UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIROMENTAL SCIENCES

School of Ocean and Earth Science

Vertical mixing and interannual variability of primary production in the North Atlantic

by

Andrew Gravelle

Thesis for the degree of Doctor of Philosophy

September 2016

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIROMENTAL SCIENCES

Ocean and Earth Sciences

Thesis for the degree of Doctor of Philosophy

Vertical mixing and the interannual variability of primary production in the North Atlantic

Andrew David Charles Gravelle

It is widely held that as the ocean becomes more intensely stratified with anthropogenic-driven climate change, marine primary productivity (PP) will decline within mid-to-low latitude nutrient-limited waters, and increase within higher latitude light-limited waters. This is consistent with projections from Earth-system models, which predict a decline in global PP over the next century for 'business-as-usual' and high-mitigation warming scenarios. However, interannual and longer-term relationships between stratification and PP are more ambiguous in observational studies. Underlying the projected changes in PP are assumptions as to the response of phytoplankton to changes in stratification and vertical mixing at these time scales. This thesis focuses upon the identification and analysis of interannual relationships between phytoplankton biomass and vertical mixing in the North Atlantic.

Satellite-derived chlorophyll (Chl) data are divided into regions of similar variability in order to assess the spatial dependence of interannual relationships with vertical mixing. A large-scale bimodal pattern represents the Chl response to the tri-pole pattern of climatic variability associated with the NAO in the North Atlantic. This Chl pattern is related to similar patterns in satellite-derived sea surface temperature (SST), wind speed, and Argo float-derived mixed layer depth (MLD) and stratification. Relationships with these proxies for vertical mixing are found to be spatially heterogeneous. However, it is general to this analysis that relationships are also spatial-scale dependent: localised variability may dominate the local-scale, but tends to cancel-out within regions of similar Chl response to reveal larger-scale relationships that dominate overall variability. Thus, while observational data tend to be noisier at the local-scale, they are consistent with Earth-system models in revealing an overall dependence of phytoplankton upon vertical mixing at larger scales.

These large-scale patterns and relationships are in good agreement with output from a hindcast biogeochemical model (NEMO MEDUSA), which is analysed to determine how indirect-relationships with vertical mixing are mediated. Interannual Chl variability is shown to depend upon nutrient-availability throughout the mid-to-low latitude North Atlantic; relationships within

the subpolar North Atlantic are undetermined, presumably due to a seasonal-dependence of relationships that are poorly represented by yearly-averaged/yearly-metric time series at these latitudes.

Lastly, the bimodal pattern of Chl variability is assessed for stability and continuity of relationships against climate warming. In the projected output of NEMO MEDUSA, under a relative concentration pathway (RCP) 8.5 'business-as-usual' warming scenario, this bimodal pattern in Chl variability is shown to weaken over the next century. While relationships with vertical mixing appear to continue unabated, the dominance of this bimodal pattern upon Chl, SST and MLD variability in the North Atlantic appears to decline over this time.

Table of Contents

List of Tables

List of Figures

List of Supplementary Figures and Tables

DECLARATION OF AUTHORSHIP

I, ANDREW GRAVELLE, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Vertical mixing and the interannual variability of primary production in the North Atlantic

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. [Delete as appropriate] None of this work has been published before submission [or] Parts of this work have been published as: [please list references below]:

Signed:.....

Date:.....

Acknowledgements

Firstly, I would like to thank my supervisors for their fantastic input and support throughout my PhD. Adrian, Stephanie, Claudie and Katya, your help and guidance has been invaluable and very much appreciated. I would also like to thank my panel chair, Martin Solan, for his input and guidance during our meetings.

I would like to thank everyone who has made studying here over the last four years such a friendly and enjoyable experience. A special thanks goes to my office mates, especially long-term office mates Victoria and Anna, for your friendship and all the uplifting tea breaks. Tim, thanks for being a great friend and the early morning badminton! Lucy, Mike and Gert-jan, thanks for all the great times going rock climbing! Sarah and Rob, thanks for being great housemates.

I would like to thank my family – especially for supporting me over the last few months. I could not have done this without you.

I dedicate this work to
my little niece – Abi.

Definitions and Abbreviations

AMSR-E	Advanced Microwave Scanning Radiometer - Earth observing system
AVHRR	Advanced Very High Resolution Radiometer
BATS	Bermuda Atlantic Time-series Study
CCI OC	Climate Change Initiative, Ocean Colour
Chl	Chlorophyll-a
CMIP3	Phase 3 of the Coupled Model Intercomparsion Project
CMIP5	Phase 5 of the Coupled Model Intercomparsion Project
CORA	COriolis ReAnalysis
CPR	Continuous Plankton Recorder
CTD	Conductivity Temperature and Depth
DCM	Deep Chlorophyll Maximum
ENSO	El Niño-Southern Oscillation
Eq.	Equation
ESA	European Space Agency
Fig.	Figure
GSW	Gibbs Seawater Routine
HadGEM2-ES	Hadley Centre Global Environment Model Version 2 – Earth System
K	Number of clusters for K-means clustering
MEDUSA	Model of Ecosystem Dynamics, nutrient Utilisation, Sequestration and Acidification
MEI	Multivariate ENSO Index
MERIS	MEdium-spectral Resolution Imaging Spectrometer
MLD	Mixed Layer Depth
MODIS	MODerate-resolution Imaging Spectroradiometer
N. Atl.	North Atlantic
NAO	North Atlantic Oscillation
NCEP	National Centre for Environmental Prediction
NEMO	Nucleus for European Modelling of the Ocean
NOAA	National Oceanic and Atmospheric Administration
NPP	Net Primary Production
PAR	Photosynthetically Active Radiation
PAR_{ML}	Photosynthetically Active Radiation averaged over MLD
PCI	Phytoplankton Colour Index
PP	Primary Production
QuickSCAT	Quick Scatterometer

R^2	Coefficient of determination
RCP	Relative Concentration Pathway
RI	Relative Importance
S	Supplementary figure or table
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SIC	Schwarz Information Criterion
SSB	Sum of Squares Between
SSMIs	Special Sensor Microwave Imager / Sounder
SST	Sea Surface Temperature
SST	Sum of Squares Total
SSW	Sum of Squares Within
TEOS-10	Thermodynamic Equation of Seawater - 2010
TMI	Tropical rainfall measuring mission's Microwave Imager
VRC	Variance Ratio Criterion
XBT	eXpendable BathyThermograph
XCTD	eXpendable Conductivity Temperature and Depth

Chapter 1: Introduction

1.1 Background

The importance of phytoplankton

Photosynthetic marine primary production (hereafter, PP) forms the base of almost the entire marine ecosystem and is fundamental in shaping and maintaining biogeochemical cycles of carbon, oxygen and nutrients within the ocean (e.g. Gruber and Sarmiento, 2002). About half of all global net primary production (NPP), which is defined as the photosynthetically-fixed carbon made available to higher trophic levels, is driven by marine phytoplankton (Field, 1998). As well as representing the primary food source of the ocean, phytoplankton are substantially responsible for balancing the global oxygen cycle that higher trophic animals, including ourselves, require to live.

Given rising atmospheric CO_2 concentrations, the key role of PP in the global carbon cycle is of particular importance at present. While much of the carbon fixed by phytoplankton will be consumed by herbivores or decomposed and regenerated within surface waters, a fraction of this carbon will be exported to the interior ocean or even to the sediments (Duarte and Cebrián, 1996). As the interior ocean and sediments represent long-term reservoirs of carbon, with residence times spanning thousands of years (interior ocean) or geological time scales (sediments), the ocean itself acts as a net sink of carbon. On average, the global mean export of PP to the interior ocean is about 15% of the total, with export to sediments being a small part of this (Laws et al., 2000). Thus, although phytoplankton only represent about 1-2% of the global plant stock of carbon (Falkowski, 1994), their fast turnover rate and the large surface area of the ocean means that they drive a large export flux of carbon from the surface ocean, and equilibrated atmosphere, into the interior ocean or sediments. This biological carbon pump (Broecker, 1982; Sarmiento and Bender, 1994; Volk and Hoffert, 1985) plays a significant role in transferring carbon out the atmosphere and into the longer-term oceanic reservoir of carbon (Gruber and Sarmiento, 2002; Post et al., 1990; Siegenthaler and Sarmiento, 1993) and maintains atmospheric CO_2 concentrations at about 150-200ppmv less than would otherwise be the case (Falkowski et al., 2000).

Vertical mixing and PP

It has long been recognised, especially since satellite ocean colour measurements have provided a synoptic view of the distribution of chlorophyll (Chl) in the ocean, that the observed climatological distribution of phytoplankton biomass can be explained in terms of nutrient availability (including macro- and micronutrients), light, grazing and stratification (Falkowski, 1994; Ryther, 1969). Thus, phytoplankton populations are driven by a number of processes including both bottom-up and top-down controls. However, to first order, the abundance of phytoplankton in the ocean is associated with nutrient concentrations (Ryther, 1969). For an overview of phytoplankton nutrient-limitation within the ocean see Moore et al. (2013). Thus, in the mid-to-low latitude open ocean (< 45° N/S), where permanent stratification inhibits the vertical mixing of nutrients up from depth, climatological phytoplankton biomass tends to remain low. This is especially true of the subtropical gyres, wherein a confluence of surface water causes a sharper density divide with deeper waters, further inhibiting the upwards mixing of nutrients. By contrast, in the higher-latitude temperate and subpolar regions of the open ocean, wintertime convective mixing acts to replenish surface nutrients, allowing for a phytoplankton bloom at the onset of stratification in spring, and higher yearly-averaged values of phytoplankton biomass. There are a number of exceptions to this generalisation, including regions of coastal and equatorial upwelling, high-nutrient low-chlorophyll regions and regions of sea ice cover. However, in general, vertical mixing plays a key role in determining climatological nutrient concentrations and phytoplankton biomass within the open ocean.

Vertical mixing also plays a key role in the variability of phytoplankton biomass. As previously alluded to, vertical mixing can drive seasonality in phytoplankton. This is more pronounced at higher latitudes, where stratification is completely broken down during winter months. However, even within the permanently-stratified mid-to-low latitude ocean, seasonal variability in surface heat flux and wind speed can drive changes in surface temperature and the strength of stratification, which in turn can result in seasonal variability of phytoplankton biomass. The shape and amplitude of this seasonal cycle is latitudinally dependent (González Taboada and Anadón, 2014).

However, while vertical mixing is instrumental in shaping the mean distribution of phytoplankton biomass within the ocean, and in driving seasonal variability, its role in driving interannual and longer-term variability in phytoplankton biomass are less certain.

Vertical mixing and interannual Chl variability

A number of studies have revealed relationships between Chl variability and physical processes related to vertical mixing in time series after seasonality has been removed (e.g. Behrenfeld et al., 2006; Kahru et al., 2010; Siegel et al., 2013). The standard approach to deseasonalising a time series is to subtract the climatological monthly (weekly, etc.) values of the time series from the corresponding monthly (weekly, etc.) values within the time series:

$$x'_{ij} = x_{ij} - \bar{x}_i \quad (1.1)$$

Where x is a given time series of which i denotes months (weeks, etc.) and j denotes years. $\bar{x}_i$ is the mean of all x_i and x' is the deseasonalised time series. This gives a time series of anomaly values that represent deviations from a fixed seasonal cycle and mean value. Such relationships should thus indicate if vertical mixing drives non-seasonal variability in Chl. Furthermore, as seasonality is ubiquitous in environmental time series, adjusting for it is required in order to test against casual associations.

Monthly anomalous Chl variability is significantly related to SST (Siegel et al., 2013), wind speed (Kahru et al., 2010) and stratification (Behrenfeld et al., 2006) within the permanently-stratified mid-to-low latitude ocean (< ~45° N/S). The predominantly positive relationships with wind speed, and inverse relationships with SST and stratification within this region, are consistent with vertical mixing driving growth via associated nutrient fluxes within these permanently-stratified waters. Furthermore, a generally negative relationship between wind speed and SST throughout the entire ocean, implies that wind speed is responsible for driving vertical mixing within this region and elsewhere (i.e. SST is a proxy for vertical mixing as temperature tends to decrease with depth within the upper ocean; Kahru et al. 2010).

At higher latitudes, anomalous Chl variability is negatively (north of ~45° N) and positively (south of ~45°S) related to SST when averaged over these regions (Siegel et al., 2013). As SST is a proxy for vertical mixing, the sign of these relationships may imply that phytoplankton growth is, on average, nutrient-limited and light-limited in the respective regions. However, these overall relationships between Chl and SST are less strong at these latitudes than in the mid-to-low latitude ocean, and when relationships are assessed at the local scale, in 1° cells, they are less frequently significant and patchier (i.e. including areas of both positive and negative sign) at these latitudes than in the mid-to-low latitude ocean (Siegel et al., 2013). This is also apparent of

relationships between Chl and wind speed at higher latitudes. In general, Chl tends to be negatively related to wind speed at latitudes polewards of 45° N/S (Kahru et al., 2010). This is consistent with phytoplankton growth being negatively related to vertical mixing and thus potentially light-limited at higher latitudes. However, the sign of relationship between Chl and wind speed is patchier at these latitudes then in the mid-to-low latitude ocean, with frequent areas of positive relationship, especially within the higher latitude southern hemisphere (> ~45°S; Kahru et al. 2010).

These studies thus imply that Chl variability in the mid-to-low latitude ocean is positively related to vertical mixing (Kahru et al., 2010; Siegel et al., 2013) and negatively related to stratification (Behrenfeld et al., 2006). Changes in stratification and Chl variability throughout this domain have also been shown to be strongly related to the multivariate El-Nino Southern Oscillation index (MEI; Behrenfeld et al., 2006).

However, when the relationship between Chl and stratification in the mid-to-low latitude ocean is considered at the local scale, in 5° cells, only 11% of this area is found to have a significant relationship between the two with $R^2 > 0.1$ (Dave, 2014; Dave and Lozier, 2013). Furthermore, using 21 years of monthly in-situ data from the long-term Hawaii Ocean Time-series (HOT) in the tropical Pacific, and the Bermuda Atlantic Time Series (BATS) in the subtropical North Atlantic, Dave and Lozier (2010) and Lozier et al. (2011) find no significant correlation between stratification and either PP or Chl measurements at the interannual scale, nor do they find any consistent association between productivity and either the ENSO or PDO indices.

Barton et al. (2015) compared 48 years of phytoplankton colour index (PCI; a proxy for phytoplankton biomass) from the Continuous Planktonic Recorder (CPR) survey in the North Atlantic to various physical processes relating to stratification and vertical mixing (including wind speed, heat flux, turbulent kinetic energy, MLD, stratification and SST). They showed that while these variables are strongly related to seasonality in Chl throughout the North Atlantic, including a general change in the sign of relationship between sub-tropical and sub-polar waters, there are very few significant relationships between them once seasonality and a long term trend have been removed. From this, they argue that these proxies for stratification and vertical mixing drive seasonality and long-term trends in Chl, but are not related to interannual variability in Chl.

A similar result is found by Feng et al. (2015), who show that once seasonality and auto-correlation are accounted for, SST, surface insolation and wind speed collectively account for only ~20% variance in Chl when averaged over all biogeographical provinces (see: Longhurst, 2007) comprising the temperate, subtropical, tropical and equatorial ocean. Again, this implies that proxies for vertical mixing (SST, surface insolation and wind speed) are not good descriptors of interannual Chl variability.

Thus, there are contrasting results as to the importance of stratification and vertical mixing in driving interannual Chl variability. Although these studies use different datasets and proxies for phytoplankton biomass and physical variables, the results are nonetheless largely contradictory, with findings that stratification and vertical mixing either explain the majority (e.g. Behrenfeld et al., 2006; Siegel et al., 2013), or very little (e.g. Barton et al., 2015; Dave and Lozier, 2013) of interannual variability in phytoplankton biomass.

A possible limitation of these studies, however, may be their reliance upon subtracting, or normalising against (Kahru, et al., 2010), climatological monthly values to deseasonalise time series (e.g: Eq. 1.1). This does not allow for variations in the timing, shape and amplitude of the seasonal cycle (Vantrepotte and Melin, 2009). Given that there is generally greater variance in Chl time series at the seasonal scale than the interannual scale (e.g. Vantrepotte and Melin, 2011), any seasonal variance left after subtracting a fixed climatology may overpower a subtler interannual or decadal signal. Furthermore, if such residual seasonality remains in anomalous time series, this can result in spurious relationships between variables when correlated. This is because seasonality is ubiquitous in environmental time series and a coincidence of seasonality does not imply a causal relationship. Given that there tends to be considerable variability in the seasonality of Chl time series (e.g. Vantrepotte and Melin, 2009), correlation of residual seasonality may present a continual problem when subtracting fixed seasonal cycles.

Another contrasting factor between these studies is in the spatial scale at which they assess relationships. For example, Dave and Lozier (2013) find a similarly strong relationship to that of Behrenfeld et al. (2006) when assessing the relationship between Chl and stratification at the same large scale, despite finding that this relationship is only representative of 11% of the mid-to-low latitude ocean when assessed at the local scale.

Increased stratification with warming

Global warming has the potential to affect PP in a number of ways (e.g. Doney et al., 2011; Gruber, 2011). However, of importance to this study is the potential impact of increasing stratification upon phytoplankton. As it preferentially warms the surface ocean, global warming acts to increase stratification. This has the potential to both limit nutrient supply and alleviate light limitation in the surface ocean. Thus, it is predicted that in lower-latitude oligotrophic regions of the ocean, where primary production (PP) is predominantly nutrient-limited, increasing stratification will result in reduced PP, while in the high latitude regions of the ocean where PP is predominantly light-limited, increasing stratification and shallower mixing will result in increased PP (Doney, 2006). Such changes have been predicted in various global ocean models (Bopp et al., 2001; Cabré et al., 2015; Sarmiento et al., 2004; Steinacher et al., 2010), or analytically using the equations that express phytoplankton growth in these models (Marinov et al., 2010).

However, such changes are not as clearly defined in observational data. For example, these projections assume a strong association between stratification (or vertical mixing) and Chl variability at interannual and longer-term time scales, which are not necessarily apparent in observational data (see §1.1 *vertical mixing and PP*). Additionally, a number of studies have looked for trends in satellite Chl records (e.g. Gregg, 2005; Irwin and Oliver, 2009; Polovina et al., 2008; Vantrepotte and Mélin, 2011). However, such records remain too short to distinguish secular trends, as expected of global warming, from amidst stronger natural variability at these time scales (Henson et al., 2010). In order to overcome this, other studies have compared satellite Chl records from different missions to extend time series, however opposing trends emerge depending upon how the datasets are reprocessed (Antoine, 2005; Gregg and Conkright, 2002). Still other studies have used long-term in situ datasets of ocean colour (e.g. Boyce et al., 2014, 2010; Edwards et al., 2001; Falkowski and Wilson, 1992; McQuatters-Gollop et al., 2011; Wernand et al., 2013), or PP (Bidigare et al., 2009; Saba et al., 2010), but these tend to suffer from uncertainties in the merging and reprocessing of datasets (e.g. Mackas, 2011; Rykaczewski and Dunne, 2011), or represent regional or point samples that under-sample the ocean.

Thus, observational datasets relating to phytoplankton biomass tend to be too short, or too localised, to assess global trends in PP with confidence. As forecast trends in PP cannot be verified in this way, an alternative is to assess the relationships between stratification (or vertical mixing) and PP at the interannual time scale for consistency with the forecast trend.

1.2 Aim of this thesis

This thesis aims to assess the relationship between vertical mixing and the interannual variability in phytoplankton biomass. Specifically, this thesis aims to determine if the strength of relationship between the two is dependent upon spatial scale. It also aims to determine how such relationships are mediated. Most studies assumed that indirect relationships between *e.g.* Chl and SST represent an underlying relationship between e.g. Chl and nutrients, or Chl and light. However, such direct relationships are often assumed and not identified to the exclusion of other possible mechanisms. It is also of interest to determine how any such relationships will change over the coming century with continued global warming. More specifically, this thesis aims to address the following questions:

Questions

- Does the spatial scale of analysis affect the identification of relationships between chlorophyll and proxies for vertical mixing?
- What physical processes are most strongly related to chlorophyll variability and how are these relationships mediated?
- Will relationships between chlorophyll and vertical mixing proxies change over the next century with global warming?

Scope of analysis: additional drivers of Chl variability

In addition to wind-driven and buoyancy-forced changes in vertical mixing, other processes may also influence Chl variability. For example, mesoscale eddies (McGillicuddy et al., 1998), advection (Williams and Follows, 1998), zooplankton consumption (Banse, 1992; Miller et al., 1991) and basin-scale changes in the depth of the main pycnocline (i.e. as opposed to the seasonal pycnocline at the bottom of the mixed layer; Martinez et al., 2009) have all been shown to drive changes in PP.

This analysis focuses upon vertical mixing as a driver of interannual Chl variability and, rather than attempting to account for all potential drivers of Chl variability, will rely upon the presence

(absence) of relationships with proxies for vertical mixing to determine where vertical mixing (or another process) drives interannual Chl variability.

Of course, the influence of various drivers upon PP may change into the future with projected climate change. For example, mesoscale activity and/or advective processes could change regionally in response to meteorological changes, or changes in large scale circulation may result in basin-scale changes to the depth of the pycnocline. However, a number of model projections have already established increased stratification as a major control upon PP into the future (Bopp et al., 2013; Cabré et al., 2015; Steinacher et al., 2010). This thesis does not aim to reassess the influence of increasing stratification upon PP within warming scenario model projections, but to assess observational data for consistency with the underlying premise that stratification and vertical mixing are dominant controls upon Chl at interannual time scales, and, where applicable, to establish the mechanisms underpinning such relationships and to assess model projections for their continued influence into the future.

1.3 Study Domain

In order to assess the large scale relationship between phytoplankton biomass and vertical mixing in the ocean, a large scale spatiotemporal dataset of density profiles is required. Argo is a global deployment of free-drifting profiling floats that collect and transmit temperature and salinity measurements of the upper 2000m of the water column. Argo represents the major source of density profiles for the global ocean and as such is the main contributor to global-scale datasets of observed ocean density (Riser et al., 2016). The Argo floats also operate throughout the year and as such their data is not seasonally biased at higher latitudes by poor weather conditions. At present there is comparatively good coverage of Argo floats throughout the ocean, however the first deployments began in late 1999 and in the earlier years of this program certain parts of the ocean had better coverage than others (Fig. 1.1). The North Atlantic represents part of the ocean having had good coverage since the earlier years of this program and as such represent a region with consistent and coherent coverage throughout the program (Fig. 1.1; Riser et al., 2016).

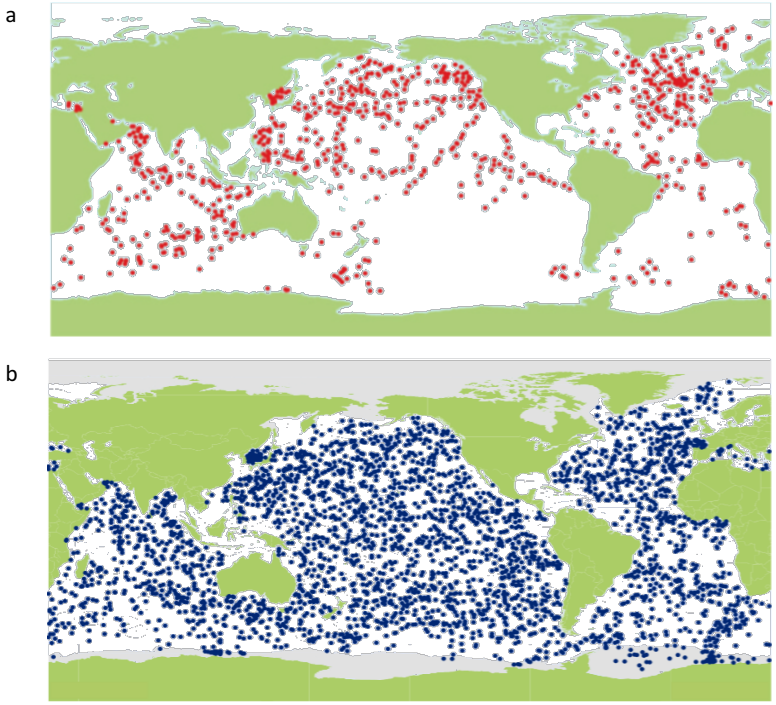


Fig. 1.1: Global distribution of Argo floats in a) April 2003 with 763 floats and b) April 2008 with 3159 floats. From: http://www.argo.ucsd.edu

Given the long-term operation of the CPR program in the North Atlantic (e.g. Reid et al., 2003), the relationship between physical forcing and phytoplankton dynamics have been particularly well studied in this ocean basin (e.g. Barton et al., 2015). As such, the North Atlantic additionally represents a well-studied region in which to build upon, and contrast results with, existing literature. This study will thus focus upon the interannual relationship between phytoplankton variability and vertical mixing in the North Atlantic.

1.4 North Atlantic

In the North Atlantic, as seen elsewhere in the ocean (§1.1 *Vertical mixing and PP*), increased mixing is observed to result in an enhanced phytoplankton bloom within subtropical waters (Follows and Dutkiewicz, 2001), and a reduced bloom within subpolar waters (Henson et al., 2009). This distribution is consistent with processes of nutrient limitation controlling growth in subtropical regions, and light limitation and restratification controlling growth in subpolar regions (Follows and Dutkiewicz, 2001; Henson et al., 2006; Sverdrup, 1953). However, while Barton et al. (2015) also identify opposing positive and negative relationships between vertical mixing and phytoplankton biomass within subtropical and subpolar waters, respectively, they attribute these relationships almost exclusively to seasonal variability and long-term trends. They find only sparse and weak associations between these processes at the interannual time scale. Thus, there is disagreement in the literature as to the importance of vertical mixing in driving interannual phytoplankton growth in the North Atlantic.

The North Atlantic Oscillation (NAO) is the dominant mode of climatic variability in the North Atlantic (Hurrell et al., 2003). As a measure of the pressure difference between the Azores High and Icelandic Low pressure systems, the NAO index is associated with the strength and position of westerly winds and storm tracks in the North Atlantic (Fig. 1.2). As such, this climatological index is related to large-scale changes in vertical mixing and advection within the North Atlantic Ocean (e.g. Hurrell and Dickson, 2004; Visbeck et al., 2003).

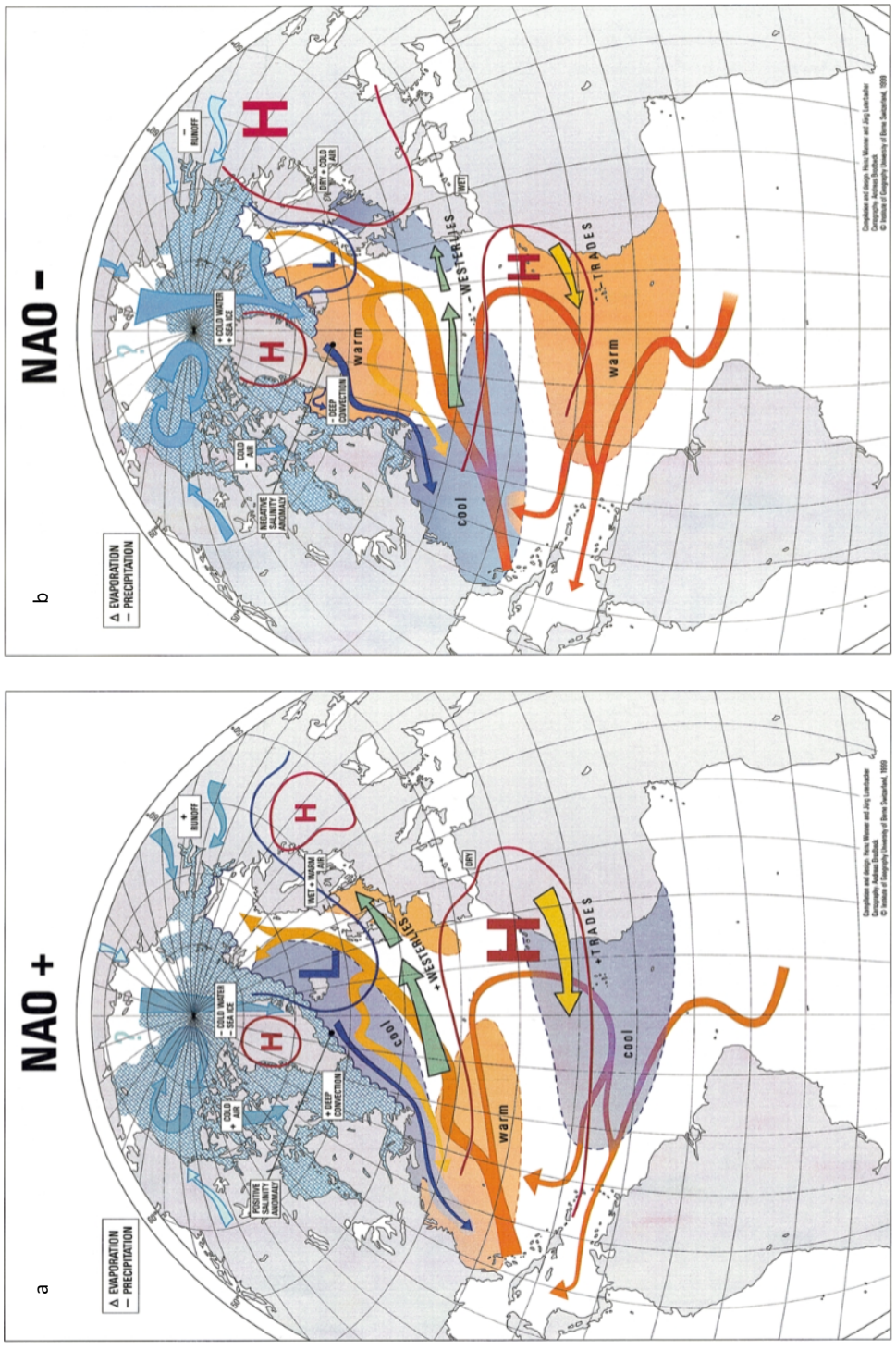


Fig. 1.2: Schematic representation of changes in mean North Atlantic conditions during a) positive and b) negative phases of the NAO index. Orange and blue areas (dashed demarcations) and arrows represent warmer- and cooler-than-usual SSTs and surface currents, respectively. Green and yellow arrows denote changes in the mean position of Westerlies and Trade winds, with size of these arrows representing relative changes in mean wind speeds. Red and blue lines demark changes in mean positions of High (H) and Low (L) atmospheric pressure systems. Light blue area shows relative extent of sea-ice. Light blue arrows represent surface currents and relative river discharge. White boxes detail changes in climatic conditions. From: Wanner et al., (2001).

The NAO has been shown to be associated with a tri-pole pattern of Chl anomalies in the North Atlantic, with a positive (negative) NAO resulting in reduced (enhanced) growth at subtropical latitudes and enhanced (reduced) growth at latitudes both north and south of this (Patara et al., 2011). Using a marine ecosystem model that broadly replicates this observed pattern, these NAO related Chl anomalies were shown to be associated with concurrent changes in vertical mixing throughout the North Atlantic at interannual time scales (Patara et al., 2011)

However, relationships between the NAO and phytoplankton dynamics are less apparent in observational analyses. Interannual variability in the timing of the spring phytoplankton bloom has been observed to be related to the NAO index within both the subpolar (Henson et al., 2009) and central (Zhai et al., 2013) North Atlantic. In both regions, a positive NAO index is associated with the delayed onset of the spring bloom. In the case of Henson et al. (2009), the two were shown to be related via anomalous wind-driven mixing, although Zhai et al. (2013) also found a strong relationship with SST and inferred that associated changes in the southwards advection of fresh Arctic waters may control stratification and the timing of the spring phytoplankton bloom. However, both these studies found only weak associations between the NAO and phytoplankton biomass. Interannual variability in phytoplankton biomass has been related to the NAO in the northeast Atlantic using PCI from the CPR survey in the North Atlantic (Edwards et al., 2001). However, this long term time series of phytoplankton biomass is only weakly related to the NAO in a few positions throughout the the North Atlantic when time series are detrended (Barton et al., 2003). There is thus uncertainty as to what extent inter annual variability in phytoplankton biomass within the North Atlantic can be attributed to changes in vertical mixing and the NAO.

1.5 Variables for vertical mixing

The studies referred to in §1.1 and §1.4 use a variety of variables to estimate vertical mixing and/or stratification in the water column. The most direct measures of mixing commonly used in large scale studies are mixed layer depth (MLD) and stratification. These two variables represent opposing processes of mixing and stability and thus tend to be inversely related. However, the two may be better suited, in term of capturing water column dynamics that affect phytoplankton, at different latitudes and localities. As stratification is referenced to a fixed depth (often 200m; e.g. Behrenfeld et al., 2006), it is likely to be better suited to measuring water column stability in mid-to-low latitude regions of the permanently stratified ocean, where seasonal and interannual changes in mixing depth do not exceed the reference depth. By contrast, MLD may be better

suited to tracking larger changes in mixing depth typical of temperate and high latitude waters, where seasonal variations in wind mixing and heat flux result in seasonal stratification and deep winter mixing/convection. (Nb. In this thesis, the term winter convection is used to denote deep winter mixed layers typical of temperate and higher latitude waters [e.g. de Boyer Montégut 2004], and not the open-ocean deep convection that is unique to regions of deep water formation [e.g. Marshall & Schott 1999]). The depth of winter convection is important as it can determine the abundance of new nutrients in the surface layer at the onset of stratification, and thus the amplitude of the spring phytoplankton bloom (2014, 2010; Martin et al., 2010; Pätsch et al., 2001; Steinhoff et al., 2010).

Proxies for vertical mixing can be divided into processes that promote mixing or stratification (e.g. wind speed and derivatives thereof; surface heat flux) and those that are a corollary of mixing and stratification (e.g. SST). Wind speed is related to wind stress (e.g. Hellerman and Rosenstein, 1983) and acts to increase turbulent mixing in the surface ocean. By contrast, the surface heat flux controls the buoyancy forcing of the surface ocean and thus can promote static stability (stratification) or instability (convection). However, the net heat flux requires a number of parameters to calculate, including radiative fluxes (incoming shortwave; outgoing longwave) and turbulent heat fluxes (both sensible and latent), with these depending upon a number of variables including cloud cover and wind speed. As such, previous observational studies have relied upon reanalysis data to estimate the surface heat flux (e.g. Barton et al., 2015), or downwards shortwave radiation (e.g. Feng et al., 2015). In contrast to these forcings, SST may respond to changes in buoyancy forcing and vertical mixing. That is, SST may correspond to changes in surface heating (and thus reflect changes in stratification), while also corresponding to changes in vertical mixing (as temperature tends to decrease with depth). As such, changes in SST may be correlated with stratification and anti-correlated with vertical mixing.

Given wind velocities and surface heat flux, previous studies have also calculated the turbulent kinetic energy (TKE) of the mixed layer (e.g. Barton et al., 2015; Follows and Dutkiewicz, 2001). TKE is a measure of turbulent mixing, however as these studies estimate TKE from wind velocities and heat flux it may also be considered a proxy for vertical mixing in these cases.

Unlike direct measurements of vertical mixing, changes in these proxies may not necessarily result in changes of mixing. For example, SST can also change with advection or near-surface heating/cooling processes that may have little effect upon vertical mixing deeper in the water

column. Furthermore, given the existing structure of the water column, changes in wind speed or heat flux may only affect near-surface waters and have little impact upon vertical mixing as measured by stratification or MLD. Thus, while each of these variables may be correlated in relation to vertical mixing, they may each emphasise different aspects of mixing that may affect phytoplankton, or may in some instances change independently of vertical mixing (e.g. SST). This study will consider a number of these variables, including both yearly-averages and winter extreme values, in order to rigorously test for relationships with vertical mixing and also to determine which variables are the best estimators of interannual phytoplankton variability.

1.6 Structure of thesis

The structure of this thesis is as follows. Chapter 1 gives a general background to the topic of this thesis and identifies key questions to be addressed. Chapters 2-4 will represent individual studies that aim to address these question sequentially. Chapter 5 will then summarise all results in the context of these questions, give a critique of the data and methodologies used and provide suggestions for possible future work.

Chapter 2: Spatial dependence of chlorophyll forcing in the North Atlantic

2.1 Introduction

2.1.1 Overview

A number of recent studies have presented conflicting results as to the importance of stratification in driving interannual chlorophyll variability in the mid-to-low latitude ocean (e.g. Behrenfeld et al., 2006; Dave, 2014). These contradictory observations appear to largely depend on the spatial scale of analysis. This chapter aims to identify the importance of spatial scale in determining the strength of relationship between chlorophyll (Chl) variability and various proxies for vertical mixing, including stratification, in the North Atlantic.

2.1.2 Background

Photosynthetic marine primary production (PP) forms the base of almost the entire marine ecosystem and is fundamental in shaping and maintaining biogeochemical cycles within the ocean (Gruber and Sarmiento, 2002; e.g. Redfield, 1958). Marine PP represents about half of total global net PP (Field, 1998). As a fraction of this fixed carbon is exported to the ocean interior (~15%; Laws et al. 2000), marine PP helps to lower atmospheric CO_2 concentration by about 200ppmv via the action of this biological carbon pump (Parekh et al., 2006).

With rising atmospheric greenhouse gas concentration, there is increasing interest in how marine PP may respond to climate warming. With warming, increased stratification and reduced vertical mixing rates are expected to hinder PP within the mid-to-low latitude nutrient-limited surface ocean, and enhance PP within the high-latitude light-limited surface ocean (e.g. Doney, 2006). Such trends are realised in forced warming-scenario coupled atmosphere-ocean general circulation models (Bopp et al., 2001; Sarmiento et al., 2004). However, observational studies are more ambiguous. A strong negative correlation is observed between stratification and PP anomalies within the permanently-stratified (i.e. year-round) Tropical Ocean (Behrenfeld et al., 2006). As these permanently-stratified waters were also found to drive the majority of the global signal in marine PP, this may project a global decline in marine PP with warming. However, when a similar analysis is applied at the local scale, considering the permanently-stratified ocean in

discrete 5x5° cells, the relationship between chlorophyll (Chl; a proxy for PP) and stratification is limited to a small fraction of the total area (Dave and Lozier, 2013). This may indicate that changes in stratification have a limited effect upon PP within the majority of the permanently-stratified ocean and that other physical processes such as advection and/or wind mixing may be more important. Thus there are contradictory results as to the importance of stratification in driving PP anomalies, depending upon the spatial scale of the analysis. Assuming both studies are representative of the real ocean, such contrary results may arise for three reasons:

- The few strong localised relationships may dominate the large-scale.
- Local-scale relationships could be obscured by low signal-to-noise ratios, while the overall relationship is revealed at the large-scale due to improved signal-to-noise ratios.
- The changes in stratification and PP that drive the large-scale relationship may not be co-located at the local scale.

Thus, there remains an open question as to the proportion of the permanently-stratified ocean that exhibits a relationship between stratification and PP: does the large-scale signal represent a majority or minority area? Understanding what physical processes drive interannual changes in marine PP is critical in order to accurately model this fundamental part of the ecosystem and predict how it will respond to future climate change. This chapter considers the importance of spatial scale in the analysis of environmental relationships.

2.1.3 Aim of this chapter

This chapter uses an intermediate-scale analysis to assess the spatial extent of relationships between the physics of vertical mixing and stratification, and the response of interannual Chl variability. The intermediate scale is defined by considering regions of the ocean that have similar Chl variability. As this analysis aims to identify the spatial extent of relationships with Chl, the benefit of defining regions of similar response in this way are potentially two-fold: improved signal-to-noise ratios over local-scale analysis, and allows for spatial heterogeneity in relationships and variables that may otherwise be obscured at the large-scale. Thus, by averaging over the full extent of an area that has similar Chl time series, we maximise the potential for improving the signal-to-noise ratio with any variable that may be driving that Chl response. This assumes regions of similar Chl variability are likely to be driven by the same physical process or processes. Another key aspect of this study is that all time series are standardised to uniform variance. This helps ensure that spatial-average relationships represent the majority of local-scale relationships

comprising that spatial-average, as opposed to representing a minority of highly variant relationships.

This chapter is not limited to investigating relationships between stratification and Chl. While the study assumes that regions of coherent Chl variability are driven by common forcing (or forcings), the underlying forcing is not assumed. A number of physical processes are considered in this study including: sea surface temperature (SST); zonal and meridional wind velocities; wind speed; stratification; mixed layer depth (MLD); and the climatic index of the North Atlantic Oscillation (NAO). However, the study is not exhaustive and the focus is upon processes that drive and/or are representative of stratification and vertical mixing. Please see introductory chapter (§1) for a general background to previous studies showing relationships between these additional physical variables and PP.

The aims of this chapter are summarised in the following three questions:

- To what spatial extent are stratification and vertical mixing related to interannual Chl variability?
- Does the spatial-scale of analysis affect identification of relationships?
- Does averaging over regions of similar variability help identify relationships?

In the remainder of this chapter, the methods section (§2.2) gives details of data used and pre-processing applied, as well as the k-means clustering technique used to define regions of similar Chl variability. The main results are then presented (§2.3) and discussed (§2.4) followed by an overall conclusion (§2.5) in respective sections.

2.2 Methodology

2.2.1 Study Area

The North Atlantic, from 0 to 70°N and 70°W to 10°E, was chosen as the study area for this analysis. This area was chosen because in-situ density profiles are required in this analysis and the North Atlantic is comparatively well sampled (§1.3). Each of the gridded datasets used in this analysis were extracted to these same geographical coordinates and re-gridded into matching 1x1° cells for consistency.

This study area includes both permanently-stratified Tropical waters and seasonally well-mixed Sub-Polar waters, and thus considers regions where PP is expected to be both nutrient-limited (permanently stratified) and light limited (well mixed; c.f. §2.1.2). In this respect, this study expands upon previous analyses (e.g. Behrenfeld et al., 2006; Dave and Lozier, 2013) wherein only permanently-stratified waters were considered.

2.2.2 Data

Satellite Chl data

Monthly Chl data spanning years 1998 through 2012 were downloaded from the European Space Agency's Ocean Colour Climate Change Initiative (OC_CCI) program at: http://www.oceancolour.org/. This globally-gridded (4 Km resolution) dataset is a merged ocean colour product including data from the European Space Agency's medium-spectral resolution imaging spectrometer (MERIS) and NASA's SeaWiFS and MODIS satellite sensors. The data were subsequently extracted for the North Atlantic and re-gridded as described in §2.2.1.

Satellite SST data

Monthly SST data spanning years 1998 through 2012 are NOAA's Optimal Interpolation SST version 2 (OISST.v2) and were downloaded from: http://www.esrl.noaa.gov/psd. This globally-gridded SST dataset is optimally interpolated and uses both satellite and in-situ SST to correct biases in the satellite data due to atmospheric effects (e.g. Reynolds et al., 2002). Satellite SST for this dataset is from the NOAA family of Advanced Very-High-Resolution Radiometer (AVHRR) satellite sensors. This gridded dataset is already at 1° resolution and so is extracted to the same area as described in §2.2.1.

Wind data

Monthly wind vector data spanning years 1998 through 2012 are Blended Sea Winds downloaded from: http://www.ncdc.noaa.gov/. This 0.25° globally gridded dataset uses optimal interpolation to blend wind speed data from multiple satellite microwave scatterometers (up to 6 from 2002 including: QuikSCAT, SSMIs, TMI and AMSR-E; Zhang et al. 2006). This is then combined with wind direction information from the National Center for Environmental Prediction (NCEP) Re-analysis 2 (NRA-2) globally gridded model output to give zonal (u) and meridional (v) wind velocities. The original blended wind speed (w) can be exactly retrieved as $w = \sqrt{u^2 + v^2}$. The velocity data (u, v) were extracted for the North Atlantic and re-gridded as described in §2.2.1. Wind speed was calculated after re-gridding the u and v components.

In-situ density data

Monthly in-situ gridded temperature and salinity profiles spanning years 1998 through 2012 are from the Coriolis Reanalysis (CORA) v4.0 dataset produced and maintained by the French Research Institute for Exploration of the Sea (IFREMER) and downloaded from: http://www.myocean.eu. The CORA dataset is complied from a yearly reanalysis of the Coriolis database, under the MyOcean2 project, which collects in-situ data from many different sources and instruments. The CORA dataset is thus comprised of temperature and salinity data from different instruments, although predominantly: Argo floats, XBT, CTD, XCTD and moorings. The data are subject to an initial validation when included in the Coriolis database (including automatic checks and objective analysis) and are further subject to a delayed-mode validation (including format and range checks, climatological tests, spike checks, duplicate checks, XBT correction and ultimate objective analysis) as they are aggregated into the CORA dataset (Cabanes et al., 2013). This gridded dataset is already 1° resolution and so is extracted to the same area described in §2.2.1.

Monthly in-situ gridded density data including years 1998 through 2012 were calculated from CORA v4.0 gridded temperature and salinity data for the study area using the thermodynamic equation of seawater 2010 (TEOS-10), as encoded within the Gibbs Seawater (GSW) Oceanographic Toolbox (available for MATLAB: www.TEOS-10.org; c.f. McDougall & Barker 2011).

NAO

Monthly NAO indices for years 1998 through 2012 were downloaded from NOAA's Climate Prediction Centre: http://www.cpc.ncep.noaa.gov/. Monthly NAO indices are calculated by

applying a rotated principal component analysis (Barnston and Livezey, 1987) to standardized monthly 500mb height anomalies, as derived from the NOAA's global data assimilation system, between 20-90°N. The NAO teleconnection pattern is one of the leading rotated modes to result from the analysis.

MLD

The gridded time series of density profiles were used to calculate gridded MLD. The MLD criteria chosen for this analysis was a surface to depth density increase of 0.125 kg m^{-3}, as this corresponds to the interface with subtropical mode water in the North Atlantic (Levitus, 1982, cited in, Kara et al., 2000) To improve vertical resolution, the density profiles were interpolated to find MLD (e.g. de Boyer Montégut, 2004).

Stratification

The gridded time series of density profiles were used to calculate stratification. In order to be consistent with previous studies (Behrenfeld et al., 2006; Dave and Lozier, 2013), a 200m stratification index was chosen. Stratification was thus calculated as the surface-to-200m-depth density increase in kg m^{-3}. As 200m is one of the reference depths included in the gridded CORA v4.0 dataset, no interpolation was required to calculate this stratification index.

Yearly averages and yearly metrics

Seasonality can be difficult to remove from environmental time series. If two time-series are correlated wherein one or both has residual seasonality, it will tend to decrease or increase the relationship, respectively. If not recognised, this can result in a weaker or stronger relationship being attributed between the two 'de-seasonalised' time series than should be the case. In order to avoid potential biasing of results, seasonality is excluded from this study by using yearly metrics of the data. The loss of data points makes it harder to establish statistically significant relationships between the time series. However, results may be more robust. Table 2.1 lists the yearly metrics chosen for each of the variables used in this study.

Variable	Yearly Metric
Chl	Yearly Average
SST	Yearly Average
Zonal wind	Winter Average (D,J,F,M)
Meridional wind	Winter Average (D,J,F,M)
Wind speed	Winter Average (D,J,F,M)
MLD_{YR}	Yearly Average
MLD_{WM}	Winter Maximum (D,J,F,M)
$Stratification_{YR}$	Yearly Average
$Stratification_{WM}$	Winter Minimum (D,J,F,M)
NAO	Winter Average (J,F,M)

Table 2.1: List of yearly metrics chosen for each variable used in this study. The left column lists variables used in this study and the right column indicate what yearly metrics has been used in defining these variables.

2.2.3 Standardisation

All data are standardised to uniform variance in order to preclude large-scale relationships being biased by strong variability within a small subset of the data points comprising that region. The Chl data were also standardised for use in the clustering analysis. Although clustering analysis is applied to a single-variate dataset here, standardising the data before clustering ensures that the Chl data are clustered according to similarities in interannual variability and not regional differences in mean and variance.

