UNIVERSITY OF Southampton

University of Southampton Research Repository
ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

http://eprints.soton.ac.uk

UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

School of Ocean and Earth Sciences

The Modelling of Mixotrophy in the Oligotrophic Atlantic

by

Siân Joscelyn Herrington

Thesis for the degree of Doctor of Philosophy

December 2012

Abstract

THE MODELLING OF MIXOTROPHY IN THE OLIGOTROPHIC ATLANTIC

By Siân Joscelyn Herrington

In the oligotrophic Atlantic Ocean small algae are the dominant fixers of inorganic carbon. *In situ* experiments have shown that a large proportion of these algae are mixotrophs - eating bacteria (bacterivory) as well as obtaining energy from sunlight. Bacterivory performed by algae has implications for our understanding of the role of ultraplankton (< 5 µm) in biogeochemical cycling. The motivation of this thesis is to explore how mixotrophy may be modelled in the subtropical Atlantic using a data driven approach.

An ecosystem model incorporating ultraplankton mixotrophy was developed, constructed and parameterised using *in situ* data, initially through network analysis and later using a µ-Genetic Algorithm technique. The model highlights the key role of mixotrophy in the cycling of nutrients, in a region where fast nutrient turnover is important for the functioning of the ecosystem. In addition, the model reveals that bacterivory is the predominant route of nutrient acquisition for these mixotrophs and suggests that mixotrophy in this low nutrient region is an adaptive rather than a survival mechanism.

This thesis also addresses wider questions related to model structure and assumptions. The need for an explicit dissolved organic phosphate variable in an ecosystem model for the oligotrophic Atlantic is questioned through *in situ* radio-nucleotide bioassay techniques. Additionally, ultraplankton spatial variability is statistically assessed and used to demonstrate that a zero-dimensional model is not necessarily applicable to an entire region, despite the ultraplankton community within that region being statistically similar according to multivariate analyses. Furthermore a comparison of *in situ* to remotely sensed data shows that ocean colour is limited in its ability to detect ultraphytoplankton, making the use of such data to calibrate and assess future models difficult.

This thesis therefore not only contributes to our ability to model the oligotrophic Atlantic but more broadly to our understanding of the role of mixotrophs within it.

Contents

List of Figures

List of Tables

DECLARATION OF AUTHORSHIP

I, **SIÂN JOSCELYN HERRINGTON**

declare that the thesis entitled

THE MODELLING OF MIXOTROPHY IN THE OLIGOTROPHIC ATLANTIC

and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- 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;
- where I have consulted the published work of others, this is always clearly attributed;
- 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;
- I have acknowledged all main sources of help;
- 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;
- none of this work has been published before submission

Signed: ……

Date:……

This PhD thesis by

SIÂN JOSCELYN HERRINGTON

has been produced under the supervision of the following persons:

Supervisors:

Dr. Adrian Martin

Prof. Meric Srokosz

Prof. Mikhail Zubkov

Chairs of the Supervisory Panel:

Prof. Harry Bryden

Dr. Thomas Anderson

Thesis for the degree of Doctor of Philosophy

December 2012

Acknowledgements

I would like to whole heartedly thank my supervisors Adrian Martin, Meric Srokosz and Mike Zubkov for their advice and support throughout this PhD. I really appreciate the many hours spent in discussion and the guidance they provided. Thanks also to my panel chairs Tom Anderson and latterly Harry Bryden, for his well-timed and priceless encouragement. Specifically within the flow cytometry team at NOCS I wish to thank Manuela Hartmann and Ross Holland, for their seemingly infinite knowledge which they were always happy and enthusiastic to share with me. Thank you also to Steph Henson for her valuable remote sensing guidance, Carolina Grob for providing cellular phosphate content data and Maureen Pagnani for her help in extracting Argo data. Particular thanks also go to the Captain, Officers, crew, technical staff and fellow scientists aboard the *RRS* James Cook during JC039 (AMT19). My appreciation also to the Challenger Society for Marine Science, whose travel award help enabled me to present at ASLO Ocean Sciences 2012, in Salt Lake City.

I would also like to extend an extra special thanks to my family and friends. Particularly to my mum, for her continual love and support over the last 29 years (and for proof reading this thesis!). I would also like to thank my lovely house mates–Alice, Anna and Katie (and more recently Chris) who were always on hand with a glass of wine. Also thank you to the many people at NOCS, specifically the tea drinking team - Adam, Leigh, Helen and Charlie and my 344/27 office mates (Chris, Sam, Simon, Ros and Alex). Finally I would like to sincerely thank Rich, without whose love and encouragement I would not have made it through the last year. Hopefully now he can finally have his girlfriend back!

This PhD was funded by a Natural Environmental Research Council Grant NE/E01612X/1 .

For my Dad

Abbreviations

AMT	Atlantic Meridional Transect
AMP	Adenosine monophosphate
APA	Alkaline Phosphatase Activities
ATP	Adenosine triphosphate
B	Bacteria
Chl *a*	Chlorophyll *a*
C-N-P	Carbon-Nitrogen-Phosphorus
CTD	Conductivity Temperature and Depth
Crypt	Cryptophytes
CZCS	Coastal Zone Colour Scanner
DIP	Dissolved Inorganic Phosphorus/Phosphate
DOP	Dissolved Organic Phosphorus/Phosphate
DPM	Disintegrations Per Minute
G	Grazers
µGA	micro-Genetic Algorithm
HB	Heterotrophic Bacteria
HNAhs	High-Nucleic Acid bacteria with a high 90° light scatter.
HNAls	High-Nucleic Acid bacteria with a low 90° light scatter
HTC	High Throughput Computing
LNA	Low-Nucleic Acid Bacteria
M	Mixotrophs
MDS	Multi-Dimensional Scaling
MicroPP	Microphytoplankton
ML	Mixed Layer
MLD	Mixed Layer Depth

MODIS	Moderate Resolution Imaging Spectroradiometer
NAG	North Atlantic subtropical Gyre
NADR	North Atlantic Drift, Longhurst Province
NAST	North Atlantic Subtropical Gyre East, Longhurst Province
NATR	North Atlantic Tropical Gyre, Longhurst Province
NPZ	Nutrient Phytoplankton Zooplankton
NPZD	Nutrient Phytoplankton Zooplankton Detritus
PAP	Porcupine Abyssal Plain
PB	Phototrophic Bacteria
PFA	Paraformaldehyde
PFT	Phytoplankton Functional Type
PicoEuk	Picoeukaryotes
PP	Primary productivity
Pro	*Prochlorococcus* spp.
SAG	South Atlantic subtropical Gyre
SATL	South Atlantic Gyre, Longhurst Province
SeaWiFS	Sea-viewing Wide Field-of-View Sensor
Syn	*Synechococcus* spp.
Ttl	Total Heterotrophic Bacteria
UMP	Uridine monophosphate
WTRA	Western Tropical Atlantic, Longhurst Province

1. Introduction

1.1. General Introduction

The nutrient poor (oligotrophic) regions of the oceans form the Earth's largest ecosystem, covering approximately 40 % of the Earth's surface (Karl 1999, Polovina et al. 2008) and accounting for roughly half of total marine primary production, in terms of the production of biomass from the uptake of CO_2 (Sarmiento and Gruber 2006). Consequently the oligotrophic regions profoundly affect global biogeochemical cycling. Planktonic plastidic protists (algae) are major fixers of inorganic carbon in the oceans (Li 1994). It is conventionally thought that in oligotrophic areas the dominant phototrophic ultraplankton (< 5 μm) are severely nutrient limited (Tyrrell 1999, Wu et al. 2000, Mather et al. 2008). Ultraplankton, however, do live and reproduce in these waters. 'How?' is a major question in biological oceanography (Mahaffey et al. 2004).

Measurements from the Sargasso Sea suggest that dissolved organic phosphate may be subsidising the limited inorganic phosphate (Casey et al. 2009, Lomas et al. 2010, Michelou et al. 2011), yet measurements from the adjacent central north Atlantic contradict this (Zubkov et al. 2007). Bacterivory has been previously documented to play a role in algae survival not only in the subtropics (Zubkov and Tarran 2008, Hartmann et al. 2012), but in other regions of the oceans (see Table 1.1, discussed in Section 1.3, and references therein). Mixotrophy in this thesis, is defined as the ability of planktonic organisms to 'mix' two modes of nutrition – gaining energy from sunlight and preying on bacterioplankton to acquire inorganic and some essential organic nutrients (this is

explained further in Section 1.3). Bacterivory by algae has significant implications for the current understanding of the role of ultraplankton in biogeochemical cycling.

The motivation of this thesis is to undertake research into mixotrophy in the open ocean surface waters of the oligotrophic Atlantic's ecosystem. More specifically the thesis will focus on how models for the region incorporating mixotrophy can be built, this will partly involve the construction and parameterisation of simple ultraplankton trophodynamical models, using *in situ* ultraplankton observations to construct, calibrate and parameterise them. However it will also address wider related questions. Necessary model structure will be addressed, such as the need to include an organic nutrient variable, and the validity of assumptions of steady state and spatial uniformity, that may be thought applicable to these 'Ocean deserts'. It should be noted that although the focus of this thesis is on the subtropical Atlantic, the work will be relevant to other oligotrophic regions.

This Chapter provides an overview of mixotrophy, the organisms of interest and the region within which they are being studied – the oligotrophic subtropical Atlantic. This is followed by a review of previous mixotrophy models. The objectives of this thesis are then outlined.

1.2. The Oligotrophic Atlantic Ocean

The oligotrophic subtropical and tropical Atlantic ocean covers in excess of 10 million square kilometres (Polovina et al. 2008). The subtropics are usually defined as being between 40° and 10° north and south of the equator, and are dominated by the vast ocean circulations known as the subtropical gyres. These are driven by wind circulation patterns (Munk 1950). These large systems of ocean currents rotate clockwise in the northern hemisphere and anti-clockwise in the southern hemisphere. Each of the subtropical oligotrophic oceanic gyres are asymmetric and bounded to the west by a pole-ward current, the western boundary current (e.g. in the north Atlantic gyre, the Gulf Steam). To the east the gyres are bounded by a current flowing towards the equator. These eastern currents are typically slower moving, wider, shallower and diffuse (Stommel 1947, Stommel 1948).

The biological distribution of plankton within the subtropical gyres is influenced by the physical environment. Water is upwelled at the edges of a gyre (equatorial or coastal margins), introducing nutrient rich water from the deep ocean to the surface for utilisation by plankton in neighbouring waters. The nutrient concentration decreases towards a minimum in the centre of the gyre (which can be identified by maximum sea surface height), where convergence resulting from Ekman transport driven by the trade winds causes the water to be down-welled, suppressing upward transport of nutrients (Colling 2002, Sarmiento and Gruber 2006). The subtropical Atlantic regions display a relatively shallow stratified mixed layer, with only small changes in the mixed layer depth between opposing seasons (Monterey and Levitus 1997, Henson et al. 2009). This lack of seasonality is also observed in remotely sensed chlorophyll concentrations, which are relatively constant all year round (Lutz et al. 2007, Cole et al. 2012).

1.2.1. Oligotrophic Atlantic Ultraplankton

In the oligotrophic ocean the majority of the planktonic organisms are small (< 5 µm) and are termed ultraplankton. *Prochlorococcus* spp. (*Pro*) cyanobacteria and the SAR11 group of heterotrophic α-proteobacteria are the most numerous microbes contained within this fraction (Chisholm et al. 1988, Morris et al. 2002). In terms of phototrophic ultraplankton, *Pro* are the most abundant followed by *Synechococcus* (*Syn*) spp. (Partensky et al. 1999a) and then single-celled picoeukaryotes (protists), which may include among others Chrysophyceae, Prasinophyceae, Prymnesiophyceae species (Lepère et al. 2009, Liu et al. 2009, Jardillier et al. 2010). Although picoeukaryotic algae are not numerically dominant, they do dominate on the basis of biomass and primary production (Li 1995). These small picoeukaryotic algae are major fixers of inorganic carbon in the oceans (Li 1994, Li 1995).

The ultraplankton within the north Atlantic subtropical gyre are understood to be primarily limited by the availability of phosphorus (Tyrrell 1999, Mather et al. 2008), with field measurements estimating the concentration as < 2 nmol l^{-1} (Wu et al. 2000, Zubkov et al. 2007). Phosphorus (P) is an essential element for all living cells, forming the structural framework of DNA and RNA and is a component of ATP and lipids in cell membranes. As such it is a critical driver of phytoplankton growth and ecosystem functioning in the ocean.

The P limitation in the NAG is thought to be as a result of iron enhanced nitrogen fixation, as the region is subject to iron deposition from Saharan dust (Duce et al. 1991, Wu et al. 2000). Bacterioplankton dominate P uptake in the gyre and outcompete protists for it due to their higher surface area to volume ratio (Zubkov et al. 2007, Hartmann et al. 2011). Despite the low P uptake by protists, that cannot satisfy their growth requirements (Hartmann et al. 2011), they are major contributors to carbon fixation (Li 1994, Jardillier et al. 2010). Studies in the Sargasso Sea, adjacent to the north Atlantic (and with differing physical forcing, see Chapter 3 for further explanation of differences), have indicated that carbon fixation may be being maintained due to planktonic organisms utilising dissolved organic P (DOP) as an alternative dissolved inorganic P (DIP) source (Casey et al. 2009, Michelou et al. 2011). However, measurements made in the north Atlantic gyre, have shown DOP to be only a minor contribution to total P uptake (Zubkov et al. 2007); Chapter 3 of this thesis investigates this further.

Zubkov and Tarran (2008) challenged the assumption that algae within the oligotrophic north Atlantic subtropical gyre are completely reliant on light and inorganic or organic nutrients. They demonstrated that a large proportion, 37 - 70 % (a comparable percentage was also reported in the temperate North Atlantic Ocean, 40 - 95 %) of bacterial grazing (bacterivory) in this region was undertaken by these small algae. Recently the significance and ubiquity of mixotrophy was shown across the whole of the subtropical and tropical Atlantic (40°N to 40°S), with algae accounting for 60 - 77 % of bacterivory (Hartmann et al. 2012, also see Table 1.1). These studies suggest that algae are using mixotrophy as a mechanism to survive in nutrient depleted waters (see Section 1.3 on mixotrophy). This observation is significant because in these waters ultraplankton account for ~ 80 % of phytoplankton biomass and chlorophyll *a* and ~ 70 % of primary production (Li and Harrison 2001). Thus, these findings have important implications for the current understanding of biogeochemical cycling and the role of ultraplankton and picoeukaryotes in particular, as the amount of primary production and export occurring in the Atlantic oligotrophic gyres may be affected by algae 'eating' in addition to photosynthesising. Firstly, algae may be able to survive at lower nutrient conditions than previously thought, as they can supplement nutrient uptake and reduce competition through predation of competitors. Secondly, the fraction of nutrients passed on to higher

trophic levels may be less than previously thought as an extra trophic link is present within the system. This decreases the efficiency of energy transfer and subsequent primary production that is passed on to higher levels, which in turn reduces the amount of export out of the surface waters.

1.2.1.1. Remote Sensing of Ultraplankton

Oceanographers were limited to *in situ* sampling for centuries, unable to study vast areas and the plankton within them. Remotely sensed observations from satellites changed this by allowing routine regional and global monitoring. From space, the differing population densities of phytoplankton reveal themselves as many blues and greens which satellites are able to measure as 'ocean colour' (seen in Figure 1.1). The principles of ocean colour algorithms have changed little since the first space borne ocean colour sensor CZCS (Coastal Zone Colour Scanner) was launched in 1978 (Feldman et al. 1989, Yoder et al. 1993). Atmospheric correction is performed (90 % of the original signal), then the blue-to-green ratio of water-leaving radiances allows the estimation of 'chlorophyll *a* concentration' - the sum of chlorophyll *a* and pheophytin *a*, which is used as a proxy for phytoplankton (algae and cyanobacteria) biomass (Martin 2004).

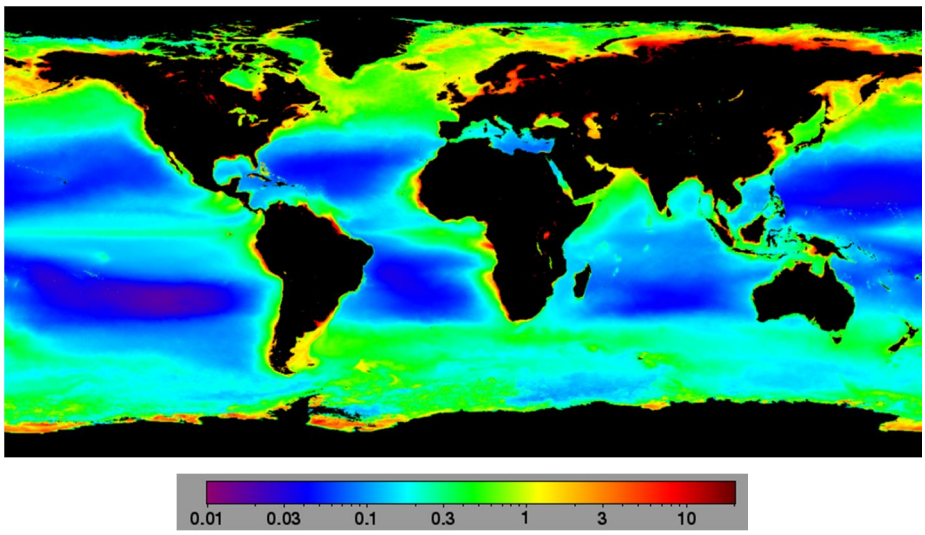


Figure 1.1. SeaWiFS mission average (September 1997 – November 2010) chlorophyll *a* concentration. Blue indicates low chlorophyll and yellow and reds the highest regions of chlorophyll concentration. Note units of chlorophyll *a* are mg m^{-3}. Downloaded from http://oceancolor.gsfc.nasa.gov.

If the oceans can be viewed remotely, at high spatial and temporal resolutions, what is the advantage of using *in situ* measurements? First, remote sensing cannot measure rates, such as grazing and bacterivory. Second, the accuracy of remotely sensed ultraplankton has been brought into question by a study undertaken in the Mozambique Channel, which did not find a strong correlation between satellite derived chlorophyll estimates and *in situ* abundance measurements of ultraplankton taxa (Zubkov and Quartly 2003). It is hitherto unknown if ultraplankton are also poorly represented by ocean colour within the oligotrophic Atlantic ocean. This will be addressed in Chapter 6 of this thesis.

1.2.2. The Atlantic Meridional Transect Programme

The Atlantic Meridional Transect (AMT) is a multidisciplinary oceanographic research programme, established in 1995, which provides a unique annual time series of surface ocean measurements along a transect throughout the mid-Atlantic Ocean. It has enabled measurements to be made regularly along an entire transect of the Atlantic ocean (~ 13,500 km). A primary aim of the AMT project since 2002 has been to "provide data for use in the development of models" with the specific objective to determine "how the structure, functional properties and trophic status of the major planktonic ecosystems vary in space and time" (Robinson et al. 2006, pp.1489). Research presented in this thesis contributes to the fulfilment of these AMT project objectives, as it utilises ultraplankton data collected on four AMT cruises to help construct, develop and constrain a model inclusive of mixotrophy. Details of the data collected from these cruises are outlined in Appendix Tables (Tables A.1 and A.2).

1.3. Mixotrophic Protists

Mixotrophy is the 'mixing' of plant and animal modes of nutrition, akin to the terrestrial Venus fly trap which obtains energy within terrestrial biomes from sunlight and the catching of insects. A mixotroph here, is more specifically a protist with the capability to utilise autotrophic and heterotrophic nutrition to varying degrees. In the context of this

thesis, autotrophy refers to phototrophy as using electromagnetic energy (light), inorganic nutrients and CO_2 to produce energy and organic matter. Heterotrophy refers to phagotrophy, where organic compounds are absorbed into food vacuoles for digestion, metabolism, growth and reproduction (Sanders 1991, Jones 1994, Raven 1997).

The ability of some protists to exist straddling two trophic levels has been observed throughout ocean and freshwater ecosystems (Sanders 1991) and appears to be a successful evolutionary strategy. Numerous species from diverse taxa (e.g. Cryptophyceae, Dinophyceae and Prymnesiophyceae) have been found to be capable of mixotrophy (see references in Sanders and Porter, 1988). As previously stated, mixotrophs have recently been found to be notably prevalent in oligotrophic regions, where it is thought bacterivory by ultraplankton gives a competitive advantage (Zubkov and Tarran 2008, Hartmann et al. 2012).

A simple technical definition of mixotrophy is difficult, as mixotrophic protists vary widely in their photosynthetic and ingestion capabilities. Existing at different points along the spectrum of nutritional strategies (Jones 1994), three broad types have been identified (see schematic in Figure 1.2): Type I, an ideal or obligate mixotroph, which combines both autotrophic and heterotrophic modes of nutrition equally; Type II, phagotrophic 'algae' that are primarily phototrophic; and Type III, phototrophic 'protozoa' that are primarily phagotrophic (Stoecker 1998). The oligotrophic Atlantic mixotrophs are primarily 'algae' that are phagotrophic, believed to subsidise their growth and reproduction by bacterivory when nutrients are limiting (Zubkov and Tarran 2008). Therefore, under Stoecker's (1998) definition they are Type IIa mixotrophs (phagotrophic when nutrients are limiting, see Figure 1.2). Throughout this thesis, unless otherwise indicated, references to mixotrophs refer to small (≤ 5 µm) Type IIa mixotrophs (phagotrophic or bacterivorous 'algae'). For a more detailed overview of mixotrophic nutritional types the reader is directed to Jones (1994) and Stoecker (1998).

The housing of organelles for both autotrophy and heterotrophy, however, is energetically expensive, which is why evolutionary forces may not have led mixotrophy to be the dominant nutritional mechanism in the photic zone (Raven 1997). Mixotrophs' nutritional flexibility, however, suggests they have a competitive advantage over obligate (pure)

autotrophs and heterotrophs when nutrients are low and particulate food and prey is limited in oligotrophic regions (Nygaard and Tobiesen 1993, Raven 1997, Stibor and Sommer 2003), which may explain their dominance in low nutrient regions (see Table 1.1 and references therein).

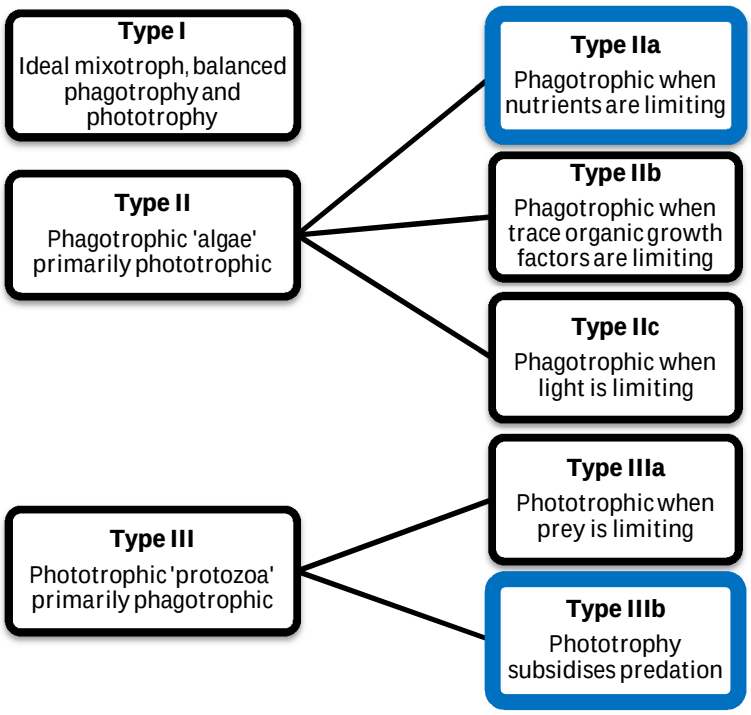


Figure 1.2. The physiological types of mixotrophy possible among planktonic protists, according to Stoecker (1998). Type IIa and IIIb, highlighted in blue, are postulated as being the most common.

1.3.1. Observations of Mixotrophy

There have been a number of studies of mixotrophy in the laboratory, *in vitro* (e.g. Rothaupt 1996a, 1996b, Stibor and Sommer 2003) and in the field, *in situ* (e.g. Bird and Kalff 1986, Arenovski et al. 1995). Despite this, the ubiquitousness of mixotrophy in the oligotrophic Atlantic has only recently been recognised (Zubkov and Tarran 2008, Hartmann et al. 2012).

Table 1.1. Percentage of total protist bacterivory by mixotrophic flagellates in different trophic marine environments, according to previously published research.

System	Depth	Trophic status	Season	Tracer	Suggested dominant mixotrophic taxa	% bacterivory by mixotrophs	References
Southwest coast of New Zealand	10m	Oligo	Winter	FLM	*	32-40 %	Hall et al. (1993)
Aegean Sea, East Mediterranean Sea	1-100m	Oligo	Spring	FLM	Haptophyceae (2-4µm)	5 %	Christaki et al. (1999)
Aegean Sea, East Mediterranean Sea	1-100m	Oligo	Summer	FLM	Phytoflagellates (8-10µm)	5 %	Christaki et al. (1999)
New Zealand, Subtropical Convergence	DCM	Oligo-meso	Summer	FLBc	Haptophyceae; Dinophyceae	55 %	Safi and Hall (1999)
New Zealand, Subtropical Convergence	DCM	Meso	Summer	FLM	Haptophyceae; Dinophyceae	40 %	Safi and Hall (1999)
Bay of Aarhus, Denmark	Surface	Meso	Spring	FLBn	Haptophyceae; Dictyochophyceae	86 %	Havskum and Riemann (1996)
Bay of Aarhus, Denmark	Below the pycnocline	Meso	Spring	FLBn	Haptophyceae; Dictyochophyceae	19 %	Havskum and Riemann (1996)
Boston Harbour	Surface	Meso?	Autumn	FLBc	Euglenophyceae; Chrysophyceae; Cryptophyceae	38 %	Epstein and Shiaris (1992)
Coast of Norway	0-4m	Meso	Bloom of *prymnesium parvum*	FLB*	Haptophyceae	0-60 %	Nygaard and Tobiesen (1993)
Norwegian fjord	Surface	Meso	Summer	FLBc	Dictyochophyceae; Haptophyceae; Dinophyceae	0 %	Havskum and Hansen (1997)
Black Sea	Mixing layer	Oligo to eutro	Summer	FLBn	Dinophyceae; Ciliate	14 %	Bouvier et al. (1996)
Blanes Bay, NW Mediterranean Sea	Surface	Oligo	Annual	FLBc	Haptophyceae; Dinophyceae; Cryptophyceae;	35-65 %	Unrein et al. (2007)
Ross Sea, Antarctica	0-49m (Surface-DCM)	Eutro - Oligo?	Spring	FLBn	*	8-42 %	Moorthi et al. (2009)
Eastern Equatorial Pacific – HNLC area	18 m	Oligo	Summer /Autumn	FLBn	*	54 %	Stukel et al. (2011)
Subtropical and Tropical Atlantic (40°N to 40°S)	20 m	Oligo	Autumn (N) /Spring (S)	RLB, RLA	*	60-77%	Hartmann et al. (2012)
Temperate North Atlantic Ocean (58.2–60°N)	7 or 47m (euphotic layer)	Eutro	Summer	RLB	Haptophyceae; Prasinophytes	40–95%	Zubkov and Tarran (2008)
Tropical North-East Atlantic Ocean (12.6–22°N)	20m	Oligo	Winter	RLB	Haptophyceae; Prasinophytes	37–70%	Zubkov and Tarran (2008)
Sargasso Sea	1-5m	Oligo	Summer	FLBn; FLM	*	74-80%	Arenovski et al. (1995)
Sargasso Sea	15m	Oligo	Summer	FLM	Dinophyceae	5-18% PNAN^ 4-15% Phago^	Sanders et al. (2000)
Sargasso Sea	10-20m	Oligo	Winter	FLB	Dinophyceae	12-18% PNAN^ 12-28 % Phago^	Sanders et al. (2000)
Georges Bank, NW Atlantic Ocean	0-2m	Eutro	Summer	FLB	Dinophyceae	7-30% PNAN^ 10-23% Phago^	Sanders et al. (2000)
Georges Bank, NW Atlantic Ocean	5-10m	Eutro	Winter	FLB	Dinophyceae	<2-38% PNAN^ 7-39% Phago^	Sanders et al. (2000)

* not specified; FLM, fluorescently labelled mini-cells; FLBc, FLB with a bacterial culture; FLBn, FLB prepared with natural bacteria; RLB, radiolabeled bacteria; RLA, radiolabeled bacteria algae; FLA, fluorescently labelled <3 µm algae; FLC, fluorescently labelled cyanobacteria; PNAN^, phototrophic nanoplankton; Phago^, total heterotrophic nanoplankton plus mixotrophic nanoplankton; Oligo, oligotrophic; Meso, mesotrophic; Eutro, eutrophic; DCM, deep chlorophyll maximum

The recent nature of this discovery, coupled with cost, logistical and methodological limitations of observations means there is a very limited dataset to study mixotrophy, that may in turn be utilised for parameterisation and calibration of models. *In situ* data are preferable to using laboratory based measurements, as the latter does not take into account 'real-world' complexity and variability (e.g. physical forcing such as the mixing between surface and deeper waters, or biological influences, such as predator-prey dynamics).

Table 1.1 provides a summary of studies of bacterivory by mixotrophs. This Table shows that mixotrophy has been found in a number of different environments, from the seasonally oligotrophic Antarctic (Moorthi et al. 2009) to the mesotrophic coast of Norway (Nygaard and Tobiesen 1993). As such there does not appear to be a correlation with nutrient availability. Table 1.1 demonstrates that there is not a consistent fraction of bacterivory by mixotrophs, with measurements varying from 5 % up to 80 % in oligotrophic regions (Arenovski et al. 1995, Christaki et al. 1999) and from < 2 % to 95 % in eutrophic regions (Sanders et al. 2000, Zubkov and Tarran 2008). It is nevertheless clear from these studies that mixotrophic strategies appear to be ubiquitous.

1.3.2. Modelling Mixotrophs

Plankton dynamics can also be explored *in silico*, using computer modelling to simulate and study marine ecosystems. Models have been applied as a quantitative and a descriptive tool for over 70 years, since the seminal studies of Fleming (1939) and Riley (1946), who used Lotka-Volterra predator-prey models to interpret temporal changes in phytoplankton abundances. Simple NPZ-type (Nutrient-Phytoplankton–Zooplankton, described below) models have formed the basis of the planktonic biological modelling for decades (Gentleman 2002) and have become common place since the seminal paper by Fasham et al. (1990). These studies are the ancestors of many large-scale marine ecosystem and biogeochemistry models (e.g. Sarmiento et al. 1993, Oschlies et al. 2000, Schartau and Oschlies 2003a). Major advances have been made possible by increasingly efficient computation, and models can now include a vast array of processes. For a detailed overview of planktonic modelling see for example reviews by Gentleman (2002) and Follows and Dutkiewicz (2011).

Direct observations of mixotrophy over time-scales longer than a few hours are limited. Therefore *in silico* studies are a useful tool to investigate the impact of mixotrophy over

longer temporal scales, such as a complete seasonal cycle. Model parameterisation is, however, limited by the available data. To date there have been no direct measurements of mixotrophs used within a model incorporating mixotrophy.

NPZ-type models incorporate a simple set of dynamics to describe oceanic plankton ecosystems (Franks 2002). NPZ-type models have few parameters and a limited number of state variables (e.g. a single nutrient, phytoplankton and zooplankton). As such they can be relatively easily parameterised with and tested against data compared with more complicated models. Therefore despite being simple they still allow for a wide range of model ecosystems dynamics to be simulated (Franks 2002). Consequently several theoretical mixotroph models have been based on this structure (simple, with few variables and minimal parameters). They follow the philosophy that it is desirable to use a simple approach, when an ecosystems structure and internal processes are not fully understood (Anderson 2005).

Plankton functional types (or PFT's), define ecosystem model variables based on their ecological traits (Hood et al. 2006a), for example by nutritional strategy, such as mixotrophy (e.g. Thingstad et al. 1996, Baretta-Bekker et al. 1998, Jost et al. 2004). Such simple mixotroph models have, despite being mainly theoretical (e.g. Thingstad et al. 1996, Hammer and Pitchford 2005) or based upon broad literature values (e.g. Taylor and Joint 1990, Stickney et al. 2000, Jost et al. 2004, Crane and Grover 2010), been able to point to the potentially important role of mixotrophic organisms as stabilisers of system dynamics, as their grazing on autotrophs reduces the competition for nutrients (Jost et al. 2004, Hammer and Pitchford 2005). In addition, such models have shown that mixotrophs are particularly important within nutrient limited systems (Baretta-Bekker et al. 1998, Floder et al. 2006, Crane and Grover 2010). However, the consequences of mixotrophy in terms of their influence upon primary productivity remain uncertain. Some models have concluded that mixotrophs have a positive influence on productivity (Baretta-Bekker et al. 1998, Hammer and Pitchford 2005), because primary production is supported by nutrients obtained from grazing. Other models report a negative influence on total primary productivity (Stickney et al. 2000), although the mixotrophs may maintain the total photosynthetic rate through the direct recycling of nutrients (this however was not considered 'primary' production as it was derived from recycling).

Although the majority of previous models incorporating mixotrophy are of the simple NPZ-type, with a preponderance for studying a single common nutrient (e.g. Taylor and

Joint 1990, Baretta-Bekker et al. 1998), there are other theoretical mixotroph models that are more complex. It is argued that constructing a detailed mechanistic physiological description is justified for ecological studies, despite the data limitations for parameterisations. The paper 'Building the "perfect beast": modelling mixotrophic plankton' by Flynn and Mitra (2009) assumed that such a model with variable C-N-P stoichiometry (on a cellular and regional level) is necessary to model the range of mixotroph configurations (i.e. type I, II or III and the gradients between, see Figure 1.2 and Stoecker 1998). However, obtaining stoichiometric measurements *in situ* is highly impractical at present.

Kooijman and colleagues have also used Dynamic Energy Budget theory to investigate evolutionary issues in mixotrophy, by examining the symbiotic merger of autotrophs and heterotrophs into a single organism, a mixotroph (Kooijman et al. 2003, Kooijman et al. 2004). These more complex studies, however, were not undertaken in the context of an ecosystem but at the physiological level of an individual organism. At present, mechanistic quantification of processes and internal elemental composition is insufficient in mixotrophs for a data-driven physiological mechanistic-stoichiometric model to be constructed. Hence a different type of model is not used in this thesis.

1.4. Research Objectives

Over the next five Chapters, the research presented aims to further current knowledge of mixotrophy, using a multidisciplinary approach, by combining a range of observations and ecological modelling. A mixotroph model is constructed and parameterised, using *in situ* data, a mixotroph model for the subtropical Atlantic region (**Chapters 4 and 5**). In addition to this, a number of fundamental model structure questions and assumptions are addressed: In terms of ultraplankton, can the Atlantic be considered a single province (**Chapter 2**)? Is a steady state model suitable (**Chapters 2 and 3**)? Is a dissolved organic P variable necessary for a model with mixotrophs (**Chapter 3**)? Do the observed variability and spatial scales of the different microbial groups support an assumption of uniformity in a mixotroph model (**Chapter 6**)? Can satellites accurately quantifiably sense ultraplankton?

The main aim of this thesis is:

'To explore how mixotrophy may be modelled in the subtropical Atlantic using a data driven approach.'

The objectives of this thesis are as follows:

Objective 1: To ascertain if the subtropical Atlantic ocean can be considered as a single oceanographic province (**Chapter 2**).

Objective 2: To investigate dissolved organic phosphate (DOP) utilisation by ultraplankton in the Atlantic oligotrophic ocean (**Chapter 3**), in order to identify if an organic component needs to be included in models representing mixotrophy (**Chapters 4 and 5**).

Objective 3: To develop and to parameterise from *in situ* observations a simple zero-dimensional model of the Atlantic microbial ecosystem incorporating mixotrophy (**Chapters 4 and 5**).

Objective 4: To explore microbial spatial distribution and variability throughout the subtropical Atlantic Ocean (**Chapter 6**).

Objective 5: To investigate if remotely sensed satellite chlorophyll *a* concentration can be used to estimate the abundances, biomass or chlorophyll *a* content of phototrophic ultraplankton (**Chapter 6**).

The above objectives form the basis of **Chapters 2 to 6**. The introduction of each Chapter provides a more in depth overview applicable to the objectives being addressed in that Chapter together with details of the relevant hypotheses. The results are then discussed in the light of the hypotheses detailed at the beginning of each Chapter. **Chapter 7** synthesises and summarises the key findings of **Chapters 2 to 6** and relates them to the objectives listed above.

2. Microbial Partitioning: Can Ultraplankton Assemblages be used to define the Biogeography of the Atlantic Ocean?

2.1. Introduction

The Oceans cover approximately 71 % of the earth's surface, within which vastly different features are found, from depths, to currents, to temperature and salinity. Distinctive biogeographic provinces have nevertheless been recognised within this array of contrasting environments and defined albeit with differing criteria, since Sverdrup et al. (1942) 70 years ago. The province concept provides a framework to enable analysis over broad regions of the oceans by aggregating or separating data. Thus provinces have been employed in many Atlantic biogeochemistry studies, focusing for example on primary production (Tilstone et al. 2009), bacterial communities (Gomez-Pereira et al. 2010) and phytoplankton size structure (Marañón et al. 2001). The most widespread set of criteria is that defined by Longhurst (1995, 1998, 2007). This Chapter will address whether these previously defined provinces, which are typically based on physical processes and remotely sensed data, are applicable to *in situ* ultraplankton (< 5 µm) assemblages. This will be done by applying multivariate analysis to a high spatial resolution data set of abundances for six ultraplankton groups. This issue is key to modelling the Atlantic subtropical region: Can the Atlantic be considered and modelled as a single unit or can models only be applied to specific smaller sub-areas?

2.1.1. Previous Approaches to defining Provinces

Remote sensing (McClain et al. 2004), modelling (Follows et al. 2007) and physical features (Longhurst 1995, 1998, 2007) have all been used previously to define

biogeographic provinces. The most widely applied criteria originate from a seminal paper by Alan Longhurst (1995), who divided the oceans using physical oceanographic processes and ocean colour images from satellite observations. Longhurst's partitioning has since been used widely, and often without question, in a range of studies from carbon flux (Boyd and Newton 1999) to zooplankton assemblages (Woodd-Walker et al. 2002). This study's sampling transect (AMT14 cruise labelled in red on Figure 2.1 and detailed in Appendix Table A.1) passes through five Longhurst defined biogeographic provinces (see Figure 2.1 and following description).

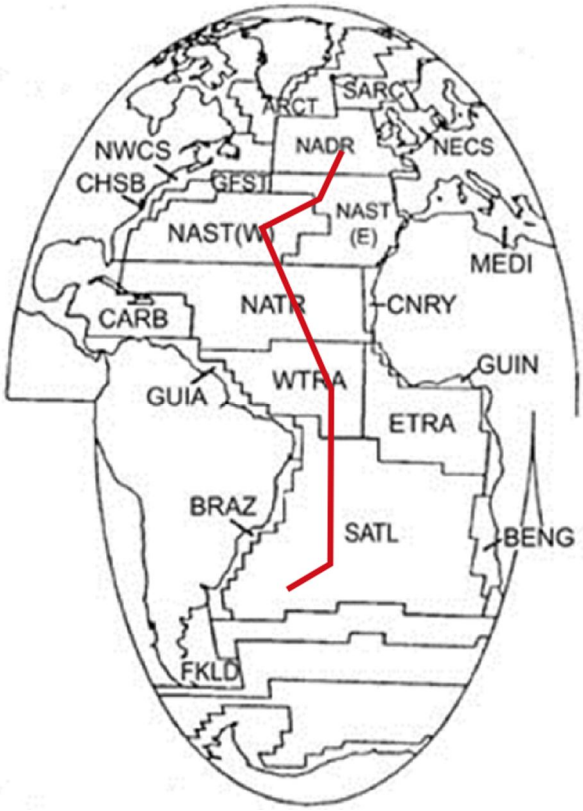


Figure 2.1. Longhurst provinces in the Atlantic (Longhurst 1998). The AMT14 sampling transect (red solid line) passes through the North Atlantic Drift (NADR), the North Atlantic Subtropical Gyre (NAST), the North Atlantic Tropical Gyre (NATR), the Western Tropical Atlantic (WTRA) and the South Atlantic Gyre (SATL) provinces.

The North Atlantic Drift (NADR) province is characterised by deep winter mixing and a strong spring bloom. The North Atlantic Subtropical Gyre - East (NAST) is, in contrast to the NADR, oligotrophic with a weak spring bloom and low levels of primary production. In the North Atlantic Tropical Gyre (NATR), phytoplankton biomass and productivity are minimal, with mixed layer depth (MLD) varying marginally with seasonality throughout the year. The Western Tropical Atlantic (WTRA) is influenced by the Equatorial current system that is driven by seasonal changes in wind forcing and during most of the year a

band of enhanced chlorophyll *a* (chl *a*) is evident. Finally, the South Atlantic Gyre (SATL) is also oligotrophic, but does accumulate phytoplankton during the austral summer and a seasonal variation in MLD is found (Longhurst 1995, 1998, 2007).

Two other ways to delineate regions considered here, are not as commonly used as Longhurst provinces. McClain et al. (2004) used remotely sensed chl *a* concentrations to delineate oligotrophic (chl *a* ≤ $0.07\ \mathrm{mg\ m^{-3}}$) and non-oligotrophic (chl *a* > 0.07 mg m^{-3}) regions. Longhurst provinces NAST, NATR and SATL regions fall into McClain et al. (2004) oligotrophic category, with the temperate (NADR) and parts of the equatorial region (WTRA) defined as non-oligotrophic regions by McClain et al. (2004) criterion.

The emergent biogeography model of Follows et al. (2007) used physiological traits of phytoplankton types in a "self-assembling" marine ecosystem and biogeochemistry model to predict distributions of plankton communities. This model identified two distinct regions in the Atlantic: one dominated by an organism analogous to *Prochlorococcus* spp. (*Pro*), extending from approximately 20° north to 20° south of the equator (covering the WTRA Longhurst defined province and parts of the SATL and NATR). The other region was dominated by small photo-autotrophs that encompassed the remaining north and south Atlantic up to ~ 55°. Photo-autotrophs are analogous to mixotrophs, because as already stated, algae in the Atlantic are primarily 'mixotrophs' (see Section 1.3.1, Zubkov and Tarran 2008, Hartmann et al. 2012). In the context of Longhurst provinces, Follows et al. (2007) analogous photo-autotroph regions covers the NADR and NAST provinces and parts of the NATR in the north Atlantic and the majority of the SATL in the south Atlantic.

The three approaches to defining provinces are clearly not always in agreement. A key question is therefore: "how do these provinces relate to ultraplankton biogeography"? There is no *a priori* reason why Longhurst provinces should apply to ultraplankton, as no microbial data apart from ocean colour (which is commonly used as a proxy of total phytoplankton, although does not separately distinguish ultraplankton) was used in their definition or validation. McClain et al. (2004) also utilised remotely sensed chl *a*, but no other ultraplankton measurements. Detrimental to these two methods is evidence that the smallest photosynthesising planktonic organisms that dominate the oligotrophic Atlantic may be being missed by ocean colour remote sensing (Zubkov and Quartly 2003). This is investigated further in Chapter 6. The emergent microbial model of Follows et al. (2007) does attempt to identify regions by investigating the microbial community and is the only

one to compare and find agreement between *in situ* measurements (Johnson et al. 2006) and modelled regions, in this case for *Pro*.

This Chapter uses ultraplankton abundances and multivariate analysis to define provinces along a transect within the Atlantic open ocean (~ 50°N to ~ 40°S). To understand the region of applicability and to model it, is important to know over what extent the community can be viewed as similar and to what extent it is constrained by physical boundaries and parameters, such as those used by Longhurst (1998, 2007). A similar stance to biogeography has been taken by Gomez-Pereira et al. (2010) as part of their study on *Flavobacteria* abundances which undertook multivariate analysis. However Gomez-Pereira et al. (2010) study only utilised 18 sample locations, along a ~30° transect, and this Chapter utilises a richer dataset in terms of sample number, spatial resolution and quantity of microbial groups used. This present study also covers a wider area of the Atlantic (Gomez-Pereira et al. was restricted to 34°N to 65°N), crossing five Longhurst (1995) defined provinces (see Figure 2.1). In addition this Chapter considers oligotrophic and non-oligotrophic regions as defined by McClain et al. (2004) and regions analogous to *Pro* and small photo-autotrophs as dictated by Follows et al. (2007) emergent microbial model.

This Chapter addresses the following hypotheses:

- Multivariate analysis of ultraplankton microbial group abundances can delineate distinct biogeographic regions within the Atlantic sharing similar community structure.

- These multivariate ultraplankton defined regions need not adhere to the provinces defined by Longhurst (1995, 1998, 2007), McClain et al. (2004) and Follows et al. (2007).

- A model incorporating mixotrophy will be valid for the entire subtropical Atlantic.

2.2. Method

2.2.1. Data Collection

Samples were taken along a northbound transect of the Atlantic Ocean from 40°S to 49°N on the AMT14 (part of the Atlantic Meridional Transect Programme) by M Zubkov on-board *RRS* James Clark Ross from 28th April to 1st June 2004, in the boreal spring and austral autumn. Water was collected along this transect from the underway seawater supply (~ 5 m depth), using a miniprep-60 autosampler, at a horizontal resolution of ~ 10 - 20 km (shown in Figure 2.3a, each point representing the location of one sample). This gave 663 samples for each ultraplankton group enumerated. A FACsort flow cytometer (Beckton Dickinson, Oxford, UK) was used to sort and count the following photosynthetic ultraplankton groups: *Prochlorococcus* spp. (*Pro*), *Synechococcus* spp. (*Syn*), Picoeukaryotes (PicoEuk) and Cryptophytes (Crypt). The majority of PicoEuk's have been found to be mixotrophic (Zubkov and Tarran, 2008). In addition, heterotrophic bacteria were enumerated and further characterised as low DNA content bacteria (LNA), high-nucleic acid bacteria with a low 90° light, or side scatter (HNAls) and a high-nucleic acid bacteria with a high 90° light scatter (HNAhs). The three heterotrophic bacterial groups (LNA, HNAls and HNAhs) were also counted as total heterotrophic bacteria (ttl). See Zubkov et al. (2007) for microbial sampling and enumeration protocol.

SeaWiFS chl *a* concentration data (level 3 mapped, rolling 32 day data at a 9 km resolution) for the duration of the cruise was provided by NASA Goddard Space Flight Centre (downloaded from http:/oceancolor.gsfc.nasa.gov/).

For each multivariate defined cluster and *a priori* defined region/province the peak and mean and associated standard error is reported for the abundances of each ultraplankton group. This aims to demonstrate the spread of the data within each cluster.

2.2.2. Multivariate Statistics applied to Abundance Data

Multivariate statistical analyses were applied to distinguish spatial groupings in ultraplankton along the oceanic Section of the AMT14 transect (samples taken over the

continental shelf, < 16°W, were excluded) following the methods described in Clarke (1993) using E-PRIMER (v. 6.0) (Clarke and Gorley 2006), see Figure 2.2.

Ultraplankton abundances were normalised to balance the contributions of common (i.e. *Pro*) and rarer groups (i.e. PicoEuk), then fourth root transformed to stabilise variance and reduce skew (Figure 2.2, Stage 2). Remotely sensed chl *a* concentration was square root transformed. The Box-Cox method was used (not shown) to determine the appropriate transformation of the data (Box and Cox 1964). Once data were normalised and transformed a pairwise Bray-Curtis similarity matrix was calculated between the *j*th and *k*th samples, S_{jk} (Figure 2.2, Stage 3).

Equation 2.1

$$S_{jk} = 100 \times \frac{\sum_{i=1}^{p} 2min(y_{ij}, y_{ik})}{\sum_{i=1}^{p} (y_{ij} + y_{ik})}$$

Here y_{ij} represents the abundance of the *i*th ultraplankton group in the *j*th sample. Likewise y_{ik} counts for the *i*th abundance in the *k*th sample. Min(.,.) represents the minimum of the two counts. The total number of samples is *p* (Clarke and Warwick 2001). To illustrate the similarity matrix, a dendrogram was constructed (Figure 2.2, Stage 4a). A similarity profile test (SIMPROF) was then used to test the null hypothesis that the resultant groups in the dendrogram (which are not *a priori* divided) did not differ from one another in multivariate structure at the 5 % level (Clarke and Gorley 2005, Clarke et al. 2008).

Multi-dimensional scaling (MDS) plots were also generated (Figure 2.2, Stage 4b). An MDS plot is a map of all samples separated by distances related to their dissimilarity. Plotting this map in two dimensions can distort it, this effect is represented by the stress function (Clarke and Warwick 2001). To ensure the minima of the stress function is reached, 25 restarts, starting at different random positions of samples in the MDS were performed.

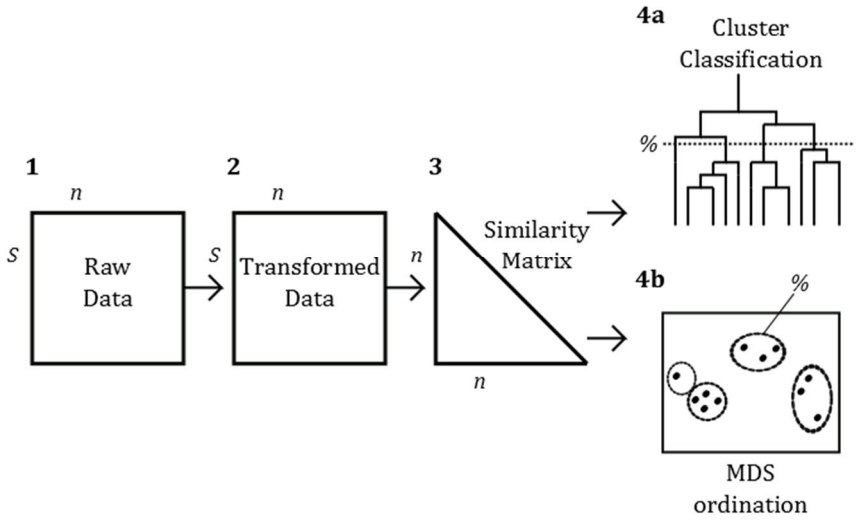


Figure 2.2. Diagrammatic summary of stages leading to cluster classification and ordination, adapted from Field et al. (1982). Stages are referred to in the text. Stage 1, raw data are presented in a matrix of samples (n) by microbial group (S). Data were normalised and transformed (Stage 2). A Bray-Curtis measure of similarity (Equation 2.1) is used to draw a comparison of each sample with every other sample resulting in a triangular similarity matrix (Stage 3). Stage 4 generates a dendrogram (Stage 4a) and multidimensional scaling plot (MDS, Stage 4b) to summarise the sample relationships.

Once clusters had been discriminated, a one-way similarity percentages routine (SIMPER) could be used on the Bray-Curtis similarity matrix to identify the key ultraplankton groups contributing to the dissimilarity *between* and the similarity *within* clusters (Clarke and Warwick 2001, Clarke and Gorley 2006).

2.2.3. Multivariate Statistics on *a Priori* Defined Regions

Preceding the *a priori* analysis, all the samples in the dataset were grouped according to the predefined biogeographic regions outlined in Section 2.1.1 – Longhurst (1998, 2007), McClain et al. (2004) and Follows et al. (2007). This enabled their relevance to the ultraplankton assemblages to be assessed. To test the null hypothesis that there were no differences in community composition in *a priori* defined regions (i.e. Longhurst provinces) a one-way analysis of similarities (ANOSIM) was used on the Bray-Curtis resemblance matrix. ANOSIM is an approximate analogue to the standard univariate ANOVA. The test statistic *R* is scaled between + 1 and - 1; a value of + 1 implies that the similarities between all samples in one province are higher than similarities between provinces. The following thresholds can be used: > 0.75 total separation; 0.75 - 0.5 weak

overlap; 0.5 - 0.25 overlap but some separation; < 0.25 no separation (Vichi et al. 2011). *R* was recomputed under the default 999 random permutations of the sample labels to achieve a significance level (Clarke 1993, Clarke and Green 1998).

2.3. Results

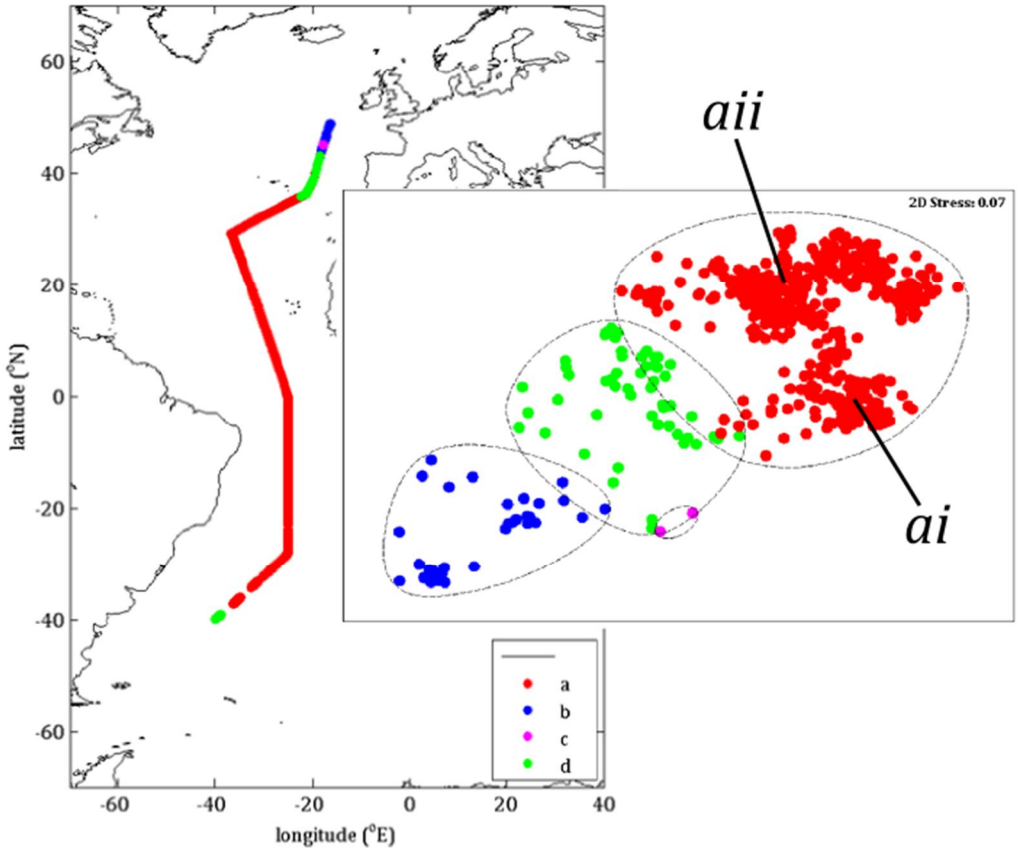
Ttl was a count of all heterotrophic bacteria (LNA, HNAhs and HNAls).Ttl was excluded from multivariate analysis as its inclusion would repeat representation of bacterial groups. It should also be noted that Crypt abundances were low throughout the transect (0 – 477 cell ml^{-1}). At low abundances, experimental error can be high, due to a decrease in the accuracy of flow cytometric counting and sorting at low abundances. Therefore Crypt is excluded from the analysis.

