcher input or final output can be described as ‘hidden’. The topological architecture of neural networks may vary depending on the task they are anticipated to perform (e.g. Hagen et al. (1996)) and their design is itself an active field of research. By exploiting different architectures, researchers can ‘train’ artificial neural networks to perform tasks such as the recognition of classes, by providing pre-classified training data. Dynamic variation of the matrices of weights inside the nodes is then applied to change the network behaviour over many iterations, until the network is optimised for the production of the ‘correct’ output for a training dataset. The complex nature of the pathways individual data may follow, potentially passing through many nodes inside hidden layers, means that direct interpretation of the operations performed on data by a neural network is often not tractable to researchers. Consequently, artificial neural networks differ from the mathematical methods applied in this thesis in being ‘black-box’ approaches.

While the artificial neural network requires initial supervision in its construction, its learning process proceeds in an autonomous fashion. Furthermore, simultaneous cross-validation schemes with both a training and ‘test’ datasets can be employed to autonomously improve algorithm robustness and reduce the risk of fitting functions to the training dataset that are not representative of data structure found more widely in nature (Bricaud et al. 2007). While this example application required researcher supervision, artificial neural networks may also undertake unsupervised learning, allowing the expansion of learning to include fully autonomous assessment of structure in data for which researchers have provided no a priori classification.

These approaches have been used by biological oceanographers to produce predictive models that can distinguish contributions of algal pigments and size-classes to $a_{ph}(\lambda)$ (Bricaud et al. 2007), and to reconstruct benthic secondary production in marine ecosystems from commonly-available suites of measurements (Brey 2012). Indeed, the latter example was directly employed by Degen et al. (2016) in their assessment of environmental and biological correlates of megabenthic secondary production in the Barents Sea. The ‘black-box’ limitation of the opacity of artificial neural networks to researchers is illustrated by the analytical choices made by Brey (2012) to assess the leading variables that explain variance in inferred secondary production. A standard multiple linear regression was used to identify the salient variables (body mass of the benthic fauna and water temperature), and the significance of the remaining variables was then established by constructing new artificial neural networks devoid of the salient variables. It can therefore be seen that the performance of the artificial neural network is being assessed empirically, and that mechanistic understanding required the application of conventional linear statistical methods.

Artificial neural networks have been discussed here as an illustrative example, but a wider repertoire of machine-learning approaches is being applied in biological oceanography and remote-sensing, such as in Lewis and Arrigo (2020), in which a genetic algorithm (the

algorithm is ‘mutated’ between iterations, new variants are compared by a cost function and the fittest variants establish the initial characteristics for the next iteration) is applied to discover the form of polynomial equations that could relate remotely-sensed ocean-colour with *in-situ* measurements of [*chl-a*], CDOM and particulate backscattering. The convergence of genetic algorithms upon the relatively fittest solution within a given landscape of parameter variation is a process that is at least in part stochastic, and therefore not immediately tractable to the researcher.

Other machine learning approaches for the assessment of remote-sensing data have included random forests classifiers (a layered set of decisions, known as a ‘tree’ can be used to classify input. In a random forest approach, many trees are produced and a consensus is taken by finding the modal or mean output, depending on the task) seeking to identify potential archaeological sites in the Shetland isles from a combination of spectral channels and topographic information, (although, in this case, LDA was found to outperform random forests) (Megarry et al. 2016).

Black-box machine learning approaches are likely to be employed increasingly in biological oceanographic studies as computational power continues to become cheaper, further open-source software tools are developed and a bigger number of increasingly multivariate large datasets become available (e.g. (Petro et al. 2020)). The utility that these approaches provide must be carefully considered against the risk that researchers develop proxies for ecosystem state that they do not fully understand. This is a prominent risk in remote-sensing applications, for which sensor degradation or processing can significantly perturb the retrieval of derived variables (Gregg et al. 2017). It is possible to imagine, for example, that a researcher should train a machine learning algorithm to infer [*chl-a*] on a historic dataset, and that changes in the underlying relationships between ocean optical properties and phytoplankton biomass (Dutkiewicz et al. 2019) may evolve with time, that sensor performance may degrade or that processing conventions might change. This will result in the

temporal evolution of increasing error in $[chl-a]$ with time, and the researcher may not possess a clear understanding of the mechanistic cause of this error if the algorithm they have contrived is not readily interpretable. There is a case for favouring simple, tractable and interrogable algorithms that can be well understood by researchers. Work by authors such as Lewis and Arrigo (2020) may represent a blue-print for the future application of machine learning in biological oceanography; while the genetic algorithm used to optimise retrieval of $[chl-a]$ from ocean-colour may be intractable, the empirical algorithms it developed were composed from a family of polynomial functions. Therefore, any change in function of the empirical algorithm will be interpretable.

Integration

The work conducted under this thesis required the explicit integration of *in-situ* measurements and remotely-sensed ocean-colour. Arctic **PRIZE** has, furthermore, already produced publications that integrate data streams from remote-sensing platforms with measurements from semi-autonomous submersible ‘gliders’ (Kostakis et al. 2020; Porter et al. 2020). This practice reflects the evolution of broad multi-disciplinary research in marine ecosystems, that has been essential to campaigns such as the Pro Mare (Sakshaug et al. 1991), N-ICE, MOSAIC (Shupe et al. 2018) and Nansen Legacy projects.

Miniaturisation of electronics has, furthermore, contributed to the ongoing development of smaller and cheaper semi-autonomous and autonomous robotic systems that can be deployed at sea. Future researchers may be tasked, therefore, with the integration of diverse data-streams from moorings, micro-satellites, ship-borne sensors, aerial drones and profiling floats (Figure 6.6). The potential benefits of such ‘heterogeneous’ data collection are being invited as a ‘paradigm shift’ (e.g. Berge et al. (2020) pages 241-275).

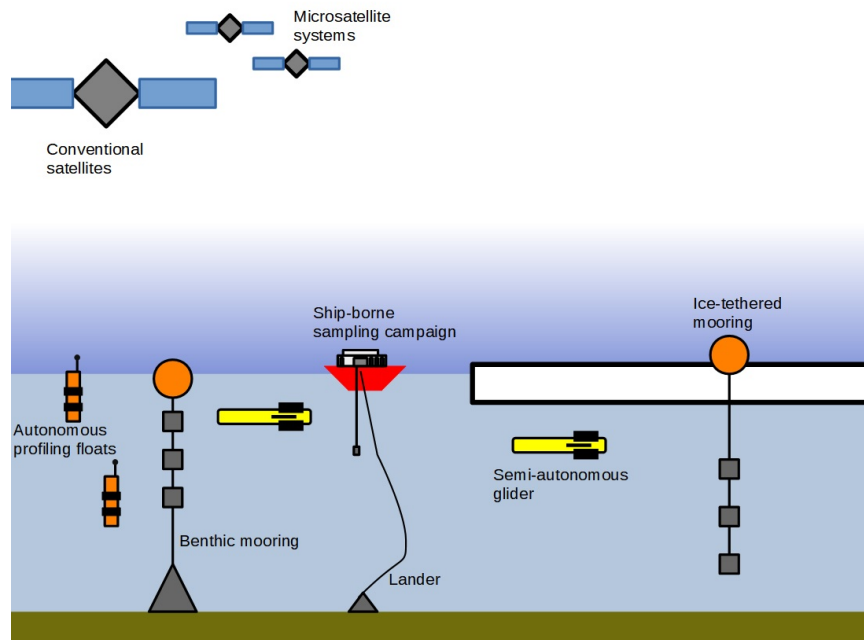


Figure 6.6: A diverse range of instruments; space-borne, ship-borne, benthic, ice-tethered or submerged, are increasingly forming the architecture for data collection in a changing Arctic. Image adapted from Figure 9.1 in Berge et al. (2020).

The management of large and diverse data streams is a task of such complexity that it may well demand the application of sophisticated machine learning approaches so that researchers can concentrate their efforts on the interpretation of datasets and not only their integration. Long-term sampling platforms, such as moorings, satellites and gliders, offer the potential to provide useful information to direct the sampling of shorter-lived missions such as ship-board sampling campaigns to procure *in-situ* samples. Semi-autonomous or remotely-operated sensor platforms, furthermore, may help obviate seasonal biases in data-collection towards the warmer and more quiescent periods of the year, because they do not require the presence of the researcher in the field.

References

- Alvain, Séverine, Cyril Moulin, Yves Dandonneau, and François-Marie Bréon. 2005. "Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery." *Deep Sea Research Part I: Oceanographic Research Papers* 52 (11): 1989–2004.
- Ardyna, Mathieu, Marcel Babin, Michel Gosselin, Emmanuel Devred, Luc Rainville, and Jean-Éric Tremblay. 2014. "Recent Arctic Ocean sea ice loss triggers novel fall phytoplankton blooms." *Geophysical Research Letters* 41 (17): 6207–6212.
- Assmy, Philipp, Mar Fernández-Méndez, Pedro Duarte, Amelie Meyer, Achim Randelhoff, Christopher J Mundy, Lasse M Olsen, Hanna M Kauko, Allison Bailey, Melissa Chierici, et al. 2017. "Leads in Arctic pack ice enable early phytoplankton blooms below snow-covered sea ice." *Scientific reports* 7 (1): 1–9.
- Astoreca, Rosa, Véronique Rousseau, Kevin Ruddick, Cécile Knechciak, Barbara Van Mol, Jean-Yves Parent, and Christiane Lancelot. 2008. "Development and application of an algorithm for detecting *Phaeocystis globosa* blooms in the Case 2 Southern North Sea waters." *Journal of Plankton Research* 31 (3): 287–300.
- Barros, MP, M Pedersén, P Colepiccolo, and P Snoeijs. 2003. "Self-shading protects phytoplankton communities against H₂O₂-induced oxidative damage." *Aquatic microbial ecology* 30 (3): 275–282.
- Berge, Jørgen, Geir Johnsen, and Jonathan H Cohen. 2020. *POLAR NIGHT marine ecology: life and light in the dead of night*. Vol. 4. Springer Nature.
- Bishop, CM, et al. 1995. *Neural networks for pattern recognition*. Clarendon.
- Brey, Thomas. 2012. "A multi-parameter artificial neural network model to estimate macrobenthic invertebrate productivity and production." *Limnology and Oceanography: Methods* 10 (8): 581–589.
- Bricaud, Annick, Carlos Mejia, David Blondeau-Patissier, Hervé Claustre, Michel Crepon, and Sylvie Thiria. 2007. "Retrieval of pigment concentrations and size structure of algal populations from their absorption spectra using multilayered perceptrons." *Applied Optics* 46 (8): 1251–1260.
- Brunet, Christophe, Geir Johnsen, Johann Lavaud, and Suzanne Roy. 2011. "Pigments and photoacclimation processes." *Phytoplankton pigments: characterization, chemotaxonomy and applications in oceanography*.
- Cael, B.B., Alison Chase, and Emmanuel Boss. 2020. "Information content of absorption spectra and implications for ocean color inversion." *Applied optics* 59 (13): 3971–3984.
- Dalpadado, Padmini, Kevin R Arrigo, Gert L van Dijken, Hein Rune Skjoldal, Espen Bågøien, Andrey V Dolgov, Irina P Prokopchuk, and Erik Sperfeld. 2020. "Climate effects on temporal and spatial dynamics of phytoplankton and zooplankton in the Barents Sea." *Progress in Oceanography* 185:102320.
- Degen, Renate, Lis Lindal Jørgensen, Pavel Ljubin, Ingrid H Ellingsen, Hendrik Pehlke, and Thomas Brey. 2016. "Patterns and drivers of megabenthic secondary production on the Barents Sea shelf." *Marine Ecology Progress Series* 546:1–16.
- Degerlund, Maria, and Hans Christian Eilertsen. 2010. "Main species characteristics of phytoplankton spring blooms in NE Atlantic and Arctic waters (68–80 N)." *Estuaries and coasts* 33 (2): 242–269.
- Devred, Emmanuel, Shubha Sathyendranath, Venetia Stuart, and Trevor Platt. 2011. "A three component classification of phytoplankton absorption spectra: Application to ocean-color data." *Remote Sensing of Environment* 115 (9): 2255–2266.
- Dutkiewicz, Stephanie, Anna E Hickman, Oliver Jahn, Stephanie Henson, Claudie Beaulieu, and Erwan Monier. 2019. "Ocean colour signature of climate change." *Nature communications* 10 (1): 578.

- Eriksen, Elena, Hein Rune Skjoldal, Harald Gjøsæter, and Raul Primicerio. 2017. "Spatial and temporal changes in the Barents Sea pelagic compartment during the recent warming." *Progress in Oceanography* 151:206–226.
- Gobler, Christopher J. 2020. "Climate change and harmful algal blooms: insights and perspective." *Harmful Algae* 91:101731.
- Gregg, Watson W, Cécile S Rousseaux, and Bryan A Franz. 2017. "Global trends in ocean phytoplankton: a new assessment using revised ocean colour data." *Remote Sensing Letters* 8 (12): 1102–1111.
- Hagen, MT, HB Demuth, and M Beale. 1996. *Neural Network Design* PWS Publishing Co.
- Hancke, Kasper, Erlend K Hovland, Zsolt Volent, Ragnhild Pettersen, Geir Johnsen, Mark Moline, and Egil Sakshaug. 2014. "Optical properties of CDOM across the Polar Front in the Barents Sea: Origin, distribution and significance." *Journal of Marine Systems* 130:219–227.
- Hansen, Kathryn, Joshua Steven, and Bjarte Stig. 2019. "Beauty and a Burden in the Norwegian Sea NASA Earth Observatory." Accessed: 2019-07-18, <https://earthobservatory.nasa.gov/images/145151/beauty-and-a-burden-in-the-norwegian-sea>.
- Hátún, Hjalmar, K Azetsu-Scott, R Somavilla, Francisco Rey, C Johnson, Moritz Mathis, Uwe Mikolajewicz, P Coupel, J-É Tremblay, S Hartman, et al. 2017. "The subpolar gyre regulates silicate concentrations in the North Atlantic." *Scientific reports* 7 (1): 1–9.
- Hebbeln, Dierk, and Gerold Wefer. 1991. "Effects of ice coverage and ice-rafted material on sedimentation in the Fram Strait." *Nature* 350 (6317): 409–411.
- Hegseth, Else Nøst. 1997. "Phytoplankton of the Barents Sea- the end of a growth season." *Polar Biology* 17 (3): 235–241.
- Heimdal, Berit Riddervold. 1983. "Phytoplankton and nutrients in the waters north-west of Spitsbergen in the autumn of 1979." *Journal of plankton research* 5 (6): 901–918.
- Johansson, N, and Edna Granéli. 1999. "Cell density, chemical composition and toxicity of *Chrysochromulina polylepis* (Haptophyta) in relation to different N: P supply ratios." *Marine Biology* 135 (2): 209–217.
- Johnsen, G, and E Sakshaug. 2000. "Monitoring of harmful algal blooms along the Norwegian coast using bio-optical methods." *South African Journal of Marine Science* 22 (1): 309–321.
- Johnsen, Geir, Runar Dalløkken, Wenche Eikrem, Catherine Legrand, Jan Aure, and Hein Rune Skjoldal. 1999. "Eco-physiology, bio-optics and toxicity of the ichthyotoxic *Chrysochromulina leadbeateri* (Prymnesiophyceae)." *Journal of Phycology* 35 (6): 1465–1476.
- Johnsen, Geir, and Egil Sakshaug. 2007. "Biooptical characteristics of PSII and PSI in 33 species (13 pigment groups) of marine phytoplankton, and the relevance for pulse-amplitude-modulated and fast-repetition-rate fluorometry 1." *Journal of Phycology* 43 (6): 1236–1251.
- Johnsen, Geir, Egil Sakshaug, and Maria Vernet. 1992. "Pigment composition, spectral characterization and photosynthetic parameters in *Chrysochromulina polylepis*." *Marine Ecology Progress Series*, 241–249.
- Johnsen, Geir, Oddgeir Samset, Lois Granskog, and Egil Sakshaug. 1994. "In vivo absorption characteristics in 10 classes of bloom-forming phytoplankton: taxonomic characteristics and responses to photoadaptation by means of discriminant and HPLC analysis." *Marine Ecology Progress Series*, 149–157.
- Kauko, Hanna M, Alexey K Pavlov, Geir Johnsen, Mats A Granskog, Ilka Peeken, and Philipp Assmy. 2019. "Photoacclimation state of an Arctic underice phytoplankton bloom." *Journal of Geophysical Research: Oceans* 124 (3): 1750–1762.
- Kostakis, Ina, Rüdiger Röttgers, Andrew Orkney, Heather A Bouman, Marie Porter, Finlo Cottier, Jørgen Berge, and David McKee. 2020. "Development of a bio-optical model for

- the Barents Sea to quantitatively link glider and satellite observations.” *Philosophical Transactions of the Royal Society A* 378 (2181): 20190367.
- Kwok, Ron. 2018. “Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018).” *Environmental Research Letters* 13 (10): 105005.
- Lalande, Catherine, Eduard Bauerfeind, Eva-Maria Nöthig, and Agnieszka Beszczynska-Möller. 2013. “Impact of a warm anomaly on export fluxes of biogenic matter in the eastern Fram Strait.” *Progress in Oceanography* 109:70–77.
- Legendre, Pierre, and Loic FJ Legendre. 2012. *Numerical ecology*. Elsevier.
- Lewis, KM, AE Arntsen, P Coupel, H Joy-Warren, KE Lowry, A Matsuoka, MM Mills, GL Van Dijken, V Selz, and KR Arrigo. 2019. “Photoacclimation of Arctic Ocean phytoplankton to shifting light and nutrient limitation.” *Limnology and Oceanography* 64 (1): 284–301.
- Lewis, KM, and KR Arrigo. 2020. “Ocean Color Algorithms for estimating chlorophyll a, CDOM absorption, and particle backscattering in the Arctic Ocean.” *Journal of Geophysical Research: Oceans* 125 (6): e2019JC015706.
- Liu, Xiaohan, Emmanuel Devred, and Catherine Johnson. 2018. “Remote sensing of phytoplankton size class in Northwest Atlantic from 1998 to 2016: bio-optical algorithms comparison and application.” *Remote Sensing* 10 (7): 1028.
- Liu, Yangyang, Emmanuel Boss, Alison Chase, Hongyan Xi, Xiaodong Zhang, Rüdiger Röttgers, Yanqun Pan, and Astrid Bracher. 2019. “Retrieval of phytoplankton pigments from underway spectrophotometry in the Fram Strait.” *Remote Sensing* 11 (3): 318.
- Megarry, William P, Gabriel Cooney, Douglas C Comer, and Carey E Priebe. 2016. “Posterior probability modeling and image classification for archaeological site prospection: Building a survey efficacy model for identifying Neolithic felsite workshops in the Shetland Islands.” *Remote Sensing* 8 (6): 529.
- Middlemas, SJ, RJ Fryer, D Tulett, and JD Armstrong. 2013. “Relationship between sea lice levels on sea trout and fish farm activity in western Scotland.” *Fisheries Management and Ecology* 20 (1): 68–74.
- Mobley, CD, and LK Sundman. 2012. “HydroLight 5.1 EcoLight 5.1 Users’ Guide, Sequoia Scientific.” *Inc., Bellevue*.
- Nöthig, Eva-Maria, Astrid Bracher, Anja Engel, Katja Metfies, Barbara Niehoff, Ilka Peeken, Eduard Bauerfeind, Alexandra Cherkasheva, Steffi Gäbler-Schwarz, Kristin Hardge, et al. 2015. “Summertime plankton ecology in Fram Strait- a compilation of long and short-term observations.” *Polar Research* 34 (1): 23349.
- Oziel, Laurent, G Neukermans, M Ardyna, C Lancelot, J-L Tison, Paul Wassmann, Jérôme Sirven, Diana Ruiz-Pino, and J-C Gascard. 2017. “Role for Atlantic inflows and sea ice loss on shifting phytoplankton blooms in the Barents Sea.” *Journal of Geophysical Research: Oceans* 122 (6): 5121–5139.
- Pante, Eric, and Benoit Simon-Bouhet. 2013. “marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R.” *PLoS ONE* 8 (9): e73051. doi:10.1371/journal.pone.0073051.
- Pavlov, Alexey K, Torbjørn Taskjelle, Hanna M Kauko, Børge Hamre, Stephen R Hudson, Philipp Assmy, Pedro Duarte, Mar Fernández-Méndez, CJ Mundy, and Mats A Granskog. 2017. “Altered inherent optical properties and estimates of the underwater light field during an Arctic under-ice bloom of *P. haeocystis pouchetii*.” *Journal of Geophysical Research: Oceans* 122 (6): 4939–4961.
- Petro, Susanna, Karen Pham, and George Hilton. 2020. “Plankton, Aerosol, Cloud, Ocean Ecosystem (PACE) Mission Integration and Testing.”
- Porter, Marie, SF Henley, Andrew Orkney, HA Bouman, B Hwang, E Dumont, EJ Venables, and F Cottier. 2020. “A polar surface eddy obscured by thermal stratification.” *Geophysical Research Letters* 47 (6): e2019GL086281.

