Phytoplankton community structure determined through remote sensing and *in situ* optical measurements

Hayley Evers-King

Thesis presented for the degree of Doctor of Philosophy

Department of Oceanography
University of Cape Town
August 2014

Declaration

The present work has been originally written by me, with the full support of my supervisors: Professor Frank Shillington of the Department of Oceanography, University of Cape Town, South Africa; Dr Stewart Bernard of the Council for Scientific and Industrial Research, South Africa; Dr Stephanie Henson of the National Oceanography Centre, Southampton, United Kingdom and A/Professor Mike Lucas of the Department of Biological Sciences, University of Cape Town, South Africa. In addition, important contributions to the approach used in this study are clearly acknowledged through referencing within the text.

Abstract

Linking variability in optical signals with phytoplankton community characteristics is important to extend the use of the vast resource that is the satellite ocean colour archive. Detection of species, functional types or size classes has been addressed through a spectrum of empirical to analytical approaches. A key step in developing these techniques is quantifying the sensitivity in reflectance, which can be attributed to phytoplankton characteristics (e.g cell size) under different optical regimes. Ultimately, highly spatially and temporally resolved information on phytoplankton characteristics can help the global scientific community to answer important questions relating to primary ecosystem variability. In the southern Benguela, Harmful Algal Blooms threaten public health and the economic viability of fishery and aquaculture industries in the region. Concurrently, the dominance of phytoplankton biomass amongst optically significant constituents in the southern Benguela makes the region ideal for assessing the extent to which phytoplankton characteristics beyond biomass can influence the ocean colour signal.

A forward and inverse approach is presented. Phytoplankton absorption and backscattering are generated from a phytoplankton particle population model coupled to two radiative transfer approaches: a reflectance approximation and the radiative transfer model, EcoLight-S. Non-linear optimisation inversion schemes are then implemented. A simulated dataset is created to investigate how much variability in reflectance can be associated with changes in phytoplankton cell size in different bio-optical water types. This dataset is inverted to investigate the errors inherent in the inversion process as a result of ambiguity. Comparison of the two radiative transfer techniques allows for consideration of the suitability of approximations for bidirectionality and subsurface propagation. The inversion algorithm is then applied to hyperspectral *in situ* radiometric data to provide validation and further assessment of errors from all sources. Results indicate that size related sensitivity in reflectance is highly dependent on phytoplankton biomass, as determined by the relative phytoplankton contribution to the Inherent Optical Property budget. The algorithm is finally applied to ten years of MERIS data covering the southern Benguela. A time series of biomass and cell size is presented and metrics developed to demonstrate the utility of this approach for identifying previously unobserved interannual variability in Harmful Algal Blooms.

Acknowledgements

For extensive academic and personal support and advice, I would like to thank my supervisors; Stewart Bernard, Stephanie Henson, Frank Shillington and Mike Lucas.

For data, advice and field work assistance, I would like to acknowledge the substantial support I received from Grant Pitcher, Trevor Probyn, Andre du Randt, Vinesh Rajpaul, Maya Pfaff, Sandy Thomalla, Stanford Hooker, Jim Brown, John Bruun and Andy Rabagliati.

For direct funding I gratefully acknowledge: The National Research Foundation (NRF), the Marine-Research (MA-RE) Institute at the University of Cape Town and the Applied Centre for Climate and Earth System Science (ACCESS).

For additional financial and technical support I gratefully acknowledge: The Council for Scientific and Industrial Research (CSIR) and the Centre for High Performance Computing (CHPC), the Nansen-Tutu centre, the Department for Agriculture Forestry and Fisheries (DAFF) and the South African Weather Service (SAWS).

For never-ending academic advice, field assistance, cups of tea and a friendly ear to listen to my frustrations; I would like to thank various fellow graduate students and friends but in particular: Emily McGregor, Lauren Biermann, Marie Smith, Lisl Robertson Lain, Brett Kuyper, Mark Matthews, Janine Basson, Christo Whittle, Flavia Flaviani, Ffion Atkins, Theoni Photopoulou and Hazel Dickens.

For tolerating my absence from too many christmas days, and their unwavering support and confidence in my academic abilities; my thanks go to my parents Joanne Evers-King and Michael Lloyd-King, their respective partners Jonathan Hill and Kath Lloyd-King, Thomas Evers-King, Annie Loveday and Lou Goodacre.