2.3.1. Multivariate Analysis

Cluster analysis of the underway flow cytometry sorted microbial groups (*Pro*, *Syn*, PicoEuk, LNA, HNAhs and HNAls) revealed four clusters that were all significantly different (SIMPROF, $p < 0.05$) at the > 90 % similarity level. These groups were taken on for further analysis. An accord was found between the dendrogram and MDS representations (% labelled, Figure 2.2, Stage 4a and 4b) and a low stress level of 0.07 was obtained; robustness is achieved when these are viewed in combination. MDS ordination (Figure 2.3a) displays the four clusters and Figure 2.3b the corresponding overlaid dendrogram. From here, multivariate defined ultraplankton groups are plotted geographically (Figure 2.3). Group *c* has only 2 sample points, both located at the very north of the transect (> 44°N). As this group is small and an overlap is seen in the MDS (Figure 2.3a), it is not clearly seen on the geographical transect plot and due to its low *n* will not be considered further here.

Cluster region *a* is the largest, grouping the north and south Atlantic and the equator together. Group *d* encompasses the north and south temperate region above and below group *a* (> 39°S) and between groups *a* and *b* to the north (36 - 42°N), despite these temperate areas being ~ 9000 km apart. Group *b* is found at the very north of the transect (42°- 49°N). The MDS stress is low, clusters are in agreement with the dendrogram (see Figure 2.3b) and the standard error for each cluster's microbial group abundance is relative low, all indicating that this partitioning is robust.

a.



b.

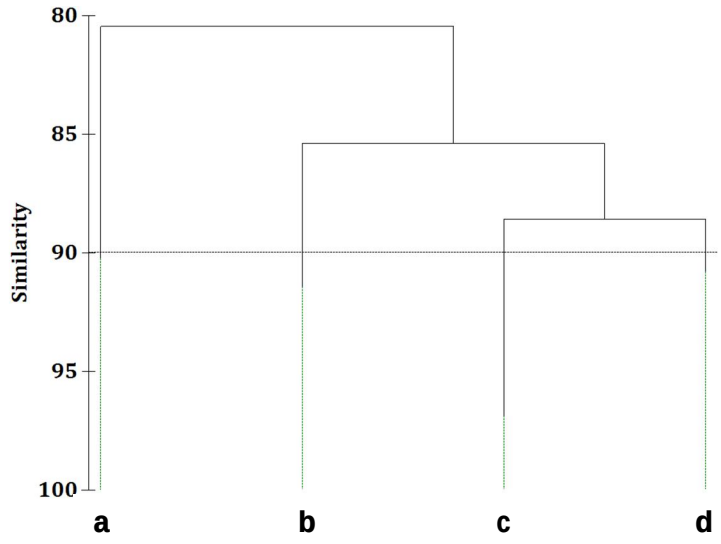


Figure 2.3. (a) AMT14 transect samples and associated multidimensional scaling (MDS) plot. Colours denote MDS > 90 % similarity groupings and dashed lines around these denote multivariate defined groups from the dendrogram in plot b. 'Sub' cluster labelled *ai* encompasses 38°S - 22°N and *aii* between 23 - 38°N. (b) Dendrogram of same data (> 90 % similarity groupings). Letters correspond to clusters in plot a.

It should also be noted that cluster region *a* could be further statistically divided into two groups if similarity was increased to 91 %: one of the north Atlantic gyre between 23 - 38°N (grouping in the bottom of MDS cluster region *a*, labelled *ai* in Figure 2.3a) and the other encompassing the whole of the south Atlantic, the equator and the southernmost part of the north Atlantic gyre, 38°S - 22°N (top grouping within MDS cluster region *a*, labelled *aii* in Figure 2.3a). This Chapter will focus on the four provinces defined from > 90 % ultraplankton similarity. The reason behind the divide in cluster region *a*, will be discussed further in the light of additional results in Chapter 6.

2.3.1.1. SIMPER Analysis of Multivariate Defined Clusters

The SIMPER test was used to identify the contribution from each ultraplankton group to the dissimilarity *between* the multivariate defined groups (Table 2.1). The greatest average dissimilarity is found between cluster region *a* and *b* (27 %, data not shown), the oligotrophic region and the northern end of the transect respectively. *Syn* and PicoEuk dominated the dissimilarity between all the groups, followed by *Pro* and HNAIs. It is also important to note that the mixotrophs (PicoEuk) were key in differentiating group *a*, from groups *b*, *c* and *d*. See Figure 2.4a, b and c for *Syn*, PicoEuk and *Pro* abundances across clusters.

Table 2.1. SIMPER analysis of variables dominating difference between clusters at the > 90 % similarity level. Percentage contribution of each variable to dissimilarity between CLUSTER defined groups is given in brackets. See Figure 2.3a.

	a	***B***	***c***	***d***
a				
b	Syn (28%) PicoEuk (26%)			
c	PicoEuk (42%) Pro (27%)	Syn (42%) HNAIs (20%)		
d	Syn (34%) PicoEuk (28%)	Pro (30%) PicoEuk (21%)	Pro (34%) PicoEuk (25%)	

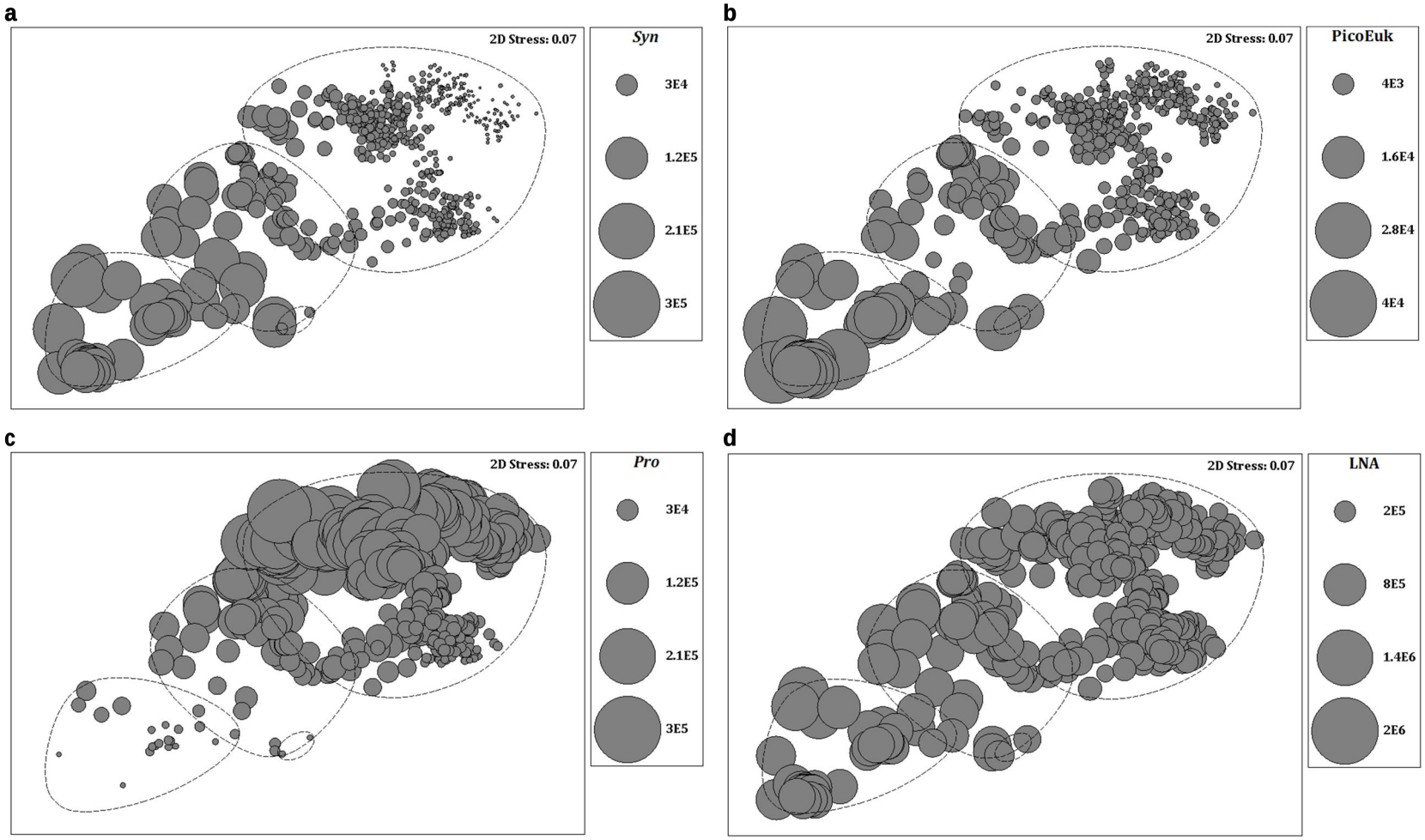


Figure 2.4. Bubble plots of ultraplankton group's abundances (cell ml^{-1}) superimposed on the MDS in Figure 2.3. Bigger circles indicate larger abundances– (a) *Synechococcus* spp (*Syn*); (b) Picoeukaryotes (PicoEuk); (c) *Prochlorococcus* spp. (*Pro*) and (d) low nuclei acid (LNA).

SIMPER was also used to identify the characteristic ultraplankton groups *within* each cluster (Table 2.2). For three of the multivariate defined clusters average similarity was typified best by PicoEuk (group *c* and *d*) or *Syn* (group *b*). The heterotrophic bacteria (LNA, HNAhs and HNAls) were most typical of cluster region *a*. LNA despite not being the most defining ultraplankton group in any of the clusters, contributed consistently (17.5 - 18.9 %) to all cluster samples similarity (this can be seen in Figure 2.4d, as the size of the bubbles, representative of abundance, are relatively consistent throughout the clusters). *Pro*, although relatively unimportant in clusters *b*, *c* and *d*, does contribute to within-cluster similarity in cluster region *a* (17.1 %) and cluster region *a* is the only cluster where its contribution to similarity is more than *Syn* (12.3 %). The contrasting distribution of *Pro* relative to PicoEuk and *Syn* can be clearly seen in Figure 2.4 a-c.

Table 2.2. SIMPER routine result for the percentage ultraplankton contribution of each microbial group to the Bray-Curtis Similarity within each four cluster regions at > 90 % similarity level, as defined in Figure 2.3a.

	a	***b***	***C***	***d***
PicoEuk	14.7	22.9	25.6	18.9
Syn	12.3	23.1	15.1	18.3
LNA	18.9	17.2	18.0	17.5
HNAls	18.2	19.5	18.1	16.0
HNAhs	18.8	15.3	16.5	16.1
Pro	17.1	2.1	6.8	13.2
Average Similarity	92.5	93.9	96.9	93.6

2.3.1.2. Microbial Abundances of Multivariate Defined Clusters

Across all samples absolute cell numbers of prokaryotic cells peaked at 1,143,200 cell ml^{-1} (LNA) and at 36,707cell ml^{-1} for eukaryotic cell. LNA was the most abundant group and PicoEuk the least. Heterotrophic bacteria (LNA, HNAhs and HNAls) abundances were consistent throughout the transect (see Figure 2.4d).

In terms of the ultraplankton defined clusters (see Figure 2.3 and Table 2.3), cluster *b* (located at the northern temperate end of the transect) had the highest mean and peak abundances for all microbial groups except for *Pro*. Cluster region *a* (the region that covers all of the oligotrophic gyres and the equatorial region) had the lowest mean abundances for all microbial groups except for HNAhs and *Pro*, whose abundances instead peaked in region *a*.

Table 2.3. Groupings defined by multivariate analysis. Peak, mean and standard error of the mean (SEM) for each ultraplankton group (*n*, sample number).

Ultraplankton Cluster		Peak abundance (cell ml^{-1})	Mean abundance (cell ml^{-1})	SEM (cell ml^{-1})
	Syn	33342	4123	± 210
	Pro	275545	92295	± 2474
a	**PicoEuk**	6751	818	± 22
***n* = 566**	**LNA**	482196	235111	± 2778
	HNAls	147564	64944	± 845
	HNAhs	354762	154393	± 2626
	Syn	205406	97222	± 5202
	Pro	25764	4098	± 987
b	**PicoEuk**	36707	18140	± 1125
***n* = 37**	**LNA**	1143228	627578	± 25357
	HNAls	653285	373868	± 21490
	HNAhs	414899	230742	± 9539
	Syn	9691	8601	± 1091
	Pro	3103	2623	± 480
c	**PicoEuk**	18193	14682	± 3511
***n* = 2**	**LNA**	398616	368721	± 29894
	HNAls	126907	117545	± 9362
	HNAhs	141008	141002	± 6
	Syn	141042	39938	± 4612
	Pro	144032	65343	± 4905
d	**PicoEuk**	12786	5530	± 264
***n* = 58**	**LNA**	1027273	487225	± 22496
	HNAls	240939	110970	± 5258
	HNAhs	364378	195132	± 7275
	Syn	205406	12466	± 1034
	Pro	275545	84745	± 2312
all	**PicoEuk**	36707	2239	± 176
***n* = 663**	**LNA**	1143228	279472	± 5467
	HNAls	653285	86370	± 3126
	HNAhs	414899	162177	± 2514

2.3.2. *A Priori* defined Provinces

2.3.2.1. ANOSIM of *a Priori* defined Provinces

An assessment of *a priori* defined provinces was conducted using ANOSIM for Longhurst (1995, 1998, 2007) provinces, McClain et al. (2004) oligotrophic and non-oligotrophic defined regions and Follow et al. (2007) emergent model defined regions. The test statistic *R* reflects the difference *between* regions, contrasted with samples from *within*. The higher the test statistic, *R*, the more contrasting the regions.

Longhurst provinces analysis excluded the area not definitively within a Longhurst provinces (i.e. those along the boundaries). These were labelled transitional (Trans) zones in Figure 2.5. The evaluation of the applicability of Longhurst (1998, 2007) provinces to the ultraplankton community, found that the NADR province is the best constrained by the ultraplankton population, as it has the greatest difference in ultraplankton relative to all other provinces (total separation, $R > 0.9$, $p < 0.001$), see Table 2.4. NAST against WTRA and SATL had a weak overlap ($R = 0.7$, $p < 0.001$) and for all other provinces ultraplankton communities overlapped with each other but had some separation ($R = 0.25 - 0.5$, $p < 0.001$), except for NATR and SATL which had no separation (i.e. difference) in ultraplankton populations ($R = 0.2$, $p < 0.001$).

Table 2.4. Longhurst (1998, 2007) ANOSIM pairwise test (*R* statistic). All significant at $p < 0.001$. Overall test statistics $R = 0.53$.

	NADR	NAST	NATR	WTRA
NAST	0.921			
NATR	1.000	0.360		
WTRA	0.999	0.701	0.447	
SATL	0.988	0.662	0.203	0.247

McClain et al. (2004) regions split the data set in two (samples > 0.07 mg m^{-3} and ≤ 0.07 mg m^{-3} chl *a*). Here it should be noted that due to cloud cover, there were data gaps in the transect. The ANOSIM was calculated only where satellite data existed (see Figure 2.6). All oligotrophically defined samples were found within ultraplankton defined cluster region *a*. Oligotrophic and non-oligotrophic regions however, compared by ANOSIM, had a test statistic of $R = 0.23$ ($p < 0.001$). This indicates no separation of oligotrophic and non-oligotrophic regions, as the ultraplankton communities barely differ in the two regions.

Follows et al. (2007) defined regions have a *R* test statistic of 0.01, a complete overlap of ultraplankton community structure. However, this was not at an acceptable level of significance and therefore not robust (see Figure 2.7).

2.3.2.2. Microbial Abundances of *a Prior* Defined Provinces

Within the Longhurst (1995, 1998, 2007) defined provinces the highest peak and mean abundances are also found at the northern end of the transect (NADR, > 44°N), except for *Pro* which peaks in abundance within the WTRA (~ 5°S - 10°N). The lowest abundances are found across the oligotrophic regions (NATR, NAST and SATL), see Figure 2.5 and Table 2.5. This Chapters' ultraplankton defined cluster regions contrast with Longhurst provinces. As cluster region *a* houses four Longhurst provinces - NAST, NATR, WTRA and SATL. Cluster region *b* is the only region to agree with a Longhurst province, the NADR (see Figure 2.5).

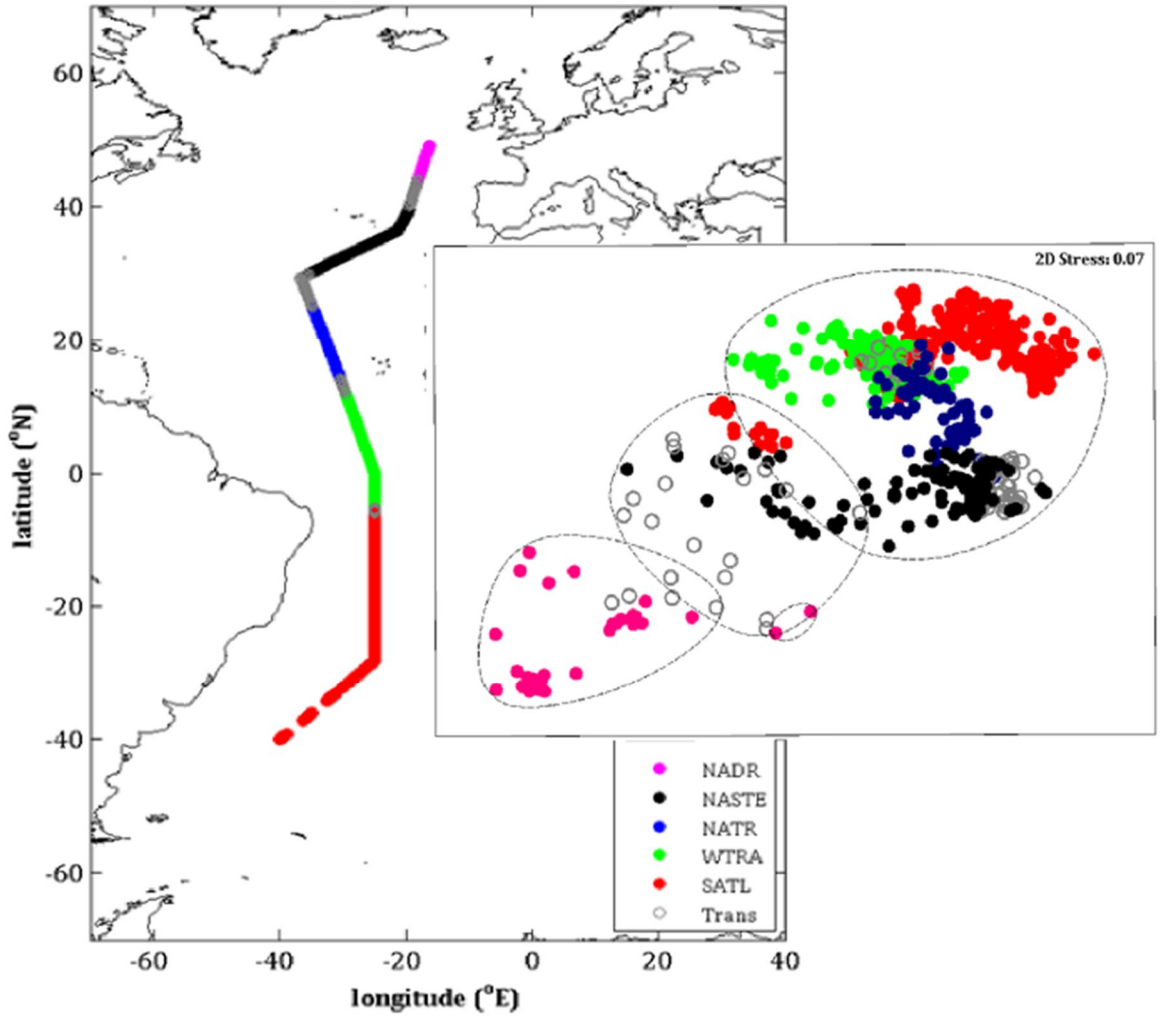


Figure 2.5. AMT14 transect samples and corresponding multidimensional scaling (MDS) plot (see Figure 2.3), with Longhurst (1998, 2007) provinces overlaid.

Table 2.5. Microbial characteristics of provinces as defined by Longhurst (1995, 1998, 2007). Peak, mean and standard error of the mean (SEM) for each ultraplankton group (n, sample number).

Longhurst Provinces		Peak abundance (cell ml^{-1})	Mean abundance (cell ml^{-1})	SEM (cell ml^{-1})
NADR *n* = 34	***Syn***	205406	91599	± 6541
	Pro	25764	3719	± 1055
	PicoEuk	36707	19513	± 985
	LNA	1143228	610701	± 28859
	HNAls	653285	387528	± 21757
	HNAhs	414899	230833	± 10428
NAST *n* = 97	***Syn***	79219	10900	± 1397
	Pro	86812	22899	± 1860
	PicoEuk	6751	2225	± 183
	LNA	1027273	280833	± 13077
	HNAls	240939	66950	± 2667
	HNAhs	364378	109293	± 4726
NATR *n* = 69	***Syn***	4679	2776	± 105
	Pro	150139	59920	± 4344
	PicoEuk	1569	980	± 26
	LNA	288976	214713	± 3048
	HNAls	93758	65045	± 1313
	HNAhs	302232	185237	± 4970
WTRA *n* = 107	***Syn***	33342	9491	± 842
	Pro	275545	149475	± 3703
	PicoEuk	2633	1003	± 37
	LNA	482196	307159	± 7569
	HNAls	147564	85257	± 2247
	HNAhs	354762	227580	± 5357
SATL *n* = 282	***Syn***	23901	3373	± 273
	Pro	214881	108374	± 2419
	PicoEuk	7847	960	± 94
	LNA	438556	226009	± 4274
	HNAls	138722	61805	± 1469
	HNAhs	280981	150475	± 2797

Using the McClain et al. (2004) criteria, oligotrophic regions (chl *a* ≤ 0.07 mg m^{-3}) had lower ultraplankton peak and mean abundance, than non-oligotrophic regions (chl *a* > 0.07 mg m^{-3}), see Figure 2.6 and Table 2.5. In terms of this Chapters' ultraplankton defined regions, cluster region *a* housed a mixture of oligotrophic and non-oligotrophic defined samples and cluster regions *b* to *d* were all classed as non-oligotrophic according to the McClain et al. (2004) criteria (Figure 2.6).

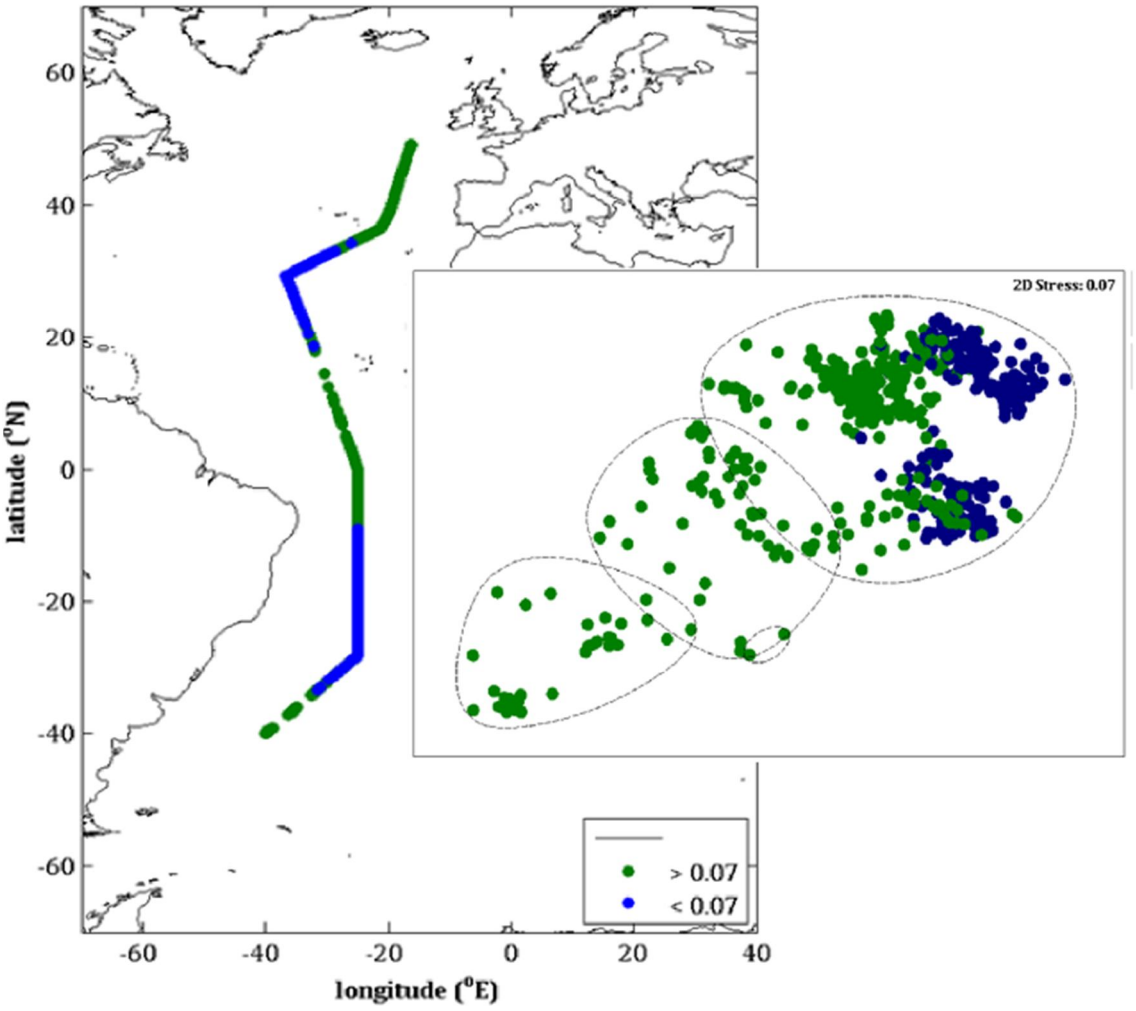


Figure 2.6. AMT14 transect samples and corresponding multidimensional scaling (MDS) plot (see Figure 2.3), with McClain et al. (2004) chlorophyll *a* concentration defined regions (> or < 0.07 mg m^{-3}) overlaid.

Table 2.6. Groupings defined using McClain, et al. (2004) criteria for oligotrophic and non-oligotrophic regions (≤ or > 0.07 mg m^{-3}, derived from SeaWiFS). Peak, mean and standard error of the mean (SEM) for each ultraplankton group (*n*, sample number).

McClain et al. (2004) regions		Peak abundance (cell ml^{-1})	Mean abundance (cell ml^{-1})	SEM (cell ml^{-1})
Oligotrophic ≤ 0.07 mg m^{-3} *n* = 225	***Syn***	5774	1708	± 93
	Pro	213282	54200	± 2986
	PicoEuk	1569	623	± 15
	LNA	276014	192960	± 2082
	HNAls	92151	54698	± 917
	HNAhs	214276	109227	± 2608
Non-Oligotrophic > 0.07 mg m^{-3} *n* = 351	***Syn***	205406	20564	± 1793
	Pro	275545	95793	± 3306
	PicoEuk	36707	3558	± 312
	LNA	1054667	334624	± 8237
	HNAls	595960	107763	± 5312
	HNAhs	374152	183054	± 3011

The Follows et al. (2007) emergent model defined grouping that is analogous to a region dominated by *Pro* had the lowest peak and mean abundances for all microbial groups, except for *Pro*. The opposite was found in the region analogous to small photo-autotrophs (see Figure 2.7 and Table 2.6). Similar to the McClain et al. (2004) criteria, cluster region *a* defined by ultraplankton in this Chapter had a mixture of photo-autotrophs and *Pro* defined samples, and cluster regions *b* to *d* were all analogous to photo-autotrophs.

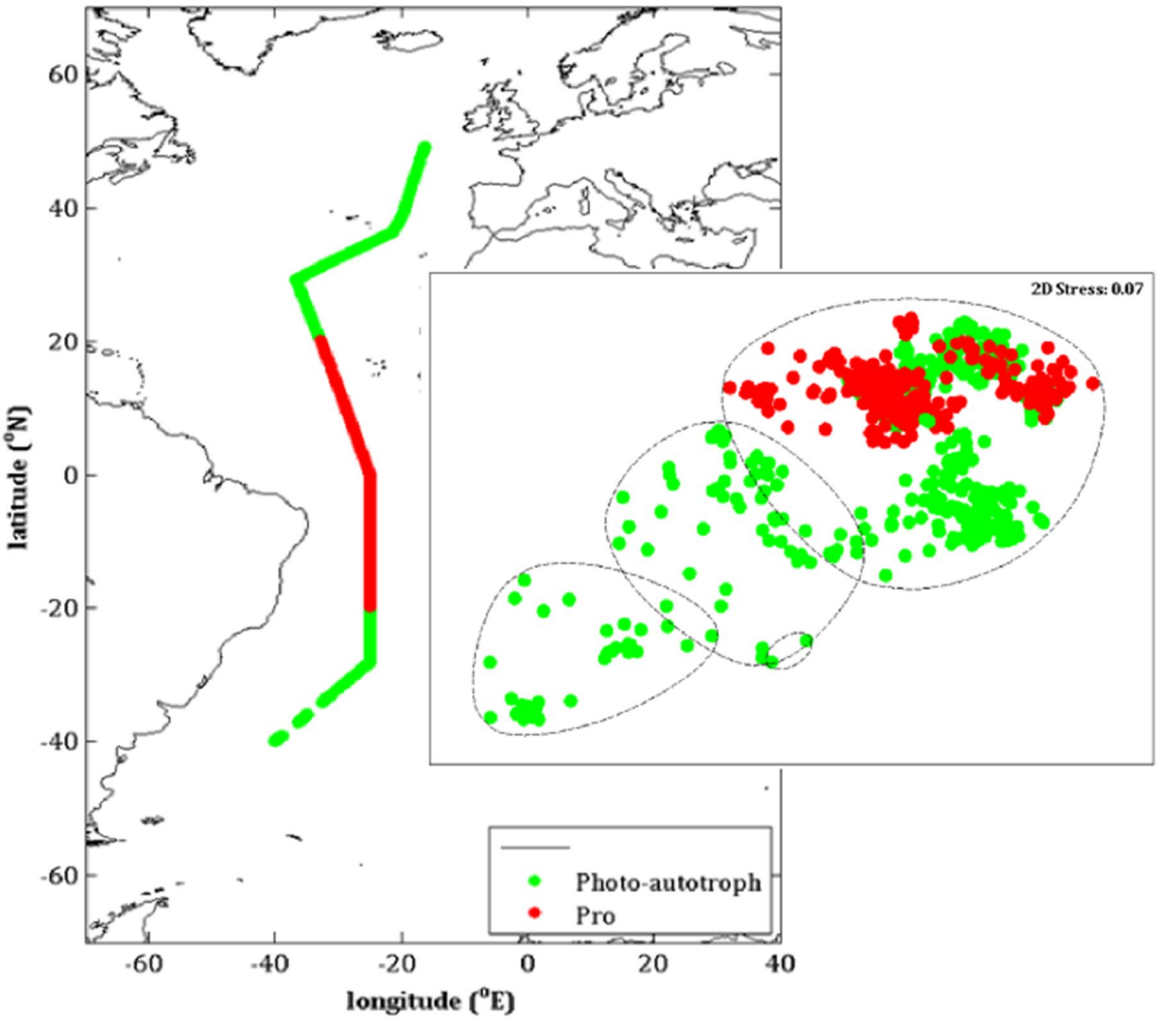


Figure 2.7. AMT14 transect samples and corresponding multidimensional scaling (MDS) plot (see Figure 2.3), with Follows et al. (2007) emergent model regions for *Pro* and Photo-autotrophs overlaid

Table 2.7. Follows et al. (2007) emergent model dictated regions defined as analogous to *Pro* spp. and to small photo-autotrophs - peak, mean and standard error of the mean (SEM) for each ultraplankton group (*n*, sample number).

Follows et al. (2007) regions analogous to:		Peak abundance (cell ml^{-1})	Mean abundance (cell ml^{-1})	SEM (cell ml^{-1})
Photo-autotrophs *n* = 431	***Syn***	205406	16278	± 1541
	Pro	214881	65788	± 2719
	PicoEuk	36707	2990	± 263
	LNA	1143228	291742	± 7889
	HNAls	653285	95174	± 4677
	HNAhs	414899	148193	± 2709
***Pro* *n* = 232**	***Syn***	33342	5382	± 467
	Pro	275545	119962	± 3154
	PicoEuk	2633	844	± 24
	LNA	482196	256676	± 5103
	HNAls	147564	70012	± 1623
	HNAhs	354762	188156	± 4676

2.4. Discussion

This Chapter defined four Atlantic Provinces on the basis of ultraplankton abundances (see Figure 2.3). This study shows that ultraplankton community structure is clearly delineated along the transect studied and that ultraplankton communities can therefore be used to define a large oligotrophic subtropical and tropical Atlantic province, characterised by similar abundances.

The large region, cluster region *a*, was found to coherently describe both the north and south oligotrophic Atlantic and the equatorial region, regardless of the transect sampling two opposing seasons (boreal spring and austral autumn). As seasonality does not affect the delineation of provinces, this Chapter gives some support to the assumption of steady state made when the mixotroph model is first developed in Chapter 4. This result is also supported by previous observational studies on the stability of the meridional distribution of picoplanktonic groups over differing seasons (Zubkov et al. 2000). Cluster region *a* differs from previous frameworks that have separately defined the north and south Atlantic (e.g. Pauly 1998, Longhurst 1998, 2007). Some support for the NATR-NAST Longhurst province defined boundaries, however, is found. If cluster region *a* was further separated into two 'sub' cluster regions (*ai* and *aii*, see Figure 2.3a), the boundary (~ 23°N) would lie close to that of the Longhurst (1998, 2007) NATR-NAST boundary (25 - 30°N) which, was defined by the presence of the subtropical convergence zone. Nevertheless this study's result is more consistent with recent emergent models, that

produce provinces that are inclusive of regions both north and south of the equator (e.g. Lewis et al. 2008, Dutkiewicz et al. 2009). This suggests that provinces defined by physical processes have limited application when describing ultraplankton distributions.

The existence of a single province encompassing both gyres for ultraplankton groups is important for future *in situ* observations and modelling. Limited samples, due to methodological, time and monetary constraints on measurements in the oceanic Atlantic, may be less of a problem than previously thought, because samples could be localised and extrapolation of results could be applied across the province. Similarly in modelling, which often requires a large amount of data for parameterisation, one area could be anticipated to be the equivalent of another within the same province, and then extrapolations for the province as a whole could be made (if spatial variability is ignored, Chapter 6 examines this assumption).

The multivariate analyses of the ultraplankton groups showed that > 50 % of the distinction between the main cluster regions (*a*, *b* and *d*) was due to the PicoEuk and either *Syn* or *Pro* microbial groups (see Table 2.1). Previous studies have shown that not only are these groups major fixers of inorganic carbon (Li 1994), but PicoEuk's have also been observed to be predominantly responsible for bacterivory (see Chapter 1) within the temperate and tropical North Atlantic Ocean (Zubkov and Tarran 2008). PicoEuk's as a key defining ultraplankton group of cluster region *a* is reassuring for later models, incorporating mixotrophy in this thesis (Chapter 4 and 5), models are based upon this multivariate defined cluster.

Results presented in this chapter also quantitatively tested the applicability of previously defined provinces to ultraplankton communities. Longhurst (1998, 2007) states that bottom up control by physical drivers defines biogeochemical provinces. Therefore, at the base of the marine ecosystem, these physical forces should directly affect ultraplankton. Consequently, one might expect that ultraplankton communities would have been given consideration. Longhurst (1998) gave a brief review of phytoplankton and zooplankton species interactions. However, bacterioplankton are absent from the first edition and in the second edition are discussed only briefly (Longhurst 2007), referring to two graphs of Atlantic transects (from Li, 1995, Zubkov et al. 1998). Nonetheless since publication, the Longhurst province framework has been used in numerous studies, including those on microbes (e.g. Boyd and Newton 1999, Gomez-Pereira et al. 2010, Freidline et al. 2012). In this study the NADR is shown to be best constrained by an ultraplankton population

distinct from other provinces (see Table 2.7). This is the northernmost part of the transect, a temperate region also distinguished by clusters (*b*, *c* and *d*, see Figure 2.5). Longhurst provinces, however, incorrectly distinguish the SATL from other subtropical and tropical Atlantic Longhurst defined regions (NAST, NATR and WTRA), as the north and south Atlantic gyre regions have no significant difference in ultraplankton populations. However, if similarity were to be increased in this study, the boundary of NATR-NAST would be consistent with the distribution of the ultraplankton (*ai* and *aii* in Figure 2.3a).

Secondly, the current multivariate analysis result was compared to McClain et al.'s (2004) oligotrophic and non-oligotrophic regions. Figure 2.6, shows relatively small gyre regions using the McClain et al. criteria (~ 21 - 32°N and ~ 10 – 30°S, similar to those found by Polovina et al. 2008), north and south of a large non-oligotrophic equatorial area. This is in contrast to the cluster results from this study where a single region was defined that covered these differing McClain regions (see cluster region *a*, in Figure 2.3a and Figure 2.6). Within cluster group *a*, approximately equal numbers of McClain et al. (2004) defined oligotrophic and non-oligotrophic samples are present. In cluster regions *b*, *c* and *d* all samples are non-oligotrophic according to McClain et al. (2004). The ANOSIM test statistic reported no difference between oligotrophic and non-oligotrophic ultraplankton communities ($R = 0.23$, $p < 0.001$) using the McClain et al. (2004) definition. This is an interesting result in two respects, of which the latter may be the cause of the former. Firstly, it may suggest that the oligotrophic 'bar' (chl *a* ≤ 0.07 mg m^{-3}) has been set too low (Antoine et al. 1996 set it at chl *a* ≤ 0.1 mg m^{-3}). Secondly, it may support previous evidence that suggests inadequacies in the remote sensing of chl *a* for phototrophic ultraplankton groups (Zubkov and Quartly 2003). This in turn brings the accuracy of determining ultraplankton communities by remote sensing into question. Phototrophic ultraplankton detection by remote sensing is investigated with this dataset, in Chapter 6 of this thesis.

Thirdly, the region defined as analogous of *Pro* by the Follows et al. (2007) emergent model is smaller than that defined by the presented ultraplankton multivariate analysis (cluster region *a*, see Figure 2.3 and 2.7). The difference between the two Follows et al. (2007) defined regions (photo-autotroph and *Pro*) could also not be shown to be statistically different on the basis of ultraplankton abundances. Dutkiewicz et al. (2009) expanded on the Follows et al. (2007) model, using similar equations except for a change in the grazing term (for variable palatability of phytoplankton and sloppy feeding). This led to an even larger area around the equator dominated by *Pro*. Therefore, this expanded

Follows et al. (2007) model may be more appropriate to describe ultraplankton community regions within the Atlantic.

Out of all the previous regional definitions discussed, none matches the regions extracted from the ultraplankton community. The multivariate cluster regions defined here bear little resemblance to Longhurst (1998, 2007) provinces. As well as enveloping the entire oligotrophic region of McClain et al. (2004) and part of the non-oligotrophic region. The ultraplankton defined cluster region *a* spans a wider area than those analogous to *Pro* defined by Follows et al. (2007). This suggests that the ultraplankton-defined clusters presented here are of more relevance to ultraplankton studies, especially those in the tropical and subtropical Atlantic, than *a priori* defined provinces. Thus, future studies of ultraplankton should not use unconditionally regions previously defined. Studies ideally ought to conduct multivariate analysis on each occasion to give confidence in region definitions.

2.5. Summary and Implications

This investigation found that samples taken from the Atlantic Meridional Transect could be grouped into four provinces (at > 90 % similarity level) using ultraplankton group abundances, the largest of which incorporates the north and south subtropical and tropical region (~ 36°N to ~ 39°S). SIMPER analysis of these multivariate defined groups, showed that > 50 % of the distinction between the main cluster regions (*a*, *b* and *d*) was due to the PicoEuk (mixotrophic algae) and either *Syn* or *Pro* microbial groups (see Table 2.1).

This multivariate defined cluster region *a*, will be used in the subsequent Chapters, in the following ways:

- Dissolved organic phosphate uptake by bacterioplankton will be compared within cluster region *a* multivariate defined provinces (Chapter 3)
- The Atlantic subtropical and tropical region (cluster region *a*) has been defined on the basis of similarity in ultraplankton assemblages. Therefore a model constructed incorporating mixotrophy, is applicable for this entire region. This also helps to the support an assumption of steady state (Chapters 4 and 5), as seasonality does not influence region similarity.

- The multivariate method in this Chapter does not reveal any information on spatial homogeneity in cluster region *a*. To evaluate this, Chapter 6 assesses microbial group spatial variability within cluster region *a* which appears to displays homogeneity at the ultraplankton community scale.

- An assessment will be made of the ability of remotely sensed chl *a* to detect ultraplankton groups (Chapter 6), this question having arisen from the comparison of McClain et al. (2004) oligotrophic regions to ultraplankton defined provinces.

From the results presented in this chapter and considering the three stated hypothesis to be tested the following can be concluded:

- Multivariate analyses of ultraplankton microbial group abundances can delineate biogeographic regions within the Atlantic.

- Ultraplankton defined regions do not equate to previously defined provinces.

- A model incorporating mixotrophy is valid for the entire subtropics and tropical Atlantic.

3. Dissolved Organic Phosphate - an Alternative Source of Phosphate in the Subtropical North and Central Atlantic?

3.1. Introduction

Phosphorus is one of the main elements controlling primary production in the oceans. It is a key macronutrient, a component of DNA, RNA and lipids and is a limiting nutrient for productivity in the subtropical and tropical north Atlantic subtropical gyre (NAG) (Tyrrell 1999, Mills et al. 2004, Mills et al. 2008). The NAG has very low concentrations of phosphorus (Zubkov et al. 2007), in its inorganic form, phosphate. Measurements of dissolved inorganic phosphate (DIP) uptake by microbes showed that DIP uptake was too low to sustain the observed picoeukaryotic (algae) biomass (Hartmann et al. 2011). In spite of this, the NAG has comparable rates of carbon fixation to the south Atlantic subtropical gyre, which is not P limited (Poulton et al. 2006). This suggests the utilisation of an alternative phosphorus source to satisfy demand.

Dissolved organic phosphate (DOP) represents a considerable fraction (~ 70 – 80 %) of the total dissolved phosphorus (comprised of DOP and DIP) pool (Karl and Björkman 2002). The composition of DOP is poorly understood, although a study has indicated a predominance of phosphonate (C-P) and P-ester (C-O-P) bond classes (Kolowith et al. 2001), the latter of which are considered more labile to primary producers (P-ester examples include ATP, UMP and AMP the nucleotides in this study, which are explained below). It has previously been suggested in studies from the Sargasso Sea (e.g. Casey et al. 2009, Lomas et al. 2010, Michelou et al. 2011) and *ex situ* laboratory cultures (Wang et al. 2011) that DOP may be being significantly utilised by plankton. Zubkov et al. (2007) demonstrated however, through a sensitive bioassay radioactive labelling technique and

flow cytometry that ATP represented only a small fraction of the DIP uptake in the NAG. Casey et al. (2009) and Michelou et al. (2011) used a similar technique to Zubkov et al. (2007) to measure assimilation of ATP by different picophytoplankton groups, concluding contrary to Zubkov et al. (2007) that all the microbial groups studied assimilated ATP at a significant rate (Casey et al. 2009) and that heterotrophic bacteria dominated ATP uptake, while cyanobacteria accounted for less than 10 % of the total ATP uptake (Michelou et al. 2011). This conclusion of significant ATP utilisation was further supported by Lomas et al. (2010), who measured DOP indirectly by subtracting DIP from total dissolved P, and concluding that DOP supported approximately 25% of annual primary production. Nevertheless these three studies (Casey et al. 2009, Lomas et al. 2010, Michelou et al. 2011), took place within the Sargasso Sea, which despite bordering the NAG has differing biogeochemical variables and are not here considered analogous to the north Atlantic gyre. Casey et al. (2009) and Michelou et al. (2011) also assumed a 1 $nmol^{-1}$ ambient concentration of ATP for calculations, based on bioassay data from the NAG from Zubkov et al. (2007) and chemical data from the North Pacific (Karl and Bossard 1985). They used this assumed concentration of ATP despite acknowledging the trans-Atlantic difference between the oligotrophic Atlantic and Sargasso Sea, and regardless of the bioassay techniques they employed being able to deduce maximum ambient concentrations. The reason they omitted to use direct observations of concentration are not explained.

Past studies of DOP have concentrated on ATP (Adenosine-5'-triphosphate, which contains three phosphate groups, a purine base - adenine and a pentose sugar - ribose) as a model compound, despite its instability in seawater, assuming it is the main component of, or proxy for, the total DOP pool (e.g. Zubkov et al. 2007, Casey et al. 2009, Orchard et al. 2010, Michelou et al. 2011, Bjorkman et al. 2012). ATP, however, is only one constituent of DOP (Karl and Björkman 2002). Here for the first time, two other nucleotides, that are also monomers of RNA are studied: Uridine monophosphate, also known as 5'-uridylic acid (UMP), and adenosine monophosphate (AMP), or 5'-adenylic acid. ATP and AMP are known to be dephosphorylated extracellularly with the phosphate groups and the adenine base then separately transported by a carrier-mediated transport system (Bengis-Garber and Kushner 1982, Bengis-Garber 1983). To the author's knowledge no studies have been undertaken on microbial UMP uptake transport. Here it is hypothesised that the uracil base is taken up in the same way as the ATP and AMP adenine bases. Therefore, the utilisation of these three phosphate nucleotide constituents of the DOP pool are investigated in the NAG and central Atlantic, to assess if any of the nucleotides tested are consistently being utilised at a significant rate by the marine microbial pool (> 0.2 µm) or

if there is a significant difference between these nucleotides, and hence to evaluate if DOP is a significant alternative source of P to DIP in this P limited region.

It has previously been stated (Chapter 1) that recent research has indicated mixotrophy (bacterivory) as a major source of phosphate to algae (Zubkov and Tarran 2008, Hartmann et al. 2012). This finding changed the basic understanding of the oligotrophic food web, as it established that plastidic protists (algae) are the foremost bacterivores as well as the main CO_2 fixers in this oligotrophic ecosystem (Li 1994, Hartmann et al. 2012). Motivated by this discovery, a central aim of this thesis is the development of a mixotroph model (in Chapters 4 and 5). Prior to doing so, however, it is necessary and prudent to examine whether DOP is an alternative source of phosphate that needs to be included in the model.

The aim of data presented in this chapter was to determine if, in the tropical and subtropical north and central Atlantic ocean, DOP is being used as a significant source of phosphorus by plankton (> 0.2 µm). In addition to assessing this by examining three different nucleotide fractions of the DOP pool (ATP, AMP and UMP), it also examines if there is inter-annual variability or seasonality in the ATP fraction of the DOP pool. The overall objective of this Chapter is to ascertain if a separate DOP pool is required in a model to capture the population dynamics of the tropical and subtropical north Atlantic ocean.

This Chapter addresses the following hypotheses:

- DOP is utilised as an alternative significant source of phosphate (to DIP) by plankton in the subtropical north and tropical central Atlantic.
- For the three DOP nucleotides measured (ATP, AMP and UMP), turnover time, concentration and uptake by plankton are not significantly different in the subtropical north and tropical central Atlantic.
- Across the three years (2004, 2008 and 2009), inter-annual variability and seasonality is not present in ATP turnover time, concentration or uptake by plankton in the subtropical north and tropical central Atlantic.
- DOP needs to be included in a model of the oligotrophic Atlantic, as an important alternative source of P.

3.2. Method

3.2.1. Sampling

Experiments were conducted in the subtropical and tropical Atlantic during three Atlantic Meridional Transect (AMT) cruises (Table 3.1). Only ATP was collected on all three cruises. (see Figure 3.1 and Table 3.1). The analysis is restricted to the ultraplankton defined area, cluster region *a* (detailed in Chapter 2).

Table 3.1. Details of cruises and respective nucleotides measured (also see Figure 3.1).

Cruise	Date	Austral Season	Ship	Nucleotides	Collected by
AMT14	May 2004	Spring	*RRS* James Clark Ross	ATP	M Zubkov
AMT18	Oct-Nov 2008	Autumn	*RRS* James Clark Ross	ATP	M Zubkov M Hartmann
AMT19	Oct-Nov 2009	Autumn	*RRS* James Cook	ATP, AMP, UMP, DIP	S Herrington M Hartmann

Seawater samples were collected at midday (AMT14) and predawn (AMT18 and AMT19) from a depth of 20 m (representing the mixed layer and the shallowest depth not affected by the ships movement) with 20 litre niskin bottles mounted on a sampling rosette with a conductivity-temperature-depth profiler. Samples were collected in acid-washed 1 litre thermos flasks, using acid soaked silicon tubing, which were rinsed with sampled seawater prior to sample collection. Radiotracer experiments were started immediately after sampling. To avoid light effects, experiments were conducted in the dark and at *in situ* temperatures.

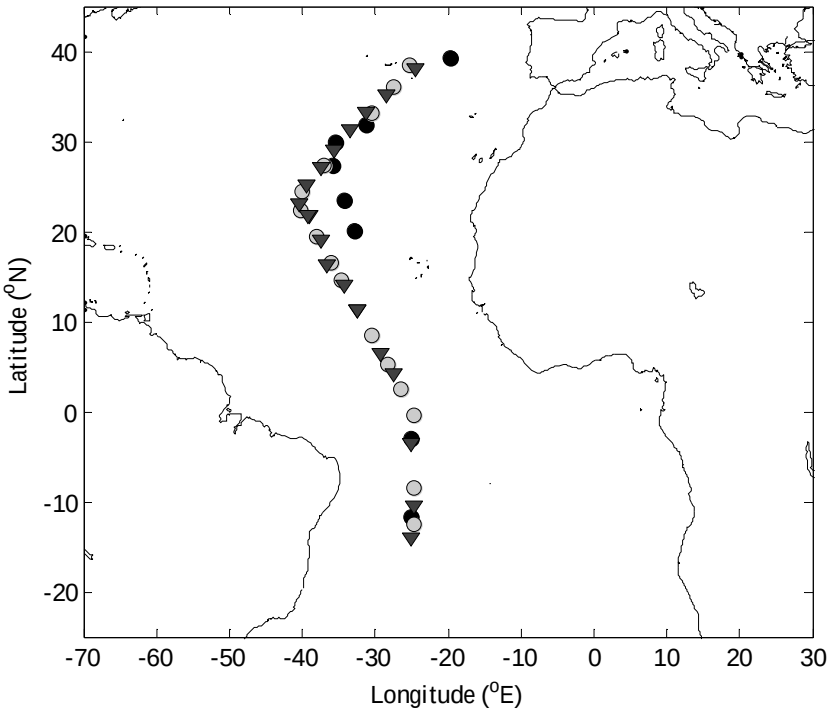


Figure 3.1. Sampling locations of AMT14 (black circles), AMT18 (light grey circles) and AMT19 (dark grey triangles) in May 2004, October/November 2008 and October/November 2009 respectively.

3.2.2. Determination of Organic and Inorganic Phosphate Ambient Concentration, Microbial Uptake Rate and Turnover Time using Bioassays

Isotopic dilution time-series incubations (Zubkov et al. 2004, Zubkov et al. 2007), referred to as bioassays, were used to measure turnover rates, ambient concentrations and microbial uptake rates. [α^{33}P]AMP, [α^{33}P]ATP (on AMT14 only [γ^{33}P]ATP) and [α^{33}P] UMP (all > 111 TBq mmol^{-1} – Hartmann Analytics GmbH, Germany) were all added to individual samples at a concentration of 0.2 nmol l^{-1} and diluted with non-labelled (cold) AMP, ATP and UMP respectively using a dilution series of 0.4, 0.6, 0.8, 1.0, 1.2 and 1.4 nmol^{-1}. [^{33}P] orthophosphoric acid (> 111 TBq mmol^{-1} – Hartmann Analytics GmbH, Germany) was added at a concentration of 0.1 nmol^{-1} and diluted with non-labelled (cold) Na_2HPO_4 using a dilution series of 0.4, 0.8, 1.6, 2.4, 3.2 and 4.0 nmol^{-1}. For all DOP and DIP samples quadruplets of 1.6 ml for each addition were incubated in 2 ml screw top sterilised polypropylene micro-centrifuge tubes. DOP samples were fixed for 1 hour using paraformaldehyde (PFA) solution (to give a final concentration of 1 %) at 15, 30, 45, and 60 min or 20, 40, 60 and 80 min (no statistically significant differences were found

between data from different timings, thus results were treated the same in further analysis). DIP samples were incubated for 20, 40, 60, and 80 min. PFA was added to give a final concentration of 1 % and fixed for one hour.

Samples were collected onto 0.2 µm (anything smaller than this size fraction was assumed to be a virus) polycarbonate filters (Poretics Corporation, autoclaved in deionised water prior to harvesting). After the sample was collected on the filter, the filter was washed twice with 2 ml of deionised water. DIP samples were also collected with 0.2 µm polycarbonate filters (Poretics Corporation, autoclaved in the following buffer) and washed with two 5-ml aliquots of a solution of 0.5 mol^{-1} LiCl and 1 $mmol^{-1}$ phosphate (LiCl-PO_4 buffer). Radioactivity remaining on the filters was measured the same day as disintegrations per minute (DPM) using a liquid scintillation counter (Tri-Carb 3100, Perkin Elmer).

Calculations of turnover time, substrate concentration and uptake rate were performed as described previously (Zubkov et al. 2007). The rate of DOP and DIP uptake was calculated as the slope of the linear regression of radioactivity against incubation time (see example in Appendix, Figure A.1a). This was then used to calculate turnover time by dividing it by total radioactivity per sample by the rate of its uptake per hour. The resulting turnover times were plotted on the *y*-axis against the corresponding concentration of DOP or DIP on the *x*-axis and extrapolated using linear regression (see example in Appendix, Figure A.1b). The slope divided by the intercept of this regression line gives an estimate of DOP or DIP microbial uptake rate. The *y*-intercept of the regression line gives an estimate of turnover time of the ambient concentration plus the transport constant (Wright and Hobbie 1966). The concentration plus the transport constant is estimated by the *x*-axis intercept. The transport constant is a measure of the affinity of the microbial uptake system for P, the lower the constant the more effective the uptake at low concentrations. The bioavailable concentration should therefore be treated as upper estimates, as oligotrophic plankton should be adapted to living at nanomolar concentrations. The transport constant is therefore assumed to be negligible compared with the concentration.