- Rodriguez, F, M Chauton, G Johnsen, K Andresen, LM Olsen, and M Zapata. 2006. "Photoacclimation in phytoplankton: implications for biomass estimates, pigment functionality and chemotaxonomy." *Marine Biology* 148 (5): 963–971.
- Roy, Suzanne, Carole A Llewellyn, Einar Skarstad Egeland, and Geir Johnsen. 2011. *Phytoplankton pigments: characterization, chemotaxonomy and applications in oceanography*. Cambridge University Press.
- Sakshaug, Egil, Christopher CE Hopkins, and Nils Are Øritsland. 1991. "Proceedings of the Pro Mare Symposium on Polar Marine Ecology, Trondheim, Norway, 12-16 May 1990." *Polar Research* 10 (1): 1–4.
- Sazhin, Andrey F, L Felipe Artigas, Jens C Nejstgaard, and Marc E Frischer. 2007. "The colonization of two *Phaeocystis* species (Prymnesiophyceae) by pennate diatoms and other protists: a significant contribution to colony biomass." In *Phaeocystis, major link in the biogeochemical cycling of climate-relevant elements*, 137–145. Springer.
- Shupe, Matthew, Gijs de Boer, Klaus Dethloff, Elizabeth Hunke, Wieslaw Maslowski, Allison McComiskey, Ola Perrson, David Randall, Michael Tjernstrom, David Turner, et al. 2018. *The Multidisciplinary Drifting Observatory for the Study of Arctic Climate (MOSAIC) Atmosphere Science Plan*. Technical report. DOE Office of Science Atmospheric Radiation Measurement (ARM) Program . . .
- Simonsen, Susanne, and Øjvind Moestrup. 1997. "Toxicity tests in eight species of *Chrysochromulina* (Haptophyta)." *Canadian Journal of Botany* 75 (1): 129–136.
- Sosik, HM, S Sathyendranath, J Uitz, H Bouman, and A Nair. 2014. "In situ methods of measuring phytoplankton functional types." *Phytoplankton Functional Types from Space. Reports of the International Ocean-Colour Coordinating Group (IOCCG)*, no. 15, 21–38.
- Wassmann, Paul, Marit Reigstad, Tore Haug, Bert Rudels, Michael L Carroll, Haakon Hop, Geir Wing Gabrielsen, Stig Falk-Petersen, Stanislav G Denisenko, Elena Arashkevich, et al. 2006a. "Food webs and carbon flux in the Barents Sea." *Progress in Oceanography* 71 (2-4): 232–287.
- Weydmann-Zwolicka, Agata, Paula Pratnicka, Magdalena Lacka, Sanna Majaneva, Finlo Cottier, and Jørgen Berge. 2021. "Zooplankton and sediment fluxes in two contrasting fjords reveal Atlantification of the Arctic." *Science of the Total Environment* 773:145599.

Chapter 7

Appendix: Arctic PRIZE

7.1 Project outline

Arctic **PRIZE** (**P**roductivity in the Seasonal **I**ce **Z**one) is a multi-disciplinary research project designed to investigate the effects of ongoing anthropogenic climate change on the hydrographic set-up and illumination of the Barents Sea, the consequences that these physical changes have for phytoplankton phenology and the cascade of effects at higher trophic levels mediated by food-web interactions. Arctic **PRIZE** therefore seeks to gain a holistic view of the Barents Sea ecosystem's baseline functions – bringing together physicists, chemists, pelagic and benthic ecologists. The sampling missions undertaken as part of **PRIZE** were distributed across multiple years and seasons in order to fully establish the role that variability in climatological and hydrographic conditions plays in structuring ecosystem response, and to ensure that our sampling was representative of the range of conditions encountered in the region.

The University of Oxford contributed to work package 3 of the project, which is focussed on the study of pelagic phytoplankton community production, phenology and composition. I

participated in four research cruises in the Western Barents Sea upon the RRS James Clark Ross of the British Antarctic Survey and the FF Helmer Hanssen of UiT (Norges arktiske universitet) (Hopkins 2018, Cottier and Berge 2018a, Cottier and Berge 2018b, Cottier 2018). Maps of the distribution of hydrographic and phytoplankton sampling completed aboard these cruises are presented in Figure 7.1. The temporal distribution of these cruises is presented in Table 7.1

Table 7.1: *Cruise durations*

Cruise	Vessel	Julian days	Year
JR16006	RRS James Clark Ross	181-220	2017
HH180101	FF Helmer Hanssen	4-18	2018
HH180423	FF Helmer Hanssen	113-125	2018
JR17006	RRS James Clark Ross	162-187	2018

The distribution of the cruises across two years means that a component of inter-annual variability is convolved in the variance of environmental conditions. This is especially evident in the variable ice conditions; while the Southern margin of the sea-ice extended to 78° North on the mid-late summer cruise JR16006 in 2017, the ice-edge had receded to beyond 82° North in 2018 on the early summer cruise JR17006.

Prohibitive ice conditions in winter and spring determined the Northern limit of sampling on cruises HH180101 and HH180423, which means that the collection of samples is biased in favour of summer conditions and the geography of samples evolves as a function of hydrography. Sympagic communities are hence under-represented in contrast to pelagic phytoplankton communities. Sampling sites were concentrated most densely along the 30th Meridian East, extending across the Polar Front into the Seasonal Ice Zone. This transect represents a region of convergence of Atlantic-derived water-masses following the Bjørnøyarennan Trough around the Eastern limit of the Spitsbergenbanken, and Arctic derived Waters entering via the Franz-Victoria Trough East of Kvitøya.

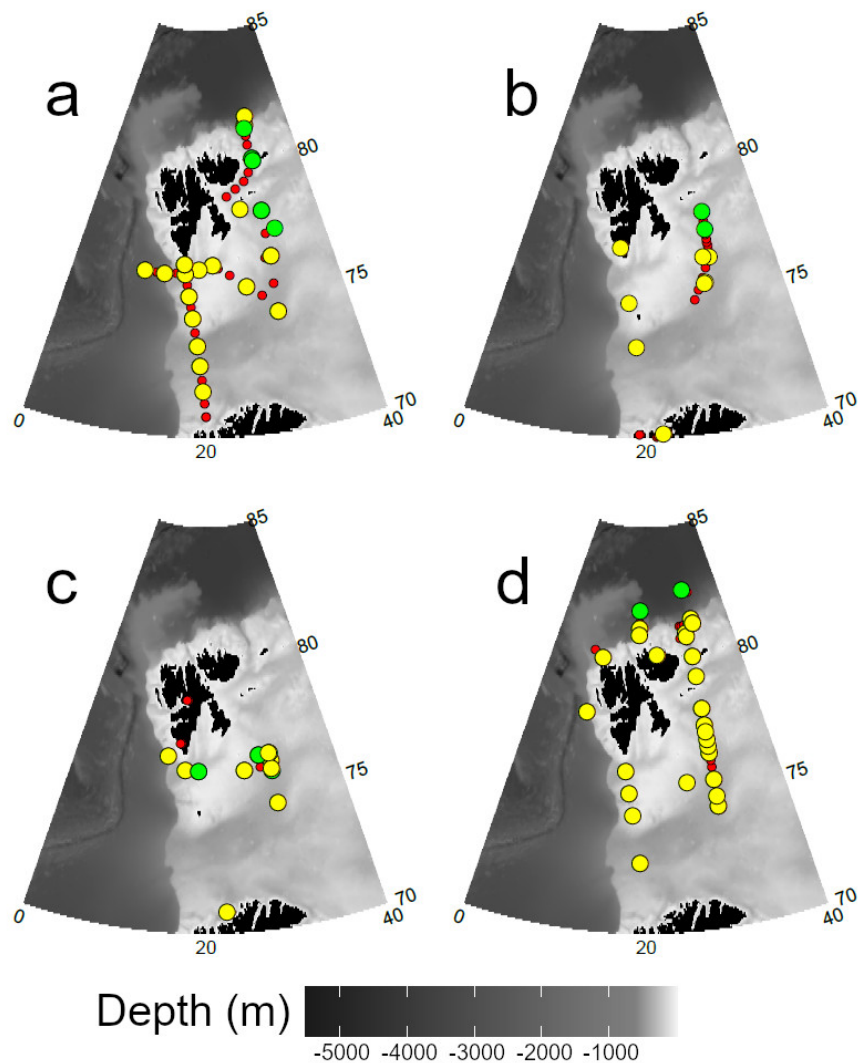


Figure 7.1: Maps of *in-situ* sampling on cruises **a)** JR16006 (July-August, 2017), **b)** HH180101 (January, 2018), **c)** HH180423 (April-May, 2018), **d)** JR17006 (June-July, 2018). Red sites represent locations where hydrography was sampled alone. Phytoplankton were sampled in open-ocean conditions at all yellow sites, whereas green sites represent sampling under ice-covered conditions.

Sampling was undertaken with Conductivity Temperature Depth (CTD) rosettes supporting an array of instruments to measure temperature, pressure, salinity, *in-vivo* chlorophyll-

a concentration (here-on referred to as $[chl-a]$) fluorescence and Photosynthetically Active Radiation (E_{PAR}). CTD devices also carried Niskin bottles for the collection of water samples for nutrient assays and the assessment of phytoplankton communities. In ice-heavy conditions, on cruise HH180423, a bucket was occasionally used to collect water and sea-ice samples when CTD deployment did not prove possible.

Instrument cluster specifications for each cruise are provided in the cruise documentation (Hopkins 2018; Cottier and Berge 2018a, 2018b; Cottier 2018).

As *in-vivo* $[chl-a]$ fluorescence may differ from real $[chl-a]$ ¹, I reviewed real-time readings of $chl-a$ fluorescence to select target depths for sampling with Niskin bottles. I took triplicate subsamples from these Niskin bottles, typically of 0.2 Litre volumes, and filtered through 25 mm GF/F filters (0.7 μ m pore diameter) under a pressure of no more than one half atmosphere. The filters were then submerged in 90% acetone solution and left to extract in the dark in a -20°C freezer for at least 10 hours prior to analysis with a Turner Designs Trilogy fluorometer to determine $chl-a$ concentration using the acidification method (Holm-Hansen et al. 1965). Samples were also filtered for the analysis of bio-optical properties and accessory pigment ratios, and were immediately frozen at -80°C until analysis. Additional seawater samples were preserved in Lugol's solution for phytoplankton species enumeration using optical microscopy. This sum of measurements was designed to provide a comprehensive representation of phytoplankton community structure, and may additionally also provide insights into life history (such as photoacclimation) and potential bio-optical signatures for recognition of functional types from remotely-sensed ocean-colour measurements.

Transects of hydrographic properties and sympathetic variation in phytoplankton biomass and nutrient stocks are presented here, in order to serve as a record of the sampling missions

1. For example, as a result of sunlight-induced fluorescence quenching (Milligan et al. 2012).

and a reference material to clarify key concepts.

7.2 Transects

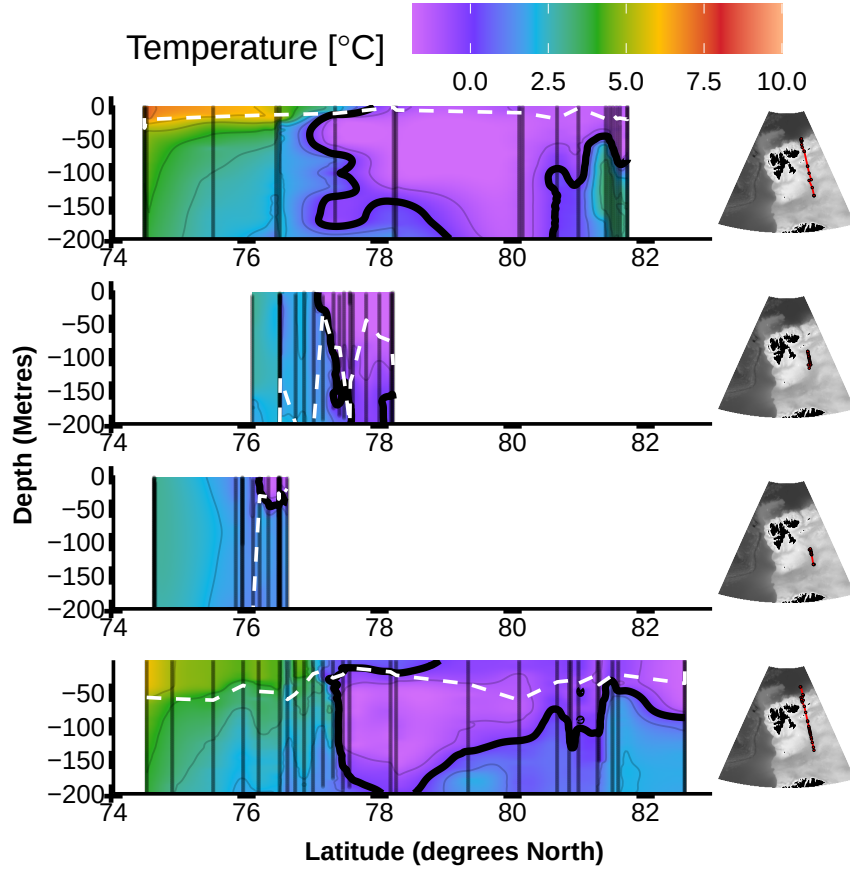


Figure 7.2: Evolution of temperature along the 30th Meridian. Isotherms are illustrated at every integer value, with a bold contour representing the 0°C isotherm. A dashed white line represents the Mixed Layer Depth. Micro-maps of meridional transects are visible at the right. Lines represent CTD profiles. Cruises from top to bottom: JR16006 (summer 2017), HH180101 (winter 2018), HH180423 (spring 2018), JR17006 (summer 2018).

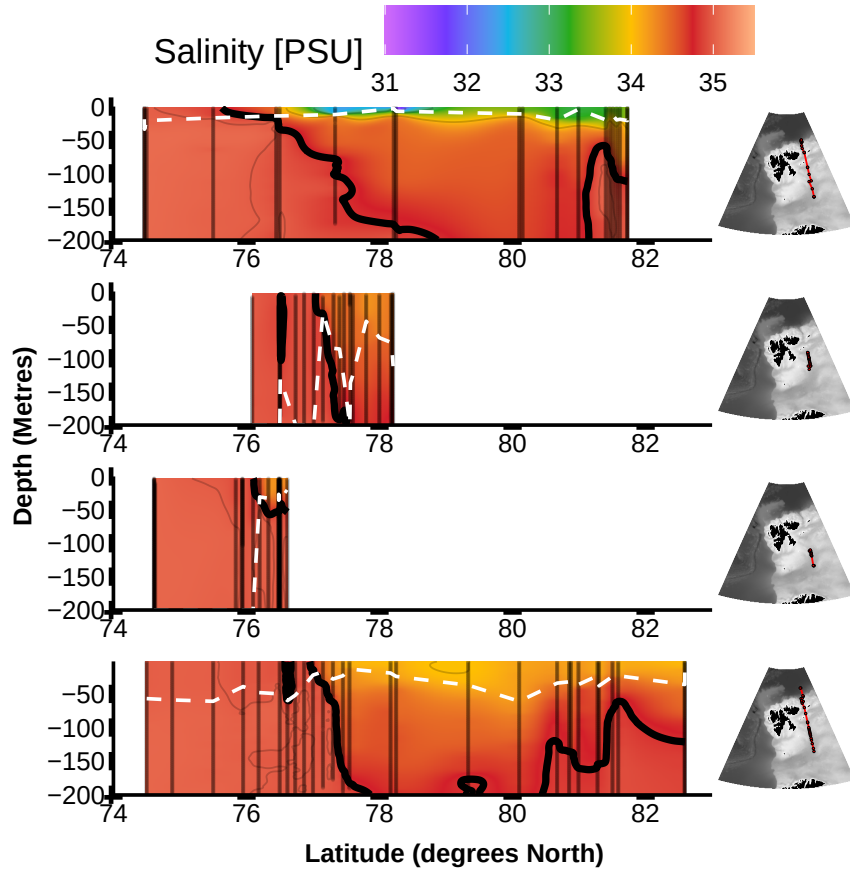


Figure 7.3: Evolution of practical salinity along the 30th Meridian. Isohalines are illustrated at every integer value, with a bold contour representing the 34.8 PSU isohaline. A dashed white line represents the Mixed Layer Depth. Micro-maps of meridional transects are visible at the right. Lines represent CTD profiles. Cruises from top to bottom: JR16006 (summer 2017), HH180101 (winter 2018), HH180423 (spring 2018), JR17006 (summer 2018).

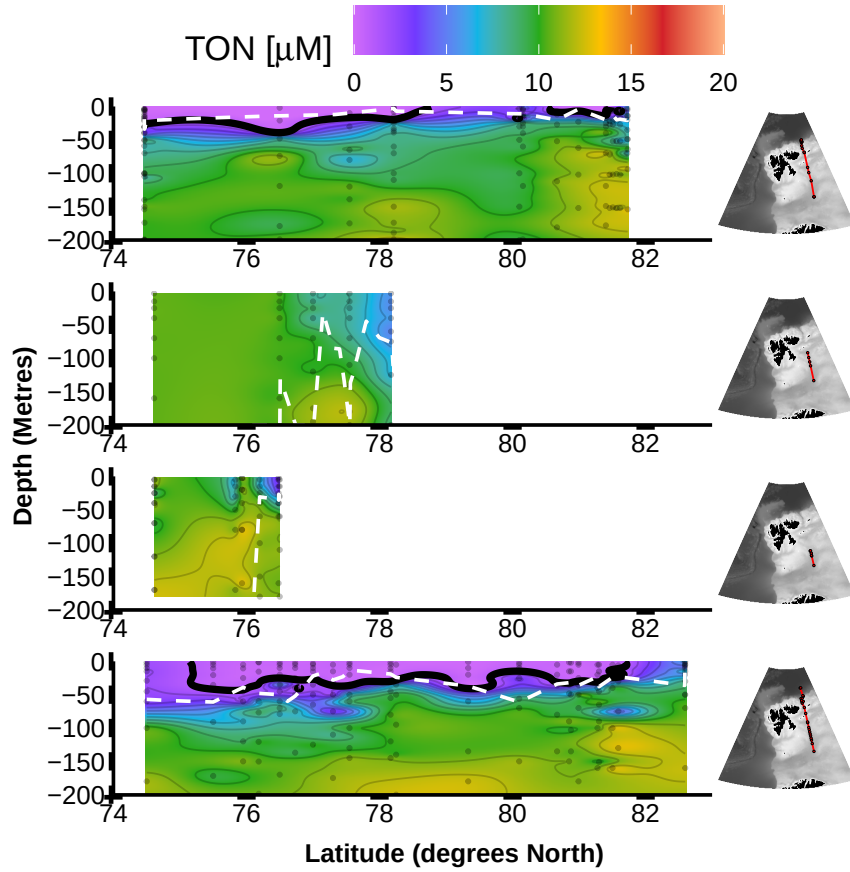


Figure 7.4: Evolution of Total inorganic Oxides of Nitrogen concentration along the 30th Meridian. Isopleths are illustrated at every integer value, with a bold contour representing the 1 μM isopleth. A dashed white line represents the Mixed Layer Depth. Micro-maps of meridional transects are visible at the right. Points represent *in-situ* sample sites. Cruises from top to bottom: JR16006 (summer 2017), HH180101 (winter 2018), HH180423 (spring 2018), JR17006 (summer 2018).

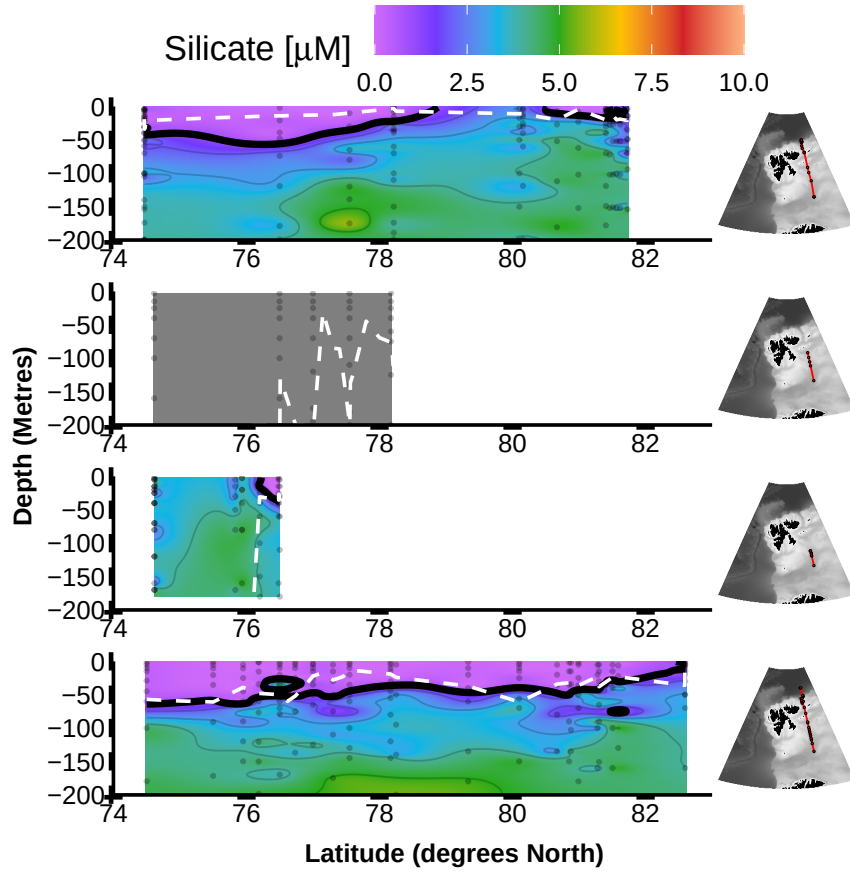


Figure 7.5: Evolution of silicic acid concentration along the 30th Meridian. Isopleths are illustrated at every integer value, with a bold contour representing the 1 μM isopleth. A dashed white line represents the Mixed Layer Depth. Micro-maps of meridional transects are visible at the right. Points represent *in-situ* sample sites. Cruises from top to bottom: JR16006 (summer 2017), HH180101 (winter 2018), HH180423 (spring 2018), JR17006 (summer 2018). Data absent for HH180101.