For encouraging my love of natural science and filling my childhood with stories of South Africa; I dedicate this work to my grandparents Marianne and Tom Evers.

Final thanks go to Ben Loveday, whose love, support and mad coding skills got me in to this and out the other side. Thank you for this adventure, I'm looking forward to many more!

List of notation

a	-	Total absorption coefficient (m^{-1})
a_{gd}	-	Combined gelbstoff and detritus absorption coefficient (m^{-1})
a_{ϕ}	-	Phytoplankton absorption coefficient (m^{-1})
ASF	-	Area scaling factor for size distribution manipulation (dimensionless)
b	-	Total scattering coefficient (m^{-1})
b_{bp}	-	Particulate backscattering coefficient (m^{-1})
b_{bs}	-	Small particle backscattering coefficient (m^{-1})
$b_{b\phi}$	-	Phytoplankton backscattering coefficient (m^{-1})
c	-	Beam attenuation coefficient (m^{-1})
c_i	-	Intracellular chlorophyll a concentration (kg m^{-3})
[Chl a]	-	Chlorophyll a concentration (mg m^{-3})
C-OPS	-	Biospherical Instruments Compact Optical Profiling System
d	-	Particle diameter (μm)
DCI	-	Inherent Optical Property basis vectors generated with dynamic c_i parameterisation
D_{eff}	-	Effective diameter (μm)
E_d	-	Downwelling irradiance (μW cm^{-2} nm^{-1} sr^{-1})
EAP	-	Equivalent Algal Population
ϵ	-	Total self shading error (%)
ϵ_{sky}	-	Direct sunlight error (%)
ϵ_{sun}	-	Diffuse skylight error (%)
ES	-	EcoLight-S radiative transfer model
f	-	Ratio of downwelling irradiance from skylight to that from direct sunlight (dimensionless)
F(d)	-	Particle size distribution (number of particles per unit volume per micrometer, # m^{-3} μm^{-1})
F*(d)	-	Chlorophyll a specific size distribution (specific number of particles per unit volume per micrometer, # m^{-3} μm^{-1} mg^{-1})
f/Q	-	Bidirectionality of upwelling radiance (sr^{-1})

FWD_A	-	Set of data simulated using the EAP forward model over four broad water types
FWD_N	-	Set of data simulated using the EAP forward model over water types representing those in the NOMAD dataset
H-TSRB	-	Satlantic Hyperspectral Tethered Surface Radiometer Buoy
JCI	-	Basis vectors generated with Jungian distribution with a [Chl a] dependent admixture
K_d	-	Downward diffuse attenuation coefficient (m^{-1})
K_u	-	Upward diffuse attenuation coefficient (m^{-1})
KL_u	-	Upward radiance attenuation coefficient (m^{-1})
L_u	-	Upwelling radiance (μW cm^{-2} nm^{-1} sr^{-1})
$\text{PFT}_{\%}$	-	Percentage of the two functional type groups used in admixture for the forward and inverse EAP approaches
ν_{eff}	-	Effective variance of the particle size distribution (dimensionless)
REFA	-	Reflectance approximation
R_{rs}	-	Remote sensing reflectance (sr^{-1})
SCI	-	Basis vectors generated with static c_i
θ	-	Solar zenith angle (^{o}, degrees)
z	-	Depth (m)

Contents

Chapter 1

Introduction

Substantial challenges are presented by climate change, declining water quality and an increasing reliance on ocean ecosystems for economic activities including fisheries and aquaculture. To understand and address these challenges requires observations of ocean biogeochemistry and ecology. Primary ecosystem variability i.e. in phytoplankton communities, is a key component of this. The structure of phytoplankton communities (in terms of cell size or functional type), has been linked with physical, chemical and biological processes of interest, such as carbon export and higher trophic level dynamics. Phytoplankton community structure varies over a broad range of spatial and temporal scales. Capturing this variability requires observations which cover this range of scales. The last two decades have seen a huge growth in the provision and use of remotely sensed data for oceanographic studies. In particular, ocean colour data have provided a way to investigate phytoplankton communities. Algorithms have been developed to estimate biomass and other characteristics of phytoplankton communities, including cell size and the presence of particular species or functional types. However, challenges in the use of this data resource remain. In optically complex regions quantifying phytoplankton biomass remains difficult and there is uncertainty surrounding exactly how phytoplankton characteristics beyond biomass influence optical signals. Combining optical theory with phytoplankton particle population and radiative transfer models should allow for a greater understanding of the sensitivity of ocean colour measurements to phytoplankton community characteristics and how this sensitivity can be used and applied through algorithms.