When a poor linear regression ($r^2 < 0.9$) between added P concentration and P turnover time was achieved using the above method (example of regression in the Appendix, Figure A.1b), the turnover time could still be calculated from the mean of the slopes of the regression of the radioactivity against incubation time (Figure A.1a), divided by total activity per sample by the rate of its uptake per hour (Pers. Comm. M Zubkov, see linear

regressions in Appendix, Figure A.1a).This is called here the t_2 method. To check that this turnover time estimate (t_2) was reliable, where available, the turnover time was calculated using the first method (t_1) and were compared to t_2. All t_1 were within three-sigma's of the mean turnover time calculation (t_2), so can be assumed here as equivalents (Pers. Comm. M Hartmann). Turnover times calculated using the t_2 method are labelled accordingly in the raw data table in the Appendix, Tables A.3 and A.4.

Bioassay concentration and uptake data included in the analysis are only those with correlation coefficients such that $r^2 > 0.9$ (see example Appendix, Figure A.1b, Pers. Comm. M Zubkov). This reduces uncertainty in the results, as those with low r^2 are assumed to be at or below the detection limit of the method (Pers. Comm. M Hartmann). All raw data used within the following analysis are shown in Appendix Tables A.3 and A.4.

3.2.3. Data Analysis

All statistical and computational analysis was completed in SigmaPlot 12.0 and GraphPad Prism 5 software. All data failed a normality test (Shapiro-Wilk, $p < 0.05$). Therefore non-parametric Mann-Whitney Rank Sum Test, Spearman Rank Correlation's and Kruskal-Wallis Analysis of Variance (ANOVA) on ranks were performed. Results were tested for statistical significance at $p < 0.05$.

The following results and discussion is on the measurements made within cluster region *a* (see Chapter 2). The turnover time and concentration are thought to be the more robust measurements (Pers. Comm. M Zubkov), as these are calculated from the intercept and slope (see Appendix Figure A.1b), whereas the uptake rate is a product of both the slope and the intercept and as such has twice the uncertainty. Therefore it will be discussed only in the context of supporting evidence.

3.3. Results

All data tested were located within cluster region *a*, defined in Chapter 2. Figure 3.2a-c details the average results and Figure 3.2d-f details the results by latitude, to show the spread of the nucleotide data within cluster region *a*

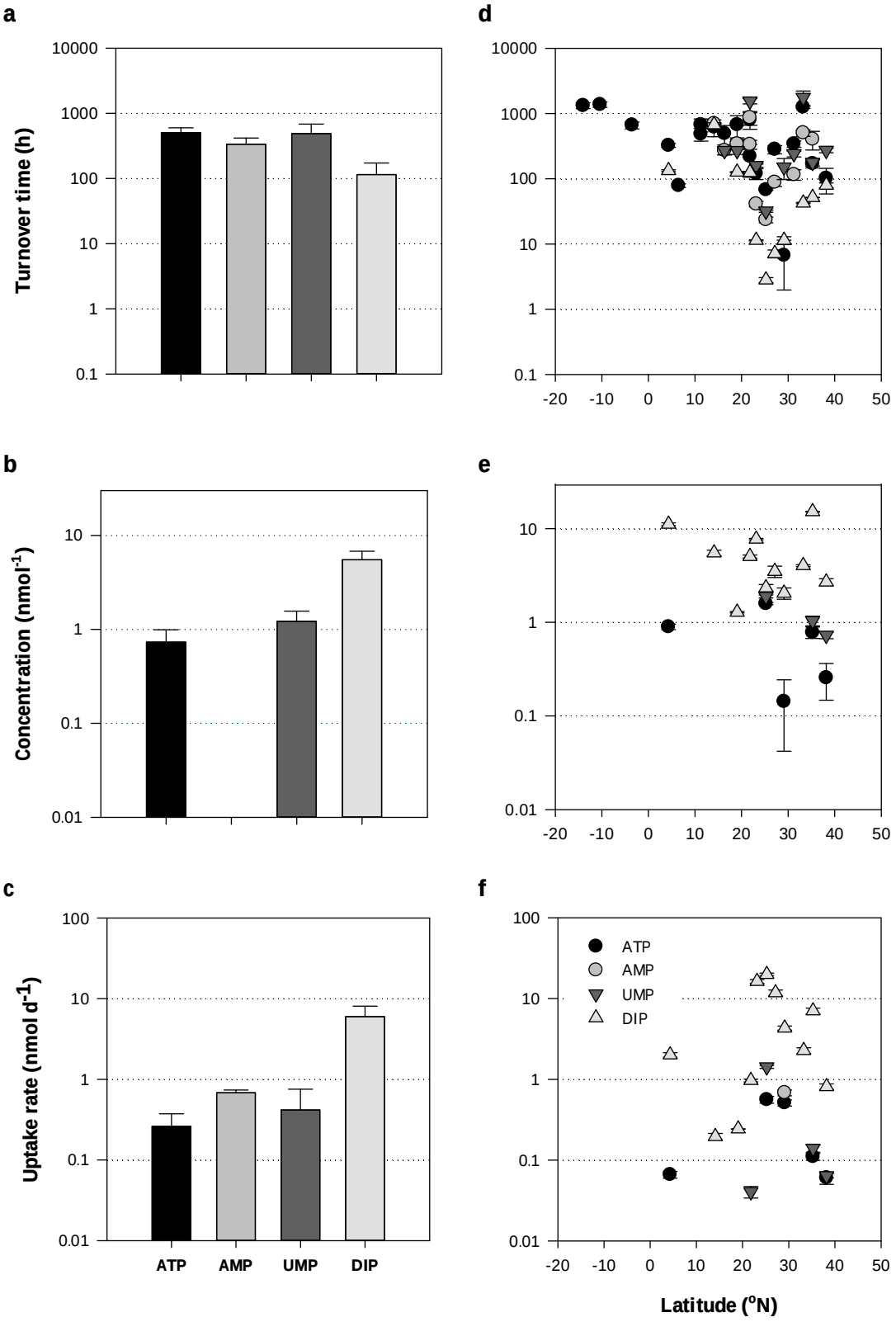


Figure 3.2. a-c Mean (a) turnover time, (b) concentration and (c) uptake rate of plankton for ATP, AMP, UMP and DIP in cluster region *a* (defined in Chapter 2), in autumn 2009. Figures 3.2 d-f Latitudinal distribution of (d) turnover time, (e) concentration and (f) uptake rate of plankton for ATP, AMP, UMP and DIP along the latitudinal AMT19 transect during autumn 2009. Standard error bars are shown. Note all *y*-axes (a-f) are logged and the symbol legend applies to all relevant plots.

The mean turnover times for samples (*n*) for DOP (ATP 504 ± 95 h, *n* = 20, AMP 336 ± 82 h, *n* = 11 and UMP 488 ± 196 h, *n* = 10, see Figure 3.2) were between 3 and 5 times longer than for DIP (115 ± 58 h, *n* = 11). The most variation in turnover time across the transect (coefficient of variation, calculated from data detailed in Appendix Table A.3) was found in DIP (1.7) and the least in AMP and ATP (0.81 and 0.84 respectively). The highest mean concentration and uptake rate measured were in DIP (respectively 5.5 ± 1.1 $nmol^{-1}$ and 6 ± 2.1 nmol d^{-1}). Tested DOP nucleotides' mean concentrations were all < 1.2 $nmol^{-1}$ and DOP mean uptake rates were all < 0.7 nmol d^{-1}.

3.3.1. DOP as an Alternative Source of Phosphate

DOP (ATP, AMP and UMP) measurements of turnover time, concentration and uptake rate were all significantly different to equivalent DIP measurements (Mann Whitney U tests, all $p < 0.05$).

By comparing the turnover times of DOP and DIP, it can be seen that DOP nucleotides (ATP and UMP) turnover time is significantly longer (often by an order of magnitude) than DIP's turnover time (Mann Whitney U tests, all $p < 0.03$ and see Figure 3.3). This is despite the mean concentration of all DOP fractions tested being lower than the mean DIP concentration (see Figure 3.2 a-b). If DOP was being exploited as an alternative phosphorus source one would expect a short turnover time in combination with a low concentration. Here, however, a long turnover time and a low concentration is seen. DOP turnover time is often an order of magnitude more than DIP turnover time (see Figure 3.2 a, d and Figure 3.3).

To support DOP turnover time being significantly longer than DIP turnover time, no significant correlation between concentration of DIP and uptake rate of DOP is observed (Spearman Correlation, all $p > 0.05$ and see Figure 3.6 in Discussion section). Thus, the evidence contradicts the hypothesis that DOP is being utilised as a significant alternative phosphate source (to DIP) by plankton.

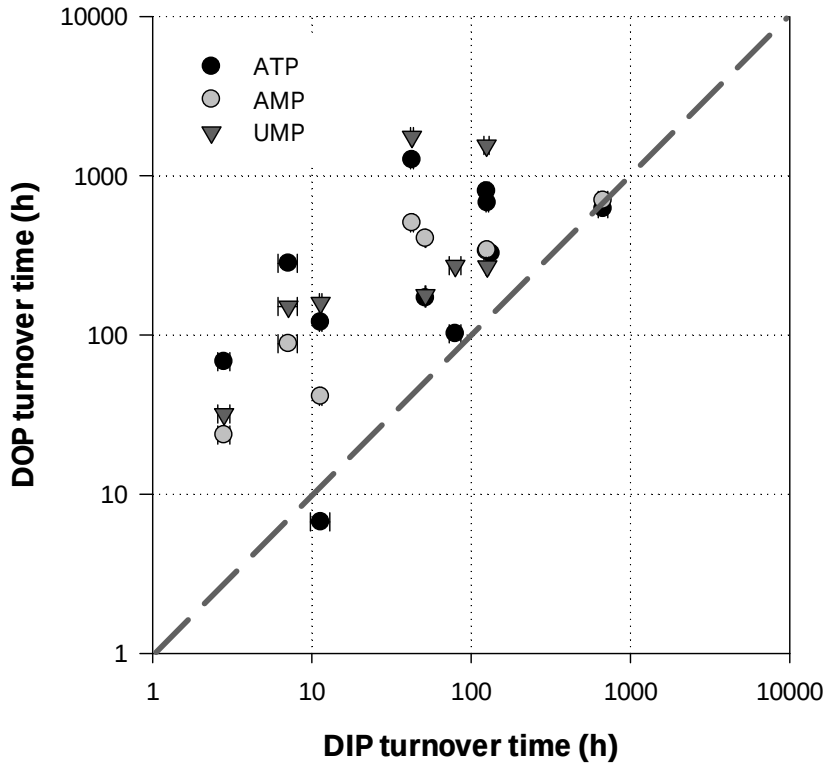


Figure 3.3. Comparison of DIP turnover time and DOP (ATP, AMP and UMP) turnover time. Bi-directional error bars show standard error of independent point measurements. The dashed line indicates the unity line. Note axes are logged.

3.3.2. Comparison of DOP Nucleotides

All comparisons between nucleotides were carried out at single stations and not between stations, therefore only concurrent stations were directly compared (see raw data in Appendix, Table A.3). For the three DOP nucleotides studied, turnover time (ATP - AMP, $p = 0.4$; ATP – UMP, $p = 0.52$ and AMP – UMP, $p = 0.86$), concentration (ATP – UMP, $p = 0.3$) and uptake rates (ATP – UMP, $p = 0.9$) are not significantly different (Mann Whitney U tests); as previously noted AMP concentration and uptake rate sample number was insufficient to run statistical analysis. Significant positive correlations (Spearman Correlations) were found between DOP nucleotide turnover times: ATP - UMP, $\rho = 0.8$; AMP - UMP, $\rho = 0.7$ (both $p < 0.04$). ATP - AMP was also positive ($\rho = 0.5$), although not significant ($p = 0.1$, graph not shown), see Figure 3.4 for significant correlations. These results therefore agree with the hypothesis that the three DOP nucleotides tested are not significantly different in utility as a substrate (as $p > 0.05$) and moreover indicate that a single nucleotide may be able to be used as a proxy for the others, due to the positive correlations.

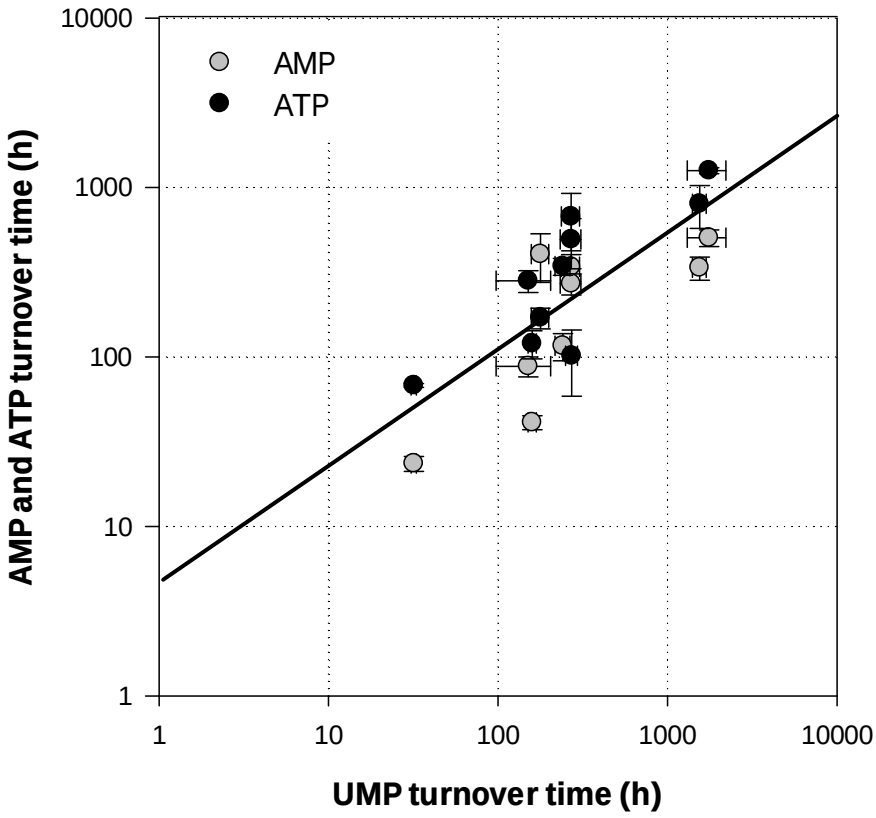


Figure 3.4. Comparison of UMP turnover time versus AMP and ATP turnover times. Error bars show standard error of independent point measurements. Turnover times of AMP and ATP are significantly positively correlated ($\rho = 0.7$ and $\rho = 0.8$ respectively, $p < 0.05$) to turnover time of UMP. ATP and AMP are also positively correlated ($\rho = 0.5$, $p = 0.1$, non-significant, data not shown). The black line indicates an independent linear regression of data in the plot ($r^2 = 0.58$, slope = 0.7 and intercept = 0.66). Note that both axes were logged.

3.3.3. ATP Inter-annual Variability and Seasonality

Examining the nucleotide ATP alone for three different years, the longest mean turnover time in ATP was observed on AMT19 (504 ± 93 h, *n* = 20) and the shortest on AMT18 (92 ± 14 h, *n* = 7). Both of these cruises took place in the autumn. The mean concentrations of ATP are between 0.7 ± 0.2 nmol^{-1} on AMT14 (*n* = 8) and 0.8 ± 0.1 nmol^{-1} on AMT19 (*n* = 5). Mean uptake rates were between 0.17 ± 0.04 nmol^{-1} on AMT14 (*n* = 8) and 0.26 ± 0.1 nmol d^{-1} on AMT19 (*n* = 5), see Figure 3.5 and raw data in Appendix, Table A.4. The highest coefficient of variation (calculated from data in Appendix Table A.4) was found in AMT19 turnover time and uptake rate (0.84 h and 0.96 nmol d^{-1} respectively), and the lowest in AMT14 turnover time (0.5 h).

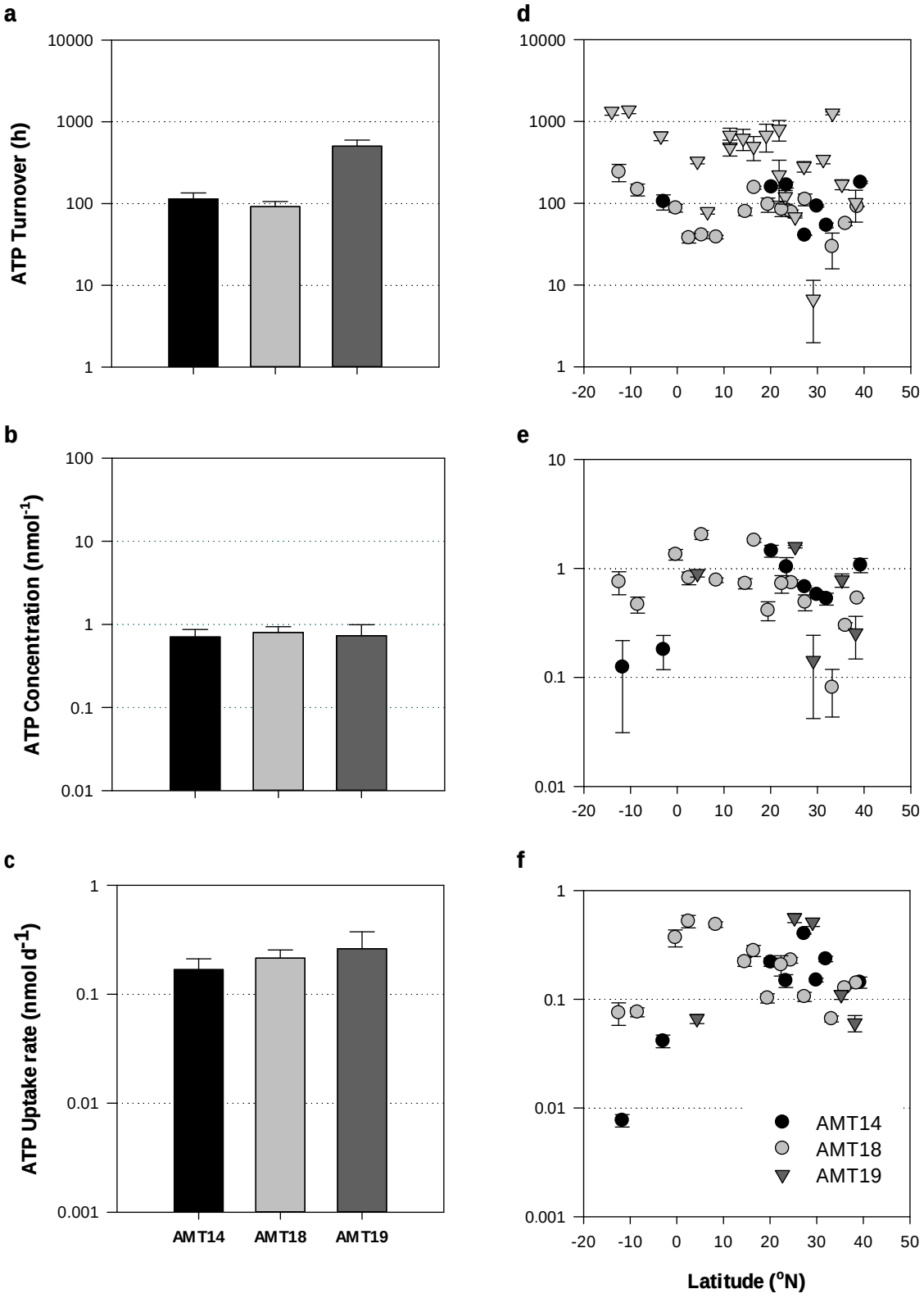


Figure 3.5. a-c Mean (a) turnover time, (b) concentration and (c) uptake rate for ATP in cluster region *a* (defined in Chapter 2), for AMT14 (spring 2004), AMT18 (autumn 2008) and AMT19 (autumn 2009). Figures d-f show latitudinal distribution of (d) turnover time, (e) concentration and (f) uptake rate along the same transects (see Figure 3.1). Standard errors bars are shown. Note that all *y*-axes (a-f) are logged and the symbol legend applies to all relevant plots.

None of ATP turnover, concentration or uptake were significantly different between any of the three years (ANOVA, $p = 0.09$, $p = 0.97$ and $p = 0.93$). ATP turnover, concentration or uptake between seasons (AMT14 versus AMT18 and AMT19), were not significantly different (Mann Whitney U test, $p > 0.05$), except for turnover time between AMT14 and AMT19 ($p = 0.01$), this was due to a large difference in mean and coefficient of variation values, which may be due to sampling error (also seen in the large coefficient of variation for this measurement). This also meant that there was no significant difference between the nucleotide labelling used ([γ^{33}P]ATP on AMT14 and [α^{33}P]ATP, on AMT18 and AMT19).

Furthermore, there was no significant difference in concentration or uptake rate between years where measurements were taken in the same season (autumn - AMT18 and AMT19), using the same nucleotide (Mann Whitney U test, $p > 0.05$). However there was a difference between AMT18 and AMT19 turnover time (Mann Whitney U test, $p < 0.05$). This again is attributed to the large difference in means between AMT18 and AMT19 turnover time and the larger coefficient of variation seen in AMT19, which is suggestive of sampling error. As this difference in turnover time was not seen between spring AMT14 and autumn AMT18, it is assumed that the data with the lowest variability is more accurate. Therefore the turnover time between seasons is not significantly different.

3.4. Discussion

3.4.1. DOP as an Alternative Source of Phosphate

This Chapter presents experimental evidence, collected from the oligotrophic subtropical and tropical north and central Atlantic Ocean, that DOP is not being readily exploited by plankton as a significant alternative to DIP. It should be noted, that the *in situ* measurements presented in this Chapter do not completely rule out DOP utilisation, nonetheless these measurements provide an indication of whether DOP is being exploited in significant amounts (or not) by plankton in the NAG. The mean turnover times of all DOP nucleotides (ATP 504 ± 95 h, AMP 336 ± 82 h and UMP 488 ± 196 h) are 3 to 5 times longer than mean DIP turnover time (115 ± 58 h), despite all DOP nucleotide concentrations being considerably lower (all < 1.2 nmol^{-1}, ATP was 0.73 ± 0.3 nmol^{-1}) than DIP concentrations (5.5 ± 1.3 nmol^{-1}). The majority of DOP turnover times are approximately an order of magnitude greater than those for DIP. Furthermore, the mean

DOP uptake rates in this study were all < 1 nmol d^{-1} (see Appendix Table A.3), over six-fold lower than the mean DIP uptake rate of 5.98 ± 2.1 nmol d^{-1} (n = 11). These measurements are in reasonable agreement with those published estimates of uptake and concentration of ATP and DIP in the same location in Spring 2004 (Zubkov et al. 2007) and illustrate that the DOP is not being significantly exploited for P.

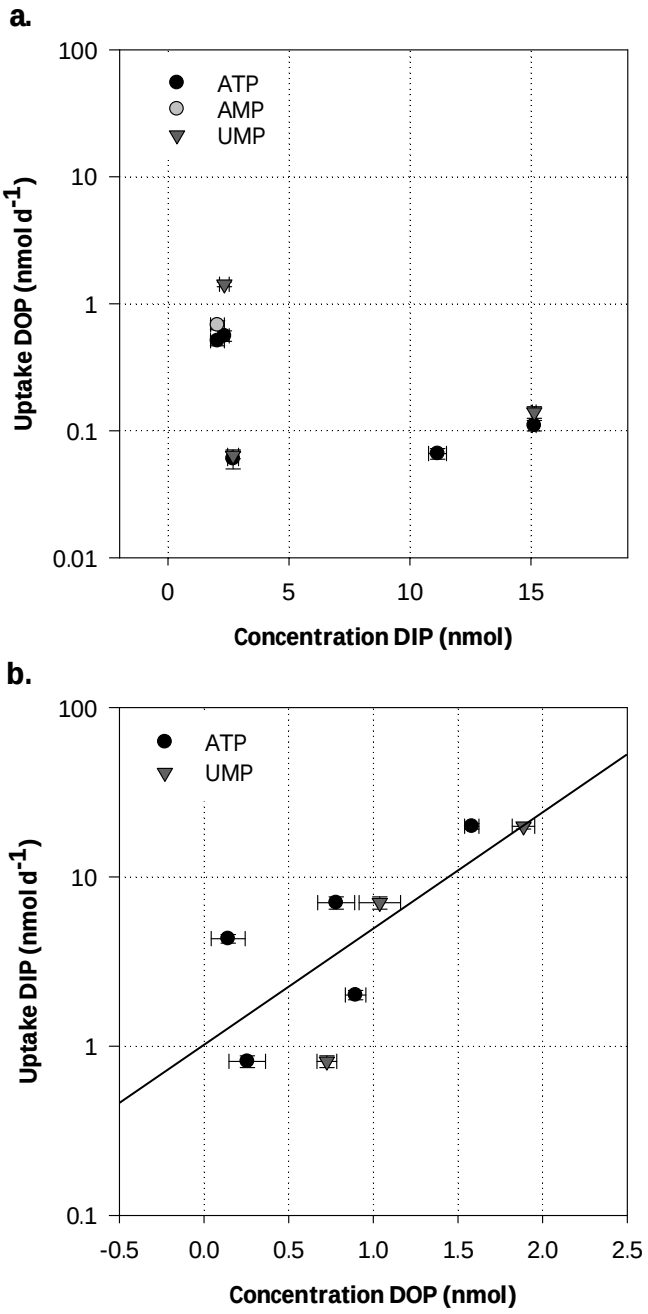


Figure 3.6. (a) Scatter plot comparisons of DIP concentration and DOP uptake rate (AMP, ATP and UMP). (b) Scatter plot comparisons of DOP concentration (ATP and UMP, no samples for AMP) and DIP uptake rate. Black line indicates a linear regression, with an r^2 value of 0.55. Error bars show standard error of independent point measurements. Note that only *y*-axis, uptake rate, is logged.

To help put DOP utilisation as a secondary source of phosphate to DIP into context, the uptake rate measurements can be examined in conjunction with concentrations. If DOP was being utilised consistently in significant (even small) amounts when DIP concentrations were low, at low concentrations of DIP, a high relative uptake of DOP would be expected to be seen (a negative correlation). Here, however, no relationship is present in the data (see Figure 3.6a). Conversely a positive correlation (although insignificant) was found between DOP concentration and DIP uptake. As DOP concentration increased, so did DIP uptake (Figure 3.6b). This increase in DOP concentration may be coincidental or due to the exudations of DOP from plankton (through messy feeding, excretion etc.), resultant from an increase in DIP uptake. Zubkov and Leakey (2009) for example found that in culture, one-third of phosphorus being consumed is remineralised.

Previously, different techniques have been used to examine phosphate concentration and planktonic uptake rate of phosphate in the central Atlantic. These include the standard colorimetric techniques which have reported large and varying concentrations of DOP (e.g. in autumn in the NAG, 210 ± 10 $nmol^{-1}$, Mather et al. 2008 and in the Sargasso Sea, 6.9 ± 2.2 $nmol^{-1}$, Lomas et al. 2010), higher than this study's autumn bioassay mean maximum ambient concentration (sum of ATP and UMP only as no AMP concentration data) of 2.1 ± 0.7 $nmol^{-1}$. These chemical techniques have also been used to infer enhanced DOP utilisation in the NAG, based upon alkaline phosphatase activities (APA, the enzyme that hydrolyses DOP, Kuenzler and Perras 1965), leading to an estimate that 12 - 30 % of autotrophic demand is being fulfilled by DOP (Mather et al. 2008). The accuracy and applicability of these measurements, however, can be questioned, as these chemically derived quantities include the entire DOP pool and not just the relevant bioavailable fraction. A strong gradient of decreasing DOP concentration from east to the central subtropical Atlantic (based on increasing APA estimates) has also been argued. It is thought that the DOP is produced in regions of high productivity (i.e. the Mauritanian upwelling) and then laterally transferred to the adjoining oligotrophic regions where it can be utilised (Reynolds et al. 2012). If these APA measurements are accurate, higher DOP utilisation in the centre of the gyre would be expected (as enzyme APA activity, that hydrolyses DOP, is higher).

ATP utilisation in this study has, however, been shown to be relatively and consistently low in all three years of this study (Figure 3.5 and raw data in Appendix Table A.4), as was found by Zubkov et al. (2007) in the region. The differences in estimates (between APA

and bioassay) are most likely due to the sensitivity of the techniques used and the assumptions underlying the calculations. Both Mather et al. (2008) and Lomas et al. (2010) calculated DOP by subtraction, as the difference between total dissolved phosphorus and DIP - measured respectively by ultraviolet photo-oxidation (Armstrong et al. 1966) and MAGIC-SRP method (the Soluble reactive phosphorus by magnesium induced co-precipitation protocol, see Karl and Tien, 1992, for details). However, these techniques are not considered sensitive enough for oligotrophic environments such as the NAG. For example, the MAGIC-SRP detection limit is reported as being above the oligotrophic concentrations reported here, of 10 - 15 $nmol^{-1}$ (Lomas et al. 2010). The methods presented in this Chapter were chosen because ambient concentrations are characteristically at or below the detection limit of these standard colorimetric methods. Dilution series radioactive bioassays more accurately measure concentrations at < 1 $nmol^{-1}$ than analytical chemical techniques and should be the method of choice for DOP studies in the NAG. Work published for the region using these chemical techniques therefore needs to be interpreted with care and future work should focus on those methods employed in this Chapter. In addition, it may be useful for further investigations to compare bioassay and alternative methods (APA and MAGIC-SRP) on *in situ* live samples to aid understanding of discrepancies between them.

Uptake of DIP alone is insufficient to satisfy the microbial physiological requirements for phosphate (Hartmann et al. 2011) and this study and that of Zubkov et al. (2007) have shown that DOP is not being readily exploited for phosphate. An alternative means of phosphate acquisition is the topic of this thesis – mixotrophy. However, it should also be noted that there are alternative methods that the microbial community may be employing to negotiate the problem of phosphate scarcity. First, symbionts could help to obtain phosphate (Cole 1982). However, these symbionts (most likely to be bacteria) would take up phosphate in a similar way to their host and therefore should be measurable with the method used here. Second, ATP may not be an adequate molecule to model other sources of phosphate, such as biogenic phosphonates which make up ~ 25% of DOP (Clark et al. 1998). However, the C-P bond is very strong and the ability to degrade it has only been found in prokaryotes (Adams et al. 2008). Phosphonates are therefore unlikely to contribute to eukaryote phosphate acquisition in large quantities. A third approach could be reducing cellular phosphorus demand by incorporation of non-phosphorus lipids (such as sulphur or nitrogen based lipids) into membranes by phytoplankton, both eukaryotes and cyanobacteria (Van Mooy et al. 2009). However, phosphorus demand still remains for other necessary cell components, such as DNA and RNA. Unfortunately, cell sorting was

not undertaken on our samples, so it is not possible to state explicitly which organisms dominate DOP uptake, to help ascertain if any were specifically using an alternative phosphate acquisition technique or source. Bacterioplankton, however, have been shown to play the dominant role (60 %) in DIP acquisition in comparison to picoplanktonic algae (0.3 %) (Zubkov et al. 2007, Hartmann et al. 2011). This is because prokaryotes are more efficient than protists at acquiring nutrients due to their higher cell surface-area-to-volume ratio. Therefore it is reasonable to expect that the majority of DOP being taken up in this study, is by the bacterioplankton. Further cell sorting experimentation would be required to confirm this. A low uptake of phosphate (both DOP and DIP) by eukaryotes is further evidence for the existence of an alternative means of phosphorus acquisition, via the predation of phosphorus-rich bacterial cells (Zubkov and Tarran 2008, Hartmann et al. 2012), i.e. mixotrophy - the overarching subject of this thesis.

3.4.2. Comparison of DOP Nucleotides

Most previous studies on DOP using the bioassay technique have concentrated on ATP as a model compound (e.g. Zubkov et al. 2007, Casey et al. 2009, Bjorkman et al. 2012). Here, two additional DOP nucleotides have been investigated *in situ*. All three show low utilisation by plankton in comparison to DIP. Furthermore, this study shows that UMP turnover time is significantly positively correlated with both ATP and AMP turnover time (see Figure 3.4). Additionally ATP and AMP are positively correlated, if not significantly. This suggests that individual nucleotides may be used as an uptake proxy for others (e.g. UMP for ATP). This is perhaps unsurprising, as the chemical structure of the three nucleotides in this study are similar. Each comprises a ribose sugar, a purine base (uracil, UMP or adenine, AMP and ATP) and one (UMP and AMP) or three (ATP) phosphate groups. Consequently the transport mechanism necessary for each nucleotide is comparable (Bengis-Garber and Kushner 1982, Bengis-Garber 1983).

The low utilisation and bioavailable concentration of the three nucleotides tested can also be considered in terms of percentages. The mean uptake rate of ATP calculated as a percentage of all the nucleotide bioavailable uptake measured (DIP, AMP and UMP, AMP data was insufficient to include) was 9 ± 3 %. To compare to previous studies, that have only measured ATP, percentage bioavailable uptake can be calculated without the additional nucleotides investigated in this Chapter (i.e. just DIP plus ATP), as such ATP uptake as a percentage of P uptake was 5 ± 2 %. These results are consistent with the

previous study in the region that estimated spring ATP uptake rate as 13 ± 6 % of the P pool (Zubkov et al. 2007). Despite this relatively low specific uptake rate, culture experiments on flagellate taxa have shown that growth can be sustained on ATP, AMP or UMP alone (Wang et al. 2011). However, these were bloom-causative phytoplankton from Chinese coastal waters and therefore the results are not necessarily applicable to open ocean oligotrophic waters. The findings of this Chapter therefore support previous conclusions that DOP (ATP and UMP) plays a secondary role in phosphate microbial dynamics (Zubkov et al. 2007, Casey et al. 2009, Michelou et al. 2011) in the subtropical and tropical Atlantic ocean.

3.4.3. ATP Inter-annual Variability and Seasonality

No significant difference was found between years or seasons for ATP concentration or uptake rate measured on the three AMT transects. This is despite AMT14 measurements being made in austral spring (Julian day -132 to 149) and AMT18 and AMT19 (Julian day – 262 to 300 and 288 to 317 respectively) in austral autumn in this study's region ($p > 0.05$). Mather et al. (2008), however, reported seasonality in DOP uptake, as APA activities were significantly greater (three times higher) in the NAG in June than in November (*t*-test, $p < 0.05$). The data presented here, estimates DOP uptake by a more sensitive procedure (bioassay), and does not support these findings.

Coincidentally by looking at inter-annual and seasonal differences of ATP utilisation, the use of different labelling of the ATP nucleotides could be addressed. On AMT14, ATP was labelled on the [γ^{33}P]ATP and on AMT18 (and AMT19) ATP was labelled on the [α^{33}P]ATP. In seawater (pH 7.9 - 8) non biological hydrolysis of P groups is believed to occur quickly (< 1 minute), however [γ^{33}P]ATP is thought to be hydrolysed slower than [α^{33}P]ATP (Bengis-Garber and Kushner 1982, Bengis-Garber 1983). To avoid underestimation of ATP uptake, α^{33}P (the phosphate group closest to the adenine base) was specifically used on AMT18 (and AMT19). Although no significant difference between the α-labelled (AMT18) and the γ-labelled (AMT14) turnover time, concentration and uptake rate is observed ($p > 0.05$), and no significant difference between AMT14 and 19 concentration and uptake rate is found. This was not unexpected, as the shortest time interval in this experiment reported in this Chapter is 10 minutes. In other words, there is no difference in uptake for the different terminal phosphate groups of ATP as the external (that furthest from the

purine base) [γ^{33}P]ATP appears to be cleaved at the same rate as the internal (that closest to the purine base) phosphate group [α^{33}P]ATP. Simultaneous station replicates were lacking however and would be required to test this further.

3.5. Summary and Implications

The ability of ultraplankton to utilise DOP as an alternative phosphate source to DIP would be an important ecological strategy for survival and competition. There is some evidence for it occurring in the Sargasso Sea (e.g. Casey et al. 2009, Michelou et al. 2011), but this Chapter (and work from Zubkov et al. 2007) indicates that it is not an important process in the oligotrophic subtropical and tropical, north and central Atlantic ocean. Therefore the plankton community must be adapted in other ways to living in this P depleted environment (Hartmann et al. 2011). The results of this Chapter support evidence for the role of mixotrophy. Therefore it can be concluded that it is not necessary to include DOP within a model representing the oligotrophic Atlantic.

From the results presented in this chapter and considering the four stated hypothesis to be tested the following can be concluded:

- This Chapter provides supporting evidence to Zubkov et al. (2007) that DOP is not significantly utilised as an alternative source of P by plankton in the subtropical and tropical north Atlantic.
- Turnover time, concentration and uptake by plankton of the three DOP nucleotides measured (ATP, AMP and UMP) were not significantly different.
- There was no significant difference between turnover time, concentration or uptake for the three years studied. There was also no significant difference between turnover time, concentration or uptake rate across two (AMT14 and AMT18) seasons. Therefore inter-annual variability and seasonality is not present in this dataset.
- DOP does not need to be included in an ultraplankton model of the oligotrophic Atlantic as an important alternative source of P.

4. Mixotrophy: Is a Simple Steady State Data Driven Model able to describe the Mixotrophic Ecosystem?

4.1. Introduction

4.1.1. Motivation for Approach to Modelling Mixotrophy

A modelling approach is useful to aid understanding of the broader implications of mixotrophy in the north Atlantic subtropical gyre, as direct observations of mixotrophic uptake are difficult to obtain; models are a useful tool to investigate the fluxes through the ecosystem. Simple ecosystem models such as NPZD (Nutrient-Phytoplankton-Zooplankton-Detritus) models (e.g. Fasham et al. 1990) segregate planktonic groups by well-defined trophic level. Within this structure, it is not clear where mixotrophs should be placed, as they straddle traditional trophic levels. Previous models incorporating mixotrophy have been largely theoretical explorations of whether coexistence of mixed populations of mixotrophs, heterotrophs and/or phototrophs is possible or of the effect of different levels of mixotrophy on fluxes through an ecosystem (see review in Chapter 1, Section 1.3.2). In addition, parameterisation is strongly limited by the lack of available data. At present quantitative observations of relevant processes and internal elemental composition in mixotrophs are insufficient for a data-driven physiological mechanistic-stoichiometric model to be configured for practical use.

The majority of previous models have been zero-dimensional, with a single common inorganic nutrient (e.g. Thingstad et al. 1996; Hammer and Pitchford 2005) or

concentrating on the modelling of a particular mixotrophic species (e.g. Hood et al. 2006b). However, as qualitative advances in microbiology have been made, complexity has also increased in theoretical physiological mixotroph models. The paper 'Building the "perfect beast": modelling mixotrophic plankton' by Flynn and Mitra (2009) argued that a physiological mechanistic stoichiometric model is necessary to model the possible types of mixotroph (i.e. type I, II or III and the gradients between, see Chapter 1 for mixotroph explanation). Quantitative data, however, is currently insufficient for parameterisation of such relatively complex models. No models thus far have been developed using an alternative approach: directly using *in situ* measurements. This Chapter attempts this for the first time, presenting a model that is simple yet data-driven.

4.1.2. Steady state

> "Make things as simple as possible, but not simpler" (Einstein 1934).
>
> "Plurality should not be posited without necessity" - Ockham's razor or the law of parsimony (attributed to William of Ockham, 14th-century, Britannica, 2012)

The present modelling approach is motivated by the above principles, as a simple model works best when variables, parameters and sensitivities of parameters are not fully understood (Anderson 2005). Specifically, the minimum number of fields and fluxes will be modelled, whilst still capturing the key components and processes, and the simplest context will be assumed valid - a zero-dimensional steady state system. A zero dimensional model is a mixed layer model which has two boxes: an upper layer assumed to be homogenously mixed and biologically active, and a lower layer with a constant nutrient concentration which is biologically inactive. This is an approximation of the system, admittedly not strongly supported by data, as ultraplankton are known to reside below the mixed layer (Tarran et al. 2006). However, this first order approach has a robust history, being one of the first approaches to plankton ecosystem modelling (Riley 1946) and it has been used consistently and successfully for over 60 years (e.g. Evans and Parslow 1985, Fasham et al. 1990, Steele and Henderson 1992) for modelling investigations of marine ecosystems.

The assumption of steady-state is also one with some history (e.g. Taylor and Joint 1990, Stickney et al. 2000, Anderson and Ducklow 2001, Anderson and Turley 2003). Steady state is further supported to some extent by observations of the region. In the Atlantic

gyres a pronounced seasonal cycle in chlorophyll is not displayed in satellite observations (ocean colour is revisited in Chapter 6), and primary production is relatively constant all year round (Lutz et al. 2007, Cole et al. 2012). In addition, in this thesis evidence has already been presented for consistent planktonic uptake rates between seasons (see Chapter 3). This minimal seasonality is echoed by the variability of the mixed layer depth (MLD). The changes between summer and winter MLD in high latitudes (~ 50 - 60°N) are tenfold, whereas in the north Atlantic subtropical gyre (~ 20 - 40°N) they are only two-fold (Monterey and Levitus 1997). As a starting point it is therefore reasonable to approximate the system as being in steady state.

The assumption of an approximate state of equilibrium (steady state) enables network analysis techniques (Wulff et al. 1989) to be applied to the available observational data, allowing algebraic solutions to be used to estimate unknown fluxes from known fluxes. This Chapter seeks to constrain the simplest model including mixotrophy based on the available direct observations.

This Chapter addresses the following hypothesis:

- It is possible to construct a simple steady state model incorporating mixotrophy from *in situ* observational data from the subtropical north Atlantic.

4.2. Observational Data

Previously collected observational data will be used to build the model. Ultraplankton samples were taken from three CTD casts in the subtropical north Atlantic gyre by M Zubkov on AMT17 along a southbound transect from Govan, Scotland to Port Elizabeth, South Africa from 15th October to 28th November 2005 aboard the *RRS* Discovery in the boreal autumn and austral spring (see Figure 4.1 and Table 4.1). All CTD casts were located within the same biogeochemical region (> 90 % similarity) as defined by ultraplankton distribution analysis in Chapter 2.

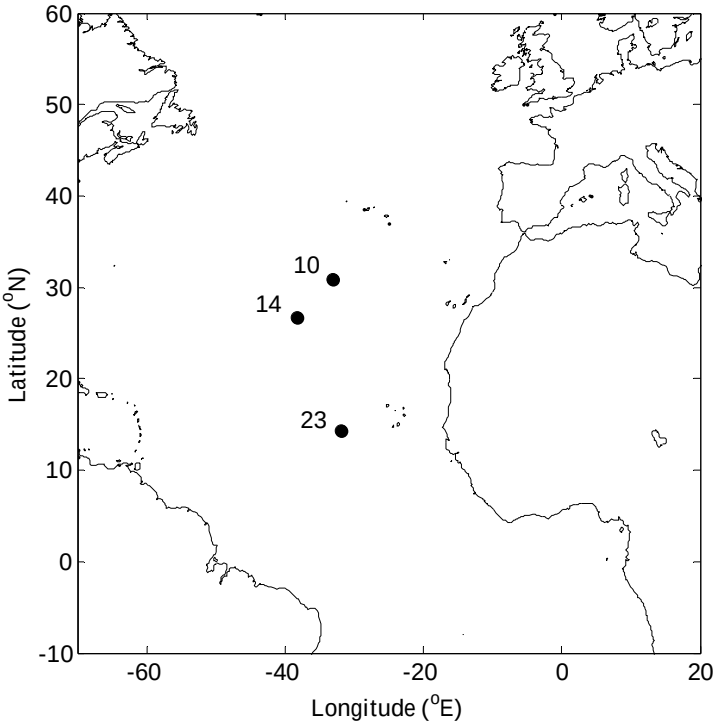


Figure 4.1. CTD station positions for cell number and uptake rate measurements, taken on AMT17, October - November 2005. See Table 4.1 for accompanying metadata.

The following ultraplankton groups were enumerated and their inorganic phosphate uptake rates measured: total bacteria (B, including *Prochlorococcus* spp., *Synechococcus* spp. and low nucleic acid heterotrophic bacteria), picoeukaryotes (following Zubkov and Tarran, 2008, attributed as mixotrophs, M) and heterotrophic flagellates (grazers, G). For data see Table 4.2. The three groupings (B, M and G) also reflect the model structure introduced in the following Section.

Table 4.1. CTD number, station position, Julian day (JD) and mixed layer depth (MLD) from which data were collected from 2 m depth *in situ* at 1100 h local time by M Zubkov (see accompanying Figure 4.1) on AMT17, October - November 2005.

CTD	Latitude (°N)	Longitude (°E)	JD	MLD (m)
10	30.85	-33.11	303	58.5
14	26.70	-38.23	305	58.5
23	14.26	-31.87	309	48.5

Concentrations of bioavailable phosphate and microbial phosphate uptake rates were estimated using a concentration series bioassay of isotopically labelled [^{33}P]orthophosphate following the methodology detailed in Zubkov et al. (2007). This combined with enumeration of groups with flow cytometry allowed specific microbial group uptake rates to be determined (see Tables 4.2 and 4.3). This method measures the

total phosphate taken up by a microbial group, without distinguishing where phosphate is taken up from (i.e. if it is of phagotrophic or autotrophic origin). MLD was measured at each CTD's location (Table 4.1).

Table 4.2. Data collected on AMT17, by M Zubkov (unpublished data) at the CTD stations (detailed in Table 4.1). Phosphate uptake rate (amol P cell^{-1} h^{-1}) for each microbial variable. SE is experimental error in amol P cell^{-1} h^{-1}. n.d, no data.

CTD	Bacteria (B) (amol P cell^{-1} h^{-1})		Mixotrophs (M) (amol P cell^{-1} h^{-1})		Grazers (G) (amol P cell^{-1} h^{-1})	
	Uptake	SE	Uptake	SE	Uptake	SE
10	0.75	± 0.04	0.38	± 0.02	0.40	± 0.12
14	3.04	± 0.28	1.36	± 0.26	3.02	n.d.
23	1.14	± 0.11	1.05	± 0.09	1.07	± 0.03

Table 4.3. Data collected on AMT17, by M Zubkov (unpublished data) at the CTD stations, detailed in Table 4.1. Bioavailable phosphate (P) concentration (Conc. nmol P l^{-1}) and cell count (Abund. cell ml^{-1}) for each model variable measured. SE for P is experimental error, SE for B, M and G is assumed to be 5 % (see Chapter 5, Section 5.2.3.4 for explanation). *Note that for CTD10 Grazer (G) abundance, the number of cells were not counted. It was estimated by dividing total uptake on CTD10 by mean specific uptake from CTD14 and 23.

CTD	P (nmol P l^{-1})		B (cell ml^{-1})		M (cell ml^{-1})		G (cell ml^{-1})	
	Concn.	SE	Abund.	SE	Abund.	SE	Abund.	SE
10	1.45	± 0.38	360387	± 18019	1343	± 67	83*	± 4
14	2.96	± 0.52	344423	± 17221	11942	± 597	316	± 16
23	4.25	± 0.35	492229	± 24611	2748	± 137	333	± 17

4.3. Theoretical Mixotroph Model Design

A four box simple steady state model was designed with the following variables – phosphate (P), bacteria (B), mixotrophs (M) and grazers (G), all modelled explicitly.

The reasoning behind this structure is as follows. The model was constructed using phosphorus (P) as its currency for several reasons. Within the gyres, P is understood to be co-limiting with nitrogen (Wu et al. 2000, Mather et al. 2008). P is also less complex than nitrogen to model. There is evidence to suggest it is not necessary to model the organic form (see this thesis Chapter 3), and there is only one inorganic form, unlike inorganic nitrogen which can be present in a number of compounds, such as ammonium and nitrate. The B variable in the model comprises phototrophic (PB) and heterotrophic bacteria (HB).

Both were included in B to ensure total bacterivory (mixotrophy) by M and G was captured and because preferences for either HB or PB by M and G are unknown. Here it should be noted that separate data (uptake rates and abundances) were available for PB and HB. However when these were modelled as separate variables the model could not be solved explicitly analytically as there were too few known fluxes to deduce unknown fluxes. Thus a simpler 4-box model was employed. This 4-box model assumption is appropriate as light is not limiting to ultraplankton within the oligotrophic region (Pers. Comm. M. Zubkov). Therefore both HB and PB can be treated the same within the model, as they both take up inorganic P to make new cells and are then both consumed by M and G. The G variable was the only obligate predator within the model. The number of grazers within the oligotrophic environment is low and the G variable encapsulated the entire grazer community, with the ability to feed on a range of sized organisms from B to M. There is no explicit obligate phytoplankton variable within the model, with the M variable representing the oligotrophic picoeukaryotic algae community which has the ability to take up nutrients and/or undertake bacterivory. For simplicity there is only one source ($\Delta(Po - P)$) and one sink (ΦG) in the model at steady state. The source represents remineralisation and the turbulent mixing (Δ) in of deep P (Po) into the mixed layer. The sink is the loss from the mixed layer, which can be due to excretion, messy feeding or death. There is no mixing of B, G and M for simplicity. Such fluxes are assumed small relative to $\Delta(Po - P)$ and if included would preclude an algebraic solution. This is revisited in Chapter 5. The mixing of individual variables due to changes in the MLD is excluded here given that the assumption of steady state precludes an annual cycle in MLD.

The simplest possible parameterisations were used to describe the fluxes in the system - first-order, non-linear Lotka-Volterra interactions, these assume implicitly that the fluxes never saturate. The flux increases linearly as either the source or the target increase (i.e. αPB in Figure 4.2, would increase as P, the source, increases and/or as B, the target, increases). This approach means that additional (unconstrained) parameters are not required, such as would be the case if Michaelis-Menten kinetics was applied. Previous mixotroph studies have also used this simple method (e.g. Hammer and Pitchford 2005).

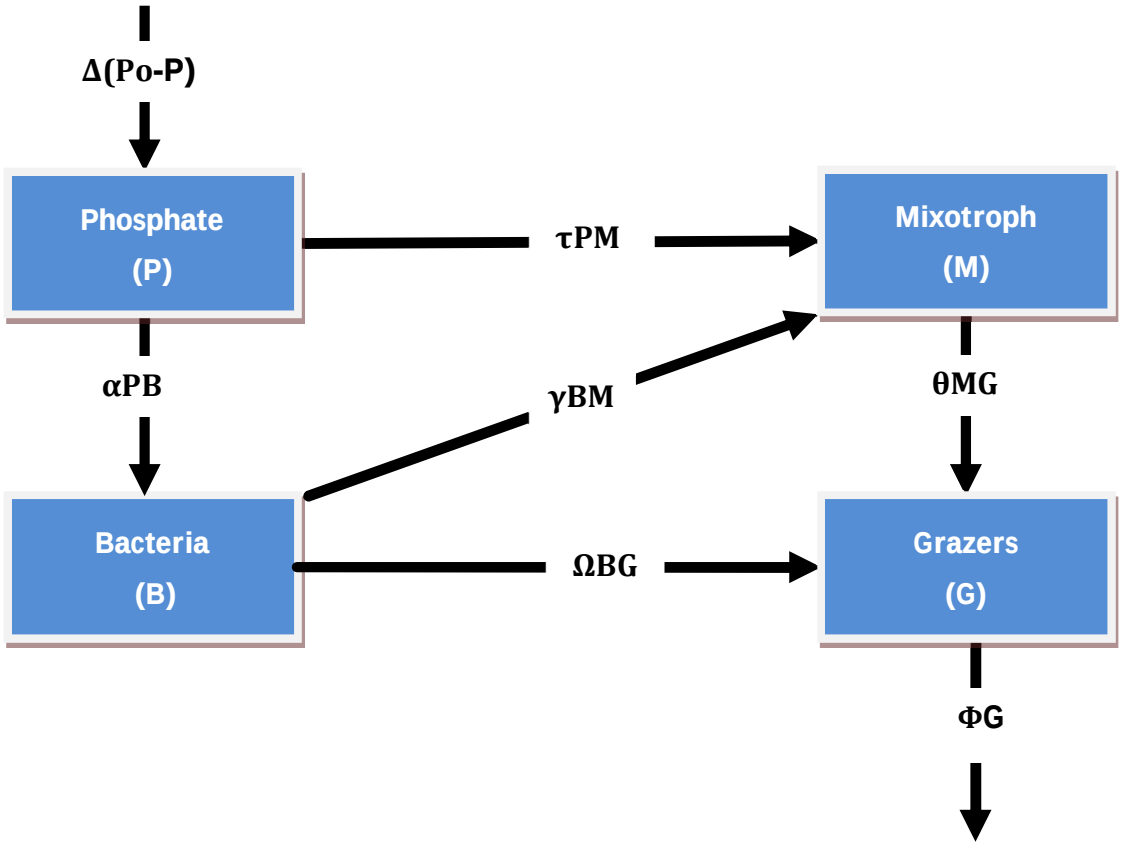


Figure 4.2. A schematic of the simple steady state model incorporating mixotrophy (see Equations 4.1 to 4.4 in the text).

The structure of the steady state mixotroph model can be described as follows (Figure 4.2 and Equations 4.1 - 4.4). Bacteria are modelled as the only variable reproducing using inorganic nutrients alone and limited by P (αPB). Mixotroph growth via their two modes of nutrition is limited by both prey (γBM) abundance and P (τPM) respectively, and bacteria and mixotrophs are both grazed on by the grazers (ΩBG and θMG respectively). As there is no detritus variable, messy feeding and excretion by the grazers are simplified into the loss term (ϕG). As larger predators are infrequent in the oligotrophic gyre, consumption of G by higher predators is taken to be negligible and so this loss flux can be used as a proxy for export. This is a suitable point to note that the flows within this system are net. Therefore although the arrows have been place in Figure 4.2 in the ecological perceived correct net direction (i.e. bacteria are predated upon by grazers, therefore the P flows predominantly from B to G), some P may also be flowing in the opposite direction (e.g. through B acting as detritivores, P could flow from G to B). The implication of this will be discussed further in Section 4.4.1. The source of P into the system is through the influx of deep phosphate (*Po*) by mixing across the base of the mixed layer (Δ). For simplicity dilution of other variables by mixing is assumed to be negligible.

At steady state, taking each variable (P, G, B and M) in turn, the fluxes into and out of each variable must sum to zero, such that:

Equation 4.1 $$0 = \Delta(Po - P) - \alpha PB - \tau PM$$

Equation 4.2 $$0 = \theta MG + \Omega BG - \Phi G$$

Equation 4.3 $$0 = \alpha PB - \gamma BM - \Omega BG$$

Equation 4.4 $$0 = \tau PM + \gamma BM - \theta MG$$

4.4. Mixotroph Model Construction

4.4.1. Flux Observations

Figure 4.3 shows the theoretical model schematic modified to demonstrate which fluxes have *in situ* observations (known measured fluxes F1, F2 and F3, solid lines, and unknown fluxes to be modelled - dashed lines).

The Equations relating modelled to observed fluxes are in summary:

Equation 4.5 $$F1 = \gamma BM + \tau PM$$

Equation 4.6 $$F2 = \alpha PB$$

Equation 4.7 $$F3 = \theta MG + \Omega BG$$

Measured from observations, flux 1 (F1, Equation 4.5) is the total amount of phosphate going into the mixotrophs (M), coming from both bacteria (γBM) and phosphate (τPM). As stated in Section 4.2, due to methodological constraints, the source of phosphate for mixotrophs cannot be distinguished. This is also true for measured flux F3. F2, is the uptake of phosphate by bacteria and is the only individual flux in the present model (αPB, Equation 4.6) that is directly constrained by the observational measurements. Finally, flux 3 (F3, Equation 4.7) is the measured net flux of phosphate taken up by the grazers, through predation on bacteria (ΩBG) and mixotrophs (θBG).