The NAO index is by default a standardised time series. Each of the other datasets are multi-dimensional and are standardised to have zero mean and uniform variance in the time dimension. As an example, standardising Chl data involved subtracting the temporal mean from every time series and then dividing the resulting anomalies by their temporal standard deviation:

$$ZChl_{ij} = \frac{\left(Chl_{ij} - \overline{Chl_i}\right)}{\sigma_{Chl_i}} \tag{2.1}$$

Where $ZChl_{ij}$ is the standardised form of Chl at observation i and year j (Chl_{ij}); $\overline{Chl_i}$ is thus the mean of the time series at observation i and σ_{Chl_i} is the corresponding standard deviation of that time series. Using z-scores puts the values in terms of number of standard deviations from the mean and helps to bring values into a comparable scale.

2.2.4 Clustering

K-means algorithm

A K-means clustering technique was used to define regions of similar Chl variability. The algorithm initially produces a random partition of N observations into K clusters and then iteratively repartitions the observations amongst the K clusters based upon minimising the Euclidean distance between observations and cluster centroids. Each time observations are repartitioned the centroids of the K clusters change. At each iteration the Euclidean distance between all N observations and K centroids is calculated and observations are subsequently assigned to the cluster with the closest centroid, or remain assigned to the cluster they are in if it has the closest centroid. Iterations can either involve the repartitioning of all observations at once, or a single reassignment of the most favourable movement. For efficiency, the algorithm can start by making bulk reassignments for a given number of iterations and then switch to making single repartitions. If this process is repeated enough times, and each cycle involves enough iterations, then this heuristic algorithm converges towards an optimal partition of N observations into K clusters without having to exhaustively calculate every possible solution (e.g. Hartigan and Wong, 1979).

The observations clustered in this study are standardised yearly-averaged Chl time series. Each observation ($ZChl_i$) is thus a vector of length v, corresponding to the number of years in the time series. The Euclidean distance (D) between the i^{th} $ZChl$ observation and k^{th} cluster centroid is thus:

$$D_{ik} = \left[\sum_{j=1}^{v}\left(ZChl_{i_j} - \overline{ZChl_{\in k}}_j\right)^2\right]^{1/2} \quad (2.2)$$

Where $\overline{ZChl_{\in k}}$ is the mean of all $ZChl$ time series that are members of cluster k, *i.e.* the centroid of the k^{th} cluster. However, as the $N \times ZChl$ time series being clustered in this study are geographical data, a novel geographical-distance penalty is introduced in order to penalise against geographical distance and thus encourage the resulting clusters to be geographically contiguous.

Instead of using K-means to cluster data according to minimising overall values of D, the data is clustered according to minimising geographically-penalised distances (D_{pnlty}) defined as:

$$D_{pnlty_{ik}} = D_{ik} + \Omega \cdot d_{ik} \tag{2.3}$$

Where d_{ik} is the great circle distance between the i^{th} Chl observation and k^{th} cluster centroid, and Ω is a scalar constant that determines the overall weighting of this geographical distance penalty: higher values of Ω will tend towards geographically coherent clusters, as minimising $D_{pnlty_{ik}}$ increasingly acts towards minimising d_{ik} within the modified K-means algorithm. The value of Ω is chosen subjectively via trial and error, seeking the minimum value of Ω that will reduce speckle and pull data into reasonably-coherent clusters. Other studies have included latitude and longitude as variables directly within the cluster analysis (e.g. Devred et al., 2007), but these do not weight the geographical terms separately. While it would be possible to include latitude and longitude as additional variables *and* weight the variables separately, the method chosen here is preferable as it uses great circle distances, which are more accurate in terms of both diagonal distances and measuring distances on a sphere.

2.2.5 Stopping Criteria

K-means clustering requires choosing K: the number of clusters to divide observations into. This can be chosen subjectively or using *a priori* knowledge of the data. However, assuming there is an optimal number of clusters to divide the data into, K can be chosen objectively using stopping criteria. Three stopping criteria were considered in this study and described as follows:

Variance Ratio Criterion

The ultimate goal of clustering may be thought of as finding a separation of the observations into K clusters that minimises total within cluster variance ($Var(w)$) and maximises total between cluster variance ($Var(b)$). Thus, the ratio of these two variances, when considered over a range of solutions for different K, may be used to identify the optimal number of clusters in the data:

$$VRC = \frac{Var(b)}{Var(w)} \tag{2.4}$$

VRC is the variance ratio criterion which is often given in the form of sum-of-squares SS between and within clusters, and called the Calinski-Harabasz criterion; the optimal number of clusters,

according to this criterion, is found by maximizing VRC with respect to K (Calinski and Harabasz, 1974). Out of 30 objective stopping criteria applied to an artificial dataset, the VRC criterion was found to be the most reliable at determining the true number of clusters (Milligan and Cooper, 1985).

Silhouette

The silhouette criterion attempts to measure how well each observation belongs to its assigned cluster. Each observation is given a silhouette value (s) which ranges between 1 and -1 and indicates how close the observation is to other points in its own cluster in comparison to how close it is to points in the next nearest cluster: $s \to 1$ indicates the observation is much closer to its own cluster; $s \to -1$ indicates the observation is much closer to its nearest neighbouring cluster. By comparing the average of these silhouette values ($S = \bar{s}$) between solutions over a range of K, this criterion may be used to identify the optimal number of clusters in the data:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{2.5}$$

Where $a(i)$ is the average distance between observation i and all other points in the same cluster, and $b(i)$ is the average distance between observation i and all other points in the nearest neighbouring cluster. The maximum value of S represents the optimal number of clusters (Rousseeuw, 1987).

Sum of Squares Within Cluster

The third stopping criterion used in this study are plots showing the decrease in mean sum-of-squares within clusters (SSW) against increasing number of clusters (K). These plots do not give an objective result but need to be interpreted to find an 'elbow': a point separating larger and smaller decreases in SSW with increasing K. SSW is calculated as:

$$SSW = \sum_{k=1}^{K} \sum_{i \in k} \sum_{j=1}^{v} \left(ZChl_{i_j} - \overline{ZChl_{\in k}}_j \right)^2 \tag{2.6}$$

2.2.6 Number of Clusters

In order to determine the optimal number of clusters into which to separate Chl data, values for the three stopping criteria described in §2.2.5 were compared over a range of Ω and K vales. To calculate these, initial clustering solutions were found by running the modified K-means algorithm (§2.2.4) for a limited number of repetitions (10 reps, 100 iterations) over a range of Ω and K values. Values of Ω from 0 to 4 in steps of 0.5 were chosen for this, as preliminary analysis showed that, depending upon K, these were the lowest values to give subjectively coherent clusters. The range of K values used (2, 3, … , 10, 15, 20, 30, 40, 50, 100) was chosen to focus on K≤15, wherein preliminary results showed the greatest rate-of-change in SSW with K, while also allowing for higher values to keep results in an overall context.

As S2.1 – S2.3 show, stopping criteria depend upon Ω. This is to be expected as Ω is a compromise between minimising within-cluster variance and improving the geographical coherence of the clusters (c.f. Eq. 2.3). Therefore, within-cluster variance tends to increase with Ω and this effect can be seen across the different stopping criteria.

With the exception of silhouette values at high K (≥30), all three criteria indicate K=2 as a preferential number of clusters. While silhouette plots imply preferential solutions for k≥15 (not shown), these results are in contrast to VRC and SSW results (which continually decline with increasing k). This would also approach a local scale analysis, which is not the aim of this study. (For comparison, the study domain equates to roughly 140 5×5° cells, which may be considered local scale.) This analysis is aimed at identifying coherent regions of variability at an intermediary-scale between local-scale and basin-scale analyses. As such, solutions at k>15 are ignored. For all values of Ω at K<15, both VRC and mean Silhouette values give best results for a two-cluster solutions (S2.1, S2.3). Neither of these criteria, however, indicate how well a two-cluster solution comparers to the un-clustered data. The third stopping criteria used in this study are plots of SSW against k, for which there is no exact objective result. Instead, the plots need to be interpreted to find an ‘elbow’: a point separating larger and smaller decreases in SSW with increasing k. However, the benefit of SSW is that the results can be compared to k=1 (i.e. no clusters). The most pronounced elbow in these plots does appear to be at K=2 for Ω≤2 (S2.2). Thus, k=2 is taken to be the optimal number of clusters for Ω≤2. Furthermore, as the data tend to form more geographically-coherent clusters at lower k (trial-and-error; not shown), no geographical distance penalty (i.e. Ω=0) is required at k=2.

Although stopping criteria give best results for two cluster solutions, further solutions are also considered in order to test for relationships at smaller spatial scales. Results from the stopping criteria are, however, fairly ambiguous for more than two clusters. For example, for $\Omega<1$, the VRC and SSW stopping criteria tend to vary smoothly with k and do not indicate an optimum value. At higher values of Ω, however, these criteria indicate slight preferences for k= 4, 7 and 9 (S2.1). Average silhouette values, by contrast, tend to imply that the least-favourable solutions are for between k=4 and 10 for $\Omega\leq1$, or k=3 and 10 for $\Omega>1$. Given such ambiguous and contradictory results, an intermediate value of k, for between k=3 and 15, was chosen subjectively. An intermediate value of k=7 was chosen as this allowed for spatial heterogeneity, while maintaining a practicable number of clusters to be analysed. There is also a plateau in VRC values around k=7 for $\Omega\geq1.5$ (S2.1). While this is not an objective value, it should be borne in mind that this secondary clustering solution is only for comparative purposes: the stopping criteria implied an optimal solution of k=2. A value of $\Omega=3$ was chosen (subjectively) to correspond with k=7 as this was found to be the lowest value to result in reasonably-coherent clusters.

2.3 Results

2.3.1 Two-Clusters Analysis

Two Clusters

The optimal division of chlorophyll variability in the North Atlantic gives two anti-correlated (r = -0.61; p=0.015) clusters having a roughly east-west separation (Fig. 2.1). The geographical coherence of these clusters, derived using Ω=0, implies this is not an arbitrary separation of the data. As the two are also strongly anti-correlated, they appear to represent a bimodal pattern of interannual Chl variability in the North Atlantic. Thus, these clusters are used to explore for potential relationships between proxies of stratification and interannual Chl variability. As the North Atlantic Oscillation (NAO) is the dominant mode of climatic variability in the North Atlantic, possible relationships between this large-scale climatic index and these clusters is also explored.

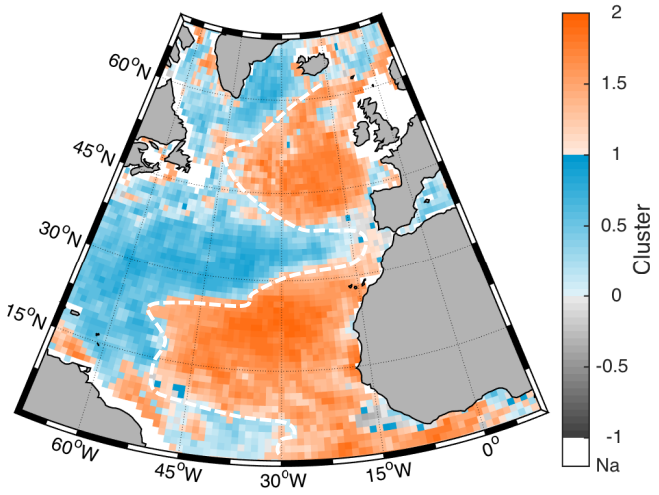


Fig. 2.1: Two cluster partition of interannual Chl variability within the North Atlantic for Ω=0. Colours denote cluster number: Cluster 1 (blue), Cluster 2 (orange). Where present, grey indicates locality is anti-correlated with cluster. Shelf sea localities less than 200m depth are omitted (Na; white). Shading (of clusters and anti-correlated localities) denotes strength of correlation (or anti-correlation) with cluster, from light (R^2 = 0) to dark (R^2 = 1).

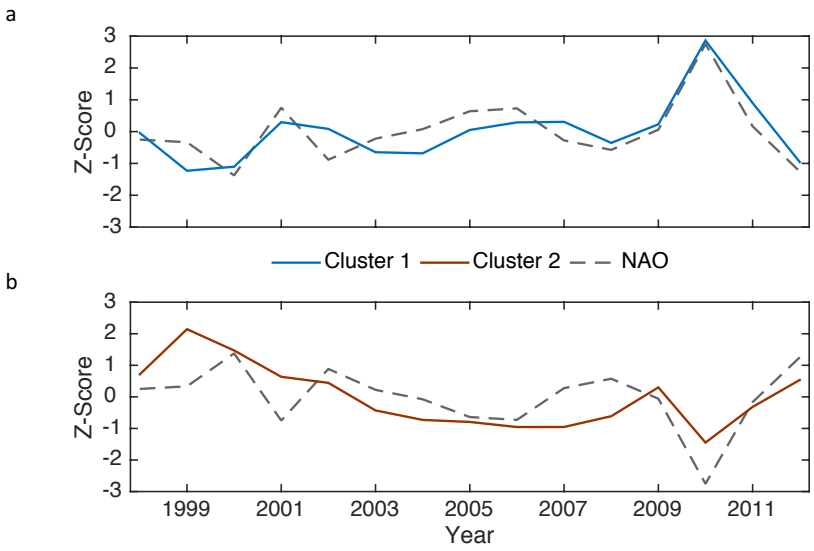


Fig. 2.2: Time series of interannual Chl variability for a) Cluster 1 and b) Cluster 2 in Fig. 2.1. Dashed lines represent NAO index (or -1 × NAO index in a) for comparison. Units are in standard deviation.

Fig. 2.2 shows the relationship between Chl variability and NAO for each cluster in Fig. 2.1. The two clusters have opposing relationships with the NAO, with Chl variability in Clusters 1 and 2 being negative and positive related to the NAO, respectively.

	1	2
SST	0.00 0.967	**-0.55** 0.002
Stratification	+0.03 0.562	-0.20 0.091
$Stratification_{WM}$	-0.11 0.222	**-0.35** 0.021
MLD	0.00 0.875	**+0.37** 0.016
MLD_{WM}	+0.17 0.128	+0.08 0.296
Wind Speed	0.00 0.999	+0.20 0.106
Zonal wind	**+0.47** 0.007	0.00 0.888
Meridional wind	-0.24 0.074	-0.13 0.205
NAO_{JFM}	**-0.71** 0.00	**+0.33** 0.024

Table 2.2: R^2 and associate p-values values (beneath) for strength of relationship between interannual Chl variability and explanatory variables (rows) for each cluster in Fig. 2.1. Columns denote cluster numbers (Cluster 1 or 2). A positive or negative symbol before each R^2 value denotes the sign of relationship. For emphasis, statistically significant R^2 values ($p<0.05$) are given in bold type and also shaded according to the sign of relationship: blue (negative) or red (positive).

Perhaps the most notable feature distinguishing the two cluster distribution in Fig. 2.1 from a simple east-west partition, is the formation of Cluster 2 into two groups originating in the east and extending westwards into the North Atlantic at latitudes of about 0-30°N and 38-60°N. These latitudes more-or-less coincide with the Trade Wind (~0-30°N) and Westerlies (~35-60°N) latitudes, with the high-pressure ridge of the Horse Latitudes (~30-35°N) situated in-between (Fig. 1.2).

As shown in Table 2.2, Cluster 1 has an especially strong correlation with NAO. This climatic index explains ~71% of the variance in interannual chlorophyll variability within this cluster. The cluster is also significantly related to zonal wind velocity, but not to any of the other proxies for stratification and vertical mixing used in this study. Cluster 2, by contrast, has an opposite relationship with the NAO and is strongly related to many of the proxies for stratification and vertical mixing (SST, MLD, $stratification_{WM}$), but is not significantly related to any of the variables for wind speed/velocity.

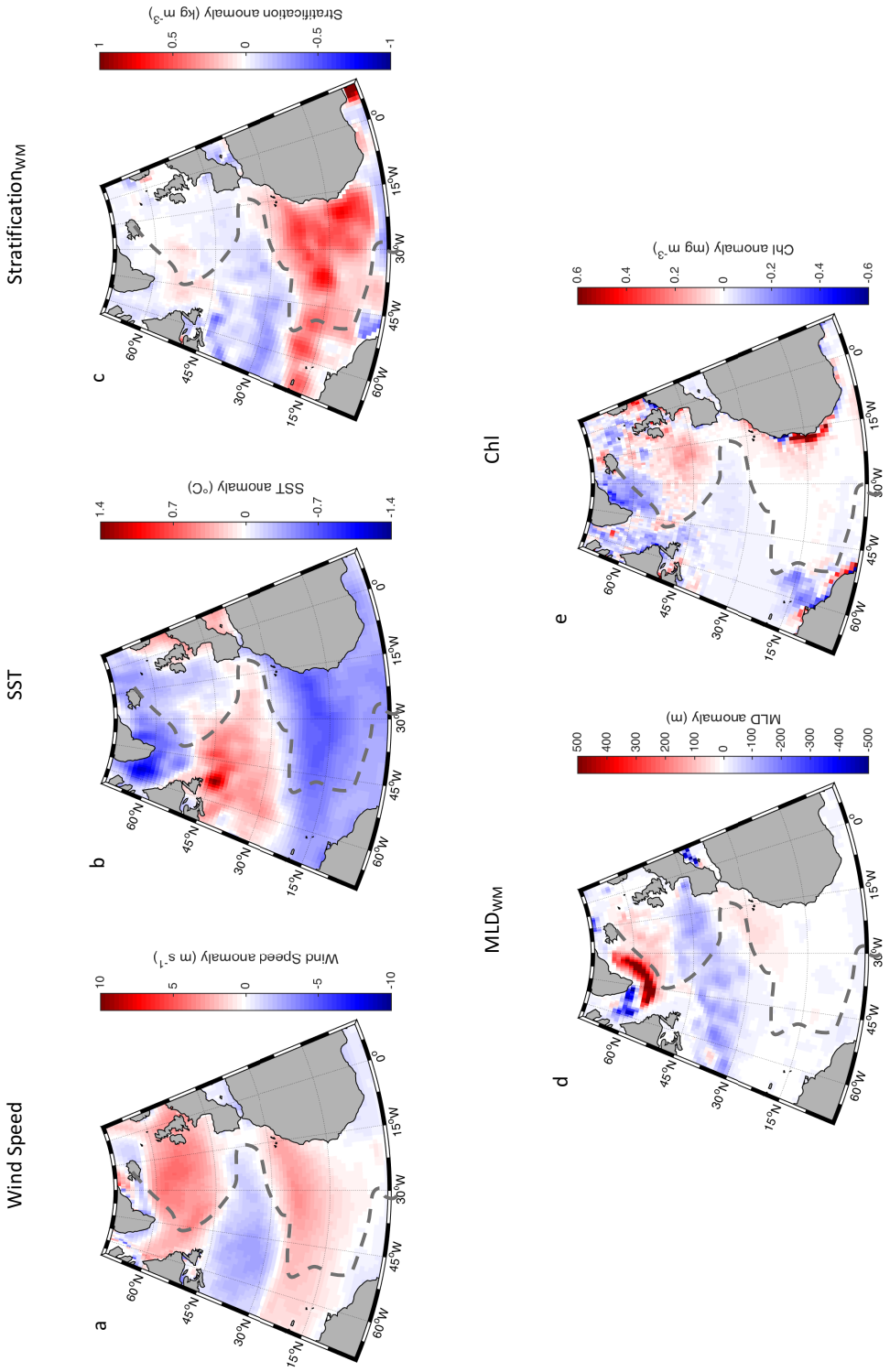


Fig. 2.3: NAO patterns of explanatory variables (as labelled; c.f .Table 2.1). Plots are a composite of years with strong NAO indices (|NAO| ≥ 1; 2000, 2010, 2012) and are plotted in aspect of positive NAO years. For each 1° cell of each variable, the mean is subtracted from the time series to give yearly anomalies and then the overall mean of positive NAO (≥+1) years and inversed (-1×) negative NAO (≤-1) years is calculated. For comparison, dashed lines shows the approximate division of the two clusters in Fig. 2.1.

Given the strong relationship identified between NAO index and interannual Chl variability within Cluster 1, NAO patterns of Chl and other variables used in this study are considered for comparison. Fig. 2.3e shows the NAO pattern of yearly averaged Chl variability in the North Atlantic. For comparison, a dashed line overlays the plot which roughly outlines where the two clusters in Fig. 2.1 are separated. The strong similarity between the cluster distribution and this NAO-Chl pattern implies that the NAO dominates the large scale forcing of interannual Chl variability identified in Fig. 2.1. In addition to Chl, wind speed (winter average), SST (yearly average) and stratification (winter minimum) also have similar NAO patterns to the two-cluster distribution of interannual Chl variability (c.f. Figs. 2.1 and 2.3). These broadscale similarities imply that the two-cluster pattern in Chl (Fig. 2.1) may be linked to climatic variability associated with NAO via changes in these physical variables. However, if these NAO patterns are assumed to drive the two-cluster Chl pattern, then the discrepancies between them either imply that different forcings dominate in different localities, or that Chl responds differently to the same forcing in different localities. Either way, separating the Chl data into more clusters may help identify relationships that could otherwise be obscured by spatial averaging.

2.3.2 Seven-Clusters Analysis

Seven Clusters

Fig. 2.4 shows the seven-cluster partition of interannual Chl variability in the North Atlantic. A comparison with the previous two-cluster partition shows that Clusters 1, 3 and 7 in this partition more-or-less subdivide Cluster 1 in Fig. 2.1, while Clusters 2, 4, 5 and 6 correspondingly tend to subdivide Cluster 2.

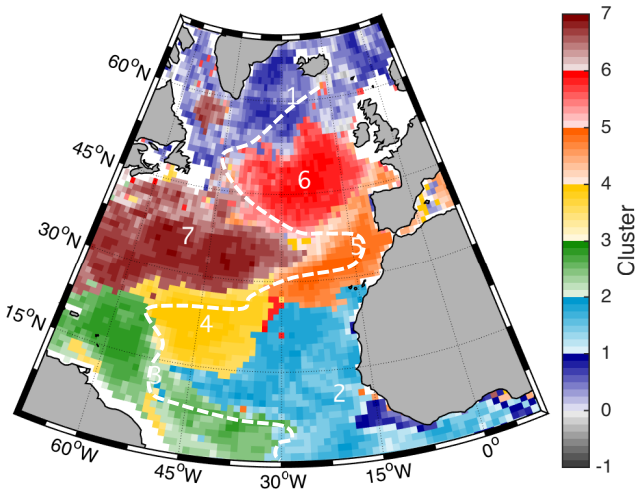


Fig. 2.4: Seven cluster partition of interannual Chl variability within the North Atlantic using Euclidean distance and Ω=3. Colour and shading of each 1×1° represents cluster and correlation to cluster centroid (mean), respectively. Where present, grey colour/shading represents anti-correlation to cluster. Shelf sea localities less than 200m depth are omitted (white). Numbers mark the geographical centre of respective clusters. For comparison, the dashed line shows the approximate division of the two clusters in Fig. 2.1.

	1	2	3	4	5	6	7
SST	-0.08 0.317	**-0.57** 0.001	-0.03 0.531	**-0.27** 0.046	-0.06 0.387	-0.08 0.297	**-0.58** 0.001
Stratification	-0.01 0.749	**-0.53** 0.002	-0.01 0.693	0.00 0.843	-0.08 0.308	0.00 0.992	0.00 0.995
$Stratification_{WM}$	-0.13 0.186	**-0.51** 0.003	0.00 0.859	0.00 0.832	**-0.61** 0.001	-0.01 0.772	-0.17 0.125
MLD	+0.03 0.547	+0.03 0.559	0.00 0.820	**+0.43** 0.008	+0.07 0.327	**+0.51** 0.003	+0.01 0.754
MLD_{WM}	+0.01 0.776	0.00 0.901	0.00 0.889	0.00 0.912	**+0.56** 0.001	+0.01 0.707	+0.26 0.054
Wind Speed	+0.03 0.574	+0.19 0.115	+0.17 0.147	0.00 0.889	+0.12 0.225	+0.16 0.163	**+0.31** 0.037
Zonal wind	0.00 0.813	**-0.38** 0.02	-0.20 0.108	0.00 0.964	-0.15 0.173	**+0.31** 0.039	**+0.37** 0.022
Meridional wind	-0.14 0.196	-0.15 0.178	-0.01 0.737	0.00 0.950	**-0.41** 0.014	-0.01 0.773	**-0.39** 0.016
NAO_{JFM}	-0.1 0.261	**+0.46** 0.006	-0.13 0.191	+0.04 0.483	+0.02 0.575	+0.18 0.114	**-0.56** 0.001

Table 2.3: R^2 and associated p-values values (beneath) for strength of relationship between interannual Chl variability and explanatory variables (rows) for each cluster in Fig. 2.4. Columns denote cluster numbers (Clusters 1 to 7). A positive or negative symbol before each R^2 value denotes the sign of relationship. For emphasis, statistically significant R^2 values ($p<0.05$) are given in bold type and also shaded according to the sign of relationship: blue (negative) or red (positive).

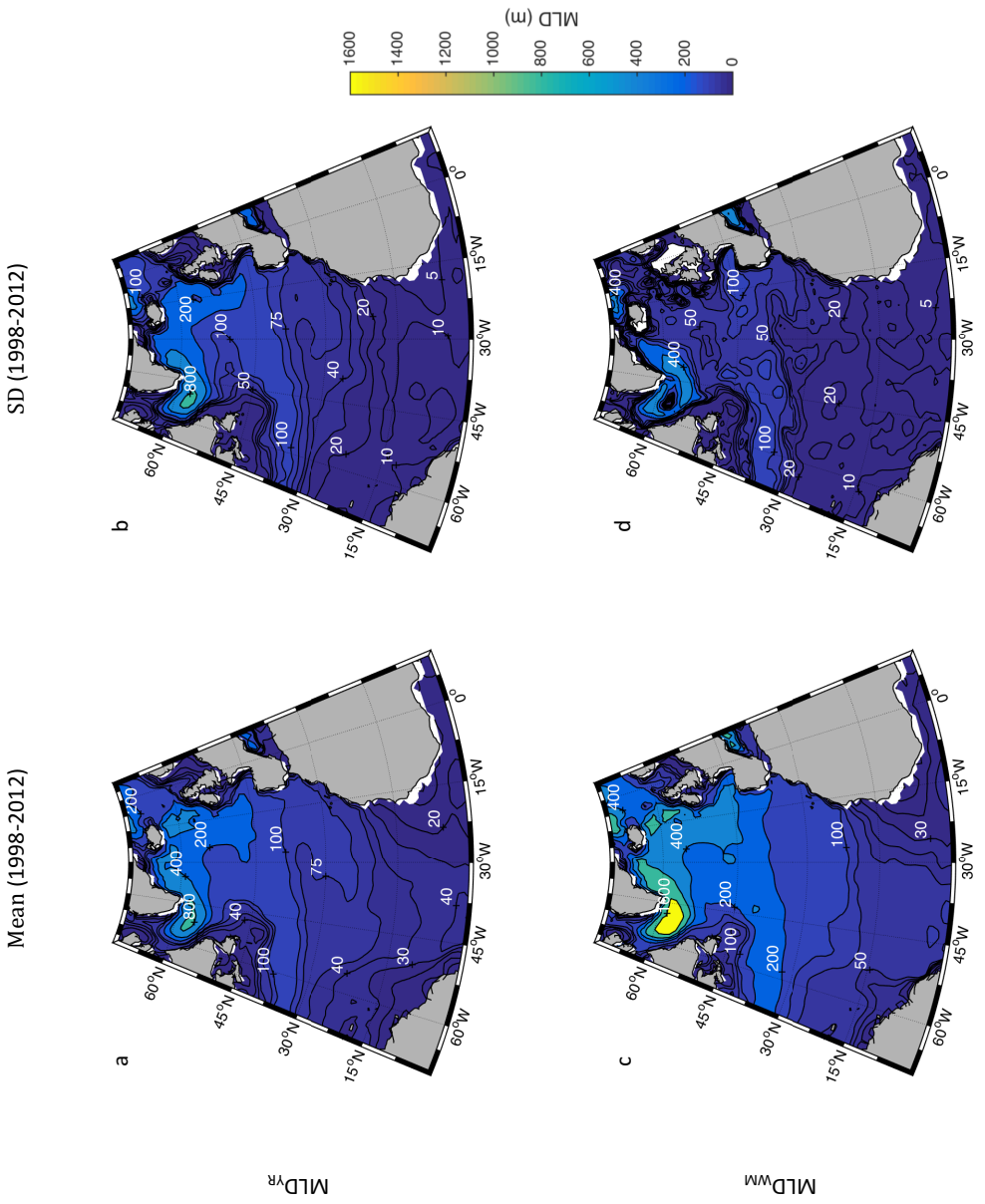


Fig. 2.5: Contour plots showing yearly composites (Mean; first column) and standard deviations (SD; second column) of both yearly averaged and winter-maximum MLD for the North Atlantic. Colour and labeling of plots corresponds to depth in meters, as per colour bar. Plots represent MLD data used in this study for years 1998-2012.

Table 2.3 gives the sign and strength of relationships between interannual Chl variability and explanatory variables for each cluster in Fig. 2.4. These relationships are briefly discussed in the following bullet points. (Nb. To avoid confusion when comparing relationships between the two and seven cluster scale, clusters maybe distinguished by using prefixes K2 [for the two-cluster partition] and K7 [for the seven-cluster partition]. Also, to avoid repetition, relationships with Chl in a particular cluster are often discussed in terms of a relationship with that cluster.)

- Cluster 1 is not significantly related to any of the variables used in this study. As this cluster represents a high-latitude region it may be expected to be light-limited and thus inversely related to yearly-averaged MLD, but no such relationship is found.

- Cluster 2 is strongly related to NAO, zonal wind velocity, winter-minimum stratification, yearly-averaged stratification and SST. This implies that the NAO index may be related to Chl via changes in wind and vertical mixing and their effects upon nutrient upwelling in this cluster. This Cluster represents a large part of the eastern Tropical North Atlantic and thus sits within the trade wind latitudes, with the intensity of these winds varying with the NAO. A comparison with NAO patterns shows that it affects wind speeds, SST and winter-minimum stratification in this region (Fig. 2.3).

- Cluster 3 is not significantly related to any of the variables used in this study. It is noteworthy, however, that this region does show strong and significant relationships with stratification and MLD if the last few years (2010-2012) are excluded from the regressions (not shown). However, this is not explored further here.

- Cluster 4 is significantly related to yearly-averaged MLD and SST. This implies a link between yearly-averaged Chl variability and vertical mixing, although the lack of a corresponding relationship with yearly-average stratification is conspicuous.

- Cluster 5 is significantly related to winter-minimum stratification, winter-maximum MLD and meridional wind velocity. These relationships suggest a dependence upon winter mixing and upwelling of nutrients in this cluster, which is consistent with the seasonal onset/breakdown of stratification at higher latitudes. This cluster is situated within a region that experiences northerly winds as air travels around the anticyclonic Azores high-pressure system. This may explain the association with meridional wind velocity, as opposed to zonal

wind velocity or wind speed. Furthermore, a comparison with Fig. 2.5a shows that this cluster coincides with a region of deeper-mixing water extending southwest from a region close to the Mediterranean.

- Cluster 6 is positively related to zonal wind velocity and yearly-averaged MLD. A comparison with Fig. 2.5 shows that this cluster coincides with a region of the northeastern North Atlantic that has a comparatively deep MLD. The deep mixing in this region may explain the lack of a corresponding association with the 200m stratification variable. Fig. 2.5 also shows that winter maximum MLD in this region is far less variable than yearly-averaged MLD. The depth and consistency of winter-maximum MLD may explain its lack of association with Chl in this region. The significant relationships with zonal wind velocity and yearly-average MLD may instead allude to the importance of sporadic mixing events, during or after shoaling of the seasonal pycnocline in this region.

- Cluster 7 is strongly related to NAO, SST and both eastwards and northwards wind velocities. It does not, however, show any significant relationships with the four stratification and MLD proxies. The association with zonal wind velocity in this cluster may be expected as it mostly represents the North Atlantic subtropical gyre, and is situated within the high-pressure 'horse latitudes' and represents a region of the North Atlantic that varies strongly in winter-average wind speed with the NAO (c.f. Fig. 2.3). The lack of significant relationships with stratification and MLD in this region is surprising as the NAO is also seen to strongly influence these variables as well as yearly-averaged Chl in this region (Fig. 2.3). It is noteworthy that this region does show strong and significant relationships with 300m winter-minimum stratification (not shown). A 200m stratification index is used in this study and the only two clusters, in this seven-cluster partition, that show a significant relationship with 200m stratification are clusters 2 and 5. As Fig. 2.5 shows, in comparison to these other two clusters, MLD is deeper in the vicinity of this cluster and thus the association with a deeper stratification index may be explained. This is not, however, explored further here.

In terms of the distribution of relationships throughout the North Atlantic, these do not appear to have any general pattern or consistency with respect to latitude, mixing depths (Fig. 2.5), or the K2 cluster pattern (Fig. 2.1). For example, it may be expected that winter-metrics of stratification and MLD would be more strongly related to Chl variability within higher latitude seasonally-stratified waters (~ >45°N), while yearly-averages of these variables would be better predictors in

lower latitude permanently-stratified waters. However, while there is limited evidence for such a distribution when comparing differences in R^2 values between yearly-averaged and winter-minimum stratification relationships (e.g. low latitude Clusters 2 and 3 are more strongly related to stratification, while higher-latitude Clusters 1, 5, 6 and 7 are more strongly related to stratification$_{WM}$), these differences in R^2 values are often small and represent comparisons between insignificant relationships. Furthermore, no such order is apparent when likewise comparing the distribution of yearly-averaged and winter-maximum MLD relationships (Table 2.3). Nor is there any consistency in terms of the type of relationships within clusters that more-or-less subdivide the larger two clusters in Fig. 2.1 (c.f. Fig. 2.4, Table 2.3). Lastly, there does not appear to be any consistency between stronger relationships with stratification (referenced to 200m depth) and/or MLD in comparison to the distribution of climatological mixing depths shown in Fig. 2.5 (c.f. Figs. 2.4, 2.5, Table 2.3). The heterogeneity of relationships at this spatial scale implies that regionality is important. That is, relationships between Chl and vertical mixing proxies are regionally dependent at this scale of analysis.

Comparison of K2 and K7 Clusters

Insofar as the seven clusters in Fig. 2.4 represent a subdivision of the two larger clusters in Fig. 2.1, then Tables 2.2 and 2.3 can be compared for consistency of relationships between these spatial scales.

The NAO is significantly related to Chl variability with opposing relationships in both clusters at the larger spatial scale (Table 2.2). However, when the seven cluster scale is considered, these opposing relationships with the NAO appear to be constrained to smaller areas within the confines of the two larger clusters (K7 Clusters 2 and 7; Table. 2.3). This may imply that the relationship between Chl and NAO is focused upon these smaller regions, but not limited to them. The use of standardised data in this analysis results in the time series of each 1° cell being evenly-weighted when averaged together. Thus, cluster-average relationships are unlikely to be biased towards smaller sub-regions in this way. Furthermore, although the relationship between Chl and NAO in K2 Cluster 2 is stronger when isolated at the smaller scale (K7 Cluster 2), it is weaker in K2 Cluster 1 when likewise isolated (K7 Cluster 7). Thus, although the other clusters comprising K2 Cluster 2 at the seven-cluster scale do not have significant relationships between Chl and NAO (K7 Clusters 4, 5, 6; Table 2.3), their inclusion in the larger-scale cluster does strengthen the overall relationship between Chl and NAO in comparison to the smaller scale. Thus, for K2 Cluster 2 at

least, it is apparent that the relationship between Chl and NAO is not limited to the smaller area implied by the seven-cluster analysis (K7 Cluster 2; Table 2.3).

In a similar manner, relationships between Chl and other variables (zonal wind, SST, $stratification_{WM}$ and MLD_{YR}) that are also significant at the two-cluster scale (Table 2.2), become both stronger and weaker when considered at smaller spatial scales (Table 2.3). Perhaps the most notable feature of this comparison is that only one of the three smaller clusters that tend to comprise K2 Cluster 1 (i.e. K7 Clusters 1, 3, 7) have any significant relationships with Chl (K7 Cluster 7; Table 2.3). Nevertheless, of the two relationships that are common between this smaller cluster (K7 Cluster 7) and the larger K2 Cluster 1, both are stronger at the larger scale (NAO, SST; c.f. Tables 2.2, 2.3). By contrast, although relationships with Chl tend to be more numerous at both scales of analysis within the confines of K2 Cluster 2, all of the relationships that are significant at the larger scale are stronger when considered at the smaller K7 spatial scale (zonal wind, SST, $stratification_{WM}$, MLD_{YR}; c.f. Tables 2.2, 2.3).

2.3.3 Local Scale Analysis

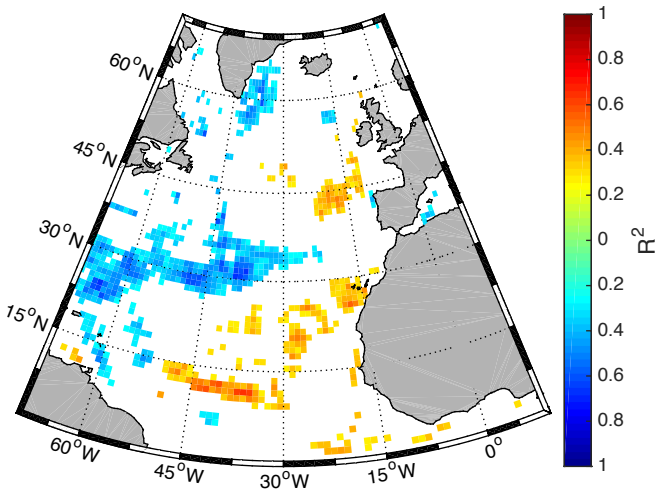


Fig. 2.6: Local-scale (1°) regression of wintertime (JFM) NAO index upon interannual Chl variability for the North Atlantic. Only significant ($p \leq 0.05$) relationships are shown. Of these, each 1° cell is coloured according to the sign of the relationship (yellow/red positive; cyan/blue negative) and the magnitude of the coefficient of determination (R^2), as per the colourbar. White areas denote insignificant relationships.

Figure 2.6 shows the local-scale (1°) distribution of relationships between Chl and the NAO index. Although there are a large number of localities where these two variables are not significantly related, the general divide between regions of positive and negative relationship agrees well with the two-cluster distribution in Fig. 2.1. This similarity in spatial distribution further implies that the two-cluster distribution of interannual Chl variability in Fig. 2.1 is associated with the NAO. However, while both clusters are significantly related to the NAO at the larger-scale (Table 2.2), less than half of the corresponding localities are significantly related when assessed at the local scale (S2.5).

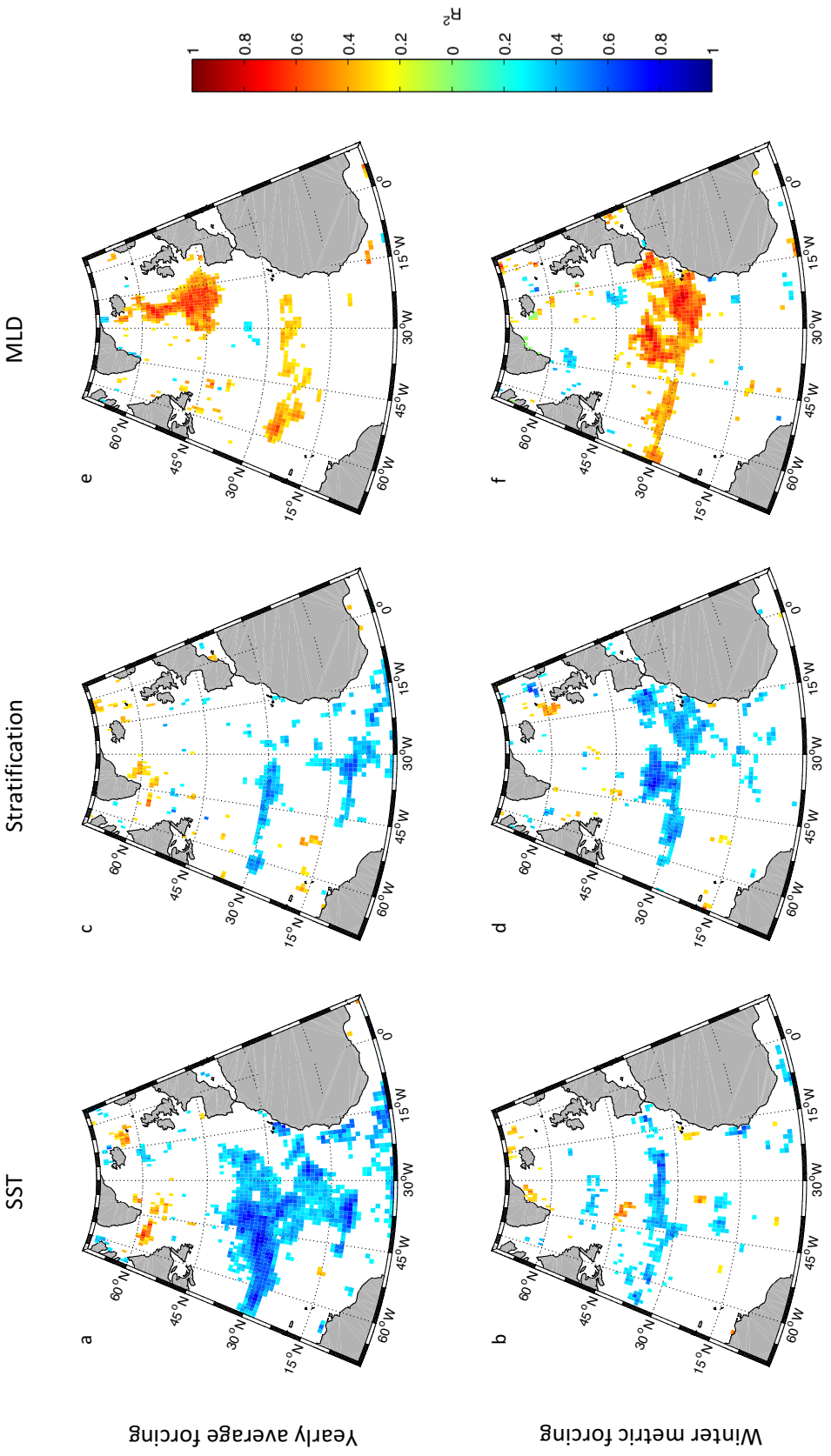


Fig. 2.7: Local-scale (1°) regressions of SST, Stratification and MLD upon interannual Chl variability for the North Atlantic. Only significant ($p \leq 0.05$) relationships are shown. Of these, each 1° cell is coloured according to the sign of the relationship (yellow/red positive; cyan/blue negative) and the magnitude of the coefficient of determination ($R2$), as per the colourbar. White areas denote insignificant relationships. Rows denote whether yearly-averages or winter-metrics (b. winter average SST, d. winter minimum stratification, f. winter maximum MLD) of the explanatory variables are used in the regressions. In each case, the response variable is yearly-averaged Chl.

Fig. 2.7 shows local scale regressions of SST, stratification and MLD upon interannual Chl variability. These plots show that large areas of the North Atlantic are strongly related to one-or-other of these proxies for vertical mixing at the local scale. As may be expected, the distribution of these local-scale relationships tends to coincide with the distribution of relationships identified in K7 clusters (c.f. Figs. 2.4, 2.7 and Table 2.3). For example, areas where yearly-averaged SST and Chl are significantly and coherently related at the local scale (Fig. 2.7a), coincide well with the positions of K7 Clusters 2 and 7 (Fig 3.4): the two K7 clusters with the strongest relationships between these variables (Table 2.3). Likewise, other regions of coherent local scale relationship in Fig. 2.7 coincide well with clusters having corresponding relationships in Fig. 2.4 and Table 2.3. A comparison of Figs. 2.4, 2.7 and Table 2.3 reveals the following commonalities: $\text{stratification}_{YR}$ (K7 Cluster 2), $\text{stratification}_{WM}$ (K7 Cluster 5), MLD_{YR} (K7 Clusters 4 and 6), MLD_{WM} (Cluster 5). However, there are also discrepancies between these analyses. For example, Chl in K7 Cluster 7 is not significantly related to $\text{stratification}_{YR}$, $\text{stratification}_{WM}$ or MLD_{WM} at the cluster scale (Table 2.3), but at the local scale there are regions within this cluster that have significant and coherent relationships between Chl and these variables at the local scale (c.f. Figs. 2.4, 2.7). Thus, as in the comparison of relationships between the two and seven cluster scales, it is apparent that local scale relationships are more spatially heterogeneous than implied by cluster-average relationships.

Local scale vs. Clusters

In the introduction to this chapter, three possible reasons were given for why a large-scale spatially averaged relationship may be confined to a limited number of cells when considered at the local scale. One reason suggested that locally significant relationships might dominate the large-scale mean. Depending upon the ratio of significant to non-significant cells comprising the larger area, this would require that the variance of the corresponding variables was higher in the significant cells than in the non-significant ones and/or that the area average of all of non-significant cells tends to cancel out or represent a white noise sequence. This study has used standardised time series in an attempt to mitigate the potential for the former of these to happen, however this does not rule it out.

In order to test if the local scale relationships observed in Figs. 2.6 and 2.7 are dominating the cluster-average relationships listed in Table 2.2, the within cluster relationships are reassessed by excluding locally-significant cells from the cluster average.

Variable	Cluster	All cells	Non-Sig. cells
NAO_{JFM}	1	**-0.75**	**-0.57**
NAO_{JFM}	2	**+0.33**	+0.24
SST	2	**-0.55**	**-0.47**
Zonal wind	1	**+0.47**	**+0.32**
$Stratification_{WM}$	2	**-0.35**	**-0.27**
MLD_{YR}	2	**+0.37**	+0.21

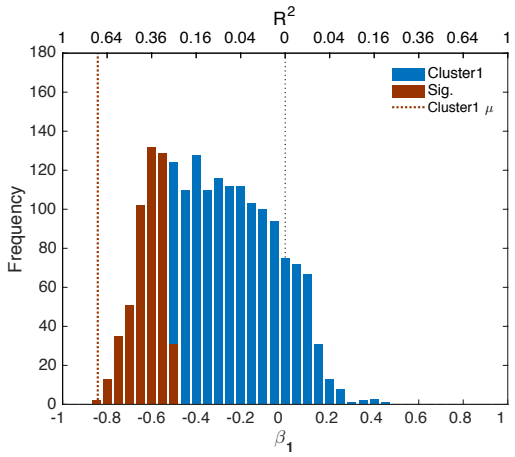
Table 2.4: Comparison of R^2 values in Clusters 1 and 2 of Fig. 2.1 before and after excluding localities with significant relationships (Figs. 2.6, 2.7) from the cluster-mean time series. A positive or negative symbol before each R^2 value denotes the sign of relationship. For emphasis, statistically significant R^2 values ($p<0.05$) are given in bold type and also shaded according to the sign of relationship: blue (negative) or red (positive).

Table 2.4 shows that the cluster-average relationships identified in Table 2.2 tend to remain significant, or at least of the same sign, when locally-significant relationships are omitted from the cluster averages. This implies that the cluster-average relationships identified in Table 2.2 are representative of the entire area of each respective cluster, and are not biased towards strong localised relationships.