1.1 Phytoplankton cell size and its importance as a metric in global ocean ecosystems

Phytoplankton cell size covers a range of nine orders of magnitude, a comparatively vast range compared to other natural and human scales (Fig. 1.1) (Finkel et al., 2010). As such, it is perhaps unsurprising that cell size has long been identified as a controlling factor in a number of essential phytoplankton life processes and can have substantial impacts on the rest of the marine ecosystem (Banse, 1976; Chisholm, 1992a; Finkel et al., 2010). From photosynthesis

(Finkel and Irwin, 2000) and nutrient physiology (Irwin et al., 2006; Weber and Deutsch, 2010), to grazing (Bergquist et al., 1985), formation of Harmful Algal Blooms (HABs) (Smayda and Reynolds, 2001a) and carbon export (Boyd and Newton, 1995), an allometric approach has been systematically applied as a parameter which can be used to understand the role of primary producers in ocean ecosystems and the wider Earth system. Enhancing this understanding has become increasingly important in the context of a changing world under anthropogenic climate change (Finkel et al., 2010).

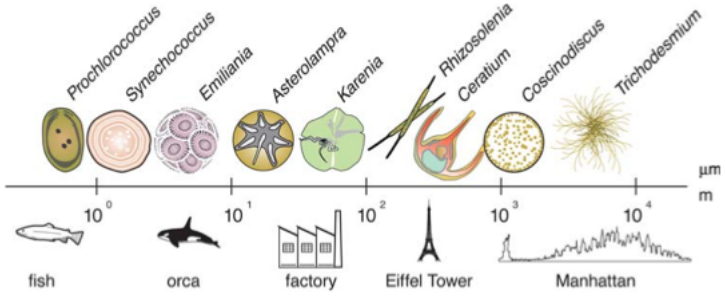

Figure 1.1: Scales of phytoplankton cell size versus other species and human constructions (using maximum linear dimension). Figure reproduced from (Finkel et al., 2010)

Phytoplankton community structure is intimately connected with the physical and chemical properties and dynamics of ocean regions (Chisholm, 1992a; Finkel et al., 2010; Munk and Riley, 1952; Parsons and Takahashi, 1973; Semina, 1972). Spatial and temporal variability has been observed in the particle size distribution attributing variability in community structures to different oceanic regions (Guidi et al., 2009). Despite the recognised role of large cells in highly productive phytoplankton blooms (Agustí et al., 1987; Irwin et al., 2006) and carbon cycling (Boyd and Newton, 1995), the small cell-dominated, oligotrophic regions represent around 50% of the global ocean and so these cells must also play a part in the global ocean carbon and other biogeochemical cycles (Richardson and Jackson, 2007). Ultimately the physical and chemical properties of different ocean regions influence the community size structure through control of the light environment and nutrient supply. It is therefore imperative to understand these interactions in order to predict how phytoplankton community structure and consequently the global carbon cycle could change with fundamental changes to ocean physics and chemistry in the future.

There are strong differences in phytoplankton community dynamics between regions, driven by the associated variations in ocean circulation, nutrient availability and top down control by the grazing community and higher trophic levels (Finkel et al., 2010). The midlatitudes are characterised by their spring phytoplankton blooms (Chiswell, 2011; Henson et al., 2009; Henson and Thomas, 2007; Siegel et al., 2002; Wasmund et al., 1998). These blooms are driven primarily by seasonal insolation changes which increase stratification and consequently light availability, giving phytoplankton the ability to use the abundant nutrients available after the winter, wind driven mixing (Chiswell, 2011; Siegel et al., 2002). The rapid growth associated with these spring blooms provides excess food for the zooplankton grazer community which