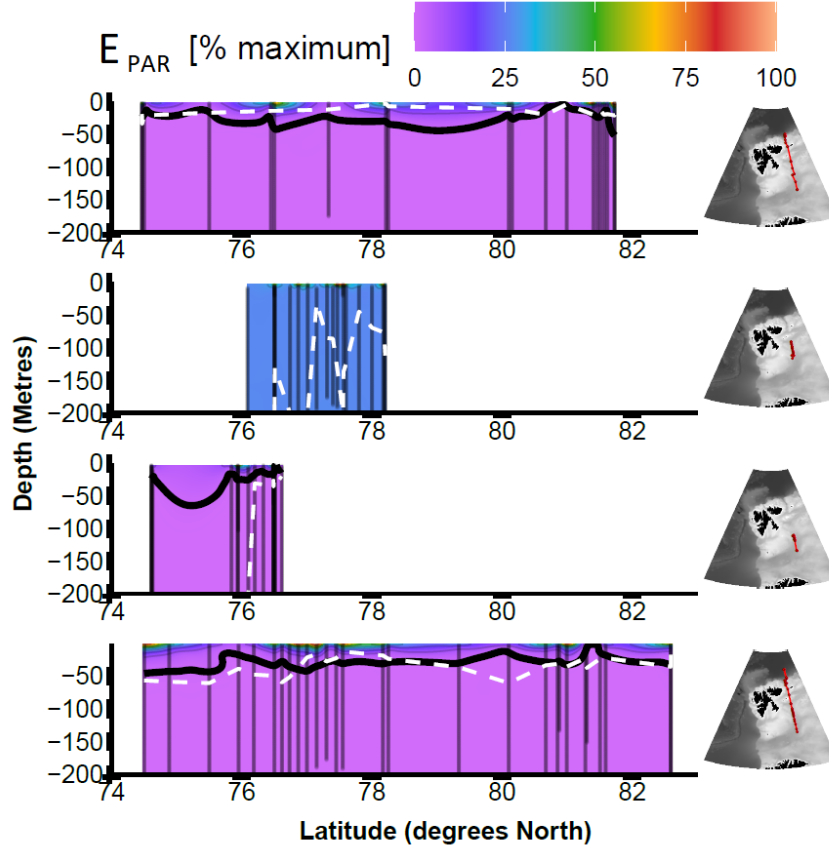


Figure 7.6: Evolution of normalised Photosynthetically Active Radiation (E_{PAR}) along the 30th Meridian. Isopleths are illustrated at every integer value, with a bold contour proxying the euphotic depth where values decline to 1% of the maximum E_{PAR} value recorded in each cruise. A dashed white line represents the Mixed Layer Depth. Micro-maps of meridional transects are visible at the right. Lines represent CTD profiles. Cruises from top to bottom: JR16006 (summer 2017), HH180101 (winter 2018), HH180423 (spring 2018), JR17006 (summer 2018). Aberrant pattern in HH180101 results from polar night conditions.

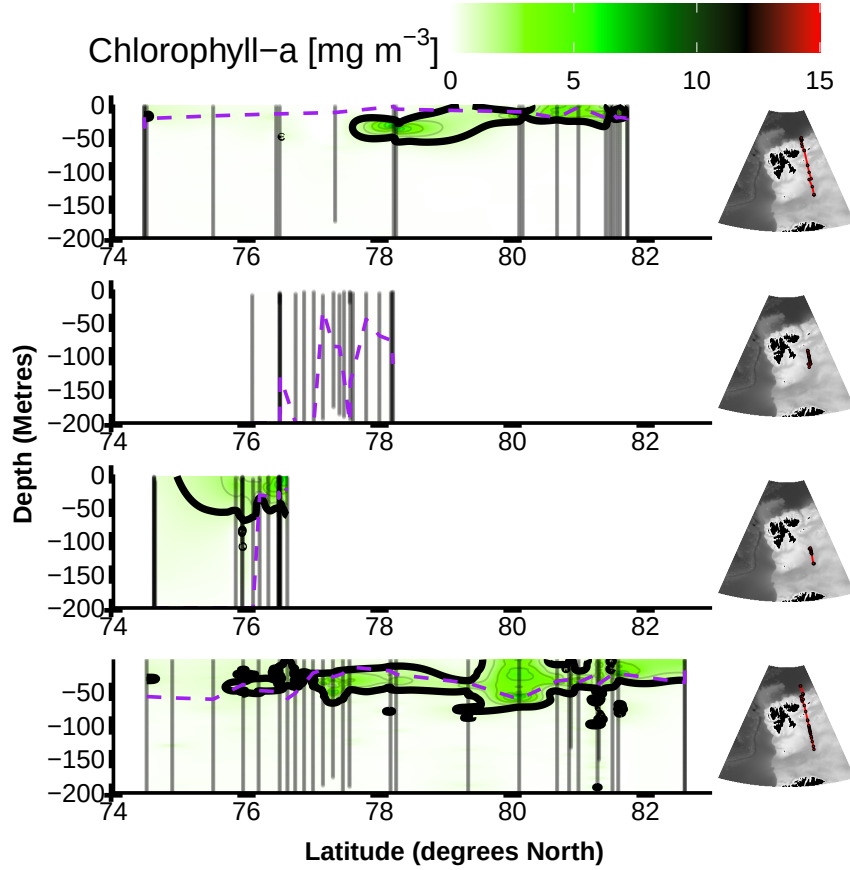


Figure 7.7: Evolution of fluorometrically-derived *in-vivo* [*chl-a*] concentration along the 30th Meridian. Isopleths are illustrated at every integer value, with a bold contour representing the 1 mg m⁻³ isopleth. A dashed white line represents the Mixed Layer Depth. Micro-maps of meridional transects are visible at the right. Lines represent CTD profiles. Cruises from top to bottom: JR16006 (summer 2017), HH180101 (winter 2018), HH180423 (spring 2018), JR17006 (summer 2018).

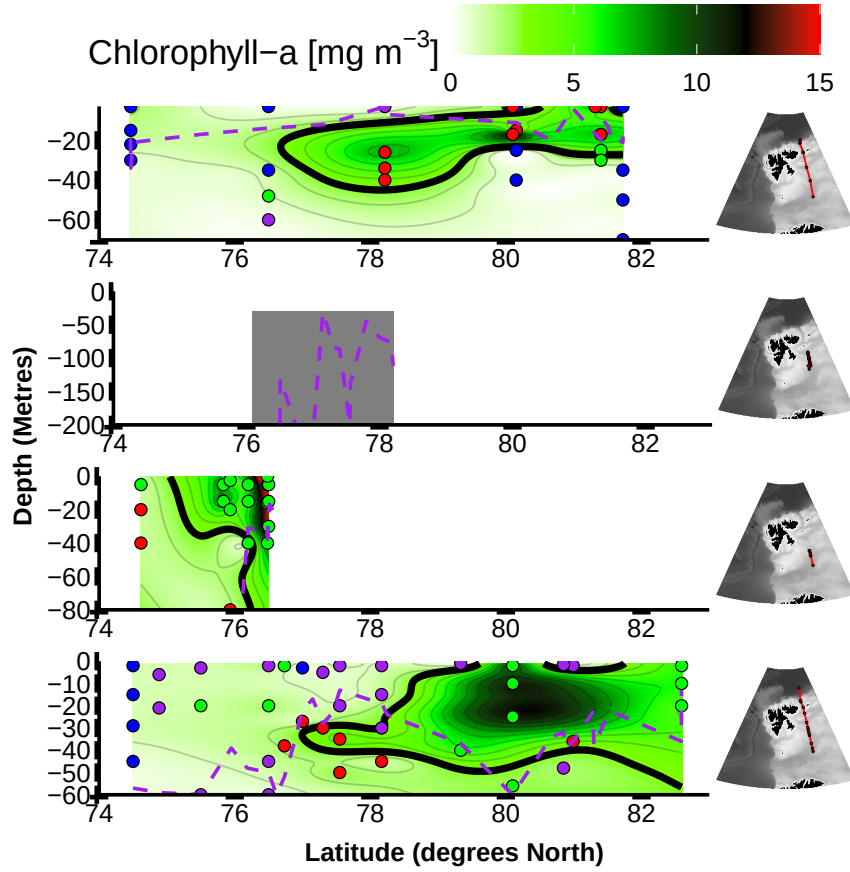


Figure 7.8: Evolution of acetone-extracted fluorometrically-derived [*chl-a*] concentration along the 30th Meridian. Isopleths are illustrated at every integer value, with a bold contour representing the 1 mg m⁻³ isopleth. A dashed purple line represents the Mixed Layer Depth. Micro-maps of meridional transects are visible at the right. Cruises from top to bottom: JR16006 (summer 2017), HH180101 (winter 2018), HH180423 (spring 2018), JR17006 (summer 2018). Points represent sample depths. Point colour indicates bio-optical group: green (diatom), red (colonial *Phaeocystis*), blue (nanophytoplankton), purple (dinoflagellate). Bio-optical group definitions are clarified in chapters 2-3.

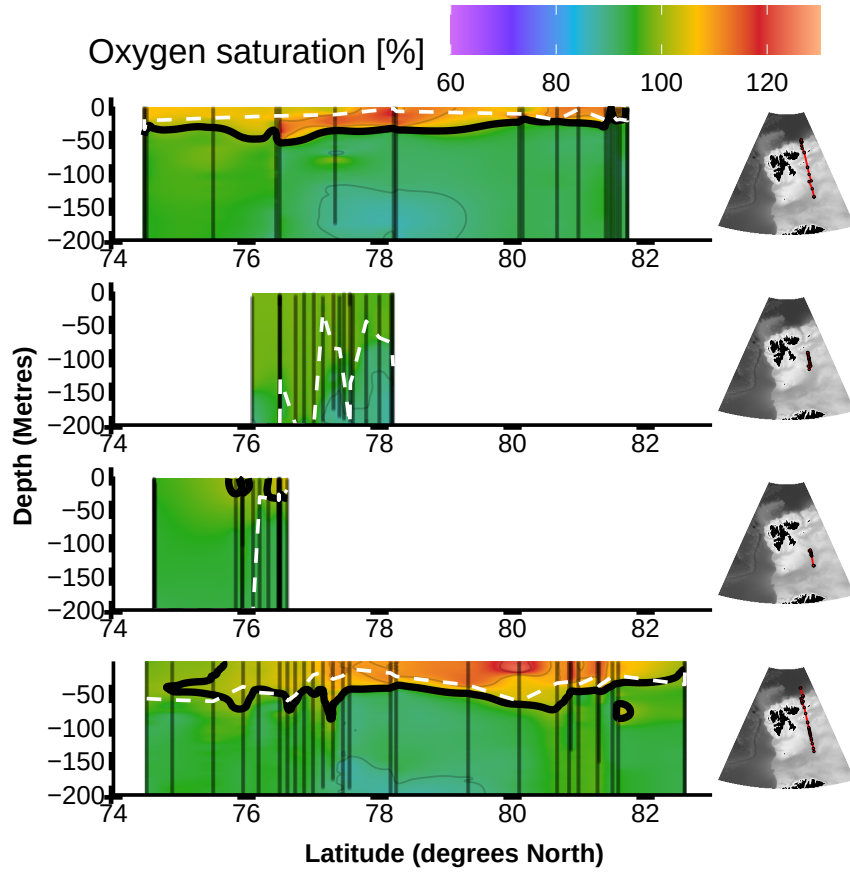


Figure 7.9: Evolution of oxygen saturation along the 30th Meridian. Isopleths are illustrated at every decile, with a bold contour representing the 100% saturation isopleth. A dashed white line represents the Mixed Layer Depth. Micro-maps of meridional transects are visible at the right. Lines represent CTD profiles. Cruises from top to bottom: JR16006 (summer 2017), HH180101 (winter 2018), HH180423 (spring 2018), JR17006 (summer 2018).

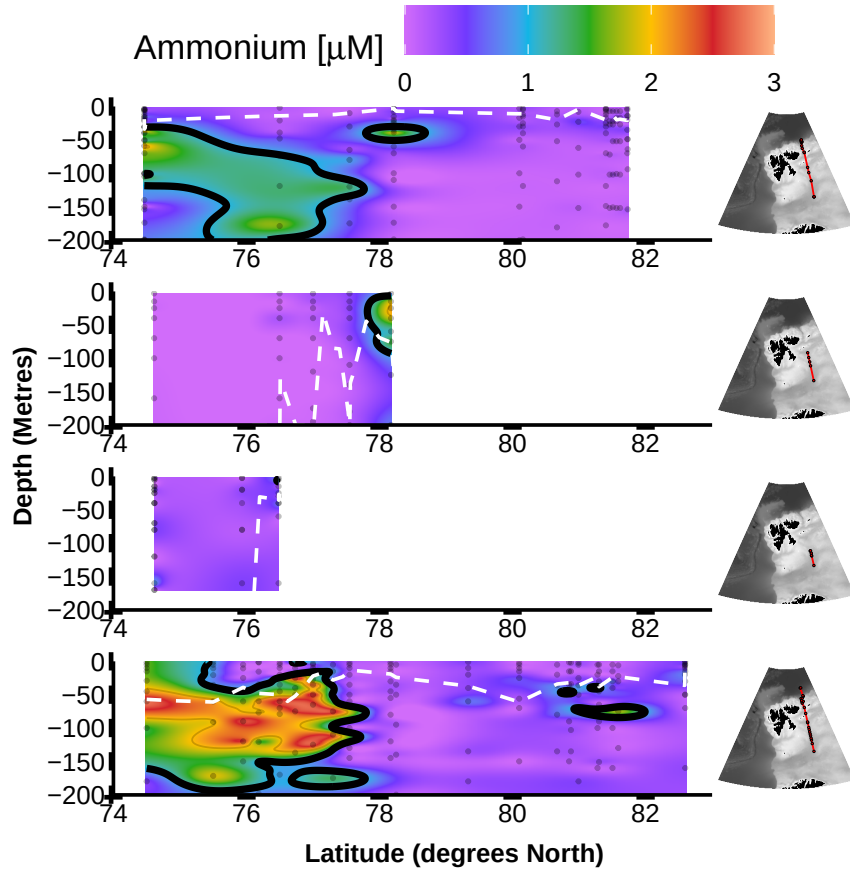


Figure 7.10: Evolution of ammonium concentration along the 30th Meridian. Isopleths are illustrated at every integer value, with a bold contour representing the 1 μM isopleth. A dashed white line represents the Mixed Layer Depth. Micro-maps of meridional transects are visible at the right. Points represent *in-situ* sample sites. Cruises from top to bottom: JR16006 (summer 2017), HH180101 (winter 2018), HH180423 (spring 2018), JR17006 (summer 2018).

References

- Cottier, Finlo R. 2018. “RRS James Clark Ross JR17006,” https://www.bodc.ac.uk/resources/inventories/cruise_inventory/report/16981/.
- Cottier, Finlo R, and Jørgen Berge. 2018a. “Helmer Hanssen HH180101,” https://www.bodc.ac.uk/resources/inventories/cruise_inventory/report/16917/.
- . 2018b. “Helmer Hanssen HH180423 (Arctic PRIZE, Arctic SIZE),” https://www.bodc.ac.uk/resources/inventories/cruise_inventory/report/17059/.
- Holm-Hansen, Osmund, Carl J Lorenzen, Robert W Holmes, and John DH Strickland. 1965. “Fluorometric determination of chlorophyll.” *ICES Journal of Marine Science* 30 (1): 3–15.
- Hopkins, Jo. 2018. “RRS James Clark Ross JR16006,” https://www.bodc.ac.uk/resources/inventories/cruise_inventory/report/16774/.
- Milligan, Allen J, Uriel A Aparicio, and Michael J Behrenfeld. 2012. “Fluorescence and non-photochemical quenching responses to simulated vertical mixing in the marine diatom *Thalassiosira weissflogii*.” *Marine Ecology Progress Series* 448:67–78.

Chapter 8

Appendix: Mathematical methods

8.1 Opening

This appendix serves as a guide to introduce readers to the mathematical methods employed in this thesis. The complex nature of variation in biological systems requires the application of sophisticated multivariate statistical methods, in order to discover and explore structure, use existing datasets to classify new input, or explain variance – the key ambitions that underlie the biological sciences. Open-source coding platforms, with pre-built statistical packages, mean that such advanced statistical methods are more accessible to researchers than ever, and there is great temptation for researchers to employ off-the-shelf statistical packages as black-boxes. However, it is necessary that a researcher should possess a sufficient understanding of these methods to ensure that they are applied carefully, with explicit knowledge of their limitations, so that the researcher does not uncover illusory results. This appendix provides a selection of illustrative examples, beginning by introducing and describing widely-used multivariate statistical methods. It is hoped that, in so doing, the nature of these mathematical methods is better clarified so that an astute reader may fully understand the work conducted in this thesis, see potential avenues by which it may be

improved upon and recognise any errors that may exist herein.

8.2 1 Principal Component Analysis

Principal Component Analysis (PCA) is one of the most widely employed multivariate statistical methods in empirically-driven research, with applications in diverse fields from sociology, geology, biology to economics and medicine.

A fundamental problem that confronts researchers in all these fields, is that a set of N samples may be described by a matrix X , with N rows and P columns, representing P variables. The researcher may wish to know in what ways the samples meaningfully differ from one another – the variance structure. However, a large number of variables, which may be intricately related to one another, can easily become intractable even to the experienced researcher – especially if the researcher does not possess knowledge a priori of the likely form of this structure.

A method is hence required to reduce the dimensionality (the total number of distinct variables over which a researcher must consider variation; effectively ‘ P ’) and present a perspective on the dataset that summarises the key hypothetical axes over which this structure in variance is distributed.

The researcher therefore desires a means to take the set of P variables describing variation, and define a new set of variables P' , organised such that they represent the variance between the N samples, while summarising any relationships that may exist between the P original variables.

Assuming that any relationships between the P variables can be approximated as linear, we can therefore state that the new set of variables P' will take the form

$$P'_i = T_1 P_1 + T_2 P_1 \dots T_P P_P, \quad (8.1)$$

where the i_{th} new variable is a linear combination of the P existing variables, weighted by a set of coefficients (T) .

In a simple case, we may imagine a varied batch of N currant buns available at a bakery that can be described by the linear combination of P ingredients – flour, sugar, fat and currants. The buns may vary widely in their total mass. The recipe of the bread mixture is 2 parts flour and fat, 1 part sugar and 1 part currants. Random processes of mixing will cause the exact balance to vary slightly across the buns, and this effect will intuitively be most pronounced in currants, as some buns will have a larger number than others. Hence we can intuitively see that deviation in the ratio of currants to dough from the expected ratio is the principal source of variance between the buns.

As the ratio of all ingredients, besides currants, are fairly constant with respect to flour, we can summarise the composition of the buns by producing a plot of currants against flour (see [8.1](#)).

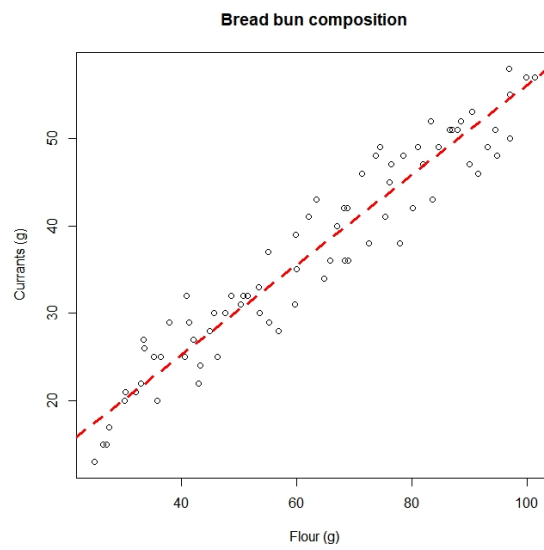


Figure 8.1: The ratio of currants to flour in a batch of bread buns. It is intuitive to us that the deviation of the ratio of currants to dough (proxied by flour) from the average ratio is the meaningful source of variance. However, conclusions like this may be more difficult to reach in less tractable and more complicated datasets.