Cluster 1

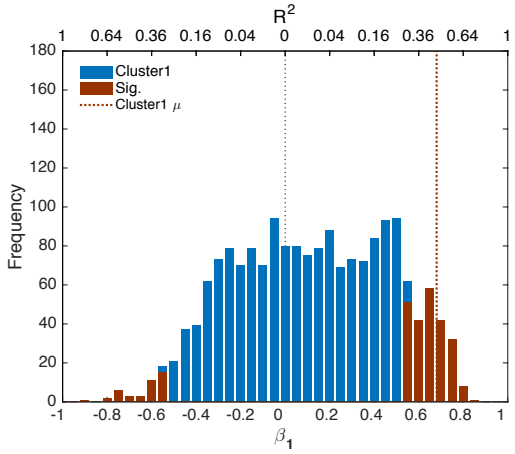
NAO

a



Zonal wind

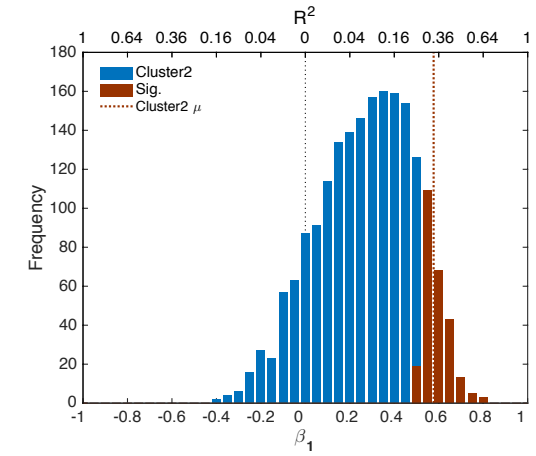
b



Cluster 2

NAO

c



SST

Stratification$_{WM}$

MLD$_{YR}$

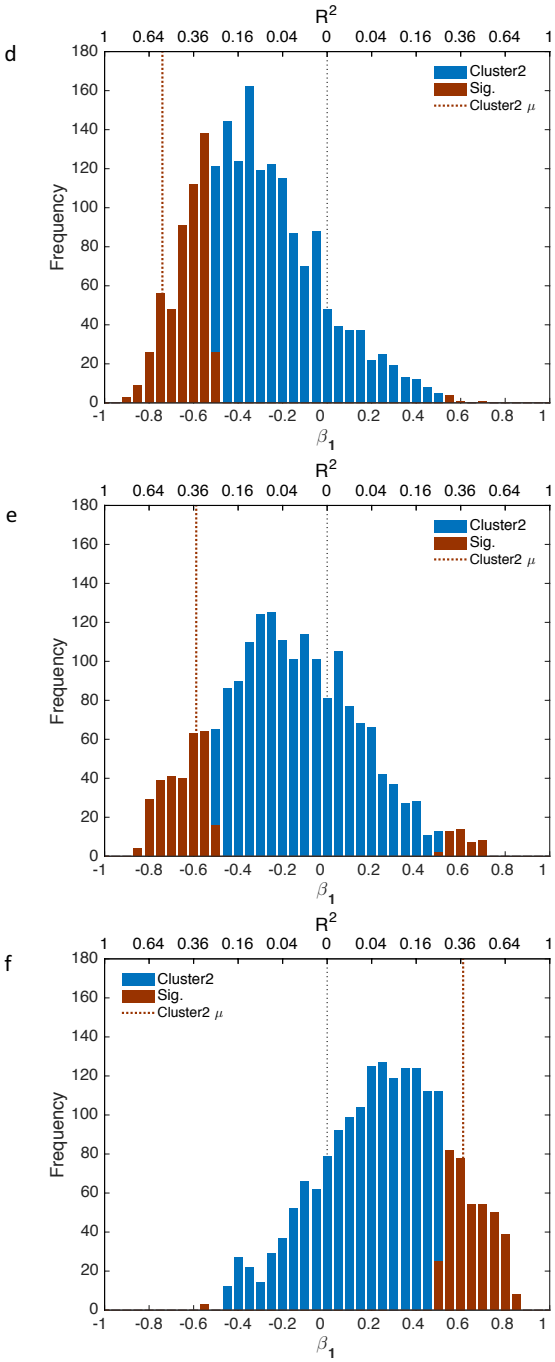


Fig. 2.8: Histograms showing the distribution of local scale (1° cells) relationships comprising a-b) Cluster 1 and c-f) Cluster 2 of Fig. 2.1. Each histogram represents the distribution of local-scale relationships within the respective cluster for the regression of a particular explanatory variable (as labelled) upon Chl. Bars represent the number of non-significant (blue) and significant (red) local-scale relationships at given ß$_1$ and R^2 values. For comparison, the dashed red line indicates ß$_1$ and R^2 values of the cluster-mean relationship. Only significant clusters-mean relationships are considered (Table 2.2). Nb. As all data are standardised to unit variance, β_1 and R^2 are proportional: r = β_1. Bars represent bin-averages of β_1 from -1 to 1 in steps of 0.05. The dashed black line delineates β_1=0 and thus distinguishes negative (left) and positive (right) relationships.

Fig. 2.8 show that cluster mean relationships identified in Table 2.2 are of the same sign as, and tend to have higher R^2 values than, the majority of local-scale relationships comprising each cluster. In other words, within these clusters, the relationship of the mean is stronger than the mean of relationships. Considered in connection with earlier results in Table 2.4, this implies that 'noise' tends to obscure relationships at the local scale, but that averaging cancels this noise out and helps to reveal underlying relationships at the larger scale. That is, the cluster-mean has improved the signal-to-noise ratio in comparison to the local scale. In this respect, the majority of cells comprising each cluster have the same underlying relationship and the difference between locally significant and non-significant cells is not due to changes in this relationship, but to the influence of localised variability that tends to cancel out over the larger area. Figure 2.8e is more complicated to interpret as wind velocities include positive and negative values. Thus, for standardised zonal wind velocity, positive values imply higher than usual eastwards wind speed (or less than usual westwards wind speed) and negative values imply higher than usual westwards wind speed (or less than usual eastwards wind speed). The resulting distribution is thus flatter than the other examples in Fig. 2.8 as both directions may have a positive influence upon Chl.

2.4 Discussion

Vertical mixing and Chl variability

When considered at different spatial scales, interannual Chl variability is significantly related to at least one-or-other of the proxies for vertical mixing included in this study within all areas of the North Atlantic. However, in contrast to other studies (Behrenfeld et al., 2006; Kahru et al., 2010; Siegel et al., 2013), none of these proxies stand out as being widely or universally related to Chl variability throughout the study domain, even if this is restricted to clusters and localities within the mid-to-low latitude (<45°N) North Atlantic (Figs. 2.6; 2.7; S2.5). Nevertheless, the lack of consistent relationships between the localities and/or clusters comprising this area does not preclude the presence of such large-scale relationships. Instead, this analysis shows that different relationships dominate at different spatial scales. At the spatial scales considered in this analysis, there is considerable spatial heterogeneity in relationships between Chl and proxies for vertical mixing throughout the North Atlantic. This spatial heterogeneity highlights the importance of considering different spatial scales if regional or local relationships are of interest.

Nevertheless, regardless of which vertical mixing proxy is related to Chl variability, the widespread presence of such relationships indicates that vertical mixing is an important driver of Chl variability throughout most of the North Atlantic (for a discussion of areas where Chl is not strongly related to vertical mixing please see §1.2 and §5). This is in contrast to recent studies that have indicated that vertical mixing plays a limited role in driving Chl (Dave and Lozier, 2013) or phytoplankton (Barton et al., 2015) variability. The discrepancy between these results is likely to be related to the use of clustering, as a technique to identify relationships, and yearly-averaged data, to avoid seasonality problems (see: §2.2.2), as well as the consideration of various explanatory variables in this study. For example, Fig. 2.8 shows that clustering helps accentuate relationships in comparison to the local scale (see: §2.4 *Clustering helps identify relationships*). However, this does not explain all discrepancies as this study also identifies widespread relationships between Chl and vertical-mixing proxies at the local scale (Fig. 2.6). By comparison, Dave (2014) finds significant relationships between Chl and stratification for only 11% of the mid-to-low latitude ocean when assessed at the local-scale (5° cells). Similarly, this analysis identifies that Chl is significantly related to yearly-averaged, or winter-minimum stratification within 14% and 17% of the mid-to-low latitude North Atlantic, respectively (Fig. 2.6; S2.5). (Although these percentages are likely to increase if averaged over 25 1° cells at a comparative scale.) However, if relationships with yearly-averaged SST are considered instead, then this percentage increases to

38% of the mid-to-low latitude North Atlantic (S2.5). Given that different relationships are stronger in different regions, then when all proxies for vertical mixing are considered, significant local-scale relationships with Chl are found to be widespread throughout the North Atlantic (Fig. 2.7).

Clustering helps identify relationships

The clusters used in this study identify regions of coherent Chl variability. Assuming these regions of similar variability are responding to a common forcing or forcings, the clusters may be isolating relationships with Chl and thereby acting to maximise their signal-to-noise ratio in comparison to averaging over arbitrary areas, or local-scale analyses. This is circumstantially evident in Fig. 2.8, wherein it is apparent that cluster-average relationships are stronger than the majority of local scale relationships within that cluster.

Furthermore, if clusters are isolating relationships in this way, they may act to emphasise different relationships at different spatial scales. Such a change in emphasis is seen when comparing relationships at both the two- and seven-cluster scales. This is well demonstrated by considering relationships with NAO. For example, the NAO is related to Chl in both clusters at the two-cluster scale. However, these large-scale relationships appear to be confined to smaller areas when the North Atlantic is divided into seven-clusters (c.f. Figs. 2.1, 2.4; Tables 2.2, 2.3). Although this could be thought to imply that the smaller-scale clusters have isolated the core regions driving the larger-scale relationships, this is unlikely to be the case. This is because the standardised data used in this analysis ensure an even-weighting amongst the time series. Furthermore, the relationship between Chl and NAO in K2 Cluster 1 is stronger than when 'isolated' to K7 Cluster 7 at the smaller scale (Tables 2.2, 2.3). Thus, the relationship between Chl and NAO in K2 Cluster 1 appears to be representative of its entire area, despite only being significantly related in one of the three K7 clusters that likewise comprise this area (c.f. Figs. 2.1, 2.4; Tables 2.2, 2.3).

This interplay of spatial scale with the emergence and obscuration of relationships is seen, in a comparison of the two- and seven-cluster partitions, for each of the other relationships between Chl and forcings considered in this study (c.f. Figs. 2.1, 2.4; Tables 2.2, 2.3).

The same is true in comparing larger-scale relationships to the local-scale. In many cases, strong relationships are apparent in clusters despite being poorly represented at the local scale (c.f. Figs.

2.1, 2.4, 2.7; Tables 2.2, 2.3). For each of the significant relationships identified in the two-cluster partition, comparing the distribution of local-scale relationships to the cluster-mean relationship reveals that, in each case, the signal-to-noise ratio of the cluster-average is improved over the local scale (Fig. 2.8). That the signal-to-noise ratio of cluster-mean relationships is improved over the local scale is also apparent in the cluster-mean relationships that are derived using only locally non-significant cells (Table 2.4). While this has not likewise been tested at the seven-cluster scale, the principle that averaging data together within regions of similar response helps to improve the signal-to-noise ratio appears to be general to this analysis. In this respect, and given that significant cluster-mean relationships tend to have comparable R^2 values to the locally-significant relationships within the cluster (Fig. 2.8), it may be interpreted that locally-significant relationships represent regions where there is less 'noise' affecting the signal, as opposed to regions where the relationships themselves are stronger.

A potential criticism of the cluster analysis used here may be the limited amount of total Chl variability explained by the cluster partitions themselves. The two- and seven-cluster patterns explain 15% and 26% of total local-scale variability within the North Atlantic, respectively (S2.4). This means that the majority of local-scale variability remains within the clusters and is unaccounted for by the cluster-mean time series. However, this analysis has demonstrated that relationships are spatially dependent. The very fact that clusters tend to represent area-average relationships that do not necessarily dominate the local scale, implies that their usefulness in explaining Chl variability may be poorly measured in terms of local-scale variability. Thus, while the majority of local-scale variability remains within clusters, most of this tends to cancel-out in favour of variability that is common to, or dominates amongst, the localities comprising each cluster. In the former case, much of this local-scale variability may be considered 'noise' at the cluster-average scale. As such, it may be more useful to assess how representative the cluster-mean relationships are of the localities comprising each cluster (e.g. Table 2.4; Fig. 2.8), than to assess them in terms of local-scale variance explained (S2.4).

The Importance of Spatial Scale

This study demonstrates that the spatial scale of analysis is important in assessing relationships with interannual Chl variability. It is shown that relationships between Chl and explanatory variables are spatially dependent: different relationships emerge depending upon the scale of analysis. As such, the scale of an analysis should be chosen in accordance with the objective of the analysis. This has implications for interpreting previous studies that give opposing results as to the

importance of stratification in driving Chl variability throughout the permanently-stratified global ocean. Behrenfeld et al. (2006) identified a strong relationship between stratification and both net primary production (R^2 = 0.73) and Chl (R^2 = not given) within the permanently-stratified ocean when domain-averaged time series were considered. By contrast, Dave and Lozier (2013), with corrections (Dave, 2014), found that only 11% of localities (5° cells) within this area had a significant relationship with $R^2 > 0.1$ between stratification and Chl. Although Dave and Lozier (2013) demonstrate that stratification is not strongly related to Chl variability at the local scale, this does not affect the conclusions of Behrenfeld et al. (2006), as it is shown here that smaller-scale processes tend to cancel out, or become less important, as the scale of analysis is increased. Indeed, Dave and Lozier (2013) recognise this, but argue that the lack of significant relationships at the local scale implies that the global-scale relationship identified by Behrenfeld el al. (2006) may be confined to a few small regions globally. This analysis has demonstrated that the cluster-average relationships identified for the two-cluster partition are not biased by the inclusion of strong localised relationships in this way, but are instead representative of the entire cluster (e.g. Table 2.4). Although Behrenfeld et al. (2006) did not use standardised data, it is suspected that this same principle will apply to the global-scale relationship they identify. However, to fully test the universality of a global-scale relationship like this, it is suggested that a bootstrap technique could be used to randomly sample and average the local-scale data, resulting in a distribution of relationships that would be reflective of larger scale averages (as determined by sample size), but unbiased by regionality or the presence of a few strong localised relationships.

NAO

The two-cluster pattern of interannual Chl variability in Fig. 2.1 appears to be related to the NAO. Chlorophyll variability is significantly related to the NAO in both clusters of this pattern, with the two having opposite signs of relationship. Furthermore, the similarity of this two-cluster distribution to the distribution of positive and negative relationships between Chl and the NAO at the local scale (Fig. 2.7) is indicative that the shape of these clusters is related to the NAO.

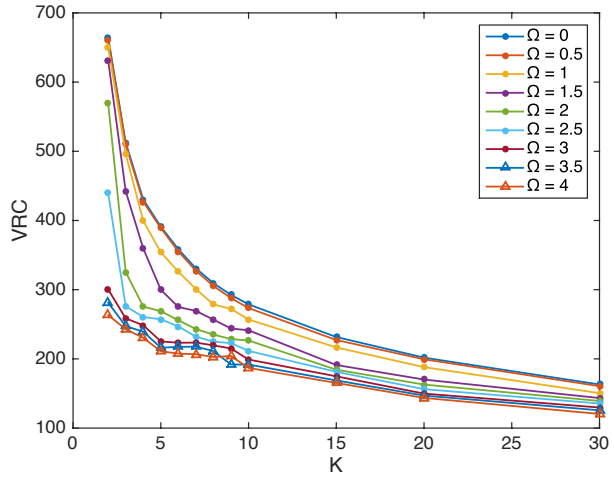
The pattern identified here resembles that previously identified by Patara et al. (2011), wherein the NAO is negatively related to Chl in the North Atlantic at latitudes between 35-40°N, and positively related to Chl at latitudes both north (35-60°N) and south (0-30°N) of this. This previous study demonstrated that the observed Chl anomalies were related to the NAO via vertical mixing and associated light and nutrient controls at the interannual time scale. However, these associations and relationships were identified using a coupled ocean-atmosphere with interactive

biogeochemistry model (Patara et al., 2011). This study identifies a strong and significant relationship between the NAO index and Chl variability using observational data. Previous observational studies have also identified relationships between Chl variability and the NAO index at the BATS mooring site within the Saragasso Sea (e.g. Bates, 2001; Lomas et al., 2010), or with phytoplankton colour index within the northern North Atlantic (e.g. Leterme et al., 2005). Yet others have failed to identify any widespread significant relationships between phytoplankton colour index and the NAO within the northern North Atlantic (Barton et al., 2015). So far as I am aware, the present study is first to present a large-scale synoptic pattern of Chl variability that is shown to be significantly related to the NAO using observational data. Furthermore, this study shows that wind speed, SST, stratification$_{WM}$ and MLD$_{WM}$ have similar NAO patterns (Fig. 2.3) to that of Chl, and that some of these (SST, stratification$_{WM}$), and other associated variables (MLD$_{YR}$, Zonal wind), are significantly related to Chl variability within the two clusters pattern (Table 2.2). As such, these vertical mixing proxies are likely to be responsible for mediating the relationship between NAO index and Chl.

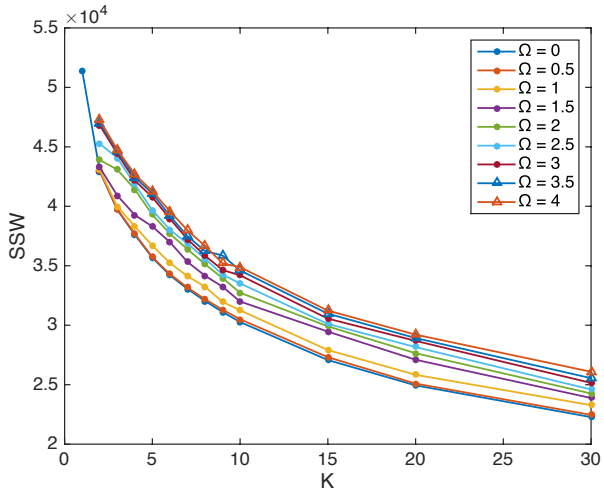
2.5 Conclusion

In this chapter, a large-scale bimodal pattern of Chl variability has been identified and shown to be related to similar patterns in SST, MLD, stratification and wind speed. The pattern is strongly correlated with the NAO index. Where area-mean relationships between Chl and these variables are present within the clusters of this pattern, these relationships are shown to be representative of the majority of localities comprising those areas. This is despite relationships having a limited representation within those localities when individually assessed. This discrepancy appears to be due to local-scale variability that can dominate locally, but tends to cancels-out when averaged over the larger cluster area. Moreover, spatial dependence is shown to be a general feature of relationships with Chl: different relationships emerge at different spatial scales.

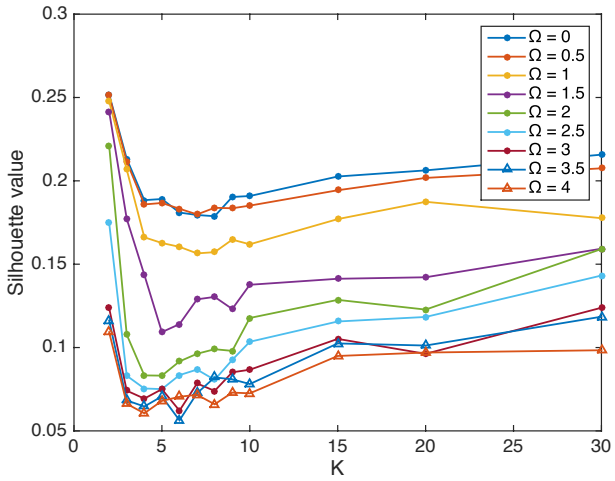
2.6 Supplementary Figures



S2.1: Plots showing variance ratio criterion (VRC) values over a range of cluster numbers (k = 2 to 30). Each line (as distinguished by colour and markings) represents variance in VRC with K for a given distant penalty value (Ω = 0 to 4; see legend). The maximum, or local maximum, VRC value is suggested as the optimal number of clusters to be chosen (Calinski and Harabasz, 1974). In this plot, the maximum VRC value occurs at k=2 for all values of Ω.



S2.2: Plots showing sum of squares within clusters (SSW) over a range of cluster numbers (k = 2 to 30). Each line (as distinguished by colour and markings) represents variance in SSW with K for a given distant penalty value (Ω = 0 to 4; see legend). SSW decreases with k and, as a rule-of-thumb, the optimal number of clusters may be identified as an inflection point or elbow in these plots. For comparison, sum of squares total (SST) has been plotted onto the Ω=0 line at k=1 (i.e. no clusters).



S2.3: Plots showing average silhouette-value (S) over a range of cluster numbers (k = 2 to 30). Each line (as distinguished by colour and markings) represents variance in S with K for a given distant penalty value (Ω = 0 to 4; see legend). According to this criterion, the maximum value of S represents the optimal number of clusters (Rousseeuw, 1987). Maximum values of S occur at k=2 for all values of Ω in this plot. However, values of S appear to increase with K after ~K=20 and these maxima may be exceeded at higher cluster numbers (K>30).

a.	1	2	N. Atl.*	
SSW ; SST*	2.2184	2.0771	5.0568	($x10^4$)
n	1846	1906	3752	
Variance	12.03	10.91	13.48	
SSB	---	---	0.7612	($x10^4$)
R^2 *(SSB/SST)*	---	---	0.15	

b.	1	2	3	4	5	6	7	N. Atl.*	
SSW ; SST*	1.0516	0.6515	0.4187	0.1989	0.3614	0.3996	0.6396	5.0568	($x10^4$)
n	882	631	402	303	371	466	697	3752	
Variance	12.02	10.44	10.60	6.72	9.93	8.71	9.27	13.48	
SSB	---	---	---	---	---	---	---	1.3355	($x10^4$)
R^2 *(SSB/SST)*	---	---	---	---	---	---	---	0.26	

S2.4: Statistics corresponding to a) the two-cluster partition in Fig. 2.1, and b) the seven-cluster partition in Fig. 2.4. Columns denote cluster numbers and the North Atlantic study domain (N. Atl.; excluding land and cells <200m depth). Statistics include sum of squares within cluster (SSW), sum of squares total in the North Atlantic (SST) and sum of squares between clusters (SSB). The number of 1x1° cells (n) and variance of each cluster and the North Atlantic are also given, as well as the fraction of SST explained by the cluster partition (R^2).

	N. Atl.	< 45°N
SST	1077 (29%)	934 (38%)
SST_{JFM}	380 (10%)	300 (12%)
Stratification	450 (12%)	344 (14%)
$Stratification_{WM}$	512 (14%)	426 (17%)
MLD	497 (13%)	215 (9%)
MLD_{WM}	608 (16%)	519 (21%)
NAO	755 (20%)	607 (25%)

S2.5: In correspondence with Figs. 2.6 and 2.7, number and percentage (in brackets) of 1x1° cells having a significant relationship between inter-annual Chl variability and explanatory variable (rows). For comparison with other studies, values are included for the mid-to-low latitude North Atlantic (< 45°N; N=2462), as well as the whole North Atlantic (N. Atl.; N=3752), wherein the North Atlantic refers to locations >200m depth within the study domain: 0 to 70°N; 70°W to 10°E.

Chapter 3: Vertical mixing and nutrient supply controls on interannual chlorophyll variability in the North Atlantic

3.1 Introduction

3.1.1 Overview

In the previous chapter, a bimodal pattern of interannual chlorophyll (Chl) variability was identified in the North Atlantic. Using observational data, this pattern and the Chl variability associated with it were found to be related to stratification, vertical mixing and the North Atlantic Oscillation (NAO). This study aims to identify the links between these variables and their association with this pattern using the additional variables available in a biogeochemical model.

3.1.2 Background

Understanding what drives interannual variability in primary production (PP) within the ocean is important in order to predict how marine PP will respond to global warming. It is expected that increased stratification will result in decreased PP in the nutrient-limited, mid-to-low latitude, ocean and increased PP in the light-limited, high latitude, ocean (Doney, 2006). This assumes that stratification is the primary control upon PP at interannual and longer-term time scales.

Climate-warming scenario biogeochemical models agree well with this hypothesis, projecting a general increase in stratification with reduced PP throughout most of the mid-to-low latitude ocean (< ~50-60°N/S) and increased PP at higher latitudes (Bopp et al., 2001; 2013; Cabré et al., 2015; Steinacher et al., 2010). Though discrepancies exist in terms of the position and extent of regions where changes in PP deviate from this broadscale pattern (Cabré et al., 2015; Steinacher et al., 2010)

The majority of observational records for marine PP are too short to distinguish the long-term trends expected of global warming from natural variability (Henson et al., 2010). Nevertheless, observational data are essential for assessing the strength of relationships between stratification and marine PP at interannual and longer time scales. A key question is thus: do changes in vertical mixing dominate interannual variability in marine PP?

3.1.3 Vertical Mixing and PP

A number of observational studies have identified relationships between chlorophyll (Chl) variability and changes in stratification (e.g. Behrenfeld et al., 2006; Corno et al., 2007), meteorological forcing (e.g. Follows and Dutkiewicz, 2001; Kahru et al., 2010; Menzel and Ryther, 1961) and SST (Siegel et al., 2013). Furthermore, the sign of these relationships between proxies for vertical mixing and Chl variability tends to change between subtropical and subpolar regions (Follows and Dutkiewicz, 2001; Kahru et al., 2010; Siegel et al., 2013). Thus, these studies are consistent with the hypothesis that vertical mixing drives Chl variability via nutrient supply in subtropical regions (Corno et al., 2007) and via the onset of stratification and light availability in subpolar regions (Sverdrup, 1953).

However, when similar observational studies are conducted at the local scale, using either globally-gridded datasets or in-situ moorings over similar time periods, they identify limited (Dave, 2014; Dave and Lozier, 2010), or non-significant (Barton et al., 2015; Lozier et al., 2011) relationships between mixing and interannual Chl variability. This discrepancy may indicate that the interannual relationship between stratification and PP is dependent upon spatial scale (§2). Thus, a commonality of weaker relationships at the local-scale may contribute to a stronger overall relationship at the larger scale. In the North Atlantic, interannual Chl variability exhibits both spatial autocorrelation and spatial heterogeneity, with regions of similar variability being separated from each other by differences in their overall variability. When relationships between interannual Chl variability and proxies for vertical mixing are considered within these regions, they tend to be stronger than when either considered locally or averaged over the North Atlantic as a whole (§2).

3.1.4 NAO and Chl variability in the North Atlantic

The North Atlantic Oscillation (NAO) is the dominant mode of climatic variability in the North Atlantic (Hurrell, 1995). It is defined as the pressure difference between the Azores High and Iceland Low pressure systems, and is associated with the position and strength of westerly winds and storm tracks in the North Atlantic. Changes in the strength of the NAO drive a basin-wide tri-pole pattern of changes in wind velocity, SST and surface pressure in the North Atlantic (Cayan, 1992).

The NAO has been associated with interannual Chl variability at the BATS site within the Sargasso Sea (Bates, 2001; Follows and Dutkiewicz, 2001; Lomas et al., 2010)and with both phytoplankton

community structure (Henson et al., 2012; Leterme et al., 2005; Zhai et al., 2013) and phytoplankton colour index (Barton et al., 2003; Leterme et al., 2005) in the northern North Atlantic. Using the synoptic view offered by satellite Chl data, positive NAO years have been shown to be associated with lower than average Chl concentrations between 35-45°N in the subtropics, and higher than average Chl concentrations north and south of this between 0-30°N and 45-60°N (Patara et al., 2011). This pattern is reproduced in a coupled ocean-atmosphere model with interactive biogeochemistry, in which the association between NAO and interannual Chl variability was found to be mediated via wind speed and MLD (Patara et al., 2011).

Ultimately, all of these relationships are considered to be mediated via vertical mixing (e.g. Follows and Dutkiewicz, 2001; Henson et al., 2012; Patara et al., 2011), and their influence upon nutrient concentrations and light availability (e.g. Patara et al., 2011). Thus, the relationship between NAO and Chl is closely related to the question of whether vertical mixing drives interannual Chl variability in the North Atlantic. As such, the NAO is also considered as a potential forcing factor in this study.

3.1.5 Aim of this chapter

This study uses a biogeochemical model to investigate how interannual relationships between Chl and proxies for vertical mixing are mediated. A basin-scale bimodal pattern of interannual Chl variability associated with the NAO has previously been identified in an observational study of the North Atlantic. The study also revealed strong relationships between interannual Chl variability and proxies for vertical mixing within the North Atlantic, both in association with the bimodal pattern and at smaller spatial scales (§2). Assuming that patterns and relationships identified in §2 are represented in the hindcast biogeochemical model used here, this study aims to identify how these relationships between physical forcings and Chl variability are mediated. While it is often assumed that interannual relationships between Chl and proxies for vertical mixing represent macro-nutrient and light limitations within the mid-to-low latitude and high latitude ocean, respectively (e.g. Behrenfeld et al., 2006; Kahru et al., 2010; Siegel et al., 2013; Wilson and Adamec, 2002). This analysis aims to identify the causal relationships mediating the indirect physical relations identified within the North Atlantic.

Thus, in addition to proxies for vertical mixing, this study also brings together all variables that have a direct influence upon PP within the model (nutrients, zooplankton, light and temperature) in order to assess how relationships with the former variables are mediated. However, in both

stages of assessing direct and indirect relationships with Chl variability, there may be multiple relationships with the explanatory variables considered. As such, this analysis aims to identify the dominant drivers of Chl variability at each stage, thus potentially tracing a causal link between NAO and vertical mixing proxies to Chl variability.

This study focuses on the following questions:

- Are observed relationships between vertical mixing and interannual Chl variability present in the hindcast biogeochemical model?

- Where such relationships are apparent, how are they mediated?

- Do large scale patterns of NAO and wind speed control interannual Chl variability in the North Atlantic?

3.2 Methodology

3.2.1 Consistency with observational analysis

This analysis is closely related to the observational analysis in Chapter 2, in which relationships between interannual Chl variability and physical processes of vertical mixing were identified using observational data. In the present chapter, the comprehensive dataset of a coupled physical-biological global ocean model is used to further investigate these relationships and to identify how relationships are mediated. This requires the model to replicate the relationships identified in the observational analysis. As such, the model is initially evaluated for consistency with relationships identified in the observational analysis (§2).

In order to remain as close as possible to the observational analysis, a hindcast simulation is chosen. This simulation forces the surface ocean with reconstructed atmospheric data to replicate the physical and biological response of the ocean to observed forcing. Furthermore, using a hindcast run ensures that the influence of large-scale climatic processes such as the NAO are represented in the model. For consistency, the analysis applied here also replicates that of §2 as closely as possible.

3.2.2 Data

All data used in this analysis are from a hindcast simulation of NEMO MEDUSA, a coupled physical–biological model. MEDUSA is an intermediate-complexity ecosystem model resolving two size classes of phytoplankton and zooplankton groups along with nitrogen (N), silica (Si) and iron (Fe) nutrient cycles (Yool et al., 2011). NEMO is an ocean general circulation model coupled to a sea-ice model (Madec, 2008). The hindcast simulation used here was forced at the surface using DFS 4.1 (DRAKKAR Group, 2007). This surface forcing dataset combines precipitation and downwards radiation fluxes from the CORE dataset (Large and Yeager, 2004), near-surface wind, air temperature and humidity from the ERA40 reanalysis dataset (Uppala et al., 2005) and uses bulk formulae to calculate air/sea and air/sea-ice energy and freshwater fluxes (Large and Yeager, 2004).

This hindcast simulation is available from 1988 to 2006 (inclusive). However, the first few years of this dataset were excluded from the analysis due to spin-up effects. Instead, this analysis was based upon years 1993-2006, as two-cluster and seven-cluster patterns of Chl variability became more stable after this date (not shown).

3.2.3 Variables

The variables extracted from NEMO MEDUSA and used in this analysis are listed below. All data were extracted as monthly averages and subsequently converted into yearly-averaged or yearly-metric values.

Variable	Processing
Chl	Surface Chl was calculated as the sum of surface Chl concentrations from both small and large phytoplankton groups.
SST	The SST field was yearly averaged.
Stratification; $Stratification_{WM}$	Salinity and Temperature fields were re-gridded to 1° horizontal resolution and used to calculate 0-200m density difference according to the Sea Water Routines (see: §2). The resulting monthly stratification fields were then averaged by year, for yearly-average stratification (stratification), or taken as minimum wintertime (Dec-Feb) values, for winter-minimum stratification ($stratification_{WM}$).
MLD; MLD_{WM}	MLD, defined within the model in terms of turbulent kinetic energy, was averaged by year, for yearly-averaged MLD (MLD), or taken as maximum wintertime (Dec-Feb) values, for winter-maximum MLD (MLD_{WM}).
Wind speed	Wind speed field was yearly averaged.
N	Surface values were yearly averaged.
Si	Surface values were yearly averaged.
Fe	Surface vales were yearly averaged.

Zooplankton	Total surface zooplankton was calculated as the sum of yearly average small and large surface zooplankton biomass.
PAR	Only surface PAR values (PAR_0) were available in the saved output of NEMO MEDUSA. Therefore, to calculate MLD average PAR, a profile of PAR values were calculated for each vertical grid cell within the MLD and then average together according to their respective heights and the height of the lower cell within the MLD. Light is attenuated with depth due to scattering and Chl absorption. Therefore, PAR at the bottom of a cell (PAR_Z) is equal to PAR at the top of the cell (PAR_{Z-H_Z}; equivalent to PAR_0 for surface cell) minus its attenuation within the cell. This attenuation is a function of the cell height (H_Z), Chl concentration within the cell (Chl_Z) and the attenuation coefficients of Chl (α_{Chl}) and water (α_W): $$PAR = \sum_{Z=0}^{MLD} H_Z' \cdot \left\{ PAR_{Z-H_Z} - \left[(H_Z \cdot \alpha_{Chl} \cdot Chl_Z) + (H_Z \cdot \alpha_W) \right] \right\}$$ where H_Z' is is the normalised height of each cell, such that: $$\sum_{Z=0}^{MLD} H_Z' = 1$$ and PAR_{Z-H_Z} is PAR at the top of the cell corresponding to Z. This is equivalent to PAR_0 in the top cell, PAR_0 minus the attenuation of the top cell for the second cell, PAR_0 minus the integrated attenuation of the top and second cell for the third cell and so on.
NAO	NAO index was not available as a save output of NEMO-MEDUSA. However, as the hindcast simulation is forced with reanalysis data, this should be equivalent to observed NAO index. As such, the monthly NAO index for corresponding years (1993-2006) was downloaded from NOAA's Climate Prediction Centre: http;//www.cpc.ncep.noaa.gov/. Winter months (December, January, February, March) of this index were then averaged together to give a yearly time series of wintertime NAO index.

Meridional and Zonal wind velocities are not available in the saved output of NEMO-MEDUSA. As such, unlike the previous observational analysis (§2), these variables are not included in this model analysis.

3.2.4 Rescaling data

All variables used in this analysis were rescaled to 1x1° resolution for comparison to results from the observational analysis in §2. The NEMO MEDUSA simulation was run at ¼° resolution and so rescaling involved averaging over each of the four cells that comprise a degree of latitude and longitude. Where one or more of the cells being averaged represented land, the average was taken from the other non-land cells.

3.2.5 Clustering

As in §2, the Chl data were clustered using the K-means algorithm to sort the time series into groups according to the similarity of their interannual Chl variability. A distance penalty coefficient (Ω) is implemented to also encourage the data to be grouped according to geographical proximity (c.f. §2.2.4). In accordance with previous analysis (§2), the number of clusters (k) and the distance penalty coefficients used were: k = 2 with Ω = 0 and k = 7 with Ω = 3.

3.2.6 Assessing relationships between variables

This analysis aims to identify the strength of relationship between Chl variability and processes of vertical mixing, and to determining how such relationships are mediated. To this end, the potential explanatory variables are divided into two groups: variables that are directly related to Chl variability in NEMO MEDUSA (direct-effect variables) and physical variables relating to vertical mixing (physical variables) which are indirectly related to Chl variability. Using these two groups of variables, relationships are considered in three stages:

1. Each of the variables in both groups are compared to Chl using simple linear regression. This provides an overview of relationships with Chl, as well as a basis for comparing the relationships identified in NEMO MEDUSA to those identified in observational data (§2).
2. According to the sign of relationships in stage 1, the direct-effect variables are narrowed down by excluding those that have non-causal relationships with Chl (see: §3.2.7). The remaining direct-effect variables are then considered in a multiple regression framework (see: §3.2.8) in order to distinguish which is most important in describing Chl variability within each cluster.
3. The dominant drivers (see: §3.2.9) of Chl variability identified in stage 2 are considered as response variables for each cluster. The indirectly-related variables (SST, Stratification, MLD and wind speed) are then regressed upon these directly-related variables in a multiple regression framework (see: §3.2.8) for each cluster. If strong relationships are identified in stages 2 and 3, this should imply how relationships between indirectly-related variables and Chl identified in stage 1 are mediated.

3.2.7 Causal or casual relationships based upon sign

As they represent a direct relationship with Chl, the directly-related variables (N, Si, Fe, PAR, SST and Zooplankton) can be separated into those that may be causally related to Chl and those that are not. This distinction is based upon the sign of relationship with Chl. For example, an increase in nutrients cannot directly result in a decrease in Chl concentration within the model (Yool et al., 2011). A negative relationship between nutrient concentration and Chl would therefore imply a non-causal relationship. The indirect variables cannot be distinguished in this way as they may be positively or negatively related to Chl via opposing influences upon directly-related variables.

	Causal Relationship
N	+'ve
Si	+'ve
Fe	+'ve
PAR*	+'ve
SST	+'ve
Zooplankton	–'ve

*The model does not account for the possible effects of photoinhibition (Yool et al., 2011).

Table 3.1: Sign of causal relationships between directly-related explanatory variables (rows) and Chl: given as positive (+'ve) or negative (–'ve) relationships.

Table 3.1 lists signs of relationships that would imply causality between Chl and each of the directly-related explanatory variables. This list is used to exclude non-casually related variables from multiple regression analysis, as outlined in §3.2.6.

3.2.8 Collinearity and Relative importance

The explanatory variables in this analysis are likely to be collinear. For example, increased wind speed may result in increased MLD and nutrients, as well as decreased stratification, PAR and SST. In this case, all these variables may be associated with Chl variability, but would tend to overlap each other in terms of the variance in Chl that they explain. That is, their unique contributions to explained Chl variability would likely be less than that indicated by their coefficients of determination (R^2) when regressed upon Chl individually. The presence of collinearity amongst explanatory variables is obvious in the case where individual R^2 values sum to more that 1.

Including the variables together in a multiple linear regression will limit the total R^2 value to the unique amount of variance explained in Chl. However, the proportioning of this R^2 value amongst the explanatory variables remains a problem. The sum of squares (SS) explained can be calculated for each regressor. However, the attribution is order specific with the first regressor considered free to fit all the variance in Chl, and the second regressor then confined to fit the residual variance and so on (see: Grömping, 2006). Furthermore, (multi)collinearity amongst explanatory variables in a linear model can result in incorrect coefficient estimates that change erratically with small changes in the model/data (e.g. Tu et al., 2005).

Thus, irrespective of whether comparisons are made between single regressions or within a multiple regression, collinearity leads to two potential problems. Firstly, it is unclear how

explained variance is proportioned amongst the explanatory variables. Secondly, the magnitude and sign of coefficient estimates may incorrect.

In the latter case, a number of techniques (including ridge regression and variable reduction by e.g. principle component analysis) exist for dealing with collinearity in designing robust predictive models (e.g. Dormann et al., 2013). However, this analysis is primarily concerned with identifying the most important driver(s) of Chl variability amongst explanatory variables. Relative Importance (RI) describes a collection of techniques aimed specifically at addressing the first of these two problems (Grömping, 2006; 2007). The technique used here involves averaging the Sum of Squared Errors (SSE) for each regressor over all orders of the regressors (Lindeman et al., 1980, pp. 119). In a comparison of different methods, this approach was recommended by Grömping (2006):

$$RI(x_k) = \frac{1}{p!} \sum_{r\ permutations} R^2([x_k] \cup s_k(r)) - R^2(s_k(r)) \quad (3.1)$$

where r symbolises different orders of the regressors $(x_1, \dots x_p)$, x_k denotes the k^{th} regressor and $s_k(r)$ all preceding regressors in the set r.

Equation 3.1 calculates the average additional R^2 contributed by each regressor over all possible orders of the regressors: from it being the only regressor, through to all combinations of it being 1 of p regressors. In this way, the technique also allows a crude check of the stability of coefficient estimates to (multi)collinearity to be made: the coefficient estimates can be checked for stability against these various orderings of the regressors.

3.2.9 Confidence intervals

In addition to R2 and RI values for multiple-regressions upon Chl, 95% confidence intervals for these values are calculated by bootstrap resampling of the data. Bootstrapping is recommended for use with RI as parametric tests are not easily applied to Eq. 3.1 (Grömping, 2006). This is done by randomly resampling the explanatory and response variables. Fourteen years of data (1993 – 2006) are included in this analysis, so bootstrap resampling will involve randomly sampling with replacement 14 years from these data (including paired years of both explanatory and response variables) and repeating the analysis upon the randomised time series. This is done repeatedly (n = 200) and 95% confidence intervals are derived from the spread of results (Grömping, 2006).

3.2.10 Identifying dominant drives of Chl variability

As outlined in §3.2.6, directly-related variables that are causally related to Chl (§3.2.7) are included in a multiple regression using RI to account for collinearity (§3.2.8). This proportions the overall R^2 for the multiple regression amongst the explanatory variables in terms of RI values, which are equivalent to R^2 in meaning and sum to overall R^2. Once collinearity has been accounted for in this way, the dominant drivers (RI > 0.5) of Chl variability are used as response variables in the next stage of analysis (§3.2.6). In the case that no directly-related explanatory variable has RI > 0.5, then either the directly-related variable with highest independent R^2 is used as the response variable, or no response variable is assigned to that cluster. This depends upon the strength of collinearity amongst the explanatory variables and the strength of their independent relationships with Chl. If one or more of the explanatory variables have high individual R^2 values (R^2 > 0.5) when regressed upon Chl, but are strongly collinear and so do not have RI > 0.5, then the explanatory variable amongst these with highest independent R^2 value is used as the dependent variable in the next stage of analysis (§3.2.6). In this case, the indirectly-related variables would be considered to be regressed upon the collinear directly-related variables collectively. This provision allows for the situation where variables are strongly related to Chl but also strongly collinear and so have low RI values.

If the RI values of explanatory variables are close to one another, it is likely that their 95% confidence intervals will overlap. While this would imply that the RI values are not significantly different, it does not affect the rankings of RI values in terms of selecting direct-effect variables for the next stage of analysis (§3.2.6). This is because the same years of data (1993-2006) are being used in both stages of analysis, and so only the RI values that related to these time series are relevant.

3.2.11 Yearly-averaged vs yearly-metric variables

As RI multiple regression is aimed at inter-comparing the explanatory variables, only yearly-average variables are included in these regressions. While yearly-metrics of some explanatory variables (e.g. $\text{stratification}_{WM}$, MLD_{WM}; c.f. §3.2.3) may be better related to yearly-averaged Chl variability, 'cherry-picking' how particular explanatory variables are defined in this way may result in them being disproportionately represented within the RI regression. In order to ensure equality amongst the explanatory variables, they should either all be included as yearly-averages, or all be included as best-related yearly-metrics. However, given the large number of combinations

possible, identifying such best-related yearly metrics for each of the explanatory variables within each cluster may require a time-consuming and extensive analysis. As such, for simplicity, only yearly-average variables are included within the RI multiple regressions.

3.3 Results

3.3.1 Basin-Scale analysis

Two-cluster pattern of North Atlantic Chl

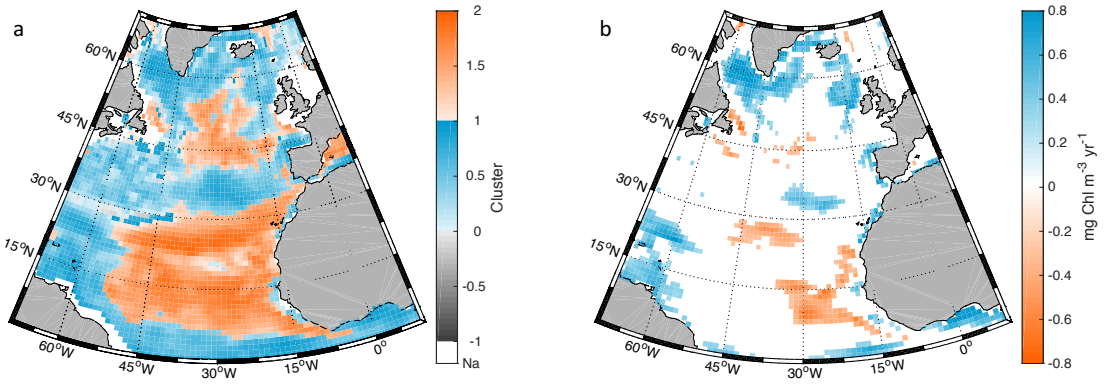


Fig. 3.1: a) Two-cluster distribution and b) linear trends of yearly-average surface Chl variability in the North Atlantic from NEMO MEDUSA 1993-2006. Only statistically significant values (P<0.05) are plotted in (b). Notice that in comparison to other figures, the colourbar in (b) has been reversed to highlight the similarity with clusters. In a) Colours denote cluster number: Cluster 1 (blue), Cluster 2 (orange). Where present, grey indicates locality is anti-correlated with cluster. Shelf sea localities less than 200m depth are omitted (Na; white). Shading denotes strength of correlation (or anti-correlation) with cluster, from light (R^2 = 0) to dark (R^2 = 1).

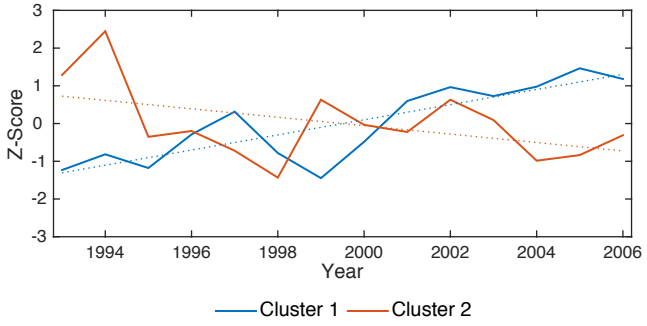


Fig. 3.2: Cluster-mean time series (solid line) of standardised surface Chl concentration for Clusters 1 (blue) and 2 (orange) in Fig. 3.1a. Dotted lines represent simple linear trends fitted to time series for Clusters 1 (blue) and 2 (orange).

Fig. 3.1 shows the two-cluster distribution of interannual surface Chl variability derived from NEMO MEDUSA. This is broadly similar to the two-cluster distribution of the observational data

(Fig. 2.1), despite the discrepancy in years represented by the two datasets (1998-2012; Fig. 3.1, 1993-2006). However, unlike the observational data, the two clusters are not significantly anti-correlated (R^2 = -0.16; p = 0.163), even though they have opposing trends in Chl concentration (Fig. 3.2). In this respect, a comparison of Fig. 3.1a and b shows that the two clusters tend to separate the data into regions of negative (cluster 1) and positive (cluster 2) Chl trend.

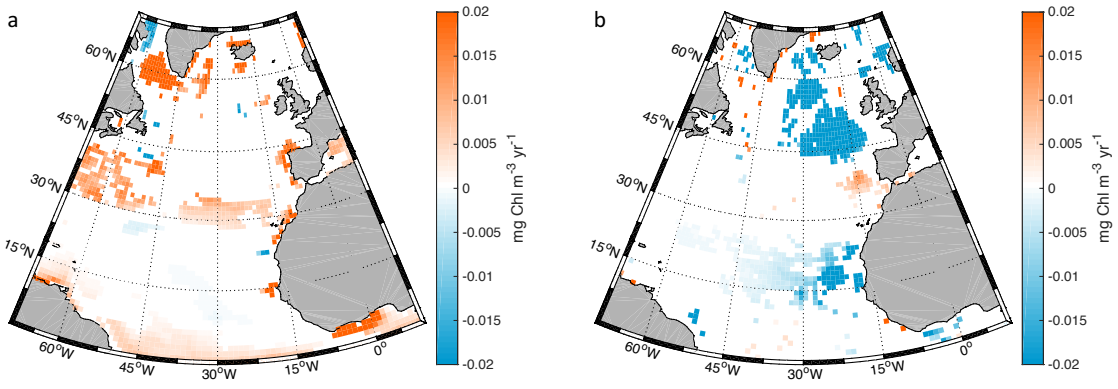


Fig. 3.3: Local scale (1x1°) trends in surface Chl concentration between 1998-2006 (inclusive) for a) NEMO MEDUSA hindcast simulation and b) ESA CCI ocean colour.

Fig. 3.3 shows the distributions of Chl trends in the North Atlantic for the overlapping years of observational data and model output. A comparison of the two shows that they broadly agree on the sign of trend within the regions defined by the two clusters in Fig. 3.1a. However, the two datasets differ in terms of how strongly they emphasise the positive or negative trends associated with these two regions: the model output tends to emphasise the positive trends that corresponds with Cluster 1 in Fig. 3.1, while the observational data more strongly emphasise the negative trends that correspond with Cluster 2.