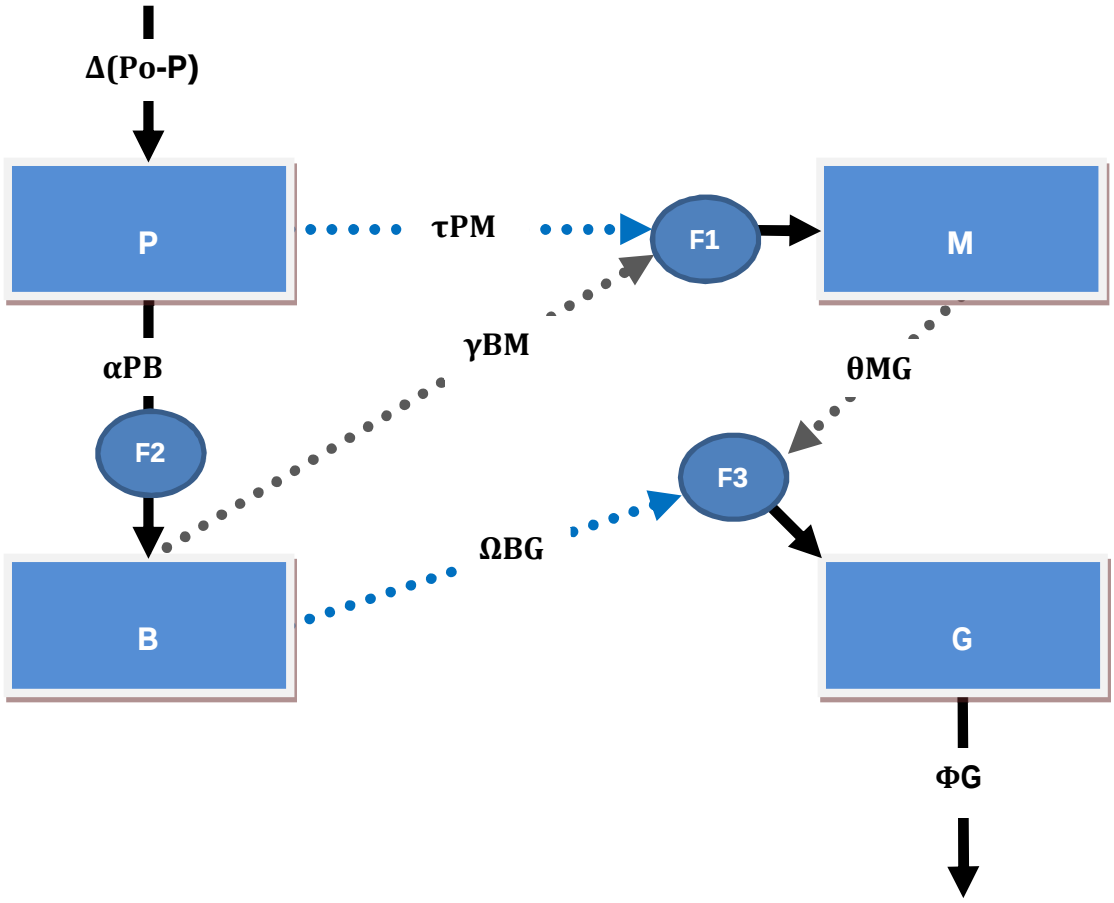


Figure 4.3. Summary of *in situ* measurements available for the steady state mixotroph model. Black solid lines and blue filled circles indicate known net fluxes (F1, F2 and F3) and dashed lines, unknown fluxes (see Equations 4.5 to 4.15 in the text for explanation).

Measured uptake (the observed flux) taken up by the variables, M, B and G (F1, F2 and F3 respectively, µmol P l^{-1} d^{-1}) are calculated from the *in situ* measurements for abundance (cell ml^{-1}) and uptake rate (amol P $cell^{-1}$ h^{-1}), according to Equation 4.8 (see Tables 4.2 and 4.3). It was necessary to obtain the uptake rate per volume (of P, G, B and M), rather than per cell, to match model units.

Equation 4.8 $\boldsymbol{uptake\ per\ cell\ x\ cell\ abundance = uptake\ per\ unit\ volume}$

Table 4.4 details the uptake observations for each CTD. The relative sizes of the observed uptake are consistent across the CTD's within their respective maximum possible uncertainty (calculated using the maximum and minimum error of both uptake per cell and abundance). F2 is consistently bigger, by at least an order of magnitude, than both F1 and F3 (e.g. CTD14, F2 = 2.5 x 10^{-2} ± 7.1 x 10^{-3} µmol P l^{-1} d^{-1}, F1 = 3.9 x 10^{-4} ± 1.9 x 10^{-4} µmol P l^{-1} d^{-1} and F3 = 2.3 x 10^{-5} ± n.d µmol P l^{-1} d^{-1}). F3 is always smaller (again by an order of

magnitude) than both F1 and F2. The impact on the relative modelled fluxes is detailed in Table 4.4.

Table 4.4. Steady state mixotroph four box model observed fluxes for each CTD, calculated using Equation 4.8 and data in Tables 4.2 and 4.3. All units µmol P l^{-1} d^{-1}. For abbreviations see the text. Maximum uncertainty (±), as correlation between errors ignored. n.d. is no data.

CTD	F1	±	F2	±	F3	±
10	1.2 x 10^{-5}	2.7 x 10^{-6}	6.5 x 10^{-3}	1.3 x 10^{-3}	8.1 x 10^{-7}	5.8 x 10^{-7}
14	3.9 x 10^{-4}	1.9 x 10^{-4}	2.5 x 10^{-2}	7.1 x 10^{-3}	2.3 x 10^{-5}	n.d.
23	7.0 x 10^{-5}	1.9 x 10^{-5}	1.4 x 10^{-2}	3.9 x 10^{-3}	8.5 x 10^{-6}	1.3 x 10^{-6}

4.4.2. Solving the Model

As the model is assumed to be in steady state (input balances output), the known observational net fluxes (F1, F2, F3, see Figure 4.2 and Equations 4.5 to 4.7) can be used to calculate the model's unknown fluxes (Equation 4.9 to 4.15). It is possible to explicitly solve the model (as sink balances source and because there are as many Equations as unknown fluxes) given the observations.

So that:

Equation 4.9 $$\alpha PB = F2$$

Equation 4.10 $$\Phi G = F3$$

Equation 4.11 $$\Delta(Po - P) = F3$$

Equation 4.12 $$\tau PM = F3 - F2$$

Equation 4.13 $$\theta MG = F1$$

Equation 4.14 $$\Omega BG = F3 - F1$$

Equation 4.15 $$\gamma BM = F2 - F3 + F1$$

Table 4.5. Model fluxes for each CTD, calculated from observations F1, F2 and F3. Calculated using Equation 4.9-4.15. All units µmol P l^{-1} d^{-1}. For abbreviations see the text. Negative modelled fluxes are highlighted in bold.

Modelled flux	CTD10	CTD14	CTD23
αPB	6.46 x 10^{-3}	2.51 x 10^{-2}	1.35 x 10^{-2}
τPM	**-6.46 x 10^{-3}**	**-2.51 x 10^{-2}**	**-1.35 x 10^{-2}**
γBM	6.47 x 10^{-3}	2.55 x 10^{-2}	1.36 x 10^{-2}
ΩBG	**-1.14 x 10^{-5}**	**-3.66 x 10^{-4}**	**-6.1 x 10^{-5}**
θMG	1.22 x 10^{-5}	3.89 x 10^{-4}	6.96 x 10^{-5}
ΦG	8.1 x 10^{-7}	2.29 x 10^{-5}	8.52 x 10^{-6}
Δ(Po-P)	8.1 x 10^{-7}	2.29 x 10^{-5}	8.52 x 10^{-6}

The network analysis shows that the model can only be solved if two of the fluxes are negative (e.g. CTD14 ΩBG = -3.7 x 10^{-4} µmol P l^{-1} d^{-1} and τPM = -2.5 x 10^{-2} µmol P l^{-1} d^{-1}, see Table 4.5), i.e. in the opposite direction to that expected. First, for uptake rate of phosphate by mixotrophs (τPM, Equation 4.12), according to the model, the excretion *to* the phosphate pool was therefore more than the uptake by mixotrophs *from* the phosphate pool. Second, phosphate obtained through predation of bacteria by grazers (ΩBG, Equation 4.14), indicated that grazers were excreting more to be remineralised by the bacteria pool than they were gaining through bacterivory (see Figure 4.3). Prior to model development, the consequences of the observed fluxes (F1, F2 and F3, Table 4.4) on the modelled fluxes (Table 4.5) were unknown.

4.5. Model Performance

Ockham's razor or the law of parsimony was stated at the beginning of this Chapter. This principle of simplicity guided by the available measurements was the motivation for this mixotroph steady state model construction. Unfortunately this approach was not successful. The simple steady state model was constructed using *in situ* observational data but two of the fluxes were found to be in the opposite direction to that expected ecologically – the phosphate uptake by mixotrophs (τPM) and the grazing of bacteria by grazers (ΩBG).

First, if the flux from phosphate to mixotrophs was negative (τPM), mixotrophs would be exuding more inorganic phosphate than they were taking up. This is inefficient and such waste is very unlikely to occur under these oligotrophic conditions. Second, the negative ΩBG flux in the model indicates that the recycling of phosphate from grazers to bacteria is higher than the phosphate gained through the predation of bacteria by the obligate

grazers, a sign that the bacteria pool may be dominated by a community of detritivores. This model result, however, contradicts previous culture work results of Zubkov and Leakey (2009), that found that approximately one third of phosphate is remineralised (by bacteria) from grazing. More than this is unfeasible in the oligotrophic system as once again it is very inefficient.

The outcome of the ecologically incorrect flux directions could be an anomalous result that has happened by chance. However, despite having so few data points here it can be argued that this is unlikely, as the relative sizes of the three fluxes are consistent between CTD's, even when accounting for the maximum uncertainty (see Tables 4.3 and 4.4). The reversal of the perceived ecologically correct directions for these fluxes is robust and shows therefore that this simple steady state model is inappropriate to describe the system and suggests that a more complex model is required to describe it accurately. This could include a departure from steady state and/or additional fluxes. The observations, however, can still be used to constrain the model. This next stage of model development (Figure 5.1) is discussed in Chapter 5.

4.6. Summary and Implications

A simple four box steady state model incorporating mixotrophy was constructed for the first time from *in situ* measurements from the subtropical north Atlantic. Two of the net fluxes in the model (mixotrophs to phosphate, and bacteria to grazers) are found to be in the ecologically perceived incorrect direction. To rectify the incorrect directionality in these fluxes, it is necessary to construct a more complex model (in Chapter 5).

From the results presented in this chapter and considering the stated hypothesis to be tested the following can be concluded:

- It is not possible to construct a representative simple steady state model incorporating mixotrophy from the *in situ* observational data from the subtropical north Atlantic.

5. Mixotrophy: Can a Simple Dynamic Model be Consistent with Observational Data?

The structure of a simple steady state ecosystem model incorporating mixotrophy was described in Chapter 4. This model forms the basis of the dynamical model presented in this Chapter. Due to methodological constraints, data on mixotrophy from oligotrophic ecosystems is very limited. Despite this small dataset, an assessment of the mixotrophic model structure can be made. The selection of model parameters in an attempt to fit to the very limited *in situ* observations of organism abundances (detailed in Chapter 4, Table 4.2-4.3) using a stochastic optimisation technique is described. Model performance is further assessed through comparison of model output to observed fluxes (described in Chapter 4, see Figure 4.1 and Table 4.2) and to previously published literature.

This Chapter addresses the following hypothesis:

- A simple dynamic zero-dimensional seasonal cycle model incorporating mixotrophy can be configured to maintain variable coexistence and be consistent with the limited *in situ* observations from the north Atlantic oligotrophic gyre.

5.1. Mixotroph Model Development

A steady state model incorporating mixotrophy was presented in Chapter 4. Through network analysis, using *in situ* observations of fluxes, two of the model fluxes were found to be in the opposite direction to those ecologically expected (phosphate uptake by mixotrophs, τPM, and bacterivory by grazers, ΩBG, see Figure 4.3). Two changes were consequently made to the model. Fluxes that had been diagnosed as going in the ecological incorrect direction were replaced with explicit fluxes in both directions to allow the net

flux to be decomposed. The model also became a dynamic one, representing an annual cycle to test whether the assumption of steady state hindered attempts to characterise the ecosystem.

5.1.1. Model Structure and Equations

This section describes the structure of a simple dynamic model developed for the study of mixotrophy. This dynamic model is also (such as the model presented in Chapter 4) a zero-dimensional mixed layer model with four components, representative of the interaction between a nutrient (phosphate, P), an obligate heterotroph (grazers, G), bacteria (both autotrophic and heterotrophic bacteria, B) and a mixotroph (M), see Chapter 4, Section 4.3 for reasons for model structure. The model presented in this Chapter, however, is now time varying, forced by a seasonally varying mixed layer. It is effectively a two box model. The upper layer is assumed to be homogenously mixed and biologically active. The lower layer is assumed to be abiotic with a constant nutrient concentration of deep phosphate (*Po*). The model conserves mass. The structure of the biological model including the flows between the four state variables is presented in Figure 5.1.

Changes in the modelled ecosystem are forced by seasonal variation in the mixed layer (ML). As the ML deepens, P is mixed into the ML from more nutrient rich deeper water, but there is a dilution of bacteria, mixotrophs and grazers as their concentrations below the ML are assumed to be zero (see Figure 5.1 and Equations 5.1-5.7). A reduction in mixed layer depth (MLD) leads to no change in nutrient or microbial groups concentration, as no new water is entrained into the ML. Unlike other similar simple plankton models (e.g. Fasham et al. 1990), when the ML shoals the grazers do not concentrate within it, as they are small within oligotrophic gyres and thus assumed to be effectively non-motile. Light forcing is also not included, as it is not considered a limiting factor within the oceanic subtropical gyres (Pers. Comm. M. Zubkov). The simple model structure allows the biological dynamics in the model to be understood without being confused by more complex physical forcing (Fasham 1993). For simplicity, horizontal advection is also not considered in the model. Consequently it is implicitly most relevant to large uniform areas of open ocean (Evans and Parslow 1985), such as the oligotrophic Atlantic. The issue of homogeneity in the gyre is returned to in Chapter 6.

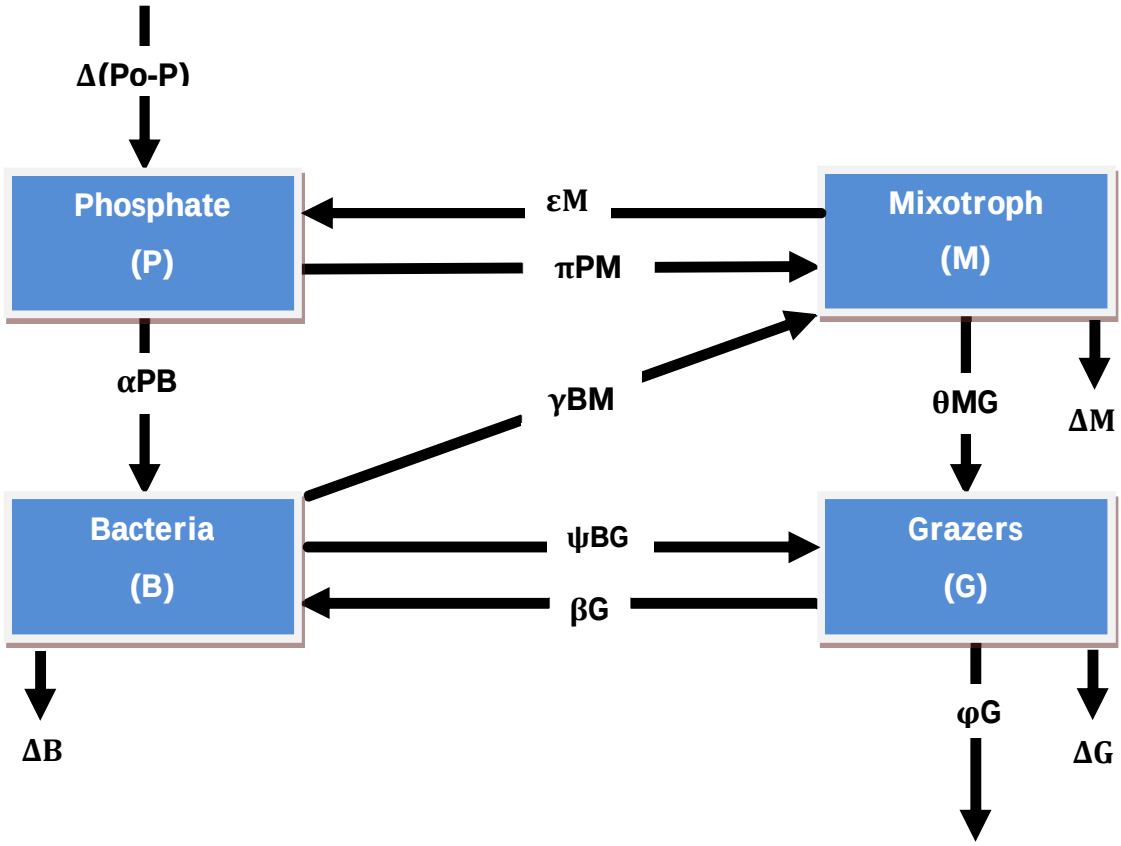


Figure 5.1. A schematic of the dynamic model incorporating mixotrophy. Note additional fluxes compared with Figure 4.3 (βG and εM). See Equations 5.1 – 5.7.

The Equations for the model are:

Equation 5.1 $$\frac{m+h^+}{MLD} = \Delta$$

Equation 5.2 $$\frac{dMLD}{dt} = h(t)$$

Equation 5.3 $$h^+(t) = \max\left(h(t), 0\right)$$

Equation 5.4 $$\frac{dP}{dt} = \Delta\left(Po - P\right) + \varepsilon M - \alpha PB - \pi PM$$

Equation 5.5 $$\frac{dG}{dt} = \theta MG + \psi BG - \beta G - \Phi G - \Delta G$$

Equation 5.6 $$\frac{dB}{dt} = \alpha PB + \beta G - \gamma BM - \psi BG - \Delta B$$

Equation 5.7 $$\frac{dM}{dt} = \pi PM + \gamma BM - \theta MG - \varepsilon M - \Delta M$$

The model was coded in Fortran 90. Equations 5.2 and 5.3 describe the changes in MLD. Mixing between the mixed and deeper constant layer due to turbulence is also represented explicitly in the Equations as m (Equation 5.1). For a variable x, the combined effects of changes in the MLD and mixing is $\frac{m+h^+}{MLD} \times x$, if x is zero in the lower layer (see Fasham et al. 1990).

Equations 5.4 to 5.7 describe the evolution of the model biogeochemical fields. The parameters α and π control the uptake rate of phosphate (P) by bacteria (B) and mixotrophs (M) respectively. ε is the rate of exudation/excretion by mixotrophs directly back to the P pool. γ and ψ control the grazing rates of mixotrophs and heterotrophic flagellates (grazers, G) on bacteria respectively. θ sets the rate of grazing of mixotrophs by grazers and β is rate of the excretion by grazers, assumed taken up instantaneously directly into the bacterial pool. Φ represents export from the system through the grazers, whether as waste, dead organisms or to higher trophic levels. A full list of parameters in the model is given in Table 5.2.

These coupled differential equations were solved using a fourth-order Runge-Kutta algorithm (see Press et al. 1992). The model was run using a time step of 0.01 days. A time step bigger than this led to instability in the system. The model was initially run for 30 years to ensure coexistence of all variables and that a repeating annual cycle was reached. However when coexistence was present, the model was found to reach equilibrium in less than 3 years. The tenth year of a model run was thereafter used for analysis for safety.

As an additional two parameters (ε, exudation of P by mixotrophs directly into the P pool and β, remineralisation of P from grazer mortality, excretion, messy feeding etc. directly into the bacterial pool) are present relative to the initial seven used in the steady state model (presented in Chapter 4), the observational flux data are insufficient to use network analysis to constrain the unknown parameters under an assumption of steady state, as was done in Chapter 4. For this reason, both changes (extra fluxes and dynamic model) were made at the same time. Parameter values were estimated instead using objective model optimisation techniques (see Section 5.2).

5.2. Model Optimisation

Parameters of planktonic ecosystem models can be especially hard to determine accurately and precisely. Many parameter values are poorly constrained (see large range of parameter values from literature in this study for example, Table 5.2) and immeasurable and/or unrepresentative of tangible physiological functions (Franks 2009). Difficulties in parameterisation are also exacerbated by the simplifications of numerous species into broad microbial functional groups (Hood et al. 2006a). The parameterisation of these groups is often based on axenic cultured species *ex situ* (Pahlow and Oschlies 2009), which may not represent the microbial group in a natural diverse *in situ* sample. Using an objective optimisation technique it is possible to use observational *in situ* data (Section 4.2) to estimate parameter values. Parameter optimisation techniques aim to adjust the model parameters to the data, by minimising the misfit between a set of observations and the model's predictions for those observations. The benefit of estimating parameters through optimisation is that it enables parameters that could not be measured directly or *in situ* to be estimated, by utilising the available observations and known parameter ranges and assuming that the structure and interactions within the model are representative (Ward et al. 2010). A number of optimisation techniques for ecological model fitting are available that enable a set of parameter values to be found that minimise the misfit between observations and model output, such as the variational adjoint technique (Friedrichs et al. 2007), genetic algorithms (Schartau and Oschlies 2003a, Schartau and Oschlies 2003b) and annealing algorithms (Kidston 2010).

5.2.1. The µ-Genetic Algorithm

In this study the parameter values were estimated using a micro-genetic algorithm (µGA) stochastic optimisation technique analogous to evolution by natural selection, following Schartau and Oschlies (2003a) and Ward et al. (2010). A flow diagram is used to illustrate the general process of µGA (see Figure 5.2) in conjunction with the following short description.

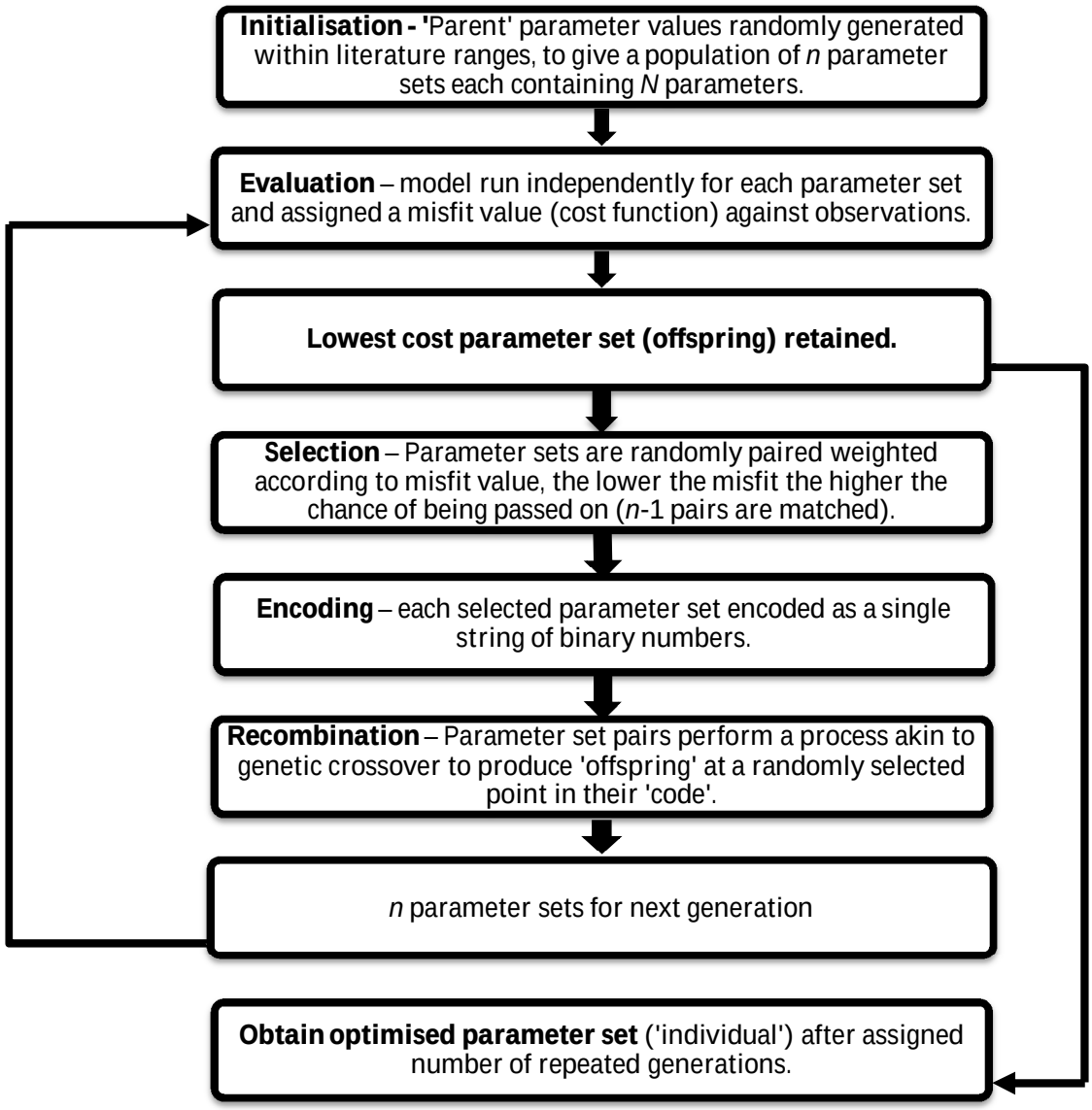


Figure 5.2. Flow diagram illustrating the µ-Genetic Algorithm (µGA).

The µGA begins with *n* (population size) randomly generated model parameter sets (parents) chosen from within the predefined parameter range (see Table 5.2 and Section 5.2.4 on applying the µGA). The model is run independently for each parameter set. Each parameter set is assigned a misfit value using a cost function (see Section 5.2.2) which evaluates model output against the observations. The best fit (lowest cost) parameter set is retained for the next generation. The parameter vectors are then randomly paired, the probability of their choice weighted according to cost (the lower the parameter vector cost, the higher the likelihood of selection) to give *n*-1 pairs. The two parameter sets in each pair are then encoded as single strings of binary digits equivalent to their numeric value, here using 6-bit accuracy.

A process analogous to genetic crossover is then applied. For each paired parameter set, a single point along the 'code' is selected at random and all points occurring after this point are exchanged between the parameter sets. One of the two new parameter sets is then chosen at random to pass to the next generation. This 'genetic' crossover ensures that subsequent generations explore parameter space more fully. The µGA cycles through a predefined number of iterations (generations), with the parameter sets of lowest cost at the end being the optimised one. In order to prevent the parameter search from getting trapped in a small region of parameter space, if the binary codes describing the *n* parameter sets (individuals) contain < 5 % difference between each other, only the currently best parameter set is retained and *n*-1 parameter sets are regenerated at random. Further information regarding this technique is given by Ward et al. (2010).

5.2.2. Cost Function

The cost function used within the µGA optimisation technique provides a non-dimensional value representing the misfit between the model output and observational data. It is a quantitative assessment of model performance and provides an indication of 'goodness of fit' (e.g. Allen et al. 2007). The cost function is defined here as the sum (over the four model components) of the weighted least square misfits between the model result and observations (Schartau and Oschlies 2003a). The nearer the cost function (J) is to zero, the better the model is said to perform.

Equation 5.8

$$J = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N}\sum_{n=1}^{N}\frac{(M_n - D_n)^2}{{\sigma_{D_n}}^2}$$

Equation 5.8 defines how the cost function J quantifies misfit between observations (D) and model output (M). Misfits are summed over the number of data types T (here there are 4; phosphate, grazers, bacteria and mixotrophs) and the number of observations (N, here there is 1) for each data type. σ_{Dn} is the error associated with the observational data. The cost function can be very sensitive to the uncertainty estimate (σ_{Dn}) on the observed data (D_n), as a large uncertainty estimate on an observation will lead to a lower cost than if there was a small uncertainty estimate for the same model output and data. Uncertainties used in the optimisation will be discussed further in Section 5.2.3.4. Note that the contributions to the cost of each of the model variables can be used individually to assess how well the model reproduces that field (see Section 5.2.3.4 for details of observations used in cost function), $\frac{1}{T}$ can be used to provide the mean cost, see Table 5.4.

5.2.3. Data Input to Model

Physical forcing of the system was provided by changes in MLD. Cellular phosphate content (necessary to convert model concentration in units of µmol P l^{-1} to cell abundance for fitting to the observations) and the parameter ranges were determined from a review of the literature. The following Sections describe these data.

5.2.3.1. Mixed Layer Depth

Seasonal variation in MLD was obtained using profiling Argo float data, sourced from www.coriolis.eu.org. Data were extracted within the predefined oligotrophic gyre area (39°N to 15°N and 40°W to 20°W) for 2005, to coincide with the CTD uptake observations (see details in Section 4.2). MLD was calculated using temperature data from extracted profiles. MLD was defined as the depth where there is 0.2°C difference in temperature from that at 10m below the surface (de Boyer Montégut et al. 2004). This has been shown to provide better approximations to MLD than the 0.5°C benchmark of Monterey and Levitus (1997). Salinity, temperature and density can all provide the measureable gradient to detect the base of the ML (de Boyer Montégut et al. 2007). Temperature was used here as salinity sensors are not present on all Argo floats, therefore potential density data was not consistently available either. The monthly averaged data have been linearly interpolated to daily values (see Figure 5.3). These interpolated data were used to force the model, repeating for all years of the model run. Mixing rate (m) was set to 0.1 m d^{-1} according to Fasham et al. (1990).

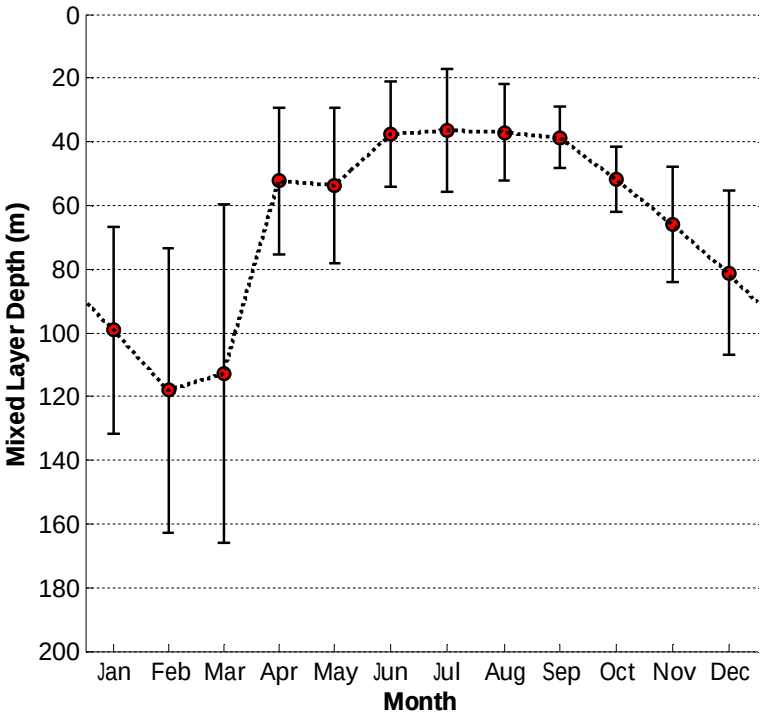


Figure 5.3. Mixed layer depth (MLD) used to force the model, using temperature data from Argo profiles. Red circles are the monthly averaged MLD within the region in 2005. Error bars indicate ± 1 standard deviation for the monthly means. Dashed line shows linearly interpolated values.

5.2.3.2. Cellular Phosphate Content

Data was required from the literature for the amount of phosphate per cell for each of the model's microbial groups (B, G and M). This was necessary to estimate cell abundance from the model output for comparison to observations, as elemental analysis was not undertaken on the AMT17 cells. Data was collated from the literature and then the mean, minimum and maximum cellular phosphate content (g P $cell^{-1}$) was calculated for each microbial group. Published estimates were taken from studies on cultures of oligotrophic origin (natural or cultured) or of organisms of an oligotrophic species size (M and G, cell diameter < 6 µm and B < 1.5 µm). Bacterial cultures had to be P-limited (defined as < 50 µmol P l^{-1}) for data to be used.

Table 5.1. Literature values for grams of phosphorus per cell for (mean, minimum and maximum) heterotrophic flagellates or Grazers (G), bacteria (B) and mixotrophs (M). SD is standard deviation and *n* is sample number, source references are shown

g P cell^{-1}	Mean	Min	Max	SD	*n*	Reference
M	3.5 x 10^{-13}	3 x 10^{-13}	4 x 10^{-13}	7.1 x 10^{-14}	2	Legrand et al. (2001)
G	9.3 x 10^{-13}	1.2 x 10^{-13}	2.2 x 10^{-12}	9.9 x 10^{-13}	4	Eccleston-Parry and Leadbeater (1995)
B	1.3 x 10^{-15}	9.8 x 10^{-16}	1.75 x 10^{-15}	2.4 x 10^{-16}	19	Grob et al. (2011); Bertilsson et al. (2003); Heldal et al. (2003); Gundersen et al. (2002)

Table 5.1 details the results of the literature review on cellular phosphate content. As expected phosphate content was highest in the largest organism, the heterotrophic flagellates, the only obligate heterotroph within the model (G, mean phosphate content of 9.3 x 10^{-13} g P cell^{-1}). The mixotrophs (M) had a mean content of 3.5 x 10^{-13} g P cell^{-1}. It was necessary to calculate a representative bacteria cellular phosphate content to reflect the fact that different types of bacteria vary in their phosphate content. Therefore the relative abundances measured at the CTD stations of *Prochlorococcus* spp., *Synechococcus* spp. and heterotrophic bacteria (average of 32%, 66% and 2% respectively) and the associated P content from the literature were used to calculate a weighted bacterial cellular phosphate content mean of 1.3 x 10^{-15} g P cell^{-1}. Bacteria, due to their small size, have the lowest cellular phosphate quotient.

5.2.3.3. µGA Parameter range

Parameter ranges were defined from the literature, see Table 5.2. It was necessary to convert the units reported by published study's to match the model units. An example of parameter calculation for α, using Michelou et al. (2011) is as follows. The paper provided a P per cell uptake rate and a cell abundance. Therefore αPB could be calculated (as, uptake per cell x cell abundance = P uptake rate, µmol P l^{-1} d^{-1}). Using the paper's specific P and B concentration (µmol P l^{-1}, calculated by multiplying the cell abundance with cellular P content), α could then be deduced (α = αPB / PB). Where the cellular phosphate content (g P cell^{-1}, Table 5.1) was required for a calculation and was not provided in the paper the parameter value was calculated using the relevant microbial groups minimum, maximum and mean g P cell^{-1} (Table 5.1) to provide a maximum parameter range. This is reflected in the large parameter ranges (some several orders of magnitude) obtained from the literature. However, as the optimisation of an oligotrophic mixotrophic ecosystem had not

be undertaken previously and the literature reported wide ranging values, it was necessary for the optimisation to explore a large amount of parameter space. All parameters in Table 5.2 were optimised (see applying the µGA, Section 5.2.4), with the exception of mixing rate at the base of the ML (*m*), which was set according to Fasham et al. (1990), where it's variation was shown to have little effect on the model output. Phosphate concentration (*Po*) below the ML (deep phosphate) was not set, as the deep phosphate observation for the region region was sourced from the World Ocean Atlas (Garcia et al. 2006, Garcia et al. 2010). For autumn at 150 m (below the ML at all times of year) this ranged between 0 and 0.6 µmol P l^{-1}.

Table 5.2. A description of model parameters and the parameter ranges defined from the literature for use within parameter optimisation. Note that *m* had a set value and was not included in the optimisation.

	Parameter	***Range***		***Unit***	***Ref*[1]**
		Min	**Max**		
α	Phosphate uptake by bacteria	32.91	1090.55	$d^{-1}(\mu mol\ P\ l^{-1})^{-1}$	9
γ	Grazing by Mixotrophs on Bacteria	0.57	258.08	$d^{-1}(\mu mol\ P\ l^{-1})^{-1}$	2, 7, 15
ψ	Grazing by Grazers on Bacteria	0.01	287.4	$d^{-1}(\mu mol\ P\ l^{-1})^{-1}$	1, 2, 7, 15
θ	Grazing by Grazers on Mixotrophs	0.01	0.51	$d^{-1}(\mu mol\ P\ l^{-1})^{-1}$	7
π	Phosphate uptake by Mixotrophs	0.7	158.77	$d^{-1}(\mu mol\ P\ l^{-1})^{-1}$	11,14
β	Material from Grazers due to natural mortality, excretion and messy feeding taken directly into Bacterial pool	0	0.3	d^{-1}	3, 4, 8
ε	Material from Mixotrophs remineralised directly into Phosphate pool	0	0.88	d^{-1}	4, 5, 8, 10, 13
Φ	Loss rate for Grazers	0	0.25	d^{-1}	4, 5, 8, 12, 16
Po	Phosphate in deeper abiotic layer	0	0.6	$\mu mol\ P\ l^{-1}$	6
m	Mixing rate	set to 0.1		$m\ d^{-1}$	5

5.2.3.4. Observations for use in Cost Function

To calculate misfit through the cost function (Section 5.2.2) *in situ* observations were compared to the modelled cell abundance (cell ml^{-1} for M, B and G) and P concentration (µmol P l^{-1}) from the same day of the year (see Section 4.2, Tables 4.2 and 4.3 for data).

[1] ***Reference key:*** 1. Barcina et al. (1992); 2. Bennett et al. (1990); 3. Edvardsen et al. (2002); 4. Fasham (1993); 5. Fasham et al. (1990); 6. Garcia et al. (2006, 2010); 7. Hall et al. (1993); 8. Kriest et al. (2010); 9. Michelou et al. (2011); 10. Obayashi and Tanoue (2002); 11. Rothhaupt (1996b); 12. Sarthou et al. (2005); 13. Six and Maier-Reimer (1996); 14. Stibor and Sommer (2003); 15. Tsai et al. (2011); 16. Turner (2002)

The flux data (F1, F2 and F3, µmol P l^{-1} d^{-1}) detailed in Section 4.4.1 and Table 4.4 were used after optimisation for independent testing of the model (this will be explained further in Section 5.2.6). As the units of the modelled variables are in µmol P l^{-1}, a conversion was required to convert M, B and G model output to cell abundance. This was done using the mass of cellular phosphate (detailed in Table 5.1) and the molecular mass according to Equation 5.9.

Equation 5.9

$$\frac{\boldsymbol{Model\ P\ per\ unit\ volume}\ (\boldsymbol{mol\ l^{-1}}) \times \boldsymbol{P\ molar\ mass}\ (\boldsymbol{30.97\ g\ mol^{-1}})}{\boldsymbol{P\ cellular\ content}\ (\boldsymbol{g\ cell^{-1}})} = \boldsymbol{Microbial\ abundance}\ (\boldsymbol{cell\ l^{-1}})$$

To calculate the cost function, an error estimate was required for each of the observations. In this study a standard error of 5 % for each microbial group cell count observation (M, B and G) was used. This was made up from the following – 2 % flow cytometer (laser alignment, flow rate); 1 % flow cytometric gating; 1 % pipette and 1 % scintillation counter error (Pers. Comm. R. Holland). For P concentration the experimental error derived from the methodology (Zubkov et al. 2007) was used (see Table 4.2). It should be noted that error may have also arisen from the g P $cell^{-1}$ used in the model variable conversion calculation (Equation 5.9). Unfortunately, for this study the optimiser could not be used to fit the g P $cell^{-1}$ as well as the variable concentrations, as one was dependent on the other. Thus, to allow the model to explore different cellular P contents, separate optimisations were run with different g P $cell^{-1}$ applied (see Section 5.2.4 for details on applying the µGA).

Despite all three locations where data were available (see Figure 4.1 and Table 4.1) being located within the same oligotrophic region, delineated by multivariate analysis in Chapter 2, model output was only compared to observational data from CTD14 for the following reasons. Firstly, CTD10 was unsuitable as it was missing grazer abundance. Secondly, there was a large variability between the CTD14 and CTD23 despite being only 4 days apart (e.g. B abundance - CTD23, 492,229 cell ml^{-1} and CTD14, 344,423 cell ml^{-1}). Consequently it would have been too difficult to 'fit' the model through both observations. This may have been due to ultraplankton spatial variability (this is discussed and investigated further in Chapter 6). At CTD14, the mixotrophs were the most abundant (CTD14 M, 11,942 cell ml^{-1} and CTD23 M, 2,748 cell ml^{-1}). Therefore the role of mixotrophy should be the most evident at this location. Consequently CTD14 was selected as the

observations to fit the model output too. Observations for just one day does not necessarily make it trivial to fit a model too, as will be shown.

5.2.4. Applying the µ-Genetic Algorithm

The µGA has a number of its own parameters, such as the number of generations used and the population size (*n*). These parameters can be adjusted to improve convergence on a solution. Standard practice is to set the µGA population size to equal the number of free parameters and that is done here, *n* = 9 (Schartau and Oschlies 2003a). The model optimisation was carried out for 10,000 generations, using a 30 year model run, more than the 5,000 generations of Ward et al. (2010) and 2,000 generations of Schartau and Oschlies (2003a). This greater number of optimisations was necessary as despite the observations being much more limited, it took more iterations to obtain coexistence of all variables and an acceptable cost function. Optimisations were initially run for 2,500 and 5,000 generations, and the optimised cost function was considerably higher. No improvement in cost was seen when more than 10,000 generations were run.

The µGA used randomly selected initial parameter sets, as a limitation of µGA is that it can be sensitive to the initial parameter values. In addition, the randomly generated initial parameters maximised the search over the parameter space (Ward et al. 2010), which was restricted to the parameter ranges in Table 5.2. In total 150 independent µGA optimisations were run each for 10,000 generations, this number were run, due to the difficulty in obtaining a low cost from the observations (see Sections 5.2.3.4). To take into account the wide range in measurements from the literature, the minimum, maximum and mean g P cell^{-1} were used to calculate model cell abundance for 50 runs each of the optimiser. An example of the difficulty fitting a model to even the very limited set of observations here is that out of the 150 optimisations, only 9 had M, B and G coexisting and had a total cost function of less than 100 (which is still high, see Section 5.3.1).

Each 10,000 generations run of the µGA took approximately 16 hours. Thus the optimisation required ~ 2,400 hours of computation. The actual time necessary was reduced considerably by the use of the Condor High Throughput Computing (HTC) environment, which enabled numerous optimisations to be run in parallel on a pool of machines (Thain et al. 2005). Notwithstanding this reduction in computation time attained by the Condor HTC, the 27 additional permutations that would be required to optimise to the different combinations of minimum, maximum and mean g P cell^{-1}, would take

~21,600 hours (non-stop optimisations for 2.5 years) if completed for 50 optimisations, this is beyond what could be realistically completed in this study's time frame.

5.2.5. Sensitivity analysis

Model sensitivity of the optimised parameter set determined by the µGA technique was also assessed. The model was run varying each of the parameters in turn, whilst all other parameters were held constant at their optimised value. Model runs were performed for ± 10 % and ± 20 % of each optimised parameter value, leading to 36 model permutations. All runs used parameter values within the literature defined range (see Table 5.2), except for α + 20 %, which was outside of this range, at 1127.35 $d^{-1}(\mu mol\ P\ l^{-1})^{-1}$, which was included for completeness. Changes in total cost and the individual contributions to this by each model variable were assessed for each permutation. A large increase in cost as a parameter is varied indicates the model is sensitive to that parameter. Here it should be noted that a stipulation was placed upon the model (here and previously when applying the µGA), such that if coexistence of all variables was not achieved a large cost of 500, was imposed.

5.2.6. Independent Model Testing

The measured *in situ* uptake rates detailed in Chapter 4 (F1, F2 and F3 - see Table 4.4 in Section 4.4.1), were not used in the µGA optimisation. The model could not be solved algebraically unlike in Chapter 4. It was therefore decided to adopt the most likely future approach to modelling these regions, optimising to observations. For reasons given, flux data will remain very sparse but abundances are much easier to obtain. For this reason abundances are used for fitting, but fluxes are still used as an independent test.

The flux data, although quantitative, present some difficulties in their use. Firstly, as detailed in Chapter 4, the uptake rate experiments measure the entire uptake rate and cannot distinguish from where it is arising. Therefore the model fluxes have to be combined for comparison. Secondly, and perhaps more importantly, the P uptake flux measured during the time period of the experiment, 45 - 60 minutes (see Zubkov et al. 2007 for methodology) may also be including excreted P or P passed to another microbial group, for example through predation. As such the F1 - F3 measurements are net, measuring P going in, as well as P going out during the time frame of the experiment, not

measuring just the total gross P uptake. It is currently not possible to separate these processes within the experiment or to say which processes are definitively occurring during the period of the experiment, although the experiments were designed to make the flux as close to gross as possible. However for completeness both 'gross' and 'net' model output is compared to the observed fluxes. For clarity this is outlined in Table 5.3. Mixing across the base of the ML was not included, as this is not present in the bioassay sample. Observed flux errors displayed in Figure 5.11 were compound (derived from the observations of cell count and uptake rate per cell ± their respective error). Correlations between errors were ignored so that maximum uncertainty is displayed.

Table 5.3. Breakdown of gross and net modelled fluxes to be compared to observed fluxes F1 - F3, see Figure 5.11.

Observed flux		Compared to: Gross Modelled Flux	Net Modelled Flux
F1	Uptake by M	$\pi PM + \gamma BM$	$\pi PM + \gamma BM - \theta MG - \varepsilon M$
F2	Uptake by B	αPB	$\alpha PB + \beta G - \gamma BM - \psi BG$
F3	Uptake by G	$\psi BG + \theta MG$	$\psi BG + \theta MG - \beta G - \phi G$

5.3. Model Assessment

5.3.1. Overall Model Performance

The best overall mean cost achieved by the 150 optimisations was 0.76 (see breakdown in Table 5.4). This was obtained using the maximum cellular phosphate content in Table 5.1. Optimal parameter values are shown in Table 5.5.

Table 5.4. Breakdown of calculated misfit (cost function) between model output and observational data for the individual model components

Phosphate	Grazer	Bacteria	Mixotroph	Mean Cost
1.98	0.06	0.63	0.35	**0.76**

This model performance was evaluated according to the Radach and Moll (2006) criteria for cost functions. The criteria are: $J < 1$ very good; 1 – 2 good; 2 – 3 reasonable; > 3 poor. The model therefore attains a very good fit (with a mean cost of 0.76), with the

microbial variables achieving a better fit (< 1, very good), than phosphate concentration (< 2, good).

Table 5.5. Optimised parameter values corresponding to the lowest cost function.

	Parameter	*Optimised value*	*Unit*
α	Phosphate uptake by bacteria	939.46	$d^{-1}(\mu mol\ P\ l^{-1})^{-1}$
γ	Grazing by Mixotrophs on Bacteria	12.83	$d^{-1}(\mu mol\ P\ l^{-1})^{-1}$
ψ	Grazing by Grazers on Bacteria	9.13	$d^{-1}(\mu mol\ P\ l^{-1})^{-1}$
θ	Grazing by Grazers on Mixotrophs	0.215	$d^{-1}(\mu mol\ P\ l^{-1})^{-1}$
π	Phosphate uptake by Mixotrophs	3.21	$d^{-1}(\mu mol\ P\ l^{-1})^{-1}$
β	Material from Grazers due to natural mortality, excretion and messy feeding taken directly into Bacterial pool	0.04	d^{-1}
ε	Material from Mixotrophs remineralised directly into Phosphate pool	0.24	d^{-1}
Φ	Loss rate for Grazers	0.14	d^{-1}
Po	Phosphate in deeper abiotic layer	0.5	$\mu mol\ P\ l^{-1}$

5.3.2. Model System Dynamics

Figures 5.4 and 5.5 show the 10th year of model output and observation for each variable and the associated physical forcing (the MLD). The model was successful in allowing coexistence, which previous mixotroph models have found difficult (e.g. Thingstad et al. 1996) with a repeating annual cycle. There is a low percentage difference between model and observations of the modelled grazer (1 %), bacteria (4 %) and mixotroph (3 %) cell counts to the abundance observations on Julian day 305 (corresponding to CTD14 observation day). The model fit to phosphate concentration is still reasonable, despite being the worst overall fit (33 % lower than the observation), though the model output is not within the uncertainty estimate of the observation.

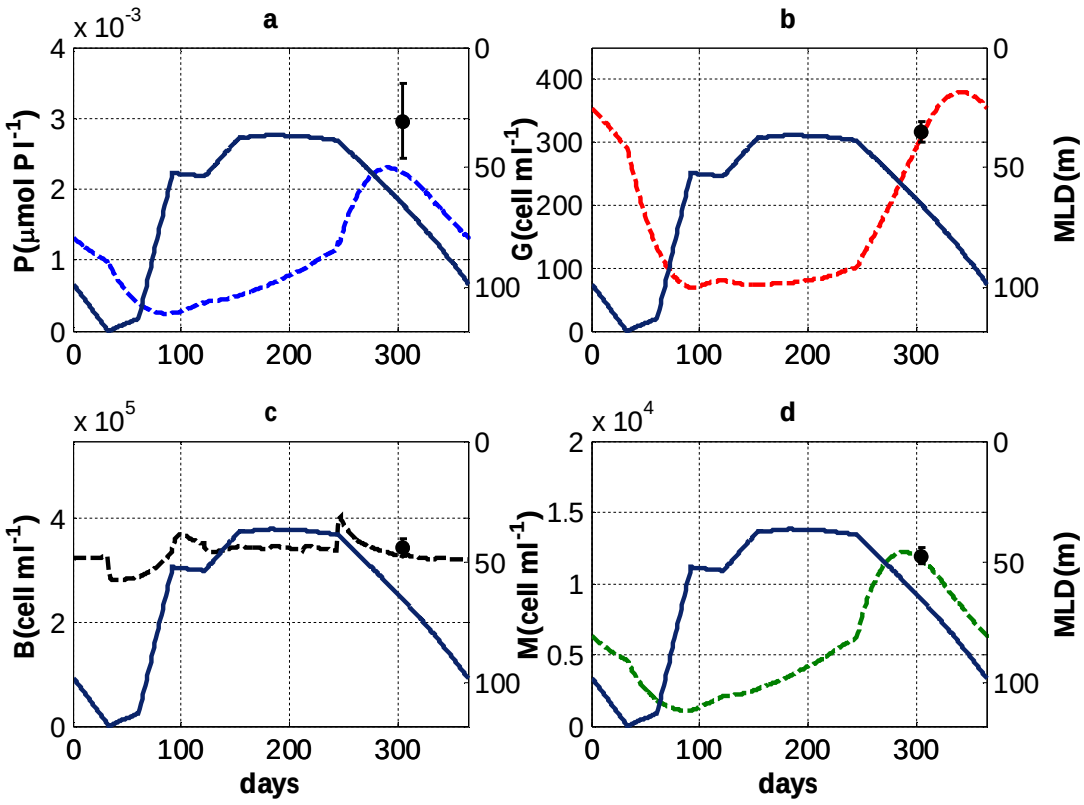


Figure 5.4. a-d Optimised model output (dashed lines). Observational data from CTD14 are also shown (observation and error described in Section 5.2.3.4), as is mixed layer depth, the solid lines (MLD, m). Note different units for (a) Phosphate concentration (µmol P l^{-1}) compared to (b) Grazers, (c) Bacteria and (d) Mixotroph abundance (cell ml^{-1}). *X*-axis, denotes Julian day.

Figure 5.5 is model output as in Figure 5.4b-d but in comparable units to P concentration (Figure 5.4a, µmol P l^{-1}) in order to demonstrate the relative variable sizes. Figures 5.4a and 5.5a-c show that the majority of the P within the system is located within the mixotroph variable and the least P is located within the modelled P pool. The majority of the P within the ecosystem is therefore within the microbial groups and not 'freely' available within the ML. The observations plotted on Figure 5.5 are calculated from microbial abundance observations converted to µmol P l^{-1} using a rearranged version of Equation 5.9. The error bars shown are as described in Section 5.2.3.4.

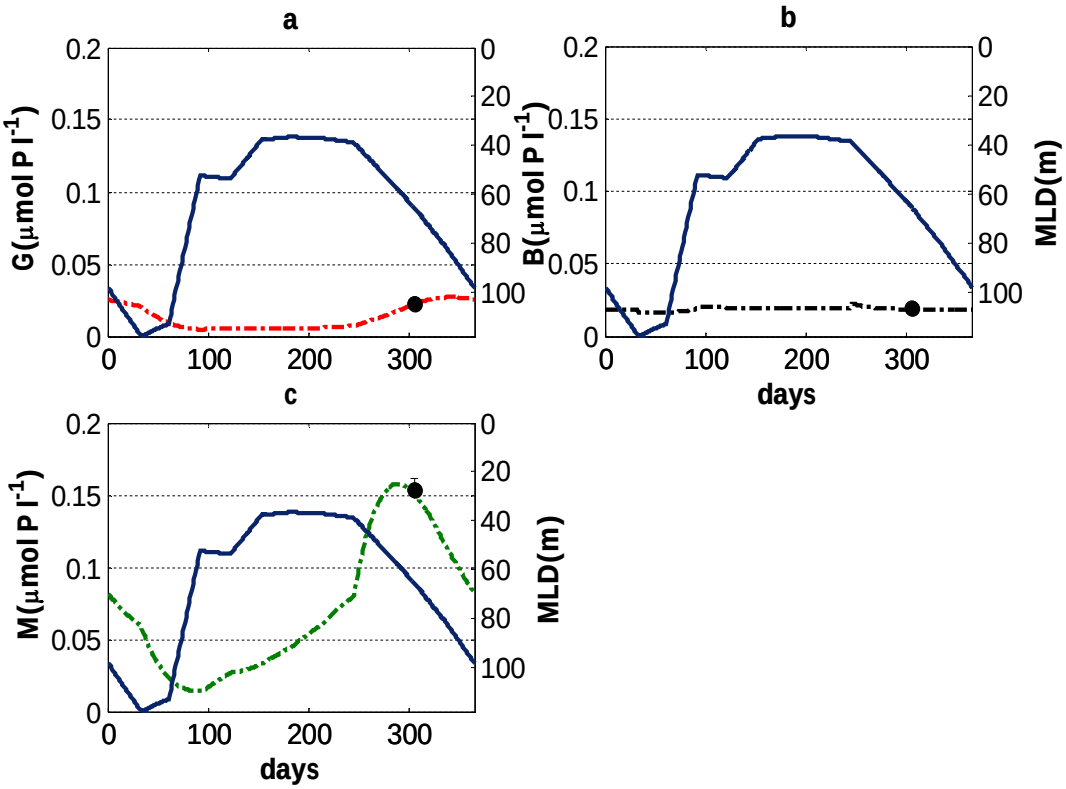


Figure 5.5. Optimised model output in µmol P l^{-1}. Also shown are observational data from CTD14 and associated error described in Section 5.2.3.4), (a) Grazers, (b) Bacteria and (c) Mixotrophs. MLD (m) on secondary *y*-axis is also shown. See Figure 5.4a for P concentration.