It is intuitive from Figure 8.1 that there is a major axis of variation between currants and flour, and that if we are interested in the flavour of buns it is actually deviation perpendicular (or orthogonal) to this axis that should preoccupy us.

The `prcomp()` function, the standard implementation of PCA in R, can decompose the variation between the ingredients, according to the singular value decomposition:

$$X = U\Sigma W^T, \quad (8.2)$$

where U is an N -size square matrix, the columns of which are known as ‘left singular vectors’, W is a P -size square matrix, the columns of which are known as ‘right singular vectors’, Σ is a rectangular $N \times P$ matrix, the diagonal entries of which $(\sigma_{i,j})$ are known as ‘singular values’ and T is a P -size square matrix of coefficients.

This description is opaque, however if we posit that the matrix of data X is a contortion, that could be represented by r unique linearly-independent columns (or ‘dimensions’), then we can state.

$$X = \sum_{i=1}^r \sigma_i U_i W_i^T, \quad (8.3)$$

from which it becomes evident that singular value decomposition is breaking down X into a weighted sum of linear transforms of $U_i W_i^T$.

This means of decomposition therefore immediately reminds us of our desire to find a new set of variables P'_i that are linear recipes of the original variables P , and we can see that W^T is a rotation matrix onto a new set of axes, and that the values Σ describe the proportion of variance of the N samples described by X along each of these new axes.

Hence the decomposition can be sorted into a set of new axes, analogous to P'_i , that can be ranked by the variance they explain in X .

This is illustrated directly by a PCA performed on the matrix of currant bun ingredients.

It is plain that the leading axes of this decomposition represent a rotation of the original data in Figure 8.1, and that one of these axes described variation in currants perpendicular to the relationship of currants to flour (demonstrated by a lack of any linear relationship of PC2 with flour; see Figure 8.2).

We can, furthermore, by reviewing W^T , see that the coefficients, also known as ‘loadings’ (see Table 8.1), that describe the rotation have discovered the recipe of our bread buns (see PC1), and the computation of the dot product of vectors PC1 and PC2 = 0 and therefore shows that the second axis is orthogonal to the recipe. Inspection shows the highest magnitude of loading on PC2 is the contribution of currants, conforming to our expectation that it is variance in currants that meaningfully distinguishes the buns from one another.

Table 8.1: *Rotation matrix from PCA performed on bread bun compositions.*

Ingredient	PC1	PC2	PC3	PC4
Flour	0.6239643	-0.29057246	0.30554521	-0.65793487
Fat	0.6263755	-0.22581044	0.09860497	0.73955424
Sugar	0.3213863	0.02328317	-0.93621242	-0.14026785
Currants	0.3391654	0.92953493	0.14291776	-0.02249813

Hence, we have arrived at an intuitive understanding that PCA is an exploratory statistical method that can rotate an existing, convoluted, matrix of continuously-distributed data onto a new set of axes that are convenient for exploring its variance structure. We have gathered the understanding, in this process, that this method will only produce meaningful decompositions for datasets where relationships between variables can be approximated as linear, and we must acknowledge that as variance may include random variation, error or relationships to variables that are not represented; that the output of PCA is not guaranteed to present a researcher with a meaningful synopsis of data structure.

That said, when the conditions necessary to use PCA are met, it represents a powerful

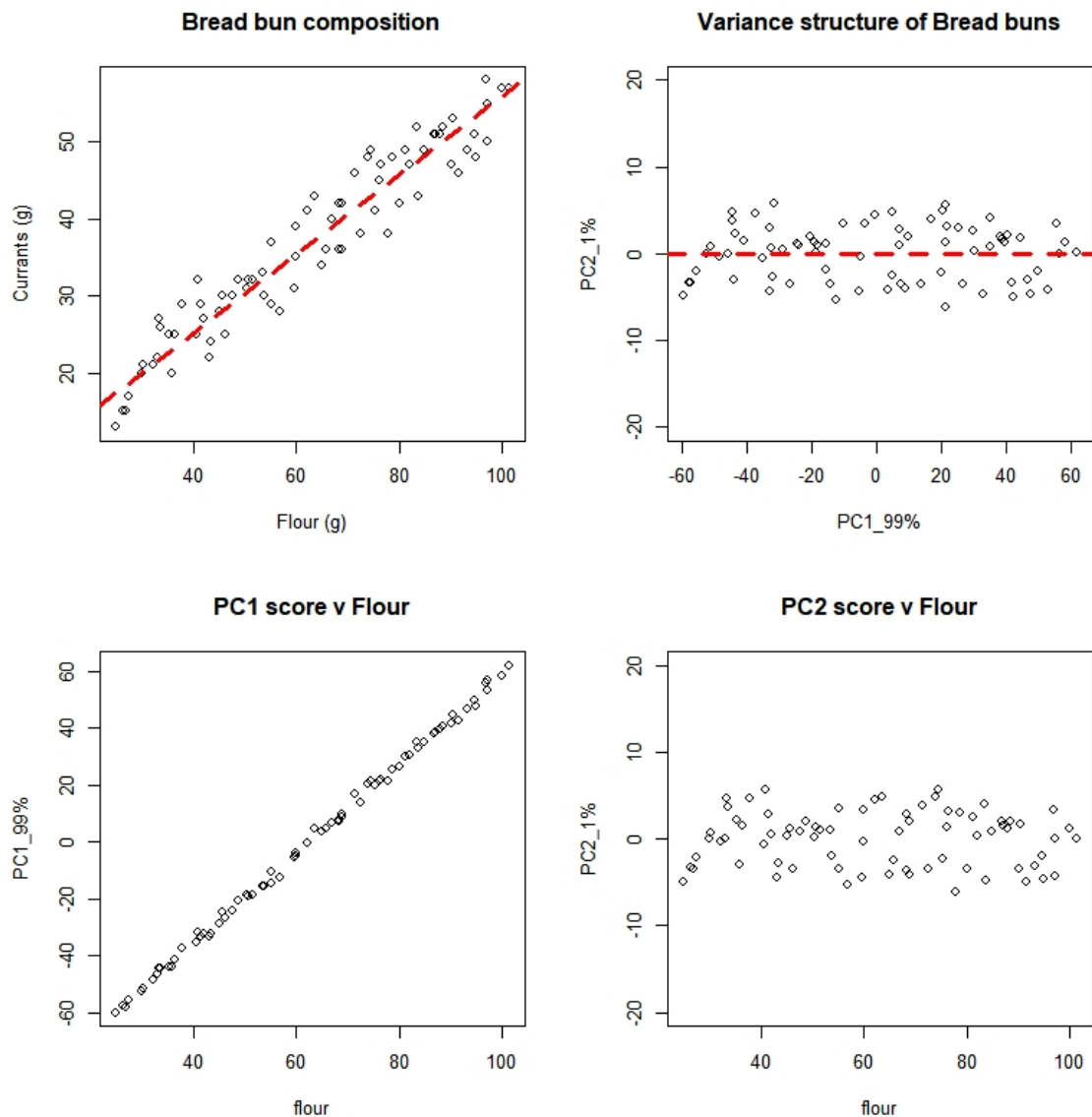


Figure 8.2: A set of diagnostic plots, demonstrating that PCA operates as a rotation, finding a series of orthogonal new axes to maximally describe variance in the original dataset. It can be seen, for example, that PC1 is reproducing the original recipe of the dough (as it correlates strongly with flour). A lack of correlation between PC2 and flour is consistent with the orthogonal construction of the new axes.

tool to make plain the salient structure of a dataset. The convenience of having a set of new derived variables, ranked by explained variance, means the researcher can choose to reduce dimensionality by discounting axes that explain less variance. Indeed, this is often a necessary step, because the constraint that the new axes must be mutually orthogonal may mean that higher axes do not themselves possess meaningful structure, but are simply an echo of lower axes that explain greater variance.

A fuller, more mathematical description of the implementation of PCA in R can be found in Venables and Ripley (2002).

I now turn to a real example, using ocean-colour data, to demonstrate the utility of PCA and the operational procedure for its use in my thesis.

The codes required to perform this example analysis are available through:

<https://github.com/aorkney/thesis>

We will begin by downloading ocean-colour data for the 6th of July, 2016, in the Barents Sea, during which there is a good example of a coccolithophore bloom (Hansen et al. 2016). The data is distributed across six spectral channels ($\lambda = 412, 443, 490, 510, 560$ and 665 nm). The requisite data can be downloaded by running the following script:

```
Download_ocean_colour_data.R
```

The code required to perform analysis is found in the following script:

```
PCA_example_and_coccolithophore_mask.R
```

A PCA decomposition upon this data shows that almost all of the variance (97%) occurs along the first axis, (Figure 8.3). The existence of several suborthogonal trends to this axis along PC2 evidence an ‘arch’ effect, when PCA fails to distinguish all modes of variance onto their own unique, mutually exclusive axes. This might occur because not all trends in the original dataset are linearly independent of one another.

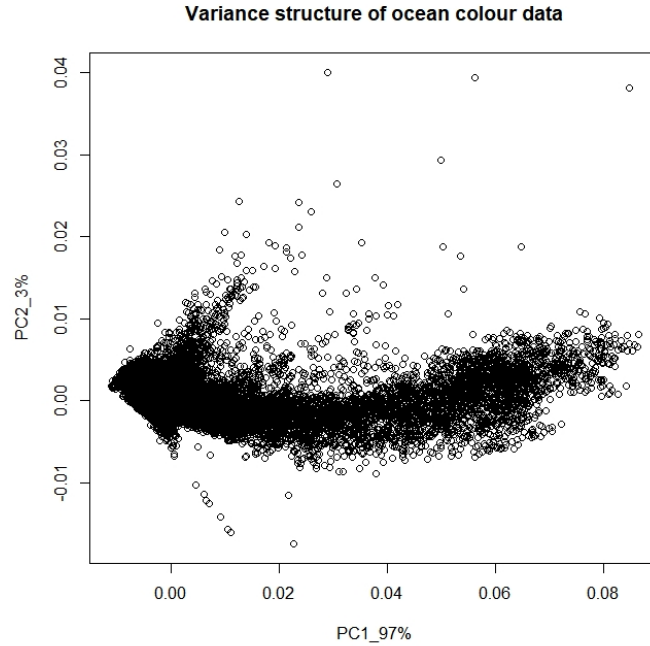


Figure 8.3: PCA decomposition performed on ocean-colour data.

This distribution in PC1-2 space provides us with an indication that the ocean-colour dataset we are exploring may contain multiple sources of variance that are convolved with one another to some degree. In particular, we should consider that we are viewing a dataset that represents an example of blooming coccolithophores; perhaps variation in coccolithophore concentration is also convolved with ocean-colour signals attributable to variable pigment biomass or sediment load in the water-column.

Let us explore this hypothesis by computing $[chl-a]$ concentration (I use the OC6_MERIS algorithm from O'Reilly and Werdell (2019)), and plotting it against PC1 (Figure 8.4).

We can easily see, from inspecting Figure 8.4, that $[chl-a]$ correlates with PC1 for some portion of the data, when PC1 exceeds 0. This implies that there is a source of variance in our dataset that could bias the inference of $[chl-a]$ – a worrisome conclusion that might

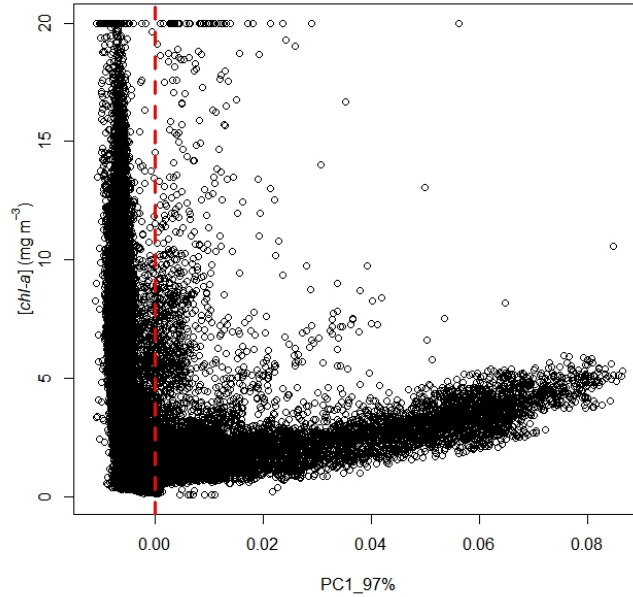


Figure 8.4: Inferred $[chl-a]$ (capped at 20 mg m^{-3}) plotted as a function of PC1. A red dashed line is placed at $PC1 = 0$.

preoccupy the mind of a biological oceanographer who wishes that such derived variables should be faithful. Let us explore further by plotting a pseudo true-colour image of our data against the PC1 score, so that we might gain a greater insight into what exactly PC1 is reflecting (Figure 8.5).

Indeed, it becomes clear that PC1 is very likely proxying the concentration of coccoliths in the surface ocean. These liths imbue the ocean with a milky colouration by scattering incident light. This scattering flattens the blue-to-green ratio of reflected light from the ocean surface. The algorithm we are using to infer $[chl-a]$ also makes use of the blue-to-green ratio, hence this probably explains the relationship in Figure 8.4. We can reason, then, that we do not wish to include pixels influenced by coccolithophores in any remote-sensing assessment of $[chl-a]$, because substances other than $[chl-a]$ are significantly biasing

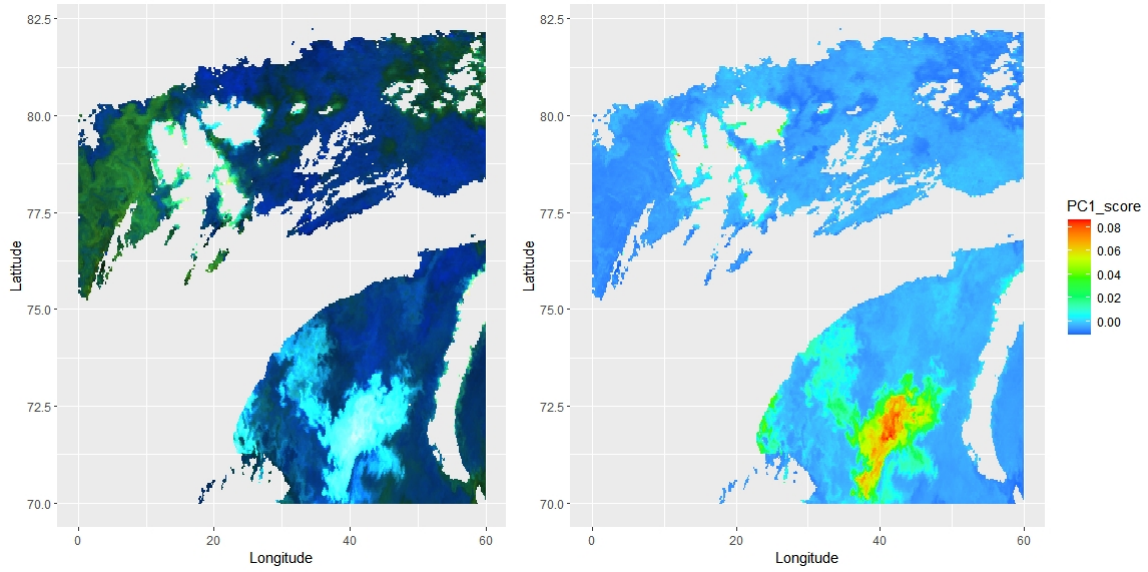


Figure 8.5: **Left:** a pseudo true-colour image of the ocean-colour data. A coccolithophore bloom is visible in the bottom-right, near Novaya Zemlya. **Right:** A filled contour plot of PC1 score, showing a close correspondence to the geographic region occupied by coccolithophores.

the inference of this derived variable. Let us then, observing that it is pixels in which PC1 exceeds 0 that PC1 and inferred $[chl-a]$ correlate, mask all pixels in which PC1 exceeds 0 (Figure 8.6).

It is evident, inspecting Figure 8.6, that an effective mask for identifying ocean-colour retrievals affected by coccolithophores has been discovered. We can therefore reflect on the utility of PCA in allowing us to rapidly process this ocean-colour dataset and come to some salient meaningful conclusions. Not only did we avoid the requirement to study the 6 spectral channels individually – greatly accelerating our analysis – but we also identified potential sources of bias that could affect $[chl-a]$. This provides us with a clear understanding for the reasons PCA is described as an ‘exploratory’ analysis.

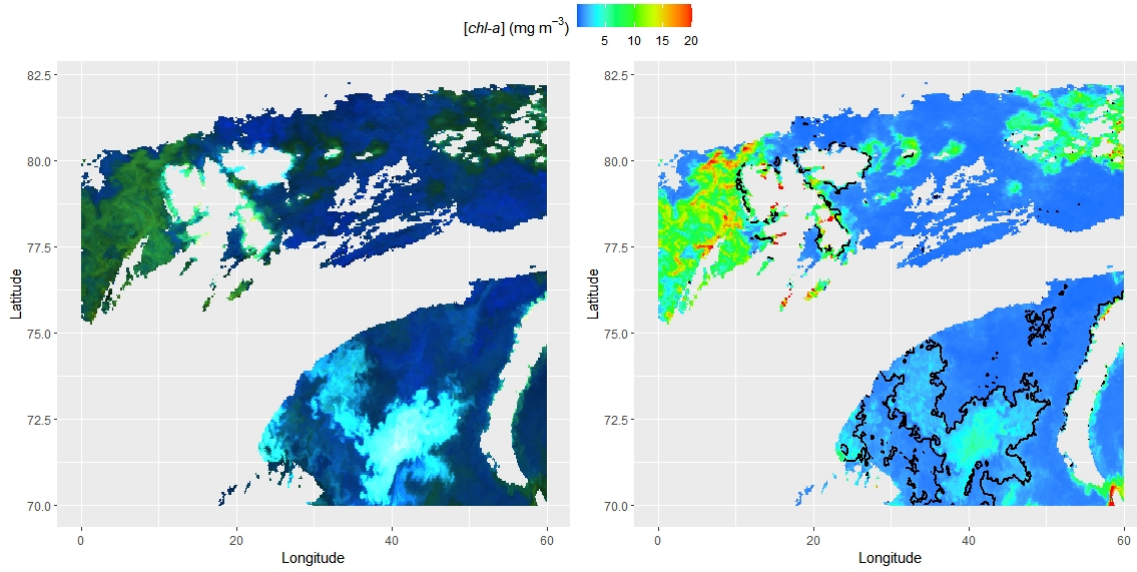


Figure 8.6: **Left:** Pseudo true-colour image of the Barents Sea. Green phytoplankton blooms are visible in the North and coccolithophores in the South. **Right:** A field of inferred $[chl-a]$, and a black contour encircling pixels that should not be considered, for they are influenced by coccolithophores.

8.3 2 Two Block Partial Least Squares

Having discussed PCA, let us now move on to discuss related exploratory statistical approaches. We shall begin by considering Two Block Partial Least Squares (2BPLS).

Let us consider the situation in which we have two different blocks of data; X_1 and X_2 , which we expect are related by some unknown structure that can be approximated as a series of independent linear functions, similar to those we have already encountered while performing PCA.

Now that we understand Singular Value Decomposition, we can progress expeditiously to using it to solve this problem.

We will first compute the covariance matrix between X_1 and X_2 , representing the pairwise covariance between all constituent variables.

Let us call this matrix R , we can reason that R can be divided into quarters as in Table

8.2:

Table 8.2

R_{11}	R_{12}
R_{21}	R_{22}

We can see that R_{11} represents covariance internal to X_1 , and that R_{22} represents covariance within X_2 . R_{12} and R_{21} are transposes of one another, and represent cross-covariance between X_1 and X_2 . This is what we care about, but we need only consider one of them.

Let us select R_{12} . We can, as we did when investigating PCA, explore its variance structure by performing Singular Value Decomposition.

$$R_{12} = U\Sigma W^T, \quad (8.4)$$

where U will now represent an orthogonal series of vectors decomposing variance in X_1 , W^T will represent the decomposition of X_2 , and Σ is, again, a rectangular matrix in which the diagonal elements describe the relative contributions to explained variance in R_{12} .

We can see, from this description, that the decomposition will retrieve a series of pairs of vectors – and that the distribution of the N samples over each of these combinations will represent the way covariance is distributed between X_1 and X_2 . Each pair of such vectors is known as a ‘salience’, and each salience is mutually orthogonal to the others.

Let us elucidate these concepts further by performing a 2BPLS decomposition on an example dataset.