Fe trend

A potential complication with this simulation of NEMO MEDUSA lies in the fact that Fe concentration is found to be increasing throughout the North Atlantic. Although Fe concentration is not strongly limiting to phytoplankton growth within the study domain (Fe concentration is above the half saturation constant for both small and large phytoplankton size-classes throughout the North Atlantic in 1993 and all subsequent years [S3.7]), the increasing trend will inevitably result in a limited increase in phytoplankton growth because of the asymptotic nature of the nutrient limitation terms in the model (Yool et al., 2011). This may result in a spurious correlation

between Chl and Fe in clusters where Chl variability is also dominated by an increasing trend. In this situation, it is unlikely that Fe would be responsible for the overall trend in Chl, but this could not be easily discounted. A potential solution to this problem would be to detrend all the time series. However, as this analysis is aimed at identifying the drivers of interannual Chl variability, including its trends, this is not a complete solution. As such, both with-trend and detrended relationships will be considered in this analysis. With-trend relationships can then be compared to detrended relationships for consistency: i.e. is the relationship robust to detrending? If a coefficient of relationship changes sign, or becomes insignificant, when detrended, it is unlikely to dominate the with-trend relationship. While it is possible to consider relationships in terms of nutrient limitations (i.e. nutrient concentrations in comparison to half-saturation constants), these limitations would tend to reflect mean conditions. Growth limitation terms are multiplicative within this model (Yool et al., 2011), and interannual variability does not necessary reflect primary-limiting nutrients, or indeed nutrients at all. (For example, PP in a given locality may be primarily limited by e.g. N, but in the absence of strong N variability may vary in accordance with some other explanatory variable, e.g. temperature.) As this analysis is concerned with interannual variability, relationships are not ranked in terms of nutrient concentrations.

Relationships with Chl in Clusters 1 and 2

	{With Trend}		*{De-trended}*	
	1	2	1	2
SST	+0.00 0.867	**-0.70** 0.000	**-0.36** 0.024	**-0.74** 0.000
Stratification	**-0.33** 0.031	**-0.78** 0.000	**-0.29** 0.046	**-0.83** 0.000
$Stratification_{wm}$	**-0.56** 0.002	**-0.68** 0.000	**-0.37** 0.022	**-0.63** 0.001
MLD	+0.27 0.057	**+0.52** 0.004	**+0.34** 0.029	**+0.59** 0.001
MLD_{wm}	+0.01 0.756	**+0.61** 0.001	+0.16 0.115	**+0.68** 0.000
Wind	+0.08 0.323	**+0.49** 0.006	+0.32 0.036	**+0.41** 0.014
NAO_{djfm}	**-0.30** 0.043	+0.19 0.118	-0.09 0.288	+0.07 0.349

Table 3.2: R^2 and associated p-values (beneath) for the strength of relationship between interannual Chl variability and explanatory variables (rows). Columns denote the corresponding clusters (1 or 2) and whether or not variables were detrended before regression. A positive or negative symbol before each R^2 value denotes the sign of relationship. For emphasis, statistically significant R^2 values ($p<0.05$) are given in bold type and shaded according to the sign of relationship: blue (negative) or red (positive).

As the model has a similar two-cluster distribution of interannual Chl variability to that of the observational data, the cluster-mean relationships in Table 3.2 can be compared to those of their respective clusters in the observational analysis for consistency (Table 2.2). Both analyses considered relationships between Chl and SST, Stratification, $Stratification_{WM}$, MLD, MLD_{WM}, wind speed and the NAO. Regardless of whether the with-trend or detrended relationships in Table 3.2 are considered, the model tends to be consistent with the observational analysis both in terms of the sign of these relationships and in the fact that, with the notable exception of the NAO, all the relationships are stronger in Cluster 2 than in Cluster 1. Relationships also tend to be stronger in the model than in the observational analysis. As such, the model has a greater number of statistically significant relationships, including all of those that were statistically significant in the observational analysis: Cluster 1 (NAO); Cluster 2 (SST, $Stratification_{WM}$, Stratification, MLD_{YR}).

The main discrepancies between the two analyses are as follows: the relationship with the NAO was stronger in the observational analysis, both for Cluster 1, when compared to the model including Chl trend, and for Clusters 1 and 2, when compared to the model excluding Chl trend; the relationship with MLD_{WM} in Cluster 1 was also stronger in the observational analysis, both compared to the model including and excluding Chl trend. The only other minor discrepancy was that the sign of the relationship with MLD was negative in the observational analysis. However, this is probably due to the poor fit of this insignificant relationship in the observational analysis, as a positive relationship is identified if a robust regression is fitted instead (not shown).

Thus, both the observational and model analyses are consistent in identifying relationships between the physical processes of vertical mixing, and interannual Chl variability within the two-cluster patterns. This is especially true for Cluster 2, where these relationships tend to be significant in both analyses. These similarities imply that the model is replicating the biophysical relationships apparent in the observational data, and that it may be interrogated further to infer how the observed indirect relationships are mediated.

Two-cluster patterns of biophysical variables

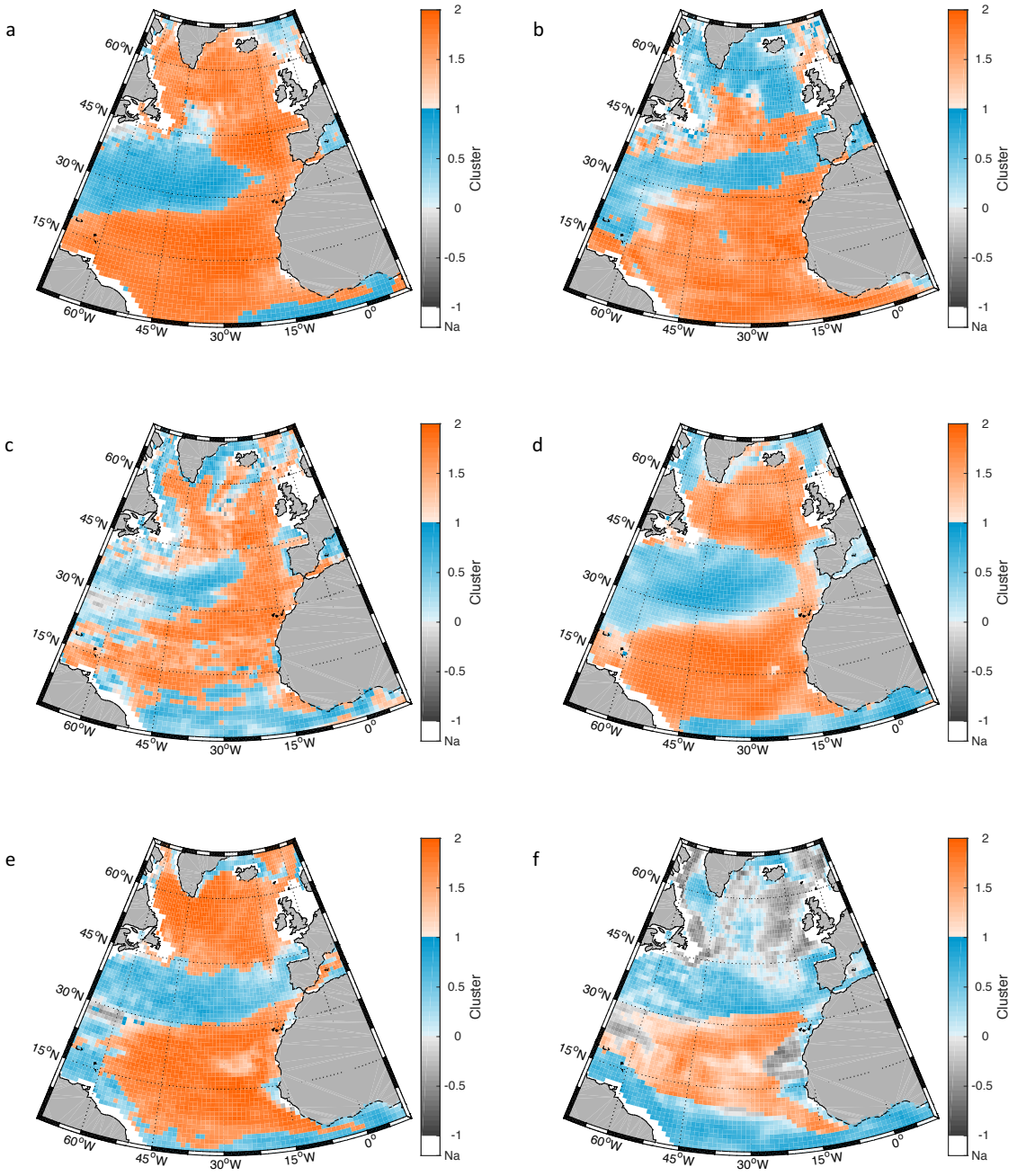


Fig. 3.4: Two-cluster distributions of a) SST, b) Stratification, c) MLD, d) wind speed, e) N and f) zooplankton variability in the North Atlantic from NEMO MEDUSA 1993-2006. Colours denote cluster number: Cluster 1 (blue), Cluster 2 (orange). Grey indicates locality is anti-correlated with cluster. Shelf sea localities less than 200m depth are omitted (Na; white). Shading denotes strength of correlation (or anti-correlation) with cluster, from light (R^2 = 0) to dark (R^2 = 1).

Fig. 3.4 shows two-cluster patterns for explanatory variables in Table 3.2. Similar plots for the variables not shown are given in supplementary figures (S3.1). Despite noticeable discrepancies, a comparison of these patterns to that of Chl in Fig. 3.1 reveals a broadscale similarity wherein a zonal band of the subtropics (~23.5-40°N) and fringes of the western, northern and southern North Atlantic are separated from large parts of the Tropical and Subpolar North Atlantic. The plot for Zooplankton is more complicated as parts of Cluster 2 are anti-correlated with its mean variability. This probably represents the need for additional clusters in representing zooplankton variability in this way. Nevertheless, the regions of anti-correlation in this plot are part of Cluster 2 and are thus more similar, in terms of minimising euclidian distances, to Cluster 2 than Cluster 1. As such, this pattern also represents a similar separation of the clusters as the other plots.

The similarity of the physical and biological variables in Fig. 3.4 to that of Chl variability in Fig. 3.1 implies that these variables are not only correlated with Chl (see: Table 3.2), but are also related to the bimodal pattern of Chl seen in Fig. 3.1.

In order to identify the main drivers of Chl variability within this pattern, and to determine how such relationships are mediated, this analysis considers these relationships with Chl in two parts. First of all, variables with the potential to directly influence Chl concentration (N, Si, Fe, Zooplankton, SST, PAR) are considered in order to identify dominant ($R^2 > 0.5$) directly-related driver(s) of Chl variability. Secondly, proxies for vertical mixing, which are not directly related to Chl variability (SST, Stratification, MLD, wind speed), are assessed for their relationships with the direct drivers. Nb. the NAO is not included amongst the explanatory variables in the latter section, as it is not directly related to nutrient variability. However, changes in NAO may correspond with changes in wind speed, SST, stratification and MLD, which are represented.

3.3.2 Direct drivers of Chl variability

This section aims to identify to what extent each of the directly-related explanatory variables (N, Fe, Si, Zooplankton, SST and PAR) are responsible for driving interannual Chl variability within the Clusters of Fig. 3.1a. These variables are first narrowed down by excluding non-causal associations, according to the criteria outlined in §3.2.7 and Table 3.1, and relationships in S3.2. The remaining variables are then assessed for their individual contributions to Chl variability using Relative Importance.

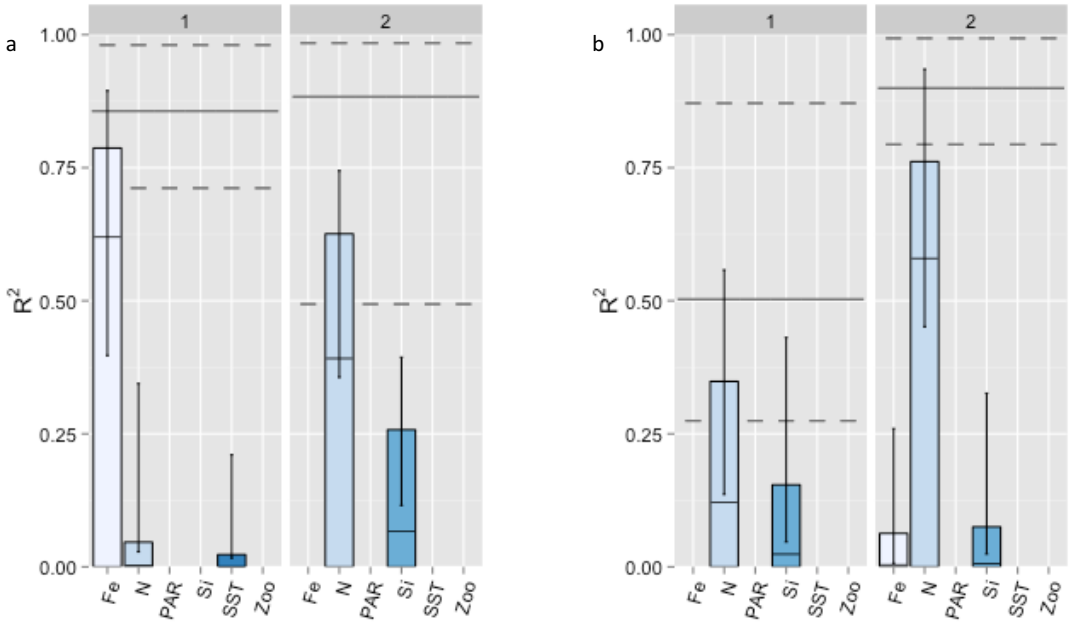


Fig. 3.5: Sections titled 1 and 2 correspond to clusters in Fig. 3.1a and each give the results of a multiple regression, including the explanatory variables for which blue bars are plotted (as labelled), upon Chl for a) with trend and b) detrended cluster-mean time series. Solid horizontal lines spanning each section indicate the R^2 of each multiple regression. Dashed horizontal lines above and below this represent associated 95% confidence intervals. Blue bars (shaded as per explanatory variable) represent the contribution of each individual explanatory variable to the total R^2 in terms of RI. Whiskers represent 95% confidence intervals of these RI values. Solid horizontal lines within each bar represent the unique contribution of that variable to total R^2, in terms of squared partial-correlations. (Nb. No confidence intervals are given for squared partial correlations.)

Fig. 3.5 shows the relative contribution (i.e. after accounting for collinearity) of each regressor to interannual Chl variability within the clusters of Fig. 3.1a. Nutrients (N, Fe and Si) and SST are the only variables found to be causally related to interannual Chl variability. Of these, Fe (N) and N (N)

appear to be the most important drivers of Chl variability for with-trend (detrended) time series in Clusters 1 and 2, respectively. However, these rankings are not statistically significant in clusters wherein the 95% confidence intervals of these RI values overlap with those of other drivers. This is true of Cluster 2, for time series with trend, and cluster 1, for detrended time series. In both these cases, N and Si are jointly the most important drivers.

In Cluster 1, Fe appears to be the dominant driver of Chl variability for time series including trends. However, this could be a spurious relationship as Fe is no longer causally related to Chl when time series are detrended (i.e. the relationship is not robust to detrending). Thus, the two are only related due to a coincidence of positive trends. This is also true of the weaker relationship with SST. Although all variables that have a direct influence upon Chl have been considered in this section, an overall positive trend in Chl could be driven by different relationships within the area of this cluster. As such, the with-trend relationship with Chl in this area may be better considered at the regional scale (§3.3.5). By contrast, the relationships with N and Si in Cluster 1 get stronger when time series are detrended. Thus, these results imply that N and Si drive the year-to-year variability in Chl within this cluster.

As an aside, given that year-to-year variability in Fe is small in comparison to its trend (S3.3), it may be argued that Fe could drive a trend in Chl while another variable drives its year-to-year variability. However, this would assume that the other variable in question (which is more strongly related when detrended) did not also have a trend. Given that N and Si drive the year-to-year variability of Chl in this cluster (Fig. 3.5b) and both have negative trends (S3.3.), this possibility can be disregarded.

In Cluster 2, N and Si are the dominant drivers of Chl variability for time series including trends. As the 95% confidence intervals of these two variables overlap one another, they are not significantly different in terms of RI. However, when time series are detrended, the explanatory power of N increases while that of Si is reduced. Fe is also revealed as a driver of Chl variability in Cluster 2 when time series are detrended. However, both Fe and Si have a small influence upon Chl ($RI < 0.1$) in comparison to N ($RI > 0.75$) when time series are detrended. This implies that the relationship between Si and Chl in Cluster 2 is largely attributable to a coincidence of negative trends, while N drives the majority of year-to-year variability in Chl within Cluster 2.

In neither cluster do Zooplankton or light availability appear to control interannual Chl variability. With the exception of Cluster 1 for time series including trends, this is also true of SST. As stipulated in §3.2.7, these relationships are considered to be non-causally related to Chl as per the sign of relationship (c.f. Table 3.1). Thus, while Zooplankton variability is strongly related to interannual Chl variability in both clusters (S3.2), these relationships appear to represent bottom-up control, with Zooplankton biomass varying in response to that of phytoplankton and Chl.

3.3.3 Indirect drivers of Chl variability

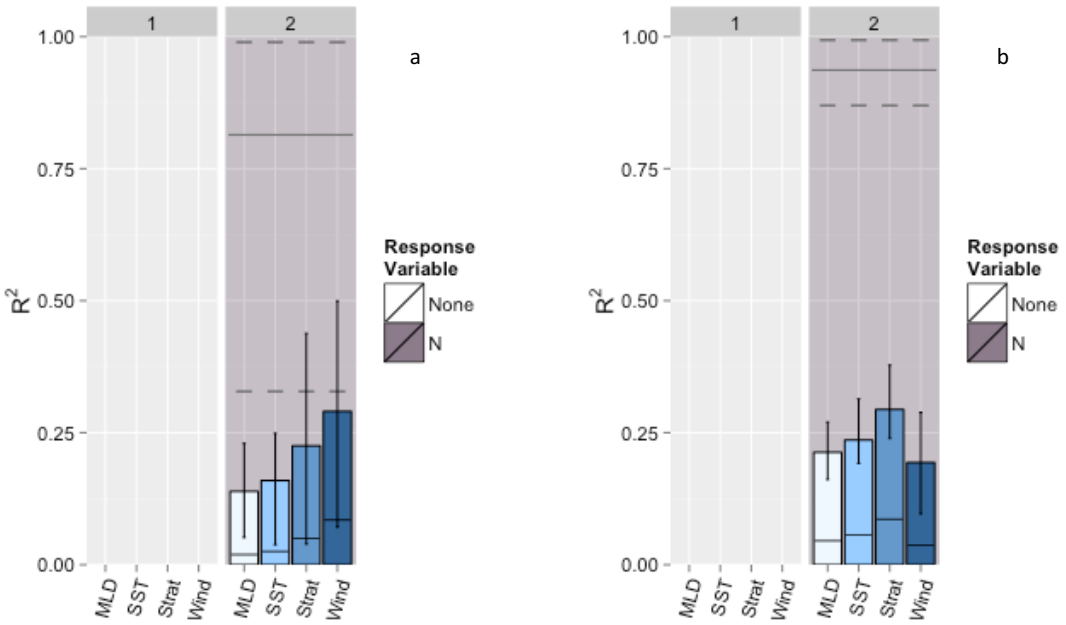


Fig. 3.6: Sections titled 1 and 2 correspond to clusters in Fig. 3.1a and each give the results of a multiple regression, including explanatory variables identified by blue bars (as labelled), upon the response variable identified by background colour (see legend) for both a) with trend and b) detrended cluster-mean time series. In each case, the response variable corresponds to the direct-driver of Chl variability identified in §3.3.2: if no direct driver was identified then a regression is not included in this section. Solid horizontal lines spanning each section indicate the R^2 of each multiple regression. Dashed horizontal lines above and below this represent associated 95% confidence intervals. Blue bars (shaded as per explanatory variable) represent the contribution of each individual variable to the total R^2 in terms of RI. Whiskers represent 95% confidence intervals of these RI values. Solid horizontal lines within each bar represent the unique contribution of that variable to total R^2, in terms of squared partial-correlations. (Nb. no confidence intervals are given for squared partial correlations.)

This section considers to what extent variability in the directly related variables identified in §3.3.2 can be explained by proxies for vertical mixing. The dominant drivers (RI > 0.5) of Chl variability identified in Fig.4.6 are used as response variables in this section in order to determine how well their variability can be explained by proxies for vertical mixing. In most cases, N and Fe are the dominant drives of Chl variability in Fig. 3.5. The exception is Cluster 1 for detrended time series, wherein N and Si were identified as co-drivers of Chl variability with $R^2 > 0.5$ when collectively regressed upon Chl. However, neither of these variables had RI > 0.5 or $R^2 > 0.5$ when individually considered (S3.2). As these co-drivers tend to vary independently of one another, multiple causal pathways must exist in driving detrended Chl variability within this cluster. As such, no response variable is included for detrended Cluster 1 in this section. Furthermore, Fe is

omitted as a potential driver of with-trend Chl variability in this section. The near-linear trends and surface accumulation of Fe reveal that it is driven by high atmospheric deposition rates within the North Atlantic in this hindcast simulation (S3.3, S3.6b). Increasing Fe trends are thus not attributable to vertical mixing and cannot be driven by the explanatory variables considered in this section. As such, the relationship with Fe in Cluster 1 with trend (Fig. 3.5) is also excluded from this section. In Cluster 2, by contrast, N is the response variable for both with-trend and detrended time series.

Figure 3.6 shows that, for both with-trend and detrended time series, the majority ($R^2 > 0.5$) of N variability in Cluster 2 of Fig. 3.1a is explained by proxies for vertical mixing. In-turn, N explains the majority (RI > 0.5) of Chl variability for both with-trend and detrended time series within this cluster (Fig. 3.5). Thus, collectively, Fig. 3.5 and 3.6 show how the relationships identified between the indirect variables and Chl in Table 3.1 may be mediated by changes in N concentration.

In terms of RI, Fig. 3.6 shows that the overall R^2 for multiple-regression is proportioned fairly evenly amongst the four explanatory variables (SST, stratification, MLD and wind speed) in Cluster 2, for both with-trend and detrended time series. This implies strong collinearity amongst these variables, as also indicated by the small proportion of RI that is entirely unique to each variable (Fig. 3.6). As the 95% confidence intervals for these RI values tend to overlap one another, there is no significant difference between these vertical mixing proxies in driving N variability within Cluster 2, irrespective of whether the time series are detrended or not.

Lastly, Fig. 3.6 shows that there are no significant changes in R^2 and RI when time series are detrended. That is, while R^2 and RI do change when time series are detrended, these changes are within the confidence bounds of their equivalent values for with-trend time series. However, despite changes not being significant, increases in overall R^2 and reduced uncertainties in both R^2 and RI values are notable when time series are detrended. Nevertheless, unlike relationships between directly-related explanatory variables and Chl, relationships between these indirectly-related explanatory variables and N are not dependent upon trends.

3.3.4 Regional-scale analysis

The two clusters capture a large-scale pattern of forcing in the North Atlantic. However, interannual Chl variability may also respond to regional forcing. It is thus informative to also consider Chl variability at smaller spatial scales.

Regional cluster distribution

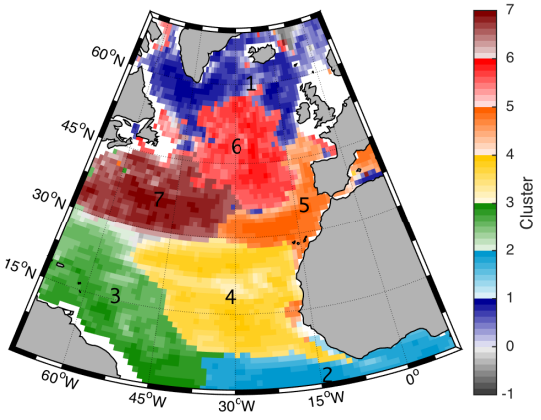


Fig. 3.7: Seven-cluster distribution of yearly-average surface Chl variability in the North Atlantic from NEMO MEDUSA 1993-2006. Colours denote cluster number (as labelled). Where present, grey indicates locality is anti-correlated with cluster. Shelf sea localities less than 200m depth are omitted (white). Shading denotes strength of correlation (or anti-correlation) with cluster, from light ($R^2 = 0$) to dark ($R^2 = 1$).

Figure 3.7 shows the seven-cluster distribution of interannual Chl variability derived from NEMO MEDUSA. Again, this distribution is broadly similar to observational data (Fig. 2.4). The most obvious discrepancy between the observational and model clusters is in the shape and position of the two most south-easterly clusters (Clusters 2 and 4). However, trial and error reveals that this region of the North Atlantic is more sensitive to the cluster analysis than the rest of the North Atlantic. For example, clusters south of ~20°N change shape and position in both datasets if logarithms are taken of the Chl data prior to clustering, while clusters in the rest of the North Atlantic remain broadly similar (not shown).

a	1	2	3	4	5	6	7
SST	**+0.55** 0.003	**-0.56** 0.002	**-0.43** 0.010	**-0.54** 0.003	**-0.64** 0.001	**-0.29** 0.046	**-0.85** 0.000
Stratification	-0.19 0.112	**-0.37** 0.021	**-0.39** 0.018	**-0.56** 0.002	**-0.61** 0.001	**-0.38** 0.019	**-0.42** 0.012
$Stratification_{WM}$	**-0.69** 0.000	-0.04 0.489	-0.07 0.359	**-0.54** 0.003	**-0.76** 0.000	**-0.37** 0.020	**-0.60** 0.001
MLD	+0.00 0.987	**+0.66** 0.000	**+0.43** 0.011	+0.20 0.108	**+0.36** 0.024	**+0.30** 0.041	**+0.48** 0.006
MLD_{WM}	-0.01 0.764	+0.16 0.152	+0.07 0.358	+0.23 0.081	**+0.65** 0.001	**+0.34** 0.028	+0.12 0.234
Wind	-0.01 0.693	**+0.33** 0.031	+0.26 0.061	**+0.40** 0.016	+0.22 0.094	+0.04 0.510	**+0.66** 0.000
NAO_{djfm}	-0.12 0.223	-0.07 0.359	-0.06 0.403	**+0.30** 0.043	-0.01 0.717	-0.01 0.684	-0.23 0.081

b	1	2	3	4	5	6	7
SST	+0.05 0.464	**-0.73** 0.000	**-0.58** 0.001	**-0.58** 0.002	**-0.71** 0.000	-0.29 0.047	**-0.88** 0.000
Stratification	+0.02 0.665	**-0.79** 0.000	**-0.54** 0.003	**-0.55** 0.002	**-0.63** 0.001	**-0.4** 0.016	**-0.42** 0.012
$Stratification_{WM}$	-0.08 0.314	-0.20 0.111	-0.22 0.092	**-0.48** 0.006	**-0.74** 0.000	**-0.38** 0.018	**-0.60** 0.001
MLD	-0.14 0.181	**+0.85** 0.000	**+0.52** 0.004	+0.27 0.058	**+0.46** 0.007	**+0.32** 0.036	**+0.47** 0.007
MLD_{WM}	+0.01 0.760	+0.19 0.119	+0.14 0.189	**+0.48** 0.006	**+0.73** 0.000	**+0.45** 0.009	+0.13 0.212
Wind	-0.17 0.137	**+0.60** 0.001	**+0.40** 0.016	**+0.30** 0.043	+0.27 0.057	+0.02 0.593	**+0.70** 0.000
NAO_{djfm}	+0.01 0.697	+0.00 0.998	+0.00 0.823	+0.15 0.176	+0.00 0.810	-0.04 0.467	-0.24 0.076

Table 3.3: R^2 and associated p-values (underneath) for a) with-trend and b) detrended relationships between interannual Chl variability and explanatory variables (rows) in each of the seven clusters of Fig. 3.7. Column numbers denote the corresponding clusters (1 to 7). A positive or negative symbol before each R^2 value denotes the sign of relationship. For emphasis, statistically significant R^2 values ($p<0.05$) are given in bold type and shaded according to the sign of relationship: blue (negative) or red (positive).

Tables 3.3a and b list the sign and strength of cluster-mean relationships with interannual Chl variability for each of the Clusters shown in Fig. 3.7. The two tables differ in representing with-trend and detrended relationships, respectively. The explanatory variables included in these two tables are the same as those included in a similar observational analysis (§2). Although the two analyses represent slightly different years (see: §3.3.1), their 7-Cluster patterns are similar. Thus,

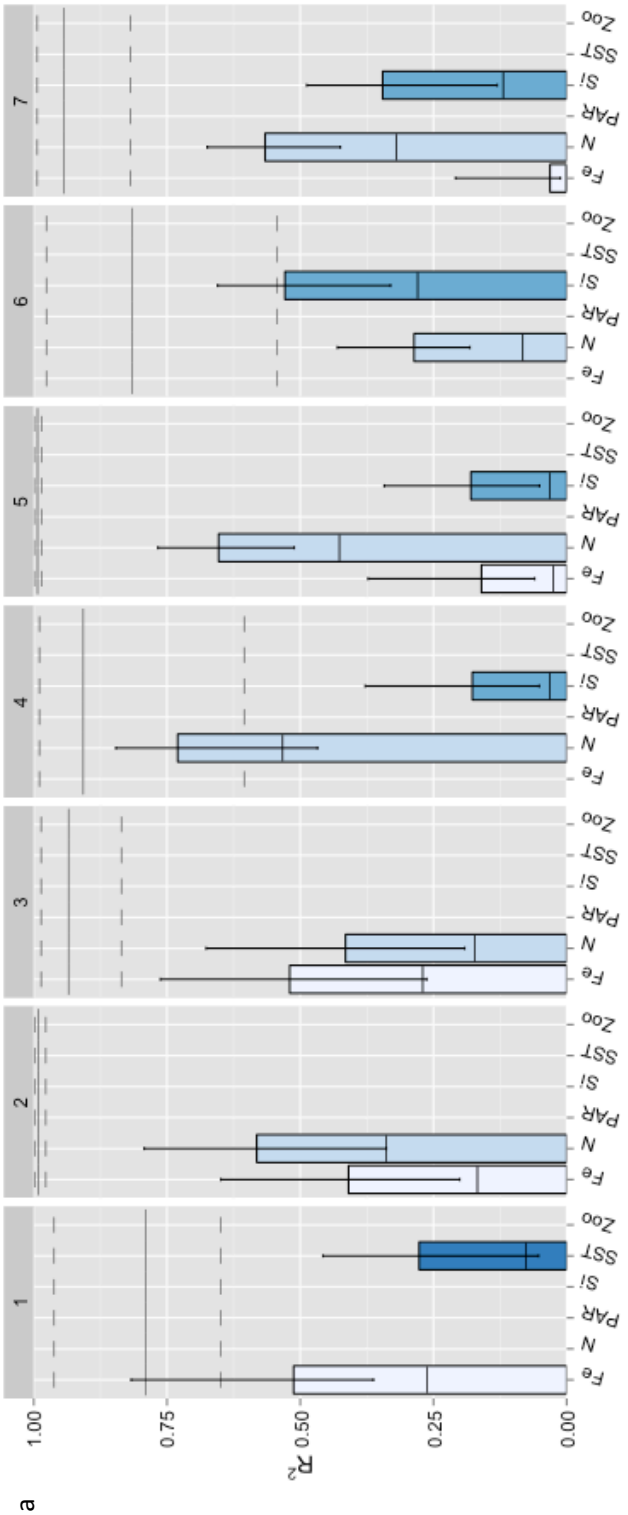
assuming that the physical processes that underlie these Chl clusters remain similar, the relationships in Tables 3.3 and 3.4 can be compared to those of the observational analysis (Table 2.3) for consistency.

Most clusters in Fig. 3.7 overlap with clusters of the corresponding number in the observational analysis (c.f. Fig 3.4). As such, comparisons of relationships between the two analyses can be made by comparing clusters of the same number in most cases. However, this is not the case for Clusters 2 and 4. While these clusters overlap with regions of the corresponding clusters in the observational analysis, the greatest overlap is between Cluster 4 in this analysis, and Cluster 2 in the observational analysis (Fig. 3.7, Fig. 2.4). As such, for the purposes of comparison, relationships of Cluster 4 are compared to those of Cluster 2 in the observational analysis, and relationships of clusters 2 (this analysis) and 4 (observational analysis) are omitted from the comparison as they do not overlap between analyses.

Table 3.3 includes both with-trend and detrended relationships. The time series are detrended in this analysis in order to assess and compare relationships within the model. For the purposes of comparing the model and observational analysis, with-trend relationships are considered.
In this case, a comparison of relationships in Table 3.3a with corresponding observational relationships (Table 2.3) reveals a number of broadscale consistencies in terms of the sign and strength of relationships between the two analyses. All relationships with an $R^2 > 0.1$ in both datasets have the same sign of relationship. Furthermore, the model exhibits 9 out of the 12 significant relationships identified in the observational analysis (Table 2.3). However, the model has many more statistically significant relationships between Chl and these explanatory variables than the observational analysis (c.f. 28/42 versus 12/42). As such, the coincidence of these 9 significant relationships may be chance. In general, the model exhibits stronger relationships with these variables than identified in the observational analysis and a better comparison of the two analyses may be to compare how the clusters are ranked in terms of strength of relationship (c.f. Table 2.3, Table 3.3a) with each of the variables. In this case, there are mixed results in terms of consistency between the two analyses. While overall rankings are not strongly concordant, there is good agreement between the clusters that have the strongest, or two strongest, relationships with each variable. All but one of the variables listed in Table 3.3 either has its strongest relationship in corresponding clusters between the two analyses (SST; $Stratification_{WM}$; MLD_{WM}; Wind), or has its two strongest relationships amongst corresponding clusters (Stratification; NAO). The only exception is MLD, whose two strongest relationships in the observational analysis

correspond to clusters having the third and fourth strongest relationships with MLD in the model. These rankings are the same irrespective of whether with-trend or detrended relationships in Table 3.3 are compared to the observational analysis. There is thus good agreement between the two analyses in terms of which clusters have the strongest relationships. Furthermore, both analyses reveal the general presence of relationships between interannual Chl variability and proxies for vertical mixing (Table 3.3; §2). These similarities imply that the model may be reproducing the biophysical relationships apparent in the observational data, and that as such, it may be interrogated further to infer how these observed indirect relationships are mediated.

3.3.5 Direct drivers of Chl variability



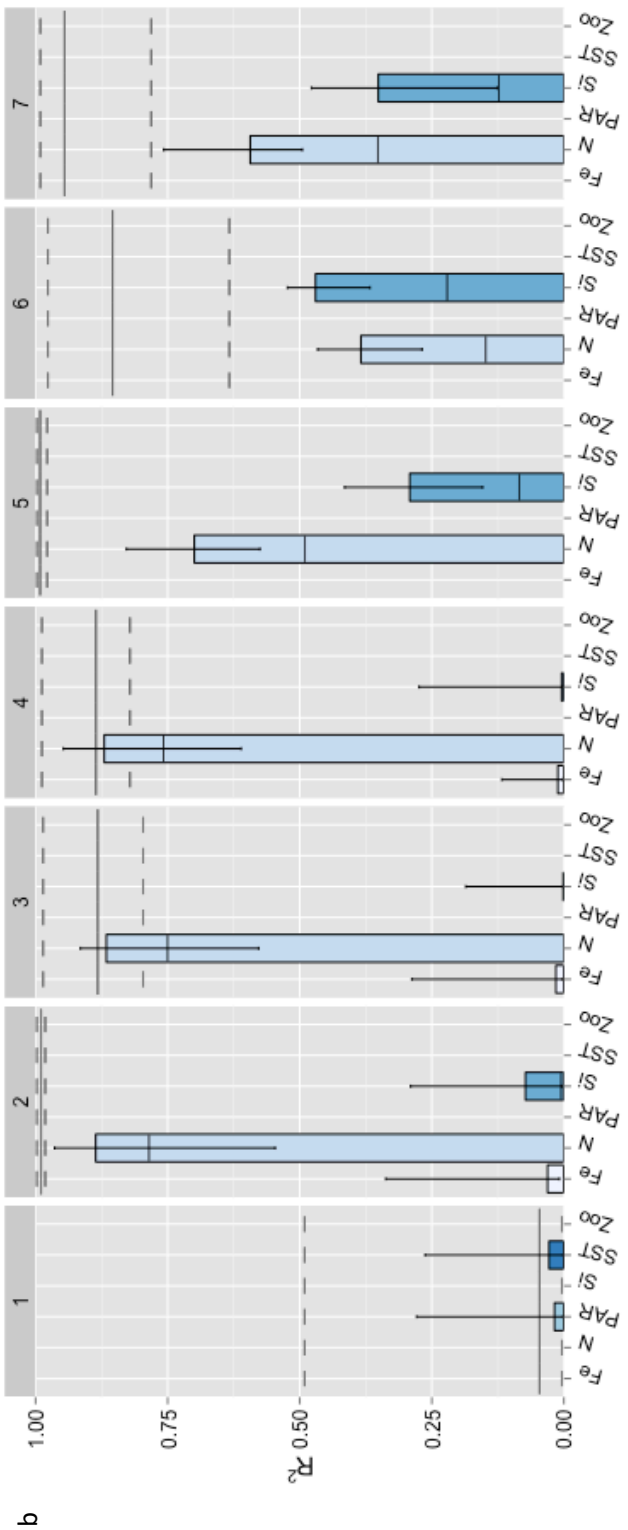


Fig. 3.8: Sections titled 1 to 7 correspond to clusters in Fig. 3.7 and each give the results of a multiple regression, including the explanatory variables for which blue bars are plotted (as labelled), upon Chl for a) with trend and b) detrended cluster-mean time series. Solid horizontal lines spanning each section indicate the R^2 of each multiple regression. Dashed horizontal lines above and below this represent associated 95% confidence intervals. Blue bars (shaded as per explanatory variable) represent the contribution of each individual explanatory variable to the total R^2 in terms of RI. Whiskers represent 95% confidence intervals of these RI values. Solid horizontal lines within each bar represent the unique contribution of that variable to total R^2, in terms of squared partial-correlations. (Nb. No confidence intervals are given for squared partial correlations.)

Figure 3.8 shows the RI of each directly-related variable upon driving interannual Chl variability within the seven clusters of Fig. 3.7. The two figures represent relationships for both with-trend (a) and detrended (b) time series. In both cases, only the direct-effect variables that are considered to be causally related to interannual Chl variability are included in the RI regression for each cluster (c.f. §3.2.7; Table 3.1). Thus, with the exception of Cluster 1, Zooplankton, PAR and SST are not considered to be causally related to phytoplankton (S3.4). In all clusters, Zooplankton appears to be related to Chl via bottom-up control. The relationships in Fig. 3.8a and b are inter-compared and discussed separately in following subsections.

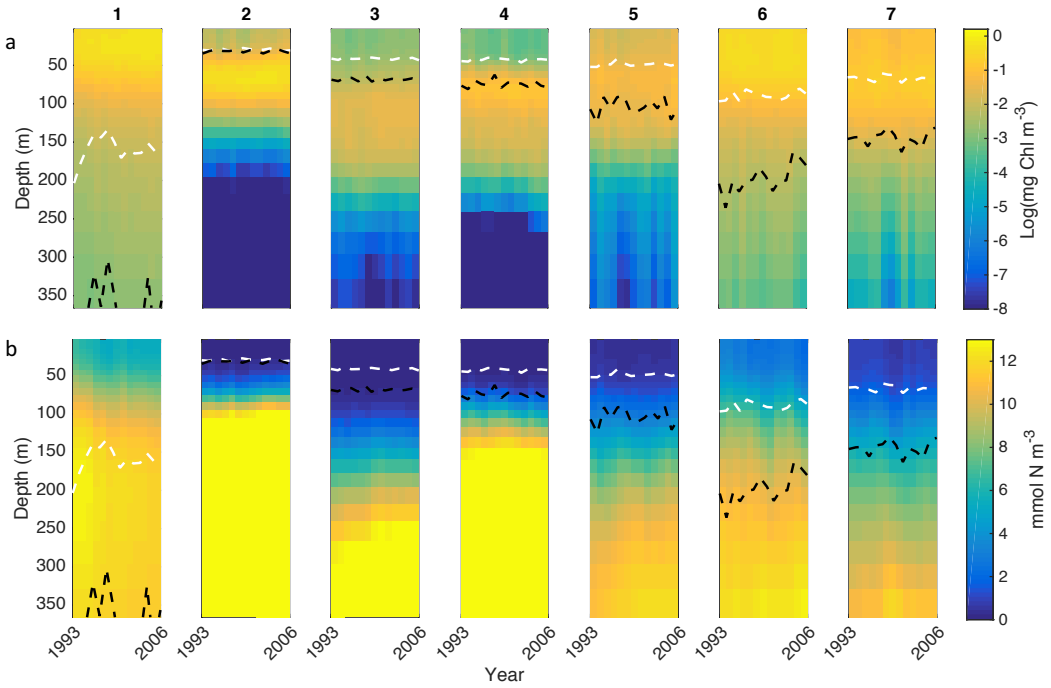


Fig. 3.9: Time series of vertical profiles for a) Chl concentration and b) N concentration, corresponding to the the seven cluster in Fig. 3.7 (as titled). Profiles represent cluster-mean concentrations. Dashed white (black) lines represent time series of yearly-averaged (winter maximum) MLD for each cluster.

With Trend vs Detrended

A comparison of Fig. 3.8a and b show that the relationships between Chl and N are stronger when the data are detrended. This implies that N drives the year-to-year variability in Chl within each of these clusters, but has less influence on the trend. By contrast, the relationship between Fe and Chl is either considerably weakened (Clusters 2 and 3), or the two are no longer causally related (Clusters 1, 5 and 7) when the time series are detrended. This implies that with-trend relationships between Chl and Fe in Clusters 1, 5 and 7 are spurious due to a coincidence of

positive trends, and that those in Clusters 2 and 3 may over-represent a weak relationship with Fe for the same reason. In any case, Fe is unrelated to Chl in Clusters 4 and 6, and has a limited or no effect upon the year-to-year variability of Chl within Clusters 1, 2, 3, 5 and 7. Unlike the uniform changes noted for relationships with N and Fe, relationships between Chl and Si either become stronger (Clusters 2, 3 and 5), weaker (Clusters 4 and 6) or remain the same (Cluster 7) when time series are detrended. The weakened relationships in Clusters 4 and 6 imply that Si drives both the year-to-year variability and negative trend in Chl within these clusters. While the stronger relationships in Clusters 2, 3 and 5 imply that Si drives only the year-to-year variability of Chl in these clusters. Lastly, SST and PAR are related to Chl variability in Cluster 1. However, the relationships with SST is considerably weakened when the time series are detrended. Again, this implies that while SST may drive the trend in Chl, it has limited or no effect upon the year-to-year variability of Chl within Clusters 1. PAR is causally related to Chl variability in cluster 1 only when the time series are detrended. However, both SST and PAR are negligibly related to detrended Chl variability in this Cluster, describing less than 10% variability ($R^2 < 0.1$) in a combined multiple regression (Fig. 3.8b).

With Trend

Fig. 3.8a shows that N is the most widespread driver of Chl variability for time series including trends, with causal relationships between the two in six out of seven clusters. Cluster 1, which has a higher N concentration than the other clusters (Fig. 3.9), is the only cluster where Chl is not causally related to N. As in the two-cluster partition, the next most prevalent drivers of interannual Chl variability are Fe and Si, having apparent causal relationships in five (Clusters 1, 2, 3, 5 and 7) and four (Clusters 4, 5, 6 and 7) of the seven clusters, respectively (Fig. 3.8a). The distribution of relationships with Fe tends to divide the seven clusters into the two-cluster pattern (c.f. Figs. 3.1a and 3.7). That is, K7 Clusters 1, 2, 3, 5 and 7 in Fig. 3.7 more-or-less represent the region of K2 Cluster 2 in Fig. 3.1a. In both partitions, these clusters represent regions of positive or negligible Chl trend (Figs. 3.1b, 3.2; S3.5) that are positively related to Fe (S3.2, S3.4a). Whereas, allowing for the exclusion of K7 Cluster 1, which is also unrelated to N, the relationships with Si appear to divide the clusters latitudinally, with relationships in all clusters except K7 Cluster 1 and the low-latitude K7 Clusters 2 and 3 (Fig. 3.8a). In addition to relationships with nutrients, SST is also causally related to interannual Chl in Cluster 1.

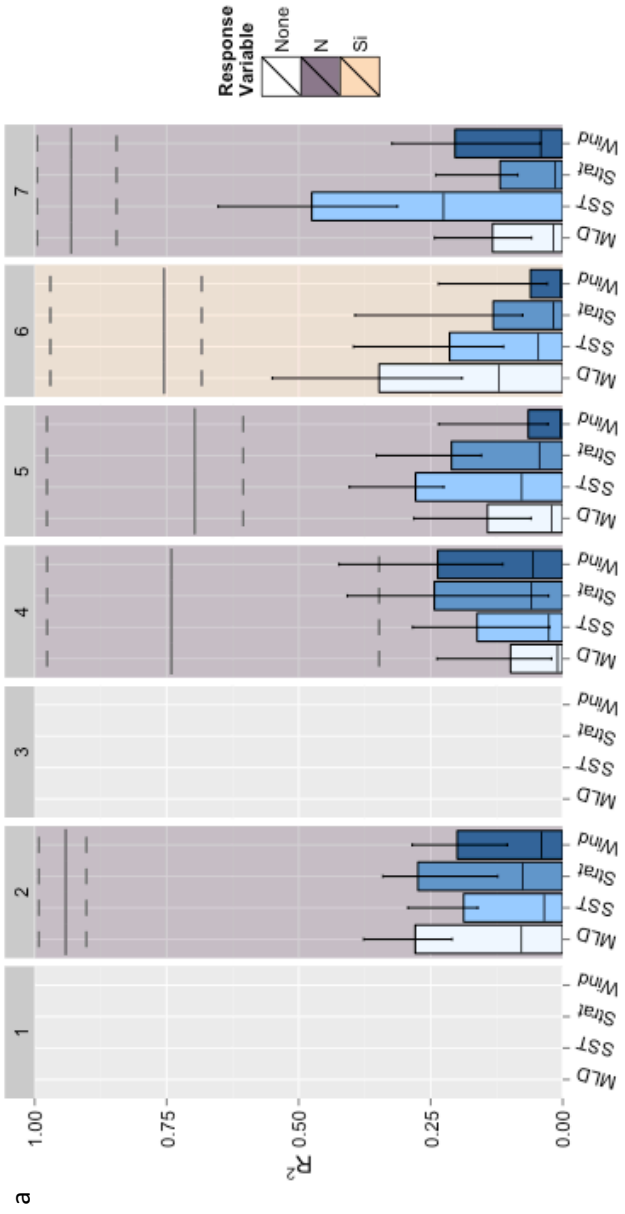
Fig.3.8a ranks the covariates in terms of their contributions to explained variance in each cluster. However, with the exception of N in Clusters 4 and 5, the 95% confidence intervals of these RI values overlap with one another. As such, these ranking are not statistically significant and may change if different years were selected for the analysis. Thus, the RI analysis does not robustly identify any single explanatory variable as being the most important descriptor of overall (including trend) Chl variability in Clusters 1, 2, 3, 6 and 7.

As regards the relationship with SST in cluster 1, it should be noted that a positive relationship between Chl and SST may imply a direct causal relationship (via the effect of temperature on metabolic rates), but could also imply an indirect-relationship via light availability (as SST is also a proxy for vertical mixing). This cannot be distinguished from the statistical relationship. However, given that the relationship is very weak when time series are detrended (Fig. 3.8b), it may more likely represent a casual or coincidental relationship in the presence of trends.