The optimised model output (Figures 5.4 and 5.5) demonstrates a 'seasonal' cycle throughout the year for P, G and M. The model output exhibits a peak in P, M and G in the autumn. Somewhat strikingly the bacteria varies little throughout the year in comparison to the other variables, see Figure 5.6.

Simulated phosphate concentrations are highest during the autumn months. The deepening of the ML in autumn causes an entrainment of deep P concentration. The mixotroph seasonal cycle very closely echoes the phosphate concentration with peaks and troughs at the same points. The peak in grazers is steeper than the peak in phosphate and mixotroph components and the grazers peak which is (< 1 month) after the mixotrophs.

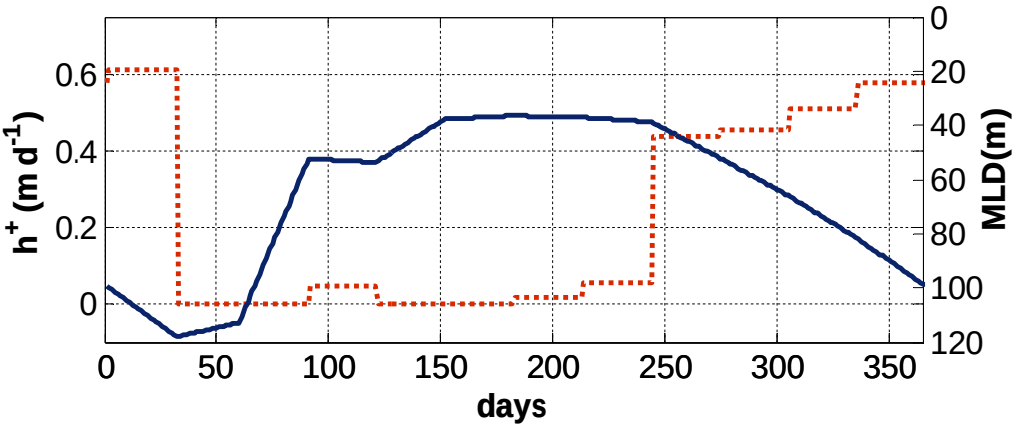


Figure 5.6. Mixed Layer Depth (MLD, solid line) and changes in h^+ (dotted line). *X*-axis is Julian day.

In addition to minimal 'seasonality' in bacteria in comparison to the other model variables (just ± 20 % over the year), the model output for bacteria also displays a 'step effect' upon even slow deepening of the ML (such as around Julian days 100 and 250, see Figure 5.4c in conjunction with Figure 5.6). This occurs due to the linear interpolation of MLD used to force the model. When the ML deepens, a step effect in h^+ is observed (see Figure 5.6 and Equations 5.1, 5.3 and 5.6). The variable h^+ controls both the influx of nutrients to the ML from the abiotic constant nutrient layer below and a dilution with bacteria free waters, hence the signal in B. It is visible in B but not in B, M and G because of the much weaker seasonal cycle.

To investigate the dynamics of the modelled ecosystem further the model fluxes are shown. Figure 5.7 shows the individual fluxes affecting each variable. The most striking observation is the symmetry in the dominant losses and gains shown in the plots on the left-hand side. The larger fluxes in the system essentially cancel each other out. The biggest loss from P is through direct bacterial uptake (αPB, Figure 5.7a), but this is balanced by the remineralisation of phosphate from mixotrophs (εM). The latter is also the largest loss out of the mixotroph variable (Figure 5.7g), despite the optimised ε parameter being in the lower half of the parameter range (see Table 5.2). This flux out of the mixotrophs is balanced by bacterivory, which in turn is the biggest loss for bacteria (γBM, Figure 5.7e). Importantly this indicates that mixotrophs are the primary control on bacteria and that, in this model, bacterivory is the mixotrophs principal phosphate source (Figure 5.7g). For B, bacterivory is balanced by P uptake (αPB, Figure 5.7e). Figure 5.7c shows that even the largest fluxes in and out of the grazers are of a relatively lower

magnitude, but the largest fluxes once again balance with export (ΦG) balancing grazing on bacteria (ψBG).

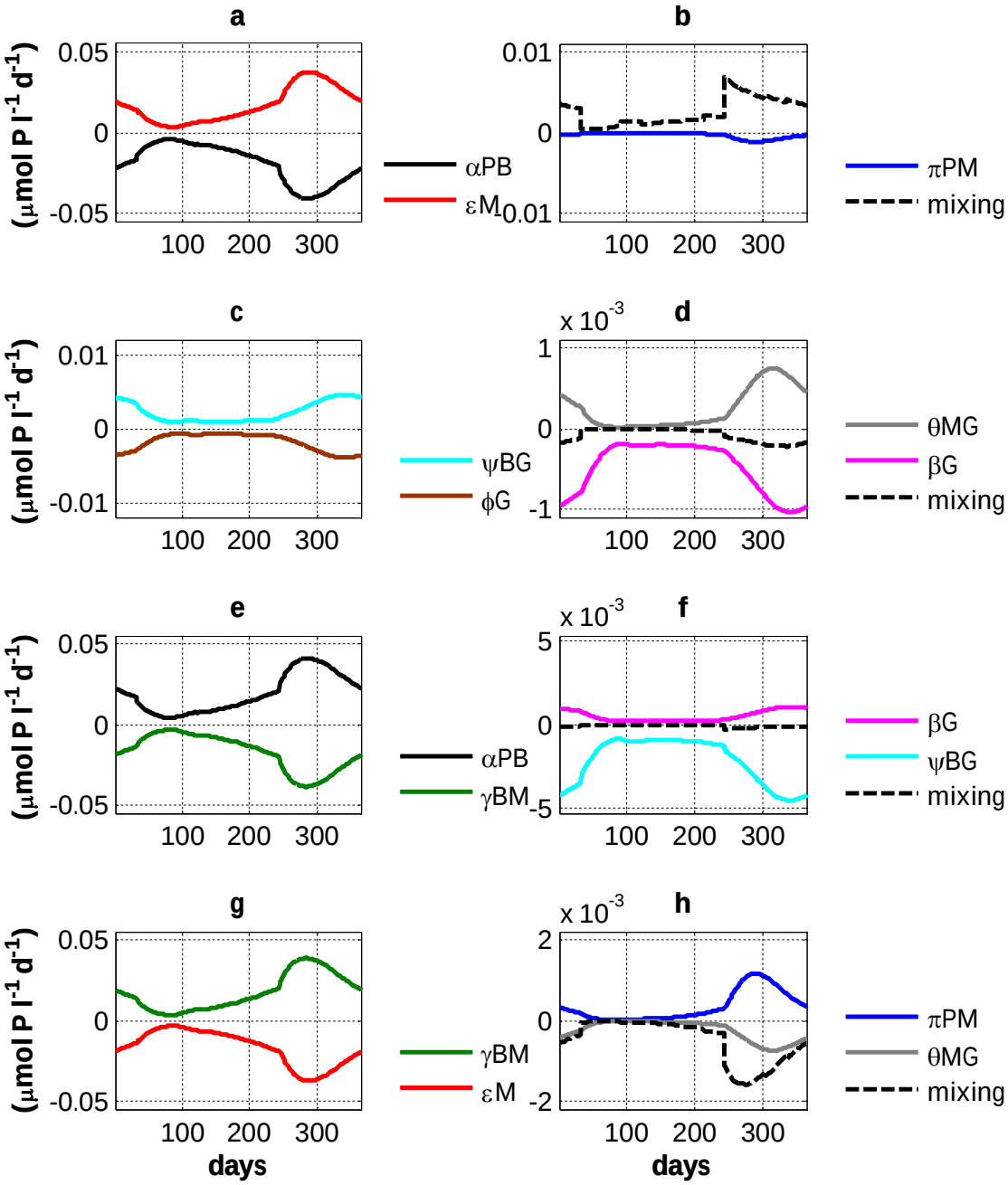


Figure 5.7. Individual model fluxes for (a, b) phosphate, (c, d) grazers, (e, f) bacteria and (g, h) mixotrophs. Each flux in the model is plotted as positive and negative depending on whether it is a source or a sink for each variable. Note *y*-axis scales in plots on the left hand side are a larger scale, than plots on the right hand side. Black dashed line is mixing across the ML.

The symmetry of the larger fluxes conceals the fact that the smaller fluxes (shown in plots on the right hand side of Figure 5.7) are crucial to the seasonal dynamics of the system. A small influx of P in autumn unbalances the system (without mixing, the model outputs

constant values for all variables and fluxes). This influx of P is due to the deepening of the ML around Julian day 240 (Figure 5.7b), providing a kick-start to the direct uptake of P by mixotrophs (πPM, Figure 5.7h) and bacteria P uptake (αPB, Figure 5.7e). Predation shortly follows and an increase in remineralisation is seen from G to B (βG, Figure 5.7d), due to the small increase in predation by grazers (ψBG, Figure 5.7d and θMG, Figure 5.7c), which is responsible for the depression in the system shortly after the peak. Therefore the system could be viewed as being both top-down (peak M depressed by predation by G and consequently remineralisation of P through G) and bottom-up (peak initiated by mixing in of P) controlled throughout the year.

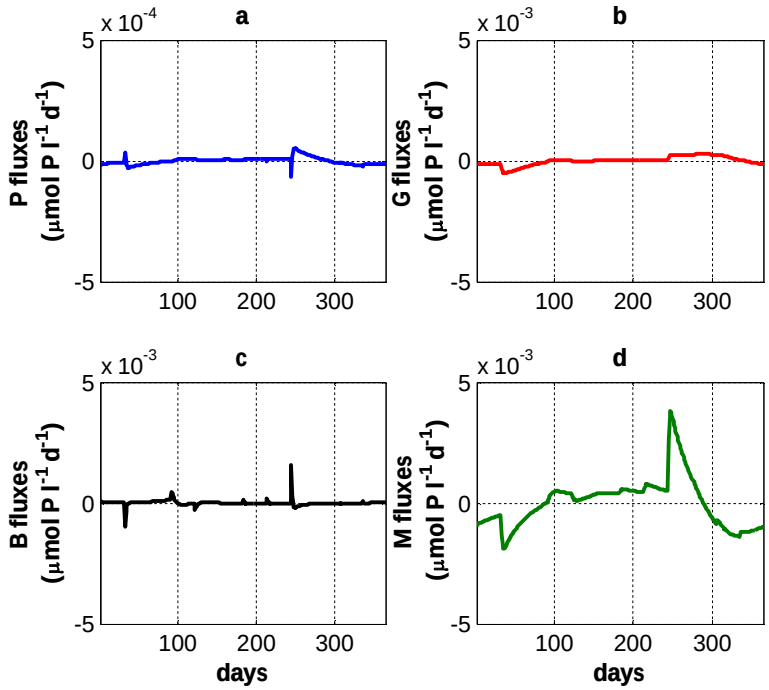


Figure 5.8. Net fluxes into each state variables (a) Phosphate, (b) Grazers, (c) Bacteria and (d) Mixotrophs. Note, P variable *y*-axis is an order of magnitude smaller than plots b-d. Also note that all axis are at least an order of magnitude smaller than the left hand side fluxes in Figure 5.7. *X*-axes are Julian days.

Figure 5.8 shows the balance of the system from the perspective of the net fluxes summed into each of the variables. The deepening of the ML causes a sharp decrease in P, because of the influx of deep P (Figure 5.8a). The introduced P is sharply taken up by B and M (Figure 5.8c-d). Any increase in the bacteria is quickly depressed by increases in M and G grazing (Figures 5.8b and 5.8d). The mixotrophs decline as P is consumed and as they are predated upon (although decreasingly rapidly) by grazers. From Figure 5.8, it can be seen that the system is tightly coupled and any P is quickly utilised, as would be expected in the

nutrient economical oligotrophic regions. Note the net fluxes are an order of magnitude smaller than the individual model fluxes, presented in Figure 5.7.

5.3.3. Model sensitivity

Sensitivity analysis was carried out on the optimised model parameters (see Table 5.2) and results are shown in Figures 5.9 and 5.10. These Figures show that the model cost is most sensitive to the parameters for grazing of bacteria by mixotrophs (γ), the grazing of bacteria by grazers (ψ), the remineralisation from mixotrophs directly into the phosphate pool (ε) and the export out of the mixed layer from grazers (ϕ). The model was not very sensitive to parameters involved in the direct phosphate uptake rate by bacteria (α) and mixotrophs (π).

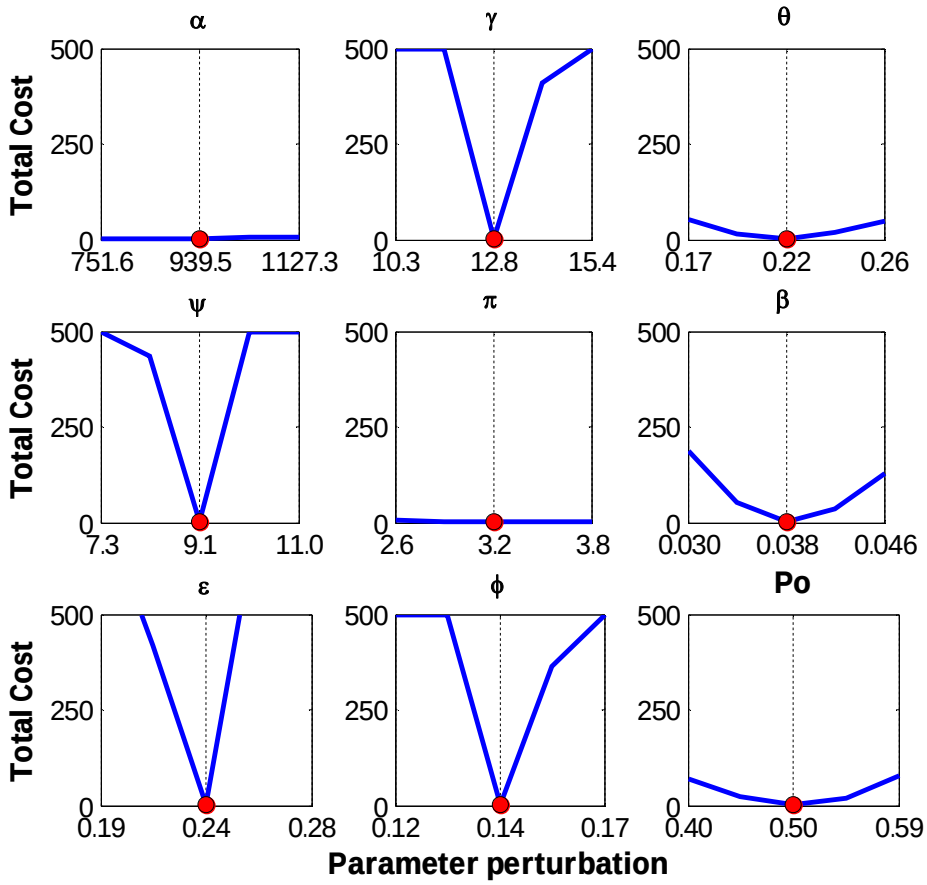


Figure 5.9. Model sensitivity to perturbation in individual parameters. While a single parameter was varied by ±10 % and ± 20 %, other parameters were held constant at the optimised value. The red dot and dashed line indicates the original optimised parameter value and blue line indicates changes in cost as the parameter was perturbed. A cost value over 500 indicates no co-existence of microbial groups (see Section 5.2.5).

The sensitivity analysis further showed that when ψ was perturbed above a value of 10 $d^{-1}(\mu mol\ P\ l^{-1})^{-1}$ or ε was increased beyond 0.26 d^{-1}, all variables could not co-exist. Likewise if γ was reduced below 11.6 $d^{-1}(\mu mol\ P\ l^{-1})^{-1}$ and ϕ below 0.13 d^{-1} co-existence for all variables could not be maintained. In each of these cases the mixotroph was outcompeted and became extinct within the model. It is interesting to note that mixotrophs and then grazers were the most sensitive variables within the model, as the largest proportion of change in total cost came from these two variables (see Figure 5.10).

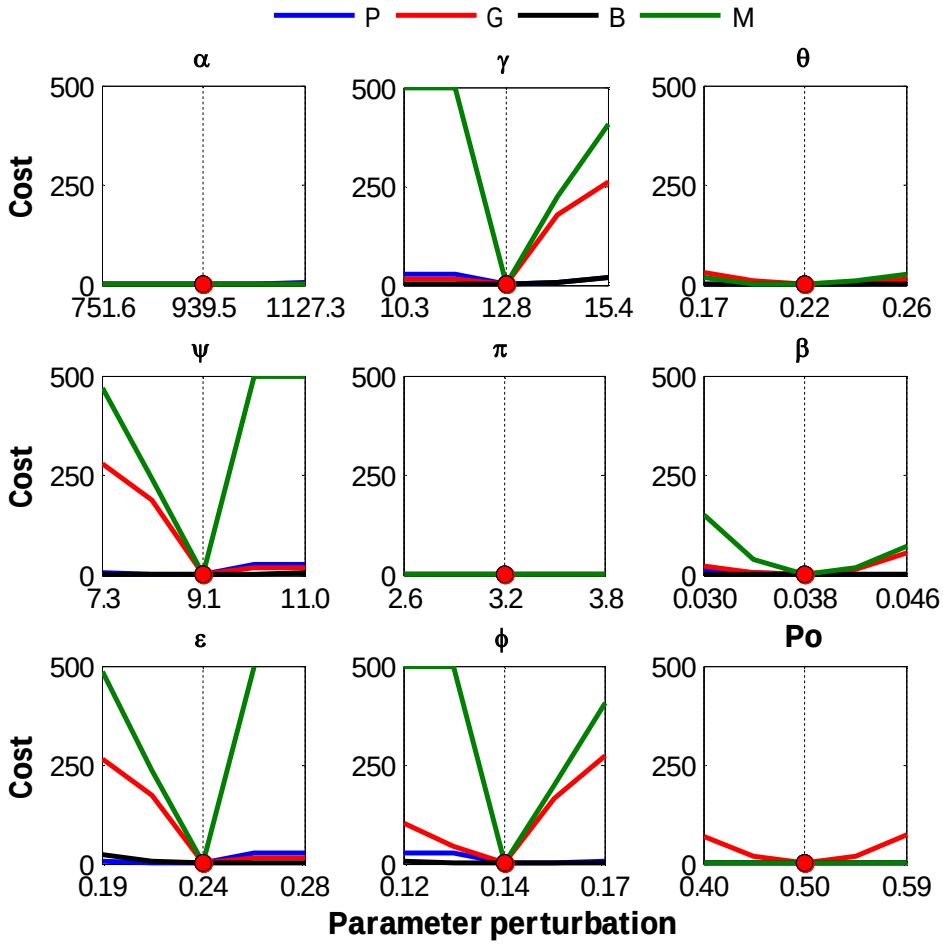


Figure 5.10. Model sensitivity to perturbation in individual parameters. While a single parameter was varied by ± 10 % and ± 20 %, other parameters were held constant at the optimised value. The red dot and dashed line indicates the original optimised parameter value, coloured solid lines indicates changes in individual variable (P, G, B and M) cost as the parameter was perturbed. A cost value over 500 indicates the variable had become extinct, therefore no co-existence was achieved in the model with those parameters.

5.3.4. Independent Model Assessment using Flux Data

Figure 5.11 presents a cross-validation of the model output to the independent flux data, described in Chapter 4 (Section 4.4.1). In Section 5.2.6, it was noted that it was uncertain which model fluxes describe the observational fluxes (F1-F3). Therefore both gross and net modelled flux output is plotted (see Table 5.3).

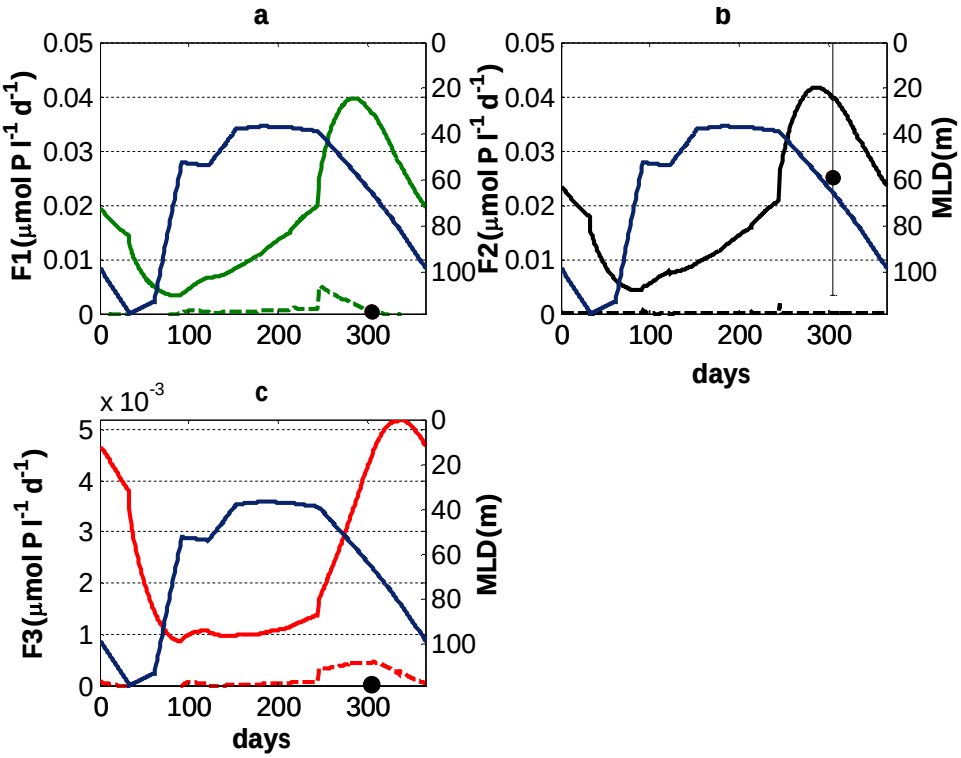


Figure 5.11. Model output in terms of fluxes (F1, F2 and F2, see Figure 4.3). Solid lines are gross uptake, dashed lines are net uptake (see Table 5.3). (a) F1, uptake by M, (b) F2, uptake by B, (c) F3, uptake by G. Black dots are the flux observations and error bars indicate compound errors for maximum uncertainty, present in plots a and c, although small. Note smaller scale for plot c.

Figure 5.11a is the uptake of P by M. The independent measurement is in reasonable agreement with the modelled net flux (dashed line). However the gross flux (solid line) is two orders of magnitude larger in places than the net flux (e.g. coinciding with observation, modelled net flux was 4 x 10^{-4} µmol P l^{-1} d^{-1}, whereas modelled gross flux was 4 x 10^{-2} µmol P l^{-1} d^{-1}). P uptake by B (Figure 5.11b), is closer to the modelled gross uptake (i.e. no losses included, the solid line). The modelled net flux is much smaller. For example, annual mean F2 net flux is 1 x 10^{-4} µmol P l^{-1} d^{-1}, while annual mean modelled gross flux is

2 x 10^{-2} µmol P l^{-1} d^{-1}. P flux into G (Figure 5.11c) is the worst fit to either the modelled net or gross fluxes, with the independent flux measurement below even the model's net flux estimation. The annual mean modelled net flux (1 x 10^{-4} µmol P l^{-1} d^{-1}) is an order of magnitude higher than the flux observation (2.3 x 10^{-5} µmol P l^{-1} d^{-1}). This may be an indication that the G processes are not being completely caught during the time scale of the experiment, as the G uptake observation (F3) is so much lower than the model predicted G uptake.

The uncertainty displayed is as large as possible (compound errors), to include the biggest possible potential range of the flux observations. Both gross and net model fluxes are plotted, as the relative rates at which processes (fluxes) are occurring over the experiments' timescale (< 60 minutes) is not known. In reality the observations are probably somewhere between the modelled gross and net uptake. This is discussed further in Section 5.4.1.

5.3.5. Modelled Bacterivory

The eating of bacteria (bacterivory) by mixotrophic algae has been shown to be ubiquitous within the oligotrophic Atlantic (Zubkov and Tarran 2008, Hartmann et al. 2012). Modelled output can be used to calculate the proportion of bacterivory undertaken by mixotrophs (algae) in the system. The percentage mixotrophic bacterivory has been calculated using Equation 5.10. The amount of phosphate obtained by mixotrophs through mixotrophy has also been calculated (see Equation 5.11), as this demonstrates how much of the nutrient acquired is through 'eating' rather than direct uptake.

Equation 5.10 $$\boldsymbol{Bacterivory\ undertaken\ by\ mixotrophs}\ (\%) = \frac{\gamma BM}{(\gamma BM + \psi BG)} \times \mathbf{100}$$

Equation 5.11 $$\boldsymbol{Mixotroph\ P\ obtained\ by\ bacterivory}\ (\%) = \frac{\gamma BM}{(\gamma BM \times \pi PM)} \times \mathbf{100}$$

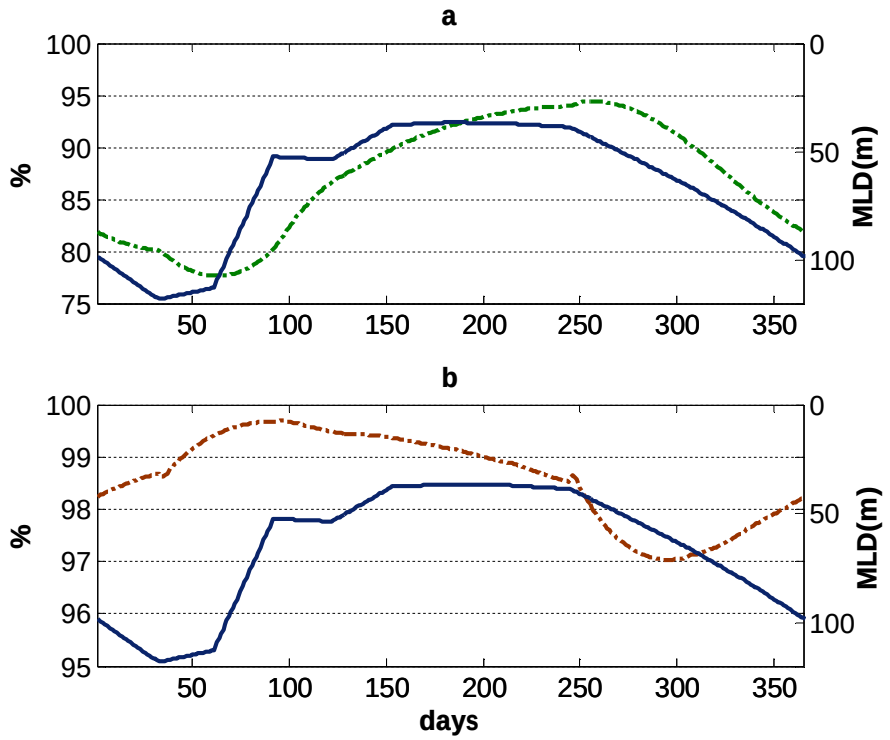


Figure 5.12. (a) The proportion of bacterivory undertaken by mixotrophs (%); (b) percentage of phosphate obtained by mixotroph through bacterivory (%). Note *y*-axis scales are different.

The fraction of bacterivory undertaken by mixotrophs varies between 77.6 and 94.4 % (17.3 %) over the year (Figure 5.12a). However the amount of P obtained by mixotrophs through bacterivory only varies 2.6% throughout the year, 97.1 to 99.7 % (Figure 5.12b). This demonstrates that obtaining P through the bacterivory pathway is the dominant process of P uptake for mixotrophs, even when the proportion of bacterivory undertaken by them declines.

In times of low phosphate concentration (around day 100), the fraction of phosphate obtained by mixotrophs from bacterivory is at its highest. The grazing of mixotrophs on bacteria may be reducing the competition (eating your competitor, as suggested by Thingstad et al. 1996) for phosphate and thus contributes to the increase in phosphate concentration thereafter. Percentage bacterivory by mixotrophs (in terms of total bacterivory) is, however, at this time at its lowest. There is strong competition from grazers for bacteria. The relatively steady-state nature of bacterial abundance throughout the year is because bacterial communities are continually under heavy grazing pressure (Jost et al. 2004), see Figures 5.4c and 5.5b.

5.3.6. Modelled Primary Production

Primary production is of interest, firstly for the reason that the calculation of primary production from the model output gives an indication of the impact of mixotrophy on the ecosystem functioning and dynamics. Secondly, because calculating primary production enables a comparison of the model output to previously published modelled estimates and observations that are and are not inclusive of mixotrophy.

Primary productivity (PP) is defined as in Stickney et al. (2000) as the total amount of dissolved inorganic nutrient uptake (P) by bacteria (αPB) and mixotrophs (πPM), see Figure 5.13. The PP in the model peaks in the autumn (echoing the peak in mixotrophs, Figure 5.4d). Throughout the year P uptake by mixotrophs is lower than by bacteria by a factor of over 10 (see Figure 5.13b). Therefore if the mixotrophs' contribution to modelled PP was only from direct P uptake (not including bacterivory), mixotrophs are contributing on average only 3.4 x 10^{-4} µmol P l^{-1} d^{-1}, almost two orders of magnitude less to PP than bacteria (average, 1.9 x 10^{-2} µmol P l^{-1} d^{-1}).

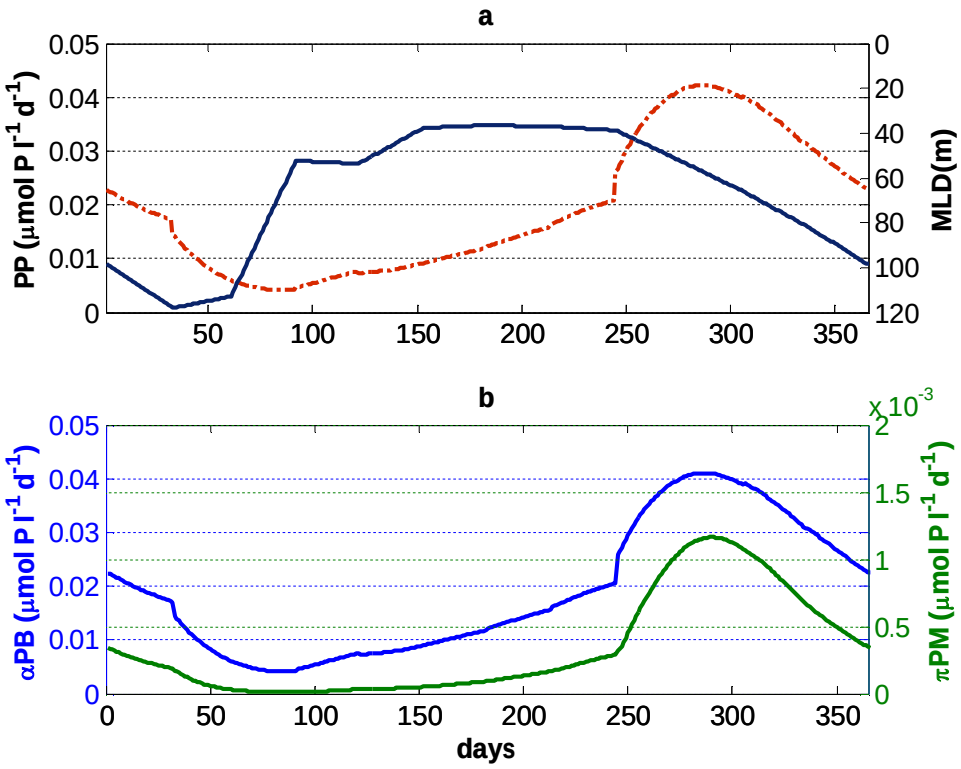


Figure 5.13. (a) PP, Primary Production; (b) PP component fluxes. The primary *y*-axis is P uptake by bacteria (αPB in blue), the secondary *y*-axis is P uptake by mixotrophs (πPM in green). Note different scales of *y*-axes and both *x*-axis are Julian day.

The presented model, however, cannot be used to determine explicitly if phosphate obtained from bacterivory is being utilised by mixotrophs for primary production (see Chapter 1). It has no light forcing component and therefore cannot model photosynthesis and production of biomass directly. Nevertheless, it is possible to consider the contribution of bacterivory by algae to primary production through a simple thought experiment.

Here *n* % of the phosphate obtained from bacterial grazing was arbitrarily included as being utilised for PP. Therefore PP labelled PP*n*, is the total amount of direct nutrient uptake by bacteria and mixotrophs plus *n* % of phosphate obtained through bacterivory, assumed used in PP. This can be compared to PP from direct uptake of P alone, in Figure 5.14.

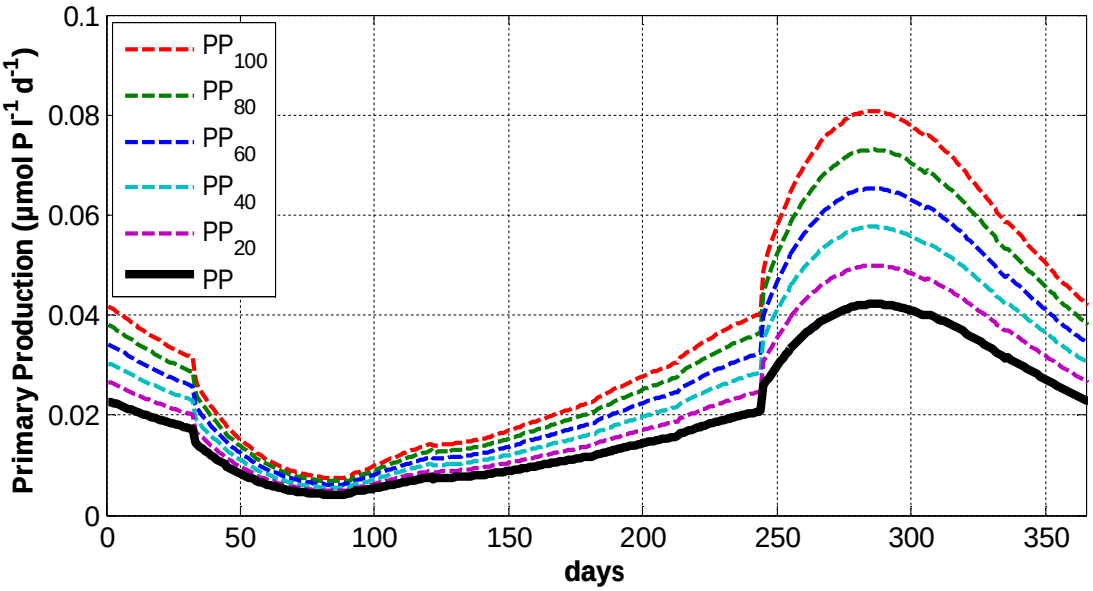


Figure 5.14. PP model output showing alternative scenarios for primary production with (dashed lines) and without (solid line) bacterivory derived phosphate utilisation. PP (µmol P l^{-1} d^{-1}, thick black line), only dissolved nutrient uptake utilised in PP. PP_{20}, PP + 20 % of bacterivory derived phosphate utilised in PP; PP_{40}, PP + 40 % of bacterivory derived phosphate utilised in PP; PP_{60}, PP + 60 % of bacterivory derived phosphate utilised in PP; PP_{80}, PP + 80 % of bacterivory derived phosphate utilised in PP and PP_{100}, PP + 100 % of bacterivory derived phosphate utilised in PP. Days, is Julian days.

Figure 5.14 shows the increase in primary production when 20, 40, 60, 80 and 100% of phosphate taken up through bacterivory contributes to primary production. If 100% of phosphate obtained by mixotrophs through bacterivory is utilised in primary production (PP_{100}), primary production almost doubles. At its peak in autumn, total primary productivity from PP is 4.2×10^{-2} compared to PP_{100} 8.1×10^{-2} µmol P l^{-1} d^{-1}. The average

PP over the year is 1.9 x 10^{-2} µmol P l^{-1} d^{-1}, just over half of the PP_{100} mean of 3.6 x 10^{-2} µmol P l^{-1} d^{-1}.

5.3.7. Modelled Export

As an initial cautionary note, grazer losses in the model (ΦG), are not explicitly export in the sense of material sequestered at depth. They include contributions from faecal matter, excretion, messy eating, mortality and consumption by higher predators. Hence here ΦG can only be used as a first order proxy for export and not as a direct estimate.

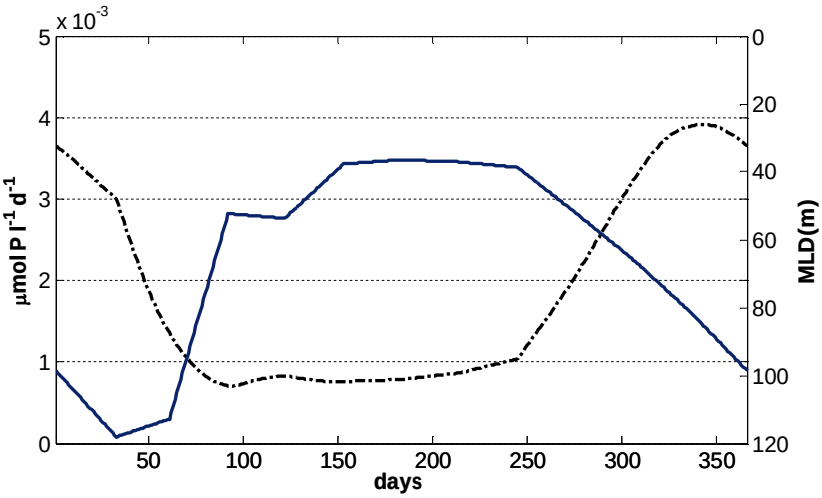


Figure 5.15. Export (ϕG), dashed line, primary *y*-axis. MLD, solid line, secondary *y*-axis. Note, timescale is Julian day.

Export (annual average, 1.8 x 10^{-3} µmol P l^{-1} d^{-1}, Figure 5.15) is low in comparison to the balanced larger fluxes in the model (e.g. αPB mean 1.9 x 10^{-2} µmol P l^{-1} d^{-1}), yet high amongst the important unbalanced smaller fluxes (e.g. θMG mean 2.4 x 10^{-4} µmol P l^{-1} d^{-1}, see all fluxes in Figure 5.7). Export varies between 0.7 x 10^{-3} and 4 x 10^{-3} µmol P l^{-1} d^{-1} throughout the year (see Figure 5.15). In comparison, the equivalent range for phosphate input into the ML by deepening and mixing is 0.4 x 10^{-3} to 6.9 x 10^{-3} µmol P l^{-1} d^{-1} with a mean of 2.5 x 10^{-3} µmol P l^{-1} d^{-1}. This is higher than export because of 'mixing' losses on B, G and M (Figure 5.1). Export peaks in the late autumn (at the same time as G and slightly after M peaks) and is depressed throughout the summer. The modelled export is dependent upon the grazer state variable, and therefore tightly follows its yearly cycle (see Figure 5.4). If a conversion to carbon is made using the Redfield ratio (C:P, 106:1)

(Redfield 1934), the mean export is 0.2 µmol C l^{-1} d^{-1} (~ 70 µmol C l^{-1} y^{-1}). This will be discussed further in the next Section.

5.4. Discussion

A simple dynamic model including mixotrophy has been constrained to limited observations from the oligotrophic north Atlantic. A µGA technique was used to optimise parameters using ranges defined from the literature.

5.4.1. Model Performance

5.4.1.1. General Performance

The optimisation process successfully generated parameters to obtain a fit to the *in situ* observations that could be classified as good to very good and enabled coexistence of all variables. It might *a priori* be thought trivial to fit 4 data points to 4 variables. However this is clearly not the case. Calibration of the model was challenging and no solution was found that matched all data simultaneously, despite there being only one observation for each variable. Only nine parameter sets arising from 150 optimisations provided a cost less than 100 and out of these, only two had a total cost of less than 10. This is despite the optimisation running for double the number of generations of previous studies using µGA, such as Ward et al. (2010). In theory it should have been easier to optimise a model to fewer observations, as the model output has less points to fit through.

The modelled inorganic phosphate had the highest cost of 1.98, when compared to the observation. Although this is still a good fit it was 33 % lower than the observation, with the model output not even being within the uncertainty estimate of the observation (see Figure 5.4). This may be due to the experimental error used on this measurement in reality being significantly larger (the likelihood of this is unknown), and an increase in uncertainty would have led to a lower cost. In retrospect the uncertainty of 5 % employed for the other data observations (B, G and M cell ml^{-1}) may also have been too small. The conversion of model output to cell ml^{-1} (Equation 5.9) utilising the cellular phosphate content, necessary to calculate the cost function, will also have introduced uncertainty, due to the range of possible cellular P contents (Table 5.1). However, only one could be

used. An equal number of optimisations were undertaken with minimum, maximum and mean cellular P quotient to attempt to account for this, allowing the model the chance to optimise to the complete spread of phosphate cellular content data (see Section 5.2.3.3). The differences between these estimates from the three optimisations were more than the 5 % uncertainty used in the cost function calculation (Table 5.1), with the standard deviation being ~20 % of the mean for M and B and 106 % for G. Conversely the best optimisation achieved was using maximum cellular P content. This was somewhat unexpected, as within the oligotrophic regions, 'survivalist' algae are expected to dominate. These have a high N:P ratio (> 30). They can sustain growth when resources are low as they contain a lot of resource-acquisition machinery within the cell (e.g. pigments/proteins which have a high N:P ratio) (Arrigo 2005). This model success of high P content cells may be explained by the cellular phosphate content taken from the literature being based primarily on oligotrophic measurements. Regardless a larger spread of data (both for estimates of phosphate cellular content and cell abundance) and refined associated uncertainties would be beneficial to future optimisations.

The difficulty shown in fitting even this very small dataset, means that the model may not have been appropriate for the system, as parameter sets with a sufficiently low cost were difficult to obtain. This problem may have arisen from potential deficiencies in the model structure, such as oversimplification (Schartau and Oschlies 2003a), in terms of the number of model variables. For example there is no obligate autotrophic algae in the model. This group is included in the model presented by Thingstad et al. (1996) to study mixotrophy. However, the Thingstad et al. (1996) model failed to allow coexistence of all variables, whereas the model structure presented here did. Another alternative is to refine the model by making it one-dimensional (1D). A 1D model would allow vertical processes to be better resolved, for example, it would no longer be necessary to assume there is no life below the ML. The further step of including horizontal spatial variability would allow advective effects to be included. Heterogeneity of ultraplankton will be investigated in the next Chapter. The main challenge in increasing the complexity (i.e. number of parameters) of the system is the increased need for data to constrain the model. *In situ* data are simply at present not available in sufficient quantities for the oligotrophic Atlantic.

The cost function should be interpreted with caution, as the mismatch between the low number of observations used to calculate the cost function and the large number of data points from the model can reduce the reliability of its result (OSPAR 2008). More data

points at different times of year would help to rectify this (whilst making optimisation even more challenging). Unfortunately these were not available for this study.

5.4.1.2. Seasonal Cycle

The timing of the peak in mixotrophs in autumn in the model is consistent with previous research in the oligotrophic north Atlantic region. SeaWiFS chlorophyll maxima have been observed over autumn and winter (Siegel et al. 2002, Henson et al. 2009) and previous models (that exclude mixotrophy) have estimated that in the subtropical regions (south of 45°N) blooms occur in autumn or winter (Dutkiewicz et al. 2001).

Although the model timing appears to be broadly correct, the amplitude of the seasonal cycle model variables varies between microbial groups (e.g. bacteria does not demonstrate seasonality, but mixotrophs and grazers do, see Figures 5.4 and 5.5). The oligotrophic regions have previously been described as relatively constant throughout the year (Lutz et al. 2007, Cole et al. 2012). This is also supported partly by the results of this Chapter (the bacteria variable has a weak seasonal cycle) and by the result presented in Chapter 3 of this thesis (no significant difference between measurements of dissolved organic phosphate in opposing seasons). A recent study has used a low coefficient of variation in satellite sensed chlorophyll to characterise a region as having a weak seasonal cycle (< 0.35, Cole et al. 2012). The model output can be assessed using the same criteria. The coefficient of variation over the model output year was almost double the seasonality criterion for the phosphate, grazer and mixotroph variables (0.63, 0.64 and 0.65 respectively). For bacteria it was much lower (0.07). Here however the coefficient of variation is calculated using the model variables. In the aforementioned study, the coefficient of variation was calculated using the remotely sensed chlorophyll *a* concentration. Given the varying seasonal response, it is therefore necessary to decide which organism is of primary interest. If it is mixotrophs then comparison of model to data requires that satellite chlorophyll *a* be closely related to the abundance of pigmented organisms represented by the model; the mixotrophic algae and the bacteria, the latter of which encompasses phototrophic cyanobacteria. There is however some evidence that satellites may not accurately detect ultraplankton abundances (Zubkov and Quartly 2003). Therefore the ultraplankton abundance coefficient of variation in the oligotrophic gyres may be higher than the ocean colour data suggests for mixotrophic algae. The reliability of remotely sensed data for measuring ultraplankton and variability in ultraplankton will be addressed in Chapter 6.

Ideally, to test if the amplitude and the timing of the model's yearly cycle is correct, *in situ* monthly observations could be obtained to constrain the parameters. However due to field expenses, logistics and methodological limitations this is not practical. Furthermore, spatial variability may cause an additional problem. Large variability was seen between CTD points ~ 1400 km apart (CTD14 and CTD23). Despite the oligotrophic Atlantic being consistent with one microbially defined region of > 90 % similarity (see Chapter 2), mixotroph cell numbers were over an order of magnitude higher at CTD14 (11,942 cell ml^{-1}) than at CTD23 (2,748 cell ml^{-1}). One explanation for this, is that if (returning to Chapter 2) similarity was increased to 91%, it is evident that CTD14 and CTD23 would be in different sub-clusters (Figure 2.3a and Figure 4.1). Therefore as the model is optimised to CTD14, the mixotroph model results may be more indicative of the subtropical north Atlantic gyre (cluster region *aii*) than the equatorial and southern Atlantic (cluster region *ai*). This has been taken into consideration when comparing the model output with previous studies model output and observations. If computing time was not a limitation (Section 5.2.4 indicated 2,400 hours to optimise to CTD14), it would be interesting to also optimise the model to CTD23. Therefore ideally any future monthly observations may be better limited to the vicinity of the same location. However there may still be variability at smaller scales, e.g. due to eddies, which could give anomalous observations at a site. The issue of spatial variability in microbial group abundances will be discussed further in Chapter 6.

5.4.1.3. Parameters

The parameter ranges from which the optimisation procedure could assign parameter values were intentionally large (Table 5.2) to ensure any potential values were not omitted. The optimised parameter value for α (direct phosphate uptake by bacteria) and *Po* were the only parameter values towards the higher end of their literature defined range (see Table 5.5). The majority of the other parameters were low within their literature defined range. This was true even for ε (remineralisation from mixotrophs to phosphate) and γ (bacterivory by mixotrophs), despite being associated with the largest fluxes within the system (see Figure 5.7). The model therefore implies that the parameter ranges described from the literature may be higher than *in situ* oligotrophic ocean observations. This might be taken to suggest that the literature ranges had to rely heavily on culture values due to the lack of *in situ* data and as a result are too broad. In the future conducting further i*n situ* grazing and uptake experiments would help to narrow the parameter ranges. An alternative explanation is that limitations in the model structure are

pushing parameter values to the edges of their ranges through inadequately representing some processes.

The optimised model was particularly sensitive to a number of parameters, relatively small independent variations in a few causing mixotrophs to become extinct during a ten year model run (Figures 5.9 and 5.10). The sensitivity analysis results show that the model is mainly sensitive to grazing of bacteria by mixotrophs (γ) and grazers (ψ) and to a lesser extent to remineralisation from mixotrophs to the phosphate pool (ε) and export of phosphate out of the model through grazers (ϕ). Two of the most sensitive parameters (γ and ε) are associated with some of the largest phosphate fluxes within the model (εM and γ BM, see Figure 5.7). When γ is decreased by 10% or more or ε is increased by 10% or more, mixotrophs become extinct within the system. This is an indication of the sensitivity of the mixotroph variable, as well as demonstrating the importance of bacterivory by mixotrophs within the system. It also helps to explain why optimisation was so difficult (9/150 optimisations allowed coexistence with a cost function, $J \leq 100$) as these two parameters (responsible for a large proportion of the cycling in the system) are extremely sensitive. As for the model's closure term (ϕ), the form and the value of the model's closure terms have been shown previously to strongly influence model dynamics (Steele and Henderson 1992a, Edwards and Yool 2000). Further measurements for model parameterisation (especially to describe the most sensitive parameters, e.g. γ, ψ, ε and ϕ) would help to rectify this, by reducing the uncertainty in the amplitude of the model variables at other times of year, thus aiding the model calibration.

5.4.1.4. Comparison to Fluxes

The optimised model output has been independently compared to the observational flux data presented in Chapter 4 (Section 4.4.1). Figure 5.11 shows the model output fit to the data. As previously stated, two model flux lines are presented, the solid line is gross, and the dashed line net flux, as it is not possible to state explicitly what processes are occurring during the timescale of the bioassay experiment (< 1 hour) used to obtain microbial group uptake rates. For example regarding Figure 5.11b, it is not known if significant grazing of bacteria (by M or G) is occurring within the same time frame as the uptake of P by bacteria. In reality the answer is probably somewhere between the two (net and gross) and may be different for each variable.

Figure 5.11a shows the uptake flux of P into M. It shows that the independent measurement (F1) is quite well fitted to the modelled net flux (dashed line), but not to the modelled gross flux. The grazer predation on mixotrophs is relatively small within the model (Figure 5.7h). If this is excluded from the modelled net flux calculation, the fit to the flux observations is very similar. Therefore it is the exudation of P back to the P pool which dominates the difference between net and gross, but unfortunately little is known of this process. Uptake of P by bacteria (F2), is the only flux observation that fits the gross flux observation, indicating that uptake of P by bacteria is accurately captured during the experiment, assuming that the model is correct. The P flux into grazers (F3) is the worst fit to the model output, with the independent flux measurement below even the modelled net flux estimation. This may be due to the majority of predation undertaken by the grazers occurring over longer time scales than the experiment (> 1 hour) used to make the observations, perhaps because of the relatively low grazer abundance (Table 4.3) causing a lower encounter rate with bacteria and mixotrophs. The independent comparison of the model with flux observations, therefore might indicate that contributing processes are caught occurring over different timescales during the experiment. Unfortunately, at present, it is not possible to separate and measure these fluxes simultaneously. Future methodological advances will hopefully rectify this, and clearly limitations of the model should not be ruled out.

5.4.1.5. Modelled Bacterivory

The optimised model enables an estimate of the amount of mixotrophy taking place within the system throughout the year. Mixotrophy is a dominant nutritional strategy within the system and this can be clearly demonstrated. The percentage bacterivory undertaken by algae (mixotrophs) is high (ranging between 78 and 95 % throughout the year) and mixotrophs obtain the majority (97 – 100 %) of their phosphate in this way (Figure 5.12) and thus have less dependence on dissolved inorganic phosphate. The percentage of bacterivory by mixotrophs is slightly higher (78 – 95 %) yet consistent with previous oligotrophic regions *in situ* observations, such as 37 - 70 % for the north Atlantic gyre, (Zubkov and Tarran 2008), 60 – 77 % across the subtropical and tropical Atlantic ocean (Hartmann et al. 2012), 35 – 65 % in the Mediterranean Sea (Unrein et al. 2007) and 40 % off the southwest coast of New Zealand (Hall et al. 1993). This model therefore supports previous observations of the importance of mixotrophy in oligotrophic regions. The model also shows that bacterivory is the main nutrient strategy even when nutrients are not strongly depleted (see Figures 5.12 and 5.13b), as bacterivory by mixotrophs is highest in

the model when phosphate concentration is at its highest. This suggests that mixotrophy is not a survival mechanism as previously suggested (Nygaard and Tobiesen 1993), but a competitive mechanism within the system. This model therefore shows that mixotrophs are not strictly the Type IIa mixotrophs, that is 'phagotrophic when nutrients are limiting' (Stoecker 1998), as they were thought to be. Perhaps another Type is required in Figure 1.2; Type IId, 'phagotrophic for an ecological advantage'?

5.4.1.6. Modelled Primary Production

Primary production in the model peaks in the autumn, following the cycle of mixotrophs, which are major contributors to it (see Figures 5.4 and 5.13). To enable comparisons to the literature to be made the Redfield Ratio (C:P, 106:1) is used (Redfield 1934). The Redfield ratio will also be used for carbon export comparison. Primary production in the model, excluding any bacterivory from mixotrophs, ranges between 5.79 and 192.78 mol C m^{-2} y^{-1} over the year. Including varying amounts of phosphate obtained from bacterivory, the values range from 6.77 (PP_{20}) to 369.12 (PP_{100}) mol C m^{-2} y^{-1} (see Figure 5.14). These estimates are much higher than the Oschlies and Garcon (1998) coupled ecosystem-circulation model (without mixotrophy) result of 0.26 to 0.62 mol C m^{-2} y^{-1} for the north Atlantic subtropical gyre. However, the lower end of the model output range is closer to *in situ* measurements of 8.02 mol C m^{-2} y^{-1} for the north Atlantic gyre (Marañón et al. 2003) and from the euphotic zone near Bermuda of 3.7 mol C m^{-2} y^{-1} (Jenkins 1988).

5.4.1.7. Modelled Export

The loss rate of grazers is used as the proxy for export (ΦG), with caveats already described. The mean modelled export is 32.89 mmol P m^{-2} yr^{-1} (see Figure 5.15) This is higher than the modelled particulate phosphorus export in the north Atlantic subtropical gyre of 5 – 12 mmol P m^{-2} yr^{-1} (Torres-Valdés et al. 2009). The discrepancies may be explained by the fact that the Torres-Valdés et al. (2009) model is (unlike this model) a 3D physical-nutrient model based on the model of Roussenov et al. (2006), and does not explicitly model the microbial community.

Carbon export can be estimated as previously for primary production by using the Redfield ratio. The model's carbon export equates to a mean of 3.5 mol C m^{-2} y^{-1}. This is slightly higher than observational estimates. For example in April/May time, Thomalla et al. (2006) calculated zero carbon export in the oligotrophic north Atlantic from the water

column $^{234}Th/^{238}U$ disequilibria, but they observed export of the same order of magnitude as that presented here in the oligotrophic south Atlantic (2.19 mol C m^{-2} y^{-1}). This mixotroph model also estimated the lowest export in April-May time (Figure 5.15). Therefore it is feasible that export may have been below their detection level. Richardson and Jackson (2007), using observations and network analysis, reported that 1.83 to 4.38 mol C m^{-2} y^{-1} is exported directly by picoplankton throughout the year in the tropical and subtropical Pacific. It should, however, be noted that Redfield ratios within microbial groups have been well documented to be variable. For example C:P ratios within picoeukaryotes can range between ≈ 70 and ≈ 200 (Arrigo 2005). Furthermore this range in stoichiometry is much wider in nutrient limited cells (Geider and La Roche 2002). Unfortunately, DNA analysis was not undertaken on this study's samples, so an estimate according to species composition cannot be made for export or primary production. Despite the nutrient cycling through the grazer variable in the presented model being relatively low in comparison to the other fluxes (e.g. ϕG, export mean flux 1.8×10^{-3} compared to γBM mean flux of 1.2×10^{-2} µmol P l^{-1} d^{-1}, see Figure 5.7), the export estimate remains similar and in some cases higher than other estimates from models and observations.