The codes required to perform this example analysis are available through:

<https://github.com/aorkney/thesis>

First, it is necessary to download the dataset

```
SST_and_storms.csv
```

This is a dataset I produced by downloading SST data from the

```
ARCTIC_REANALYSIS_PHYS_002_003
```

dataset, and wind-speed from the

```
WIND_GLO_WIND_L3_REP_OBSERVATIONS_012_005
```

dataset. Both of these datasets are available on

```
https://marine.copernicus.eu/
```

I averaged data for September, over a series of different years, within 1 degree latitude-longitude boxes, so that the dataset is small and manageable. The csv file contains columns for latitude, longitude, year, SST, and storminess. Storminess is computed as the average number of days, within a grid cell, that wind speeds exceed 10 metres per second.

The following analysis can be performed by running the code

```
2BPLS_script.R
```

In the Barents Sea, we expect a South-west to North-east decline in Sea Surface Temperature. We do not necessarily expect stormy days to have a strong relationship to geography in a given year. Perhaps there will be more in the warmest waters, closest to the Atlantic inflow region, where storms might enter the Barents Sea. A plot of SST and storminess in September 2002 conforms to our expectations (Figure 8.7).

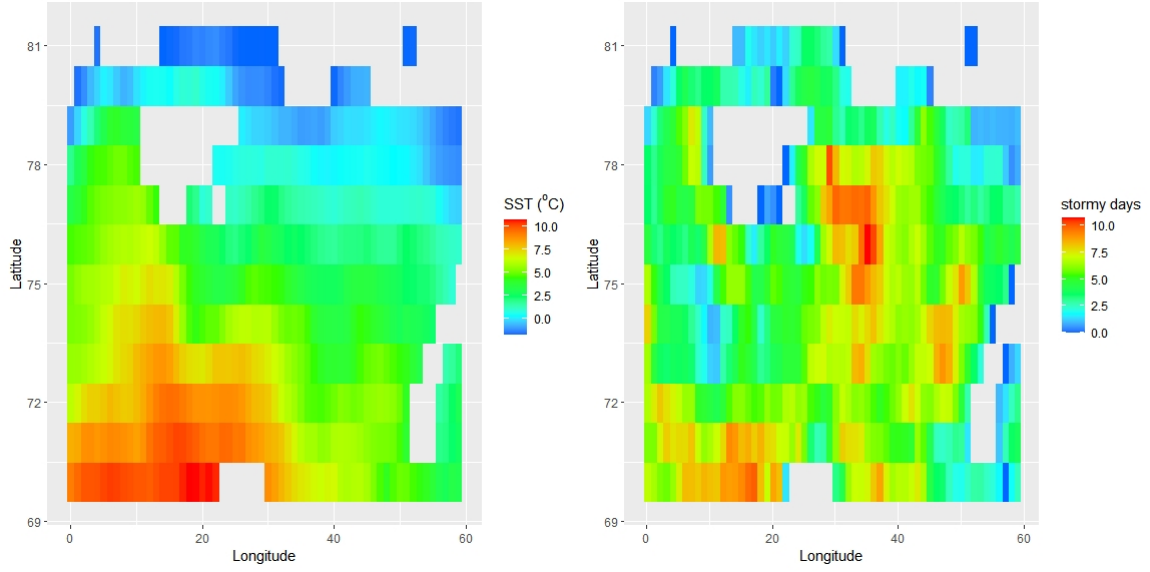


Figure 8.7: **Left:** SST in the Barents Sea September 2002. **Right:** Storminess in the Barents Sea September 2002.

Let us explore the covariance structure between the environmental conditions and geography; we will begin by separating our data in X_1 , a matrix of latitude and longitude values, and X_2 ; a matrix of corresponding SST and storminess. We then perform the 2BPLS decomposition, and we can view the resulting structure of the leading mode of covariance (salience) by plotting the first columns of U and W^T against one another (Figure 8.8).

It is evident, in Figure 8.8, that the first salience explains variance in temperature. Reviewing salience 2 shows that it explains variance in storminess (Figure 8.9).

Furthermore, computation of the explained variance of these saliences;

$$E_i = \sigma_i^2 / \sum_{i=1}^i \sigma_i^2, \quad (8.5)$$

shows that the first salience explains almost all of the covariance (99%). If we view the left and right PLS vectors (Table 8.3), we can see that the left component of first salience (PLS1) is composed of increasing latitude and longitude, while the corresponding right component

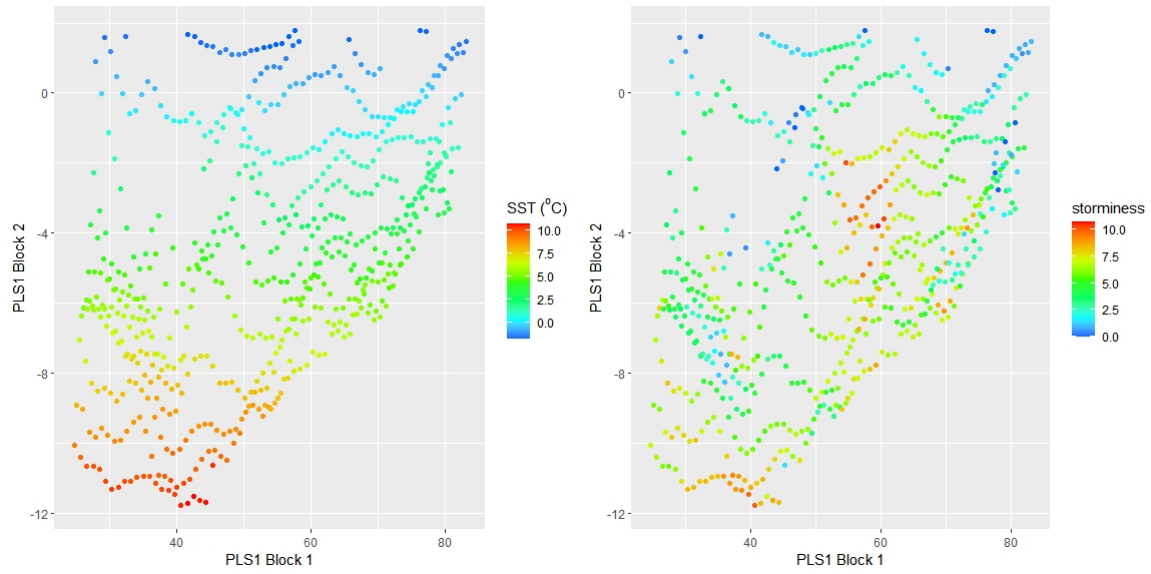


Figure 8.8: **Left:** Saliense pair 1, showing a strong linear distribution of temperature across the axes. **Right:** The distribution of storminess is less well organised across these axes.

of the first salience (Table 8.4) is dominated by declining SST.

Table 8.3

variable	PLS1	PLS2
latitude	0.3524933	0.9358143
longitude	0.9358143	-0.3524933

Table 8.4

variable	PLS1	PLS2
storminess	-0.1327878	-0.9911445
SST	-0.9911445	0.1327878

The second salience is shaped by declining storminess with increasing latitude.

Hence, we have used 2BPLS as an exploratory analysis, confirming our expectations that

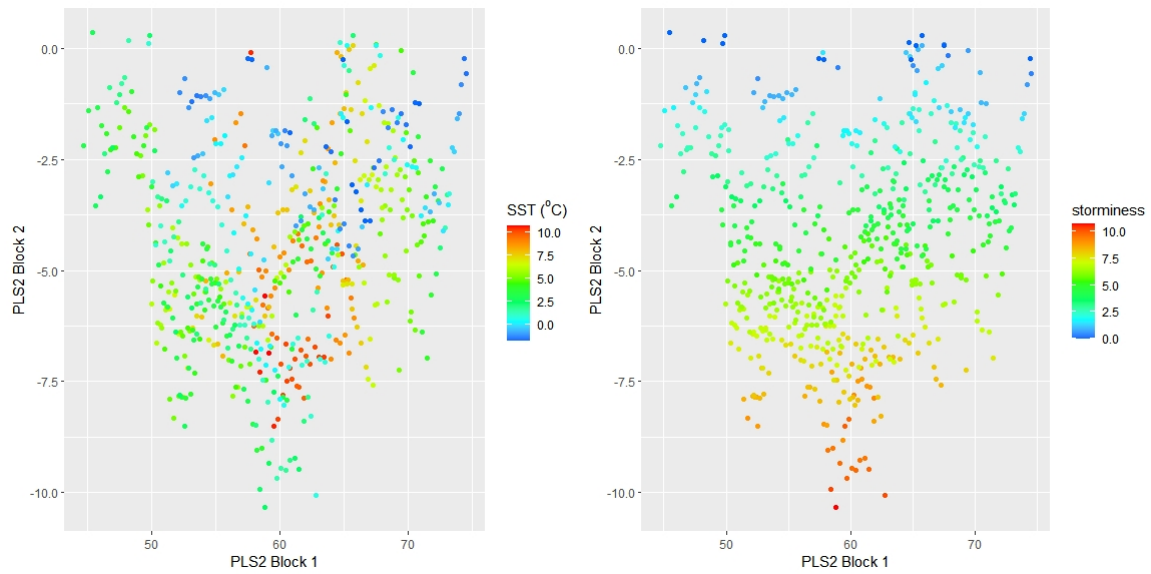


Figure 8.9: **Left:** Saliense pair 2, showing a disorganised distribution of temperature across the axes. **Right:** The distribution of storminess is more well organised across these axes.

there should be a dominating South-west to North-east gradient in SST, and that there should be a semi-independent pattern of storminess super-imposed upon this, with greater frequencies of storms in more Southern waters.

We can, furthermore, envision computing goodness-of-fit metrics or effect-sizes on the resultant saliences, to quantify their significance.

This example should emphasise the utility of 2BPLS as an exploratory method to probe relatedness between pairs of input data blocks. A more involved description, complete with biological morphological examples, can be found in Rohlf and Corti (2000).

8.4 3 Redundancy Analysis

Thus far our exploratory analyses have not posited any causal relationships between variables – exploring only the structure of variance across them.

However, if we consider our dataset of SST, storms and geography, we can easily see that there are dependent (affected causally) and independent (not causally affected) variables. For instance, we understand that a gradient in radiative forcing with latitude will partially cause the gradient in SST from the Southern to Northern Barents Sea. Conversely, we are well aware that the temperature of a grid cell of seawater in the Barents Sea cannot cause that grid cell to change its latitude; those are processes underlain by tectonic drift and are agnostic to variation in the weather.

Let us then use this a priori knowledge of the network of causal relationships in our dataset to our advantage, to explore it more rigorously.

In the trivial case, we can imagine that we might surmise that some portion of change in SST is caused by variation in latitude, and we can solve the equation

$$Y = mX + C, \quad (8.6)$$

where Y reflects the dependent variable SST, X reflects our independent variable, latitude, and m and C are constants to fit the relationship. We are assuming, in this construction, that the relationship can be approximated as linear.

We could even generalise our modelled relationship to include variation in an arbitrary number of independent variables:

$$Y = m_1X_1 + m_2X_2 + \dots m_iX_i + C, \quad (8.7)$$

and it would be relatively straight-forward to find solutions.

Consider, however, the more complex scenario where there may be multiple dependent variables, which may have some structure of covariance.

$$Y_1, Y_2 \dots Y_i \alpha X_1, X_2 \dots X_i, \quad (8.8)$$

In this case, we need a more sophisticated analysis than simple multiple regressions; we need some means of performing multiple regressions with both multiple dependent and independent variables. Redundancy Analysis (RDA) offers us that capability. Let us begin by simply undertaking what we know is trivial; let us perform multiple linear regressions of each variable represented by a multivariate matrix Y against some other multivariate matrix X :

$$\hat{Y} = x[x'x]^{-1}x'Y. \quad (8.9)$$

We can, thereafter, conclude that we have produced some matrix of fitted values of Y ($\hat{Y}$), that vary linearly with combinations of X , and that there is some residual of unexplained variance.

However, we do not yet have a helpful synopsis of the relationship between our dependent and independent variables.

We should stop and remind ourselves that we have already learned how to decompose matrixes of data to understand and simplify the variance structure inside them; Singular Value Decomposition. By performing PCA upon $\hat{Y}$ we can gain an insight into the variance structure in a Y that is explained by X . We can also, by reviewing the residuals of ($\hat{Y}$) (which can similarly be decomposed), understand the portion of variance that we have not been able to constrain with the explanatory variables in X . The results of RDA lend themselves to simple visualisation; variables in X can be projected as vectors across the distribution of the first pair of Principal Components of $\hat{Y}$, producing a ‘biplot’.

Let us clarify this mathematical description with an example, applying RDA to better understand the dataset we previously reviewed with 2BPLS.

The codes required to perform this example analysis are available through:

<https://github.com/aorkney/thesis>

First, it is necessary to download the dataset

```
SST_and_storms.csv
```

This is a dataset I produced by downloading SST data from the

```
ARCTIC_REANALYSIS_PHYS_002_003
```

dataset, and wind-speed from the

```
WIND_GLO_WIND_L3_REP_OBSERVATIONS_012_005
```

dataset. Both of these datasets are available on

```
https://marine.copernicus.eu/
```

I averaged data for September, over a series of different years, within 1 degree latitude-longitude boxes, so that the dataset is small and manageable. The csv file contains columns for latitude, longitude, year, SST, and storminess. Storminess is computed as the average number of days, within a grid cell, that wind speeds exceed 10 metres per second.

The following analysis can be performed by running the code

```
RDA_script.R
```

The results of RDA, applied to the weather data in the Barents Sea September 2002, are presented in Figure 8.10. I have scaled the axes in these plots to unity, and stated that the explanatory variable projections should have line thickness indicating the degree of variance they explain.

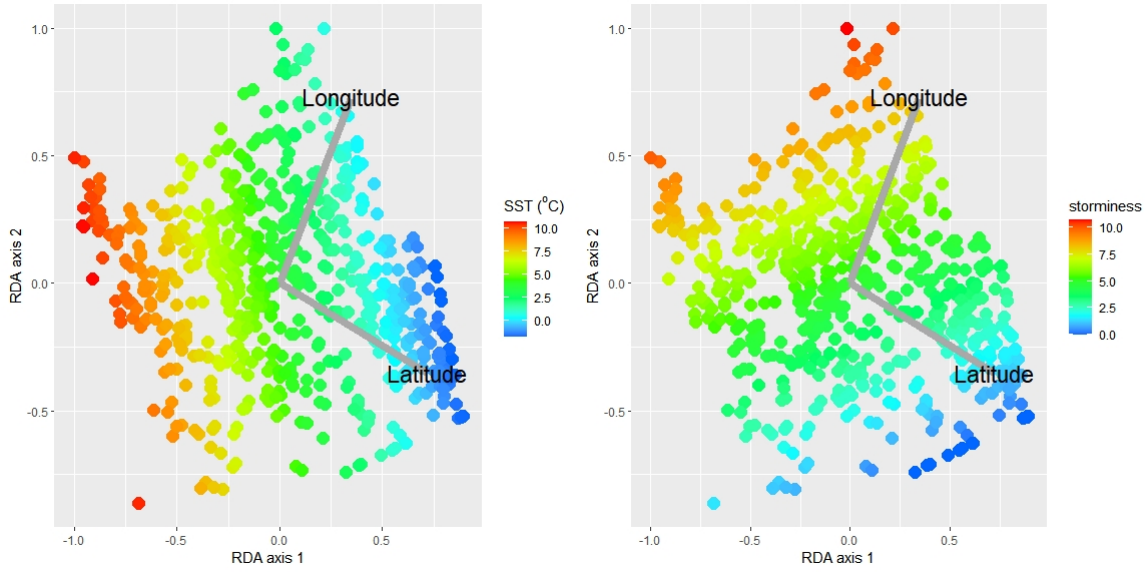


Figure 8.10: **Left:** A projection of latitude and longitude across the first two RDA axes, with the point distribution coloured by SST. **Right:** A projection of latitude and longitude across the first two RDA axes, with the point distribution coloured by storminess.

It is evident, from Figure 8.10, that RDA has retrieved the expected salient structure in the dataset; that variation in SST is strongly explained by latitude, with a minor longitudinal component. Similarly, storminess is also observed to decline with increasing latitude. Latitude and Longitude have similar contributions to explained variance in this dataset. The first axis explains 99% of the constrained variance in $\hat{Y}$. Overall 64% of variance in Y is explained. Analysis of the principal components of unconstrained variance shows that variation in storminess dominates, which is in keeping with our earlier observations that the distribution of storms in a given year is not a strict deterministic function of geography, but reflects stochastic variation in the weather (Table 8.5).

Let us now consider a more niche application of this statistical method. Let us say we wish to know whether the structure of explained variance of weather by geography in September 2003 is more similar to September 2002 than it is to September 2018. We might expect this is true, because a background of secular change has altered the hydrographic

Table 8.5

variable	PC1	PC2
storminess	-5.3449	-0.4257
SST	-0.8896	2.5578

conditions in the Barents Sea over this interval, with more Atlantic water flowing into the shelf (potentially relaxing latitudinal variation in SST for example).

We can achieve this by computing RDA for each of these years. Then we can summarise the relationships between the point distributions and the explanatory variables as a matrix of angles between the centre of mass and the projected explanatory vectors. This information can be presented in a variety of ways; a hierarchically-clustered dendrogram is one such means. An example dendrogram, describing the similarity of 10 randomly chosen grid-cells in September 2002, is presented in Figure 8.11.

A dendrogram (tree diagram) is a useful means to represent the data, because the similarity of different dendrograms can be computed by comparing their average cophenetic distances – the average number of branch differences between the same sample on two different dendrograms (see Galili (2015) for a more involved description of this process). For example, the dendrogram in Figure 8.11 can be compared to the same sampling of grid-cells from September 2003 (Figure 8.12), and it can be shown that – for this particular selection of a very small number of grid-cells – there is a low degree of agreement about the intersection between the sample ordination and the explanatory vectors.

An example application of this method is presented in Figure 8.13. This analysis was produced by running the script:

```
RDA_script2
```

Note that I have used permutation analysis in order to compute a density distribution for the cophenetic distance based correlation indices in Figure 8.13 a. I achieved this by re-

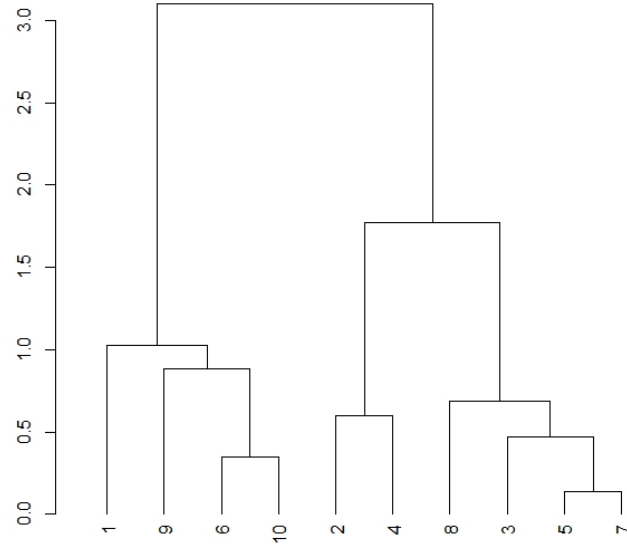


Figure 8.11: An example of a dendrogram, produced by hierarchical clustering on a distance matrix relating the angles between individual sampled grid-cells and explanatory environmental vectors (latitude and longitude).

sampling the number of individual grid-cells I considered to 300 (a much larger sample than the 10 previously considered). This means, of the several hundred grid-cells in the original image, I randomly select which to consider some 300 times and I allow grid-cells to be chosen more than once. I re-ran this analysis with fresh re-samplings 100 times, allowing me to generate a range of probable correlation indices. This re-sampling procedure is known as ‘boot-strapping’, and is a useful means to generate statistics that describe the accuracy of an estimate about a population of data without any necessary assumptions about the true nature of the density distribution.