Detrended

Fig. 3.8b shows that N and Si are the most widespread drivers of Chl variability, being causally related to Chl in six out of seven clusters when the time series are detrended. Cluster 1, which has a higher N concentration than the other clusters (Fig. 3.9), is the only cluster in which Chl is not causally related to either N or Si. However, a number of these relationships with Si, and other variables, have a negligible (RI < 0.1) effect upon Chl when time series are detrended (Fig. 3.8b). This includes relationships with SST and PAR in Cluster 1 and relationships with Si and Fe in Clusters 2, 3 and 4. As this analysis is aimed at establishing the dominant drivers of Chl variability, these weaker relationships are not of interest. The dominant drivers of Chl variability are thus N in Clusters 2, 3 and 4, and both N and Si in Clusters 5, 6 and 7 (Fig. 3.8b). Of these, N is the dominant control upon interannual Chl variability within Clusters 2, 3, 4, 5 and 7 having RI > 0.50 and error-bars that do not overlap with other variables. By contrast, Si has the highest RI value in Cluster 6, however this does not exceed 0.5 and is not significantly different to the RI value of N (Fig. 3.8b). The lower RI value in this cluster is not due to a lower overall R^2 value, as Cluster 6 has a similar R^2 value to the other clusters (Fig. 3.8b). Instead, the lower RI value in this cluster is due to N and Si both representing strong drivers of Chl variability, both in terms of unique and collinear variance (Fig. 3.8b), and thus neither variable represents a dominant control upon detrended Chl variability. Lastly, no dominant driver is identified for detrended Chl variability in Cluster 1. This is due to a low overall R^2 in this cluster, as SST and PAR have $R^2 < 0.1$ for a multiple regression upon Chl in this cluster when time series are detrended (Fig. 3.8b).

3.3.6 Indirect drivers of Chl variability



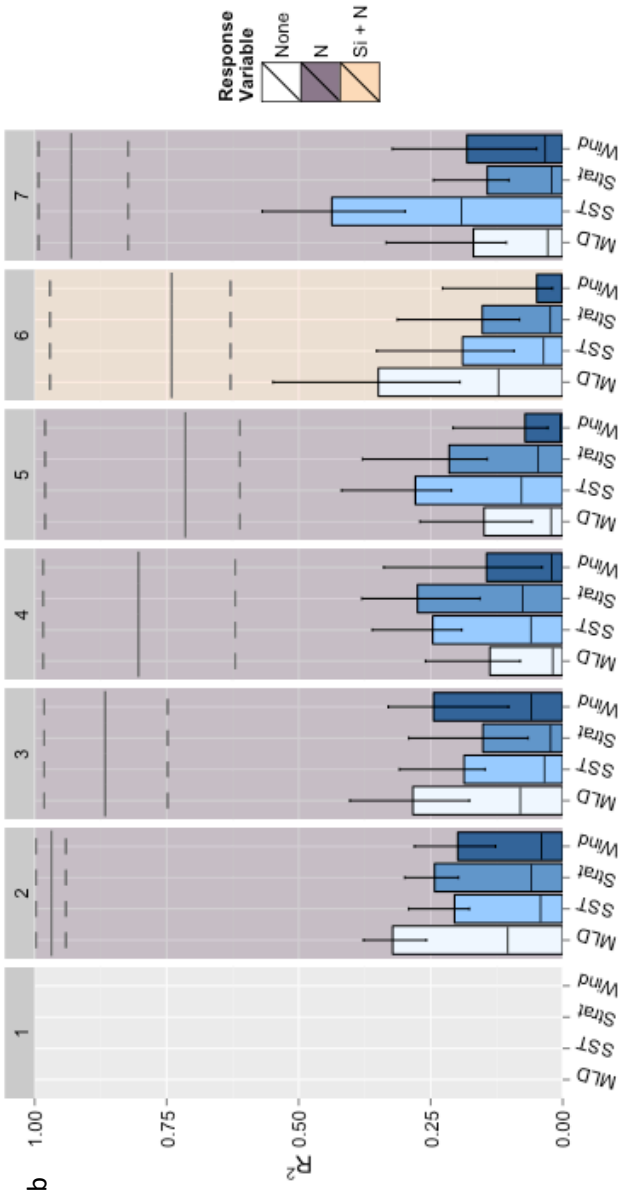


Fig. 3.10: Sections titled 1 to 7 correspond to clusters in Fig. 3.7 and each give the results of a multiple regression, including explanatory variables identified by blue bars (as labelled), upon the response variable identified by background colour (see legend) for both a) with trend and b) detrended cluster-mean time series. In each case, the response variable corresponds to the direct-driver of Chl variability identified in §3.3.5: if no direct driver was identified then a regression is not included in this section. Solid horizontal lines spanning each section indicate the R^2 of each multiple regression. Dashed horizontal lines above and below this represent associated 95% confidence intervals. Blue bars (shaded as per explanatory variable) represent the contribution of each individual variable to the total R^2 in terms of RI. Whiskers represent 95% confidence intervals of these RI values. Solid horizontal lines within each bar represent the unique contribution of that variable to total R^2, in terms of squared partial-correlations. (Nb. no confidence intervals are given for squared partial correlations.)

This section considers to what extent variability in the directly-related variables identified in §3.3.5 can be explained by proxies for vertical mixing. As outlined in §3.2.10, the dominant drivers (RI > 0.5) of Chl variability identified in Fig. 3.8 are used as response variables in this section in order to determine how well their variability can be explained by proxies for vertical mixing. Nitrate, Si and Fe are the dominant drivers of Chl variability in the clusters of Fig. 3.7 (Fig. 3.8). However, Fe is omitted as a potential driver of with-trend Chl variability in this section. The near-linear trends and surface accumulation of Fe reveal that it is driven by high atmospheric deposition rates within the North Atlantic in this hindcast simulation (S3.5; S3.6b). Increasing Fe trends are thus not attributable to vertical mixing and cannot be driven by the explanatory variables considered in this section. Thus, depending upon whether or not the time series are detrended, N, Si or none (i.e. no response variable) represent the various response variables for the clusters of Fig. 3.7 in this section.

Fig. 3.10 shows that, within most of the clusters of Fig. 3.7, the majority ($R^2 > 0.5$) of N and/or Si variability can be explained by proxies for vertical mixing. In-turn, these N and/or Si response variables explain the majority of Chl variability within each of their respective clusters (Fig. 3.8). Thus, collectively, Figs. 3.8 and 3.10 show how the relationships identified between indirect variables and Chl in Table 3.3 may be mediated by changes in nutrient (N and Si) concentrations.

In terms of RI, Fig. 3.10 shows that the overall R^2 for multiple-regression in each cluster is proportioned fairly evenly amongst the four explanatory variables (SST, Stratificatin, MLD and wind speed). This implies strong collinearity amongst these variables, as also indicated by the small proportion of R^2 that is unique to each variable (Fig. 3.10). As the 95% confidence intervals for RI values tend to overlap one another, there is no significant difference between these vertical mixing proxies in driving Si and/or N variability within each cluster, irrespective of whether the time series are detrended or not. Fig. 3.10 also shows there is little difference in R^2 and RI values between the with trend and detrended relationships in those clusters for which comparisons can be made. For example, while the order in which the RI values of explanatory variables are ranked changes in some clusters when detrended, these ranking are not significant in either case. Nor do the RI values or R^2 values change significantly from their with-trend values when detrended (Fig. 3.10). Thus, unlike between directly-related explanatory variables and Chl, relationships between indirect explanatory variables and N/Si are not dependent upon trends.

3.4 Discussion

3.4.1 Nutrients drive interannual Chl variability

The locality and strength of relationships between nutrients and Chl are better considered at the 7-cluster scale. This is because in the 2-cluster distribution, K2 Cluster 2 appears to represent a cohesive region responding to similar forcing, while K2 Cluster 1 does not. A comparison of Tables 3.2 and 3.3 shows that K7 Clusters 1, 2, 3, 5 and 7, which more-or-less comprise the area of K2 Cluster 1 (c.f. Figs. 3.1a, 3.7), tend to have stronger individual relationships with explanatory variables than K2 Cluster 1 itself. Furthermore, the sign of relationships in K7 Cluster 1, which represents the sub-polar part of K2 Cluster 1, tends to have opposing signs of relationship between macro nutrients (N) and Chl in comparison to the other clusters (K7 Clusters 2, 3, 5, 7) comprising K2 Cluster 1 (S3.4). As such, there is heterogeneity in the sign and strength of relationships comprising the larger cluster when considered at smaller spatial scales.

Detrended relationships

Of all the variables that have a direct influence upon Chl variability in the model, N is the most widely related (Fig. 3.8). Furthermore, when time series are detrended, N has the strongest causal relationship in all clusters except K7 Cluster 6, for which N and Si are jointly the strongest drivers of Chl variability, and K7 Cluster 1, in which N is not causally related to Chl (Fig. 3.8). With the exception of K7 Cluster 1, N and Si represent the only strong (RI > 0.3) drivers of Chl variability when time series are detrended. The strength of Si as a co-driver of Chl variability is latitudinally dependent: from having a limited (RI < 0.3) influence upon Chl in clusters below 30°N (K7 Clusters 2, 3 and 4), a strong but non-dominant (RI < 0.5) influence in clusters between ~30-45°N (K7 Clusters 5 and 7) to being jointly the strongest driver of Chl variability along with N in K7 Clusters 6, at about ~40-55°N (Fig. 3.7; Fig. 3.8). Iron also plays a limited role (RI < 0.3) in driving Chl variability in clusters below 30°N (K7 Clusters 2, 3 and 7) when time series are detrended (Fig. 3.7; Fig. 3.8). These results show that in the tropical, sub-tropical and parts of the sub-polar North Atlantic, year-to-year (i.e. excluding trend) variability in Chl is driven by nutrients (Fig. 3.8).

In the higher-latitude K7 Cluster 1 (~50-70°N), year-to-year Chl variability is not causally related to nutrients (e.g. N, Si and Fe). The negative relationship between nutrients and Chl in this region implies that it is not nutrient limited, and instead responds negatively to processes that drive

nutrient concentration. This is consistent with light-limitation, as light availability and nutrient-availability tend to be inversely related through vertical mixing and MLD. However, while PAR is causally related to Chl in this cluster its influence is limited (RI < 0.3), even after allowing for uncertainties. The only other directly-related variable that is causally related to Chl in this cluster for detrended time series is SST (Fig. 3.8). However, this also has a limited influence upon Chl (RI < 0.3). The positive relationships with SST in this cluster may imply a direct association via the effects of temperature upon metabolic rates (Eppley, 1972), or could also be indicative of a relationship with PAR, as SST is also a proxy for stratification. The weak relationships in this cluster may be a short-coming of using yearly-averages as seasonality is increasingly important at higher latitudes. This is discussed further in §3.4.3.

These results are thus partly consistent with previous studies using coupled physical-biological ocean models wherein nutrient availability limits growth in Tropical and Sub-Tropical regions of the ocean, while light availability limits growth in the well-mixed polar regions of the ocean (e.g. Bopp et al., 2013, 2001; Cabré et al., 2015).

With-trend relationships

When time series including trends are considered, the strength of relationships with N tends to decrease in most clusters (K7 Clusters 2-7), as corresponding relationships with Fe (K7 Clusters 1, 2, 3, 5 and 7) or Si (K7 Clusters 4 and 6) tend to become stronger in given clusters. While not all these changes in RI values are significant (i.e. outside uncertainty bounds), these tendencies towards stronger relationships with Fe or Si more-or-less partition the seven clusters into the two-cluster pattern (c.f. Figs. 3.1a; 3.7). As this two cluster pattern represents regions of positive and negative Chl trend (Fig. 3.1), these relationships with Fe and Si likewise coincide with regions of positive and negative Chl trend, respectively. Thus, these results imply that changes in Fe and Si concentrations have driven the opposing Chl trends represented by K2 Clusters 1 and 2. However, as the relationships with Fe are not robust to detrending, either changing sign (Clusters 1, 4, 5, 7) or becoming negligible (Clusters 2, 3, 4) when detrended (Fig. 3.8; S3.4), the corresponding relationships may be spurious due to a coincidence of trends. Given the limited degrees of freedom when correlating trends, this is not determined here.

3.4.2 Vertical mixing controls nutrient variability

Given that nutrients drive the majority interannual Chl variability throughout the tropical and subtropical North Atlantic (§3.4.1), it is important to understand what controls their variability. Within each of the K7 clusters comprising the tropical and subtropical North Atlantic (Fig. 3.7), vertical mixing is strongly related ($R^2 > 0.6$) to the nutrient(s) that drives Chl variability within each cluster when time series are detrended (Fig. 3.10). Each of these R^2 values represents a multiple regression of yearly-averaged SST, stratification, MLD and wind speed upon the nutrient(s) in question. There is strong collinearity amongst these variables, as indicated by their low individual RI values in comparison to their overall R^2 (Fig. 3.10). That each of these proxies for vertical mixing are strongly collinear in relation to nutrient variability implies that vertical mixing is the underlying mechanism driving these relationships. Although this study does not exclude the possibility that advection, or another process, may control nutrient variability, the strong collinearity between these explanatory variables makes this less likely as each of them would also have to be a proxy for (i.e. correlated with) the process in question.

In K7 Cluster 1, representing part of the sub-polar North Atlantic (Fig. 3.7), Chl variability was found to be causally related to PAR and SST when time series are detrended (§3.4.1). However, as neither variable explained $R^2 > 0.5$ of Chl variability (Table 3.3b; S3.4b), no response variable was included to regress vertical mixing proxies upon in this cluster (Fig. 3.10b). As such, these results do not indicate what controls Chl variability in the subpolar North Atlantic.

These results are similar to those for time series including trends, however uncertainties in R^2 and RI values increase when trends are included (Fig. 3.10). As such, R^2 values can be as low as $R^2 > 0.3$ (K7 Cluster 4), or $R^2 > 0.5$ (K7 Clusters 2, 5, 6 and 7) in the presence of trends. Nevertheless, each of the explanatory variables (SST, stratification, MLD and wind speed) remain strongly collinear in explaining nutrient variability within each of the clusters. The most obvious discrepancy between the with-trend and detrended results, is that no response variable is included in regressions for K7 Clusters 1 and 3 when time series include trends. This is due to Fe being the strongest (RI > 0.5) driver of Chl variability in these clusters (Fig. 3.8). As the increasing trend in Fe is not related to vertical mixing processes in this model, it is omitted as a response variable in these clusters.

Relationships are more ambiguous in the subpolar North Atlantic. In contrast to other K7 clusters, none of the directly-related variables appear to dominate Chl variability within the subpolar K7

Cluster 1 (§3.4.1). This may be due to spatial heterogeneity, or seasonally-varying relationships at these latitudes (§3.4.1). In any case, relationships between indirect and direct explanatory variables are not considered for K7 Cluster 1 as they do not relate to Chl variability. Nevertheless, relationships between Chl and proxies for vertical mixing are identified within this cluster, and these relationships and the way in which they may be mediated are discussed in §3.4.3.

3.4.3 Vertical mixing drives Chl variability

Mid-to-low latitude North Atlantic

Given strong relationships between vertical mixing and nutrients (§3.4.1), and nutrients and Chl (§3.4.2), these results imply that vertical mixing drives Chl variability within the tropical and subtropical North Atlantic. This is consistent with relationships between indirect explanatory variables (SST, Stratification, MLD, wind speed) and Chl identified for these clusters in Tables 3.2 and 3.3.

These results are thus consistent with the hypothesis that vertical mixing controls growth in the permanently-stratified (year-round SST > 15°C) ocean (Behrenfeld et al., 2006). However, in comparison to previous studies that have identified such indirect relationships between Chl and proxies for vertical mixing (Behrenfeld et al., 2006; Siegel et al., 2013; §2), this study also shows that relationships are mediated by nutrient availability. This is the case for all clusters within the tropical and subtropical North Atlantic. This is at odds with studies that have identified vertical mixing as a poor descriptor of interannual Chl variability (Barton et al., 2015), or shown relationships between Chl and stratification to be limited to less than 11% of the permanently-stratified ocean (Dave, 2014; Dave and Lozier, 2013). Discrepancies between these results and those found here are likely to be due to the spatial and/or temporal scales of analysis.

As shown in §2, relationships can be weak at local spatial scales, but dominate at larger spatial scales. A comparison of the K2 and K7 clusters in this analysis shows that relationships are spatially dependent. Perhaps the most obvious example of spatial dependence is that the NAO is significantly related to Chl within K2 Cluster 1, but not in any of the smaller K7 Clusters (1, 2, 3, 5 and 7) that tend to comprise its area (Table 3.1, 3.3). K2 Cluster 1 includes the area of the subpolar North Atlantic and thus comprises regions of opposing sign and strength of relationship with physical variables. Thus, with the exception of the NAO, relationships in K2 Cluster 1 may be better considered at smaller K7 spatial scales (c.f. K2 Cluster 1, Table 3.1; K7 Clusters 1, 2, 3, 5 and

7, Table 3.3). By contrast, nearly all relationships with physical variables in K2 Cluster 2 are stronger at that larger scale than when considered at smaller spatial scales (c.f. K2 Cluster 2, Table 3.1; K7 Clusters 4 and 6, Table 3.3). The exception in this case is again for the relationship with the NAO, which is not significant at the larger scale (Table 3.1). This spatial dependence implies that relationships with Chl should be assessed at the spatial scale of interest. Thus, while relationships between stratification and Chl may not be widespread at the local scale (Dave, 2014), this does not preclude the two being strongly related at global scale (Behrenfeld et al., 2006). It is also noteworthy that stratification and Chl are strongly related in both K2 clusters at the larger scale of analysis in this study (Table 3.1).

Discrepancies with previous studies may also be related to the difference of using yearly-averaged and yearly-metric time series instead of deseasonalised monthly data (e.g. Barton et al., 2015). Comparing deseasonalised monthly time series assumes a constant relationship between variables throughout the seasonal cycle, which may not represent dynamics within seasonally stratified waters. For example, yearly-averaged phytoplankton growth in temperate and higher-latitude waters may be related to the opposing processes of deep winter convection and seasonal stratification throughout the seasonal cycle. Thus, yearly-metrics and yearly-averages may represent processes that drive interannual Chl variability within seasonally-stratified waters better than deseasonalided monthly data. Furthermore, given that seasonality increases with latitude, such yearly-metrics may become increasingly important with latitude. This is partly supported by the distribution of K7 relationships identified here: for both with-trend and detrended relationships, Chl variability tends to be more strongly related to yearly-average stratification in clusters south of ~30°N, and to winter-minimum stratification in clusters north of this. Although, such a latitudinal separation is not apparent amongst yearly-averaged and winter-maximum MLD relationships (c.f. Fig. 3.7, Table 3.3).

High latitude North Atlantic

Results are less conclusive for the subpolar North Atlantic. Aside from trend-dependent relationships, yearly-averaged Chl variability in K7 Cluster 1 is not causally and significantly related to any of the explanatory variables (direct or indirect) considered in this analysis (Fig. 3.8). However, although they are non-significant, all yearly-averaged (i.e. excluding winter metrics) relationships with nutrients, PAR and vertical mixing proxies in K7 Cluster 1 are of the opposite sign to relationships within the other nutrient-limited K7 clusters (§3.4.1; Table 3.3; S3.4b). While only circumstantially, this indicates that yearly-averaged growth is light-limited in this cluster. In

this case, the weakness of these relationships may be due to the reliance upon yearly-averaged and winter-extreme time series in a seasonally stratified region. Deep winter convection and increased wind speeds may bias yearly-averaged time series towards wintertime values, yet phytoplankton growth at these latitudes occurs during seasonally stratified periods ranging between Spring/Summer/Autumn months (although see: Behrenfeld, 2010; Sverdrup, 1953). Thus, light-limited relationships may be better identified using yearly-metrics corresponding to the seasonally stratified period. Spatially heterogeneity may also obscure relationships within the subpolar North Atlantic (e.g. Follows and Dutkiewicz, 2001). In this case, phytoplankton growth may be forced, or related to large-scale forcing, differently throughout the region. For example, yearly-averaged Chl variability may be both light and nutrient limited within the area K7 Cluster 1. This would appear to be in contradiction to clusters representing regions of coherent interannual Chl variability. However, the clusters were not defined using detrended time series and as such may be less well correlated in this case.

When with-trend relationships are considered, interannual Chl variability in K7 Cluster 1 is significantly and causally related to Fe, SST and winter-minimum stratification (Table 3.3; S3.4a). Ignoring the relationship with Fe, which changes sign after detrending (S3.4), negative associations with N and Si in this cluster imply that interannual phytoplankton growth is not nutrient limited (S3.4a). Instead, a positive relationship with SST could either imply a direct-relationship via temperature and enhanced metabolic rates (Eppley, 1972), or an indirect-relationship via enhanced stratification and light availability. In the latter case, it is questionable why Chl is not also significantly related to yearly-averaged PAR, stratification or MLD. Either way, in contrast to this yearly-averaged relationship with SST, the negative relationship with winter-minimum stratification may imply a positive response to enhanced winter mixing and new nutrient availability at the onset of stratification. In this case, phytoplankton growth may be light-limited or temperature dependent when yearly averaged, but still respond to total nutrient availability seasonally. Such a dichotomy in relationships throughout the seasonal cycle may explain the overall lack of yearly-averaged or winter-metric relationships in this cluster. Although, in this case, it is questionable why Chl is not also positively related to winter-maximum MLD (Table 3.3).

In addition to spatial and temporal variability, the lack of strong and consistent relationships in K7 Cluster 1 could also reflect a dependence upon dynamics that have not been considered (e.g.

re-stratification dynamics and bloom timing [although see: Henson et al. 2009]; phytoplankton-zooplankton decoupling [Behrenfeld 2010]). For a discussion of other drivers of Chl variability see §1.2 and §5.2.1).

3.4.4 Bimodal pattern

The two-cluster distribution of Chl variability identified in this analysis (Fig. 3.1) is broadly similar to patterns identified in previous studies (§2; Patara et al. 2011). Although there are discrepancies, the patterns are similar in separating the sub-tropics (30-45°N), and much of the western North Atlantic, from the remainder of the North Atlantic. The forcing of this broadscale pattern has previously been linked to the NAO index, SST, N and wind (Patara et al., 2011), or the NAO index, SST, stratification, MLD and zonal wind speed (§2) within the North Atlantic. However, while previous studies have identified associations with other variables, they have not accounted for collinearity amongst them, or demonstrated how indirect relationships with Chl are mediated. Here it is shown that interannual Chl variability in this two-cluster pattern is driven by N variability in K2 Cluster 2, which is, in turn, driven by vertical mixing (Figs. 3.5, 3.6). No directly-related variable is found to dominate Chl variability in K2 cluster 1, this is likely because this cluster includes most of the subpolar North Atlantic, wherein yearly-averaged relationships tend to be of opposing sign to other sub-regions within the cluster (see: §3.4.3). Thus, while the whole of K2 Cluster 1 may respond similarly to indirect-drivers of Chl variability (Table 3.2), these relationships may be mediated differently within the confines of the cluster, e.g. via nutrient- and light-limitation.

Furthermore, many of the indirect (SST, MLD, wind speed) and direct (N, Zooplankton and detrended Fe and Si) drivers of Chl variability tend to have similar bimodal patterns of variability to that of Chl within the North Atlantic (Fig. 3.4; S3.1). Thus, given the relationships between Chl and these explanatory variables, it is likely these bimodal patterns are interrelated with that of Chl variability in Fig. 3.1a. As such, the two cluster division in Fig. 3.1a may be driven by changes in meteorological forcing associated with NAO (Follows and Dutkiewicz, 2001; Patara et al., 2011). This association is further supported by the similarity of NAO patterns of wind speed, SST, stratification and MLD (§2, Fig. 2.3) to corresponding two-cluster patterns of these variables and Chl in the North Atlantic (Fig. 3.4). Discrepancies between these patterns and that of Chl may thus be attributable to regions of opposite response, especially within the confines of K7 Cluster 1.

However, despite these similarities, only one of the two Chl clusters in Fig. 3.1a is significantly related to the NAO index when time series include trends, and neither is significantly related when time series are detrended (Table 3.2). Nevertheless, the opposing sign of relationship with NAO between the two clusters implies a nominal association with the index. These poor relationships with the NAO index may be due to the index itself representing a crude proxy of large-scale changes in this climatic oscillation (including manifestations in wind speed, SST, stratification and MLD variability) that are better related to Chl variability themselves. Poor associations could also be due to discrepancies between the reanalysis data used to force this hindcast modal and the observational data from which the NAO index used in this analysis was calculated (§3.2.2; §3.2.3). Relationships with the NAO index are even less-well represented at the the 7-clusters scale. In only one of seven clusters is Chl significantly related to NAO for time series including trends, and there are no significant relationships when time series are detrended (Table 3.3). The lack of association with NAO in these clusters may be related to the different spatial scales of analyses. As in §2, the two-cluster pattern may represent large-scale NAO-related Chl variability while smaller clusters identify relationships the dominate regionally.

A distinguishing feature of the two clusters in this analysis is their opposing trends. K2 Clusters 1 and 2 have opposing positive and negative Chl trends, respectively (Fig. 3.1). Indeed, the negative trend in K2 Cluster 2 is consistent with observational data from a similar period (1997-2007; Henson et al. 2010; Vantrepotte & Mélin 2011). The positive Chl trend in K2 Cluster 1 may be driven by an increasing Fe concentration throughout the North Atlantic in this simulation. However, both yearly-averaged and winter-minimum stratification are also strongly related to the trend in Chl variability within this cluster (Table 3.2). These are the only explanatory variables to have similar year-to-year variability and trends to that of Chl in this cluster. The increasing trend in Chl could thus be related to weakening stratification over this period. However, with the exception of Fe, none of the other directly-related variables (N, Si, PAR, SST, Zooplankton) are likewise related to the trend in Chl. If the relationship with stratification is mediated by increased vertical nutrient flux, then the lack of yearly-averaged accumulation could imply that additional nutrients are being taken-up in supporting the enhanced phytoplankton growth. Long-term changes in nutrient concentration may thus be more apparent in comparing winter/spring monthly values. However, the lack of a cluster-average relationship could also be due to spatial heterogeneity of intermediary-relationships with stratification within this cluster (see: §3.1.3). As such, the cause of the long-term trend in Chl within the cluster requires further investigation,

perhaps considering relationships locally using monthly data. In K2 Cluster 2, the only variables that tend to be more strongly causally-related to Chl variability when time series include trends are Si and wind speed (Table 3.2). These are also the only explanatory variables to have similar year-to-year variability and trends to that of Chl in this cluster. A comparison of Figs. 3.1a and 1.2 shows that K2 Cluster 2 underlies the track of westerly and trade winds belts within the North Atlantic. Thus, K2 Cluster 2 may be responding to changes in mean wind speeds within the North Atlantic. The NAO is indicative of changes in the speed and direction of wind tracks in the North Atlantic. However, this analysis identifies stronger relationships with NAO-related variables than the NAO index itself.

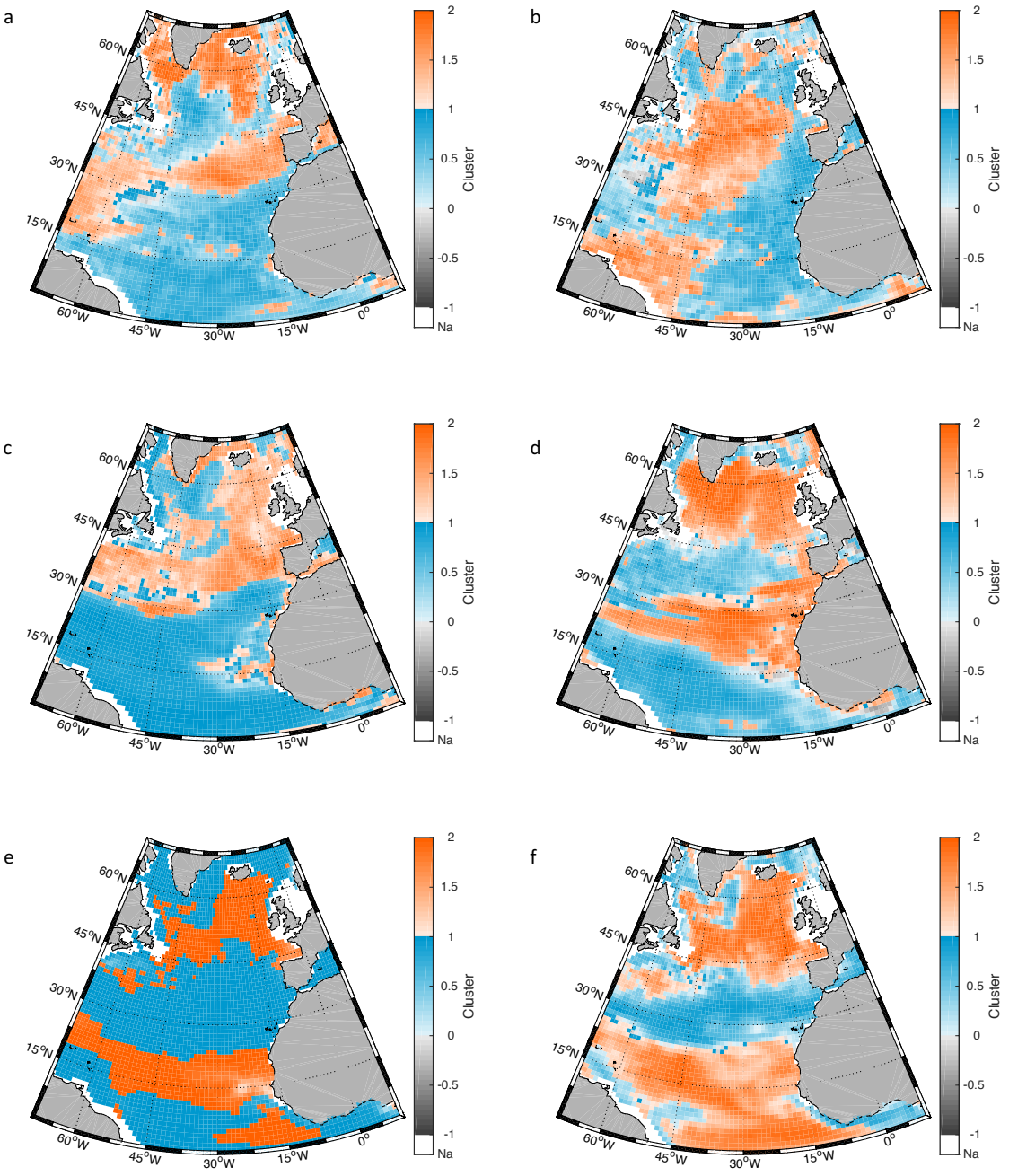
3.5 Conclusion

This chapter used a biogeochemical model to investigate how previously observed relationships between interannual Chl variability, vertical mixing and NAO are mediated (§2). Similar patterns of variability imply that Chl is responding to the same processes within the hindcast model as to observed satellite data. The model also reveals the majority of relationships between Chl and vertical-mixing explanatory variables identified in the observational analysis, although relationships tend to be stronger and less infrequent within the model.

While relationships with the NAO index are less apparent, similar bimodal patterns of variability in Chl and explanatory variables are consistent with previously observed NAO patterns (Fig. 2.3). The model reveals that interannual Chl variability is dominated ($R^2 > 0.5$) by N variability within the eastern part of this pattern, with strong collinearity amongst the vertical mixing variables (SST, stratification, MLD, wind speed) that drive this macro-nutrient variability. Relationships are also strong between Chl variability and the vertical mixing variables themselves. In the western part of this pattern, spatial variability appears to weaken overall relationship with directly-related explanatory variables, although Chl is significantly related to N variability when time series are detrended ($R^2 = 0.49$; $p=0.005$; positive). Chlorophyll is also strongly related to vertical mixing variables in this cluster when time series are detrended. Thus, although overall relationships are weaker, year-to-year Chl variability in this cluster is also largely controlled by vertical mixing and macro-nutrient availability. However, with-trend Chl variability in this Cluster is strongly related to Fe. Given that Fe has a high concentration throughout the North Atlantic in this hindcast simulation, and that this relationship changes sign when detrended, this may represent a spurious relationship due to a coincidence of trends. The only other variable that is strongly related to un-detrended Chl variability in this cluster is stratification.

At the smaller seven-cluster spatial scale, strong interannual relationships with vertical mixing variables are mediated by N availability in the mid-to-low latitude ocean, with increasing dep\endence upon Si with latitude. Interannual Chl variability appears to be light-limited in the subpolar North Atlantic, although these relationships tend to be insignificant. This could be due to spatially or seasonally-varying relationships within this region.

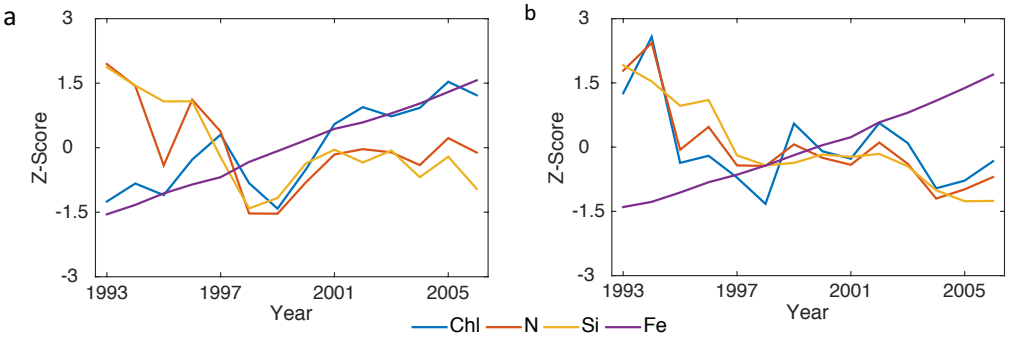
3.6 Supplementary figures



S3.1: Two-cluster distributions of a) stratification$_{WM}$, b) MLD$_{WM}$, c) Si, d) detrended-Si, e) Fe and f) detrended-Fe variability in the North Atlantic from NEMO MEDUSA 1993-2006. Colours denote cluster number: Cluster 1 (blue), Cluster 2 (orange). Grey indicates locality is anti-correlated with cluster. Shelf sea localities less than 200m depth are omitted (Na; white). Shading denotes strength of correlation (or anti-correlation) with cluster, from light (R^2 = 0) to dark (R^2 = 1).

	{With Trend}		*{De-trended}*	
	1	2	1	2
N	+0.00 0.980	**+0.81** 0.000	**+0.49** 0.005	**+0.89** 0.000
Fe	**+0.69** 0.000	-0.21 0.102	-0.24 0.078	+0.17 0.145
Si	-0.13 0.213	**+0.44** 0.010	**+0.30** 0.043	+0.19 0.122
PAR	**-0.38** 0.018	**-0.43** 0.011	**-0.43** 0.011	**-0.48** 0.006
Zooplankton	**+0.76** 0.000	**+0.84** 0.000	**+0.89** 0.000	**+0.88** 0.000

S3.2: R^2 and associated p-values (beneath) for the strength of relationship between interannual Chl variability and explanatory variables (rows). Columns denote the corresponding clusters (1 or 2) and whether or not variables were detrended before regression. A positive or negative symbol before each R^2 value denotes the sign of relationship. For emphasis, statistically significant R^2 values ($p<0.05$) are given in bold type and shaded according to the sign of relationship: blue (negative) or red (positive).

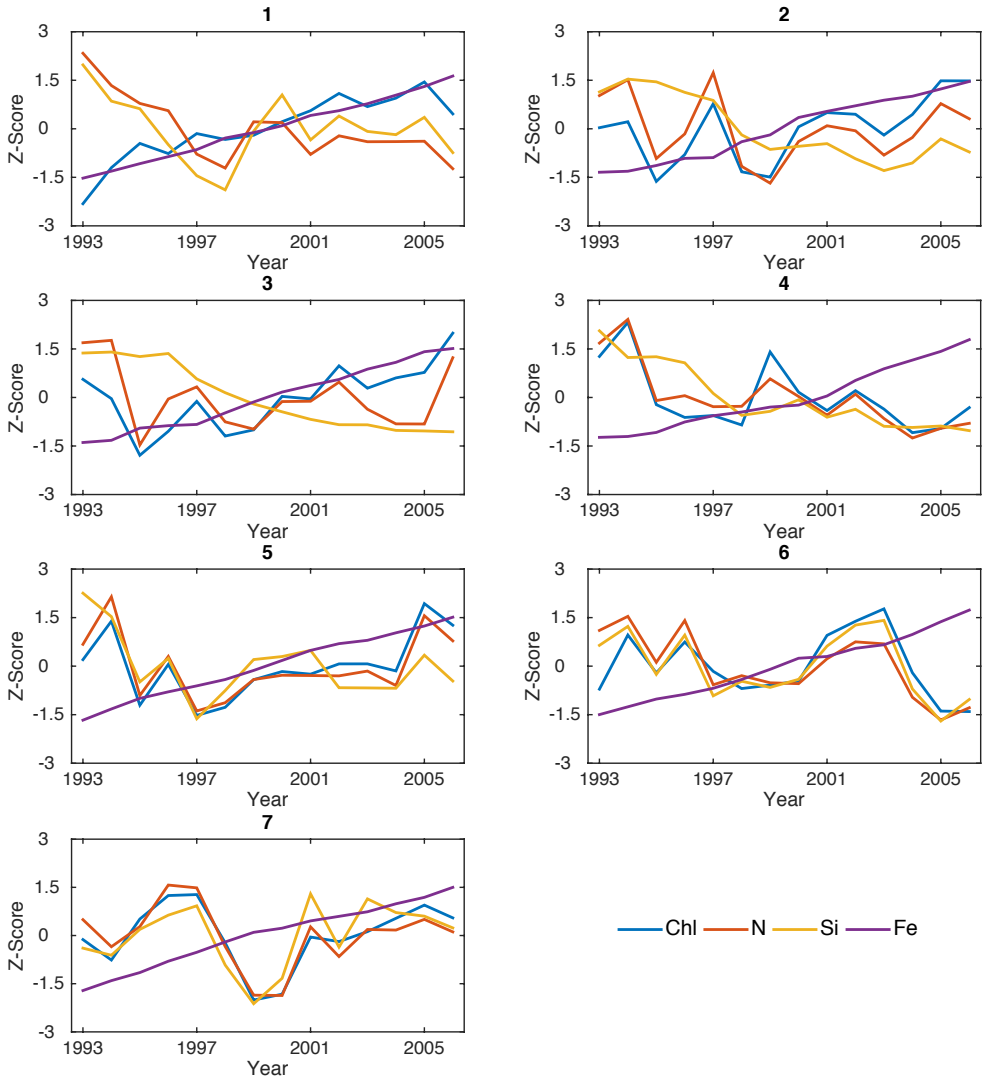


S3.3: cluster-mean time series of standardised surface Chl (blue), N (orange), Si (yellow) and Fe (purple) concentrations for a) Cluster 1 and b) cluster 2 in Fig. 3.1a.

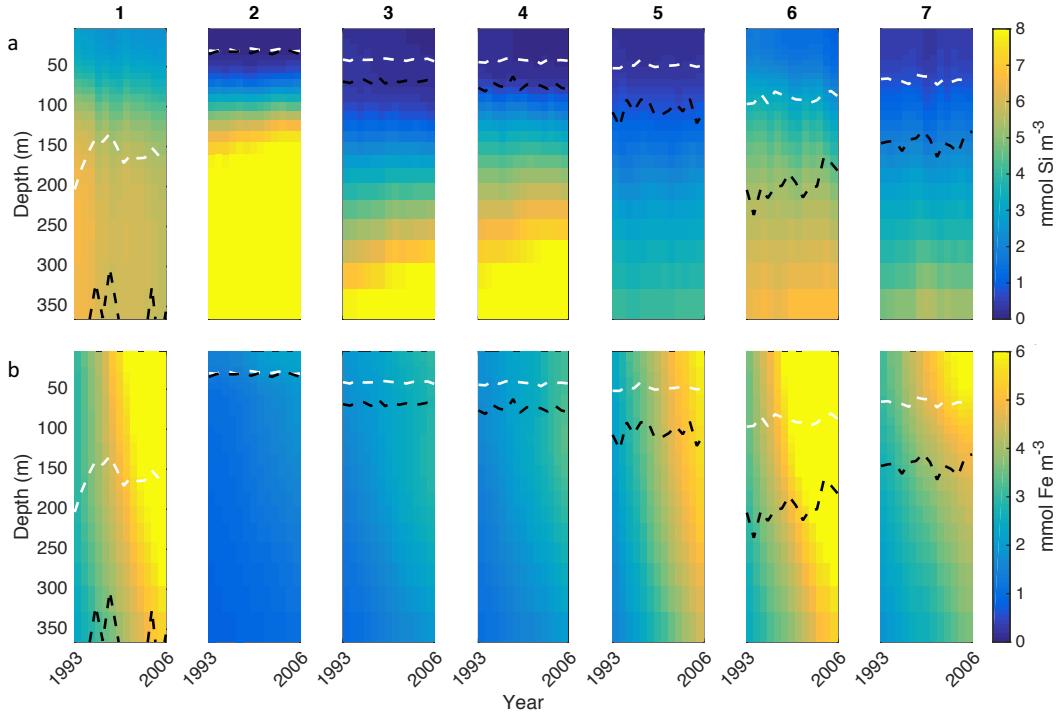
a	1	2	3	4	5	6	7
N	**-0.56** 0.002	**+0.48** 0.006	**+0.30** 0.044	**+0.84** 0.000	**+0.87** 0.000	**+0.52** 0.004	**+0.88** 0.000
Fe	**+0.78** 0.000	**+0.31** 0.039	**+0.40** 0.015	-0.28 0.054	+0.13 0.214	-0.04 0.499	+0.00 0.815
Si	-0.11 0.255	-0.07 0.364	**-0.34** 0.029	**+0.29** 0.048	+0.25 0.071	**+0.76** 0.000	**+0.68** 0.000
PAR	-0.11 0.247	**-0.88** 0.000	**-0.37** 0.022	-0.20 0.110	**-0.37** 0.020	**-0.39** 0.017	**-0.55** 0.002
Zooplankton	+0.05 0.465	**+0.95** 0.000	**+0.96** 0.000	**+0.94** 0.000	**+0.98** 0.000	**+0.60** 0.001	**+0.94** 0.000

b	1	2	3	4	5	6	7
N	-0.11 0.244	**+0.96** 0.000	**+0.86** 0.000	**+0.87** 0.000	**+0.99** 0.000	**+0.72** 0.000	**+0.94** 0.000
Fe	**-0.42** 0.012	+0.03 0.579	+0.01 0.753	+0.02 0.662	**-0.44** 0.010	**-0.39** 0.017	-0.24 0.073
Si	-0.05 0.434	+0.14 0.196	+0.00 0.893	+0.00 0.958	**+0.58** 0.001	**+0.81** 0.000	**+0.70** 0.000
PAR	+0.03 0.579	**-0.86** 0.000	**-0.53** 0.003	-0.24 0.077	**-0.40** 0.015	**-0.39** 0.017	**-0.57** 0.002
Zooplankton	**+0.50** 0.005	**+0.98** 0.000	**+0.95** 0.000	**+0.93** 0.000	**+0.98** 0.000	**+0.75** 0.000	**+0.94** 0.000

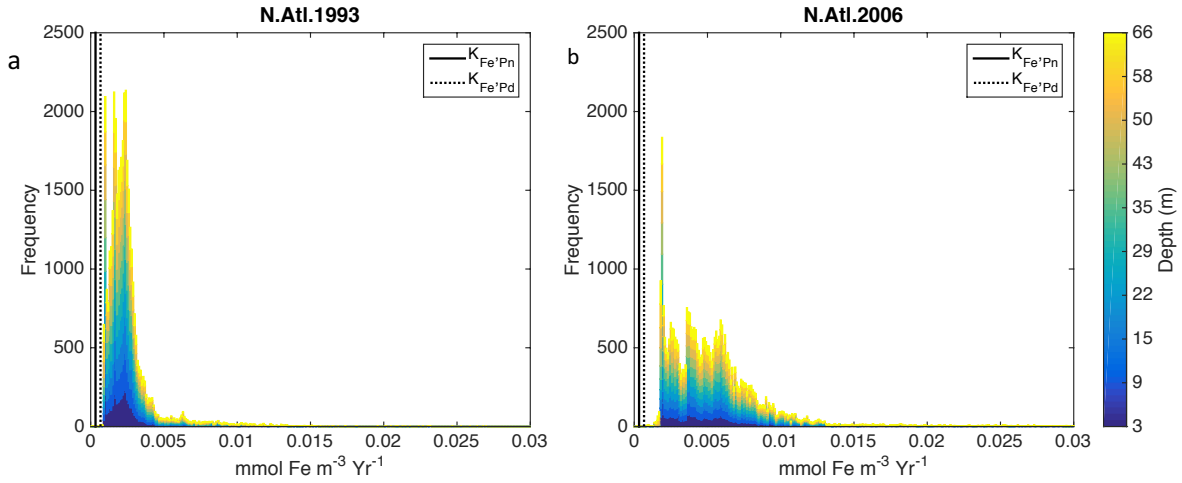
S3.4: R^2 and associated p-values (underneath) for a) with-trend and b) detrended relationships between interannual Chl variability and explanatory variables (rows) in each of the seven clusters of Fig. 3.7. Column numbers denote the corresponding clusters (1 to 7). A positive or negative symbol before each R^2 value denotes the sign of relationship. For emphasis, statistically significant R^2 values ($p<0.05$) are given in bold type and shaded according to the sign of relationship: blue (negative) or red (positive).



S3.5: cluster-mean time series of standardised surface Chl (blue), N (orange), Si (yellow) and Fe (purple) concentrations for Cluster 1 to 7 (as labelled: 1, 2, …, 7) in Fig. 3.7.



S3.6: Time series of vertical profiles for a) Si concentration and b) Fe concentration, corresponding to the the seven cluster in Fig. 3.7 (as titled). Profiles represent cluster-mean concentrations. Dashed white (black) lines represent time series of yearly-averaged (winter maximum) MLD for each cluster.



S3.7: distribution of Fe concentration in a) 1993 and b) 2006 within the study domain of the North Atlantic. Histograms represent concentration of each 1° cell for the ten near-surface vertical-grid cells (0 to 66m; as per colourbar). For comparison, vertical black lines represent half-saturation state of small non-diatom (solid line; $K_{Fe'Pn}$) and large diatom (dashed line; $K_{Fe'Pd}$) phytoplankton size classes (Yool et al., 2011).

Chapter 4: The interplay of Chl variability and forcing over the next century in the North Atlantic

4.1 Introduction

4.1.1 Overview

In previous chapters, stratification and vertical mixing were found to be related to interannual Chl variability in the North Atlantic. Furthermore, a bimodal pattern of Chl variability was identified that was similar to those of stratification and vertical mixing and helped accentuate these relationships. This chapter aims to assess if these relationships will hold over the next century, and what consequences they have for Chl concentration and variability under a warming climate.

4.1.2 Background

The interannual relationship between Chl and stratification in the ocean is of interest as it underpins a key question in oceanography: how will marine primary production respond to a warming climate? It is generally expected that a warming climate will increase stratification and inhibit (stimulate) productivity in regions that are nutrient (light) limited (Doney, 2006). This is consistent with broad-scale relationships observed globally (e.g. Behrenfeld et al., 2006; Kahru et al., 2010). However, the importance of such relationships has been questioned due to weak, or insignificant associations at the local scale (Dave, 2014; e.g. Dave and Lozier, 2013). In the North Atlantic, interannual Chl variability can be divided into two regions that have similar variability amongst themselves, but that are anti-correlated with each other. Relationships between Chl and processes of mixing are stronger when averaged within these regions than when assessed at the local scale (§2). This chapter assesses how these relationships, and the pattern of variability they describe respond to a warming climate. In addition, the long-term mean response of Chl to warming over the next century is also assessed.