5.4.2. Comparison to Previous Models: Incorporating Mixotrophy

For the model presented here, mixotrophy is a significant link in the phosphorus cycle, as the majority of bacterivory within the system is being performed by mixotrophs (see Figure 5.7). This extra trophic link (in comparison to a standard Nutrient-Phytoplankton-Zooplankton model) reduces the amount of phosphate being passed on to higher trophic levels (the average grazing of mixotrophs, θMG, is 2.5×10^{-4} µmol P l^{-1} d^{-1}), as a large amount of the phosphate is recycled back around the system through the bacterivory performed by mixotrophs and ensuing remineralisation (average εM is 1.7×10^{-2} µmol P l^{-1} d^{-1}, Figures 5.7 and 5.8).

The theoretical Baretta-Bekker et al. (1998) model also presented mixotrophy as a significant link in the carbon and phosphorus budgets with mixotrophs responsible for more than 40 % of the grazing of bacteria in low nutrient waters. They concluded that explicitly including mixotrophs within a model, rather than incorporating them within the heterotrophic or autotrophic functional groups, actually increased total primary

production levels in waters of oligotrophic status, as primary production was supported by nutrients obtained through bacterivory. Hammer and Pitchford (2005) also theoretically modelled mixotrophs, in that case as a fraction of the predation in the system being involved in primary production. Their simple model also found that a system becomes more productive in the presence of mixotrophy.

In this Chapter we cannot explicitly state that mixotrophy is increasing primary production, but if even a small percentage of the phosphate obtained through bacterivory was being utilised for primary production, it would substantially increase the inferred primary production (see Figure 5.14). The Stickney et al. (2000) theoretical model (a 5-box model with a dissolved inorganic nitrogen variable, obligate phytoplankton and zooplankton, detrital pool and mixotrophs), however, suggests that the inclusion of mixotrophs decreases overall net primary production, measured by dissolved inorganic nitrogen uptake. Stickney et al. (2000) does state that the total photosynthetic rate may be maintained through direct recycling of organic nitrogen (i.e. predation on phytoplankton and mixotrophs themselves). Yet, the Stickney et al. (2000) model does not include bacteria as a food source, with mixotrophs instead predating on phytoplankton. Mixotrophic bacterivores are recognised to dominate the oligotrophic Atlantic region (Zubkov and Tarran 2008, Hartmann et al. 2012).

A number of previous theoretical mixotroph models have also suggested that mixotrophs act as a stabilising link in the ocean planktonic system (Thingstad et al. 1996, Stickney et al. 2000, Jost et al. 2004), as their grazing on autotrophs reduces the competition for nutrients. This is also seen in the present model, in the relative stability of bacteria throughout the year and the coexistence of all the microbial variables. As the bacterial community is under continued grazing pressure, it does not appear to be able to take advantage of the increase in phosphate concentration (see Figures 5.4c and 5.7e-f). This supports the steady state assumption made in Chapter 4 for bacteria, as for bacteria microbial growth approximately equals death (Fuhrman and Hagström 2008). Thingstad et al. (1996), however, used a model to argue that mixotrophic persistence (and thus stability) occurs only when mixotrophs have a high affinity for nutrient uptake (i.e. phosphate) combined with an intermediate affinity for bacteria. In this Chapters model almost the opposite is demonstrated, with coexistence and stability in bacteria when mixotrophs have a high affinity for bacteria and a low affinity for phosphate uptake (γ, grazing on bacteria by mixotrophs, is approximately four times larger than π, P uptake by mixotrophs, see Table 5.5). The difference in results may arise from Thingstad et al.

(1996) having an additional and obligate autotroph (phytoplankton) in their system, whereas here we assume all the picoeukaryotes have the ability to be mixotrophic (supported by the work of Zubkov and Tarran 2008).

5.5. Summary and Implications

A simple zero-dimensional model including mixotrophy has been used to simulate a yearly cycle in the ecosystem of the north Atlantic oligotrophic gyre in 2005. The model gives a good agreement with the very limited observational data. It has been shown that recycling between mixotrophs, bacteria and the phosphate variable is a dominant process within the model (see Figures 5.7 and 5.8). This is an important process as fast nutrient turnover is crucial to sustaining the oligotrophic ecosystem. In addition, this optimised model supports previous evidence (Zubkov and Tarran 2008, Hartmann et al. 2012) that the majority of phosphate obtained by mixotrophs is through bacterivory. The model further indicates that bacterivory by mixotrophs need not be a survival mechanism (Nygaard and Tobiesen 1993), as they undertake it in significant quantities throughout the year, even when phosphate concentrations are at their highest (> 95 %, see Figure 5.12). Primary production estimates, including bacterivory by mixotrophs is slightly higher than suggested by previous models and observations (see Figure 5.13 and 5.14), an indication that it may be important to include mixotrophy in future ecosystem models of the region. Modelled export is similar, although the upper range is slightly higher than previous observed and modelled estimates for this region (see Figure 5.7 and Figure 5.15).

From the results presented in this chapter and considering the stated hypothesis to be tested the following can be concluded:

- A simple zero-dimensional dynamic mixotroph model was created that maintained variable coexistence and was consistent with *in situ* abundance observations from the north Atlantic oligotrophic gyre.

6. Ultraplankton Spatial Variability – Flow Cytometry and Remotely Sensed Ocean Colour

6.1. Introduction

Spatial variability in phytoplankton is not a recently discovered phenomenon. Since the time of Captain James Cook's *HMS* Resolution voyage (in 1773) localised patches of coloured water have been observed and attributed to microscopic organisms (Bainbridge 1957). More recently satellite images of ocean colour and ship-borne surveys have affirmed the 'patchiness' of phytoplankton distributions (see for example review by Martin, 2003). However remotely sensed images only infer indirect estimates of chlorophyll concentrations, that in particular may not reflect ultraplankton (plankton < 5 µm) abundances (Zubkov and Quartly 2003). Strong spatial variability of ultraplankton has, however, been observed through *in situ* flow cytometry measurements in shelf (Martin et al. 2005, Martin et al. 2008) and temperate seas (Martin et al. 2010), but little is currently known about the degree and manner of ultraplankton spatial variability in tropical and subtropical regions that cover > 60 % of the ocean surface (Longhurst et al. 1995). There is some evidence that these small important organisms display as much spatial variability as other larger planktonic organisms (Zubkov et al. 2002, Martin et al. 2005). The subtropics have been previously considered homogenous and stable habitats, but increasingly research has shown that these environments actually display considerable variability on a variety of time and space scales (e.g. Karl 1999, Marañón et al. 2003, McClain et al. 2004). Yet looking at the remotely sensed chlorophyll *a* (chl *a*) concentration in Figure 6.1 this spatial heterogeneity seems hard to discern.

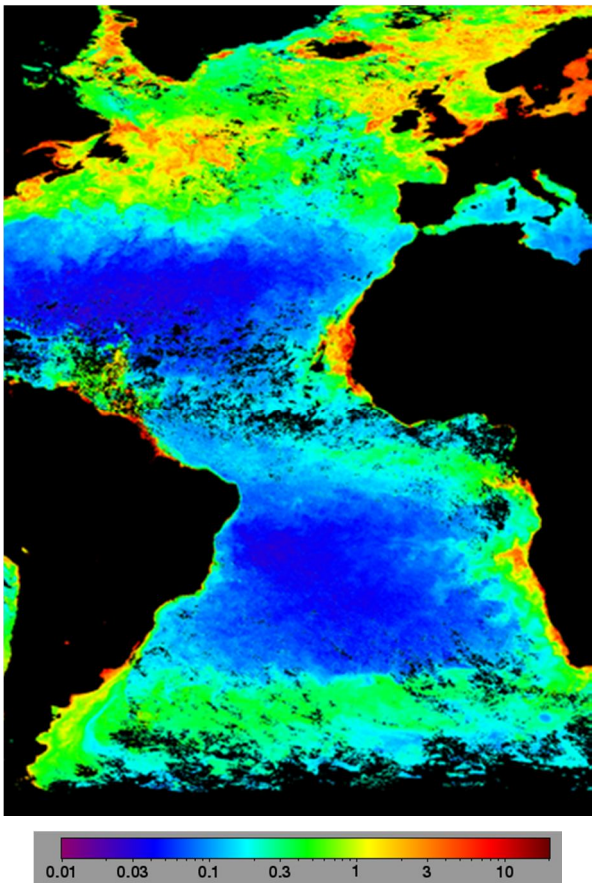


Figure 6.1. An ocean colour image of chlorophyll concentration (colour key in mg m^{-3}) from the SeaWiFS satellite. Temporal resolution is a 32-day composite (30th April – 31st May 2004) and spatial resolution is 9 km, Source, oceancolor.gsfc.nasa.gov.

The issue of ultraplankton variability is not only key to ocean primary productivity and global cycling estimates, but also has relevance in relation to previous assumptions made in this thesis. One of the model assumptions was of steady state and stability of the Atlantic Oligotrophic gyres (Chapter 4). This assumption was partly based on Chapter 2's multivariate analysis of ultraplankton abundances, the result in Chapter 3 of no significant difference in DOP between seasons, and supported by remotely sensed ocean colour images (such as that in Figure 6.1), which suggest homogeneity (in terms of chl *a* concentration) throughout the gyres (see Section 4.1.2 for justification of steady state). This Chapter will use a high resolution *in situ* sampling dataset to examine ultraplankton variability in the tropical and subtropical Atlantic Ocean and also assess the suitability of using satellite data for estimating the concentration and variability of ultraplankton groups in the surface waters.

Spatial variability at the mesoscale (1 to 100 km) is difficult to measure *in situ*, as ship and time-series sites are often spatially and/or temporally limited and advanced sampling and

measurement techniques (such as flow cytometry) are time consuming, yet required to observe the smallest scale organisms (ultraplankton, < 5 μm) in the oceans. Until recently the degree of variability in abundance around a station or along a transect was rarely measured, as the assumption was made that local heterogeneity was not as great as daily or seasonal variability. Studies in the Celtic sea (Martin et al. 2005, Martin et al. 2008) and at the temperate Porcupine Abyssal Plain (PAP) site (Martin et al. 2010) have disputed this, showing strong spatial variability. They also indicate that different ultraplankton groups may be spatially distibuted in significantly different ways from one another. For example, strong interactions (e.g. between a predator and a prey) may not manifest itself as a strong spatial correlation. Thus spatial variability needs to be regarded as a potential source of error when analysing time series and transect data, especially where averaging or extrapolation for estimates have been employed. Also estimates (such as for export or primary production) derived from models of the region that assume homogenity (such as those presented in Chapters 4 and 5) may not be appropriate, if horizontal variability is present. In previous Chapters of this thesis the tropical and subtropical Atlantic ocean has been considered as one region (cluster group *a* determined in Chapter 2, and used to analyse dissolved organic uptake in Chapter 3 and build models in Chapters 4 and 5). The definition of this cluster did not probe spatial variability within it. However, previous evidence for spatial variability in shelf and temperate seas suggests that to have faith that uniformity can be assumed when building a model this assumption should be tested. Therefore results presented in this chapter attempt to examine this conjecture.

The population densities of phytoplankton can be remotely measured as ocean colour by satellites. 'Chlorophyll *a* concentration' is estimated as the sum of chl *a* and pheophytin *a*, which is used as a proxy for phytoplankton (algae and cyanobacteria) biomass (Martin 2004). Picoeukaryotes (PicoEuk) and cyanobacteria (*Pro* and *Syn*) contribute equally to ultraphytoplankton biomass (Zubkov et al. 2000). Therefore it would be reasonable to expect that these organisms (algae and cyanobacteria) would both affect the surface water's light attenuation and reflection (Morel et al. 1993) and therefore that their presence and variability can be observed by ocean colour satellites (see Figure 6.1.). However, despite ultraphytoplankton dominance and influence in tropical and subtropical oceanic waters (Li and Harrison 2001, Marañón et al. 2001) there has only been one study to the author's knowledge that has compared satellite measurements with *in situ* ultraplankton abundance data. This research, in the Mozambique Channel, compared flow cytometry data with SeaWiFS daily and weekly chl *a* composites. Zubkov and Quartly (2003) reported a high correlation of *Syn* chl *a* concentration and satellite data, but a poor

correlation to *Syn* abundance and biomass and found no relationship at all between *Pro* abundance or chl *a* concentration and satellite data. One conclusion drawn is that SeaWiFS may not be detecting chl *a* from *Pro*, which is believed to be the most abundant photosynthetic organism on Earth (Partensky et al. 1999b) and a significant contributor to primary production (Goericke and Welschmeyer 1993, Li 1994, Liu et al. 1997). This potentially means a significant underestimate in oceanic primary production from calculations based on remotely sensed chl *a* (e.g. Behrenfeld et al. 2005). More specifically *Pro* is the numerically dominant phototroph in the subtropics

This Chapter addresses the following hypothesis:

- The ultraplankton community in the surface waters of the subtropical and tropical Atlantic are not as spatially variable as in shelf and temperate seas and so homogeneity can be assumed when building a model.
- Different ultraplankton groups in the surface waters of the subtropical and tropical Atlantic, do not vary at significantly different spatial scales
- Remotely sensed satellite chl *a* concentration can be used to estimate the abundances, biomass or chl *a* of phototrophic ultraplankton (*Pro*, *Syn* and PicoEuk) in the surface waters of the tropical and subtropical Atlantic (cluster region *a*)

6.2. Method

6.2.1. Data collection

Ultraplankton samples were collected, as stated in Chapter 2, by M Zubkov's team along a northbound transect from the Falkland Islands to the UK (~ 40°S to 49°N) on AMT14 (part of the Atlantic Meridional Transect Programme) from 28th April to 1st June 2004 on-board *RRS* James Clark Ross, in boreal spring and austral autumn. The following ultraplankton groups were sorted and enumerated (see Zubkov et al. 2007 for sampling and microbial enumeration protocol) and will be discussed in this Chapter: *Prochlorococcus* (*Pro*) spp., *Synechococcus* spp. (*Syn*), picoeukaryotes (PicoEuk) and heterotrophic bacteria that was separately characterised as low DNA content bacteria (LNA), high-nucleic acid bacteria with a low 90^{o} light scatter (HNAls) or high-nucleic acid bacteria (HNAhs) with a high 90^{o}

light scatter. Beads of 0.5µm diameter were used as an internal standard of red fluorescence, which was consequently used as a substitute for mean cellular chlorophyll (chl *a*) content (Li 1995). Biomass was calculated using the following approximations multiplied by the sample abundances of each microbial group: 29 fg C per cell for *Pro*; 100 fg C per cell for *Syn*; 11.5 fg C per cell for heterotrophic bacteria (LNA, HNAhs, HNAls and ttl) and 1.5 pg C per cell for PicoEuk (Zubkov et al. 2000). Previous studies conducted in the oligotrophic Atlantic gyres have found consistent estimates of biomass through the size fractionation method (Zubkov et al. 1998, Zubkov et al. 2000). In addition, to confirm constituent weights were applicable for this study, variance of size scatter (an indirect size measurement) was calculated and found to be very low for all groups (variance < 0.005), except for PicoEuk were it was moderately low (variance < 0.5). As biomass was a simple conversion from abundance for *Syn, Pro* and PicoEuk (i.e. multiplied by a constant weight per cell) it will only be discussed in terms of total measured ultraphytoplankton biomass (summed *Syn, Pro* and PicoEuk). A thermosalinograph (Sea-Bird, SBE45) was used underway to measure sea temperature and salinity at a depth of 6 metres (calibrated against CTD data).

To investigate the accuracy of estimating ultraphytoplankton abundances from ocean colour, two satellites are used in this study. The SeaWiFS (Sea-viewing Wide Field-of-view Sensor) launched in 1997 gathered data for thirteen years before failing. MODIS (or Moderate Resolution Imaging Spectroradiometer) was launched aboard the Aqua (EOS PM) satellite in 2002 (Esaias et al. 1998) and is still running today. The sea surface was imaged approximately once daily by SeaWiFS at a 1.1 km resolution in 8 frequency bands spanning 412 to 865 nm (Esaias et al. 1998, Robinson 2004). MODIS images the entire earth every 1 to 2 days and utilises 9 frequency bands from 411 nm to 866 nm at a 1 km resolution (Esaias et al. 1998). In this study, daily and weekly (8 days) and monthly (32 days) composites of chl *a* concentrations computed by the NASA Goddard Space Flight Centre at resolutions of 4 km (MODIS) and 9 km (SeaWiFS and MODIS) are used, see review in Robinson (2004) for further details. Data were downloaded for the time period of the cruise (26th April to 2nd June 2004) from http:/oceancolor.gsfc.nasa.gov/.

6.2.2. Data analysis

The analysis performed in this Chapter focuses solely on those samples defined in Chapter 2 as in cluster group *a* (> 90 % similarity, *n* = 566) which encompassed the north and

south Atlantic gyres and the equatorial region (see Chapter 2, Figure 2.3). As in Chapter 2, all abundance, chl *a* content and biomass per ml^{-1} data were normalised, then fourth root transformed to improve gaussianity. Temperature, salinity and satellite chl *a* were square root transformed prior to analysis (see Section 2.2.2 for explanation). The coefficient of variation was calculated without transforming the data, as this was used to assess if variability was present within cluster group *a* and transforming it would dampen this variability. Running coefficient of variations were calculated (over 25, 50 and 100 concurrent data points) using MATLAB.

Spearman's rank correlation coefficients were calculated in GraphPad Prism 5 for each ultraplankton group. The following interpretation of strength of correlation was used – < 0.2 very weak to negligible; 0.2 – 0.4 weak, low correlation; 0.4 - 0.7 moderate correlation; 0.7 – 0.9; strong correlation; > 0.9 very strong correlation (Fowler et al. 1998). This scale is also used in the interpretation of the correlation of satellite data and ultraphytoplankton abundances.

6.2.2.1. Autocorrelation

Sample autocorrelation measures the correlation between the same series of observations (e.g. ultraplankton groups) offset by a given distance (Chatfield 2004). This was used to assess length scales of variability for each of the groups and to examine if the ultraplankton groups had similar dominant length scales. Autocorrelation also gives an indication of the appropriateness of sample frequency on the AMT transect (average spatial resolution 18 km), by comparing it to full resolution satellite data (4 km and 9 km).

The autocorrelation coefficients were calculated for all sample variables for group *a*: the six ultraplankton groups, temperature, salinity and satellite chl *a* concentration (SeaWiFS and MODIS 4 and 9 km for 1, 8 and 32 day composites). It was necessary to interpolate the transect data to the regular intervals spaced by the mean separation between samples prior to computation, which may have had a smoothing effect. The sample autocorrelation was completed in MATLAB with 95 % confidence following the formula of Box et al. (1994).

6.2.2.2. Semivariogram analysis

Semivariograms are a measure of the amount of variability expected between two samples a given distance apart. This provides information on how a parameter (e.g. an ultraplankton group) varies spatially, such as fluctuating more at shorter than at larger length scales (Martin et al. 2008). To allow a clear comparison between semivariograms data (the ultraplankton groups, temperature, salinity and satellite chl *a* concentrations) are only used from positions where all variables were measured. Analysis was also carried out using matched satellite observations. Sample number, *n* were as follows for: 1 day, *n* = 24; 8 day composite, *n* = 224 and 32 day composites, *n* = 459.

Semivariograms were calculated for each ultraplankton group, temperature, salinity and 1, 8 and 32 day satellite chl *a* composites (MODIS 4 km and 9 km and SeaWiFS 9 km resolution) in Fortran 90. Semivariograms have been utilised rather than power spectra because of the unequal spacing of samples along the AMT14 transect (ranged ~ 10 – 20 km). To calculate the semivariograms a modified form of that detailed in Cressie (1993) was used (see Equation 6.1). This form reduces the effect of outliers.

Equation 6.1 $$\boldsymbol{\gamma(r) = [\sum_{ij}|Z(x_i) - Z(x_j)|^{\frac{1}{2}}/N(r)]^4/[0.457 + 0.494/N(r)]}$$

The semivariogram, $\gamma(r)$, is the cumulative sum of variability, at scales less than or equal to *r*. The field of interest is *Z* and the sum is over all pairs of points (x_i, x_j) that are *r* distance apart. $N(r)$ is the number of paired points (Cressie 1993). As it is a cumulative sum, the semivariogram should increase or remain static with increasing length scales.

A linear regression is then carried out on the log of the distance (r) and the semivariable (γ). Assuming the relationship $\gamma(r) = ar^b$, *b* is the slope of the regression line of $\ln(r)$ against $\ln(\gamma)$, whilst *a* is the intercept. Several maximum distance (100 to 200km) and resolution combinations (10 to 20 km) were tried and 40 km $\leq r \leq 100$ km, with 20 km resolution was found to be the best, giving highest r^2 for the regressions. The steeper the slope, the larger *b*, indicating a greater contribution to the variability from larger spatial scale. If *b* is equal to zero, there is no variability at spatial scales larger than that where the semivariogram becomes horizontal. Prior to semivariogram analysis, data were also de-trended by removing a multiple linear regression for time, latitude and longitude of the form $y = a_1 + a_2 * \text{time} + a_3 * \text{latitude} + a_4 * \text{longitude}$. This removed the effect of any large-scale structures.

A bootstrap method was used for robust estimations of the semivariogram. A number of data points equal to 90 % of the total dataset were drawn with repetition allowed. The calculation was repeated with 10,000 such 'bootstrapped' datasets. Semivariograms were not very sensitive to number of repetition of bootstraps or the number of points used (analysis was also conducted on 95 % and 90 % of the dataset bootstrapped 1,000 times and 95 %, 10,000 times, data not shown). Pairs of points separated by more than 1 day in time were excluded, to minimise the effect of water mass movement over intervening time. Semivariogram slopes of all data sets were calculated concurrently for each pair of variables and a *t*-test in MATLAB was used to calculate significance of the difference between the slopes.

6.2.2.3. Correlation coefficients

Spearman's rank correlation were performed for comparison of satellite chl *a* and flow cytometry data sets at 95 % confidence, using MATLAB and GraphPad Prism 5. Satellite measurements were compared only with matched daytime measurements of ultraplankton groups, to minimise the effect of diel variability. Night time data were removed according to the sunrise and sunset times detailed in the cruise report (Holligan 2004). Ocean colour was assumed not to change during the daylight period (Zubkov and Quartly 2003).

Daily, 8 day and 32-day composites of SeaWiFS and MODIS 4k and 9km resolution were compared to *Syn*, *Pro*, PicoEuk, phototrophic bacteria (*Syn* and *Pro*) and ultraphytoplankton (sum of *Syn*, *Pro* and PicoEuk) abundances, cellular chlorophyll (chl *a*) content (measured by flow cytometry FL3 measurement multiplied by abundance) and biomass (abundance multiplied by ultraplankton group weight, see Section 6.2.1).

6.3. Results

6.3.1. Spatial Variability

The region's bacterioplankton (LNA, HNAls and HNAhs) are observed in Figure 6.2 to be relatively stable in terms of abundance, in comparison to the phototrophic ultraplankton (ultraphytoplankton, - *Syn*, *Pro* and PicoEuk). The coefficient of variation across the whole

of the region (cluster region *a*) is highest in *Syn, Pro* and PicoEuk (1.21, 0.64 and 0.64 respectively, Table 6.1) and lowest in LNA (0.28, Table 6.1). Temperature and salinity also have relatively low coefficients of variation in comparison to ultraplankton (0.13 and 0.02, Table 6.1).

Table 6.1. The coefficient of variation for the entire cluster region *a* for each ultraplankton group, temperature, salinity and satellite chl *a* concentration (1D, daily data; 8D, 8 day composites; 32D, 32-day composites. See associated Figures 6.2 and 6.3

		Coefficient of Variation
	Syn	1.21
	Pro	0.64
	Pico	0.64
	LNA	0.28
	Hhs	0.31
	Hls	0.40
	Temp	0.13
	Sal	0.02
32 day	**Mod4**	0.51
	Mod9	0.51
	Sea	0.89
8 day	**Mod4**	0.49
	Mod9	0.48
	Sea	1.05
1 day	**Mod4**	0.46
	Mod9	0.48
	Sea	0.57

MODIS variability is approximately 0.5 for each temporal composite, however SeaWiFS is higher in 32-day (0.89) and 8-day (1.05) temporal composites, than in 1 day (0.57), this is likely due to the effect of smoothing.

Figure 6.2 shows the variability throughout the transect of the different variables. With salinity depressed at the equator and *Syn* and PicoEuk increasing slightly. Showing that even though they are in the same defined cluster region (Chapter 2), they are still variable (this is also shown in the coefficient of variation, Table 6.1).

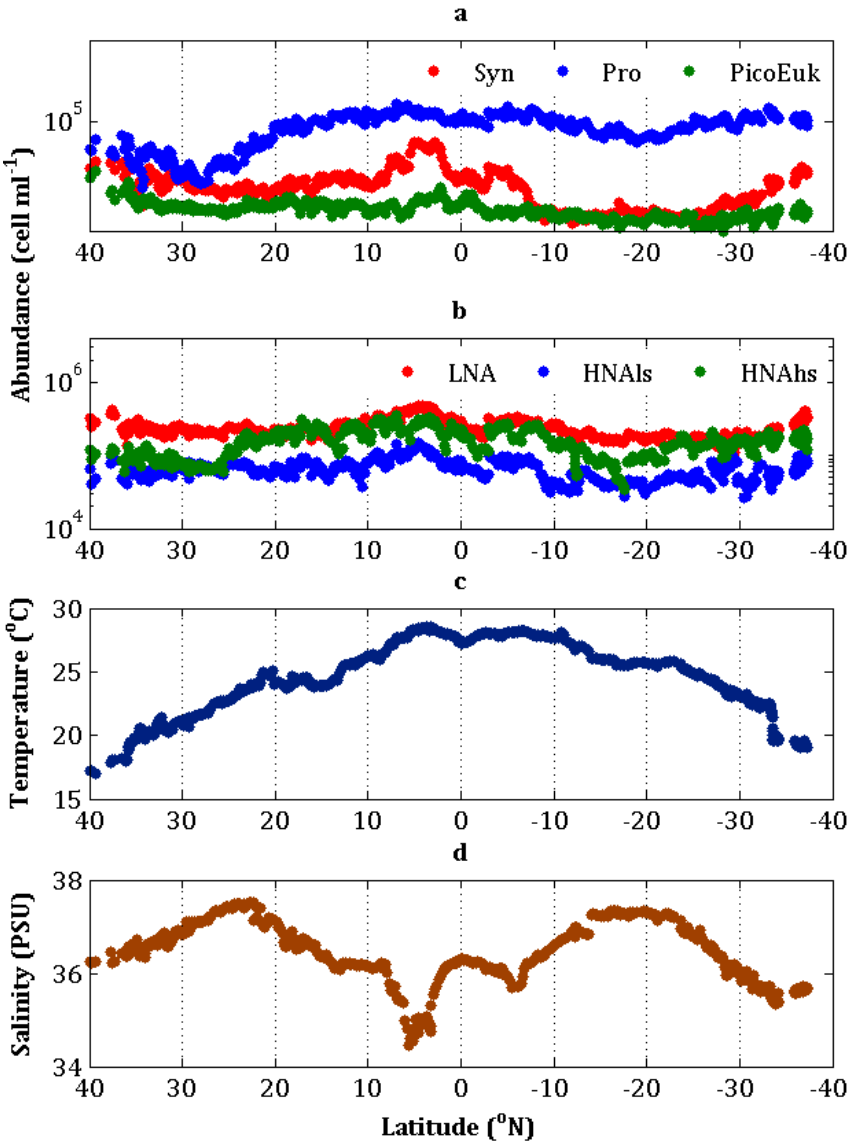


Figure 6.2. Latitudinal plot of (a) phototrophic ultraplankton abundances, (b) bacterioplankton abundances, (c) temperature and (d) salinity.

Figure 6.3 displays satellite data corresponding to the AMT14 transect. MODIS (4 and 9km) coefficient of variation is high (for 4 km and 9 km respectively – 0.46 and 0.48 for daily data; 0.49 and 0.48 for 8 day composites and both 0.51 for the 32-day composite), but not as high as for the ultraphytoplankton. SeaWiFS coefficient of variation is higher than MODIS and similar to ultraphytoplankton abundances (0.57 for daily data; 1.05 for 8 day composites and 0.89 for 32 day composites).

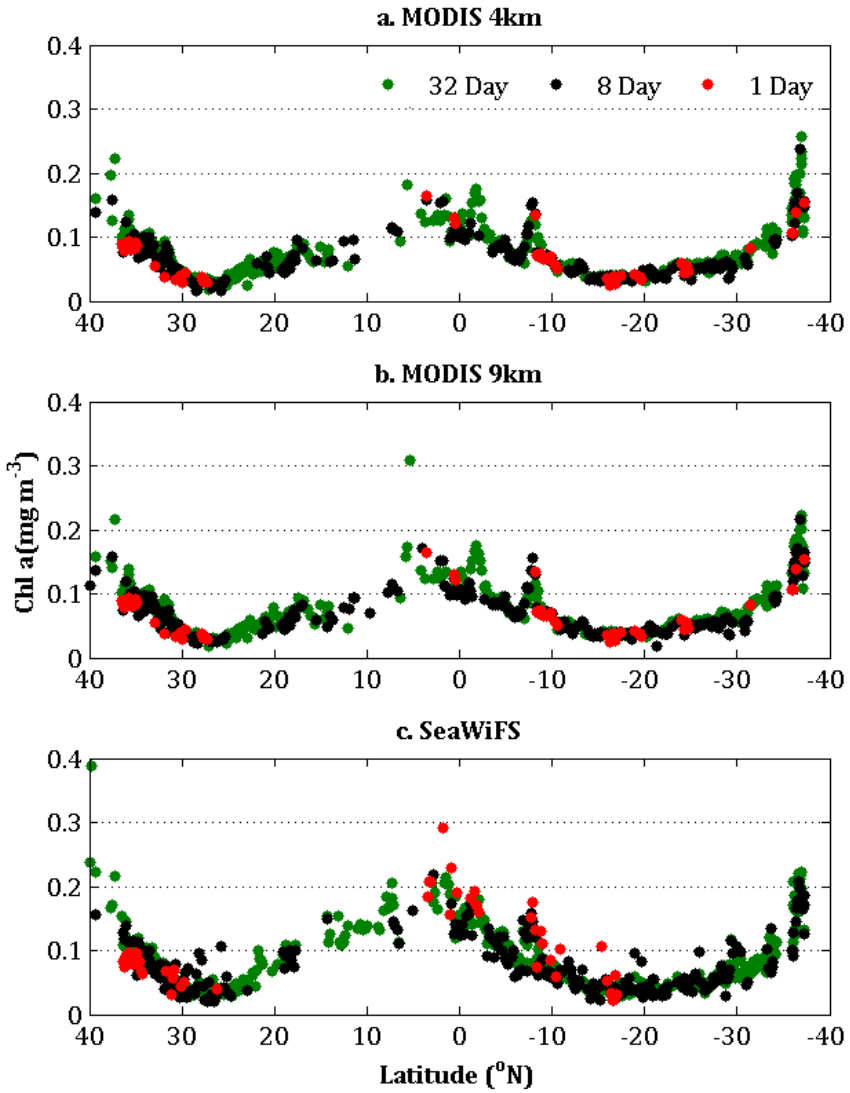


Figure 6.3. Latitudinal plot of chlorophyll *a* concentration measured from (a) MODIS 4 km resolution, (b) MODIS 9 km resolution and (c) SeaWiFS satellites. Different colour symbols indicate 1 day, 8 day and 32 day composites.

A running coefficient of variation was calculated for every 25, 50 and 100 samples (450, 900 and 1800 km). This highlighted the increased degree of variability in both *Syn* and PicoEuk towards the ends of the region and of the *Syn* increase around the equator and *Pro* increase at the northern end of the transect region (Figure 6.4a, c and e). The coefficient of variation was lowest between latitudes 10◦N to 20◦N and 12◦S to 30◦S. Heterotrophic bacteria is also plotted to show the lower coefficient of variation, that is similar to that of PicoEuk (< 0.5), see Figure 6.4b, d and f.

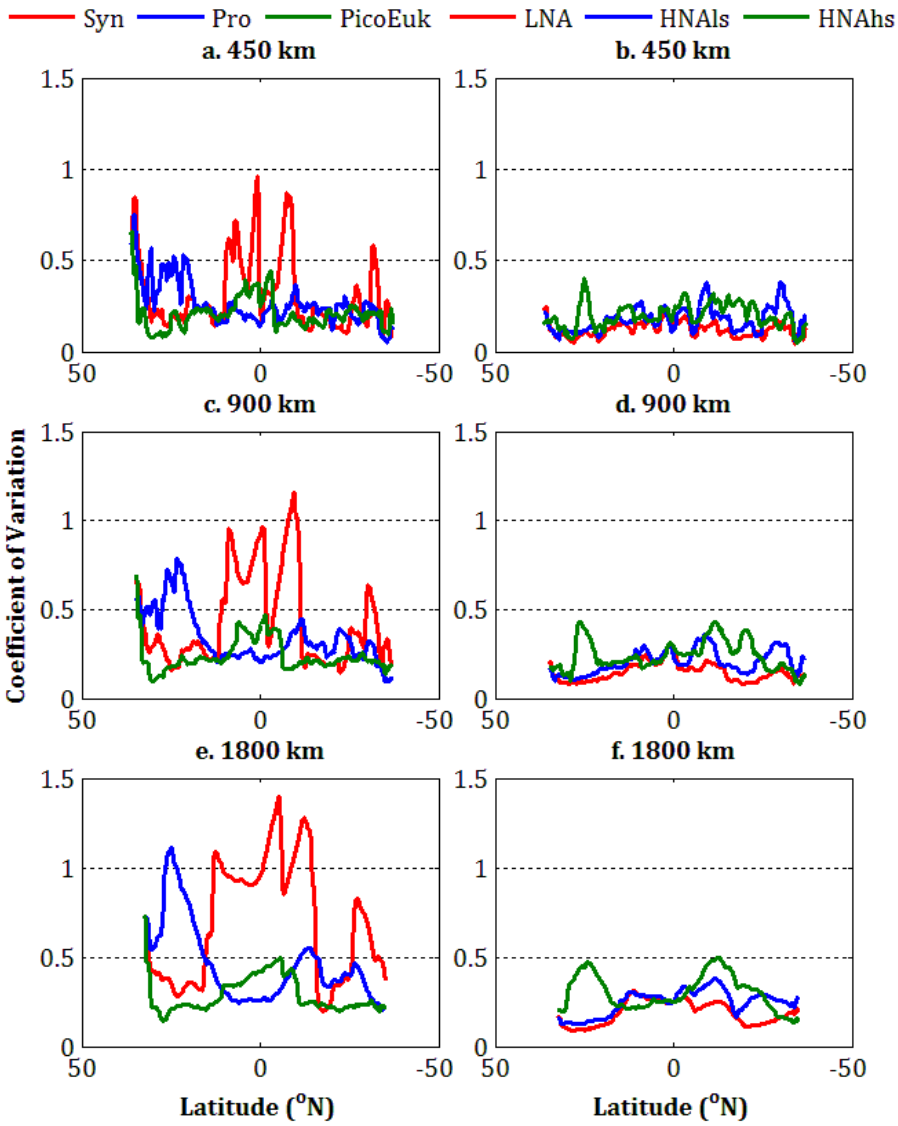


Figure 6.4. Running coefficients of variations calculated for abundances of (a, c and e) phototrophic ultraplankton and (b, d and f) bacterioplankton, running over (a, b) 450 km, (c,d) 900 km and (e, f) 1800 km.

Spearman correlation coefficients were also calculated for all pairs of variables within the cluster region *a* (see Table 6.2). PicoEuk was moderately significantly correlated with *Syn* (0.57) and moderately to weakly positively correlated with LNA and HNAls bacteria (0.43 and 0.33 respectively). The correlations of physical properties (temperature and salinity) were negative or weak with the ultraplankton groups. Salinity was significantly moderately to strongly negatively correlated to *Pro* and *Syn* (-0.78 and -0.57). Other strong correlations between the ultraplankton groups were *Syn* - LNA (0.72), HNAls - LNA (0.72) and *Pro* - HNAhs (0.73). No obvious predator-prey correlations (e.g. PicoEuk versus bacteria) were seen.

Table 6.2. Spearman rank correlation coefficient (ρ) for ultraplankton group abundances from cluster region *a*. Those in bold are statistically significant ($p < 0.01$).

	Syn	Pro	PicoEuk	LNA	HNAls	HNAhs	Temp	Sal
Syn								
Pro	***0.32***							
PicoEuk	***0.57***	***-0.17***						
LNA	***0.72***	***0.38***	***0.43***					
HNAls	***0.61***	***0.28***	***0.33***	***0.72***				
HNAhs	***0.40***	***0.73***	***0.18***	***0.53***	***0.51***			
Temp	***-0.19***	***0.33***	-0.06	0.03	0.08	***0.36***		
Sal	***-0.57***	***-0.78***	-0.02	***-0.5***	***-0.35***	***-0.55***	***0.11***	

6.3.1.1. Autocorrelation

Autocorrelation length scales were calculated with the aid of correlogram's (where correlation *r* is plotted against the lag *k*).

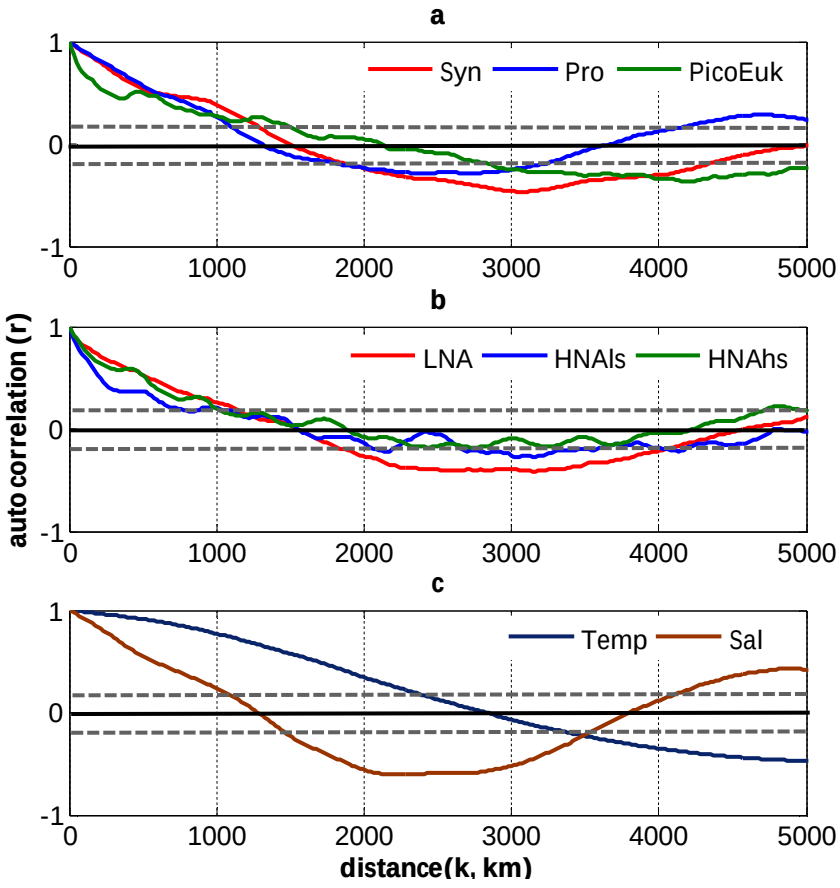


Figure 6.5. Correlograms of (a) phototrophic ultraplankton abundances, (b) bacterioplankton abundances and (c) temperature and salinity. *r* = 0 is indicated by a thick black line. Grey dashed lines are the mean 95 % confidence limits (0.18), values outside of these lines are significantly different from zero.

Correlations are strong over short distances. The length at when $r_k \simeq 0$, is where the samples are no longer correlated (Chatfield 2004), gives a typical length scale of variability for the organism or field. For reference the distances spanned by cluster region *a* is ~ 8,500km

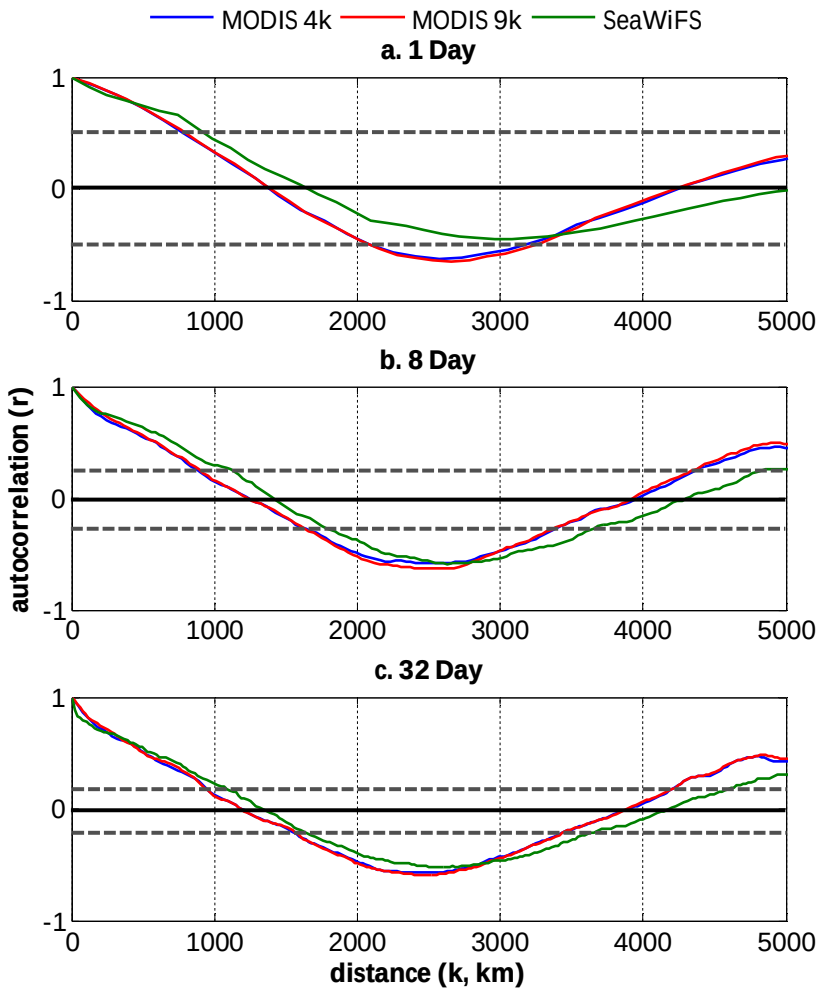


Figure 6.6. The correlogram for different frequency satellite chlorophyll *a* concentrations - (a) daily data, (b) 8-day and (c) 32-day composites for SeaWiFS and MODIS 4 km and 9 km resolution. All data were interpolated to AMT14 spatial resolutions. $r_k = 0$ indicated by thick black line. Grey dashed lines are the mean 95% confidence limits, values outside of these lines are said to be significantly different from zero.

The longest ultraplankton autocorrelation length scale (see Figure 6.5 and Table 6.3) was PicoEuk (2121 km), longer than all satellite data by ~ 700 km. The shortest was seen in *Pro* (1312 km). This was similar to the satellite autocorrelations. The temperature length scale (2841 km) was higher than all of the ultraplankton groups and satellite chl *a* data and salinity (1294 km) had a similar resolution to *Pro* and the ocean colour data (32 day data ranged from 1190 to 1596 km). Different temporal composites of satellite (MODIS 4

km, 9 km and SeaWiFS) chl *a* concentrations length scales did not significantly differ (see Figure 6.6 and Table 6.3).

Satellite chl *a* concentration (32 day composites) extracted to be coincident with the AMT14 transect samples were used to check if the AMT14 sampling resolution was sufficient to capture the ultraplankton structure. This showed that the sampling resolution (~ 18 km) was slightly too low to capture the complete picture, as the full resolution autocorrelation $r_k = 0$ was just outside of the AMT resolution confidence bounds (see Figure 6.7 and Table 6.3), this is most likely due to fewer samples ($n \leq 513$) for interpolation in the lower resolution (AMT) datasets than the full resolution satellite datasets ($n \leq 2036$). A higher sampling resolution would have been more robust, but it was not possible to do this here, therefore it will not be discussed further.

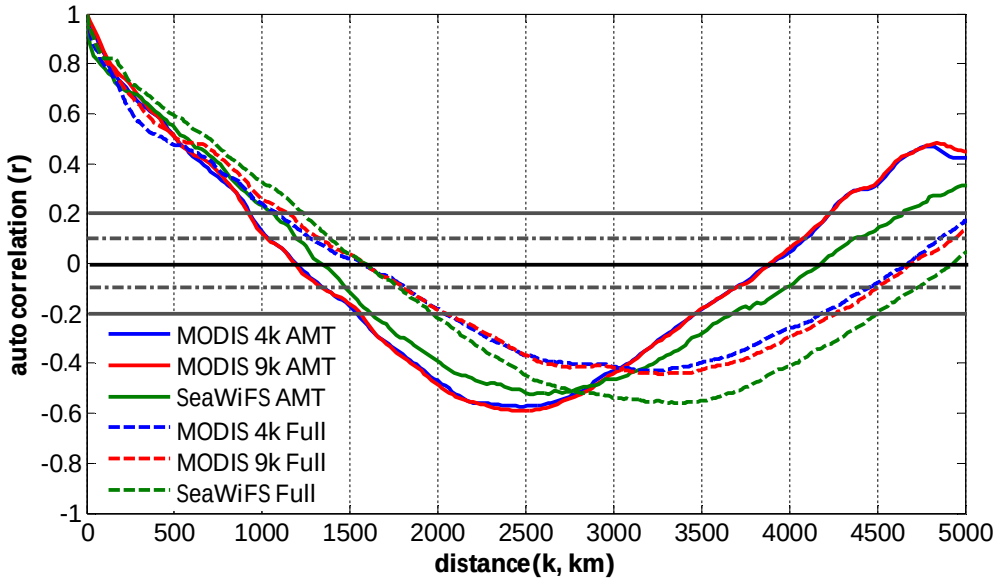


Figure 6.7. The correlogram of chlorophyll satellite concentrations for SeaWiFS and MODIS satellites extracted at full resolution (Full, dashed coloured lines - 4 km MODIS, 9 km MODIS and SeaWiFS) and for the same satellites extracted at AMT14 sampling resolution (AMT, solid coloured lines). $r_k = 0$ indicated by a thick black line. Grey solid lines are the mean 95 % confidence bounds for AMT sampling resolution and grey dashed lines are the mean 95 % confidence bounds for full satellite resolution, values outside of these lines are said to be significantly different from zero.

Table 6.3. The autocorrelation length scale for each ultraplankton group, temperature, salinity and satellite chl *a* concentration (1D, daily data at AMT14 sampling resolution; 8D, 8 day composite at AMT14 sampling resolution; 32D AMT, 32-day composite at AMT14 sampling resolution; 32D Full, 32 day composite at full satellite sampling resolution of 4 or 9 km). The 95% confidence intervals (C.I.) are also reported, as is the upper and lower distances of the C.I. See associated Figures 6.5 to 6.7.

		$r_k = 0$ (km)	95% C.I.	C.I. distance (km)	
				Upper	*Lower*
	Syn	1510	± 0.18	1294	1843
	Pro	1312	± 0.18	1081	1861
	PicoEuk	2121	± 0.18	1510	2810
	LNA	1546	± 0.18	1120	1852
	HNAls	1564	± 0.18	1043	2103
	HNAhs	1888	± 0.18	1025	2427
	Temp	2841	± 0.18	2330	3236
	Sal	1294	± 0.18	1115	1460
MODIS 4k	**1D**	1445	± 0.52	740	2095
	8D	1257	± 0.25	907	1640
	32D AMT	1212	± 0.18	924	1555
	32D Full	1572	± 0.1	1277	1830
MODIS 9k	**1D**	1391	± 0.47	742	2096
	8D	1232	± 0.24	905	1638
	32D AMT	1190	± 0.18	922	1560
	32D Full	1587	± 0.13	1325	1805
SeaWiFS	**1D**	1726	± 0.53	885	-
	8D	1433	± 0.23	1109	1755
	32D AMT	1333	± 0.18	1060	1625
	32D Full	1596	± 0.13	1380	1805

The autocorrelations are different for all the ultraplankton variables. Therefore unmeasured physical factors (e.g. eddies) are not controlling there variability, otherwise the ultraplankton would have corresponding length scales.

6.3.1.2. Semivariogram analysis

The 32 day dataset had the highest sample number and the best mean goodness of fit (indicated by high r^2 value, see Table 6.4) and lowest mean standard deviation for the slope. Thus the 32 day dataset was used for comparison to *in situ* data. The 1 day datasets sample number, however, was too low for acceptable accuracy. Therefore the 1 day dataset will not be referred to in terms of semivariograms.

Here the semivariogram slope *b* is used to quantify the changes in spatial variability with length scale for each of the variables, Figures 6.8 and 6.9. Table 6.4 gives the relevant

statistics for the slope, *b* and an r^2 value to indicate the strength of the scaling relationship between 40 and 100 km.

PicoEuk has the shallowest slope of all the ultraplankton groups (0.4) and *Pro* had the steepest (1.4). *Syn* had a slope of 0.8 and heterotrophic bacteria (LNA, HNAhs, HNAls) have slopes between 0.7 and 1.1. The ultraplankton groups (except for PicoEuk) and temperature and salinity have a high r^2 (> 0.8). The 32 day satellite data have shallow (~ 0.2 - 0.4) slopes similar to PicoEuk, but have a low r^2 of ~ 0.5. The 8 day chl *a* satellite data r^2 is higher (0.5 SeaWiFS and 0.8, MODIS 9k) and the slope ranges between 0.5 (SeaWiFS) and 1.2 (MODIS 9k).

Table 6.4. Mean and standard deviation calculated from 10,000 bootstraps for slope (*b*) and r^2 for a fit to line $\gamma(r)=ar^b$ to plot of log $\gamma(r)$ versus log(r), where *r* is the separation distance and $\gamma(r)$ is the semivariogram (see Figures 6.8 and 6.9). The amount of increase in variablity over 40 – 100 km is labelled as the increase in length scale.

		Slope (*b*)		r^2		↑ length scale
		mean	SD	mean	SD	
	Syn	0.80	0.30	0.83	0.19	2.08
	Pro	1.42	0.29	0.93	0.08	3.67
	Pico	0.41	0.24	0.59	0.29	1.46
	LNA	0.74	0.24	0.83	0.17	1.97
	Hhs	1.12	0.26	0.89	0.10	2.79
	Hls	0.84	0.29	0.82	0.18	2.16
	Temp	1.61	0.43	0.90	0.12	4.37
	Sal	1.31	0.29	0.87	0.11	3.32
32 day (*n* = 459)	**Mod4**	0.24	0.30	0.50	0.33	1.25
	Mod9	0.20	0.31	0.42	0.32	1.20
	Sea	0.38	0.31	0.46	0.31	1.42
8 day (*n* = 224)	**Mod4**	0.92	0.43	0.67	0.27	2.32
	Mod9	1.17	0.41	0.77	0.20	2.91
	Sea	0.48	0.40	0.50	0.32	1.56
1 day (*n* = 24)	**Mod4**	0.94	1.19	0.47	0.34	2.37
	Mod9	0.96	1.24	0.48	0.34	2.41
	Sea	0.89	1.44	0.44	0.34	2.25

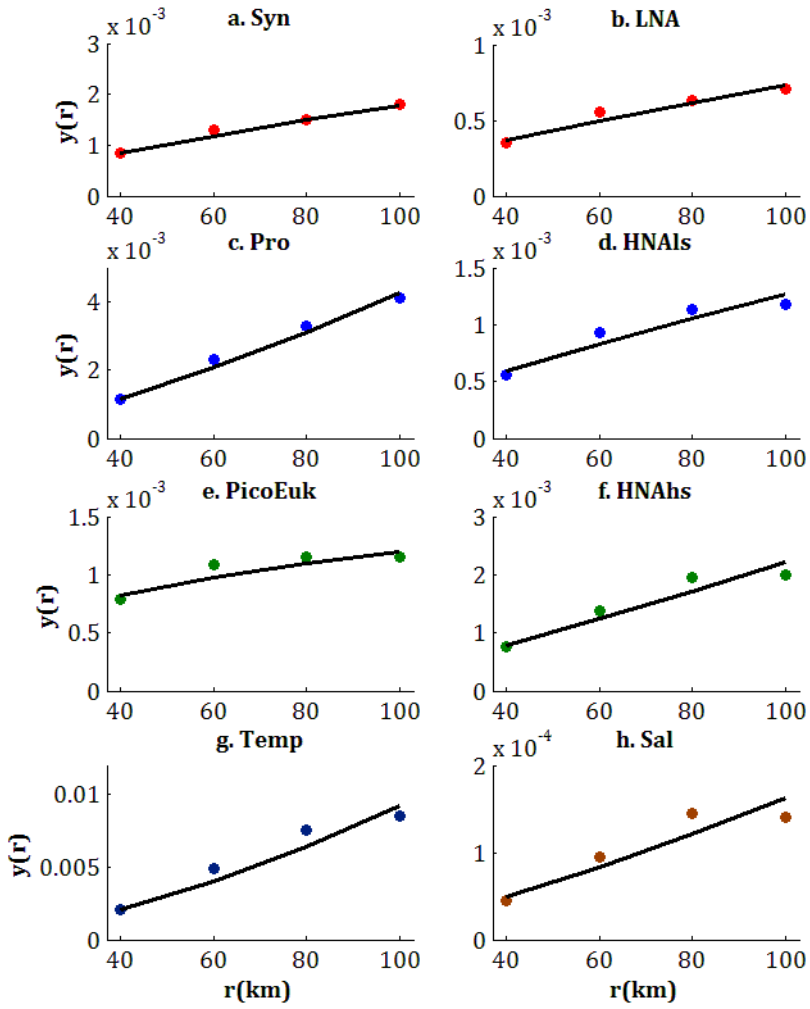


Figure 6.8. Semivariograms for the six ultraplankton groups (abbreviation in the text) and salinity (Sal) and temperature (Temp) for the range of 40 km to 100 km. The line represents the least-square fit with the form $\gamma(r)=ar^b$. See slope *b* estimates in Table 6.3. Note differing *y* axis.

In essence, the slope is a measure of how much more variable each group is at separation scales of 100 km and smaller, than at 40 km or less. For example, *Pro* has 3.7 times more variability at 100 km, whilst PicoEuk only has an increase in variability of 1.5 from 40 km to 100 km (see final column of Table 6.4). The slopes of each variable were compared and all were significantly different, including satellite data (t test, $p < 0.005$).