A different form of permutation analysis has also been used to assess the significance

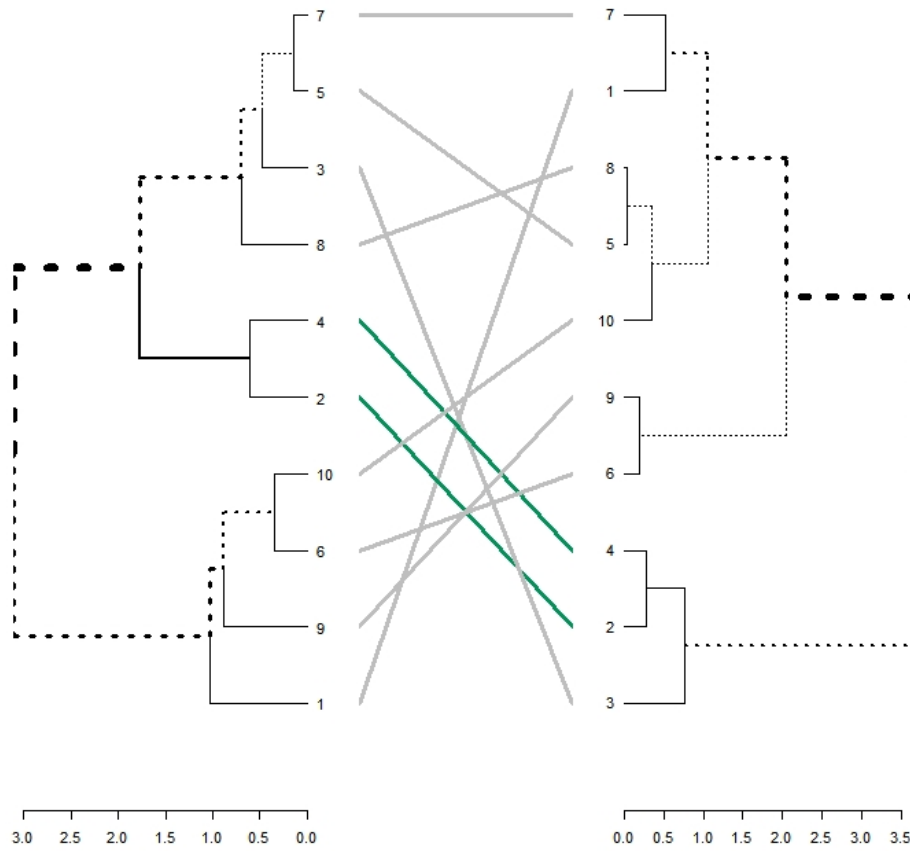


Figure 8.12: A tanglegram relating dendrograms for September 2002 (left) and September 2003 (right). Relatively few hierarchical relationships are conserved between these two example dendrograms (green) and the cophenetic-distance based correlation index is low (0.20).

of the capacity of the explanatory vectors to constrain variance (presented in Figure 8.13 b). In this case, permutation is employed to produce randomised data structures, for which it can be assumed that there are no significant relationships between the dependent and explanatory variables. The reader is referred to Legendre et al. (2011) for a full description of these methods. Simply, three statistical tests are presented here; ‘marginal’, ‘term-wise’ and ‘pure’ effects and all of these compute the capacity of individual explanatory variables

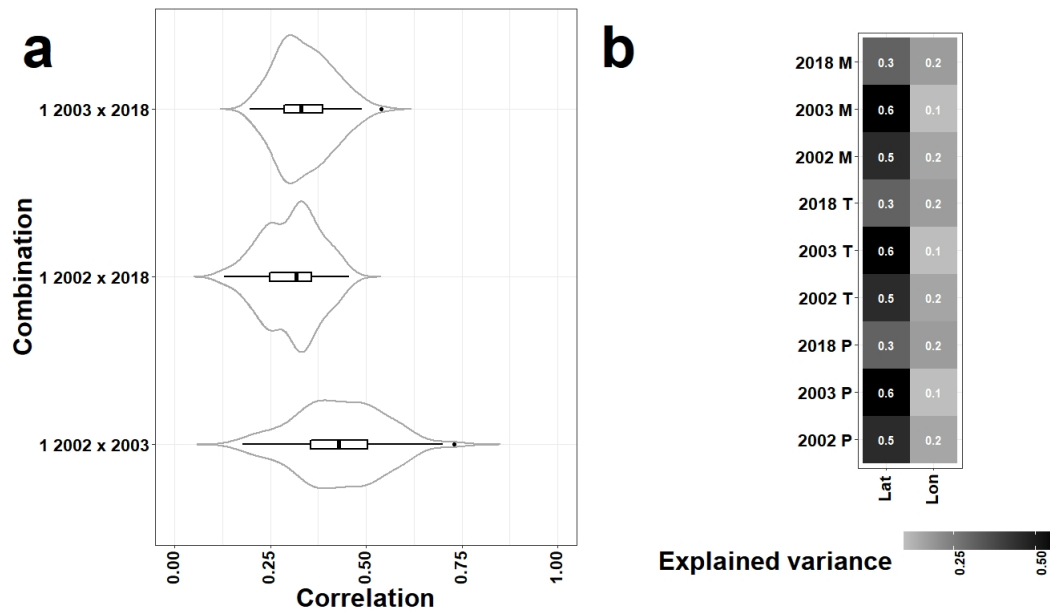


Figure 8.13: **a:** A violin plot of correlation indices relating different years in the dataset. **b:** Variance explained by either latitude or longitude, as calculated by using permutation-based marginal (M), term-wise (T) and pure-effects (P) tests. Only indices with p -values below 0.05 are shown.

to constrain variance, compared with randomly permuted data structures.

Term-wise calculation of explained variance explores the contributions of each of i explanatory variables sequentially. The explained variance for the i^{th} explanatory variable can then be divided by the residuum of unexplained variance, to arrive at an estimate known as the ‘pseudo-F-statistic’. A range of corresponding F-statistics can also be computed for the permuted data structure, in which the rows of the dependent data matrix Y are scrambled. Thereafter the researcher can simply ask what the proportion of permuted analyses were that equal or exceed the pseudo-F-statistic for the non-permuted data structure, to arrive at a facsimile of the p -value – the probability that a test statistic is equalled or exceeded by chance. The process for the calculation of marginal effects is analogous, with the exception that all i explanatory variables are included, so that the only portion of additional variance

explained by the i^{th} variable can be considered its ‘marginal’ effect. Pure effects, by contrast, are computed by first performing analysis with all variables other than i , and then exploring the capacity of i to explain structure in the matrix of residual variation. This therefore explores the capacity for the i^{th} variable to explain variance in Y that could not be explained by any combination of the remaining variables.

In our simple example in Figure 8.13 b, we know that latitude and longitude are strictly orthogonal to each other, and that geographic patterns in SST and storminess, which they appear to describe well, can essentially be considered independent of one another (Figure 8.10). Hence we are not surprised that the term-wise, marginal and pure effect indices all result in significant test results for both latitude and longitude, and we are not surprised that the explained variance is similar across all analyses within each year.

Our examples have focussed on structures of Y for which data are continuously distributed and described well by linear relationships to the explanatory variables in X . Redundancy Analysis should not be used if there is no good reason to expect such ‘clinal’ patterns of variation, or if variance in Y is dominated by outliers that could bias the fitting of linear relationships. Highly non-normal distributions of the data in Y in multivariate space, or non-linear relationships to X , might both be ameliorated by subjecting the data in Y or X to transforms. For example, a matrix of cell-abundances might possess large outliers, but the aberration of these outliers will appear comparatively reduced if the dataset is log-transformed.

8.5 4 Linear Discriminant Analysis

Let us consider a situation in which we do not wish to explore variance in a dataset to discover salient structure, but where we already know that salient structure exists and that it can be approximated as classes. Common examples of this might be that we have some raft of diagnostic medical procedures, and we know that they can be summarised as describing

patients as either well or diseased. Numerous other scenarios might be considered in the biological sciences, but the most notable that have been discussed in this thesis concern the identification of distinct functional assemblages of phytoplankton from some matrix of data that may describe their pigment compliments, enumerations of cellular abundances, or the reflected light from the ocean surface measured by some remote-sensing platform.

We are, in these scenarios, not interested in simply exploring the variance structure of a dataset X , but in considering the variance in X that explains the difference between classes in Y . These two properties may differ significantly. Consider, for example, a dataset in which we wish to distinguish red and white wines by a collection of their properties, including sugar content, tannin concentration and age. The greatest portion of variance between the wines may actually be in their ages, but this has zero relationship to their colour – which would be better proxied by tannin inventory.

How then should we proceed to investigate X ?

The explanation that follows is abridged from the ‘canonical correlations analysis’ chapter of Legendre and Legendre (2012).

Let us reflect on the fact that we are interested in maximising the relative differences between the properties of classes, compared to the range of properties within then. We therefore wish to maximise among-group variance and minimise within-group variance. For any set of groups of Y we can find the corresponding rows in X and compute matrices of the dispersion of the class properties as follows:

$$W_g = [X_g - \text{mean}(X_g)]'[X_g - \text{mean}(X_g)], \quad (8.10)$$

where W_g is the sum-of-squares dispersion of the variables in X for samples n that belong to class g known from Y .

Thereafter we can simply state that the entire within-group variance for all g must be the sum for each individual within-group dispersion;

$$W = \sum_{g=1}^g W_g. \quad (8.11)$$

Computation of the total sum-of-squares dispersion of X is trivial;

$$T = [X - \text{mean}(X)][X - \text{mean}(X)], \quad (8.12)$$

and it is intuitive that the difference of T and W provides the variance between groups (among-group) that we wish to maximise;

$$B = T - W. \quad (8.13)$$

We can then construct the eigenvalue decomposition:

$$(W^{-1}B - l_k I)u_k = 0, \quad (8.14)$$

where I is the identity matrix (diagonals of 1, off-diagonals of 0), l_k is a matrix of eigenvalues and u_k is a rectangular matrix of eigenvectors that describes the $g - 1$ axes discriminating the g groups of Y .

This decomposition can be thought of as retrieving a matrix of ‘F-statistics’, comparing among to within-group variance, to describe the multivariate differences between classes g .

As $W^{-1}B$ is rectangular, the eigenvectors of u_k may not be orthogonal, which should make intuitive sense; the orthogonality constraint made sense in PCA when we were simply exploring different modes of variance, but we have no grounds to expect the axes that discriminate different classes in X to be distributed in a mutually orthogonal fashion.

It is clear that the decomposition is retrieving a set of new principal axes that separate groups g by linear combinations of their properties in X . From this point axes may be re-scaled for plotting, and discriminant functions can be produced to distinguish different

classes. This information may be used either to quantify the structure of difference between classes, or classify new data – for example by computing the posterior probabilities that a new row of X belongs to different classes of Y .

Let us work through a simple example.

The codes required to perform this example analysis are available through:

<https://github.com/aorkney/thesis>

First, it is necessary to download the dataset

`SST_and_storms.csv`

This is a dataset I produced by downloading SST data from the

`ARCTIC_REANALYSIS_PHYS_002_003`

dataset, and wind-speed from the

`WIND_GLO_WIND_L3_REP_OBSERVATIONS_012_005`

dataset. Both of these datasets are available on

<https://marine.copernicus.eu/>

I averaged data for September, over a series of different years, within 1 degree latitude-longitude boxes, so that the dataset is small and manageable. The csv file contains columns for latitude, longitude, year, SST, and storminess. Storminess is computed as the average number of days, within a grid cell, that wind speeds exceed 10 metres per second.

The following analysis can be performed by running the code

LDA_script.R

We previously saw that there was a gradient of SST variation with latitude. Let us, then, say we are interested in knowing whether a randomly chosen grid-cell from our dataset belongs to either the Northern ($> 76^\circ\text{N}$) or Southern Barents Sea. We should expect that SST should allow us to discriminate these two regions (Figure 8.14).

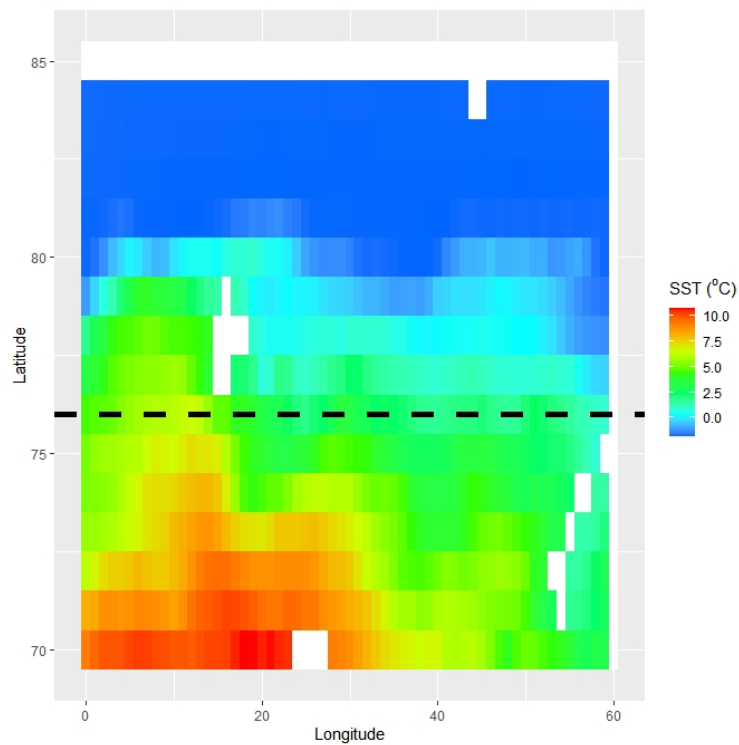


Figure 8.14: Map of SST in the Barents Sea in September 2002, with a dashed horizontal line at 76°N .

Let us construct a data matrix X , composed of SST, storminess and longitude. We know SST should be predictive of latitude, and that longitude should not be. We may suspect that storminess is not particularly well organised over latitude and therefore that it may or may not be a useful classifier of latitude.

We can use a permutation-based procedure to ascertain the capacity of variation in X to predict the desired classes. Let us perform LDA on a subset of 100 randomly chosen samples (with replacement) of X , and then use the resultant Linear Discriminant to predict the classes of the remaining grid-cells. We can perform this process repeatedly, to generate a density distribution of the predictions. We can then review how many of the predictions were correct and incorrect for each class. The means of these density distributions can be presented simply in a Table known as a ‘confusion matrix’ (Figure 8.15). In such a Table, the ‘actual’ classes can be plotted along 1 axis (here I select the y-axis), and the predicted classes across the second. The pairwise combinations then represent the proportion of correct or incorrect guesses for each class.

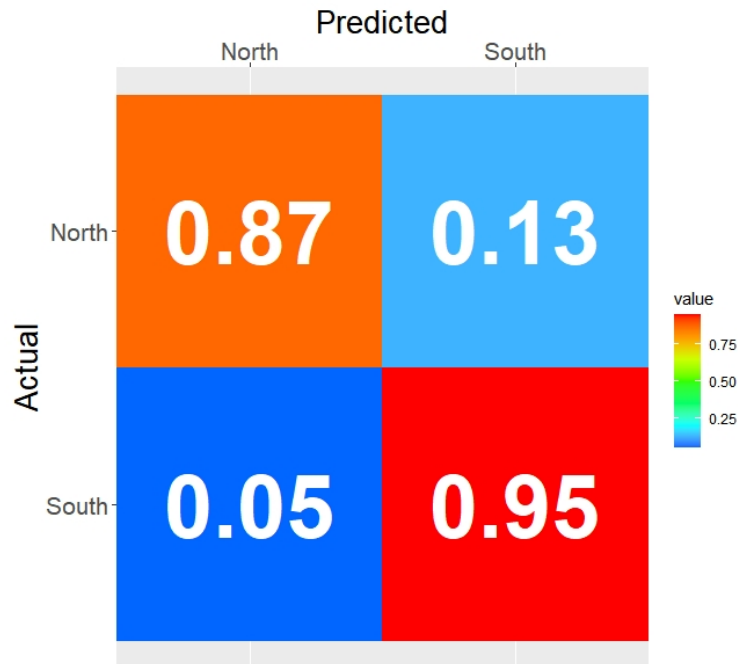


Figure 8.15: A confusion matrix, representing our capacity to predict whether grid cells originate from $>76^{\circ}\text{N}$.

It can be seen, from Figure 8.15, that there is a very good capacity to predict the

latitude of the grid-cells. It is also evident that we had more success predicting membership to the Southern Class. Perhaps this is because SST is more variable in the Northern than the Southern class; a review of Figure 8.14 shows us that the region west of Spitsbergen possesses SST values above 5°C , which is more typical of Southern waters.

We can visualise any individual attempt to predict the class membership of the grid-cell subsets, by plotting the posterior probabilities of the prediction (Figure 8.16)

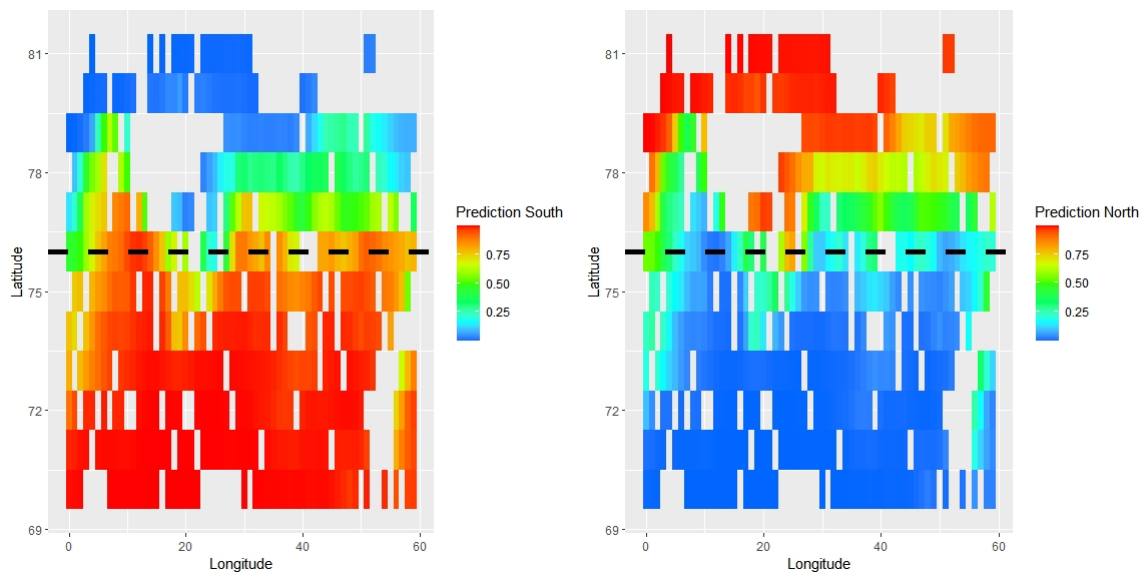


Figure 8.16: Posterior probabilities estimating class membership of a subsample of the original dataset.

This shows us that the region west of Spitsbergen is indeed driving a higher failure to recognise the Northern class.

Let us now inspect the coefficients of the variables in X that have been found to be useful in distinguishing Northern and Southern classes. The density distributions of these properties from the permutation-based procedure are presented in Figure 8.17:

The high magnitude of SST is not a surprise. Nor is the near-zero magnitude of stormi-

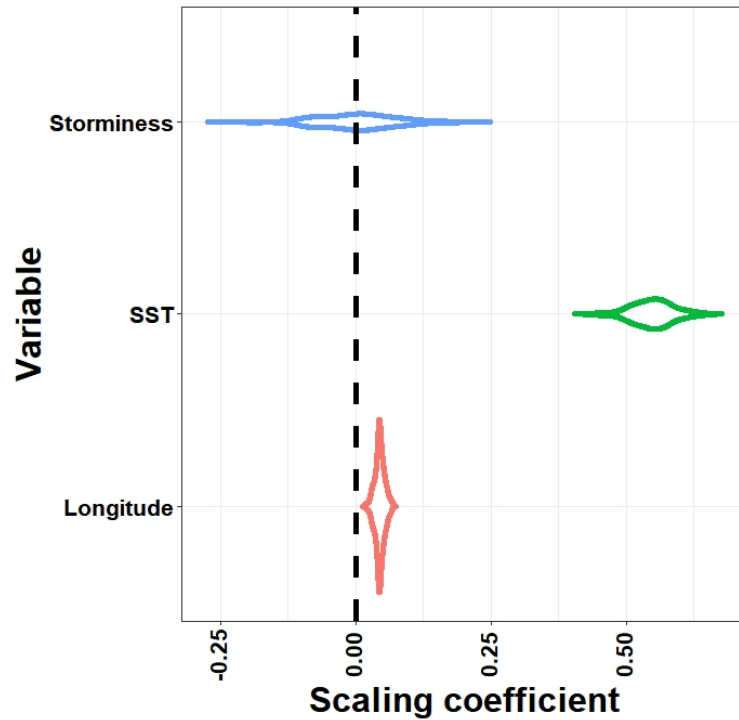


Figure 8.17: Density distributions of the coefficients of the P variables in X that make up the Linear Discriminant for separating Northern and Southern classes. A dashed line marks a coefficient of zero magnitude.

ness; the distribution of storms was not very well organised over latitude and it therefore does not confer great predictive power for membership to Northern or Southern classes. However, the non-zero magnitude for longitude is intriguing, given that we know this property is agnostic to latitude.

Inspection of Figure 8.14 provides us with some understanding; SST varies as a function of longitude as well as latitude. Therefore, while longitude has no relationship to latitude by itself, the combination of longitude and SST possess a stronger relationship to latitude than SST alone. Linear regressions can be applied to confirm this, but an alternative way of illustrating this is to re-run the permutation-based LDA with longitude excluded from X (Figure 8.18).