4.1.3 Stratification and Vertical Mixing

Stratification and vertical mixing are thought to regulate primary productivity in the ocean due to their opposing effects on nutrient and light availability. Deeper mixing can replenish nutrient-depleted surface waters, but may also inhibit growth if phytoplankton are mixed to light-limiting depths (Sverdrup, 1953). Mixing may also have indirect-effects upon phytoplankton via its association with, and their sensitivity towards, surface temperatures (Eppley, 1972) or via the concentration of phytoplankton in the water column and how this interplays with zooplankton predation (Behrenfeld, 2010). A number of studies have identified global relationships between phytoplankton biomass, either measured as Chl or primary production (PP), and stratification (Behrenfeld et al., 2006; Martinez et al., 2009), sea surface temperature (SST; Siegel et al. 2013), wind stress (Kahru et al., 2010), and mixed layer depth (MLD; Patara et al. 2011). Most of these studies find a general change in sign of relationships between the permanently-stratified mid-to-low latitude ocean, and the seasonally-stratified higher latitude ocean, and attribute relationships to controls upon growth of nutrient and light availability, respectively. However, such relationships have been criticized for only explaining a small or insignificant amount phytoplankton variability when assessed at the local scale (Dave, 2014; Dave and Lozier, 2013). Thus, it appears that while stratification and vertical mixing may be less important than other processes at driving Chl variability locally, they represent important drivers of interannual Chl variability when averaged over larger-spatial scales (§2). However, these studies have tended to assess interannual relationships over recent observational time scales (< ~14 years). While a number of studies have used biogeochemical ocean models to project the impact of global warming upon phytoplankton growth (e.g. Cabré et al., 2015; Steinacher et al., 2010), less attention has been given to how interannual variability and relationships with phytoplankton may change with warming.

4.1.4 North Atlantic Bimodal Pattern

There is a large-scale bimodal pattern of interannual Chl variability in the North Atlantic (§2). This pattern is associated with concomitant changes in stratification, vertical mixing and nutrient availability (§3), and has been linked to the North Atlantic Oscillation (§2, Patara et al. 2011). However, it is unknown how the phytoplankton and Chl response to this pattern and associated forcings will change into the future with continued warming. For example, as stratification increases does the phytoplankton response to this climatic variability change? Does the bimodal pattern of Chl variability change shape or position?

4.1.5 Aim of this chapter

This study is fundamentally concerned with the question of how stratification and vertical mixing affect Chl variability within the North Atlantic. However, previous analysis has shown that these relationships act over large spatial scales and are predominantly related via a large-scale pattern of variability (§2). As such, this chapter assesses how these relationships, and this pattern of variability, respond to a warming climate over the next century. This chapter focuses upon the following questions:

- Does a warming climate affect Chl concentration and variability via changes in stratification and mixing?

- Do relationships between interannual Chl variability and proxies of stratification and vertical mixing in the North Atlantic remain consistent throughout the next century?

- Does the bimodal pattern of interannual Chl variability in the North Atlantic (§2) remain consistent throughout the next century?

- Do the bimodal patterns of interannual stratification and vertical mixing proxies (§3) impose the bimodal pattern upon Chl variability?

4.2 Methodology

4.2.1 Overview

A number of the pre-processing and analysis techniques used in this chapter are common to those used elsewhere in this thesis. Where applicable, this methodology refers the reader to methodologies of previous chapters in this thesis (§2; §3).

4.2.2 Data

This analysis is based upon output from a forecast simulation of NEMO MEDUSA. NEMO MEDUSA is a coupled physical–biological model. MEDUSA is an intermediate-complexity ecosystem model resolving two size classes of phytoplankton and zooplankton groups along with Nitrogen, Silica and Iron nutrient cycles (e.g. Yool et al., 2011). NEMO is an ocean general circulation model coupled to a sea-ice model (Madec, 2008). This 1/4° ocean-only forecast was forced at the surface with HadGEM2-ES (Hadley Centre Global Environment Model Version 2 – Earth System; Collins et al. 2011) output using CMIP5 (Phase 5 of the Coupled Model Intercomparsion Project) setup (Jones et al., 2011) under the Relative Concentration Pathway (RCP) 8.5 future warming scenario (e.g. Moss et al., 2010). This represents a 'business as usual' emissions scenario without mitigation, resulting in an 8.5 times increase on the pre-industrial level of global radiative forcing by 2100 (Riahi et al., 2011). For a discussion of the skill of CMIP3 and CMIP5 climate models, including HadCM3 (a climate model that precedes HadGEM2-ES used here), to simulate climatic variability associated with the NAO see: Flato et al. (2013), Randall et al. (2007) and citations therein.

4.2.3 Pre-processing

All variables used in this analysis were extracted as 3-dimensional (latitude, longitude, time) yearly-average values. The study domain of the North Atlantic was chosen for consistency with previous chapters (0 to 70°N and 70°W to 10°E). All variables were either surface values (SST, Chl), or referenced to the surface (MLD, wind speed). These data were then re-gridded to 1x1° resolution and arranged into n by p matrices, comprising n observations for p years.

4.2.4 Clustering

The forecast data used in this analysis includes trends that dominated overall variability. The data were clustered both including and excluding these trends in three ways:

Fixed Clusters: including trend

The entire time series (1993-2099) of pre-processed data were standardised to unit variance and k-means clustered into two groups (k=2; see: §2.2.3; §2.2.4).

Fixed Clusters: excluding trend

The previous section (*fixed clusters: including trend*) divided the North Atlantic into regions with and without a dominant negative trend. From this, the cluster-mean time series representing those regions including strong negative trends showed a potential change point around 2042. This was tested by finding the minimum Schwarz Information Criterion (SIC) value between fitting a single linear trend to the time series, or a combination of two independent linear trends at all possible divisions of the time series (Beaulieu et al., 2016). This showed that the negative trend in Chl was best represented by two individual linear trends: 1993—2042 and 2043—2099. As such, the pre-processed data were first divided into these two periods, then individually detrended and standardised before clustering. To detrend the data, a simple linear regression of Chl against a linearly increasing series (1,2,…,n) was fitted and subtracted from each individual time series comprising the two periods. The two periods of detrended data were then standardised and k-means clustered into two groups (k=2), as described in §2.2.3 and §2.2.4.

Dynamic Clusters

To allow clusters to move dynamically with time, the data are clustered (as per: §2.2.3; §2.2.4) for smaller sections of overlapping years and then strung together into a continuous series. This requires choosing a window length (the number of years of each section) with choices between 1 (i.e. no overlap) and the whole time series (i.e. fixed clusters). Assuming a certain consistency in the data, this method also requires a procedure for labeling the clusters (i.e: Cluster 1; Cluster 2), such that the most similar clusters are labelled in the same way between consecutive clustering solutions. These choices are discussed below.

Labelling Clusters: Chl

Two methods for labeling (i.e. numbering) clusters were considered in this analysis: one based upon choosing the maximum temporal similarity between consecutive clusters, and one based upon choosing the maximum spatial similarity between consecutive clusters. To keep the labeling consistent with previous studies, the first clustering solution was seeded with the cluster centroids and labels from §3 in the k-means clustering algorithm (§2.2.4). This resulted in identical solutions to random seeding (not shown), but meant that the first clustering solution had labels consistent with the previous chapter. Subsequent clusters were then labelled according to their temporal or spatial correlation with the preceding (already labelled) clustering solution. In the temporal case, this involved correlating the cluster-mean (or centroid) time series with those of the corresponding clusters from the preceding clustering solution for overlapping years, and averaging the two R^2 values together. Then repeating this process having swapped the cluster numbers of the present clustering solution. This gives two R^2 values for the correlation of present and previous clustering solutions: one for the correlation of clusters with the present clusters labelled as default, and another where the present cluster-numbers are swapped. The present clusters are then labelled in accordance with the highest average R^2 value. In the spatial case, this involved labeling the clusters to give the greatest geographical overlap with the preceding clustering solution. As the clustering process applied throughout this thesis is based upon temporal similarity (§2.2.4), the temporal labeling solution was chosen for consistency. However, in the case of Chl data, it is noted that both solutions gave the same results when considered over various window lengths (14-30 years; not shown).

Labelling Clusters: Physics

The clusters of physical variables (SST, MLD and wind speed) were labelled to match-up temporally with already-labelled Chl clusters. This was done by labelling the physical clusters so as to give the highest (MLD, wind speed) or lowest (SST) overall R^2 relationship with pre-sorted Chl cluster for each time window. For each window, physical clusters were regressed upon corresponding Chl clusters and the average of R^2 values was taken. This was done with the physical clusters both labelled as default and the other way around. For each window, the permutation of labelling giving the highest/lowest average R^2 value was chosen.

Window Length

The choice of window length affects the stability of the clusters between sections. In general, a longer window should result in the clusters being more similar and changing less quickly through the time series. This is because more of the same overlapping years are being considered in consecutive sections. However, this may also act to postpone gradual change and result in 'jumps' from one set of similar clusters to another. Figure S4.1 shows the spread, in terms of percentage of observations being reassigned between consecutive clusters, for various window lengths. The percentage change has a maximum of 50%, as differences greater than this would have resulted in the clusters being relabelled (§4.2.4) and shows results for 14 years (chosen for consistency with §3) to 30 years. For this analysis, it was desirable to avoid large changes as this could lead to clusters swapping positions in accordance with the criterion chosen for labeling clusters (§4.2.4). Figure S4.1 shows that window lengths of 17, 27 and 29 years result in the smallest maximum-changes throughout the Chl time series. In order to allow the clusters to be as dynamic as possible within these confines, a 17-year window was chosen.

4.2.5 Assessing Relationships

Differences Between Relationships: fixed clusters

In order to test if relationships change significantly with time, a two-tailed t-test is used to test against the null hypotheses that regression coefficients remain the same (e.g. $\beta_{1_i} = \beta_{1_{i+1}}$):

$$t = \frac{\beta_{1_i} - \beta_{1_{i+1}}}{s_{\beta_{1_i} - \beta_{1_{i+1}}}} \tag{4.1}$$

Where $s_{\beta_{1_i} - \beta_{1_{i+1}}}$ is defined using the corresponding standard errors of the coefficients $(s_{\beta_{1_1}}, s_{\beta_{1_{i+1}}})$:

$$s_{\beta_{1_i} - \beta_{1_{i+1}}} = \sqrt{s^2_{\beta_{1_1}} + S^2_{\beta_{1_{i+1}}}} \tag{4.2}$$

Eq. 4.1 is described by Paternoster et al. (1998). This t-statistic (t) is then referenced against a t-distribution with n-4 degrees of freedom to give a corresponding p-value. Where n is the combined number of observations comprising each β. P-values < 0.05 are assumed significant.

Differences Between Relationships: fixed clusters

In order to test for a trend in relationships within dynamic (time varying) clusters, linear trends are fitted to time series of R^2 relationships and tested against the null-hypothesis of zero gradient

($ß_1 = 0$). For relationships between variables, this requires that corresponding clusters are representative of the same underlying relationship throughout the forecast time series. In this respect, Chl clusters are numbered to remain consistent with preceding clusters, and physical clusters are numbered to be consistent with concurrent Chl clusters (c.f. §4.2.4). Linear regressions cannot be applied if dynamic clusters transition between relationships during the forecast time series.

4.3 Results

4.3.1 Fixed Clusters

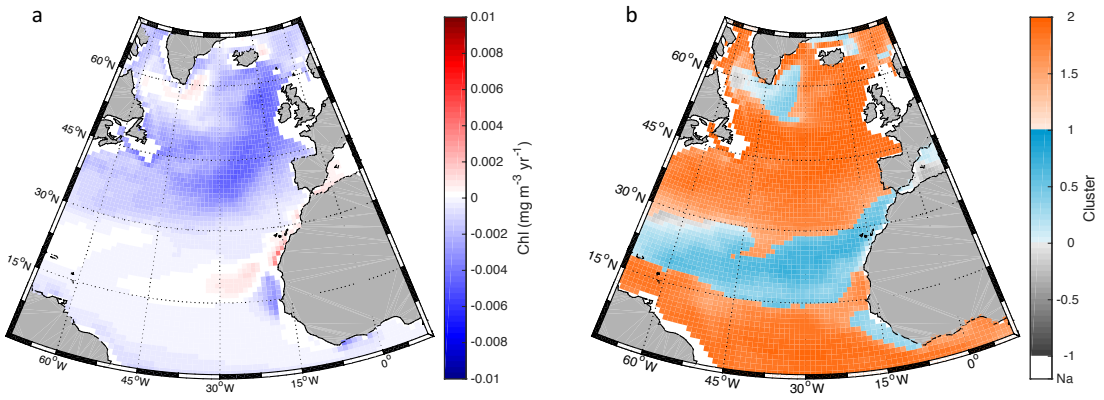


Fig. 4.1: Forecast of NEMO MEDUSA Chl variability (1993-2099) in the North Atlantic represented as a) local-scale 1x1° Chl trends over this period and b) two-clusters according to similarity of interannual Chl variability over this period. In a) colouring and shading indicates sign (red = positive; blue = negative) and strength (light to dark = weak to strong; white = insignificant) of trend. In b) colouring and shading indicates cluster number (1 = blue; 2 = orange) and strength of relationship with cluster mean (from light shading at $R^2 = 0$ towards dark shading at $R^2 = 1$); anti-correlated points are coloured grey, irrespective of cluster number. Shelf sea localities less than 200m depth are omitted (Na; white).

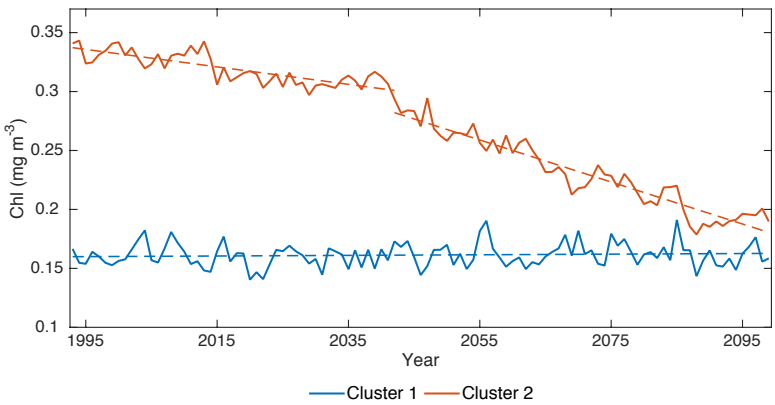


Fig. 4.2: Cluster-mean time series of interannual Chl variability (solid line) corresponding to Fig. 4.1b, and fitted linear trends (dashed lines). Two trends have been fitted to the time series of Cluster 2 (red dashed lines: 1993-2042 and 2043-2099) as there is a change point in trend between 2042 and 2043. Both trends for cluster 2 are significant ($p<0.05$) and had lower overall SIC value than fitting a single trend (see: §4.3.4). Trend for Cluster 1 is not statistically significant.

With the exception of a zonal band between ~15-30°N, and certain high-latitude regions including areas of the Labrador and Irminger Seas, Fig. 4.1a shows that the majority of the North Atlantic is forecast to have a negative trend in Chl over the next century. Where present, this negative trend tends to dominate overall variability (S4.2c) and clustering thus separates the North Atlantic into regions with (Cluster 2) and without (Cluster 1) a dominant negative trend in Chl (Figs. 4.1b and 4.2). Furthermore, there is a distinct shift in this negative trend between 2042 and 2043, and the overall decline is better represented, in terms of SIC (see: §4.3.4), by two trends than one. This reflects a greater rate-of-decline and lower mean value of Chl concentration after the shift (Fig. 4.2).

	Cluster 1	Cluster 2
N	**+0.65**	**+0.97**
	0.000	0.000
SST	-0.04	**-0.95**
	0.053	0.000
MLD	**+0.37**	**+0.80**
	0.000	0.000
Wind	**+0.09**	**+0.79**
	0.001	0.000

Table 4.1: R^2 and associated p-values (underneath) for the strength of relationship between Chl variability and each of SST, MLD and wind speed for both Clusters 1 and 2 in Fig. 4.1b. A positive or negative symbol before each R^2 value denotes the sign of relationship. Associated p-values are given underneath. For emphasis, statistically significant R^2 values ($p<0.05$) are given in bold type and also shaded according to the sign of relationship: blue (negative) or red (positive).

Strong relationships between Chl variability and nitrate (N), SST, MLD and wind speed in Cluster 2 (Table 4.1; Fig. 4.1) imply that the negative trend in Chl is due to reduced mixing and nutrient availability with warming. Relationships with Zooplankton and other nutrients (silica and iron) are also strongly related, but omitted for brevity as they tend to co-vary with N and Chl (S4.3). Of the physical variables, the strongest relationship is with SST. This may imply that SST is a better proxy for the effects of vertical mixing and nutrient availability upon Chl than stratification or MLD. However, it could also represent an additional relationship between phytoplankton and temperature (Eppley, 1972). Irrespective, the strong relationships of these physical variables with Chl and their co-linearity with N indicates that the majority of explained variance is due to the effects of stratification and vertical mixing upon nutrient availability.

In order to investigate underlying interannual variability, the long-term trend in Chl and explanatory variables needs to be removed. Despite overall good fits, however, the change point

in Chl (Fig. 4.2) is not explained by SST, MLD or wind speed (S4.4). As such, the data are detrended and assessed in two sections, both before (1993-2042) and after (2043-2099) the change point. This avoids introducing artefacts into residuals by subtracting less well-fitted linear trends from the relevant time series.

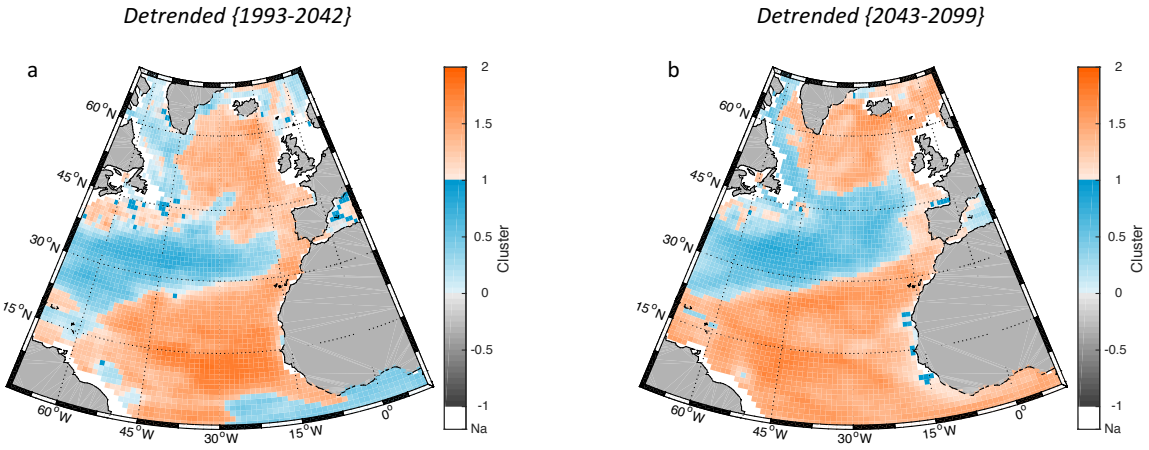


Fig. 4.3: Two-cluster partition of NEMO MEDUSA detrended interannual Chl variability for a) 1993-2042 and b) 2043-2099. Colouring and shading indicates cluster number (1 = blue; 2 = orange) and R^2 strength of relationship with cluster centroid (from light shading at R^2 = 0 towards dark shading at R^2 = 1); anti-correlated points are coloured grey irrespective of cluster number. Shelf sea localities less than 200m depth are omitted (Na; white).

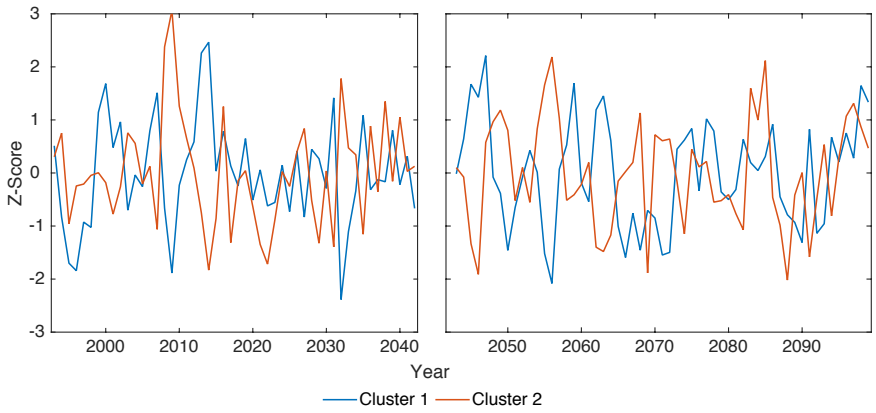


Fig. 4.4: Cluster-mean time series of standardised detrended Chl variability for Clusters 1 (blue) and 2 (red) in Fig. 4.3a (left) and Fig. 4.3b (right).

Figures 4.3a and b show the dominant pattern of interannual Chl variability in the North Atlantic after detrending the time series. The overall pattern is similar between the two periods 1993-2042 and 2043-2099, and with that of previous studies (§2; §3), suggesting that this bimodal pattern is characteristic of the North Atlantic. At the broadest scale, the pattern tends to divide the North Atlantic latitudinally by separating a region at mid-latitudes from regions both north and south of it. However, this broadscale pattern is complicated by including fringes of the northern, southern and especially western North Atlantic into the cluster representing mid latitudes (Cluster 1 in Fig. 4.3a and b).

Figure 4.4 shows cluster-mean time series of Chl concentration for detrended periods 1993-2042 (Fig. 4.3a) and 2043-2099 (Fig. 4.3b). A visual inspection of these time series appears to show the clusters are anti-correlated at shorter time scales (~2-3 years). However, this is not strongly evident in the overall correlations of these two periods (1993-2042; R^2 = 0.20; negative sign; p = 0.001 and 2043-2099; negative sign; R^2 = 0.09; p= 0.022), presumably because they also appear to be positively correlated at longer time scales (>~10 years).

	{1993 – 2042}		{2043 – 2099}	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2
SST	**-0.24** 0.000	**-0.48** 0.000	-0.09 0.029	**-0.53** 0.000
MLD	+0.11 0.021	**+0.46** 0.000	**+0.37** 0.000	**+0.53** 0.000
Wind	**+0.16** 0.003	**+0.42** 0.000	**+0.21** 0.000	**+0.37** 0.000

Table 4.2: R^2 and associated p-values (beneath) for strength of relationship between detrended Chl variability and detrended SST, MLD and wind speed (rows). Columns denote cluster and period of each relationship, as represented in Fig. 4.3a (1993-2042) and b (2043-2099). A positive or negative symbol before each R^2 value denotes the sign of relationship. For emphasis, statistically significant R^2 values ($p<0.05$) are given in bold type and also shaded according to the sign of relationship: blue (negative) or red (positive).

	{With Trend}	{Detrended}		
	Time	SST	MLD	Wind
Cluster 1				
t-statistic	-1.45	-0.83	-1.40	+0.20
p-value	0.155	0.414	0.161	0.843
Cluster 2				
t-statistic	**+8.76**	+0.27	+0.28	+0.36
p-value	0.000	0.792	0.778	0.718

Table 4.3: Test statistics for null hypothesis that relationships with Chl do not change between the two periods: 1993-2042 and 2043-2099. Column 2 (Time) relates to Fig. 4.1 and tests for a significant change in Chl trend between the two periods (Fig. 4.2). Columns 3-5 relate to Table 4.2 and Figs. 4.3a and b, and tests for significant changes in relationships between respective clusters for the two periods (c.f. Eq. 4.1). Statistically significant ($p<0.05$) t-statistic values are in bold type.

Table 4.2 lists relationships between detrended interannual Chl variability and detrended interannual SST, MLD and wind speed for the fixed (i.e. time-invariant) Chl clusters shown in Figs. 4.3a and b. Most noticeably, interannual Chl variability is more strongly related to these explanatory variables in Cluster 2 than Cluster 1. The strongest relationship is with SST in Cluster 2, with this variable explaining about 50% of the variance in Chl between the two periods. This is followed in order by MLD and wind speed, with the latter explaining about 40% variance in Chl between the two periods. By comparison, the only variable that has a significant relationship with Chl in Cluster 1 for both time periods is wind speed, which explains about 18% variance in Chl between the two periods.

In association with these relationships, Table 4.3 lists results of t-tests for the hypothesis that the coefficients for the relationships listed in Table 4.2 do not change significantly between the two periods. (Note that these do not test for a significant change in R^2, but for a significant change in the individual coefficients.) These results show that none of the detrended relationships with Chl change significantly between the two periods. This implies that the underlying relationships between Chl and processes of vertical mixing remain unchanged between the two periods. The only significant change applies to the negative trend in Chl between the two periods in clusters prior to detrending (Fig. 4.1b). This justifies detrending the Chl data in two separate periods (Fig. 4.2).

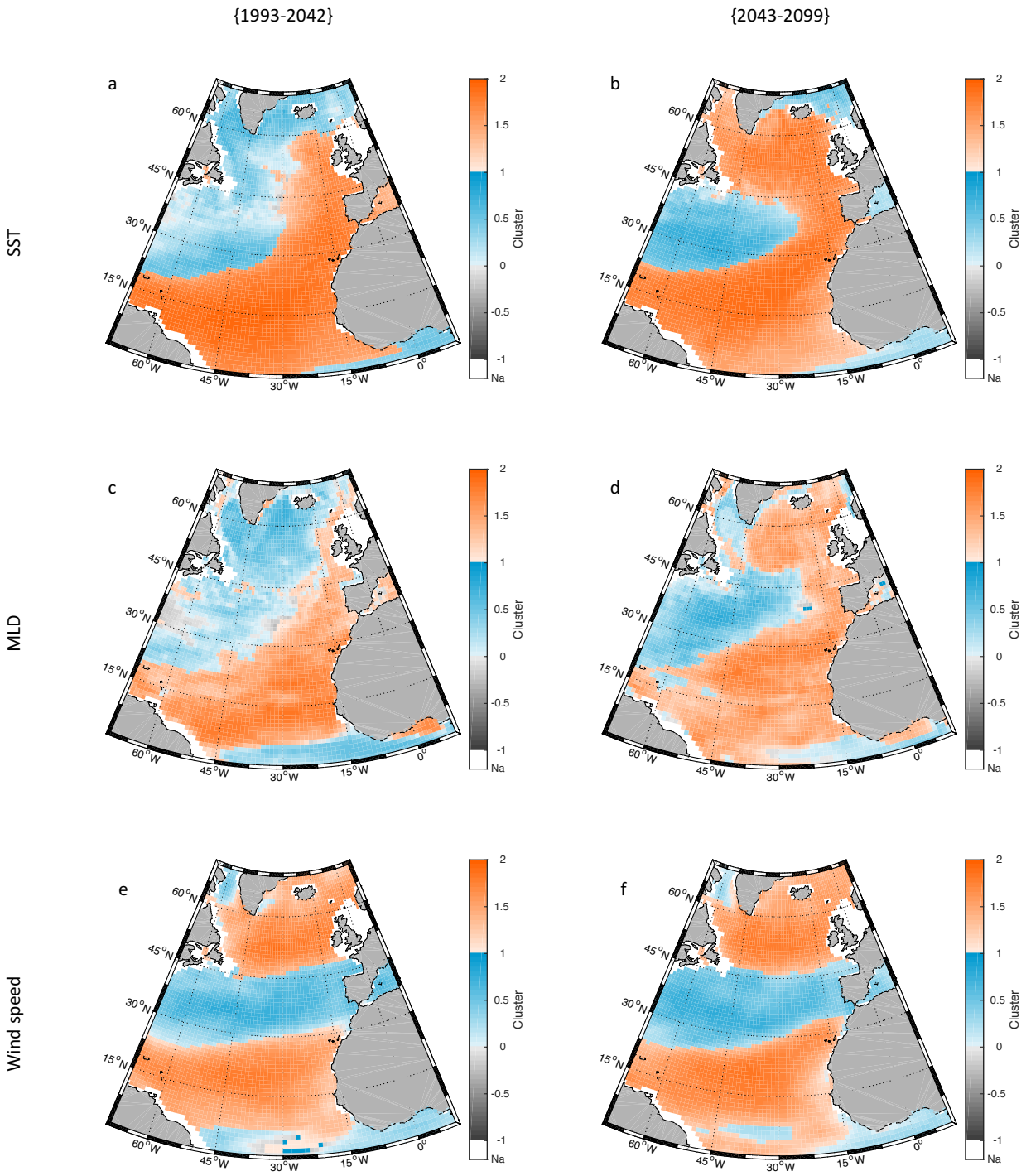


Fig. 4.5: Two-cluster partitions of NEMO MEDUSA detrended SST, MLD and wind speed (rows; as labelled) for periods 1993-2042 and 2043-2099 (columns; as labelled). Colouring and shading indicate cluster number (1 = blue; 2 = orange) and R^2 strength of relationship with cluster centroid (from light shading at $R^2 = 0$ towards dark shading at $R^2 = 1$). Anti-correlated points are coloured grey, irrespective of cluster number. Shelf sea localities less than 200m depth are omitted (Na; white).

Figure 4.5 shows detrended clusters of SST, MLD and wind speed over the two periods used for the Chl clusters (1993-2042 and 2043-2099; Fig. 4.2). A comparison of these patterns with those of the Chl clusters (Fig. 4.5) reveals good overall similarity between them. This is most apparent for SST and MLD in the latter period (2043-2099), wherein Cluster 2 tends to wrap North, South and East around Cluster 1, which itself represents the mid-latitudes and western fringes of the North Atlantic. These similarities imply that one or more of these physical variables may be responsible for driving the corresponding patterns in Chl variability.

	{1993 – 2042}	{2043 – 2099}
SST	74.21%	73.95%
MLD	65.48%	76.61%
Wind	68.75%	80.06%

Table 4.4: Spatial overlap (%) between detrended Chl clusters (Fig. 4.3) and detrended SST, MLD and wind speed clusters (Fig. 4.5). Columns denote the period of each relationship.

	{1993-2042}	{2043-2099}
Chl	**-0.20**	**-0.09**
	0.001	0.022
SST	+0.01	-0.02
	0.546	0.257
MLD	**-0.31**	**-0.20**
	0.000	0.000
Wind	**-0.13**	**-0.08**
	0.012	0.039

Table 4.5: R^2 and associated p-values (beneath) for strength of relationship between Clusters 1 and 2 of detrended Chl, SST, MLD and wind speed (rows) in Figs. 4.3 and 4.5. Columns denote the corresponding period of each relationship. A positive or negative symbol before each R^2 value denotes the sign of relationship. Statistically significant ($p<0.05$) R^2 values are in bold type.

Table 4.4 lists the spatial similarity, in terms of percentage of overlapping 1° cells, between the clusters of detrended Chl (Fig. 4.3) and those of SST, MLD and wind speed (Fig. 4.5). Notice that similarity does not drop below 50%, this being the minimum possible similarity when comparing two two-cluster patterns. The best matches are with SST and wind speed. The clusters of these two variables having an average similarity of about 74% between the two periods. However, these measures do not take into account how similar the overall shape of the clusters being compared are, only in the number of overlapping cells. As such, they should be considered in concert with a

comparison of Figs. 4.3 and 4.5, wherein it can be seen that the general patterns are also broadly similar.

Table 4.5 gives R^2 relationships between Clusters 1 and 2 for detrended two-cluster patterns of Chl, MLD, SST and wind speed. These relationships are given for both periods of each detrended two-cluster partition (Figs. 4.3 and 4.5). In similarity with Chl, it is apparent that MLD and wind speed also tend to be anti-correlated between their respective clusters. Thus, given the corresponding spatial similarity of these physical clusters to that of Chl (Table 4.4), these two variables may underlie the anti-correlation and bimodal pattern of Chl variability in the North Atlantic.

4.3.2 Dynamic Clusters

In the previous section, clusters were defined using entire time series and were spatially fixed in time. As such, they were useful in identifying and assessing overall patterns of Chl variability. However, such analysis gives no indication of how much the patterns and relationships may vary with time, or how representative the overall patterns are of any single time. By considering the time series in a series of overlapping windows, it is possible to define a continuous series of clusters that can behave dynamically with time.

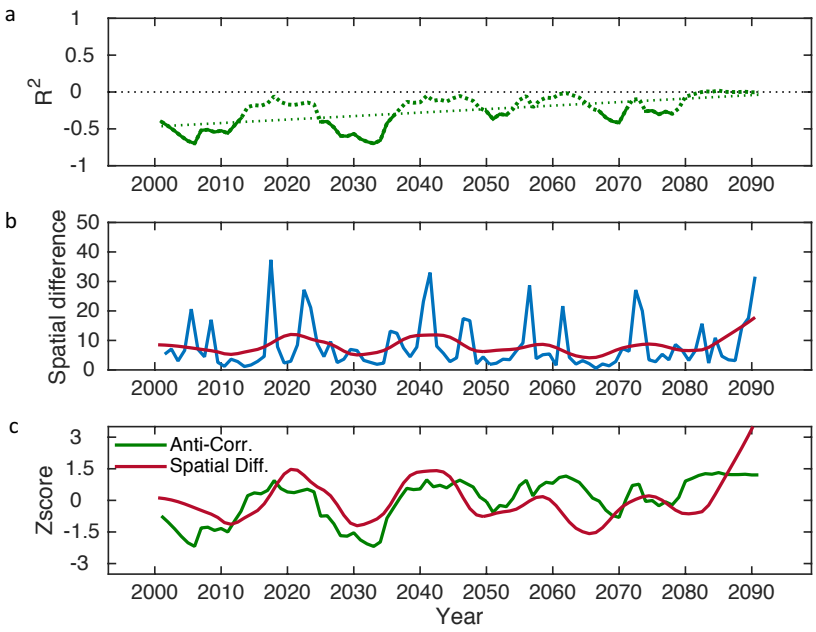


Fig. 4.6: Time series showing statistical properties for a running series (1993-2099) of 17-year, 2-cluster partitions of NEMO MEDUSA interannual Chl variability. Time series go from 2001 to 2090 with each year representing properties from successive 17-year clusters. Statistical properties represent a) significant (solid green) and non-significant (dashed green) correlations between Chl clusters, plus fitted linear trend (dotted green) significant at $p<0.05$ (see: S4.12), b) spatial difference (%) between clusters (solid blue) superimposed with a 17-year running average (solid red), and c) standardised forms of time series in a (solid green) and 17-year running average in b (solid red) superimposed upon each other for comparison.

When considered statically, Chl clusters appeared to be anti-correlated at shorter time scales (Fig. 4.4), although this was not apparent in overall correlations (see: §4.3.1). Fig. 4.6a reveals that the two clusters are anti-correlated when considered in 17-year segments, but are both more- and less-strongly anti-correlated at given intervals. Moreover, there is a weakening trend to this anti-

correlation with time (c.f. S4.12). Fig. 4.6b shows the spatial difference (100 minus % overlap) between each successive Chl cluster for the dynamic time series. A comparison of Figures 4.6a and b shows that the greatest spatial differences tend to occur at intervals when the clusters are less strongly anti-correlated. Fig. 4.6c further illustrates this by plotting standardised forms of these time series together. Clusters are defined in term of minimising sum of squared differences (§2.2.4). Thus, if the bimodal pattern is considered in terms of anti-correlation, then it is unclear if the movement of clusters reflect movement in the bimodal pattern, which is less spatially-stable when weakly anti-correlated, or if the movement of clusters reflect a varying dominance of the bimodal pattern upon interannual variability, with the strength of anti-correlation between clusters reflecting alignment with the bimodal pattern, or a combination of both these factors.

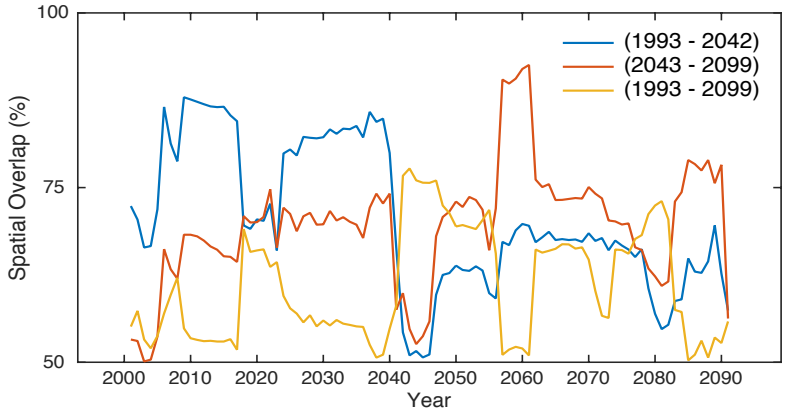


Fig. 4.7: Spatial overlap (%) between dynamic 17-year clusters and (blue) 1993-2042 detrended Chl clusters (Fig. 4.3a), (red) 2043-2099 detrended Chl clusters (Fig. 4.3b) and (yellow) 1993-2099 Chl clusters (Fig. 4.1b). Notice minimum absolute spatial similarity of 50%. For comparative purposes, values plotted here represent absolute similarities: irrespective of whether corresponding cluster-numbers overlap, or opposing cluster-numbers overlap. Dynamic clusters actually switch position with respect to the fixed detrended clusters (Fig. 4.10) in 2042.

Figure 4.7 shows the spatial overlap of dynamic clusters (e.g. Fig. 4.6a) with both un-detrended (Yellow line; Fig. 4.1b) and detrended (Blue and Red lines; Figs. 4.3a and b, respectively) static clusters. From this, it is immediately apparent that the bimodal pattern of interannual Chl variability in the North Atlantic is more spatially variable than indicated by the static clusters in Figs. 4.3a and b. Nevertheless, they share an on-average spatial similarity of about 78% with one or other of the detrended static clusters. As may be expected, the dynamic clusters are most similar to the 1993-2042 detrended clusters (Fig. 4.3a) before 2043, and to the 2043-2099 detrended clusters (Fig. 4.3b) thereafter. However, this generality is interrupted at intervals where the dynamic clusters more strongly overlap with the pattern of negative Chl trend (1993-

2099) in the North Atlantic (Fig. 4.1b). The most pronounced of these occurrences is from 2042 to 2048, which coincides with the shift in Chl trend identified in Fig. 4.2. It may be reasoned that the negative trend pattern may dominate dynamic clustering at intervals either when the negative trend in Chl is particularly strong (Fig. 4.2), or when the bimodal pattern of interannual variability is particularly weak (Fig. 4.6a).

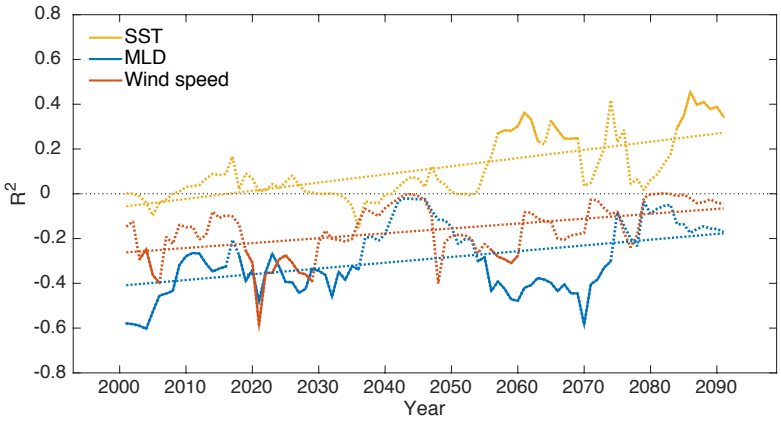


Fig. 4.8: R^2 relationships between Clusters 1 and 2 of dynamic 17-year SST (yellow), MLD (blue) and wind speed (red) two-cluster patterns. Time series are of yearly resolution (2001-2090) with each point representing values for a successive 17-year cluster partition spanning 1993-2099. Line styles denotes significant (solid) and non-significant (dotted) R^2 values. Dotted straight lines represent linear regressions fitted to correlation time series. These trends are each significant at $p<0.005$ (see: S4.12).

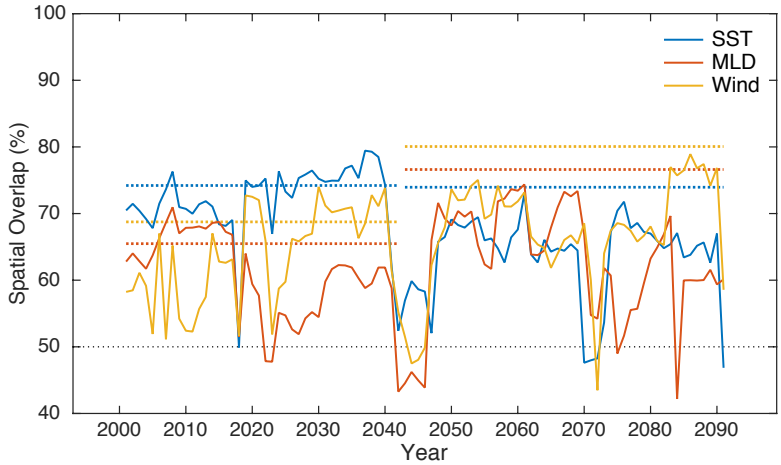


Fig. 4.9: Spatial overlap (%) between dynamic Chl clusters and dynamic SST (solid blue), MLD (solid red) and wind speed (solid yellow) clusters. For comparison, dotted lines denote corresponding relationships between fixed detrended Chl clusters and fixed detrended SST (dotted blue), MLD (dotted red) and wind speed (dotted yellow) clusters in Table 4.4. Dotted black line denotes minimum absolute spatial similarity. Values above 50% indicate increasing correspondence between similar cluster numbers (e.g. Cluster 1 with 1; Cluster 2 with 2). Values below 50% indicate increasing correspondence between opposite cluster numbers (e.g. Cluster 1 with 2; Cluster 2 with 1).

In addition to Chl, the explanatory variables are also clustered dynamically: Fig. 4.8 shows that MLD and wind speed also tend to form anti-correlated patterns within the North Atlantic, while that of SST tends to be correlated after ~2055. Figure 4.9 shows the spatial overlap between dynamic Chl clusters and those of SST, MLD and wind speed. As a temporal criterion was chosen for pairing the cluster-numbers of these physical variables to those of Chl (see: §4.2.4), the spatial overlap between them can drop below 50%. However, while this implies that at intervals the clusters are inversely matched in a spatial sense, their corresponding time series are more strongly correlated in a temporal sense when numbered this way.

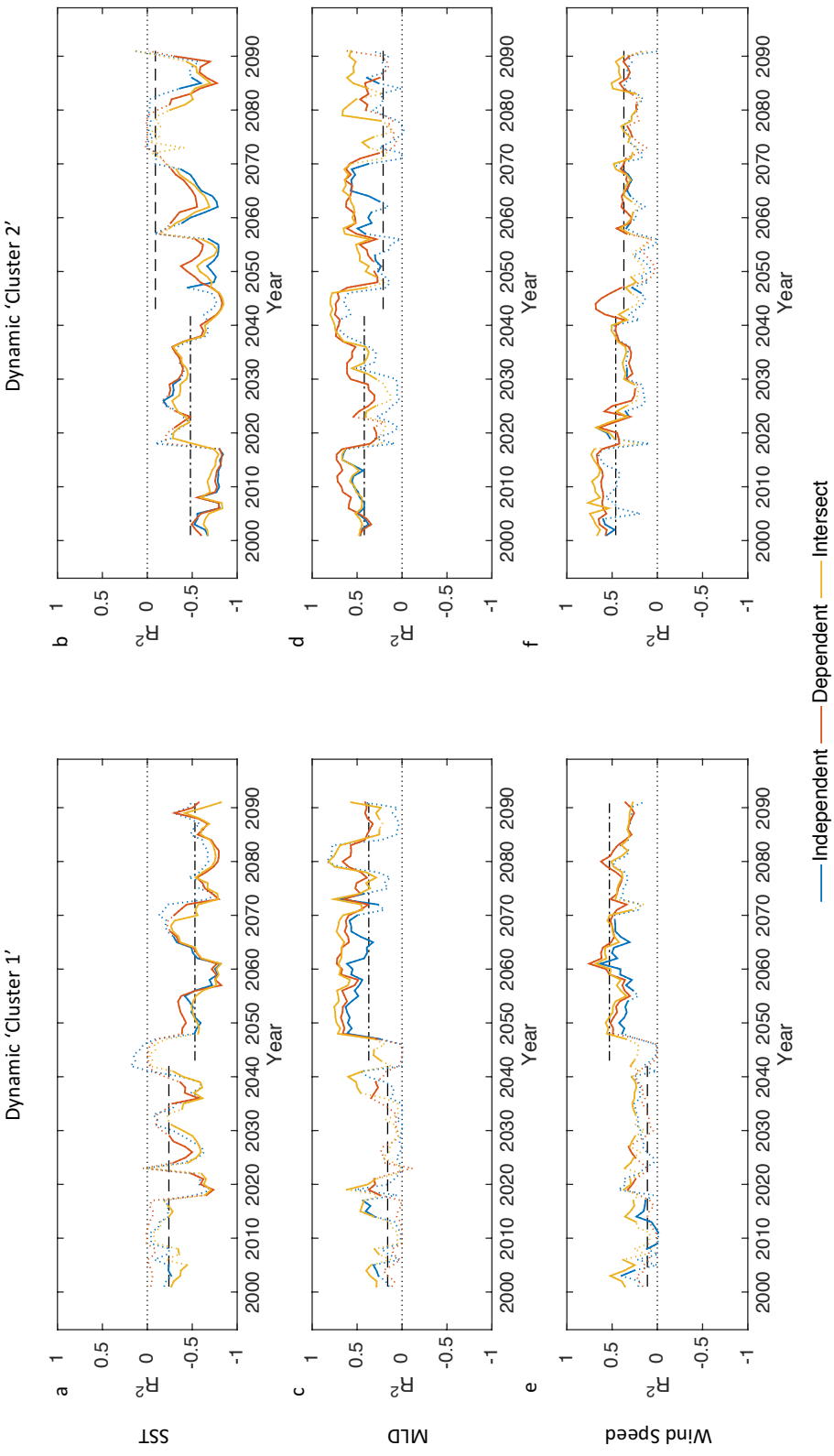


Fig. 4.10: Time series showing correlations between Chl and a-b) SST, c-d) MLD and e-f) wind speed for clusters 1 (a, c and e) and 2 (b, d and f), respectively. Colours relate to different methods of defining cluster-average time series before relating the two variables: dependent (red) refers to defining both Chl and SST/MLD/wind speed time series based upon the Chl clusters; independent (blue) refers to clustering both Chl and SST/MLD/wind speed individually and deriving the cluster-average time series from their respective clusters; intersect (yellow) refers to clustering both Chl and SST/MLD/wind speed individually and deriving the cluster-average time series from only the spatial overlap between corresponding clusters. Line style denotes significant (solid colours) and non-significant (dashed colours) R^2 values. For comparison, dashed black lines denote corresponding relationships between Chl and a-b) SST, c-d) MLD and e-f) wind speed for Clusters 1 (dash) and 2 (dot-dash) of detrended fixed clusters in Table 4.2. Notice that the dynamic Clusters 1 (a, c, e) and 2 (b, d, f) swap positions with respect to the fixed Clusters 1 (black dashed lines) and 2 (black dashed lines) after 2042.

Figure 4.10 shows R^2 relationships between Chl and physical variables. Relationships are shown for various choices of defining time series between corresponding dynamic clusters. In addition to spatial variability, it is apparent from these plots that the dynamic clusters also have stronger and weaker temporal relationships with Chl than when considered statically. This could be due to temporal variability and/or spatial variability owing to regional relationships (i.e. clusters moving between regions with different relationships).

No obvious trends or shifts are apparent in the relationships with Chl in this warming-scenario forecast model (Fig. 4.10). These dynamic relationships have not been rigorously tested for a trend or change in mean value however, as the clusters within these patterns switch positions in the middle of the time series (e.g. Fig. 4.7). As this transition happens smoothly with time, artefacts would be introduced into these time series if cluster numbers were switch back after the transition. Nevertheless, a comparison of these dynamic-cluster relationships with those of the fixed clusters they spatially approximate reveals that while these relationships are more variable, they remain broadly consistent with the latter (Fig. 4.10). In accordance with the fixed clusters, the strongest overall relationship appears to be with SST. Likewise, MLD and wind speed also have consistently stronger relationships with Chl in the dynamic cluster corresponding with fixed Cluster 2 than that of Cluster 1. This is not the case for SST however, which switches between the dynamic clusters in terms of strongest relationship with Chl. This is most likely because SST clusters tend towards being correlated, while those of MLD and wind speed tend towards being anti-correlated (Fig. 4.8).