flourishes and curtails the growth of the phytoplankton community. Early in the spring bloom when there is still turbulent mixing and relatively low light, large celled diatoms out typically compete other species and dominate the phytoplankton assemblage. As the bloom progresses, nutrients, in particular silica which diatoms require for growth, are depleted and diatoms are out competed by smaller, less nutrient demanding phytoplankton (Allen et al., 2005). Succession of phytoplankton species in relation to turbulence and nutrient supply was described by Margalef (Margalef, 1978; Smayda and Reynolds, 2001a) (Fig. 1.2). Other forms of species succession have been observed, leading to distinct assemblages outside of the traditional succession described by Margalef's mandala. Harmful Algal Blooms (HABs) or "red tides" in coastal upwelling systems, are one example. These occur frequently in highly productive upwelling regions during periods of wind relaxation when previously upwelled, nutrient rich water becomes stratified, allowing succession and dominance of HAB species (Pitcher and Calder, 2000; Smayda and Reynolds, 2001a). With regards to HABs, size is one characteristic that can play a key role in determining under which ecological niches particular species can proliferate (Smayda and Reynolds, 2001a; Smayda, 2008). Size can also play a role in the ultimate impacts of HABs, where large celled phytoplankton typically dominate the high biomass blooms that lead to anoxia and size can cause mismatches with prey at higher trophic levels and physical impacts such as gill clogging (Smayda, 1997b).

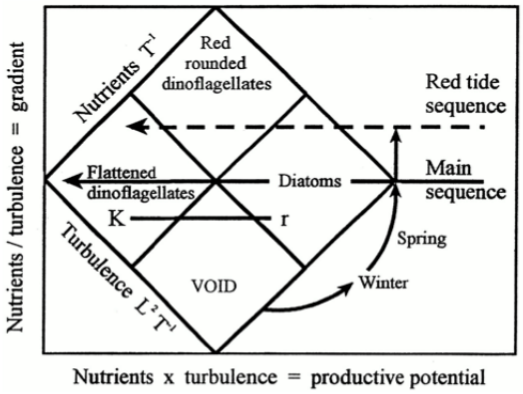

Figure 1.2: Margalef's Mandala simply describes the links between turbulence, nutrient supply and main phytoplankton assemblage species. From Margalef et al. (1979), reproduced from Smayda and Reynolds (2001a)

These physical, chemical and biological interactions are captured to varying degrees within a range of ocean models. The use of phytoplankton functional types (PFTs), has been heavily debated within the biogeochemical modelling community and whether or not the extra complexity results in better calculation of bulk biogeochemical variables is still uncertain (Anderson, 2005). A number of studies have highlighted where physical-biological relationships deviate from the typical size based relationships typically described by power laws (Finkel and Irwin (2000) and Finkel et al. (2010) and references within).

Observing and understanding the causes of variability in phytoplankton community structure benefits from suitably temporally and spatially resolved data. *In situ* methods can only

be applied where and when ship based measurements are made or within the scope of float and autonomous vehicle deployments, and whilst local studies are valuable for observing highly detailed variability within individual ecosystems, global methods of observation and quantification are required to fully constrain the ocean carbon budget for use in considerations of global change. This has led to the continually advancing use of remotely sensed data from satellites, to better address questions about phytoplankton community structure, primary productivity and global carbon export. However, a thorough understanding of the methods used and the ecosystem processes involved is needed to close the gaps between *in situ*, remotely sensed and modelled data. Recent work has shown that some information on phytoplankton community structure can be accessed using remote sensing techniques centred around ocean colour.

1.2 Answering global questions: The role of optics

Bio-optical measurements can provide quantitative and qualitative data about variability in phytoplankton biomass, community structure and physiology. Instrumentation and techniques exist to observe this variability on a range of temporal and spatial scales. Routine and synoptic measurements from satellites present the opportunity to capture variability which *in situ* studies can miss due to their limited temporal and spatial sampling scales (Joint and Groom, 2000). The main approaches to deriving information on phytoplankton dynamics have been based on the use of surface chlorophyll *a* concentrations ([Chl *a*]) estimated from satellite derived water-leaving radiance or equivalents such as remote sensing reflectance. A variety of algorithms have been developed for the suite of satellite ocean colour sensors in use over the last four decades (McClain, 2009). Many of these have used similar ratios of blue:green normalised water-leaving radiance to determine [Chl *a*] (Alvain et al., 2008; O'Reilly et al., 1998). In "Case 1" waters, where [Chl *a*] represents a reasonable biomass proxy that controls the variability in optical signals to first order (Morel and Prieur, 1977), these empirical band ratio approaches will typically perform well (Morel, 2009). However, estimating [Chl *a*] has historically been more difficult in optically complex, "Case 2" waters, where [Chl *a*] may not be the dominant cause of variability (Morel and Prieur, 1977) and alternative band ratios and semi-analytical approaches have improved performance (Doerffer et al., 2002; Ruddick et al., 2001). Even in Case 1 waters, knowledge of second order variability, attributed to other characteristics of phytoplankton (e.g. cell size or accessory pigments) and other optically significant constituents has the potential to improve derivation of [Chl *a*] and provide new information which can be exploited to learn more about phytoplankton communities (Brown et al., 2008; Sauer et al., 2012). Semi-analytical approaches have sought to address these concerns, by incorporating an understanding of the optical characteristics of various ocean water constituents (including diverse phytoplankton populations) and how they ultimately influence the signals measured by ocean colour satellites.