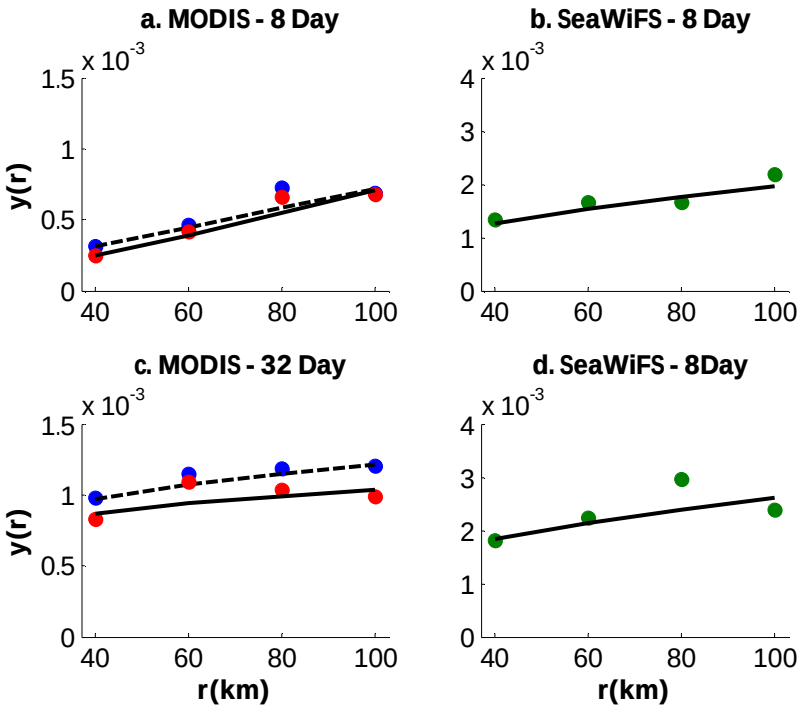


Figure 6.9. Semivariograms for 8 day and 32 day chl *a* composites from MODIS 4km resolution (blue dots, dashed line) and 9 km resolution (red dots, solid line) and SeaWiFS satellites (green dots, solid line). The lines represent the least-square fit with the form $\gamma(r)=ar^b$. See slope *b* estimates in Table 6.3. Note differing *y* axis for SeaWiFS and MODIS.

6.3.2. Satellite and Ultraplankton Correlation

SeaWiFS and MODIS (4 and 9 km) chl *a* daytime direct measurements (1 day) and daytime 8 and 32 day composite chl *a* measurements were compared to abundance (cell ml^{-1}), biomass (fg C ml^{-1}) and chl *a* content (chl *a* mg m^{-3}) for ultraplankton. To allow a clear comparison between satellites, data were only used for days when data from both satellites were available. Only MODIS 9 km resolution is presented here, as MODIS 4 km and 9 km results were very similar.

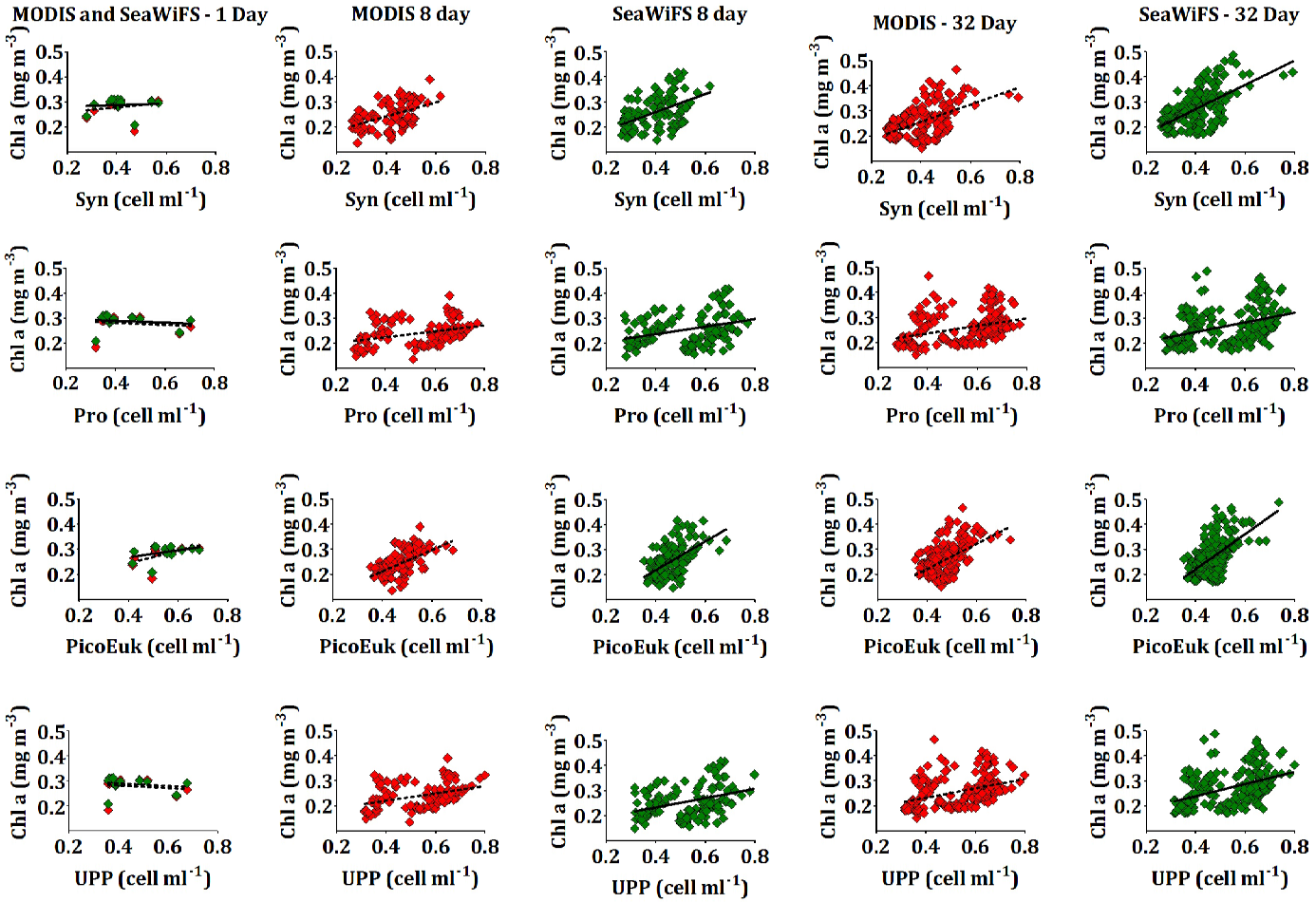


Figure 6.10. Correlation plots of Chl *a* (*y* axis) from MODIS 9 km (red diamonds, black dashed lines) and SeaWiFS (green diamonds, solid black lines) satellites versus *Synecoccocus* (Syn), *Prochloroccus* (Pro), PicoEukaryotes (PicoEuk) and total ultraphotoplankton abundances (UPP, cell ml^{-1}, standardised fourth root, *x* axis). Lines indicate linear regression. See Table 6.4. for correlation coefficients, significance and r^2

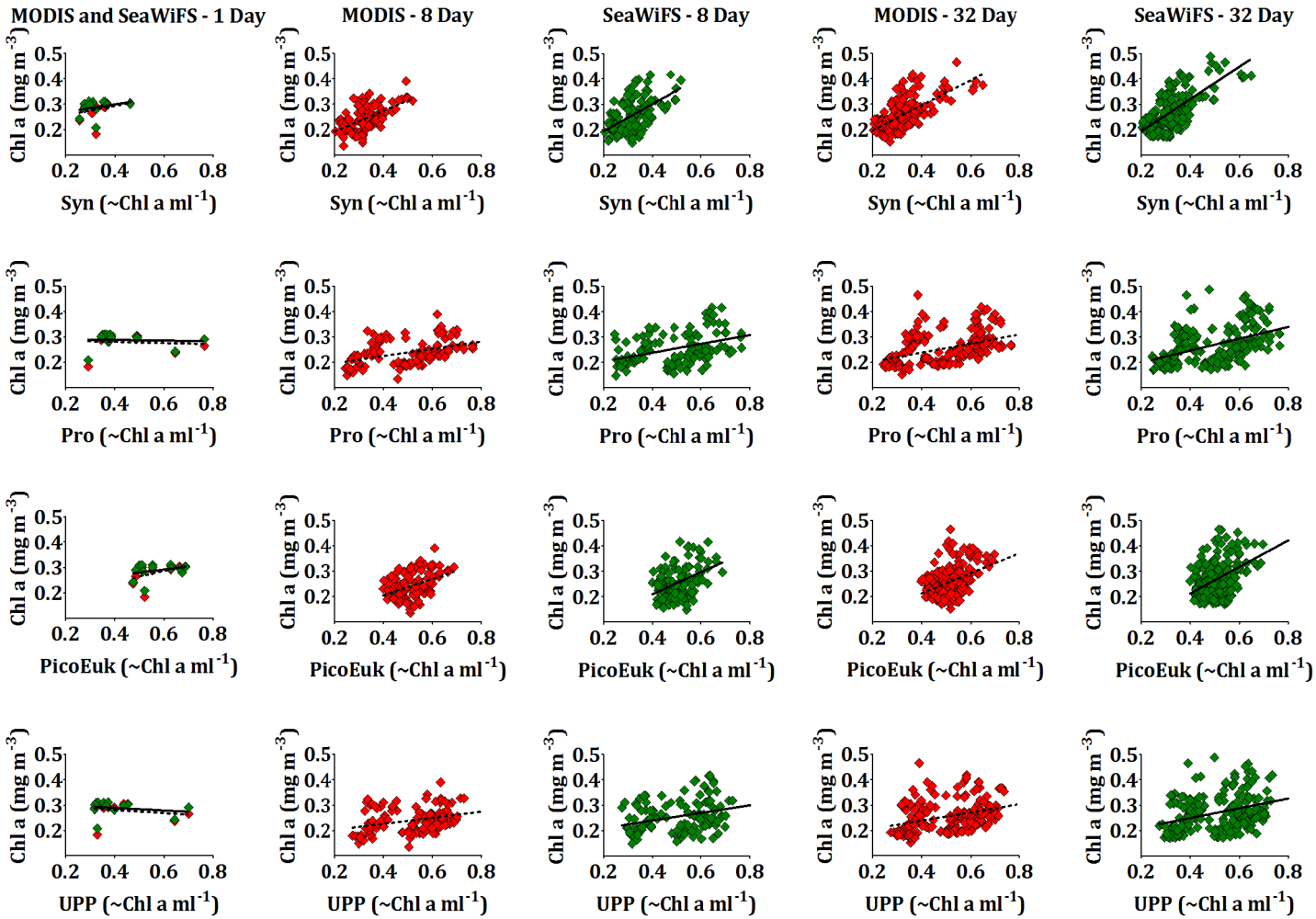


Figure 6.11. Correlation plots of satellite Chl *a* (*y* axis) from MODIS 9 km (red diamonds, black dashed lines) and SeaWiFS (green diamonds, solid black lines) versus *Synecoccocus* (Syn), *Prochloroccus* (Pro), PicoEukaryotes (PicoEuk) and total ultraphotoplankton (UPP) chl *a* content (~Chl *a* ml^{-1}, *x*-axis). Lines indicate linear regression. See Table 6.4. for correlation coefficients, significance and r^2.

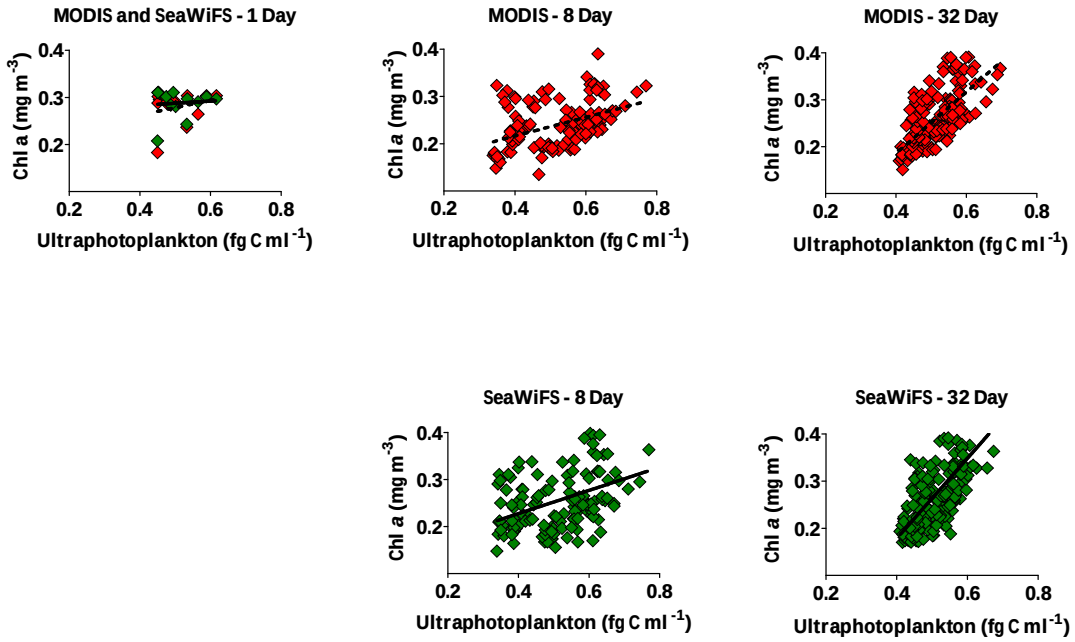


Figure 6.12. Correlation plots of Chl *a* concentrations (mg m^{-3}, *x*-axis) from MODIS 9km (red dots, black dashed lines) and SeaWiFS (green, solid black lines) satellites versus total ultraphotoplankton biomass weighted to abundance (fg C ml^{-1}, *x* axis). Lines indicate linear regression. See Table 6.4 for correlation coefficients, significance and r^2.

The Spearman correlation coefficients (ρ) for abundance versus satellite chl *a* concentration were typically positive and significant in all but SeaWiFS daily composites (see Table 6.5 and Figure 6.10.). The MODIS and SeaWiFS 32 day composites had slightly higher ρ than 8 day composites and both these temporal resolutions gave a weak to modest correlation for all the ultraphytoplankton groups. However the highest significant correlation overall was found in abundance of PicoEuk and *Syn* versus MODIS 1-day data ($\rho = 0.75$ and $\rho = 0.6$, respectively, $p < 0.05$). *Pro* abundance had consistently the lowest significant correlation to satellite data across all of the temporal ranges ($\rho = 0.37$ - 0.45).

Correlation was also high in 1-day MODIS versus PicoEuk chl *a* content ($\rho = 0.62$, $p < 0.01$), see Figure 6.11. Spearman correlation with biomass was also calculated. As this was essentially a simple clear conversion (multiplying femtograms of C by abundance), the result echoes abundance correlations with satellite chl *a*. Therefore only correlation with total ultraphytoplankton biomass is shown here (Figure 6.12) and this gave some of the highest correlations, with a strong significant correlation for 8 and 32 day (1 day was not significant) composites, $\rho = 0.6$-0.7, $p < 0.001$.

Table 6.5. Spearman correlation coefficients (ρ) and significance (*p*), see associated Figures 6.10, 6.11 and 6.12. Blank indicates not significant (*p* > 0.05), Significance indicated by stars - * *p* < 0.05, ** *p* < 0.005 and * *p* < 0.001. UPP, is ultraphytoplankton (summed *Syn*, *Pro* and PicoEuk) and PB is phototrophic bacteroa (summed *Syn* and *Pro*).**

	MODIS 9k						SeaWiFS					
	Daily		8 day		32 day		Daily		8 day		32 day	
n =	*13*		*129*		*200*		*13*		*129*		*200*	
	ρ	*p*	P	*p*	ρ	*p*	ρ	*p*	ρ	*p*	ρ	*p*
Abundance												
Syn	0.60	*	0.47	***	0.47	***	-0.05		0.44	***	0.53	***
Pro	0.16		0.45	***	0.45	***	-0.31		0.37	***	0.43	***
PicoEuk	0.75	**	0.50	***	0.43	***	0.17		0.50	***	0.52	***
PB	0.13		0.46	***	0.47	***	-0.37		0.38	***	0.45	***
UPP	0.14		0.47	***	0.47	***	-0.37		0.39	***	0.46	***
Chl *a* content												
Syn	0.44		0.52	***	0.65	***	0.21		0.48	***	0.62	***
Pro	0.14		0.51	***	0.52	***	-0.22		0.44	***	0.52	***
PicoEuk	0.62	*	0.35	***	0.36	***	0.09		0.34	***	0.37	***
PB	0.11		0.36	***	0.45	***	0.00		0.28	***	0.39	***
UPP	0.10		0.37	***	0.42	***	-0.22		0.30	***	0.36	***
Biomass												
UPP	0.43		0.68	***	0.71	***	-0.22		0.60	***	0.71	***

Here it should also be noted for *Pro* and ultraphytoplankton that in Figures 6.10, 6.11 and 6.12 a split in the data can be seen. The cluster closer to the *y*-axis (lower abundance, chl *a* content or biomass) corresponds to samples located approximately > 25°N and those to the right are samples from approximately < 24°N. This is the same split that can be seen in cluster region *a* (Chapter 2, Section 2.3.1) when similarity is increased to 91%. This will be discussed further in the following Section.

6.4. Discussion

6.4.1. Ultraplankton Spatial Variability

Are ultraplankton communities in the surface waters of the subtropical and tropical Atlantic as spatially variable as in the shelf and temperate seas? The ultraplankton variability over the region as a whole was highest in phototrophic ultraplankton abundances (Syn, Pro and PicoEuk, coefficient of variation 1.21, 0.64 and 0.64 respectively) and lowest in LNA (0.28). The variability within the region is slightly higher

here than for measurements of ultraphytoplankton abundances made at the PAP site (0.51-0.65) or in the Celtic Sea (0.38-0.78) (Martin, unpublished data). However, the heterotrophic bacteria coefficient of variation is a little smaller than at the PAP site (0.4), and substantially higher than the Celtic Sea (0.6) (Martin, unpublished data). Therefore variability could be said to be higher for phototrophic ultraplankton in the subtropical and tropical Atlantic than in the temperate Atlantic and shelf Celtic Sea

To put this further into context with studies from shelf and temperate seas, the *n*-fold range over the region can be considered (by taking the maximum abundance and dividing it by the minimum abundance for each ultraplankton group). In the Celtic Sea region (Martin et al. 2005), the *Syn* population fluctuated by more than 60-fold (spatial resolution of study 1.5 km) and in the Mozambique channel (resolution ~ 19 km) the maximum fluctuation in *Syn* was 13-fold (Zubkov and Quartly 2003). This study shows a greater fluctuation when looking at the entire subtropical and tropical gyre region. *Syn* fluctuated by 95-fold and *Pro* even more (105-fold) and PicoEuk by 29-fold despite being in a coherent grouping (see Chapter 2). Conversely heterotrophic bacteria only varied by 4-fold over the entire region, around the same amount as the 3-fold fluctuation measured in the Celtic Sea (Martin et al. 2005). Thus, there is higher variability in phototrophic ultraplankton in the Atlantic subtropical and tropical ocean than in shelf and temperate seas.

This comparison of variability nevertheless could be argued to be irrelevant as it took place over a ~ 10,000 km transect. This can be disputed however for two reasons. Firstly, the area has previously been defined as having 90% ultraplankton similarity (see Chapter 2 for details) and secondly, even adjacent samples were found to vary. In other words, variability is inherent even between adjacent samples and fluctuations are present throughout the transect. Neighbouring samples vary by as much as 135,000 cell ml^{-1} for *Pro* (a 150 % increase from the mean), 10,000 cell ml^{-1} for *Syn* (240 %) and 5,000 cell ml^{-1} for PicoEuk (a 600 % increase). These point increases are supported by sustained spikes in the running coefficient of variation over these locations (see Figure 6.4), indicating that an increase in abundance occurred over a number of samples, giving confidence that these spikes are genuine. It is impossible however to tell from this study if the full variability has been captured, as it is a linear transect (like the Mozambique Channel study by Zubkov and Quartly, 2003) and it is not possible to determine if the variability seen along the transect would have been witnessed at the smallest scales if a more detailed local survey had taken place (such as that in the Celtic Sea study sampling every ~ 1.5 km, by Martin et

al. 2008). These results nevertheless demonstrate that spatial ultraplankton variability is a persistent feature and, although more focused in the equatorial region, does extend well into the subtropical gyres. These findings indicates that the initial assumption of spatial uniformity in the subtropical and tropical gyres in the modelling Chapters 4 and 5 may have been inappropriate.

Correlation of ultraplankton groups were used to examine the source of this variability within this statistically similar cluster region. A correlation between predator and prey (i.e. mixotrophic PicoEuk and bacteria) might be expected. A strong significant correlation (0.57) is seen between PicoEuk and *Syn*, with the remaining correlations of PicoEuk to bacterial groups moderate to weak (except for *Pro* where the relationship is weakly negative, see Table 6.2). Correlations of this nature have been observed previously in the temperate regions, between PicoEuk and *Syn* (0.8) (Martin et al. 2010). Temperature and salinity correlations have also be tested, to see if variability is associated with large scale physical gradients . No significant strong correlations were found between temperature and any of the ultraplankton groups (see Table 6.2). However *Pro* and salinity were strongly negatively correlated (-0.78). This is thought to be due to the dip in the salinity from the south equatorial current and the increase in *Pro* towards the equator. However, the reason for such a relationship is not clear. Correlations between ultraplankton groups and physical fields, go some way to explaining the variability within a statistically similar region. However, physical processes on the whole clearly do not dictate their variability (Table 6.3). This supports findings in Chapter 2, which argued that physical features should not be used to define distinct ultraplankton provinces.

The first law of geography, or Tobler's Law, is that "Everything is related to everything else, but near things are more related than distance things" (Tobler 1970). Autocorrelation, bears this out (see Figure 6.5 and Table 6.3), although the ultraplankton groups have different length scales. It was expected that all ultraplankton groups would autocorrelate over a similar distance as they are all within cluster region *a*. However, PicoEuk distance (2121 km) was longer than all other ultraplankton groups (1312 - 1888 km). This may be due to the lower abundance causing a higher sampling error, but as yet no robust conclusions can be draw from this. Temperature and salinity autocorrelation supported the correlation coefficient results as temperature autocorrelation length bore little resemblance to those of ultraplankton abundances (2841 km), but salinity was similar to the ultraplankton groups (1219 km). In terms of satellite, data 8-day and 32-day temporal resolution had the most comparable (1232 - 1433 km and 1190 – 1333 km

respectively) large scale structure to phototrophic ultraplankton abundances except for PicoEuk (2121 km). Therefore autocorrelation showed that large scale ultraplankton structure, salinity and remotely sensed satellite chl *a* were relatively comparable (except for PicoEuk, although even this, at the lower end of its autocorrelation confidence interval was similar to the other ultraplankton groups) and all had a dissimilarity with temperature.

Semivariogram analysis was conducted to see how the spatial distribution of variability contrasts between the various ultraplankton groups (Table 6.4, Figure 6.8). In the range 40 - 100 km, all variables (except PicoEuk) show a strong relationship of $\gamma = ar^b$ (see Section 6.3.1.2) indicated by the high r^2 value. Out of all the ultraplankton groups, PicoEuk had the shallowest slope (~ 0.4) and *Pro* the largest (~ 1.4). Amongst the over variables, physical fields exhibited some of largest slopes (1.3 – 1.6, similar to *Pro*) and the 32-day satellite composites the smallest (similar to PicoEuk, 0.2 - 0.4). To put these differences in slopes into perspective, a few examples will be given. With a slope of $b = 0.41$, the PicoEuk distribution has 1.5 times more variability at separation scales of 100 km or smaller than at 40 km or smaller. In contrast *Pro* has 3.7 times more variability at 100 km and smaller, than at 40 km and smaller. To appreciate this difference, if *Pro* and PicoEuk had the same degree of variability at a separation distance of 40 km, then *Pro* distribution would have almost two and a half times more variability than PicoEuk at scales from 40 to 100 km. These increases in variability with distance are much smaller than those reported by Martin et al. (2008) in the Celtic Sea, for example that study reported PicoEuk as having 2.3 times more variability at 15 km than at 3 km and *Syn* (*Pro* was not measured) had 15 times more variability at 15 km than at 3 km, making *Syn* over five times more variable than PicoEuk. In contrast this study found *Syn* only 1.4 times more variable than PicoEuk although over a larger distance (40 to 100 km). All ultraplankton groups in this study had smaller mean slopes than the Celtic Sea study (Martin et al. 2008), indicating that variability was less at the longer length scales measured here. For example, the mean slope of *Syn* in this study was 0.80 (see Table 6.4), less than the slope for *Syn* found within the Celtic Sea of 1.68 (Martin et al. 2008). This implies that variability at scales 40 – 100 km (this study) contributes less to the total variability, than at smaller scales of 3 – 15 km (Celtic Sea study). However, this is only an indication, as these studies cannot be directly compared due to the very different physical environments.

The use of semivariograms complements the preceding analysis of spatial variability. Ultraplankton abundance varies considerably within the tropical and subtropical gyre

(shown by the coefficient of variation). Although this variability autocorrelate over similar scales, the distribution of spatial variability over 40 to 100 km is significantly different for all variables tested. A study on the Celtic shelf (Martin et al. 2008) found little evidence for difference in slope between variables, although there was a significant difference between *Syn* and PicoEuk. The fact that differences are found here may be an indication that some of the ultraplankton groups in temperate regions have a closer relationship (in terms of spatial variability over distance) and are more tightly coupled than those seen here, in subtropical and tropical regions, with the caveat that these observations were made at different scales. The main result of the Celtic Sea study was that microbial groups can have a very tight relationship in a model and yet show weak correlations and variability on different scales (Martin et al. 2008), therefore a model with 'tight' relationships, such as that presented in Chapters 4 and 5, may not necessarily be inappropriate to simply model these spatially varying ultraplankton, whereas here it may be. Also the investigation in this Chapter uses abundances, whereas uptake rates were used to constrain and independently test the models in Chapters 4 and 5 respectively, therefore, it would be necessary to investigate if uptake rates vary at the same spatial scale as abundance to completely disregard this structure. Unfortunately current methodological and practical constraints prevent a high resolution transect of uptake rates to be measured, therefore abundances have been used as a proxy.

6.4.2. Satellites and Ultraplankton Correlation

Oceanographers were limited to physical sampling for centuries, unable to study vast areas and variability within them. Remotely sensed observations from satellites changed this. The oceans could be viewed as dynamic, changing and varying spatially and temporally on daily to decadal time scales and many studies have relied upon them (e.g. Platt and Sathyendranath 1988, Longhurst et al. 1995, McClain et al. 2004, Tilstone et al. 2009). A previous study in the Mozambique channel (Zubkov and Quartly 2003) has suggested that ocean colour (SeaWiFS specifically) may not be adequate to assess the contribution of ultraplankton groups. The data set in this Chapter therefore provided a unique opportunity to examine whether this is true for the subtropical and tropical Atlantic (Figures 6.10 and 6.11 and Table 6.5). The correlations between phototrophic ultraplankton groups abundance and chl *a* concentration and 8-day and 32-day satellite temporal resolutions were modest to strong (0.35 - 0.65) and total biomass of phototrophic ultraplankton strongly correlated with 8 day (0.6 - 0.7) and 32 day satellite

data (0.71). Strong significant correlations were also seen in the MODIS 1 day data and abundances of PicoEuk (0.75) and Syn (0.6). However, very weak or negative correlations of phototrophic ultraplankton were also observed with the 1 day data (e.g. *Pro*). Although the strong 1-day correlations were significant, the paucity of the data (n = 13) means that further surveys are needed to ensure 1 day data is robust. The biomass calculation is a summed conversion (using the abundance multiplied by the weights detailed in the method). Therefore this is not a direct measurement, but seems to provide a stronger correlation than abundance or chl *a* concentration for longer temporal resolution data (8 and 32 day). This result may indicate that the smoothing in temporal variability of the satellite composites reduces the correlations as the variability that this study has shown to be present, is being smoothed out (also supported by linear regression, see raw data in Appendix, Table A.5, where r^2 was low, all < 0.5). The correlation results found here are similar to those in the Mozambique Channel (Zubkov and Quartly 2003) which showed a strong correlation between daily SeaWiFS composite and PicoEuk and *Syn* chl *a* and a low correlation with *Pro*. Contrastingly, although they found a negative correlation to abundance of *Pro* and a strong correlation with weekly composites and *Syn* and PicoEuk chl *a* whereas here that was not found. Therefore ultraplankton abundances, biomass and chl *a* concentration can at best be modestly well predicted and the key ultraphotoplankton *Pro*, only poorly from satellite concentrations in the subtropical and tropical Atlantic.

Figures 6.10, 6.11 and 6.12 show a divide in the scatter of *Pro* and total ultraphytoplankton (summed *Syn, Pro* and PicoEuk) in MODIS and SeaWiFS satellite data, most pronounced in the 8 day resolution. The sample points located close to the *y*-axis (lower abundance/biomass/chl *a* content and higher chl *a* concentration) are from samples taken towards the northern end of the transect > 25°N. Those to the right of the plot, are samples from < 24°N of the equator and in the southern hemisphere. This is the same cluster 'split' that was seen in Chapter 2 (Figure 2.3). The cause of this gap cannot be determined for certain here, but it can be speculated upon. It may be that seasonality is affecting the ultraplankton at the northern end of the transect more than previously thought, specifically the seasonal effects on cell pigmentation (specifically *Pro*). *Pro* cells in spring have been shown to have more pigment (Veldhuis and Kraay 2004). In other words the satellite is over estimating the amount of *Pro*. A correction for this would move this group down the *y*-axis. Alternatively satellites could be under estimating *Pro* cell abundance elsewhere, as culture studies have shown that *Pro* cells under higher irradiances (in this case those cells in the tropics and subtropics) decrease in size and chlorophyll pigment (Morel et al. 1993), so that they could be 'missed' by the satellite but

'caught' by the flow cytometer. If they were 'seen' by the satellite it would shift this grouping up the *y*-axis. Both these processes add to the uncertainty in the correlation and lead to the conclusion that *Pro* is not being accurately measured by satellites (this was also concluded by Zubkov and Quartly, 2003). If this is the case, why is this discrepancy also not seen in PicoEuk or *Syn*? What *Pro* pigmentation (or variability therein) leads to the difficulty in its satellite measurement? Unfortunately that is outside the remit of this work.

The semivariogram results of this study also raise concerns over use of satellite data to study the subtropics. The slopes of semivariograms (spatial variability over 100 km) for MODIS and SeaWiFS at all temporal resolutions were significantly different to all ultraplankton groups (see Figure 6.9 and Table 6.4), thus suggesting that the satellites are not sufficiently capturing the ultraplankton variability. The semivariogram's r^2 also varied between temporal resolutions and perhaps unsurprisingly the 8 day composites had the highest r^2. As the smoothing effect on fronts and eddies in the longer (32-day) temporal scale data would have caused more scatter around the slope. The 1-day data does not improve the situation as the number of measurements were too low (n = 24) to robustly produce a semivariogram. Thus, from this study it can be established that current ocean colour algorithms do not provide enough information about ultraplankton groups that are of key importance to ocean biogeochemical cycles (Li and Harrison 2001). This study emphasises the need for *in situ* observations and experiments at small spatial scales to be used to guide future ocean sensor development.

6.5. Summary and Implications

The ultraplankton abundances in the Atlantic subtropical and tropical gyre region have been shown to have as high variability as ultraplankton abundance in temperate and shelf regions. These groups also vary at different spatial scales, although these differences may not be as great as those in shelf seas. The marked spatial variability in ultraplankton abundances are present despite being within the same ultraplankton defined region (> 90 % similar, Chapter 2). The presence of such heterogeneity suggests that the simple zero-dimensional model constructed and parameterised in Chapter 5, may not be sufficient to describe the system. A model that can explicitly capture this variability might be required.

This Chapter also demonstrated that satellite detection of ultraphytoplankton is somewhat limited (Section 6.3.2). The choice of satellite (MODIS or SeaWiFS) made little difference to

the correlations; 8-day and 32-day correlations were similar and modest. One-day direct estimates were slightly better for MODIS detection of PicoEuk. However the limited number of data points due to cloud cover prevented a robust conclusion. Therefore models, provinces and primary production estimations based upon remotely sensed chlorophyll in the subtropics should be interpreted with care.

From the results presented in this chapter and considering the three stated hypothesis to be tested the following can be concluded:

- The ultraplankton community in the surface waters of the subtropical and tropical Atlantic have as high spatial variability as ultraplankton in the shelf and temperate seas and so homogeneity cannot be assumed.
- Different ultraplankton groups in the surface waters of the subtropical and tropical Atlantic, vary at significantly different spatial scales.
- Remotely sensed satellite chl *a* concentration cannot be reliably used to estimate the abundances, biomass or chl *a* of phototrophic ultraplankton (*Pro*, *Syn* and PicoEuk) in the surface waters of the tropical and subtropical Atlantic.

7. Overall Conclusions

The work presented in this thesis aimed to **'to explore how mixotrophy may be modelled in the subtropical Atlantic using a data driven approach'** by constructing and parameterising a mixotroph model for the subtropical Atlantic using *in situ* measurements (**Chapters 4 and 5**). This thesis further addressed questions and assumptions related to model structure (**Chapters 2, 3 and 6**). This investigation has been undertaken using a combination of observations and computer modelling.

This Chapter summarises and synthesises the results from each Chapter (**Chapters 2 to 6**), relating them back to the overarching aim and the objectives of this thesis set out in **Chapter 1**. The wider implications of the main findings are then considered. Finally, the limitations of the study and possible future directions for further research are outlined.

7.1. Summary of Research Findings

7.1.1. Objective 1

To ascertain if the subtropical Atlantic ocean can be considered as a single oceanographic province.

Ultraplankton in the subtropical Atlantic includes the picoeukaryotic algae, which have been found to be predominantly bacterivorous (Zubkov and Tarran 2008, Hartmann et al. 2012). Picoeukaryotes, their prey (cyanobacteria and low and high nucleic acid bacteria) and their predators (heterotrophic flagellates) were used to assess oceanographic regions.

In **Chapter 2** oceanographic provinces were delineated through multivariate analysis based on ultraplankton abundances. Samples taken between 49°N and 40°S were found to

group statistically (at > 90 % similarity level, $p < 0.05$) into four provinces. The largest ultraplankton defined biogeographic province (*a*) incorporated both the north and south subtropical gyres, as well as the central tropics, 36°N to 39°S. These findings indicate that in terms of ultraplankton assemblages this region can be considered distinct. This analysis did not reveal any seasonality within the subtropical gyres, as the opposing hemispheres (in opposing seasons) were considered statistically similar.

An assessment of *a priori* defined oceanographic provinces that had been based on the modelling of plankton (Follows et al. 2007), remotely sensed chlorophyll *a* concentrations (McClain et al. 2004) and physical boundaries (Longhurst 1998, 2007) revealed that the ultraplankton community was not different *between* or similar *within* these regions (except for in the temperate north Atlantic). In particular, ultraplankton assemblages between regions defined using the McClain et al. (2004) criteria barely differed, this led to further investigations of the suitability of using remotely sensed chlorophyll *a* concentrations to determine phototrophic ultraplankton abundances in **Chapter 6**.

The multivariate defined province *a* (**Chapter 2**), was used in subsequent analysis to demarcate the study region of interest (**Chapters 3 and 6**), and to motivate the development of a single model for the region (**Chapters 4 and 5**).

7.1.2. Objective 2

To investigate dissolved organic phosphate (DOP) utilisation by ultraplankton in the Atlantic oligotrophic ocean.

In nutrient poor environments it has previously be thought that organic nutrients may be an important alternative to inorganic nutrients (Casey et al. 2009, Lomas et al. 2010). This was investigated to identify if an organic nutrient (specifically dissolved organic phosphate) component needed to be included explicitly in models of the subtropical Atlantic (**Chapters 4 and 5**).

In **Chapter 3,** measurements collected *in situ* were used to determine if DOP was being utilised as a significant alternative source of P. Plankton uptake rates of three different DOP nucleotides – ATP, AMP and UMP – were measured simultaneously for the first time. The results showed that none of the DOP nucleotides tested were being significantly

utilised ($p > 0.05$), supporting previous evidence of insignificant planktonic ATP uptake by Zubkov et al. (2007). Therefore it was not considered necessary to include DOP in a model focusing on mixotrophy in this region (**Chapters 4 and 5**). ATP uptake and concentration were also compared for three different years which encompassed differing seasons. No significant differences between season or years were found ($p > 0.05$), giving further tentative support to the assumption of steady state made in **Chapter 4**.

7.1.3. Objective 3

To develop and to parameterise from *in situ* observations a simple zero-dimensional model of the Atlantic microbial ecosystem incorporating mixotrophy.

A simple model structure of dissolved inorganic P, bacteria, mixotrophs and grazer variables was used due to the limited amount of data. The insignificant utilisation of DOP by plankton, found in **Chapter 3**, meant the model did not need to explicitly contain a dissolved organic phosphate variable. Initially in **Chapter 4**, a simple zero-dimensional mixotroph model was constructed, assuming steady state. This was solved analytically using *in situ* observations of uptake fluxes. The model assumption of steady state was initially supported by the lack of seasonality found in the results from **Chapter 2 and Chapter 3**.

Through network analysis of the steady state mixotroph model, two of the net fluxes constrained by the data were found to be in the incorrect direction from an ecological perspective. It was possible that representing the two fluxes as a single net flux, rather than as independent components in each direction, may have placed too strong a constraint. Therefore, in the revised model (**Chapter 5**), each was modelled explicitly. However this meant the system could no longer be solved analytically. The model was further developed to be time-varying, allow a seasonal cycle to be described and for the measurements to be not in equilibrium. The steady state model developed in **Chapter 4**, therefore formed the basis of the dynamic model presented in **Chapter 5**.

Unable to use flow analysis on the time-varying model, the model parameters were instead estimated using a stochastic optimisation technique and the *in situ* observations (**Chapter 5**). Despite the limited number of observations, maintaining the coexistence of

all model variables and achieving a fit to data was difficult to attain. Nevertheless, a good fit to the data was eventually achieved.

The resulting model allowed the role of mixotrophy in an oligotrophic ecosystem to be explored. The dynamics of the modelled ecosystem were dominated by the mixotrophs, with the largest flux of P flowing through the mixotroph variable from bacterivory. The recycling of P through mixotrophs and bacteria to the P pool was also a central feature of the system. The majority of bacterivory performed in the system was by mixotrophs, which agrees with previous observations (Zubkov and Tarran 2008, Hartmann et al. 2012). The model further indicated that bacterivory was not a survival mechanism, as even when P concentrations were at their highest, bacterivory was a significant mechanism for obtaining nutrients by mixotrophs. Despite the bacterivory being performed by algae, primary production and export were slightly higher than previous estimates for the region. It was previously thought that the inclusion of mixotrophy would decrease energy transfer efficiency because of the extra trophic link in the system.

A lack of data to constrain the model, however, makes it difficult to determine how sensitive the model results were to its structure and parameterisation. To help assess if a simple zero-dimensional mixotroph model without horizontal variations was appropriate, microbial spatial variability was assessed in **Chapter 6**.

7.1.4. Objective 4

To explore microbial spatial distribution and variability throughout the subtropical Atlantic Ocean.

In **Chapter 6**, the spatial variability of microbial groups within the region, defined when Objective 1 was addressed (**Chapter 2**), was explored. This demonstrated that despite the ultraplankton assemblages being > 90 % similar, significant internal spatial variability was present. The phototrophic ultraplankton were shown (by coefficient of variation and *n*-fold increase comparisons) to be as spatially variable as in temperate shelf seas. The microbial groups were also shown to vary significantly over different spatial scales, indicating that there was no single factor causing their variability. The variability observed in the ultraplankton abundances was not seen in remotely sensed data for the region. This

led to the investigation of the applicability of satellite observations to detect ultraplankton (also in **Chapter 6**).

The inherent heterogeneity observed may explain the difference seen in the microbial uptake rate measurements (**Chapter 4**). This suggests that a larger number of observations would be needed to obtain an accurate mean and associated uncertainty for the region for key fluxes. Unfortunately current methodological constraints make this impractical.

This spatial heterogeneity also implies that the dynamic mixotroph model may have been too simple to describe the ecosystem. The model was a single box and did not describe the inherent ultraplankton variability found in **Chapter 6**. As such a spatial model with horizontal resolution capable of representing the physical processes that cause the advection and diffusion of plankton horizontally would be desirable to reproduce this heterogeneity within the system. Although differing length scales for the ultraplankton groups suggest biological processes may be equally important. Nevertheless, the model still provides information about the ecosystem, if viewed as a specific point model, i.e. a model for this specific location (CTD14, 27°N -32°E). However, the model result cannot necessarily be extrapolated over the wider region.

7.1.5. Objective 5

To investigate if remotely sensed satellite chlorophyll *a* concentration can be used to estimate the abundances, biomass or chlorophyll *a* content of phototrophic ultraplankton.

As length scales for remotely sensed satellite chlorophyll concentrations were significantly different from ultraphotoplankton variability and did not display the heterogeneity that was seen in the *in situ* observations the adequacy of ultraphotoplankton detection by satellites is questioned.

In **Chapter 6**, satellite measurements of chlorophyll *a* concentration (at different temporal resolutions and from different satellites) were evaluated against *in situ* abundances, biomass and chlorophyll *a* content of phototrophic ultraplankton groups and types (*Pro*, *Syn*, Picoeukaryotes, total phototrophic bacteria and total ultraphotoplankton). Remotely

sensed measurements were overall only modestly correlated with *in situ* observations. *Pro* correlations were consistently low and weak (all $\rho < 0.45$, $p < 0.001$), despite being probably the most abundant photosynthetic organism in the oceans (Partensky et al. 1999a). *Syn* and PicoEuk *in situ* measurements showed a stronger correlation with satellite observations. Overall the strongest correlation found in this study was between MODIS 1-day estimates and picoeukaryotic abundance ($\rho = 0.75$, $p < 0.005$), although this measurement had a low number of samples ($n = 13$). These findings support a previous study of SeaWiFS (1 and 8 day composites) comparison to observations in the Mozambique channel (Zubkov and Quartly 2003). The conclusion is that remotely sensed chlorophyll observations may be incorrectly estimating ultraphotoplankton in the oligotrophic Atlantic.

7.2. Wider implications

7.2.1. Ultraplankton Biogeographic Provinces

The multivariate analysis used to define the ultraplankton provinces in this thesis demonstrated that ultraplankton communities do not always adhere to physical boundaries or coincide with remotely sensed chlorophyll distributions (**Chapter 2**). In addition, demonstrated differences between the results of three alternative means of defining provinces (Longhurst 1998, McClain et al. 2004, Follows et al. 2007), show there is a lack of consistency and applicability to ultraplankton in previous province delineation. The findings of this thesis therefore suggest that for future oceanic studies, where a large enough data set is available, previously published regions (e.g. Longhurst Provinces) should not be used for inter/intra specific data comparisons without question. Ideally a multivariate analysis technique would be applied to confirm or deny the region is coherent and the boundaries are relevant to the data of interest (i.e. fish distributions should not be used to define boundaries to investigate bacteria). However, it is appreciated that this is not normally possible due to field expense, logistics and methodological limitations.

7.2.2. Mixotroph Modelling

Mixotrophy within the oligotrophic Atlantic is a dominant and ubiquitous process (Zubkov and Tarran 2008, Hartmann et al. 2012). Here it has been shown for the first time that it is possible to fit a simple dynamic model incorporating mixotrophy to albeit very limited *in situ* data and maintain coexistence of all components of the ecosystem (**Chapter 5**).

In theory the inclusion of mixotrophy in an ecosystem model should reduce nutrient export and maintain or even increase nutrient turnover in the mixed layer, as the fraction of nutrients passed on to higher trophic levels decreases due to the extra trophic link in the system. The model presented here demonstrated that the majority of P was being recycled and not passed on to a higher trophic level. However, estimates of export (1 - 18 mol C m^{-2} y^{-1}) were slightly higher than previous observations and modelled estimates. The model also has shown the potential importance of mixotrophs within the oligotrophic ecosystem, with the largest fraction of nutrients being cycled through them, relative to the other variables. Consequently, future ecosystem models of oligotrophic regions should consider including mixotrophs as a basic ecological element.

7.2.3. Ultraplankton Spatial Variability in the Subtropics

This thesis has presented evidence for spatial variability in ultraplankton in the subtropics and tropics of the same order of magnitude as that observed in temperate and coastal seas (Martin et al. 2005, Martin et al. 2008, Martin et al. 2010), with PicoEuk varying 29-fold, *Syn* 95-fold and *Pro* 105-fold within this region. Variability of this size is not apparent in remotely sensed data, possibly due to the modest ability of satellites to detect ultraphytoplankton. Different dominant length scales of variability were seen between different ultraplankton groups, implying that no single process (such as eddies) was causing the patchiness. This variability is potentially a significant source of error for observations, as to estimate a mean abundance for a region a large number of samples may be required. It is also an indication that zero-dimensional models (such as those used in **Chapters 4 and 5**) may not be entirely appropriate to model the oligotrophic Atlantic ecosystem; despite a conventional view of such regions as homogenous, variability is widespread and marked. To accurately model plankton this patchiness, horizontal advection and other physical processes that can influence population dynamics locally

needs to be taken into account. It is as yet unknown if the variability in abundances will be seen in uptake rate measurements (such as those measurements used in model constraining in **Chapter 4** and verification in **Chapter 5**). Although such uptake measurements would take a large amount of time and effort, it would considerably aid the understanding of the consequences of spatial variability.

7.2.4. Remote Sensing of Ultraplankton

In **Chapter 6**, it was shown that MODIS and SeaWiFS satellites at 8 and 32-day temporal resolutions only have a modest correlation with *in situ* ultraphytoplankton abundances ($\rho = 0.37 – 0.53$, $p < 0.001$) and chlorophyll concentration ($\rho = 0.28 – 0.65$, $p < 0.001$). SeaWiFS 1-day and 8-day composites had previously been investigated in the Mozambique Channel (Zubkov and Quartly 2003). However, to the author's knowledge this is the largest study of its kind to date, using two satellites and three temporal resolutions. One of the most concerning results of this study was the poor ability of the satellites to detect *Pro*, which is thought to be the most abundant phototrophic organism on earth (Partensky et al. 1999b) and responsible for a significant fraction of total primary production, particularly in the subtropics and equatorial regions (Liu et al. 1997). As *Pro* is not being accurately detected by satellites, any models or estimates of primary production (and in turn export, carbon sequestration or oxygen production) based upon satellite chlorophyll concentration are potentially significantly in error in the oligotrophic oceans which are dominated by these organisms and cover vast areas (the Atlantic oligotrophic region is in excess of 10 million square kilometres, Polovina et al. 2008). Attempts to fit models should therefore allow for this discrepancy and interpret their results with care. Further technological and algorithm developments are necessary in satellite remote sensing to ensure all ultraphytoplankton functional types are observed accurately.

7.3. Study Limitations and Future Directions

This Section provides an outline on the main limitations of the study. In the light of these, it also highlights priorities for future research.

7.3.1. Data

The foremost limitation of the mixotroph model and its further development is the lack of *in situ* observational data available to constrain the model parameters and to independently validate the model output. Further experimental analysis of individual processes, especially those shown to be sensitive within the model (for example bacterivory undertaken by grazers and mixotrophs, and remineralisation of P by mixotrophs) would improve the model and the confidence in its results. Enhancements in modelling the mixotrophic dynamics within an oligotrophic ecosystem will in turn advance forecasting of primary production and export in these vast regions.

Although time-consuming and harder to perform in the field, uptake rate measurements are stronger than state variable concentrations in constraining a model (Franks 2009). Simply having more rate measurements would constitute a major step forward in constraining microbial ecosystem models, not least those including mixotrophs. Observations throughout the year would considerably help to verify that the ecosystem dynamics are being correctly represented within the model. Ideally an increased temporal resolution dataset of uptake rates would be collected, with monthly means and errors calculated for provinces. This, however, would clearly be an expensive operation.

An alternative way to obtain uptake rate data could be flow cytometry derived abundances. Although they require *in situ* ship based measurements they are less time consuming and can be collected at a higher spatial resolution. However, as already stated, variable concentrations are not as strong in constraining models (Franks 2009). In addition, they also require a conversion factor (in terms of P content of the cell) for comparison to models, which can introduce another level of uncertainty into the data. In the future, advances in satellite detection of plankton functional groups may allow the detection of ultraphytoplankton (Aiken et al. 2009) concentrations that could then be used for constraining models. However this is not currently possible using the data in this thesis (**Chapter 6**).

7.3.2. Model Structure

At present the simplicity of structure of the mixotroph model may hinder its predictive skill. The model currently has a very simple structure of a single nutrient, P, a single pool

of carnivores, G, with an additional bacterial component (B) and phytoplankton modelled as mixotrophs (M). It therefore lacks a detrital pool, with the waste material of the organisms being directly remineralised into the P pool or assumed to be consumed by bacteria, acting as detritivores. This may be restricting the accurate simulation of nutrient recycling (Anderson 2010), which is important in oligotrophic regions where the availability of nutrients is dependent upon the recycling of nutrients within the surface layer (also seen in the present mixotroph model). Therefore the inclusion of a detrital pool in the mixotroph model could be an informative development. The simple model presented here also did not have an obligate photosynthetic organism, which previous models incorporating mixotrophy have included (e.g. Thingstad et al. 1996, Crane and Grover 2010). This would introduce further competition and help understand why mixotrophs are dominant. In addition, there is evidence suggesting that models have greater portability when multiple phytoplankton types (e.g. obligate phytoplankton and mixotrophs) are included (Friedrichs et al. 2007). However, the inclusion of an obligate phototroph does not guarantee coexistence of all variables in a model incorporating mixotrophy (Thingstad et al. 1996). The possible predictive benefits gained from increasing the number of parameters and variables in the model need to be weighed against the cost (in time and finances) of constraining the model, as discussed above.

There are also physical structure limitations to the model. It assumes homogeneity in the horizontal resolution and nothing below the mixed layer. In addition it excludes the influence of neighbouring or deeper populations and this prevents the model from directly simulating all but the simplest physical processes of mixed layer deepening and turbulent mixing across the mixed layer. The physics are a key factor controlling nutrient supply and advection of plankton. More detailed physics could help to describe the plankton patchiness observed and the role it plays in the ecosystem.

In future the integration of a mixotroph model into a 3D ocean general circulation model (GCM) might address these limitations and help to elucidate the wider and global impact of mixotrophy. As this thesis has shown, due to the degree of spatial heterogeneity, observations from one point in the subtropical Atlantic (**Chapter 6**) cannot be easily extrapolated to its entirety. Added complexity will accomplish nothing if the resulting model is poor (Fulton et al. 2003), or the data is insufficient. Therefore striking the balance between complexity and data will be the key to progress.

7.4. In Conclusion

Modelling of mixotrophy in oligotrophic regions is extremely challenging, primarily due to the lack of *in situ* data available from these remote regions of the world. The mixotroph model presented in this thesis is a considerable oversimplification of reality. Nonetheless, such simple models have been shown numerous times to be able to accurately represent first-order dynamics and biogeochemical cycling (Hood et al. 2008) and this thesis has demonstrated this for mixotrophy. A dynamic mixotrophic model was able to describe the dynamics of the system, albeit constrained by very few data. The difficulty even of this illustrates that even so few data exert strong constraints on a model including mixotrophy. This model highlighted the key role of mixotrophy in the cycling of nutrients within the surface of the subtropical ocean and this fast nutrient turnover is crucial for the sustained functioning of the oligotrophic ecosystem. The model also indicated that mixotrophy was not a survival mechanism, and was more likely an adaptive strategy, as bacterivory is the predominant route of P acquisition even when P concentrations are relatively high.

In addition to explicitly building a model incorporating mixotrophy, this thesis has also tested the assumptions underlying model structure to give guidance to future models building on this work. For example, dissolved organic P has been ruled out as a necessary inclusion in future models. Also the relatively high degree of ultraplankton spatial variability has indicated that zero-dimensional models are not suited to draw conclusions for this entire oligotrophic region. Furthermore the limited ability of remote sensing to detect ultraphytoplankton has been recognised making use of remote sensing data in model building difficult. It is therefore hoped that these findings will significantly contribute to future development of models for the Atlantic oligotrophic ecosystem and our ability to understand and predict the role of this vast region in global biogeochemical cycles.

Appendix

Table A.1. AMT cruise data used within this thesis; dates and principle scientists. NOC, National Oceanography Centre, Southampton, UK. PML, Plymouth Marine Laboratory

Cruise	Route	Dates	Principle scientist	Cruise report
AMT14	Falkland Islands–UK	26/04/04–02/06/04	Patrick Holligan, NOC	Holligan (2004)
AMT17	UK–South Africa	12/10/05–22/11/05	Patrick Holligan, NOC	Holligan (2005)
AMT18	UK-Falkland Islands	04/10/08-10/11/08	Malcolm Woodward, PML	Woodward (2008)
AMT19	UK-Chile	13/10/09-01/12/09	Andy Rees, PML	Rees (2009)

Table A.2. AMT cruise data utilised in this thesis, with associated Chapter reference and acknowledgement of collectors. In addition to those specific data collected (as listed below), measurements of temperature, salinity and mixed layer depth were also used. See specific Chapter referenced for details. Abbreviations here are as follows. ATP, Adenosine triphosphate; AMP, Adenosine monophosphate; UMP, Uridine monophosphate and DIP, dissolved inorganic phosphate. See Table A.1.2 for affiliation abbreviation.

Cruise	Data collected	Thesis Chapter	Collected by
AMT14	Ultraplankton microbial group abundances	2, 6	M Zubkov, NOC
AMT14	Plankton ATP uptake, concentration and turnover time	3	M Zubkov, NOC
AMT17	Specific ultraplankton groups phosphate uptake rates and associate cell counts	4, 5	M Zubkov, NOC
AMT18	Plankton ATP uptake, concentration and turnover time	3	M Zubkov, NOC; M Hartmann, NOC
AMT19	Plankton ATP, AMP, UMP and DIP uptake, concentration and turnover time	3	S Herrington, NOC; M Hartmann, NOC

Figure A.1. Characteristic bioassay estimation of maximum DOP (in this instance ATP) concentrations and uptake rates from AMT19 (November 2009), CTD31. (a) Time series at different ATP concentrations with corresponding regression lines (dashed lines). ATP uptake was estimated in a dilution series, in which [α^{33}P] ATP at 0.2 nmol^{-1} was diluted with non-labelled ATP (amounts indicated). (b) The relationship between added ATP concentration and ATP uptake time. Error bars are single standard errors. The y-axis intercept of the regression is an estimate of turnover time (6.7 h) at maximum ambient concentration of ATP which is the x-intercept (0.14 nmol^{-1}). See details in text.