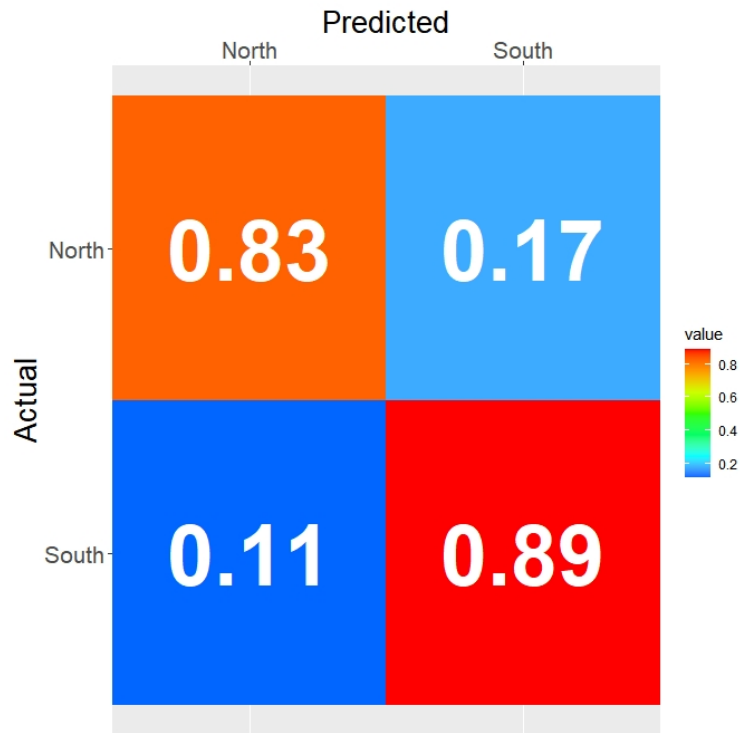


Figure 8.18: A confusion matrix, representing our capacity to predict whether grid cells originate from $>76^{\circ}\text{N}$. when longitude is excluded as a predictive factor.

The diagonal values of the confusion matrix in Figure 8.18 are lower than those in Figure 8.15, demonstrating that longitude was a useful part of the discriminant function.

This discussion of LDA has shown us that multivariate approaches to discriminate classes possess the capacity to retrieve relationships that may, at first, not be apparent to the researcher. The expression of discriminant functions as linear combinations of explanatory variables furthermore makes LDA an appealing mathematical method because the researcher can investigate the causes of structure in the dataset. This can motivate further hypothesis generation. Researchers employing LDA must consider, however, that this is a method that assumes a continuously-distributed dataset of X , and that some re-scaling

choices to find discriminant functions may assume that the distribution of clusters can be approximated as multi-variate Gaussians with equal dispersions.

8.6 5 Fuzzy Clusters

In the previous section, we saw that it was possible to identify salient data structure distinguishing classes (or ‘clusters’) of data in a multivariate context, and that we were able to adapt these methods to predict the identify of new data.

LDA required us, however, to have a priori knowledge of the clusters we wished to study. There are many scenarios we may encounter for which we do not possess knowledge a priori of the structure of clusters in a dataset. For example, we may have some dataset $a_{ph}(\lambda)$ of measurements of the absorption spectra of N phytoplankton samples. We may even have corresponding observations for some of these samples, which lead us to suspect that the dataset can be partitioned between several different prevailing communities of phytoplankton. Perhaps we know there is a diatom-rich community and a dinoflagellate-rich community, for example, but we do not know the cluster identity of most of the samples and we are not confident that we have a large enough number of known samples to assign them by a means such as LDA.

We therefore need a means of discovering clustered structure in the dataset.

For our example, of phytoplankton $a_{ph}(\lambda)$, we might also reason that $a_{ph}(\lambda)$ is an additive function of a multivariate set of constituent phytoplankton in the sample. Therefore, it is possible that some samples might be best described as a linear mixture of different clusters.

Fuzzy logic clustering can achieve both the goals we desire to meet here, by discovering clusters in a dataset and assigning individual samples a vector of partial weights to those clusters. Let us state that all samples n have membership to some number of clusters, c , that varies between 0 and 1. We must also necessarily state that, since we wish to express memberships as linear mixtures, that the sum of all memberships for any n must equal 1.

Let us follow an abridged version of the description of fuzzy c-means clustering found in Moore et al. (2001):

$$J_m = \sum_{g=1}^c \sum_{n=1}^N (u_{ng})^m |X_n - v_g|^2, \quad (8.15)$$

where u_{nk} is the membership of the n^{th} observation to the g^{th} cluster, $|X_n - v_g|$ is the Euclidean distance between the vectors x_n and v_g , m is some weighting exponent greater than 1 (a measure of ‘fuzziness’ of the memberships), and c is our total number of clusters, and N the total number of observations. v_g represents guesses for cluster means.

J_m represents a function to be minimised. It is apparent from inspection of this equation, that the function computes the distances of the samples n (described by X) from a guessed cluster centre. The memberships of those samples to the cluster can be varied, and this process is performed for the full number of clusters.

Therefore, minimising this function is achieved by exploring variety in u_{nk} and v_g , and the minimum will describe a scenario in which samples’ networks of memberships to clusters is optimised so that the weight of memberships is tightly aggregated around each cluster centre. We can imagine the guessed centres migrating through a space, while sample memberships vary, until this solution is found, where the partitioning of memberships among all clusters has the lowest ‘cost’ (because it is intuitive that a sample which is very distance from a cluster centre, but which has been attributed a high membership to it, is not being well-described).

Let us work through a simple example.

The codes required to perform this example analysis are available through:

<https://github.com/aorkney/thesis>

Phytoplankton absorption spectra can be obtained from the British Oceanographic Data Centre BODC:

[doi:10.5285/97daa7ea-8792-6cff-e053-6c86abc0dd46] (JR16006)

[doi:10.5285/982b6da2-7e11-060a-e053-6c86abc09389] (HH180423)

[doi:10.5285/982b6da2-7e12-060a-e053-6c86abc09389] (JR17006)

You will need to also download the csv file

`microscope_alone.csv`

from the github repository. This dataset contains sufficient metadata to be stitched with the data from the BODC. I have decided not to upload the fully compiled stitched dataset to the github repository, because these are datasets which I was obliged to upload to the BODC and researcher traffic should be directed towards their repository.

The following analysis can be performed by running the code

`Fuzzy_clusters_script3.R`

Let us begin by taking $a_{ph}(\lambda)$ between 400 and 700 nm and normalising to its integral. PCA can then be performed to display the variance structure, and fuzzy clustering can be performed on the data. The results of both are presented in Figure 8.19. I reasoned that there are 3 phytoplankton assemblages that were recognisable from ship-board microscopy, (diatoms, *Phaeocystis*, dinoflagellates), and an additional group that corresponded to a lack of any observed microphytoplankton (and therefore probably represented a nanophytoplankton community). I therefore solved for fuzzy clusters under the conditions $c = 4$ and $m = 2$ (following Moore et al. (2001)).

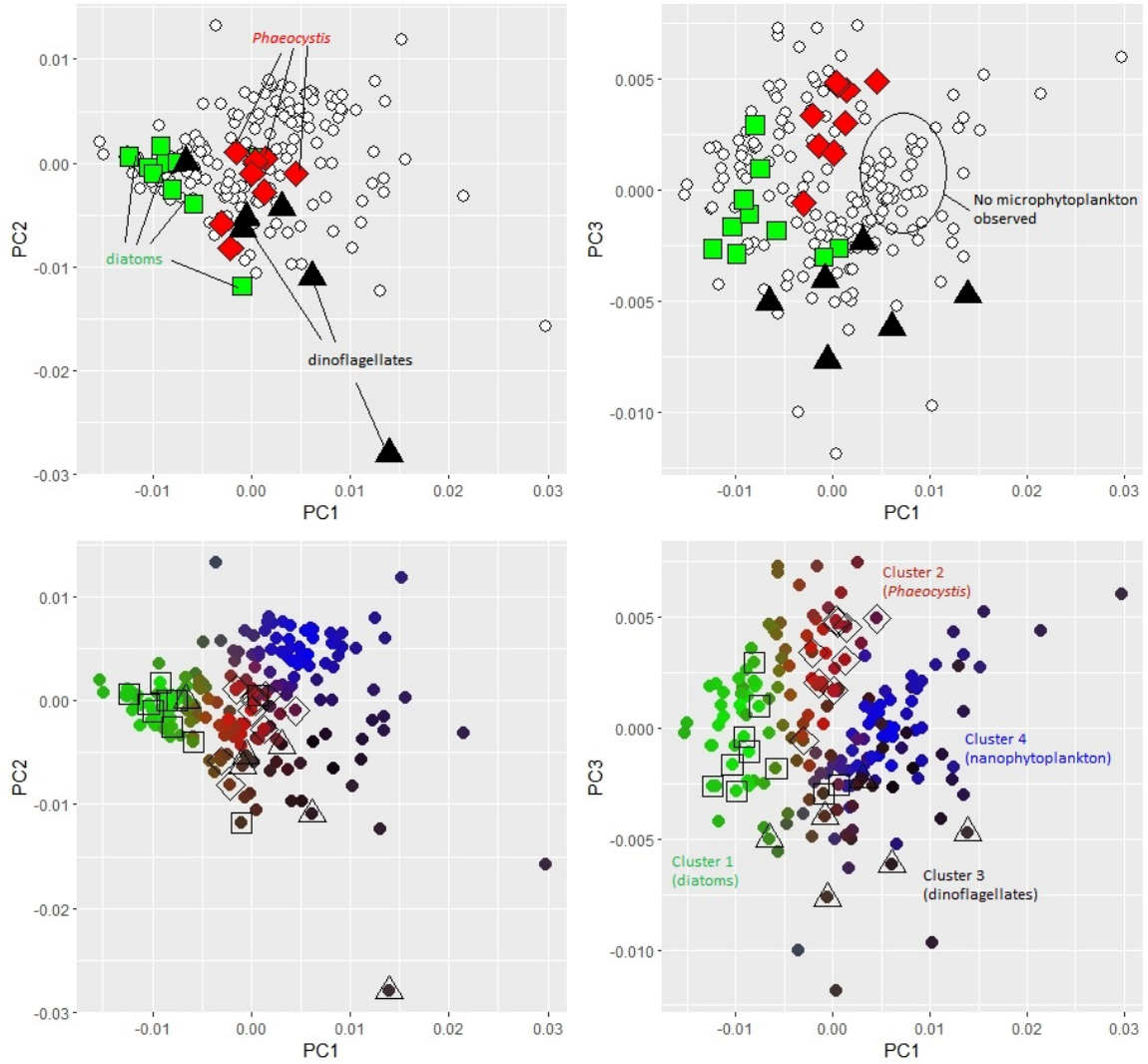


Figure 8.19: **Top:** PC1-2-3 space with clear clustering of ship-board microscopy observations. **Bottom:** Fuzzy clusters with red, green, blue and black values indicating memberships. The pattern of clustering corresponds well with the microscopy observations.

Hence, fuzzy clustering has allowed the retrieval of a data structure in the Barents Sea $a_{ph}(\lambda)$ dataset that was only partially evident from ship-board microscopy. This was a supervised clustering technique and it was necessary to pay close attention to the correspondence between expected and retrieved structure. This required subjective decision-making

on the part of the researcher.

The mean $a_{ph}(\lambda)$ for each cluster, weighted by fuzzy memberships, is presented in Figure 8.20

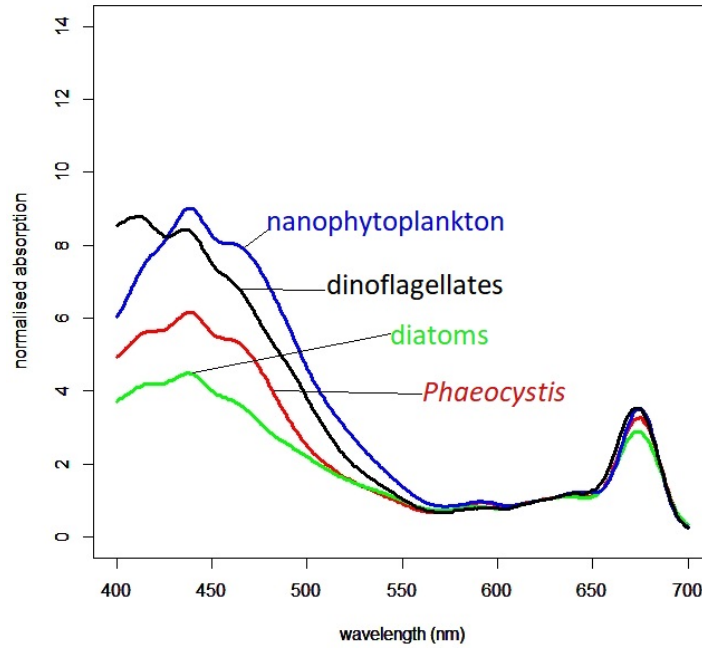


Figure 8.20: Mean spectra for the different clusters identified in Figure 8.19 (normalised to absorption at 621 nm).

Now let us consider a cautionary example of the potential pitfalls of clustering methods. Let us say that we did not normalise the $a_{ph}(\lambda)$ values to their integrals. This means that variation in the overall magnitude of absorption will dominate the variance in $a_{ph}(\lambda)$. We can see that, if we attempt to solve for clusters in this dataset, that we simply partition variance along PC1 (which is almost certainly an axis representing overall magnitude) and that this clustering has little relations to the phytoplankton assemblages we were aiming to characterise (Figure 8.21).

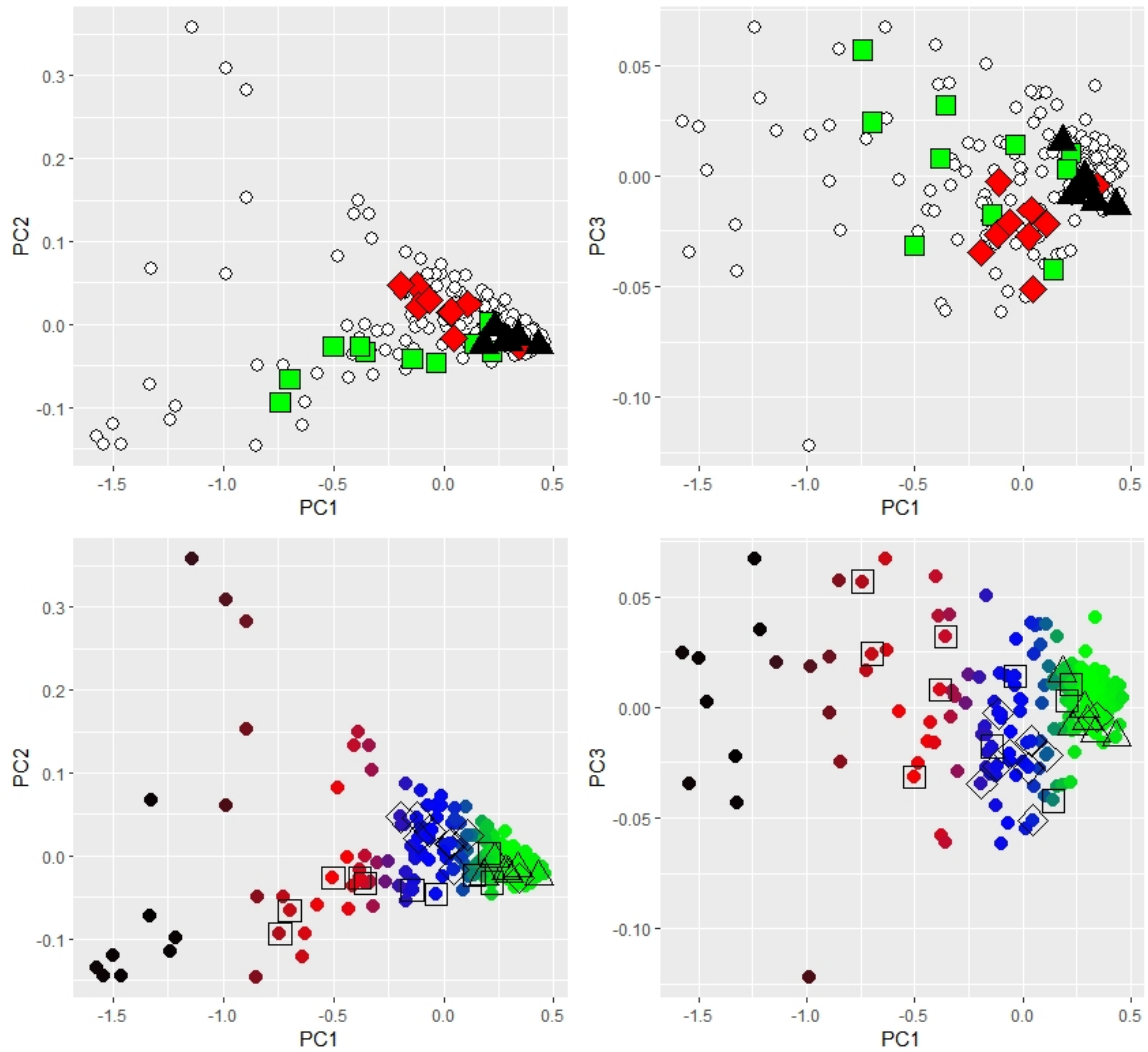


Figure 8.21: Top: PC1-2-3 space with clustering of ship-board microscopy observations (consult Figure 8.19 for identities). **Bottom:** Fuzzy clusters with red, green, blue and black values indicating memberships. The pattern of clustering does not correspond well with the microscopy observations.

We have therefore gained an appreciation of the utility of clustering analysis, but also some limitations to commonly-employed clustering approaches. Many clustering approaches assume that individual clusters have Gaussian distributions across the major axes of vari-

ation, for example. It is also routine for clustering approaches to assume that each cluster is populated by a similar number of samples. Therefore, should sampling originate from a population with extremely uneven diversity, or possess a non-Gaussian variance structure, researchers may struggle resolving clusters. Ultimately, clustering analysis is often an art form that requires explicit supervision from the researcher and visualisation of data structure. Researchers should not be misled into the belief that the resolution of clusters with an algorithm means that those clusters are ‘scientifically’ supported or meaningful for downstream analyses.

8.7 6 Temporal Autocorrelation

Temporal autocorrelation describes the scenario in which a system at time t is itself a function of the state of the system at some earlier time $t - t_{lag}$, where ‘lag’ is some time-scale over which autocorrelation is occurring. Time series describing natural systems often exhibit autocorrelation, as a result of periodicity in natural cycles (e.g. seasonality), or feedback loops. For example, in the Barents Sea, there is a multi-annual feedback loop between the velocity and pathways of inflowing currents and the wind-field at the ocean-surface (Smedsrud et al. 2013).

Temporal autocorrelation concerns us, because it is assumed in ‘least squares’ type regressions, that errors (the difference between fitted and actual values in the dependent variable) are uncorrelated. Strong positive feedback loops can result in the errors of a fitted function themselves possessing strong autocorrelation. For example, imagine a series of daily cycles of air temperature through the summer. While the overall temperature grows hotter through the summer, the instantaneous temperature follows a daily cycle; peaking in the day and cooling at night. Therefore, given a temperature at night on one day, we can predict that the temperature the following noon will be higher. It is intuitive, then, that our n measurements in such a dataset are not independent of one another.

So, why does this pose a problem? Structure in the residuals can result in flawed assessments of significance for longer-term trends.

In least squares linear regression, where we are seeking to solve an equation of the form

$$y = mx + c, \quad (8.16)$$

one of our objectives might be to test whether the residual sum of squares of the fit (the difference between the fit and the actual measurements) is significantly smaller than the residual sum of squares of an invariant model ($y = 0x + c$). We can compare the residual sum of squares error (RSS) for models 1 (invariant) and 2, by computing the F-statistic:

$$F = \frac{\frac{RSS_1 - RSS_2}{p_2 - p_1}}{\frac{RSS_2}{n - p_2}}, \quad (8.17)$$

Where RSS_1 is the residual sum of squares of the invariant model, RSS_2 is that of the full model, p represents the number of model parameters (just 1 for the invariant model, and 2 for the full model), and n is a set of statistically independent samples.

if the full model achieves smaller residuals than the invariant model, then the difference between RSS_1 and RSS_2 grows, and the size of the denominator decreases, which means that F grows. Hence larger F indicates greater favour for the full model over the invariant model.

For a scenario, where y exhibits autocorrelation over x , and the n samples are not independent; the ‘effective’ n is smaller. This means that, if we have failed to notice autocorrelation, we have overestimated n . A larger n decreases the size of the denominator, which inflates F , making it appear as though the full model is minimising residual error more effectively than it genuinely is. Hence, this will result in an over-estimate of the significance of the relationship between y and x .

Let us illustrate this with a synthetic example:

The codes required to perform this example analysis are available through:

<https://github.com/aorkney/thesis>

The first code to run is:

```
Autocorrelation.R
```

In this example, I have produced two time series of imaginary temperature over time; one that is correlated with itself (y) at the previous time-step (lag time of 1) and one without autocorrelation (x).

$$x = \frac{t}{20} + \epsilon, \quad (8.18)$$

where ϵ is random variation, and

$$y = (x + x_{lag=1}) - mean(x), \quad (8.19)$$

Random error superimposed on the data make it difficult to spot the autocorrelation from visual inspection. However, visual assessment of the residuals from fitted models makes autocorrelation clear: Figure 8.22.