Figures 4.10 a-f show that relationships between Chl and explanatory variables remain consistent, irrespective of whether relationships are based upon Chl clusters (Dependent), both Chl clusters and explanatory-variable clusters (Independent) or only the overlapping regions between respective clusters (Intersect). Thus, despite spatial differences (Fig. 4.9), SST, MLD and wind speed tend to express similar variability in the Chl clusters as they do within their respective clusters. Given that Chl, MLD and wind speed clusters tend to be anti-correlated (Fig. 4.6 a; Fig. 4.8 b-c), this implies that although the clusters include different regions, they are interrelated and express a common pattern of bi-modal variability.

4.4 Discussion

4.4.1 Warming Trend

Chl Trend

The global ocean biogeochemical model used in this study projects a decline in Chl concentration throughout the majority of the North Atlantic over the next century. This negative trend is well explained by contemporaneous changes in N, Si, SST, MLD and wind speed (Table 4.1; S4.3). Zooplankton and Fe are also well correlated with the projected trend, but phytoplankton are declining despite an increase in Fe concentration and decrease in zooplankton population, both of which would result in increased phytoplankton biomass if they were controlling the trend (c.f. §3.2.7). Of the physical variables, the strongest association is with increasing SST which is likely to reflect increased stratification and reduced mixing with continued warming (e.g. S4.5e and f). This is consistent with the declining trend in MLD and reduced surface nutrient concentrations of N and Si. However, while each of these trends may be explained in terms of surface warming and increased stratification, they may also be explained by reduced mixing due to the concomitant decline in wind speed. Both surface heating and wind stress affect the stability and vertical mixing of the water column and this analysis does not differentiate between them as drivers of the projected trends in SST, MLD, N, Si and Chl. Most attention in the literature concerns the possible effects of increased surface heating and stratification upon phytoplankton growth with continued warming (e.g. Doney, 2006). However, this analysis highlights changing wind speeds as a possible co-driver of these projected changes in the North Atlantic. The negative trend in Chl and corresponding relationships are projected for the whole of the North Atlantic, excluding isolated regions within the high-latitude subpolar North Atlantic and a zonal band corresponding to latitudes between 15-30°N (Fig. 4.1). Although there are discrepancies between models, this distribution agrees well with other global ocean biogeochemical model projections (e.g. Cabré et al., 2015; Steinacher et al., 2010).

Positive Chl trends are projected for discrete high-latitude regions, most noticeably within areas of the Labrador and Irminger Seas, and within the Canary current upwelling system (Fig. 4.1). As regions within the former are subject to deep wintertime convective mixing, they are not dependent upon wind-driven mixing *per se* for surface nutrient flux. As such, these regions may be truly light limited with production benefiting from earlier onset and sustained stratification. In the latter case, increased offshore winds are predicted to increase upwelling and production

within the Canary current upwelling system (Barton et al., 2013), however this has not been tested in the model. Lastly, the zonal band between ~15-30°N, excluding the Canary current upwelling system, has a negligible negative trend (S4.2b). This area corresponds to the central subtropical gyre in the model and has a low Chl concentration when averaged 1993-2001 (S4.2a). Indeed, NEMO-MEDUSA is recognized as having a low annual-mean Chl concentration within subtropical gyres (Yool et al., 2011). As such, the lack of Chl trend in this region may be attributed to low initial values, with little scope for further decline.

It is also noteworthy that in both this and other model projections (e.g. Cabré et al., 2015; Steinacher et al., 2010), the extent of subpolar regions where Chl responds positively to vertical mixing are limited in comparison to many 'present-day' observational studies of inter-annual variability (e.g. Kahru et al., 2010; Siegel et al., 2013; Wilson and Adamec, 2002). There are two potential reasons for this. Firstly, the regions where Chl responds positively to mixing tend to retreat in size with continued warming due to changing conditions (S4.7). Thus, these regions may be less expansive in centennial projections than in present-day observations. Secondly, given the limited expanse of positive relationships even at near-present times within the biogeochemical model used here (S4.7a), discrepancies may also be due to differences of using yearly-averaged versus monthly data. Seasonal cycles of phytoplankton growth and vertical mixing tend to be in-phase within the permanently stratified ocean and out-of-phase within the seasonally stratified ocean. Thus, residual seasonality can result in a distribution of relationships that reflect seasonal responses and not inter-annual variability (e.g. Barton et al., 2015). The close similarity of positive and negative subpolar relationship identified here to those identified in observational studies likewise using yearly-averaged data is consistent with this (Fig. 2.3). It appears to be general to the literature that long-term model projections are analysed using annual or multi-annual means, while shorter satellite-derived Chl studies are analysed using monthly-averages.

Chl Shift

There is a significant shift in the negative trend of Chl between 2042 and 2043 in the North Atlantic (Fig. 4.2; Table 4.3). This shift results in both a lower mean value and a greater rate of decline of Chl concentration in 2043 onwards. However, despite overall good fits between trends (Table 4.1), residual plots show that the explanatory variables do not explain this shift (S4.4). Nevertheless, other similarities do allude towards a connection. When the physical variables are averaged over the same areas as the Chl clusters (Fig. 4.1), there are significant shifts towards increased rates of SST warming both before (~2035 onwards; Cluster 1) and after (~2057 onwards;

Cluster 2) the shift in Chl. While the timing of the Chl shift more closely corresponds to an apparent decrease in the MLD variability and maximum mixing depth of both clusters in ~2046 onwards (S4.5). Although these dates do not exactly coincide with the 2042/2043 shift in Chl trend (Fig. 4.2), they may represent a more prolonged shift in forcing that the Chl has responded to. The decrease in MLD variability is most pronounced in Cluster 1, which includes areas of open ocean deep convective mixing in the subpolar North Atlantic (c.f. Figs. 2.5, 4.1b). As such, this shift may represent a sudden decline of open ocean deep convective mixing and overturning circulation in the North Atlantic, resulting in a more strongly stratified and nutrient limited surface waters (S4.5). This should be investigated in future analysis.

4.4.2 Interannual variability

Bimodal Pattern

If the long-term trend in Chl is removed, or negated by analysing the data in 17-year segments, then the pattern of yearly-averaged Chl variability in the North Atlantic over the next century is similar to that identified from observational analysis (c.f. Fig. 4.3; Fig. 2.1). As such, the pattern appears to be a consistent feature of Chl variability in the North Atlantic, representing two regions of anti-correlated variability. However, the strength of this anti-correlation tends to weaken over the next century (Table 4.5; Fig. 4.6a; S4.12). This is consistent with similar trends in the bimodal patterns of MLD and wind speed (Fig. 4.8; S4.12), and thus appears to be a feature of large-scale forcing, as opposed to the Chl response to it.

However, the trends do not necessary imply a cessation of bimodal variability in the North Atlantic. Given that the clusters are defined in terms of minimising within-cluster variance, the trends could also imply that the bimodal pattern and underlying forcing no longer dominate North Atlantic variability into the future. That is, the clusters could be moving apart from the bimodal pattern (or vice versa) due to it having a diminishing influence upon Chl, MLD and wind speed variability in the North Atlantic. However, a negative trend is also apparent in clusters that are fixed over the entire model projection (S4.10; S4.12). While fixed clusters cannot track changes in patterns of climatic variability, this does imply a decline in anti-correlation centred about the mean position of bimodal variability in the North Atlantic. In any case, these trends either imply diminishing bimodal variability or diminishing influence of that variability. Assuming these bimodal patterns are manifestations of large-scale climatic variability related to the NAO (e.g. §2),

this implies that its influence upon inter-annual Chl, MLD and wind speed variability in the North Atlantic will weaken with continued warming.

Whilst most attention in the literature is given to possible long-term changes in mean Chl concentration with warming (e.g. Doney, 2006), changes to inter-annual variability may have consequences upon ecosystem dynamics, including the timing, abundance, spatial distribution and interactions of phytoplankton, zooplankton and higher trophic levels in the North Atlantic (e.g. Ottersen et al., 2001; Stenseth and Ottersen, 2004). Variability associated with the NAO has also been linked to changes in phytoplankton community structure and carbon export flux within the North Atlantic (Henson et al., 2012).

A general increase in the spatial homogeneity of interannual Chl variability is also projected for the North Atlantic (S4.6). This increasing uniformity may be associated with an expansion of nutrient-limited permanently-stratified conditions with warming (S4.7). However, it does not appear to be related to the decline in bimodal variability identified here, as this decline in spatial variability is reflected in variance within-clusters, and not between them (S4.6).

Relationships

A number of significant relationships are identified between Chl variability and proxies for stratification and vertical mixing (SST, MLD and wind speed) within Clusters 1 and 2 when detrended or considered in 17 year segments (Table 4.2; Fig. 4.10). In both clusters, Chl variability is positively related to MLD and wind speed, and negatively related to SST. Thus, these results are consistent with the hypothesis that vertical mixing and nutrient flux limit PP (e.g. Behrenfeld et al., 2006). This is consistent with projected trends in this analysis, wherein continued warming and increases stratification result in reduced PP throughout the majority of the North Atlantic (§4.4.1).

However, relationships tend to be stronger in Cluster 2, representing the temperate and tropical eastern North Atlantic (Fig. 4.3 a-b), than in Cluster 1. This appears to be due to a greater consistency in the sign and strength of the local relationships comprising Cluster 2. For example, a comparison of Fig. 4.3 and S4.7 shows that the confines of Cluster 1 include many localities where Chl is negatively related to MLD over longer-term projections (c.f. §4.4.1). These localities are mostly confined to the fringes of the north-eastern North Atlantic and likely represent regions where phytoplankton are primarily light limited (e.g. Dutkiewicz et al., 2001; Siegel et al., 2002;

Sverdrup, 1953). While comparable local-scale analyses have not been conducted for inter-annual relationships *per se* in this chapter, the presence of light-limited relationships within the polar/sub-polar confines of Cluster 1 is consistent with pervious findings (§3). This is also consistent with a greater overall variance within Cluster 1 than Cluster 2 (not shown). In any case, overall interannual Chl variability in Clusters 1 and 2 responds negatively to increasing stratification and reduced macro-nutrient availability.

Interannual vs. Long-Term Relationships

This study shows that the strength of relationships with Chl can be highly variable when considered over shorter time periods (17 years), and yet still emerge as the dominant influence upon Chl concentration over the next century (Fig. 4.10; Table 4.1). This implies that while stratification and vertical mixing may not always dominate interannual Chl variability over shorter time periods, their control upon nutrient availability does emerge as the dominant control upon Chl concentration in the long-run. As observational studies of less than 17 years are often relied upon to test whether or not primary production is strongly dependent upon vertical mixing (e.g. Behrenfeld et al., 2006; Dave and Lozier, 2013), these results caution against the potential of concluding upon false-negative results when extrapolating shorter time series for long term projections. Of course, these results and the emergence of stratification as a dominant control upon Chl over the next century depends upon projected warming and the skill of NEMO MEDUSA to accurately forecast future climate change and the marine biological response to it. This is discussed in the concluding chapter of this thesis (§5.2.1).

Relationships get weaker with time?

Despite a weakening of the bimodal pattern in Chl variability (Fig. 4.6a; S4.12), there is no apparent weakening of the relationships between cluster-mean Chl concentration and SST, MLD or wind speed (Table 4.3; Fig. 4.10). There is thus no evidence that the control of vertical mixing upon Chl variability will change into the future.

Forcing

While the identification of cluster-average relationships between Chl and SST, MLD and wind speed is consistent with the hypothesis that vertical mixing and stratification drive interannual Chl variability, it does not necessarily imply that the bimodal pattern in Chl variability is associated with these variables. For example, there may be a general relationship between Chl and SST that is independent of the bimodal pattern in Chl variability.

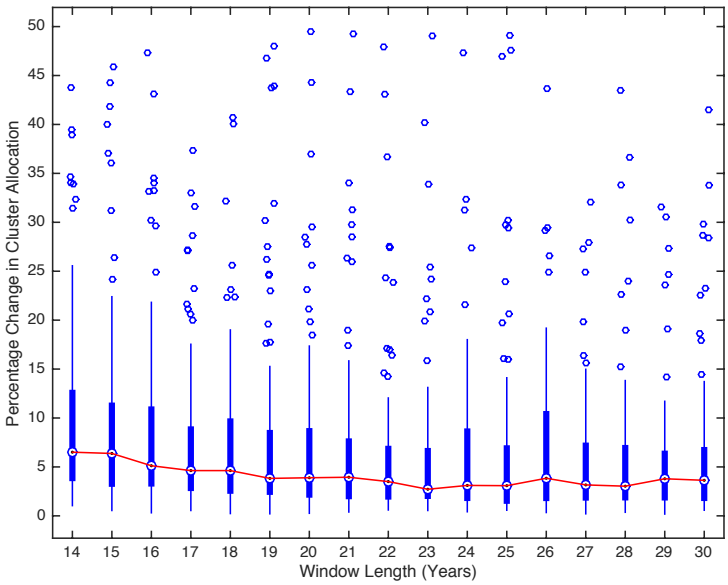
The sign of relationships between Chl and the three explanatory variables used in this study are consistent between Clusters 1 and 2. Therefore, if any one of these variables is singularly responsible for driving the anti-correlated bimodal pattern of Chl variability, it must also be anti-correlated between the two regions. Furthermore, assuming that the sign of these cluster-mean relationships are widely reflected at the local scale, then the variable in question should also have a similar pattern of bimodal variability to that of Chl. All three explanatory variables considered in this study have primary divisions that are broadly similar to that of Chl when detrended or considered over 17-year time periods (Table 4.4; Fig. 4.9). However, only MLD and wind speed are significantly anti-correlated between their respective clusters (Table 4.5; Fig. 4.8). Of these, MLD is most strongly anti-correlated, and is also most strongly related to Chl variability (Table 4.2). Therefore, of the variables considered, MLD is the best descriptor of the bimodal pattern in Chl. However, given the general similarity of all the variables, they may be interrelated and represent the same underlying forcing. In this case, wind speed may represent the underlying driver of the bimodal pattern as it is likely to be driving the associated changes in MLD. Thus, while Chl variability may be responding directly to interannual changes in MLD and nutrient availability, the overall bimodal pattern may be attributed to the effects of wind speed. Furthermore, although the bimodal patterns of Chl and physical variables identified in this chapter have not been related to the NAO index, their similarity to patterns of variability in Chl (§2; Patara et al. 2011) and physical variables (e.g. Visbeck et al., 2003) expected of this large-scale climatic oscillation implies their association. In this case, these patterns of variability would ultimately be driven by atmospheric forcing, including wind speed, and the large-scale response of phytoplankton to this climatic index would be consistent with other studies that have identified large-scale relationship between PP/Chl and climatic indices (Behrenfeld et al., 2006; Martinez et al., 2009; Patara et al., 2011).

4.5 Conclusion

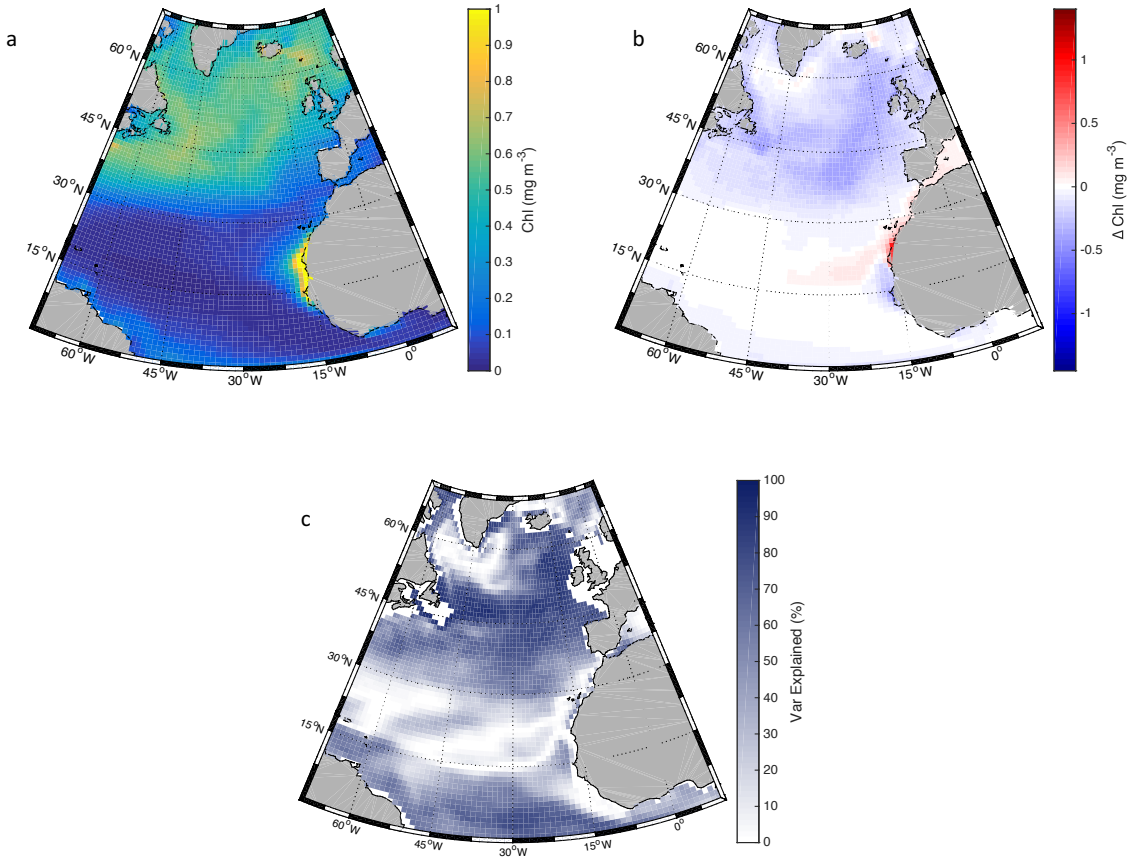
This study identifies a negative trend in Chl concentration throughout the majority of the North Atlantic over the next century. The decline is strongly related to reduced nutrient concentrations that coincide with decreasing wind speeds, MLDs, and increasing SSTs. This decline in Chl accelerates after ~2043 and is preceded by an accelerated increase in SST and increased wind speed variability (~2035). It also coincides with reduced MLD variability and mixing depths after ~2045.

When this long term trend is removed or negated by considering the time series in shorter sections, this study reveals a large-scale bimodal pattern of interannual Chl variability in the North Atlantic. The two clusters of this pattern are strongly associated with macro-nutrient variability, and, allowing for variability in the strength of relationships, are strongly related to SST, MLD and wind speed in both clusters. These explanatory variables also represent similar patterns of variability to that of Chl when clustered in the same way. Of the physical variables, SST and MLD have the strongest overall relationships with Chl. However, MLD and wind speed are also anti-correlated and therefore the best descriptors of the bimodal pattern in Chl. More generally, all of these variables represent similar two-cluster patterns and tend to be collinear in explaining Chl variability. As such, they likely represent an inter-related response to climatic variability associated with the NAO. Variability in the NAO is measured in terms of the pressure gradient between Azores-High and Icelandic-Low pressure systems. Although this has not been extracted and calculated for this model projection, the similarity of these bimodal patterns to that expected of the NAO implies that these patterns may be manifestations of this large-scale climatic oscillation. This is significant because the bimodal patterns in Chl, MLD and wind speed identified here tend to weaken over the next century. This may thus imply a weakening of variability associated with the NAO or its influence upon these variables in the North Atlantic with projected warming.

4.6 Supplementary Figures



S4.1: Box plot showing the spread, in percentage of observations being reassigned between consecutive clusters, for a range of different window lengths from 14 to 30 years. Each distribution is represented by a median (dot in circle), 25th and 75th percentiles (bottom and top of box, respectively) and whiskers extending out to 1.5 times the interquartile range below and above each box. Outliers are plotted individually (circles). The red line emphasises the variability in median between window lengths.



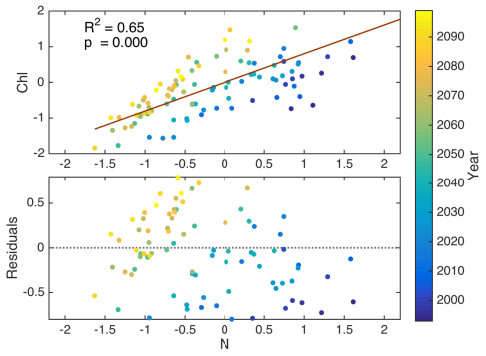
S4.2: Plots show: a) mean yearly average Chl concentration (1993-2002) in the North Atlantic, b) net change in yearly average Chl concentration between periods 1993-2002 and 2090-2099 and c) Percentage of overall yearly-average Chl variability (1993-2099) explained by local scale (1x1°) linear trends: 100(R^2).

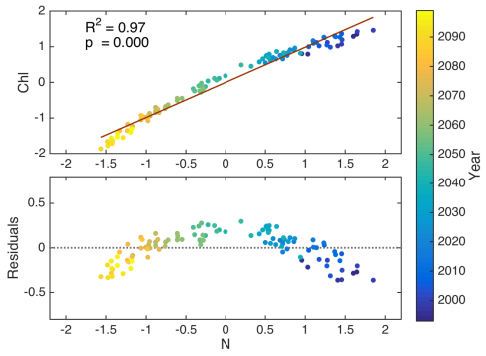
	Cluster 1	Cluster 2
Si	**+0.05**	**+0.80**
	0.019	0.000
Fe	**-0.25**	**-0.97**
	0.000	0.000
Zoo	**+0.91**	**+0.99**
	0.000	0.000

S4.3: R^2 values for the strength of relationship between Chl variability and each of Si and Fe for both Clusters 1 and 2 in Fig. 4.1. A positive or negative symbol before each R^2 value denotes the sign of relationship. Associated p-values are given underneath. Statistically significant ($p<0.05$) R^2 values are in bold type.

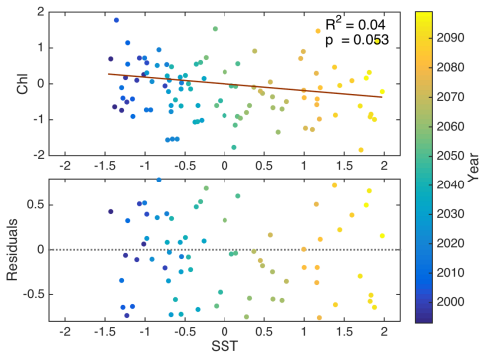
Cluster 1 Cluster 2

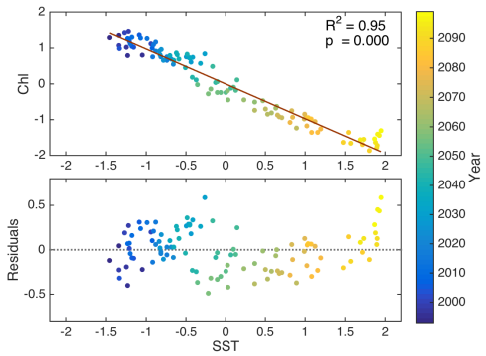
N



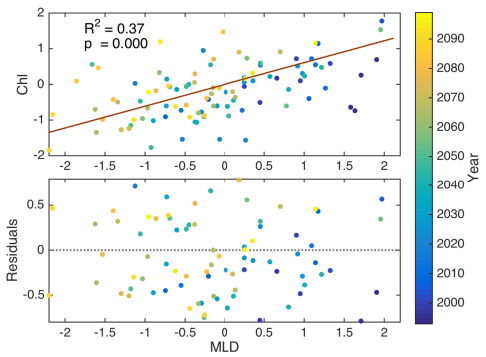


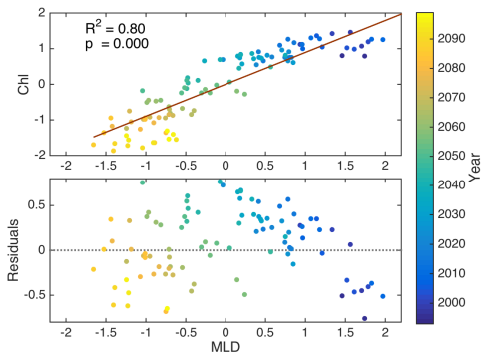
SST



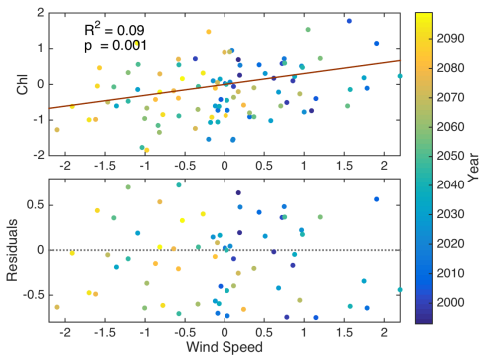


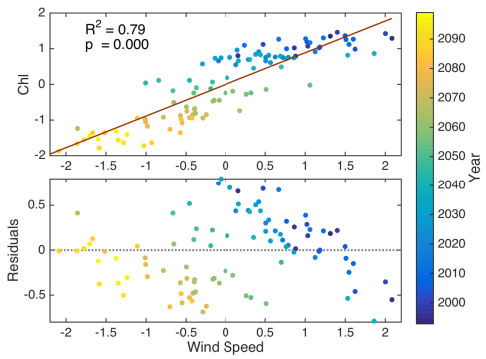
MLD

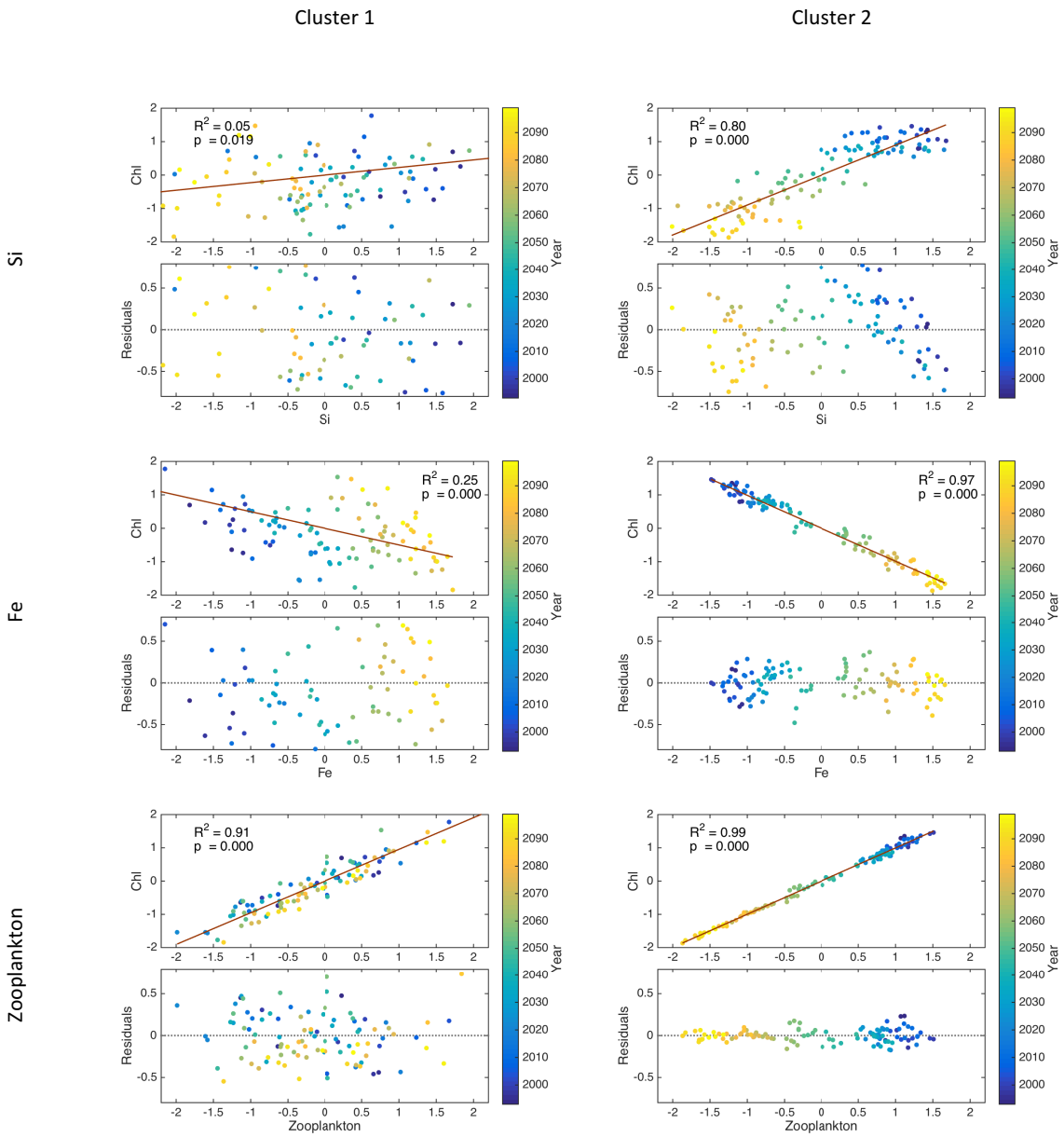




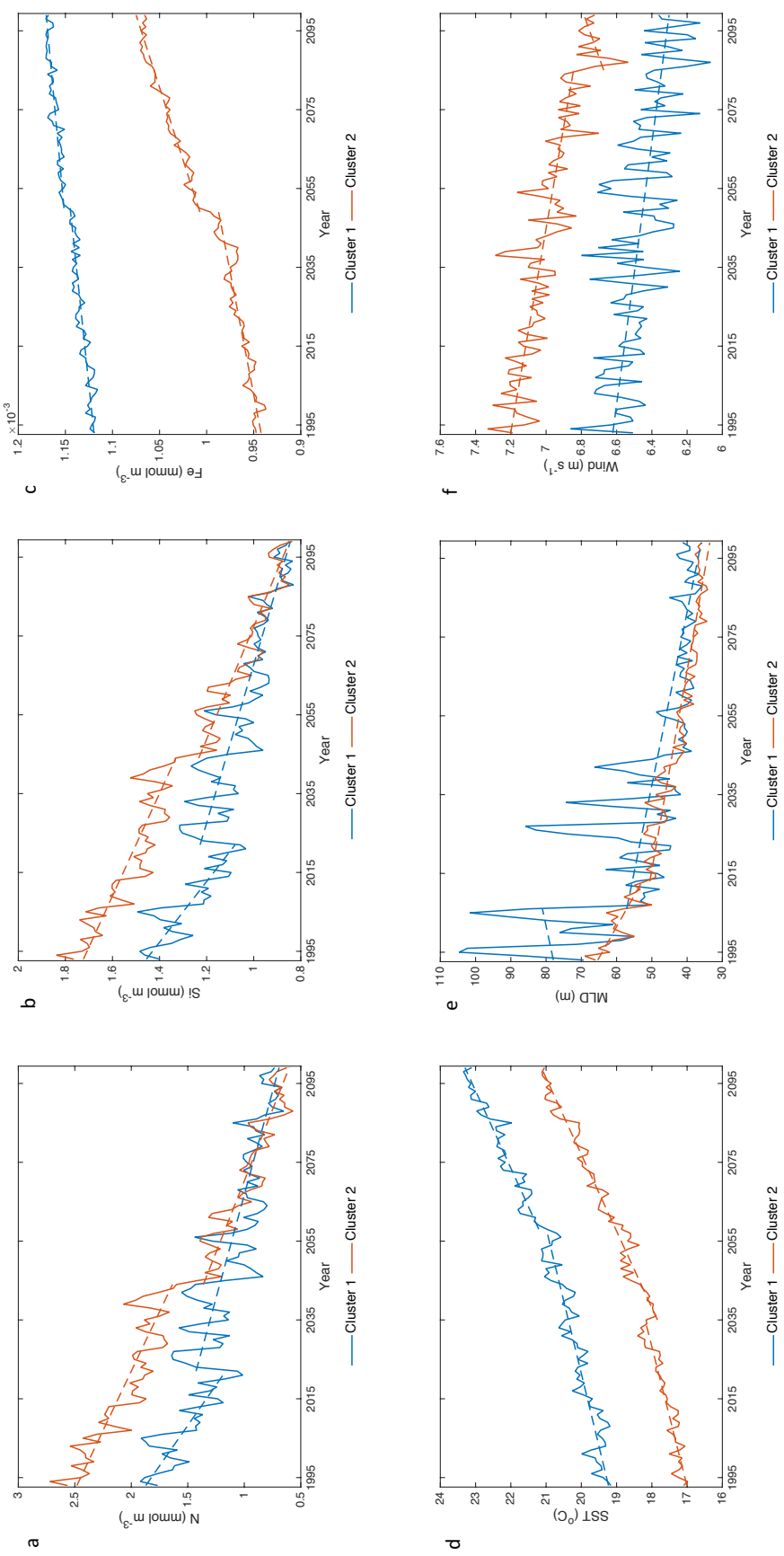
Wind Speed



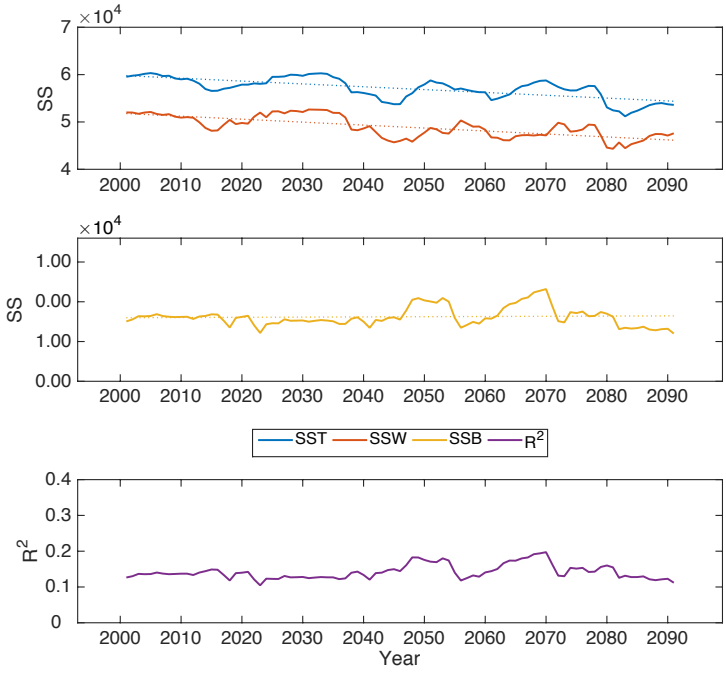




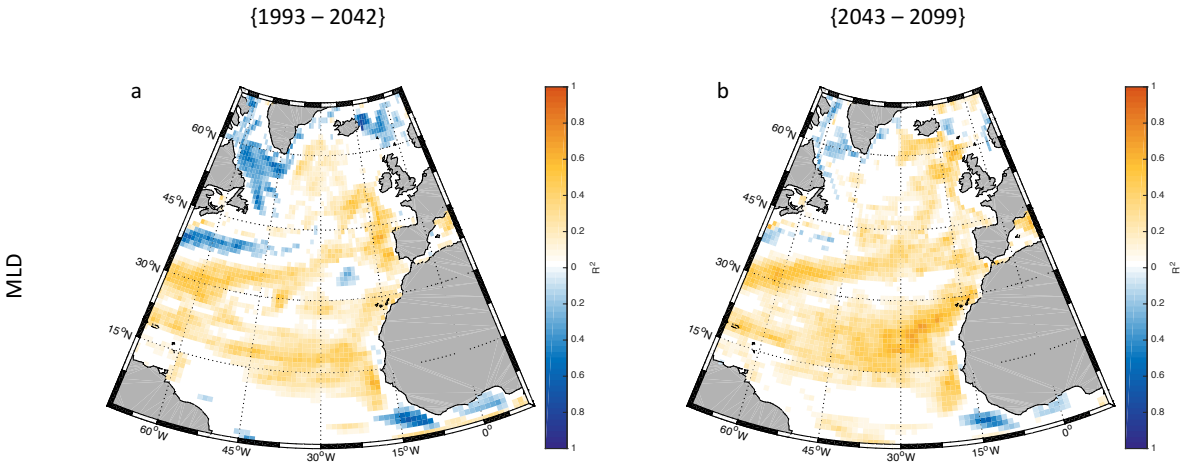
S4.4: Scatter plots and fitted simple linear regressions (red lines) showing relationships between Chl variability and each of N, SST, MLD, wind speed, N, Fe and zooplankton (rows; as labelled), for Clusters 1 and 2 (columns; as labelled) in Fig. 4.1. Associated R^2 and p-values are given in upper corners of plots, these correspond to those in Table 4.1 and S4.3. Residuals from each linear regression are plotted underneath the corresponding scatter plot.



S4.5: Cluster-mean time series of interannual a) N, b) Si, c) Fe, d) SST, e) MLD and f) Wind speed variability from Fig. 4.1b (solid line) superimposed with fitted trends (dashed lines). Two trends are fitted to the time series where this results in an overall lower SIC value then fitting a single trend to the times series.



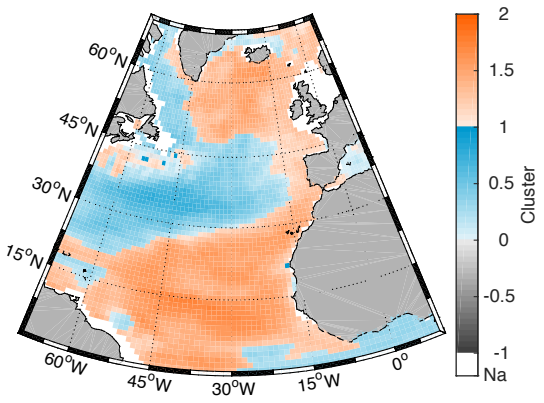
S4.6: Time series showing total (solid blue line; SST), within cluster (solid red line; SSW) and between cluster (solid yellow line; SSB) sum of squared differences in Chl variability, as well as SSB/SST (purple; R^2), for the dynamic two-cluster pattern in. Fig. 4.6. Each point is centred upon a 17-year window and represents values for that period Fig. 4.6. SST relates to the spatial variance in local scale (1x1°) time series (17 years) of interannual Chl variability throughout the North Atlantic. Likewise, SSW and SSB relate to within clusters and between clusters variance, respectively. As such, SSB/SST (R^2) indicates the total amount of local-scale North Atlantic Chl variability that is explained by the two-cluster pattern identified in Fig. 4.6. Dotted lines represent linear trends fitted to respective time series. Corresponding coefficients for the slope of trend ($ß_1$; SS/year) and p-values for significance test against $ß_1$=0 are: SST (dotted blue; $ß_1$ = -59.97; p = 0.000), SSW (dotted red; $ß_1$ = -62.55; p=0.000) and SSB (dotted yellow $ß_1$ = +2.57; p = 0.596).



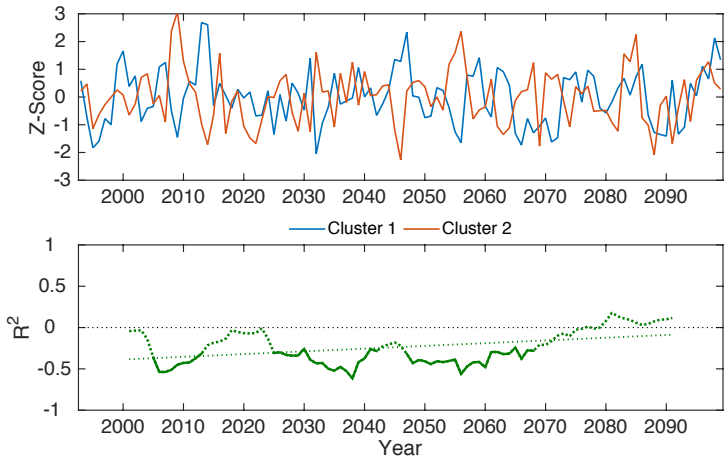
S4.7: Local-scale (1x1°) simple linear relationships between interannual Chl variability and MLD (a, b) over the periods 1993-2042 (a) and 2043-2099 (b). Colouring and shading indicates sign (red = positive; blue = negative) and strength (light to dark = weak to strong; white = insignificant) of R^2 relationship.

	{1993 – 2042}		{2043 – 2099}	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2
N	**+0.72**	**+0.91**	**+0.88**	**+0.90**
	0.000	0.000	0.000	0.000
Si	**+0.10**	+0.02	**+0.45**	**+0.12**
	0.025	0.313	0.000	0.009
Fe	**-0.49**	**-0.37**	**-0.72**	**-0.33**
	0.000	0.000	0.000	0.000

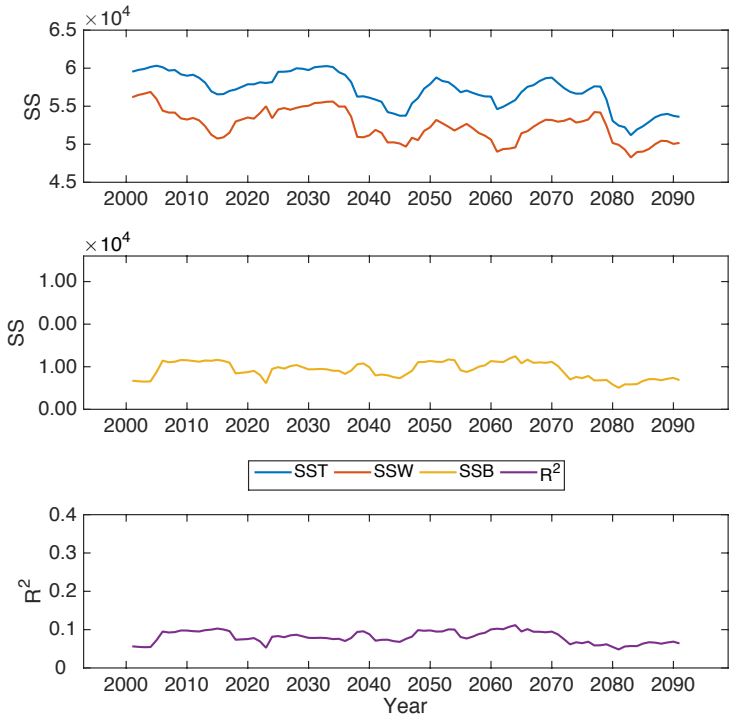
S4.8: lists R^2 values for the strength of relationship between detrended Chl variability and detrended N, Si and Fe (rows). Columns denote the corresponding cluster and period of each relationship, as represented in Fig. 4.3a (1993-2042) and b (2043-2099). A positive or negative symbol before each R^2 value denotes the sign of relationship. Associated p-values are given underneath. Statistically significant ($p<0.05$) R^2 values are in bold type.



S4.9: Two-cluster partition of NEMO MEDUSA detrended interannual Chl variability for 1993-2099. The time series is detrended by fitting separate linear trends to the two periods 1993-2042 and 2043-2099 (c.f. Fig. 4.2). Colouring and shading indicates cluster number (1 = blue; 2 = orange) and R^2 strength of relationship with cluster centroid (from light shading at R^2 = 0 towards dark shading at R^2 = 1); anti- correlated points are coloured grey irrespective of cluster number. Shelf sea localities less than 200m depth are omitted (Na; white).



S4.10: Cluster-mean time series of standardised detrended Chl variability for Clusters 1 (blue) and 2 (red) in S4.9, and the correlation between them (green). The correlation is calculated for 17-year segments of the cluster mean time series. Dotted sections of this time series indicted that the relationship between the clusters is insignificant Strait dotted line represents a linear regression fit to the correlation time series. This trend is significant at p<0.005 (S4.11).



S4.11: Time series showing total (solid blue line; SST), within cluster (solid red line; SSW) and between cluster (solid yellow line; SSB) sum of squared differences in Chl variability, as well as SSB/SST (purple; R^2), for the static two-cluster pattern in. S4.9. Each point is centred upon a 17-year window and represents values for that period.

	ß1	p-val
Dynamic Clusters		
SST	+0.0037	0.000
MLD	+0.0026	0.000
Wind speed	+0.0022	0.000
Chl	+0.0047	0.000
Stationary Clusters		
Chl	+0.0033	0.000

S4.12: Coefficients of gradient ($ß_1$) and corresponding p-values for linear trends fit to time series of between-cluster correlations of dynamic clusters (Figs. 4.6 and 4.8) and a stationary Chl cluster (S4.10). Trends are in unites of R^2/year and p-values represent significance against $ß_1 = 0$.

Chapter 5: Conclusion

5.1 Summary of conclusions

5.1.1 Overview

This study used a novel approach of clustering Chl data into regions of similar variability in order to investigate relationships with vertical mixing. Relationships were assessed at a number of spatial scales, using both observational data and output from an intermediate-complexity biological ocean model. In both observational and model studies, the Chl data sorted into similar two-cluster and seven-cluster patterns. At the two-clusters scale, these clusters form a familiar tri-pole pattern previously associated with NAO (Patara et al., 2011). This pattern and the Chl variability it represents are similar to, and correlated with, physical variables for vertical mixing (e.g. SST, stratification, MLD, wind speed). In both the observational and model data, these correlations tend to be stronger in the cluster representing the eastern Tropical and Temperate North Atlantic than in the cluster representing the western Equatorial, Subtropical and Subpolar North Atlantic. This difference appears to be due to the inclusion of areas of opposing relationship included within the latter cluster, most notably within the subpolar North Atlantic and western Atlantic between 0-20°N. Relationships pertaining to this cluster are thus better considered at smaller spatial scales. In the former cluster (representing the eastern Tropical and Temperate North Atlantic), model analysis shows that interannual Chl variability is dominated ($R^2 > 0.5$) by macro-nutrient variability (represented as Nitrate in the model), while in turn this is well explained ($R^2 > 0.75$) by proxies for vertical mixing (SST, stratification, MLD, wind speed), although strong collinearity amongst these physical variables means that none of them can be singled out as a choice explanatory variable.

As expected (e.g. Cabré et al., 2015; Steinacher et al., 2010), long-term changes in these physical variables result in a long-term negative trend in Chl concentration throughout most of the North Atlantic, with the notable exception of regions of deep convection in the subpolar North Atlantic (c.f. Fig. 2.5 and 4.1) and low-Chl regions of the subtropical gyre. However, if the projected time series are detrended, then the bimodal pattern of Chl variability in the North Atlantic is apparent throughout the next century. This pattern appears to become less anti-correlated with time. This is coincident with a weakening of similar bimodal patterns in wind speed and MLD, and may reflect a decline in the underlying bimodal variability (e.g. of large-scale atmospheric forcing

related to the NAO) and/or its dominance upon these variables in the North Atlantic. Nevertheless, there is no evidence for a change in the relationships between Chl and physical variables within this bimodal pattern over the next century. There is also an increase in the spatial homogeneity of interannual Chl variability in the North Atlantic, which may be associated with an expansion of permanently-stratified conditions with warming. This is unassociated with the trend in anti-correlation, however, as the decline in overall spatial variability is reflected within the clusters, and not between them.

The following three subsections more specifically address the aims set out in the first chapter of this thesis.

5.1.2 Importance of Spatial Scale

This study used a novel approach of clustering Chl data into regions of similar variability in order to investigate relationships with vertical mixing at intermediate spatial scales. It is shown that this method represents a compromise between ignoring spatial heterogeneity at larger scales and under-representing the importance of widespread coherent relationships at the local-scale. The importance of accounting for large-scale relationships is demonstrated as cluster-average relationships tend to remain significant despite excluding all similar locally-significant relationships from the cluster average. Thus, Chl can be strongly related to proxies for vertical mixing at large spatial scales, despite not being apparent in any of the locations that comprise the area. On the other hand, the importance of spatial heterogeneity is demonstrated as different relationships tend to dominate a given area/locality depending upon the spatial scale of analysis. Thus, relationships are dependent upon the spatial scale and significant regional/local relationships may be overlooked if analysis is only conducted at large/global scale. These results imply that the scale of analysis should be chosen to suit the aim of a study: e.g. a global/intermediate/local scale analysis should be chosen if a global/intermediate/local scale relationship is of interest. This study also shows that clustering can be used as a technique to emphasise relationships at different spatial scales. This is especially useful in the case of relationships with Chl, as large-scale climatic oscillations such as the NAO, MEI, Atlantic Multidecadal Oscillation and Pacific Decadal Oscillation tend to drive large-scale anti-correlated patterns of Chl variability within the North Atlantic and Global Ocean (e.g. §2; Behrenfeld et al. 2006; Martinez et al. 2009; Yoder & Kennelly 2003).