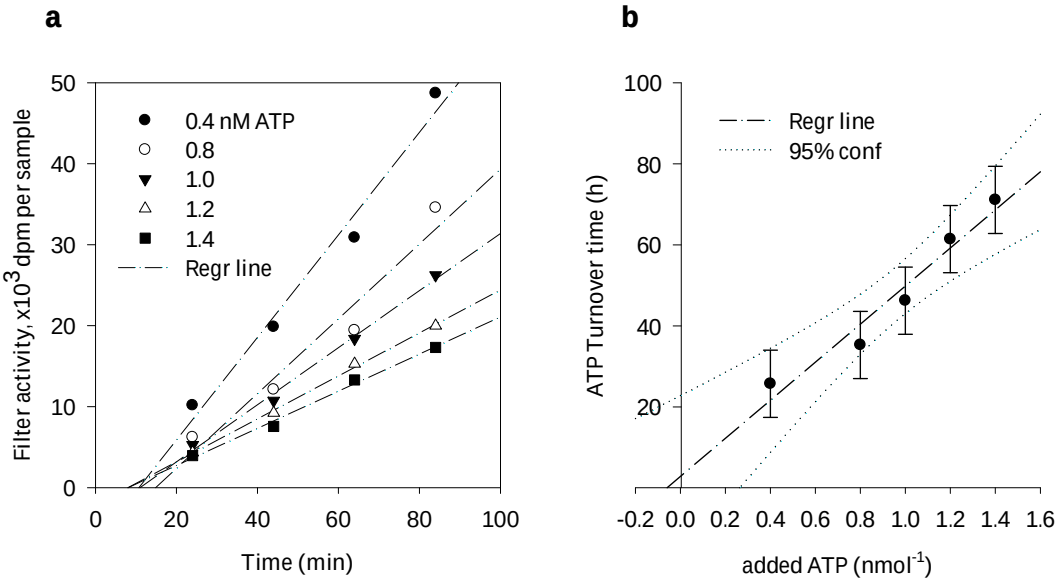


Table A.3. Concentration, uptake rate and turnover time for ATP, AMP, UMP and DIP (abbreviations in text). Note ExpE is individual experimental error, if * ExpE is standard deviation and SE is standard error of. Blank, are no data or below detection limit

			ATP						AMP						UMP						DIP					
			Concentration		Uptake rate		Turnover time		Concentration		Uptake rate		Turnover time		Concentration		Uptake rate		Turnover time		Concentration		Uptake rate		Turnover time	
CTD	Lat (°N)	Long (°E)	*nmol^{-1}*	*ExpE*	*nmol d^{-1}*	*ExpE*	*h*	*ExpE*	*nmol^{-1}*	*ExpE*	*nmol d^{-1}*	*ExpE*	*h*	*ExpE*	*nmol^{-1}*	*ExpE*	*nmol d^{-1}*	*ExpE*	*h*	*ExpE*	*nmol^{-1}*	*ExpE*	*nmol d^{-1}*	*ExpE*	*h*	*ExpE*
19	**38**	**-24**	0.26	0.11	0.06	0.01	101	43							0.73	0.06	0.06	0.00	272	22	2.69	0.22	0.81	0.07	80	6.6
22	**35**	**-28**	0.78	0.11	0.11	0.01	170	24					404	129*	1.04	0.12	0.14	0.01	178	21	15.14	0.08	7.03	0.59	52	0.3
25	**33**	**-31**					1257	52*					505	57*					1760	452*	4.02	0.08	2.27	0.19	43	0.9
28	**31**	**-34**					343	39*					116	21*					241	25*						
31	**29**	**-36**	0.14	0.10	0.51	0.04	7	5			0.68	0.06									2.04	0.28	4.31	0.25	11	1.6
34	**27**	**-38**					282	41*					88	12*					151	53*	3.49	0.48	11.76	0.96	7	1.0
37	**25**	**-39**	1.58	0.04	0.56	0.05	68	2					24	2*	1.89	0.07	1.42	0.06	32	1	2.33	0.20	19.90	0.75	3	0.2
40	**23**	**-40**					120	23*					41	4*					159	9*	7.72	0.14	16.31	0.91	11	0.2
42	**22**	**-39**					800	226*					336	52*					1547	143*	5.04	0.19	0.97	0.04	125	4.7
46	**22**	**-39**					221	115*					866	204*			0.04	0.01								
49	**19**	**-38**					674	251*					340	61*					270	33*	1.28	0.02	0.24	0.00	126	2.2
52	**16**	**-37**					493	162*					271	38*					270	38*						
55	**14**	**-34**					621	180*					699	13*							5.50	0.38	0.20	0.02	674	46.5
58a	**11**	**-32**					485	107*																		
58b	**11**	**-32**					674	152*																		
62	**7**	**-29**					79	5*																		
65	**4**	**-27**	0.90	0.06	0.07	0.01	325	22													11.14	0.37	2.01	0.13	133	4.4
73	**-4**	**-25**					662	80*																		
79	**-10**	**-25**					1374	131*																		
86	**-14**	**-25**					1326	134*																		
Average ± SE			**0.73±0.3**		**0.26±0.1**		**504±95**				**0.68±0.06**		**336±82**		**1.22±0.3**		**0.42±0.3**		**488±196**		**5.49±1.3**		**5.98±0.25**		**115±58**	

Table A.4. Concentration, uptake rate and turnover time for ATP for locations (lat, latitude and long, longitude) along the AMT14 transect in spring 2004 and AMT18 and AMT19 transects during autumn 2008 and 2009 respectively. Note ExpE is individual experimental error, except where * it is standard deviation. SE is standard error of the mean. Blanks are no data

AMT14								AMT18								AMT19							
		Concentration		Uptake rat		Turnover time				Concentration		Uptake rat		Turnover time				Concentration		Uptake rat		Turnover time	
Lat (°N)	Long (°E)	*nmol^{-1}*	*ExpE*	*nmol d^{-1}*	*ExpE*	*h*	*ExpE*	Lat (°N)	Long (°E)	*nmol^{-1}*	*ExpE*	*nmol d^{-1}*	*ExpE*	*h*	*ExpE*	Lat (°N)	Long (°E)	*nmol^{-1}*	*ExpE*	*nmol d^{-1}*	*ExpE*	*h*	*ExpE*
39.3	**-19.7**	1.07	0.16	0.14	0.02	180	6	**38.5**	**-25.2**	0.53	0.00	0.14	0.00	91	1	**38.2**	**-24.4**	0.26	0.11	0.06	0.01	101	43
31.9	**-31.2**	0.53	0.07	0.24	0.01	54	4	**36.0**	**-27.4**	0.30	0.02	0.13	0.00	56	3	**35.3**	**-28.4**	0.78	0.11	0.11	0.01	170	24
29.9	**-35.5**	0.58	0.04	0.15	0.01	92	3	**33.2**	**-30.5**	0.08	0.04	0.07	0.00	29	14	**33.3**	**-31.2**					1257	52*
27.3	**-35.8**	0.68	0.01	0.40	0.01	41	0.3	**27.4**	**-37.0**	0.49	0.08	0.11	0.01	111	18	**31.3**	**-33.4**					343	39*
23.4	**-34.2**	1.03	0.23	0.15	0.02	167	14	**24.4**	**-40.1**	0.74	0.05	0.23	0.01	77	5	**29.1**	**-35.6**	0.14	0.10	0.51	0.04	7	5
20.1	**-32.8**	1.45	0.18	0.22	0.02	158	6	**22.4**	**-40.2**	0.73	0.13	0.21	0.04	84	16	**27.1**	**-37.5**					282	41*
-3.0	**-25.0**	0.18	0.06	0.04	0.01	105	22	**19.4**	**-38.1**	0.41	0.08	0.10	0.01	97	19	**25.2**	**-39.4**	1.58	0.04	0.56	0.05	68	2
-11.7	**-25.0**	0.12	0.09	0.01	0.00			**16.5**	**-36.1**	1.81	0.08	0.28	0.03	155	6	**23.1**	**-40.4**					120	23*
								14.6	**-34.6**	0.73	0.08	0.22	0.02	79	8	**21.8**	**-39.1**					800	226*
								8.4	**-30.4**	0.78	0.04	0.49	0.03	38	2	**21.8**	**-39.2**					221	115*
								5.2	**-28.3**	2.04	0.19			41	4	**19.1**	**-37.5**					674	251*
								2.5	**-26.5**	0.82	0.11	0.52	0.07	38	5	**16.3**	**-36.6**					493	162*
								-0.4	**-24.6**	1.34	0.15	0.37	0.07	87	10	**14.1**	**-34.3**					621	180*
								-8.5	**-24.6**	0.47	0.08	0.08	0.01	147	25	**11.3**	**-32.4**					485	107*
								-12.5	**-24.6**	0.75	0.18	0.08	0.02	241	58	**11.3**	**-32.4**					674	152*
																6.5	**-29.3**					79	5*
																4.3	**-27.4**	0.90	0.06	0.07	0.01	325	22
																-3.5	**-25.1**					662	80*
																-10.4	**-24.6**					1374	131*
																-14.0	**-25.1**					1326	134*
Average ± SE		**0.70±0.2**		**0.17±0.04**		**114±21**				**0.80±0.1**		**0.21±0.04**		**92±14**				**0.73±0.3**		**0.26±0.1**		**504±95**	

Dynamic Mixotroph Model Code

```
Module ODE

 Use accuracy
 Use Parameters
 Use nrtype
 Use Mixedlayer
 Use Force

 Implicit None

 Contains

   Subroutine Derivs (x,y,dydx)!x is time

     Real(SP), Dimension (:), Intent(in):: y
     Real(SP), Dimension (:), Intent(out):: dydx
     Real(SP),intent(in) :: x
     Real(SP) ::  M, M2, t, yrlen
     Real(SP) :: h, hplus, mixcoP,mixcoG,mixcoB,mixcoM
     Real(SP):: aPB, yBM,psiBG,thetaMG,pieMP,bG,eM,phiG,source

      yrlen=365

      t=ceiling(x)

     call MLDepth(t,M,M2)

     h=(M2-M)/1
     hplus=max(h,0.0)


! yin(1)=P, yin(2)=G, yin(3)=B, yin(4)=M

      aPB = alpha*y(1)*y(3) ! Bacterial P uptake
      yBM = gamma*y(3)*y(4) ! M bacterivory
      psiBG = psi*y(3)*y(2) ! G bacterivory
      thetaMG = theta*y(4)*y(2) ! G grazing on M
      pieMP = pie*y(4)*y(1) ! M autotrophy
      bG = beta*y(2) ! G excretion/exudation of P taken up by B
      eM = episilon*y(4) ! M excretion/exudation of P directly toP
      phiG = phi*y(2) ! G excretion/sink of P from G


!Calculation of mixing coefficients
! NB all too small to be motile

      mixcoG=(y(2)*(mix+hplus))/M !G
      mixcoB=(y(3)*(mix+hplus))/M !B
      mixcoM=(y(4)*(mix+hplus))/M !M
      mixcoP=((Po-y(1))*(mix+hplus))/M !P


!Biological Equations

     dydx(1)=eM-aPB-pieMP !Phosphate
     dydx(2)=thetaMG+psiBG-bG-phiG !Grazers
     dydx(3)=aPB+bG-yBM-psiBG !Bacteria
```

```
     dydx(4)=pieMP+yBM-thetaMG-eM !Mixotrophs


!Addition of effects due to entrainment (mixing coefficient) with
mixed layer and background mixing
      dydx(1)=dydx(1)+mixcoP
      dydx(2)=dydx(2)-mixcoG
      dydx(3)=dydx(3)-mixcoB
      dydx(4)=dydx(4)-mixcoM


      return

   End Subroutine Derivs

End Module ODE
```

References

Adams, M. M., M. R. Gómez-García, A. R. Grossman, and D. Bhaya. 2008. Phosphorus Deprivation Responses and Phosphonate Utilization in a Thermophilic *Synechococcus* sp. from Microbial Mats. Journal of Bacteriology **190**:8171-8184.

Aiken, J., Y. Pradhan, R. Barlow, S. Lavender, A. Poulton, P. Holligan, and N. Hardman-Mountford. 2009. Phytoplankton pigments and functional types in the Atlantic Ocean: A decadal assessment, 1995-2005. Deep Sea Research Part II **56**:899-917.

Allen, J. I., J. T. Holt, J. Blackford, and R. Proctor. 2007. Error quantification of a high-resolution coupled hydrodynamic-ecosystem coastal-ocean model: Part 2. Chlorophyll-*a*, nutrients and SPM. Journal of Marine Systems **68**:381-404.

Anderson, T. R. 2005. HORIZONS: Plankton functional type modelling: running before we can walk? Journal of Plankton Research **27**:1073-1081.

Anderson, T. R. 2010. Progress in marine ecosystem modelling and the "unreasonable effectiveness of mathematics. Journal of Marine Systems **81**:4-11.

Anderson, T. R. and H. W. Ducklow. 2001. Microbial loop carbon cycling in ocean environments studied using a simple steady-state model. Aquatic Microbial Ecology **26**:37-49.

Anderson, T. R. and C. M. Turley. 2003. Low bacterial growth efficiency in the oligotrophic eastern Mediterranean Sea: a modelling analysis. Journal of Plankton Research **25**:1011-1019.

Antoine, D., J-M. André, and A. Morel. 1996. Oceanic primary production: 2. Estimation at global scale from satellite (Coastal Zone Colour Scanner) chlorophyll. Global Biogeochemical Cycles **10**:57-69.

Arenovski, A. L., E. L. Lim, and D. A. Caron. 1995. Mixotrophic nanoplankton in oligotrophic surface waters of the Sargasso Sea may emply phagotrophy to obtain major nutrients. Journal of Plankton Research **17**:801-820.

Armstrong, F. A. J., P. M. Williams, and J. D. H. Strickland. 1966. Photo-oxidation of Organic Matter in Sea Water by Ultra-violet Radiation, Analytical and Other Applications. Nature **211**:481-483.

Arrigo, K. R. 2005. Marine microorganisms and global nutrient cycles. Nature **437**:349-355.

Bainbridge, R. 1957. The size, shape and density of marine phytoplankton concentrations. Biological Reviews **32**:91-115.

Barcina, I., B. Ayo, M. Unanue, L. Egea, and J. Iriberri. 1992. Comparison of Rates of Flagellate Bacterivory and Bacterial Production in a Marine Coastal System. Applied and Environmental Microbiology **58**:3850-3856.

Baretta-Bekker, J. G., J. W. Baretta, A. S. Handen, and B. Riemann. 1998. An improved model of carbon and nutrient dynamics in the microbial food web in marine enclosures. Aquatic Microbial Ecology **14**:91-108.

Behrenfeld, M. J., E. Boss, D. A. Siegel, and D. M. Shea. 2005. Carbon-based ocean productivity and phytoplankton physiology from space. Global Biogeochemical Cycles **19**:GB1006.

Bengis-Garber, C. 1983. Uptake to adenosine in a marine bacterium is not an active transport process. Federation of European Biochemical Societies Letters **160**:31-36.

Bengis-Garber, C. and D. J. Kushner. 1982. Role of membrane-bound 5'-nucleotidase in nucleotide uptake by the moderate halophile Vibrio costicola. Journal of Bacteriology **149**:808-815.

Bennett, S. J., R. W. Sanders, and K. G. Porter. 1990. Heterotrophic, autotrophic and mixotrophic nanoflagellates - seasonal abundances and bacterivory in a eutrophic lake. Limnology and Oceanography **35**:1821-1832.

Bertilsson, S., O. Berglund, D. M. Karl, and S. W. Chisholm. 2003. Elemental Composition of Marine *Prochlorococcus* and *Synechococcus*: Lmplications for the Ecological Stoichiometry of the Sea. Limnology and Oceanography **48**:1721-1731.

Bird, D. F. and J. Kalff. 1986. Bacterial grazing by planktonic lake algae. Science **231**:493-495.

Bjorkman, K. M.,S. Duhamel, and D. M. Karl. 2012. Microbial group specific uptake kinetics of inorganic phosphate and adenosine-5-triphosphate (ATP) in the North Pacific Subtropical Gyre. Frontiers in Microbiology **3**: 189

Bouvier, T., S. Becquevort, and C. Lancelot. 1996. Biomass and feeding activity of phagotrophic mixotrophs in the northwestern Black Sea during the summer 1995. Pages 289-301 *in* International PELAG Symposium. Kluwer Academic Publication, Helsinki, Finland.

Box, G., G. Jenkins, and G. Reinsel. 1994. Time Series Analysis: Forecasting and Control. 3rd edition. Prentice-Hall, Upper Saddle River, New Jersey.

Box, G. E. P. and D. R. Cox. 1964. An Analysis of Transformations. Journal of the Royal Statistical Society. Series B (Methodological) **26**:211-252.

Boyd, P. W. and P. P. Newton. 1999. Does planktonic community structure determine downward particulate organic carbon flux in different oceanic provinces? Deep Sea Research Part I **46**:63-91.

de Boyer Montégut, C., G. Madec, A. S. Fischer, A. Lazar, and D. Iudicone. 2004. Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. Journal of Geophysical Research **109**:C12003.

de Boyer Montégut, C., J. Mignot, A. Lazar, and S. Cravatte. 2007. Control of salinity on the mixed layer depth in the world ocean: 1. General description. Journal of Geophysical Research **112**:C06011.

Britannica, E. 2012. Ockham's razor. Encyclopædia Britannica Online. Encyclopædia Britannica Inc.

Casey, J., M. Lomas, V. Michelou, S. Dyhrman, E. Orchard, J. Ammerman, and J. Sylvan. 2009. Phytoplankton taxon-specific orthophosphate (Pi) and ATP utilization in the western subtropical North Atlantic. Aquatic Microbial Ecology **58**:31-44.

Chatfield, C. 2004. The analysis of time series: an introduction, sixth edition. Chapman & Hall/CRC.

Chisholm, S. W., R. J. Olson, E. R. Zettler, R. Goericke, J. B. Waterbury, and N. A. Welschmeyer. 1988. A novel free-living prochlorophyte abundant in the oceanic euphotic zone. Nature **334**:340-343.

Christaki, U., F. Van Wambeke, and J. R. Dolan. 1999. Nanoflagellates (mixotrophs, heterotrophs and autotrophs) in the oligotrophic eastern Mediterranean: standing stocks, bacterivory and relationships with bacterial production. Marine Ecology Progress Series **181**:297-307.

Clark, L. L., E. D. Ingall, and R. Benner. 1998. Marine phosphorus is selectively remineralized. Nature **393**:426-426.

Clarke, K. R. 1993. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology **18**:117-143.

Clarke, K. R. and R. N. Gorley. 2005. PRIMER: Getting started with v6. PRIMER-E Ltd, Plymouth, UK.

Clarke, K. R. and R. N. Gorley. 2006. PRIMER v6: user manual/tutorial. PRIMER-E Ltd, Plymouth, UK.

Clarke, K. R. and R. H. Green. 1998. Statistical design and analysis for a "biological effects" study Marine Ecology Progress Series **46**:213-226.

Clarke, K. R., P. J. Somerfield, and R. N. Gorley. 2008. Testing of null hypotheses in exploratory community analyses: similarity profiles and biota-environment linkage. Journal of Experimental Marine Biology and Ecology **366**:56-69.

Clarke, K. R. and R. M. Warwick. 2001. Changes in marine communities: an approach to statistical analysis and interpretation, 2nd edition. PRIMER-E Ltd, Plymouth.

Cole, H., S. A. Henson, A. P. Martin, and A. Yool. 2012. Mind the gap: The impact of missing data on the calculation of phytoplankton phenology metrics. Journal of Geophysical Research Oceans **117**:C08030

Cole, J. J. 1982. Interactions Between Bacteria and Algae in Aquatic Ecosystems. Annual Review of Ecology and Systematics **13**:291-314.

Colling, A., editor. 2002. Ocean Circulation. Second edition. The Open University and Butterworth-Heinemann, Milton Keynes.

Crane, K. W. and J. P. Grover. 2010. Coexistence of mixotrophs, autotrophs, and heterotrophs in planktonic microbial communities. Journal of Theoretical Biology **262**:517-527.

Cressie, N. A. C. 1993. Statistics for Spatial Data Revised Edition. John Wiley & Sons, Inc, New York, USA.

Duce, R. A., P. S. Liss, J. T. Merrill, E. L. Atlas, P. Buat-Menard, B. B. Hicks, J. M. Miller, J. M. Prospero, R. Arimoto, T. M. Church, W. Ellis, J. N. Galloway, L. Hansen, T. D. Jickells, A. H. Knap, K. H. Reinhardt, B. Schneider, A. Soudine, J. J. Tokos, S. Tsunogai, R. Wollast, and M. Zhou. 1991. The atmospheric input of trace species to the world ocean. Global Biogeochemical Cycles **5**:193-259.

Dutkiewicz, S., M. Follows, J. Marshall, and W. W. Gregg. 2001. Interannual variability of phytoplankton abundances in the North Atlantic. Deep Sea Research Part II **48**:2323-2344.

Dutkiewicz, S., M. J. Follows, and J. G. Bragg. 2009. Modeling the coupling of ocean ecology and biogeochemistry. Global Biogeochemical Cycles **23**.

Eccleston-Parry, J. and B. Leadbeater. 1995. Regeneration of Phosphorus and Nitrogen by Four Species of Heterotrophic Nanoflagellates Feeding on Three Nutritional States of a Single Bacterial Strain. Applied and Environmental Microbiology **61**:1033-1038.

Edvardsen, A., M. Zhou, K. S. Tande, and Y. Zhu. 2002. Zooplankton population dynamics: measuring in situ growth and mortality rates using an Optical Plankton Counter. Marine Ecology Progress Series **227**:205-219.

Edwards, A. M. and A. Yool. 2000. The role of higher predation in plankton population models. Journal of Plankton Research **22**:1085-1112.

Einstein, A. 1934. On the Method of Theoretical Physics. Philosophy of Science **1**:163-169.

Epstein, S. S. and M. P. Shiaris. 1992. Size-selective grazing of coastal bacterioplankton by natural assemblages of pigmented flagellates, colorless flagellates, and ciliates. Microbial Ecology **23**:211-225.

Esaias, W. E., M. R. Abbott, I. Barton, O. B. Brown, J. W. Campbell, K. L. Carder, D. K. Clark, R. H. Evans, F. E. Hoge, H. R. Gordon, W. M. Balch, R. Letelier, and P. J. Minnett. 1998. An overview of MODIS capabilities for ocean science observations. Geoscience and Remote Sensing, IEEE Transactions on **36**:1250-1265.

Evans, G. T. and J. S. Parslow. 1985. A model of annual plankton cycles. Biological Oceanography **3**:327-347.

Fasham, M. 1993. Modelling the Marine Biota. Pages 457-504 *in* M. Heimann, editor. The Global Carbon Cycle. Springer-Verlag, Berlin.

Fasham, M. J. R., H. W. Ducklow, and S. M. McKelvie. 1990. A nitrogen-based model of plankton dynamics in the oceanic mixed layer. Journal of Marine Research **48**:591-639.

Feldman, G., N. Kuring, C. Ng, W. Esaias, C. McClain, J. Elrod, N. Maynard, D. Endres, R. Evans, J. Brown, S. Walsh, M. Carle, and G. Podesta. 1989. Ocean color: Availability of the global data set. Eos Trans. AGU **70**:634-641.

Field, J. G., K. R. Clarke, and R. M. Warwick. 1982. A Practical Strategy for Analysing Multispecies Distribution Patterns. Marine Ecology Progress Series **8**:37-52.

Fleming, R. H. 1939. The Control of Diatom Populations by Grazing. Journal du Conseil **14**:210-227.

Floder, S., T. Hansen, and R. Ptacnik. 2006. Energy-dependent bacterivory in Ochromonas minima - A strategy promoting the use of substitutable resources and survival at insufficient light supply. Protist **157**:291-302.

Flynn, K. J. and A. Mitra. 2009. Building the "perfect beast": modelling mixotrophic plankton. Journal of Plankton Research **31**:965-992.

Follows, M. J. and S. Dutkiewicz. 2011. Modeling Diverse Communities of Marine Microbes. Annual Review of Marine Science **3**:427-451.

Follows, M. J., S. Dutkiewicz, S. Grant, and S. W. Chisholm. 2007. Emergent biogeography of microbial communities in a model ocean. Science **315**:1843-1846.

Fowler, J., L. Cohen, and P. Jarvis. 1998. Practical Statistics for Field Biology, Second Edition. John Wiley and Sons

Franks, P. 2002. NPZ Models of Plankton Dynamics: Their Construction, Coupling to Physics, and Application. Journal of Oceanography **58**:379-387.

Franks, P. J. S. 2009. Planktonic ecosystem models: perplexing parameterizations and a failure to fail. Journal of Plankton Research **31**:1299-1306.

Friedline, C. J., R. B. Franklin, S. L. McCallister, and M. C. Rivera. 2012. Microbial community diversity of the eastern Atlantic Ocean reveals geographic differences. Biogeosciences Discussion **9**:109-150.

Friedrichs, M. A. M., J. A. Dusenberry, L. A. Anderson, R. A. Armstrong, F. Chai, J. R. Christian, S. C. Doney, J. Dunne, M. Fujii, R. Hood, D. J. McGillicuddy, Jr., J. K. Moore, M. Schartau, Y. H. Spitz, and J. D. Wiggert. 2007. Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. Journal of Geophysical Research **112**:C08001.

Fuhrman, J. A. and A. Hagström. 2008. Bacterial and Archael community structure and its patterns. Page 593 *in* D. Kirchman, editor. Microbial Ecology of the Oceans. John Wiley and Sons.

Fulton, E. A., A. D. M. Smith, and C. R. Johnson. 2003. Effect of complexity on marine ecosystem models. Marine Ecology Progress Series **253**:1-16.

Garcia, H. E., R. A. Locarnini, T. P. Boyer, and J. I. Antonov, editors. 2006. World Ocean Atlas 2005: Nutrients (phosphate, nitrate, silicate). U.S. Government Printing Office, Washington, D.C.

Garcia, H. E., R. A. Locarnini, T. P. Boyer, J. I. Antonov, M. M. Zweng, O. K. Baranova, and D. R. Johnson. 2010. World Ocean Atlas 2009: Nutrients (phosphate, nitrate, and silicate). U.S. Government Printing Office, Washington, D.C.

Geider, R. and J. La Roche. 2002. Redfield revisited: variability of C:N:P in marine microalgae and its biochemical basis. European Journal of Phycology **37**:1 - 17.

Gentleman, W. 2002. A chronology of plankton dynamics *in silico*: how computer models have been used to study marine ecosystems. Hydrobiologia **480**:69-85.

Goericke, R. and N. A. Welschmeyer. 1993. The marine prochlorophyte *Prochlorococcus* contributes significantly to phytoplankton biomass and primary production in the Sargasso Sea. Deep Sea Research Part I **40**:2283-2294.

Gomez-Pereira, P. R., B. M. Fuchs, C. Alonso, M. J. Oliver, J. E. E. van Beusekom, and R. Amann. 2010. Distinct flavobacterial communities in contrasting water masses of

the North Atlantic Ocean. Journal of the International Society for Microbial Ecology **4**:472-487.

Grob, C., M. Hartmann, M. V. Zubkov, and D. J. Scanlan. 2011. Invariable biomass-specific primary production of taxonomically discrete picoeukaryote groups across the Atlantic Ocean. Environmental Microbiology **13**:3266-3274.

Gundersen, K., M. Heldal, S. Norland, D. A. Purdie, and A. H. Knap. 2002. Elemental C, N, and P Cell Content of Individual Bacteria Collected at the Bermuda Atlantic Time-Series Study (BATS) Site. Limnology and Oceanography **47**:1525-1530.

Hall, J. A., D. P. Barrett, and M. R. James. 1993. The importance of phytoflagellate, heterotrophic flagellate and ciliate grazing on bacteria and picophytoplankton sized prey in a coastal marine-environment. Journal of Plankton Research **15**:1075-1086.

Hammer, A. C. and J. W. Pitchford. 2005. The role of mixotrophy in plankton bloom dynamics, and the consequences for productivity. Journal of Marine Science **62**:833-840.

Hartmann, M., C. Grob, D. J. Scanlan, A. P. Martin, P. H. Burkill, and M. V. Zubkov. 2011. Comparison of phosphate uptake rates by the smallest plastidic and aplastidic protists in the North Atlantic subtropical gyre. Federation of European Microbiological Societies Microbiology Ecology **78**: 327-335.

Hartmann, M., C. Grob, G. A. Tarran, A. P. Martin, P. H. Burkill, D. J. Scanlan, and M. V. Zubkov. 2012. Mixotrophic basis of Atlantic oligotrophic ecosystems. Proceedings of the National Academy of Sciences **109**:5756-5760.

Havskum, H. and A. S. Hansen. 1997. Importance of pigmented and colourless nano-sized protists as grazers on nanoplankton in a phosphate-depleted Norwegian fjord and in enclosures. Aquatic Microbial Ecology **12**:139-151.

Havskum, H. and B. Riemann. 1996. Ecological importance of bacterivorous, pigmented flagellates (mixotrophs) in the Bay of Aarhus, Denmark. Marine Ecology Progress Series **137**:251-263.

Heldal, M., D. J. Scanlan, S. Norland, F. Thingstad, and N. H. Mann. 2003. Elemental Composition of Single Cells of Various Strains of Marine *Prochlorococcus* and *Synechococcus* Using X-Ray Microanalysis. Limnology and Oceanography **48**:1732-1743.

Henson, S. A., J. P. Dunne, and J. L. Sarmiento. 2009. Decadal variability in North Atlantic phytoplankton blooms. Journal of Geophysical Research-Oceans **114**: C04013.

Holligan, P. 2004. AMT14 Cruise Report. Atlantic Meridional Transect. Natural Environment Research Council.

Holligan, P. 2005. AMT17 Cruise Report. Atlantic Meridional Transect. Natural Environment Research Council.

Hood, R. R., V. J. Cole, and T. F. Gross. 2008. Marine Models Pages 2217-2225 *in* S. E. Jorgensen and B. Fath, editors. Encyclopedia of Ecology. Elsevier, Oxford, UK.

Hood, R. R., E. A. Laws, R. A. Armstrong, N. R. Bates, C. W. Brown, C. A. Carlson, F. Chai, S. C. Doney, P. G. Falkowski, R. A. Feely, M. A. M. Friedrichs, M. R. Landry, J. K. Moore, D. M. Nelson, T. L. Richardson, B. Salihoglu, M. Schartau, D. A. Toole, and J. D. Wiggert. 2006a. Pelagic functional group modeling: Progress, challenges and prospects. Deep-Sea Research Part II **53**:459-512.

Hood, R. R., X. Zhang, P. A. Glibert, M. R. Roman, and D. K. Stoecker. 2006b. Modeling the influence of nutrients, turbulence and grazing on Pfiesteria population dynamics. Harmful Algae **5**:459-479.

Jardillier, L., M. V. Zubkov, J. Pearman, and D. J. Scanlan. 2010. Significant CO_2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. The International Society for Microbial Ecology Journal **4**:1180-1192.

Jenkins, W. J. 1988. Nitrate flux into the euphotic zone near Bermuda. Nature **331**:521-523.

Johnson, Z. I., E. R. Zinser, A. Coe, N. P. McNulty, E. M. S. Woodward, and S. W. Chisholm. 2006. Niche Partitioning Among *Prochlorococcus* Ecotypes Along Ocean-Scale Environmental Gradients. Science **311**:1737-1740.

Jones, R. I. 1994. Mixotrophy in planktonic protists as a spectrum of nutritional strategies. Marine Microbial Food Webs **8**:87-96.

Jost, C., C. A. Lawrence, F. Campolongo, W. van de Bund, S. Hill, and D. L. DeAngelis. 2004. The effects of mixotrophy on the stability and dynamics of a simple planktonic food web model. Theoretical Population Biology **66**:37-51.

Karl, D. M. 1999. A sea of change: Biogeochemical variability in the North Pacific Subtropical Gyre. Ecosystems **2**:181-214.

Karl, D. M. and K. M. Björkman. 2002. Chapter 6 - Dynamics of DOP. *in* A. H. Dennis and A. C. Craig, editors. Biogeochemistry of Marine Dissolved Organic Matter. Academic Press, San Diego.

Karl, D. M. and P. Bossard. 1985. Measurement and significance of ATP and Adenine-nucleotide pool turnover in microbial cells and environmental samples. Journal of Microbiological Methods **3**:125-139.

Karl, D. M. and G. Tien. 1992. Magic: A Sensitive and Precise Method for Measuring Dissolved Phosphorus in Aquatic Environments. Limnology and Oceanography **37**:105-116.

Kidston, K. 2010. Parameter Optimisation of a Marin Ecosystem Model in the Southern Ocean. Doctor of Philosophy. University of New South Wales, Australia.

Kolowith, L. C., E. D. Ingall, and R. Benner. 2001. Composition and Cycling of Marine Organic Phosphorus. Limnology and Oceanography **46**:309-320.

Kooijman, S., T. Andersen, and B. W. Kooi. 2004. Dynamic energy budget representations of stoichiometric constraints on population dynamics. Ecology **85**:1230-1243.

Kooijman, S. A. L. M., P. Auger, J. C. Poggiale, and B. W. Kooi. 2003. Quantitative steps in symbiogenesis and the evolution of homeostasis. Biological Reviews **78**:435-463.

Kriest, I., S. Khatiwala, and A. Oschlies. 2010. Towards an assessment of simple global marine biogeochemical models of different complexity. Progress In Oceanography **86**:337-360.

Kuenzler, E. J. and J. P. Perras. 1965. Phosphatases of Marine Algae. The Biological Bulletin **128**:271-284.

Legrand, C., N. Johansson, G. Johnsen, K. Y. Borsheim, and E. Graneli. 2001. Phagotrophy and Toxicity Variation in the Mixotrophic Prymnesium patelliferum (Haptophyceae). Limnology and Oceanography **46**:1208-1214.

Lepère, C., D. Vaulot, and D. J. Scanlan. 2009. Photosynthetic picoeukaryote community structure in the South East Pacific Ocean encompassing the most oligotrophic waters on Earth. Environmental Microbiology **11**:3105-3117.

Lewis, J. M., P. M. Hull, K. Q. Weinberger, and L. K. Saul. 2008. Mapping Uncharted Waters: Exploratory Analysis, Visualization, and Clustering of Oceanographic Data. Pages 388-395. Proceedings of the 2008 Seventh International Conference on Machine Learning and Applications. IEEE Computer Society.

Li, W. K. W. 1994. Primary Production of Prochlorophytes, Cyanobacteria, and Eucaryotic Ultraphytoplankton: Measurements from Flow Cytometric Sorting. Limnology and Oceanography **39**:169-175.

Li, W. K. W. 1995. Composition of ultraphytoplankton in the central North Atlantic. Marine Ecology Progress Series **122**:1-8.

Li, W. K. W. and W. G. Harrison. 2001. Chlorophyll, bacteria and picophytoplankton in ecological provinces of the North Atlantic. Deep Sea Research Part II **48**:2271-2293.

Liu, H., H. A. Nolla, and L. Campbell. 1997. *Prochlorococcus* growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean. Aquatic Microbial Ecology **12**:39-47.

Liu, H., I. Probert, J. Uitz, H. Claustre, S. Aris-Brosou, M. Frada, F. Not, and C. de Vargas. 2009. Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. Proceedings of the National Academy of Sciences **106**:12803-12808.

Lomas, M. W., A. L. Burke, D. A. Lomas, D. W. Bell, C. Shen, S. T. Dyhrman, and J. W. Ammerman. 2010. Sargasso Sea phosphorus biogeochemistry: an important role for dissolved organic phosphorus (DOP). Biogeosciences **7**:695 - 710.

Longhurst, A., S. Sathyendranath, T. Platt, and C. Caverhill. 1995. An estimate of global primary production in the ocean from satellite radiometer data. Journal of Plankton Research **17**:1245-1271.

Longhurst, A. R. 1995. Seasonal cycles of pelagic production and consumption. Progress In Oceanography **36**:77-167.

Longhurst, A. R. 1998. Ecological Geography of the Sea. Academic Press, London.

Longhurst, A. R. 2007. Ecological Geography of the Sea, Second Edition. Academic Press, London.

Lutz, M. J., K. Caldeira, R. B. Dunbar, and M. J. Behrenfeld. 2007. Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. Journal of Geophysical Research **112**:C10011.

Mahaffey, C., R. G. Williams, G. A. Wolff, and W. T. Anderson. 2004. Physical supply of nitrogen to phytoplankton in the Atlantic Ocean. Global Biogeochemical Cycles **18**: GB1034.

Marañón, E., M. J. Behrenfeld, N. González, B. Mouriño, and M. V. Zubkov. 2003. High variability of primary production in oligotrophic waters of the Atlantic Ocean: uncoupling from phytoplankton biomass and size structure. Marine Ecology Progress Series **257**:1-11.

Marañón, E., P. Holligan, R. Barciela, N. González, B. Mouriño, M. J. Pazó, and M. Varela. 2001. Patterns of phytoplankton size structure and productivity in contrasting open-ocean environments. Marine Ecology Progress Series **216**:43-56.

Martin, A. P. 2003. Phytoplankton patchiness: the role of lateral stirring and mixing. Progress In Oceanography **57**:125-174.

Martin, A. P., M. V. Zubkov, P. H. Burkill, and R. J. Holland. 2005. Extreme spatial variability in marine picoplankton and its consequences for interpreting Eulerian time-series. Biology Letters **1**:366-369.

Martin, A. P., M. V. Zubkov, M. J. Fasham, P. H. Burkill, and R. J. Holland. 2008. Microbial spatial variability: An example from the Celtic Sea. Progress In Oceanography **76**:443-465.

Martin, A. P., M. V. Zubkov, R. J. Holland, G. Tarran, and P. Burkill. 2010. Variability in ultraplankton at the Porcupine Abyssal Plain study site. Deep Sea Research Part II **57**:1336-1345.

Martin, S. 2004. An introduction to Ocean Remote Sensing. Cambridge University Press, Cambridge, UK.

Mather, R. L., S. E. Reynolds, G. A. Wolff, R. G. Williams, S. Torres-Valdes, E. M. S. Woodward, A. Landolfi, X. Pan, R. Sanders, and E. P. Achterberg. 2008. Phosphorus cycling in the North and South Atlantic Ocean subtropical gyres. Nature Geoscience **1**:439-443.

McClain, C. R., S. R. Signorini, and J. R. Christian. 2004. Subtropical gyre variability observed by ocean-color satellites. Deep-Sea Research Part II **51**:281-301.

Michelou, V. K., M. W. Lomas, and D. L. Kirchman. 2011. Phosphate and adenosine-5'-triphosphate uptake by cyanobacteria and heterotrophic bacteria in the Sargasso Sea. Limnology and Oceanography **56**:323-332.

Mills, M. M., C. M. Moore, R. Langlois, A. Milne, E. Achterberg, K. Nachtigall, K. Lochte, R. J. Geider, and J. La Roche. 2008. Nitrogen and phosphorus co-limitation of bacterial productivity and growth in the oligotrophic subtropical North Atlantic. Limnology and Oceanography **53**:824-834.

Mills, M. M., C. Ridame, M. Davey, J. La Roche, and R. J. Geider. 2004. Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. Nature **429**:292-294.

Monterey, G. and S. Levitus. 1997. Seasonal variability of mixed layer depth for the worlds ocean. U.S. Department of Commerce, Washington, D.C.

Moorthi, S., D. A. Caron, R. J. Gast, and R. W. Sanders. 2009. Mixotrophy: a widespread and important ecological strategy for planktonic and sea-ice nanoflagellates in the Ross Sea, Antarctica. Aquatic Microbial Ecology **54**:269-277.

Morel, A., Y.-H. Ahn, F. Partensky, D. Vaulot, and H. Claustre. 1993. *Prochlorococcus* and *Synechococcus*: A comparative study of their optical properties in relation to their size and pigmentation. Journal of Marine Research **51**:617-649.

Morris, R. M., M. S. Rappe, S. A. Connon, K. L. Vergin, W. A. Siebold, C. A. Carlson, and S. J. Giovannoni. 2002. SAR11 clade dominates ocean surface bacterioplankton communities. Nature **420**:806-810.

Munk, W. H. 1950. On the wind-driven ocean circulation. Journal of Meteorology **7:** 79–93.

Nygaard, K. and A. Tobiesen. 1993. Bacterivory in algae - A survival strategy during nutrient limitation. Limnology and Oceanography **38**:273-279.

Obayashi, Y. and E. Tanoue. 2002. Growth and mortality rates of phytoplankton in the northwestern North Pacific estimated by the dilution method and HPLC pigment analysis. Journal of Experimental Marine Biology and Ecology **280**:33-52.

Orchard, E. D., J. W. Ammerman, D. A. Lomas, and S. T. Dyhrman. 2010. Dissolved inorganic and organic phosphorus uptake in Trichodesmium and the microbial community: The importance of phosphorus ester in the Sargasso Sea. Limnology and Oceanography. **55**:1390-1399.

Oschlies, A. and V. Garcon. 1998. Eddy-induced enhancement of primary production in a model of the North Atlantic Ocean. Nature **394**:266-269.

Oschlies, A., W. Koeve, and V. Garcon. 2000. An eddy-permitting coupled physical-biological model of the North Atlantic 2. Ecosystem dynamics and comparison with satellite and JGOFS local studies data. Global Biogeochemical Cycles **14**:499-523.

OSPAR, C. 2008. Nutrient reduction scenarios for the North Sea: Environmental consequences for problems areas with regard to eutrophication following nutrient reductions in models scenarios.*in* OSPAR Commission, Den Haag, The Netherlands.

Pahlow, M. and A. Oschlies. 2009. Chain model of phytoplankton P, N and light colimitation. Marine Ecology Progress Series **376**:69-83.

Partensky, F., J. Blanchot, and D. Vaulot. 1999a. Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters : a review. Musée océanographique, Monaco, MONACO.

Partensky, F., W. R. Hess, and D. Vaulot. 1999b. *Prochlorococcus*, a Marine Photosynthetic Prokaryote of Global Significance. Microbiology and Molecular Biology Reviews **63**:106-127.

Pauly, D. 1998. Large marine ecosystems: analysis and management. South African Journal of Marine Science **19**:487-499.

Platt, T. and S. Sathyendranath. 1988. Oceanic Primary Production: Estimation by Remote Sensing at Local and Regional Scales. Science **241**:1613-1620.

Polovina, J. J., E. A. Howell, and M. Abecassis. 2008. Ocean's least productive waters are expanding. Geophysical Research Letters **35**:5.

Poulton, A. J., P. M. Holligan, A. Hickman, Y.-N. Kim, T. R. Adey, M. C. Stinchcombe, C. Holeton, S. Root, and E. M. S. Woodward. 2006. Phytoplankton carbon fixation, chlorophyll-biomass and diagnostic pigments in the Atlantic Ocean. Deep Sea Research Part II **53**:1593-1610.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical Recipes in FORTRAN The Art of Scientific Computing (Second Edition). Cambridge University Press, United States of America.

Radach, G. and A. Moll. 2006. Review of three-dimensional ecological modelling related to the north sea shelf system. Part II: Model validation and data needs. CRC Press, Boca Raton, FL, ETATS-UNIS.

Raven, J. A. 1997. Phagotrophy in phototrophs. Limnology and Oceanography **42**:198-205.

Redfield, A. C. 1934. On the proportions of organic derivations in sea water and their relation to the composition of plankton. Pages 177-192 *in* R. J. Daniel, editor. James Johnstone Memorial Volume. University Press of Liverpool.

Rees, A. P. 2009. AMT19 Cruise Report. Atlantic Meridonal Transect. Natural Environment Research Council.

Reynolds, S., C. Mahaffey, and R. G. Williams. 2012. Dissolved organic matter, an important Phosphorus source in oligotrophic gyres. Ocean Sciences Meeting 2012, Salt Lake City, Utah USA.

Richardson, T. L. and G. A. Jackson. 2007. Small Phytoplankton and Carbon Export from the Surface Ocean. Science **315**:838-840.

Riley, G. A. 1946. Factors controlling phytoplankton populations in Georges Bank. Journal of Marine Research **6**:54-73.

Robinson, C., A. J. Poulton, P. M. Holligan, A. R. Baker, G. Forster, N. Gist, T. D. Jickells, G. Malin, R. Upstill-Goddard, R. G. Williams, E. M. S. Woodward, and M. V. Zubkov. 2006. The Atlantic Meridional Transect (AMT) Programme: A contextual view 1995-2005. Deep Sea Research Part II **53**:1485-1515.

Robinson, I. S. 2004. Measuring the Oceans from Space. The principles and methods of satellite oceanography. Springer-Praxis, Chichester, UK.

Rothhaupt, K. O. 1996a. Laboratory experiments with a mixotrophic chrysophyte and obligately phagotrophic and phototrophic competitors. Ecology **77**:716-724.

Rothhaupt, K. O. 1996b. Utilization of Substitutable Carbon and Phosphorus Sources by the Mixotrophic Chrysophyte Ochromonas Sp Ecology **77**:706-715

Roussenov, V., R. G. Williams, C. Mahaffey, and G. A. Wolff. 2006. Does the transport of dissolved organic nutrients affect export production in the Atlantic Ocean? Global Biogeochemical Cycles **20**:GB3002.

Safi, K. A. and J. A. Hall. 1999. Mixotrophic and heterotrophic nanoflagellate grazing in the convergence zone east of New Zealand. Aquatic Microbial Ecology **20**:83-93.

Sanders, R. W. 1991. Mixotrophic protists in marine and freshwater ecosystems. Journal of Protozoology **38**:76-81.

Sanders, R. W., U. G. Berninger, E. L. Lim, P. F. Kemp, and D. A. Caron. 2000. Heterotrophic and mixotrophic nanoplankton predation on picoplankton in the Sargasso Sea and on Georges Bank. Marine Ecology Progress Series **192**:103-118.

Sanders, R. W. and K. G. Porter. 1988. Phagotrophic phytoflagellates. Pages 167-192 *in* K. C. Marshall, editor. Advances in Microbial Ecology. Plenum, New York.

Sarmiento, J. L. and N. Gruber. 2006. Ocean Biogeochemical Dynamics. Princeton University Press, Princeton, New Jersey and Woodstock, Oxfordshire.

Sarmiento, J. L., R. D. Slater, M. J. R. Fasham, H. W. Ducklow, J. R. Toggweiler, and G. T. Evans. 1993. A seasonal three-dimensional ecosystem model of nitrogen cycling in the North Atlantic Euphotic Zone. Global Biogeochemical Cycles **7**:417-450.

Sarthou, G., K. R. Timmermans, S. Blain, and P. Tréguer. 2005. Growth physiology and fate of diatoms in the ocean: a review. Journal of Sea Research **53**:25-42.

Schartau, M. and A. Oschlies. 2003a. Simultaneous data-based optimization of a 1D-ecosystem model at three locations in the North Atlantic: Part I-Method and parameter estimates. Journal of Marine Research **61**:765-793.

Schartau, M. and A. Oschlies. 2003b. Simultaneous data-based optimization of a 1D-ecosystem model at three locations in the North Atlantic: Part II-Standing stocks and nitrogen fluxes. Journal of Marine Research **61**:794-820.

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

Six, K. D. and E. Maier-Reimer. 1996. Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model. Global Biogeochemical Cycles **10**:559-583.

Steele, J. H. and E. W. Henderson. 1992a. The role of predation in plankton models. Journal of Plankton Research **14**:157-172.

Steele, J. H. and E. W. Henderson. 1992b. A simple model for plankton patchiness. Journal of Plankton Research **14**:1397-1403.

Stibor, H. and U. Sommer. 2003. Mixotrophy of a Photosynthetic Flagellate viewed from an Optimal Foraging Perspective. Protist **154**:91-98.

Stickney, H. L., R. R. Hood, and D. K. Stoecker. 2000. The impact of mixotrophy on planktonic marine ecosystems. Ecological Modelling **125**:203-230.

Stoecker, D. K. 1998. Conceptual Models of Mixotrophy in Planktonic Protists and some Ecological and Evolutionary Implications. European Journal of Protistology **34**:281-290.

Stommel, H. 1947. A summary of the theory of Convection Cells. Annals of the New York Academy of Sciences **48**:715-726.

Stommel, H. 1948. The westward intensification of wind-driven ocean currents. Transactions American Geophysical Union **29:**202–206.

Stukel, M. R., M. R. Landry, and K. E. Selph. 2011. Nanoplankton mixotrophy in the eastern equatorial Pacific. Deep Sea Research Part II **58**:378-386.

Sverdrup, H. U., M. W. Johnson, and R. H. Fleming. 1942. The Oceans Their Physics, Chemistry, and General Biology. Prentice-Hall, Inc, New York.

Tarran, G. A., J. L. Heywood, and M. V. Zubkov. 2006. Latitudinal changes in the standing stocks of nano- and picoeukaryotic phytoplankton in the Atlantic Ocean. Deep Sea Research Part II **53**:1516-1529.

Taylor, A. H. and I. Joint. 1990. A steady-state analysis of the 'microbial loop' in stratified systems. Marine Ecology Progress Series **59**:1-17.

Thain, D., T. Tannenbaum, and M. Livny. 2005. Distributed computing in practice: the Condor experience. Concurrency and Computation: Practice and Experience **17**:323-356.

Thingstad, T. F., H. Havskum, K. Garde, and B. Riemann. 1996. On the Strategy of "Eating Your Competitor": A Mathematical Analysis of Algal Mixotrophy Ecological Society of America **77**:2108-2118.

Thomalla, S., R. Turnewitsch, M. Lucas, and A. Poulton. 2006. Particulate organic carbon export from the North and South Atlantic gyres: The $^{234}Th/^{238}U$ disequilibrium approach. Deep Sea Research Part II **53**:1629-1648.

Tilstone, G., T. Smyth, A. Poulton, and R. Hutson. 2009. Measured and remotely sensed estimates of primary production in the Atlantic Ocean from 1998 to 2005. Deep Sea Research Part II: Topical Studies in Oceanography **56**:918-930.

Tobler, W. R. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography **46**:234-240.

Torres-Valdés, S., V. M. Roussenov, R. Sanders, S. Reynolds, X. Pan, R. Mather, A. Landolfi, G. A. Wolff, E. P. Achterberg, and R. G. Williams. 2009. Distribution of dissolved organic nutrients and their effect on export production over the Atlantic Ocean. Global Biogeochemical Cycles **23**:GB4019.

Tsai, A. Y., G. C. Gong, R. W. Sanders, W. H. Chen, C. F. Chao, and K. P. Chiang. 2011. Importance of bacterivory by pigmented and heterotrophic nanoflagellates during the warm season in a subtropical western Pacific coastal ecosystem. Aquatic Microbial Ecology **63**:9-18.

Turner, J. T. 2002. Zooplankton fecal pellets, marine snow and sinking phytoplankton blooms. Aquatic Microbial Ecology **27**:57-102.

Tyrrell, T. 1999. The relative influences of nitrogen and phosphorus on oceanic primary production. Nature **400**:525-531.

Unrein, F., R. Massana, L. Alonso-Saez, and J. M. Gasol. 2007. Significant year-round effect of small mixotrophic flagellates on bacterioplankton in an oligotrophic coastal system. Limnology and Oceanography **52**:456-469.

Van Mooy, B. A. S., H. F. Fredricks, B. E. Pedler, S. T. Dyhrman, D. M. Karl, M. Koblizek, M. W. Lomas, T. J. Mincer, L. R. Moore, T. Moutin, M. S. Rappe, and E. A. Webb. 2009. Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. Nature **458**:69-72.

Veldhuis, M. J. W. and G. W. Kraay. 2004. Phytoplankton in the subtropical Atlantic Ocean: towards a better assessment of biomass and composition. Deep Sea Research Part I **51**:507-530.

Vichi, M., J. I. Allen, S. Masina, and N. J. Hardman-Mountford. 2011. The emergence of ocean biogeochemical provinces: A quantitative assessment and a diagnostic for model evaluation. Global Biogeochemical Cycles **25**:GB2005.

Wang, Z. H., Y. Liang, and W. Kang. 2011. Utilization of dissolved organic phosphorus by different groups of phytoplankton taxa. Harmful Algae **12**:113-118.

Ward, B. A., M. A. M. Friedrichs, T. R. Anderson, and A. Oschlies. 2010. Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models. Journal of Marine Systems **81**:34-43.

Woodd-Walker, R. S., P. Ward, and A. Clarke. 2002. Large-scale patterns in diversity and community structure of surface water copepods from the Atlantic Ocean. Inter-Research, Oldendorf, ALLEMAGNE.

Woodward, M. 2008. AMT18 Cruise Report. Atlantic Meridonal Transect. Natural Environment Research Council.

Wright, R. T. and J. E. Hobbie. 1966. Use of glucose and acetate by bacteria and algae in aquatic ecosystems. Ecology **47**:447-464.

Wu, J., W. Sunda, E. A. Boyle, and D. M. Karl. 2000. Phosphate Depletion in the Western North Atlantic Ocean. Science **289**:759-762.

Wulff, F., J. G. Field, and K. H. Mann, editors. 1989. Network Analysis in Marine Ecology. Springer-Verlag, Berlin.

Yoder, J. A., C. R. McClain, G. C. Feldman, and W. E. Esaias. 1993. Annual cycles of phytoplankton chlorophyll concentrations in the global ocean: A satellite view. Global Biogeochemical Cycles **7**:181-193.

Zubkov, M. V., B. M. Fuchs, G. A. Tarran, P. H. Burkill, and R. Amann. 2002. Mesoscale distribution of dominant bacterioplankton groups in the northern North Sea in early summer. Aquatic Microbial Ecology **29**: 135–144

Zubkov, M. V. and R. J. G. Leakey. 2009. Evaluation of the efficiency of metabolism of dinoflagellate phosphorus and carbon by a planktonic ciliate. European Journal of Protistology **45**:166-173.

Zubkov, M. V., I. Mary, E. M. S. Woodward, P. E. Warwick, B. M. Fuchs, D. J. Scalan, and P. H. Burkhill. 2007. Microbial control of phosphate in the nutrient-depleted North Atlantic subtropical gyre. Environmental Microbiology **9**:2079-2089.

Zubkov, M. V. and G. D. Quartly. 2003. Ultraplankton distribution in surface waters of the Mozambique Channel‹flow cytometry and satellite imagery. Aquatic Microbial Ecology **33**:155-161.

Zubkov, M. V., M. A. Sleigh, P. H. Burkill, and R. J. G. Leakey. 2000. Picoplankton community structure on the Atlantic Meridional Transect: a comparison between seasons. Progress In Oceanography **45**:369-386.

Zubkov, M. V., M. A. Sleigh, G. A. Tarran, P. H. Burkill, and R. J. G. Leakey. 1998. Picoplanktonic community structure on an Atlantic transect from 50°N to 50°S. Deep Sea Research Part I **45**:1339-1355.

Zubkov, M. V. and G. A. Tarran. 2008. High bacterivory by the smallest phytoplankton in the North Atlantic Ocean. Nature **455**:224-226.

Zubkov, M. V., G. A. Tarran, and B. M. Fuchs. 2004. Depth related amino acid uptake by *Prochlorococcus* cyanobacteria in the Southern Atlantic tropical gyre. Federation of European Microbiological Societies Microbiology Ecology **50**:153-161.