The fits of generalised least squares regressions indicate the p -value for the slope of the model fit to the autocorrelated data is far lower (0.000025) than the fit the standard data (0.0019), which might be attributable to the greater slope of the model, but could also be influenced by a reduced number of independent samples.

The details of fitting a generalised least squares regression with an autocorrelated structure are involved (extensive discussion is found in the stochastic model building chapter of Box et al. (1994)). Suffice to say that a correlation structure representing autocorrelation can be incorporated into a generalised least squares regression.

When autocorrelation exists over a lag time of 1 time-step, as is the case here, this can be achieved using the ‘*corAR1()*’ function of the *nlme* R package. This was undertaken

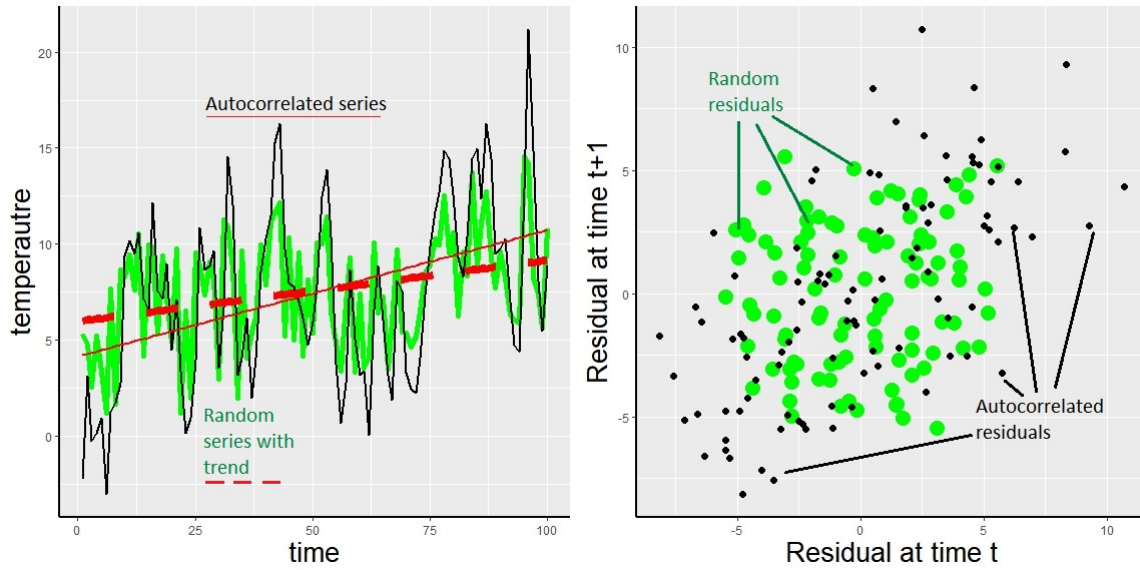


Figure 8.22: **Left:** x (green series, dashed fit) and y (black series, solid fit) It is not immediately clear whether either series exhibits autocorrelation. **Right:** Fit residuals and lagged fit residuals for x (green) and y (black). Residuals for x form a bivariate Gaussian, whereas residuals for y are distributed as an ellipsoid, indicating autocorrelation.

for both x and y . These regressions, adjusted for temporal autocorrelation, did not differ considerably in their intercepts or slopes, compared to regressions assuming no temporal autocorrelation. However, their p -values are much larger. The autocorrelated data is now associated with a value of 0.036, while the data free of autocorrelation is associated with a value of 0.0052 (note I have applied an autocorrelation correction to x here even though we know this is not an appropriate choice).

This indicates that the original assessment of significance, using generalised least squares regressions that did not incorporate temporal autocorrelation, provided an inflated estimate of the significance of the relationship of the autocorrelated data to time. We can also see that wrongly applying a correction for autocorrelation results in an underestimate of significance (higher p -value for x results).

We therefore need a means of deciding whether it is appropriate to incorporate auto-

correlation into a regression. We could visually inspect residuals and lagged residuals of standard regressions, as in Figure 8.22, and could even automate this by computing correlations and deciding a threshold value. A perhaps better alternative would be to compute standard and autocorrelated regressions, and then compare the regression fits to see which is a superior description of the series.

The `anova()` function of R can be used to compare regression models, rewarding smaller residuals and penalising greater numbers of model parameters (an intuitive decision vis-à-vis Occam's razor). The application of this function to generalised least squares regressions of x and y , with and without autocorrelation represented, demonstrates that a standard regression is preferred for x ($p = 0.24$ for preference of autocorrelated model over standard), while an autocorrelated model is favoured for y ($p < 0.001$). Therefore, we can see that it is possible to contrive a code that takes time series input, without knowledge of temporal autocorrelation, and tests whether an autocorrelated fit with a lag time of 1 time step is preferred.

Such a code is presented in

```
Autocorrelation2.R
```

and applied to assess relationships between SST, storminess and open water area in the Barents Sea in September 2002-2019. These series are presented in Figure 8.23:

Open water area was defined as the total number of observed lat-lon ocean-colour measurements in a given September (see chapter 5). I will explore potential geographic relationships between these time series. Generalised least squares models will be fitted to each lat-lon grid-cell, with and without lag-1 temporal autocorrelation represented, and the preferred model will be presented. The relationship between SST and storminess is presented in Figure 8.24

The results of Figure 8.24 are consistent with what we know; in the Northern Barents

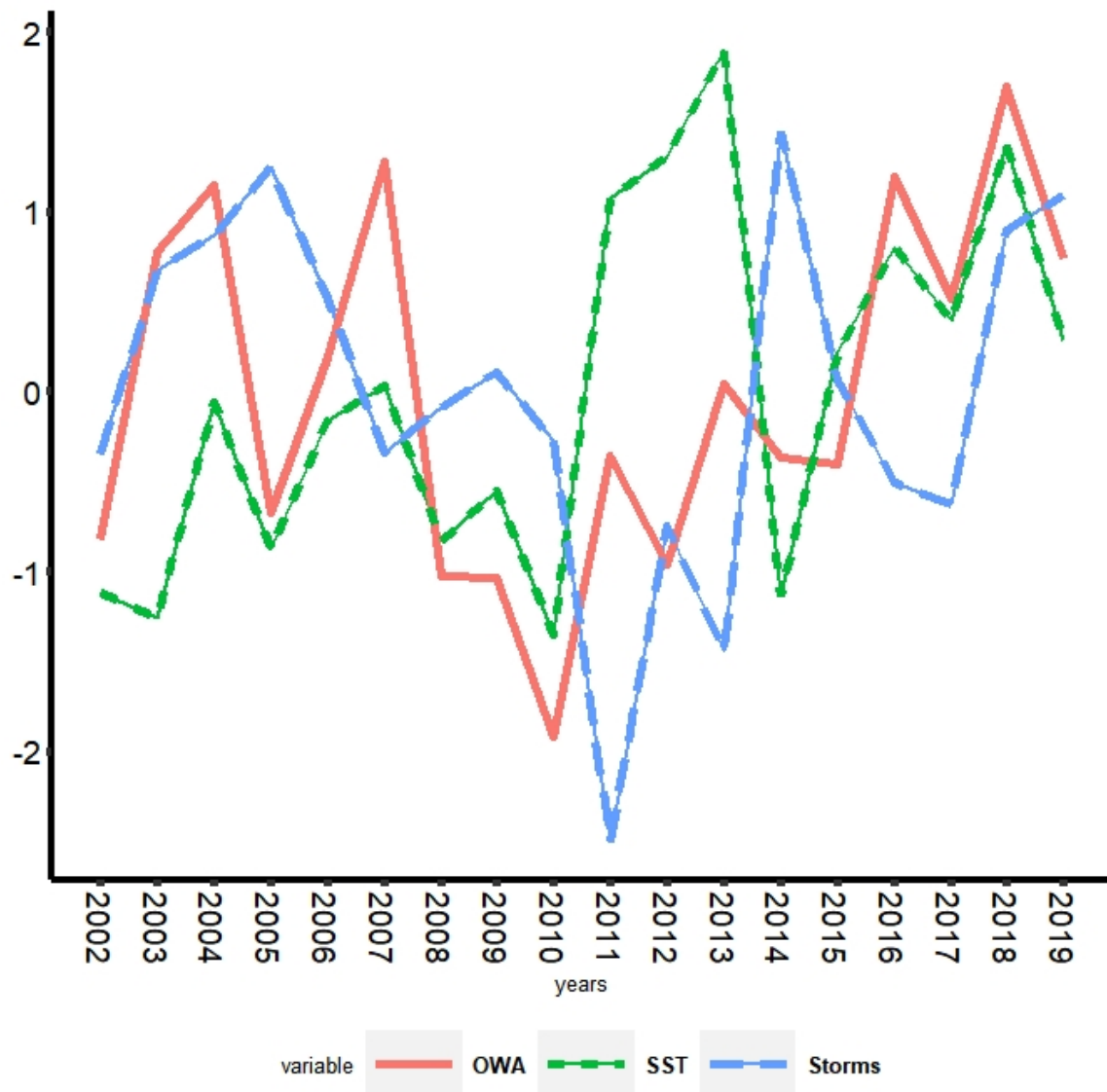


Figure 8.23: Barents-wide variance-normalised evolution of September open water area (OWA), Sea Surface Temperature (SST) and storms (winds above 10 m^{-1}).

Sea, Atlantic waters are overlain by a lid of Polar waters, therefore mixing by winds increases SST. In the Southern Barents Sea, the warmest waters are at the top of the water column, so mixing causes cooling of SST. The relationships uncovered indicate that models

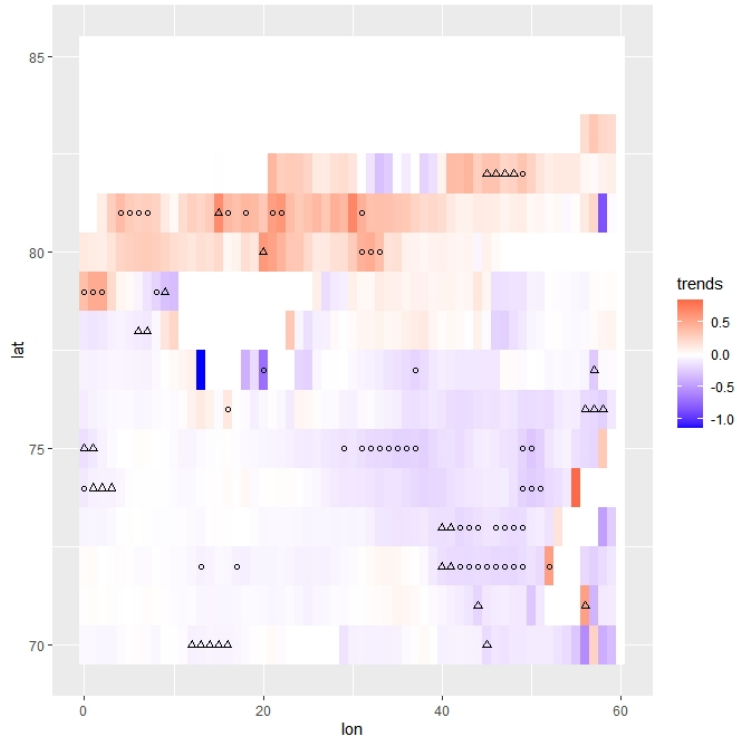


Figure 8.24: Circles indicate a standard fit is preferred and has $p < 0.05$. Triangles indicate an autocorrelated fit is preferred and has $p < 0.05$. We can see that SST is modified by wind in different ways in different regions of the Barents Sea.

incorporating autocorrelation tend not to be preferred.

The relationship between SST and open water area is presented in Figure 8.25:

Figure 8.25 appears to show that a significant number of grid-cells prefer autocorrelated models; SST and open water area in the South-east Barents Sea in particular, appear to be associated. A mechanical explanation for this relationship is not as obvious as that between SST and storms, but this observation could guide further investigation.

Finally, the relationship between storms and open water area is presented in Figure 8.26:

Even though Figure 8.23 appears to show evidence of a lagged relationship between storms and open water area between the years 2010 and 2013, there is a lack of support for autocorrelated models. The observation of a strong relationship of storms to open water

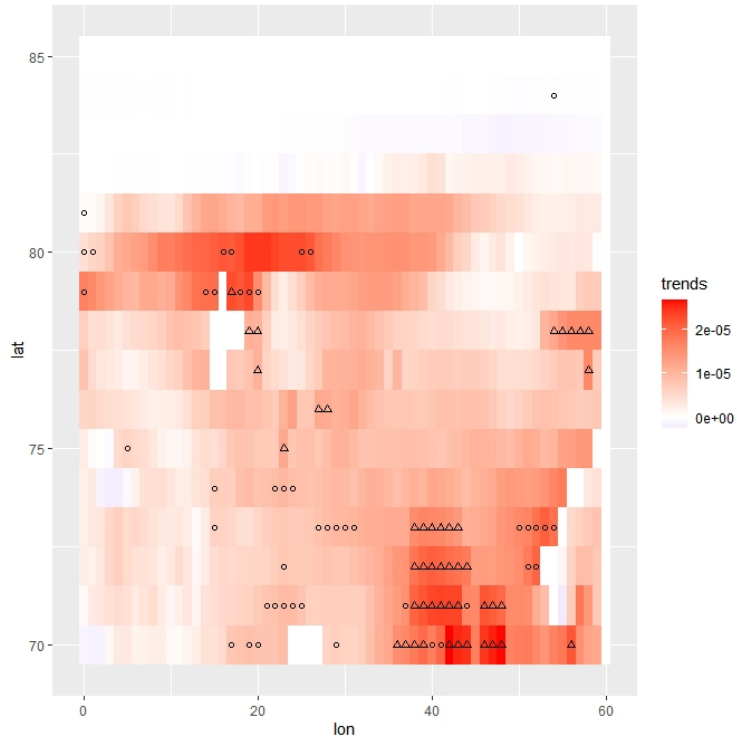


Figure 8.25: Circles indicate a standard fit is preferred and has $p < 0.05$. Triangles indicate an autocorrelated fit is preferred and has $p < 0.05$. It looks like there is a relationship between September SST and open water area.

area in the South-west could be consistent with greater delivery of Atlantic-derived air-masses or winds from the South-west. Such warm air-masses or south-westerly winds could reduce sea-ice extents by encouraging melt or by forcing the ice to migrate to the North-east.

Let us investigate the relationship between wind and the open water area the previous year; we have reason to believe this should be significant (Smedsrud et al. 2013). The results are presented in Figure 8.27:

Figure 8.27 shows the expected pattern; the open water area appears to predict the storminess in the subsequent autumn over much of the Barents Sea, consistent with a positive ‘wind-loop’ feedback. We must therefore reflect cautiously that significant lagged

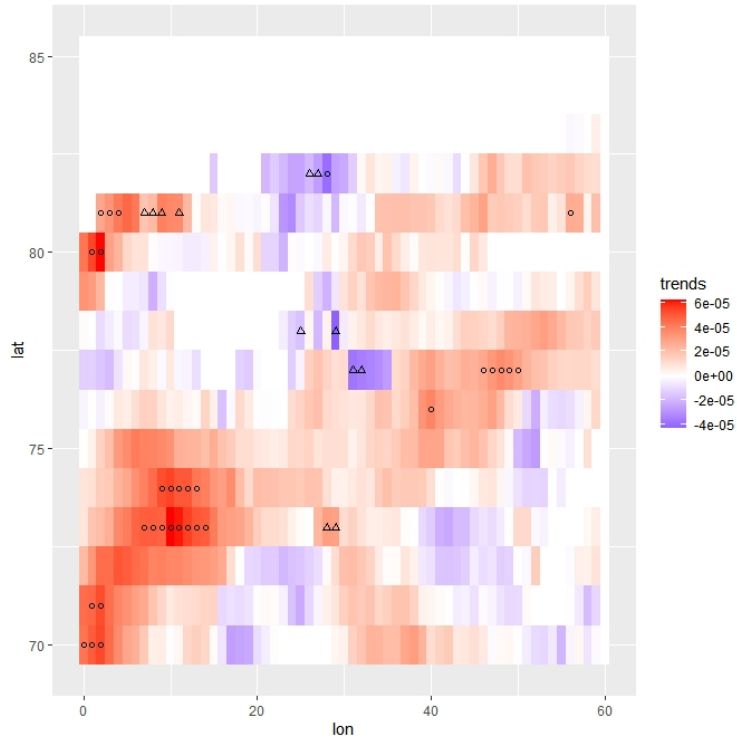


Figure 8.26: Circles indicate a standard fit is preferred and has $p < 0.05$. Triangles indicate an autocorrelated fit is preferred and has $p < 0.05$. Years with greater open water area tend to be associated with strong winds in the South-western Barents Sea.

relationships can exist in datasets, but that these may not result in autocorrelated models being preferred (Figure 8.26). Incorporating an autocorrelation structure of the investigated variables to time is not a test for lagged relationships between variables.

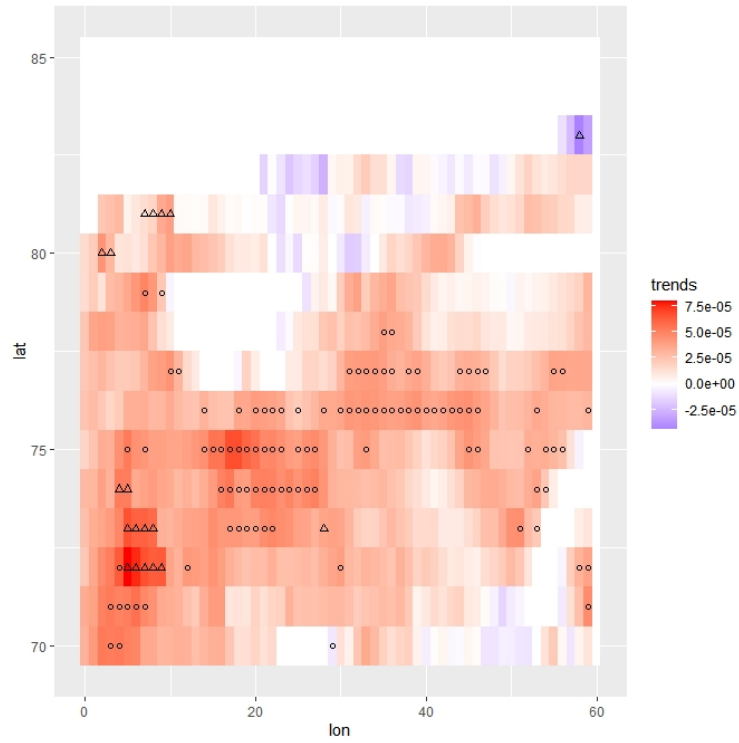


Figure 8.27: Circles indicate a standard fit is preferred and has $p < 0.05$. Triangles indicate an autocorrelated fit is preferred and has $p < 0.05$. Years with greater open water area tend to precede years with more frequent autumn storms.

References

- Box, George EP, Gwilym M Jenkins, and Gregory C Reinsel. 1994. *Time series analysis, forecasting and control*. Englewood Cliffs.
- Galili, Tal. 2015. “dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering.” *Bioinformatics*, <https://doi.org/10.1093/bioinformatics/btv428>. eprint: <https://academic.oup.com/bioinformatics/article-pdf/31/22/3718/17122682/btv428.pdf>. <https://academic.oup.com/bioinformatics/article/31/22/3718/240978/dendextend-an-R-package-for-visualizing-adjusting>.
- Hansen, Kathryn, Jeff Schmaltz, and Joshua Stevens. 2016. “The Barents Sea Abloom.” Accessed: 2020-03-03, <https://earthobservatory.nasa.gov/images/88316/the-barents-sea-abloom>.
- Legendre, Pierre, and Loic FJ Legendre. 2012. *Numerical ecology*. Elsevier.
- Legendre, Pierre, Jari Oksanen, and Cajo JF ter Braak. 2011. “Testing the significance of canonical axes in redundancy analysis.” *Methods in Ecology and Evolution* 2 (3): 269–277.
- Moore, Timothy S, Janet W Campbell, and Hui Feng. 2001. “A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms.” *IEEE Transactions on Geoscience and Remote sensing* 39 (8): 1764–1776.
- O’Reilly, John E, and P Jeremy Werdell. 2019. “Chlorophyll algorithms for ocean color sensors-OC4, OC5 & OC6.” *Remote sensing of environment* 229:32–47.
- Rohlf, F James, and Marco Corti. 2000. “Use of two-block partial least-squares to study covariation in shape.” *Systematic Biology* 49 (4): 740–753.
- Smedsrud, Lars H, Igor Esau, Randi B Ingvaldsen, Tor Eldevik, Peter M Haugan, Camille Li, Vidar S Lien, Are Olsen, Abdirahman M Omar, Odd H Otterå, et al. 2013. “The role of the Barents Sea in the Arctic climate system.” *Reviews of Geophysics* 51 (3): 415–449.
- Venables, WN, and BD Ripley. 2002. “Modern Applied Statistics, Springer.” *New York* 5.