5.1.3 Interannual Chl variability and vertical mixing

A number of recent studies have reported a lack of widespread relationship between phytoplankton biomass and proxies relating to stratification and vertical mixing at the interannual time scale (see: §1.1). Understanding to what extent stratification controls interannual variability in phytoplankton biomass may have implications for the assumptions underlying long term model projections of PP in a warming climate (e.g. Steinacher et al., 2010). As such, one objective of this analysis was to assess the strength of relationship between interannual Chl variability and stratification/vertical mixing.

This analysis has identified strong and significant relationships between interannual Chl variability and proxies for vertical mixing at various spatial scales within the North Atlantic. A premise of this analysis was that although such relationships may be scarce at the local scale (e.g. Dave and Lozier, 2013), they may nevertheless dominate interannual Chl variability over large areas of the North Atlantic. This has been demonstrated in this study (see: Importance of Spatial Scale), although even at the local scale (1x1° cells) relationships between Chl and proxies tend to be well represented within large parts of the North Atlantic. These results are thus in contrast to recent studies that have reported a lack of relationship between Chl and stratification at the local scale. This discrepancy at the local scale may be a consequence of using yearly-mean and yearly-metric values in this thesis, as this does not assume that phytoplankton respond similarly to physical changes throughout the seasonal cycle, as e.g. monthly data would.

The large scale relationship with NAO implies a widespread association between Chl and vertical mixing. This appears to be mediated via wind speed and MLD that have similar anti-correlated patterns of variability to that of Chl. Over the 100 years of data available in a forecast model output, these variables are shown to be strongly related to Chl variability within the two clusters comprising the pattern, although correlations over the 14 years of observational data tend to be weaker or insignificant. In any case, this broad relationship appears to be ultimately mediated by the effect of vertical mixing upon nutrient flux within the more easterly of these two large-scale clusters and similarly within the subtropical band comprising part of the other more westerly cluster. However, relationships tend to be weaker or non-apparent within the western North Atlantic between 0-30°N and subpolar North Atlantic also comprising part of this western cluster. Given that these regions are strongly related to the NAO at the larger scale, it is apparent that Chl variability is responding to some physical forcing within these areas. However, these relationships are unidentified here. Weak or negative relationships between Chl and vertical mixing in parts of

the subpolar North Atlantic imply light limitation is controlling interannual variability in these areas, although such relationships with MLD-averaged PAR are only weakly and non-significantly apparent in model analysis. As discussed previously, this may be a limitation of using yearly-averages to capture a seasonal control. In the case of the western North Atlantic between 0-30°N, regional forcings remain unidentified.

5.1.4 Relationships over the next century

As identified in previous studies, Chl concentration is projected to decline throughout the majority of the North Atlantic over the next century due to increasing stratification and reduced nutrient availability. The only regions where Chl does not strongly decline are within areas of deep winter convection in the North Atlantic (c.f. Fig. 2.5; 4.1) and within a subtropical band corresponding to the subtropical gyre. However, while previous studies have attributed this increase in stratification to a warming of the surface ocean with climate change (e.g. Doney, 2006), this study shows that a long term decline in surface wind speeds may also be a contributing factor.

Despite this long-term change, there does not appear to be any significant change in the relationships between interannual Chl variability and proxies for vertical mixing over this period. While there is marked spatial variability in the position of the two-cluster Chl pattern over the next century, and an apparent long-term decline in the strength of anti-correlation between these clusters, there is no evidence for a significant change in the relationships with vertical mixing variables that underlie this pattern.

5.2 Evaluation of Data and Methods

This study used a number of observational datasets in addition to both hindcast and forecast model output to investigate the relationships and dynamics linking interannual Chl variability and vertical mixing in the North Atlantic. The most notable limitations of these datasets and the analysis applied to them are discussed in the following sections.

5.2.1 Analysis

This analysis has primarily focused upon interannual relationships over 14 and 17-year time scales. However, as highlighted in the introduction, one of the motivations for the analysis was to assess relationships in observational data for consistency with the premise that changes in vertical mixing may drive interannual and, via the same mechanisms, longer-term changes in PP.

This analysis has shown that most areas of the North Atlantic have a strong relationship ($R^2 \geq 0.43$) between inter-annual Chl variability and one or more variable for vertical mixing (Tables 2.2, 2.3) in observational time series; the notable exceptions being the subpolar and western Tropical North Atlantic. However, while such relationships imply that vertical mixing is a key or dominant driver of inter-annual Chl variability throughout large regions of the North Atlantic, they do not necessary imply that increasing stratification will dominate Chl variability into the future. As well as increasing stratification, other physical processes may also change in the future with projected climate change. Thus, vertical mixing may become more or less important in driving interannual Chl variability regionally as it, and other processes that control interannual Chl variability, change into the future. Such processes may include changes in horizontal advection, eddy driven vertical advection, predation or large-scale changes in circulation and pycnocline depth.

Without observational data spanning similar long-term changes in climate, the best forecasts for future change come from increasingly sophisticated and higher resolution global ocean biogeochemical models. As in this analysis, such models project an overall decline in PP/Chl within the North Atlantic and global mid-to-low latitude ocean, and increased PP/Chl at higher latitudes over the next century (§4; Bopp et al., 2013; Cabré et al., 2015; Steinacher et al., 2010). These projected trends are coincident with increasing stratification and reduced vertical mixing at all latitudes (Cabré et al., 2015; Steinacher et al., 2010), and are strongly related ($R^2 \geq 0.79$) to declining wind speeds, MLD's, macro-nutrient availability and increasing SST's throughout the area of projected Chl decline in this study (Table 4.1; S4.5). Thus, at present state-of-art model projections, which account for large-scale changes in wind forcing and meridional overturning circulation (e.g. McInnes et al., 2011; Weaver et al., 2012), appear to forecast an overall control of vertical mixing upon PP into the future. Although, reduced circulation and increasing temperature may also contribute to changes in PP within the mid-to-low and high latitude ocean, respectively (Steinacher et al., 2010).

5.2.2 Data

Observational

The observational analysis in this study included satellite-derived Chl, SST and wind speed datasets along with in situ density profile data. The length of these time series was primarily limited by the availability of satellite Chl data (1998 to 2012 at the time of analysis), although the spatiotemporal coverage of density profile data (and therefore validity of globally-gridded density profile datasets) increased markedly after the introduction of the Argo program in late 1999 (Riser et al., 2016). This 14-year duration is identified as a primary limitation in this analysis. Longer times series allow for relationships to be more easily and robustly identified. This is especially true when time series are short, as regression analysis will be more sensitive to outliers. The 14 years available here proved sufficient to identify strong, meaningful and statistically significant relationships with Chl at a range of spatial scales, despite the limited number of years (N=14). However, it is expected that these interannual relationships will become more apparent as the duration of observational time series increases.

Satellite Chl data are also limited to the surface ocean and as such do not account for the deep Chl maximum. This occurs when an exhaustion of surface nutrients results in phytoplankton accumulating around the top of the pycnocline (where nutrient concentrations increase with depth), and can be a perennial feature of the permanently-stratified ocean, or a seasonal feature of the seasonally-stratified ocean, provided the pycnocline is within the euphotic depth (Cullen, 1982). A potential solution to this problem is to use estimated values of depth integrated PP (e.g. Behrenfeld and Falkowski, 1997). However, such estimates can differ widely between different models (Friedrichs et al., 2009; Saba et al., 2010). These models also use SST as an input which may bias relationships with proxies for vertical mixing that tend to be collinear with SST. Instead, this study assumes that changes in phytoplankton biomass within the deep Chl maximum and surface mixed layer will co-vary together with changes in stratification and vertical mixing.

Model

A further limitation of this study may be the reliance upon a single biogeochemical model. Although not necessarily pertinent to the study of annual-mean variability, NEMO-MEDUSA (the coupled physical-biological model used in this study) hindcast output is similar to observed climatological distributions and seasonal variability of nutrients and PP (Yool et al., 2011), as well

as interannual variability in the phenology of PP (Henson et al., 2013). However, the model does tend to underestimate (overestimate) total PP in oligotrophic gyres (Fe limited regions) [Yool et al., 2011]. As different models have individual strengths and weaknesses in simulating observed physical (e.g. Flato et al., 2013) and biogeochemical (e.g. Cabré et al., 2015; Steinacher et al., 2010) ocean properties, a multi-model analysis may be more robust as it accounts for uncertainties between models. The reliance upon a single model is thus acknowledged as a limitation of the present study. Nevertheless, the geographical similarity of Chl clusters between the model and observational analyses indicates that the model is capturing the spatial variability in Chl, and thus presumably the underlying forcing, that is important to this analysis.

In the hindcast simulation of NEMO-MEDUSA used in this study, Fe concentration tends to be increasing throughout the North Atlantic. Unlike other nutrients in the model, Fe is added to the surface ocean as a new nutrient via aeolian dust deposition (Yool et al., 2011). The near-linear increase (S3.3; S3.5) and surface accumulation of Fe (S3.6b) in this hindcast simulation show that this is the origin of the increasing Fe trend within the North Atlantic. While its high concentration (above the half-saturation state for both phytoplankton groups throughout the North Atlantic in 1993 – the first year of output used in this analysis) means that Fe should have a limited effect upon phytoplankton growth, the use of multiple nutrient limitation terms in the model means that it will nevertheless have a positive effect upon growth. Given the linear increase in Fe concentration, it has not been possible to differentiate between Fe and other possible drivers of a Chl trend in regions where there is a positive trend in Chl. The solution to this in this analysis was to assess relationships between variables both before and after de-trending time series. However, as much of the difference between the two clusters was in their opposing trends, this may be missing the main drivers of the bimodal pattern.

5.2.3 Methods

Potential limitations of the analysis of data in this thesis are as follows:

- Strong relationships were identified between the NAO and the bimodal pattern of Chl variability in observational data. However, this association was weaker in the hindcast model analysis, and was not assessed in the forecast model. These relationships may have been stronger (could have been assessed) if the NAO index had been individually calculated for the hindcast (forecast) model. These could have been calculated using the reanalysis data, used to force the hindcast model, and the atmospheric output from the forecast model. However, time constraints prohibited this further analysis which may have identified/given greater support to an association between the bimodal pattern and NAO in the model analyses.

- While circumstantial evidence alludes to light limitation controlling interannual Chl variability within deep mixing regions of the subpolar North Atlantic, these relationships have not been significantly identified in this analysis. While MLD-average PAR was calculated for the hindcast simulation used in this analysis, no significant relationships were identified with Chl for the subpolar North Atlantic. This is most likely because relationships in the hindcast analysis were only assessed at the intermediary scale using cluster average time series, while the local scale analysis conducted using observational data (which did not include a variable for light availability) shows that inverse relationships between Chl and vertical mixing are limited to areas of open-ocean deep convection in the subpolar North Atlantic (c.f. Figs. 2.5, 2.6). The subpolar North Atlantic thus includes regions of both nutrient and light limitation, and the lack of significant relationships for the subpolar cluster in both the observational and hindcast analyses probably reflects this. A further limitation in identifying relationships with MLD-average PAR may also be in using yearly-average values. As phytoplankton growth at these latitudes coincides with seasonal stratification, relationships may be better identified by averaging PAR over the months that support phytoplankton growth. This could be easily calculated from assessing seasonal climatologies of Chl at each locality or within different latitudinal bands.

5.3 Future work

While this analysis has progressed the understanding of how vertical mixing is linked to interannual Chl variability in the North Atlantic, there are still questions and uncertainties to be addressed. Some of these lead directly on from this analysis and may be best addressed using a similar method of clustering data. Suggestions for future work are as follows:

- It would be interesting to extend this analysis to the global scale. This should reveal how widespread the relationship between Chl and stratification is throughout the ocean (e.g. Behrenfeld et al., 2006; Dave and Lozier, 2013). The interannual relationships between phytoplankton and other large-scale climatic indices can be assessed within the global ocean. This may reveal teleconnections in Chl variability between different ocean basins related to large-scale wind patterns and climatic indices. Plus, the stability of these relationships and patterns of variability to continued warming can also be assessed for the global ocean.

- Further investigation is also required to identify what drives Chl variability in the western Atlantic between 0-30°N. Chlorophyll variability in this region appears to be strongly related to NAO variability at both the local scale (Fig. 2.7) and at the large scale (Table 2.2). However, while this appears to be related to SST, stratification and MLD in the hindcast model analysis (Table 3.3), there is little evidence for any relationships between Chl and vertical mixing in this region within the observational analysis (Fig. 2.6; Fig. and Table 2.3). Longer observational time series may reveal such relationships. However, other variables such as advection, or even Amazon fluvial discharge should perhaps be considered as possible forcings in this region.

- This analysis also raises questions relating to the forecast trend in Chl over the next century. While a long term decline in Chl is forecast for the majority of the North Atlantic (excluding subpolar areas of open-ocean deep convection and subtropical latitudes) over the next century, the rate of decline increases after 2043 (Fig. 4.2). This shift appears to be concomitant with similar shifts in the rate of change of nutrients. The shift also coincides with an apparent reduction in MLD variability and is preceded, by about 9 years, by an increase in the rate of SST warming (S4.5). Such an increased rate of decline in Chl/PP has obvious impacts for biogeochemical cycling, food available to higher trophic levels and and carbon export in the ocean. The cause of these changes and their

relationship to Chl should thus be further explored in future studies. For example, is this increased rate of decline unique to NEMO-MEDUSA in the North Atlantic? What causes the shift?

- Aside from the shift in trend, the long term decline in Chl is concomitant with increasing stratification and declining MLDs and wind speeds. Each of these processes have been shown to be related to Chl variability at the interannual time scale, and it would be interesting to partition their effects upon Chl into the future. The canonical view is that enhanced surface warming and stratification will drive a decline in Chl, but this analysis suggests that declining wind speeds and mixing may also play a role.

5.4 Closing summary

While a number of questions relating to drivers of interannual variability in phytoplankton biomass remain, this thesis has shown that processes of vertical mixing tend to be strong drivers of interannual Chl variability over large regions of the North Atlantic. It is shown that while these processes may only be weakly related at local scales, they may nevertheless accumulatively dominate overall variability at larger scales. It is shown that these relationships are mediated by nutrient availability throughout most of the mid-to-low latitude North Atlantic; seasonal-dependence and spatial variability appear to obscure the identification of interannual relationships at subpolar latitudes. Furthermore, a basin-scale bimodal pattern of Chl variability in the North Atlantic is shown to be related to similar patterns in wind speed, SST, stratification and vertical mixing. These patterns appear to be interrelated and to express climatic variability associated with the NAO. However, the strength of these bimodal patterns, in terms of their anti-correlation and/or their influence upon North Atlantic variability, is found to diminish over the next century with continued warming. This thesis thus highlights high-latitude relationships with phytoplankton, and the stability of large-scale patterns of interannual phytoplankton biomass, as major uncertainties to the understanding of phytoplankton dynamics and their response continued warming over the next century.

Chapter 6: References

Antoine, D., 2005. Bridging ocean color observations of the 1980s and 2000s in search of long-term trends. J. Geophys. Res. 110, C06009. doi:10.1029/2004JC002620

Banse, K., 1992. Grazing, Temporal Changes of Phytoplankton Concentrations, and the Microbial Loop in the Open Sea BT - Primary Productivity and Biogeochemical Cycles in the Sea, in: Falkowski, P.G., Woodhead, A.D., Vivirito, K. (Eds.), . Springer US, Boston, MA, pp. 409–440. doi:10.1007/978-1-4899-0762-2_22

Barnston, A.G., Livezey, R.E., 1987. Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns. Mon. Weather Rev. 115, 1083–1126. doi:10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2

Barton, A.., Greene, C.., Monger, B.., Pershing, A.., 2003. The Continuous Plankton Recorder survey and the North Atlantic Oscillation: Interannual- to Multidecadal-scale patterns of phytoplankton variability in the North Atlantic Ocean. Prog. Oceanogr. 58, 337–358. doi:10.1016/j.pocean.2003.08.012

Barton, A.D., Lozier, M.S., Williams, R.G., 2015. Physical controls of variability in North Atlantic phytoplankton communities. Limnol. Oceanogr. 60, 181–197. doi:10.1002/lno.10011

Barton, E.D., Field, D.B., Roy, C., 2013. Canary current upwelling: More or less? Prog. Oceanogr. 116, 167–178. doi:http://dx.doi.org/10.1016/j.pocean.2013.07.007

Bates, N.R., 2001. Interannual variability of oceanic CO2 and biogeochemical properties in the Western North Atlantic subtropical gyre. Deep Sea Res. Part II Top. Stud. Oceanogr. 48, 1507–1528. doi:http://dx.doi.org/10.1016/S0967-0645(00)00151-X

Beaulieu, C., Cole, H., Henson, S., Yool, A., Anderson, T., de Mora, L., Buitenhuis, E.T., Butenschön, M., Totterdell, I.J., Allen, J.I., 2016. Marine regime shifts in ocean biogeochemical models: a case study in the Gulf of Alaska. Biogeosciences 13, 4533–4553. doi:10.5194/bg-13-4533-2016

Behrenfeld, M.J., 2010. Abandoning Sverdrup's Critical Depth Hypothesis on phytoplankton blooms. Ecology 91, 977–989. doi:10.1890/09-1207.1

Behrenfeld, M.J., Falkowski, P.G., 1997. Photosynthetic rates derived from satellite-based chlorophyll concentration. Limnol. Oceanogr. 42, 1–20. doi:10.4319/lo.1997.42.1.0001

Behrenfeld, M.J., O'Malley, R.T., Siegel, D.A., McClain, C.R., Sarmiento, J.L., Feldman, G.C., Milligan, A.J., Falkowski, P.G., Letelier, R.M., Boss, E.S., 2006. Climate-driven trends in contemporary ocean productivity. Nature 444, 752–755.

Bidigare, R.R., Chai, F., Landry, M.R., Lukas, R., Hannides, C.C.S., Christensen, S.J., Karl, D.M., Shi, L., Chao, Y., 2009. Subtropical ocean ecosystem structure changes forced by North Pacific climate variations. J. Plankt. Res. 31, 1131–1139. doi:10.1093/plankt/fbp064

Bopp, L., Monfray, P., Aumont, O., Dufresne, J.-L., Le Treut, H., Madec, G., Terray, L., Orr, J.C., 2001. Potential impact of climate change on marine export production. Global Biogeochem. Cycles 15, 81–99. doi:10.1029/1999GB001256

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

Boyce, D.G., Dowd, M., Lewis, M.R., Worm, B., 2014. Estimating global chlorophyll changes over the past century. Prog. Oceanogr. 122, 163–173. doi:10.1016/j.pocean.2014.01.004

Boyce, D.G., Lewis, M.R., Worm, B., 2010. Global phytoplankton decline over the past century. Nature 466, 591–596.

Broecker, W.S., 1982. Ocean chemistry during glacial time. Geochim. Cosmochim. Acta 46, 1689–1705. doi:10.1016/0016-7037(82)90110-7

Cabanes, C., Grouazel, A., von Schuckmann, K., Hamon, M., Turpin, V., Coatanoan, C., Paris, F., Guinehut, S., Boone, C., Ferry, N., de Boyer Montégut, C., Carval, T., Reverdin, G., Pouliquen, S., Le Traon, P.-Y., 2013. The CORA dataset: validation and diagnostics of in-situ ocean temperature and salinity measurements. Ocean Sci. 9, 1–18. doi:10.5194/os-9-1-2013

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

Calinski, T., Harabasz, J., 1974. A dendrite method for cluster analysis. Commun. Stat. 3, 1–27.

Cayan, D.R., 1992. Latent and Sensible Heat Flux Anomalies over the Northern Oceans: Driving the Sea Surface Temperature. J. Phys. Oceanogr. 22, 859–881. doi:10.1175/1520-0485(1992)022<0859:LASHFA>2.0.CO;2

Collins, W.J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C.D., Joshi, M., Liddicoat, S., 2011. Development and evaluation of an Earth-system model–HadGEM2. Geosci. Model Dev. 4, 1051–1075.

Corno, G., Karl, D.M., Church, M.J., Letelier, R.M., Lukas, R., Bidigare, R.R., Abbott, M.R., 2007. Impact of climate forcing on ecosystem processes in the North Pacific Subtropical Gyre. J. Geophys. Res. Ocean. 112, n/a-n/a. doi:10.1029/2006JC003730

Cullen, J.J., 1982. The Deep Chlorophyll Maximum: Comparing Vertical Profiles of Chlorophyll a. Can. J. Fish. Aquat. Sci. 39, 791–803. doi:10.1139/f82-108

Dave, A.C., 2014. Correction to "Examining the global record of interannual variability in stratification and marine productivity in the low-and mid-latitude ocean." J. Geophys. Res. Ocean. 119, 2121–2128. doi:10.1002/2013JC009723

Dave, A.C., Lozier, M.S., 2013. Examining the global record of interannual variability in stratification and marine productivity in the low-latitude and mid-latitude ocean. J. Geophys.

Res. Ocean. 118, 3114–3127. doi:10.1002/jgrc.20224

Dave, A.C., Lozier, M.S., 2010. Local stratification control of marine productivity in the subtropical North Pacific. J. Geophys. Res. 115, C12032. doi:10.1029/2010JC006507

de Boyer Montégut, C., 2004. Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. J. Geophys. Res. 109, C12003. doi:10.1029/2004JC002378

Devred, E., Sathyendranath, S., Platt, T., 2007. Delineation of ecological provinces using ocean colour radiometry . Mar. Ecol. Prog. Ser. 346, 1–13.

Doney, S.C., 2006. Oceanography: Plankton in a warmer world. Nature 444, 695–696.

Doney, S.C., Ruckelshaus, M., Emmett Duffy, J., Barry, J.P., Chan, F., English, C.A., Galindo, H.M., Grebmeier, J.M., Hollowed, A.B., Knowlton, N., Polovina, J., Rabalais, N.N., Sydeman, W.J., Talley, L.D., 2011. Climate Change Impacts on Marine Ecosystems. Ann. Rev. Mar. Sci. 4, 11–37. doi:10.1146/annurev-marine-041911-111611

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography (Cop.). 36, 27–46. doi:10.1111/j.1600-0587.2012.07348.x

DRAKKAR Group, 2007. Eddy-permitting Ocean Circulation Hindcasts Of Past Decades. CLIVAR Exch. 12, 8–14.

Duarte, C.M., Cebrián, J., 1996. The fate of marine autotrophic production. Limnol. Oceanogr. 41, 1758–1766. doi:10.4319/lo.1996.41.8.1758

Dutkiewicz, S., Follows, M., Marshall, J., Gregg, W.W., 2001. Interannual variability of phytoplankton abundances in the North Atlantic. Deep Sea Res. Part II Top. Stud. Oceanogr. 48, 2323–2344. doi:10.1016/S0967-0645(00)00178-8

Edwards, M., Reid, P., Planque, B., 2001. Long-term and regional variability of phytoplankton biomass in the Northeast Atlantic (1960–1995). ICES J. Mar. Sci. J. du Cons. 58, 39–49. doi:10.1006/jmsc.2000.0987

Eppley, R., 1972. Temperature and phytoplankton growth in the sea. Fish. Bull.

Falkowski, P., Scholes, R.J., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Högberg, P., Linder, S., Mackenzie, F.T., Moore III, B., Pedersen, T., Rosenthal, Y., Seitzinger, S., Smetacek, V., Steffen, W., 2000. The Global Carbon Cycle: A Test of Our Knowledge of Earth as a System. Science (80-.). 290, 291–296.

Falkowski, P.G., 1994. The role of phytoplankton photosyntesis in global biogeochemical cycles. Photosynth. Res. 39, 235–258.

Falkowski, P.G., Wilson, C., 1992. Phytoplankton productivity in the North Pacific ocean since 1900 and implications for absorption of anthropogenic CO2. Nature 358, 741–743.

Feng, J., Durant, J.M., Stige, L.C., Hessen, D.O., Hjermann, D.Ø., Zhu, L., Llope, M., Stenseth, N.C., 2015. Contrasting correlation patterns between environmental factors and chlorophyll levels in the global ocean. Global Biogeochem. Cycles 29, 2095–2107. doi:10.1002/2015GB005216

Field, C.B., 1998. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. Science (80-.). 281, 237–240. doi:10.1126/science.281.5374.237

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S.C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., Rummukainen, M., 2013. Evaluation of Climate Models, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. pp. 741–866. doi:10.1017/CBO9781107415324

Follows, M., Dutkiewicz, S., 2001. Meteorological modulation of the North Atlantic spring bloom. Deep Sea Res. Part II Top. Stud. Oceanogr. 49, 321–344. doi:10.1016/S0967-0645(01)00105-9

Friedrichs, M.A.M., Carr, M.-E., Barber, R.T., Scardi, M., Antoine, D., Armstrong, R.A., Asanuma, I., Behrenfeld, M.J., Buitenhuis, E.T., Chai, F., Christian, J.R., Ciotti, A.M., Doney, S.C., Dowell, M., Dunne, J., Gentili, B., Gregg, W., Hoepffner, N., Ishizaka, J., Kameda, T., Lima, I., Marra, J., Mélin, F., Moore, J.K., Morel, A., O'Malley, R.T., O'Reilly, J., Saba, V.S., Schmeltz, M., Smyth, T.J., Tjiputra, J., Waters, K., Westberry, T.K., Winguth, A., 2009. Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean. J. Mar. Syst. 76, 113–133. doi:http://dx.doi.org/10.1016/j.jmarsys.2008.05.010

González Taboada, F., Anadón, R., 2014. Seasonality of North Atlantic phytoplankton from space: impact of environmental forcing on a changing phenology (1998-2012). Glob. Chang. Biol. 20, 698–712. doi:10.1111/gcb.12352

Gregg, W.W., 2005. Recent trends in global ocean chlorophyll. Geophys. Res. Lett. 32, L03606. doi:10.1029/2004GL021808

Gregg, W.W., Conkright, M.E., 2002. Decadal changes in global ocean chlorophyll. Geophys. Res. Lett. 29, 20-1-20–4. doi:10.1029/2002GL014689

Grömping, U., 2007. Estimators of Relative Importance in Linear Regression Based on Variance Decomposition. Am. Stat. 61, 139–147. doi:10.1198/000313007X188252

Grömping, U., 2006. Relative Importance for Linear Regression in R: The Package relaimpo. J. Stat. Software; Vol 1, Issue 1.

Gruber, N., 2011. Warming up, turning sour, losing breath: ocean biogeochemistry under global change. Philos. Trans. R. Soc. London A Math. Phys. Eng. Sci. 369, 1980–1996.

Gruber, N., Sarmiento, J.L., 2002. Chapter 9 . LARGE-SCALE BIOGEOCHEMICAL –

PHYSICAL INTERACTIONS IN ELEMENTAL CYCLES, in: Robinson, A.R., McCarthy, J.J., Rothschild, B.J. (Eds.), The Sea. John Wiley & Sons, Inc., New York, New York.

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. J. R. Stat. Soc. Ser. C (Applied Stat. 28, 100–108. doi:10.2307/2346830

Hartman, S.E., Hartman, M.C., Hydes, D.J., Jiang, Z.-P., Smythe-Wright, D., González-Pola, C., 2014. Seasonal and inter-annual variability in nutrient supply in relation to mixing in the Bay of Biscay. Deep Sea Res. Part II Top. Stud. Oceanogr. 106, 68–75. doi:10.1016/j.dsr2.2013.09.032

Hartman, S.E., Larkin, K.E., Lankhorst, M., Hydes, D.J., 2010. Seasonal and inter-annual biogeochemical variations in the Porcupine Abyssal Plain 2003–2005 associated with winter mixing and surface circulation. Deep Sea Res. Part II Top. Stud. Oceanogr. 57, 1303–1312. doi:10.1016/j.dsr2.2010.01.007

Hellerman, S., Rosenstein, M., 1983. Normal Monthly Wind Stress Over the World Ocean with Error Estimates. J. Phys. Oceanogr. 13, 1093–1104. doi:10.1175/1520-0485(1983)013<1093:NMWSOT>2.0.CO;2

Henson, S.A., Dunne, J.P., Sarmiento, J.L., 2009. Decadal variability in North Atlantic phytoplankton blooms. J. Geophys. Res. 114, C04013. doi:10.1029/2008JC005139

Henson, S.A., Robinson, I., Allen, J.T., Waniek, J.J., 2006. Effect of meteorological conditions on interannual variability in timing and magnitude of the spring bloom in the Irminger Basin, North Atlantic. Deep Sea Res. Part I Oceanogr. Res. Pap. 53, 1601–1615. doi:10.1016/j.dsr.2006.07.009

Henson, S.A., Sarmiento, J.L., Dunne, J.P., Bopp, L., Lima, I., Doney, S.C., John, J., Beaulieu, C., 2010. Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. Biogeosciences 7, 621–640. doi:10.5194/bg-7-621-2010

Henson, S., Cole, H., Beaulieu, C., Yool, A., 2013. The impact of global warming on seasonality of ocean primary production. Biogeosciences 10, 4357–4369. doi:10.5194/bg-10-4357-2013

Henson, S., Lampitt, R., Johns, D., 2012. Variability in phytoplankton community structure in response to the North Atlantic Oscillation and implications for organic carbon flux. Limnol. Oceanogr. 57, 1591–1601. doi:10.4319/lo.2012.57.6.1591

Hurrell, J.W., 1995. Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation. Science (80-.). 269, 676–679.

Hurrell, J.W., Dickson, R.R., 2004. Climate variablity over the North Atlantic, in: Stenseth, N.C., Ottersen, G. (Eds.), Marine Ecosystems and Climate Variation: The North Atlantic. A Comparative Perspective. Oxford University Press, pp. 15–31.

Hurrell, J.W., Kushnir, Y., Ottersen, G., Visbeck, M., 2003. An Overview of the North Atlantic Oscillation, in: The North Atlantic Oscillation: Climatic Significance and Environmental Impact. American Geophysical Union, pp. 1–35. doi:10.1029/134GM01

Irwin, A.J., Oliver, M.J., 2009. Are ocean deserts getting larger? Geophys. Res. Lett. 36, L18609. doi:10.1029/2009GL039883

Jones, C.D., Hughes, J.K., Bellouin, N., Hardiman, S.C., Jones, G.S., Knight, J., Liddicoat, S., O'Connor, F.M., Andres, R.J., Bell, C., Boo, K.-O., Bozzo, A., Butchart, N., Cadule, P., Corbin, K.D., Doutriaux-Boucher, M., Friedlingstein, P., Gornall, J., Gray, L., Halloran, P.R., Hurtt, G., Ingram, W.J., Lamarque, J.-F., Law, R.M., Meinshausen, M., Osprey, S., Palin, E.J., Parsons Chini, L., Raddatz, T., Sanderson, M.G., Sellar, A.A., Schurer, A., Valdes, P., Wood, N., Woodward, S., Yoshioka, M., Zerroukat, M., 2011. The HadGEM2-ES implementation of CMIP5 centennial simulations. Geosci. Model Dev. 4, 543–570. doi:10.5194/gmd-4-543-2011

Kahru, M., Gille, S.T., Murtugudde, R., Strutton, P.G., Manzano-Sarabia, M., Wang, H., Mitchell, B.G., 2010. Global correlations between winds and ocean chlorophyll. J. Geophys. Res. 115, C12040. doi:10.1029/2010JC006500

Kara, A.B., Rochford, P.A., Hurlburt, H.E., 2000. An optimal definition for ocean mixed layer depth. J. Geophys. Res. Ocean. 105, 16803–16821. doi:10.1029/2000JC900072

Large, W., Yeager, S., 2004. Diurnal to decadal global forcing for ocean and sea-ice models: the data sets and flux climatologies. NCAR Tech. Note NCAR/TN-460+STR. doi:10.5065/D6KK98Q6

Laws, E.A., Falkowski, P.G., Smith, W.O., Ducklow, H., McCarthy, J.J., 2000. Temperature effects on export production in the open ocean. Global Biogeochem. Cycles 14, 1231–1246. doi:10.1029/1999GB001229

Leterme, S.C., Edwards, M., Seuront, L., Attrill, M.J., Reid, P.C., John, A.W.G., 2005. Decadal basin-scale changes in diatoms, dinoflagellates, and phytoplankton color across the North Atlantic. Limnol. Oceanogr. 50, 1244–1253. doi:10.4319/lo.2005.50.4.1244

Levitus, S., 1982. Climatological atlas of the world ocea. NOAA Prof. Pap. 13, 173.

Lindeman, R., Merenda, P., Gold, R., 1980. Introduction to Bivariate and Multivariate Analysis. Scott, Foresman, Glenview, IL.

Lomas, M.W., Steinberg, D.K., Dickey, T., Carlson, C.A., Nelson, N.B., Condon, R.H., Bates, N.R., 2010. Increased ocean carbon export in the Sargasso Sea linked to climate variability is countered by its enhanced mesopelagic attenuation. Biogeosciences 7, 57–70. doi:10.5194/bg-7-57-2010

Longhurst, A.R., 2007. Chapter 7 – PROVINCES: THE SECONDARY COMPARTMENTS, in: Ecological Geography of the Sea. pp. 103–114. doi:10.1016/B978-012455521-1/50008-5

Lozier, M.S., Dave, A.C., Palter, J.B., Gerber, L.M., Barber, R.T., 2011. On the relationship between stratification and primary productivity in the North Atlantic. Geophys. Res. Lett. 38, n/a-n/a. doi:10.1029/2011GL049414

Mackas, D.L., 2011. Does blending of chlorophyll data bias temporal trend? Nature 472, E4–E5.

Madec, G., 2008. NEMO reference manual: "NEMO ocean engine." Note du Pole Model. Inst. Pierre Simmon Laplace, Fr. 27.

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

Marshall, J., Schott, F., 1999. Open-ocean convection: Observations, theory, and models. Rev. Geophys. 37, 1–64. doi:10.1029/98RG02739

Martin, A.P., Lucas, M.I., Painter, S.C., Pidcock, R., Prandke, H., Prandke, H., Stinchcombe, M.C., 2010. The supply of nutrients due to vertical turbulent mixing: A study at the Porcupine Abyssal Plain study site in the northeast Atlantic. Deep Sea Res. Part II Top. Stud. Oceanogr. 57, 1293–1302. doi:10.1016/j.dsr2.2010.01.006

Martinez, E., Antoine, D., D'Ortenzio, F., Gentili, B., 2009. Climate-Driven Basin-Scale Decadal Oscillations of Oceanic Phytoplankton. Science (80-.). 326, 1253–1256.

McDougall, T.J., Barker, P.M., 2011. Getting started with TEOS-10 and the Gibbs Seawater (GSW) oceanographic toolbox, SCOR/IAPSO WG.

McGillicuddy, D.J., Robinson, A.R., Siegel, D.A., Jannasch, H.W., Johnson, R., Dickey, T.D., McNeil, J., Michaels, A.F., Knap, A.H., 1998. Influence of mesoscale eddies on new production in the Sargasso Sea. Nature 394, 263–266.

McInnes, K.L., Erwin, T.A., Bathols, J.M., 2011. Global Climate Model projected changes in 10 m wind speed and direction due to anthropogenic climate change. Atmos. Sci. Lett. 12, 325–333. doi:10.1002/asl.341

McQuatters-Gollop, A., Reid, P.C., Edwards, M., Burkill, P.H., Castellani, C., Batten, S., Gieskes, W., Beare, D., Bidigare, R.R., Head, E., Johnson, R., Kahru, M., Koslow, J.A., Pena, A., 2011. Is there a decline in marine phytoplankton? Nature 472, E6–E7.

Menzel, D.W., Ryther, J.H., 1961. Annual variations in primary production of the Sargasso sea off Bermuda. Deep Sea Res. 7, 282–288. doi:http://dx.doi.org/10.1016/0146-6313(61)90046-6

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

Milligan, G.W., Cooper, M.C., 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50, 159–179. doi:10.1007/BF02294245

Moore, C.M., Mills, M.M., Arrigo, K.R., Berman-Frank, I., Bopp, L., Boyd, P.W., Galbraith, E.D., Geider, R.J., Guieu, C., Jaccard, S.L., Jickells, T.D., La Roche, J., Lenton, T.M., Mahowald, N.M., Maranon, E., Marinov, I., Moore, J.K., Nakatsuka, T., Oschlies, A., Saito, M.A., Thingstad, T.F., Tsuda, A., Ulloa, O., 2013. Processes and patterns of oceanic nutrient limitation. Nat. Geosci 6, 701–710.

Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., van Vuuren, D.P., Carter,

T.R., Emori, S., Kainuma, M., Kram, T., Meehl, G.A., Mitchell, J.F.B., Nakicenovic, N., Riahi, K., Smith, S.J., Stouffer, R.J., Thomson, A.M., Weyant, J.P., Wilbanks, T.J., 2010. The next generation of scenarios for climate change research and assessment. Nature 463, 747–756.

Ottersen, G., Planque, B., Belgrano, A., Post, E., Reid, P.C., Stenseth, N.C., 2001. Ecological effects of the North Atlantic Oscillation. Oecologia 128, 1–14. doi:10.1007/s004420100655

Parekh, P., Dutkiewicz, S., Follows, M.J., Ito, T., 2006. Atmospheric carbon dioxide in a less dusty world. Geophys. Res. Lett. 33, L03610. doi:10.1029/2005GL025098

Patara, L., Visbeck, M., Masina, S., Krahmann, G., Vichi, M., 2011. Marine biogeochemical responses to the North Atlantic Oscillation in a coupled climate model. J. Geophys. Res. 116, C07023. doi:10.1029/2010JC006785

Paternoster, R., Brame, R., Mazerolle, P., Piquero, A., 1998. Using the correct statistical test for the equality of regression coefficients. Criminology 36, 859–866. doi:10.1111/j.1745-9125.1998.tb01268.x

Pätsch, J., Kühn, W., Radach, G., Santana Casiano, J.M., Gonzalez Davila, M., Neuer, S., Freudenthal, T., Llinas, O., 2001. Interannual variability of carbon fluxes at the North Atlantic Station ESTOC. Deep Sea Res. Part II Top. Stud. Oceanogr. 49, 253–288. doi:10.1016/S0967-0645(01)00103-5

Polovina, J.J., Howell, E.A., Abecassis, M., 2008. Ocean's least productive waters are expanding. Geophys. Res. Lett. 35, L03618. doi:10.1029/2007GL031745

Post, W.M., Peng, T.-H., Emanuel, W.R., King, A.W., Dale, V.H., DeAngelis, D.L., 1990. The Global Carbon Cycle. Am. Sci. 78, 310–326. doi:10.2307/27828530

Randall, D.A., Wood, R.A., Bony, S., Colman, R., Fichefet, T., Fyve, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R.J., Sumi, A., Taylor, K.E., 2007. Climate Models and Their Evaluation, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. pp. 591–662. doi:10.1016/j.cub.2007.06.045

Redfield, A.C., 1958. The biological control of chemical factors in the environment. Am. Sci. 46, 230A–221.

Reid, P.C., Colebrook, J.M., Matthews, J.B.L., Aiken, J., 2003. The Continuous Plankton Recorder: concepts and history, from Plankton Indicator to undulating recorders. Prog. Oceanogr. 58, 117–173. doi:10.1016/j.pocean.2003.08.002

Reynolds, R.W., Rayner, N.A., Smith, T.M., Stokes, D.C., Wang, W., 2002. An Improved In Situ and Satellite SST Analysis for Climate. J. Clim. 15, 1609–1625. doi:10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., Rafaj, P., 2011. RCP 8.5—A scenario of comparatively high greenhouse gas emissions. Clim.

Change 109, 33–57. doi:10.1007/s10584-011-0149-y

Riser, S.C., Freeland, H.J., Roemmich, D., Wijffels, S., Troisi, A., Belbeoch, M., Gilbert, D., Xu, J., Pouliquen, S., Thresher, A., Le Traon, P.-Y., Maze, G., Klein, B., Ravichandran, M., Grant, F., Poulain, P.-M., Suga, T., Lim, B., Sterl, A., Sutton, P., Mork, K.-A., Velez-Belchi, P.J., Ansorge, I., King, B., Turton, J., Baringer, M., Jayne, S.R., 2016. Fifteen years of ocean observations with the global Argo array. Nat. Clim. Chang. 6, 145–153.

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

Rykaczewski, R.R., Dunne, J.P., 2011. A measured look at ocean chlorophyll trends. Nature 472, E5–E6.

Ryther, J.H., 1969. Photosynthesis and Fish Production in the Sea. Science (80-.). 166, 72–76.

Saba, V.S., Friedrichs, M.A.M., Carr, M.-E., Antoine, D., Armstrong, R.A., Asanuma, I., Aumont, O., Bates, N.R., Behrenfeld, M.J., Bennington, V., Bopp, L., Bruggeman, J., Buitenhuis, E.T., Church, M.J., Ciotti, A.M., Doney, S.C., Dowell, M., Dunne, J., Dutkiewicz, S., Gregg, W., Hoepffner, N., Hyde, K.J.W., Ishizaka, J., Kameda, T., Karl, D.M., Lima, I., Lomas, M.W., Marra, J., McKinley, G.A., Mélin, F., Moore, J.K., Morel, A., O'Reilly, J., Salihoglu, B., Scardi, M., Smyth, T.J., Tang, S., Tjiputra, J., Uitz, J., Vichi, M., Waters, K., Westberry, T.K., Yool, A., 2010. Challenges of modeling depth-integrated marine primary productivity over multiple decades: A case study at BATS and HOT. Global Biogeochem. Cycles 24, n/a-n/a. doi:10.1029/2009GB003655

Sarmiento, J.L., Bender, M., 1994. Carbon biogeochemistry and climate change. Photosynth. Res. 39, 209–234. doi:10.1007/BF00014585

Sarmiento, J.L., Slater, R., Barber, R., Bopp, L., Doney, S.C., Hirst, A.C., Kleypas, J., Matear, R., Mikolajewicz, U., Monfray, P., Soldatov, V., Spall, S.A., Stouffer, R., 2004. Response of ocean ecosystems to climate warming. Global Biogeochem. Cycles 18. doi:10.1029/2003GB002134

Siegel, D.A., Behrenfeld, M.J., Maritorena, S., McClain, C.R., Antoine, D., Bailey, S.W., Bontempi, P.S., Boss, E.S., Dierssen, H.M., Doney, S.C., Eplee, R.E., Evans, R.H., Feldman, G.C., Fields, E., Franz, B.A., Kuring, N.A., Mengelt, C., Nelson, N.B., Patt, F.S., Robinson, W.D., Sarmiento, J.L., Swan, C.M., Werdell, P.J., Westberry, T.K., Wilding, J.G., Yoder, J.A., 2013. Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission. Remote Sens. Environ. 135, 77–91. doi:10.1016/j.rse.2013.03.025

Siegel, D.A., Doney, S.C., Yoder, J.A., 2002. The North Atlantic Spring Phytoplankton Bloom and Sverdrup's Critical Depth Hypothesis. Science (80-.). 296, 730 LP-733.

Siegenthaler, U., Sarmiento, J.L., 1993. Atmospheric carbon dioxide and the ocean. Nature 365, 119–125. doi:10.1038/365119a0

Steinacher, M., Joos, F., Frölicher, T.L., Bopp, L., Cadule, P., Cocco, V., Doney, S.C., Gehlen, M., Lindsay, K., Moore, J.K., Schneider, B., Segschneider, J., 2010. Projected 21st century decrease in marine productivity: a multi-model analysis. Biogeosciences 7, 979–1005. doi:10.5194/bg-7-979-2010

Steinhoff, T., Friedrich, T., Hartman, S.E., Oschlies, A., Wallace, D.W.R., Körtzinger, A., 2010. Estimating mixed layer nitrate in the North Atlantic Ocean. Biogeosciences 7, 795–807. doi:10.5194/bg-7-795-2010

Stenseth, N.C., Ottersen, G., 2004. Marine Ecosystems and Climate Variation: The North Atlantic. A Comparative Perspective. Oxford University Press, Oxford.

Sverdrup, H.U., 1953. On Conditions for the Vernal Blooming of Phytoplankton. J. du Cons. 18, 287–295. doi:10.1093/icesjms/18.3.287

Tu, Y.-K., Kellett, M., Clerehugh, V., Gilthorpe, M.S., 2005. Problems of correlations between explanatory variables in multiple regression analyses in the dental literature. Br. Dent. J. 199, 457–461. doi:10.1038/sj.bdj.4812743

Uppala, S.M., KÅllberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V.D.C., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., Berg, L. Van De, Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A.E.M., Jenne, R., Mcnally, A.P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., Woollen, J., 2005. The ERA-40 re-analysis. Q. J. R. Meteorol. Soc. 131, 2961–3012. doi:10.1256/qj.04.176

Vantrepotte, V., Mélin, F., 2011. Inter-annual variations in the SeaWiFS global chlorophyll a concentration (1997–2007). Deep Sea Res. Part I Oceanogr. Res. Pap. 58, 429–441. doi:10.1016/j.dsr.2011.02.003

Visbeck, M., Chassignet, E.P., Curry, R.G., Delworth, T.L., Dickson, R.R., Krahmann, G., 2003. The Ocean's Response to North Atlantic Oscillation Variability, in: The North Atlantic Oscillation: Climatic Significance and Environmental Impact. American Geophysical Union, pp. 113–145. doi:10.1029/134GM06

Volk, T., Hoffert, M.I., 1985. Ocean Carbon Pumps: Analysis of Relative Strengths and Efficiencies in Ocean-Driven Atmospheric CO2 Changes, in: The Carbon Cycle and Atmospheric CO2: Natural Variations Archean to Present. American Geophysical Union, Washington, DC, pp. 99–110. doi:10.1029/GM032p0099

Wanner, H., Brönnimann, S., Casty, C., Gyalistras, D., Luterbacher, J., Schmutz, C., Stephenson, D.B., Xoplaki, E., 2001. North Atlantic Oscillation – Concepts And Studies. Surv. Geophys. 22, 321–381. doi:10.1023/A:1014217317898

Weaver, A.J., Sedláček, J., Eby, M., Alexander, K., Crespin, E., Fichefet, T., Philippon-Berthier,

G., Joos, F., Kawamiya, M., Matsumoto, K., Steinacher, M., Tachiiri, K., Tokos, K., Yoshimori, M., Zickfeld, K., 2012. Stability of the Atlantic meridional overturning circulation: A model intercomparison. Geophys. Res. Lett. 39, n/a-n/a. doi:10.1029/2012GL053763

Wernand, M.R., van der Woerd, H.J., Gieskes, W.W.C., 2013. Trends in Ocean Colour and Chlorophyll Concentration from 1889 to 2000, Worldwide. PLoS One 8, e63766.

Williams, R.G., Follows, M.J., 1998. The Ekman transfer of nutrients and maintenance of new production over the North Atlantic. Deep Sea Res. Part I Oceanogr. Res. Pap. 45, 461–489. doi:http://dx.doi.org/10.1016/S0967-0637(97)00094-0

Wilson, C., Adamec, D., 2002. A global view of bio-physical coupling from SeaWiFS and TOPEX satellite data, 1997–2001. Geophys. Res. Lett. 29, 94–98. doi:10.1029/2001GL014063

Yoder, J.A., Kennelly, M.A., 2003. Seasonal and ENSO variability in global ocean phytoplankton chlorophyll derived from 4 years of SeaWiFS measurements. Global Biogeochem. Cycles 17, n/a-n/a. doi:10.1029/2002GB001942

Yool, A., Popova, E.E., Anderson, T.R., 2011. Medusa-1.0: a new intermediate complexity plankton ecosystem model for the global domain. Geosci. Model Dev. 4, 381–417. doi:10.5194/gmd-4-381-2011

Zhai, L., Platt, T., Tang, C., Sathyendranath, S., Walne, A., 2013. The response of phytoplankton to climate variability associated with the North Atlantic Oscillation. Deep Sea Res. Part II Top. Stud. Oceanogr. 93, 159–168. doi:10.1016/j.dsr2.2013.04.009

Zhang, H.-M., Bates, J.J., Reynolds, R.W., 2006. Assessment of composite global sampling: Sea surface wind speed. Geophys. Res. Lett. 33, L17714. doi:10.1029/2006GL027086