1.2.1 Phytoplankton from space: empirical and semi-analytical approaches.

Although derivation of chlorophyll *a* concentration has often been the focus and validation point for satellite ocean colour products, additional information about phytoplankton community characteristics can be gained from other remotely sensed bio-optical variables. Indeed for many applications, [Chl *a*] maybe ecologically inappropriate as an indicator. Toxic effects associated with Harmful Algal Blooms (HABs) for example, may manifest at biomass levels which are indistinguishable from non-harmful blooms without ancillary information on species or functional type (IOCCG, 2009). In addition to this, regions with highly variable optical properties, such as those in the coastal zone, may require more sophisticated analytical techniques to deal with atmospheric corrections, adjacency effects, and to produce products within sufficiently constrained confidence limits (IOCCG, 2000, 2010; Moore et al., 1999; Moses et al., 2012; Schiller and Doerffer, 2005).

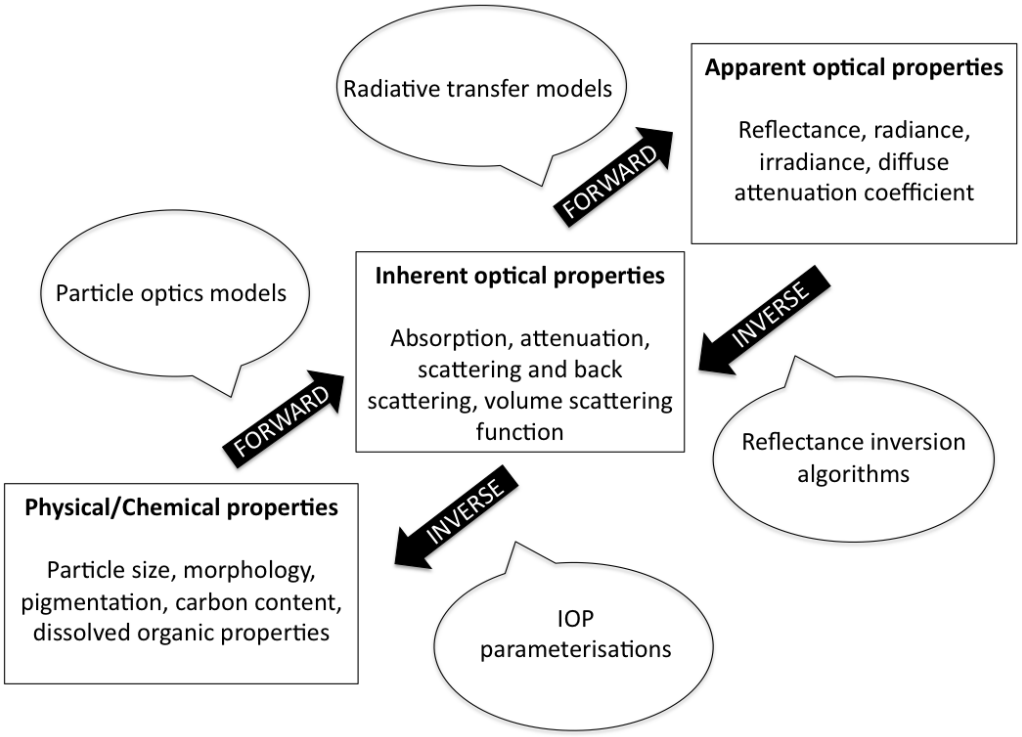

Figure 1.3: Forward and inverse approaches in ocean bio-optics. It should be noted that radiance and irradiance are not classically defined as Apparent Optical Properties (AOPs) due to their high frequency variability, but are commonly measured and used to derive reflectance and/or as a starting point for inversion approaches.

Figure 1.3 describes the forward and inverse methods used in ocean bio-optics to establish quantitative frameworks to evaluate and apply the relationships between inherent optical properties (IOPs), apparent optical properties (AOPs) and biogeochemical parameters. Inherent Optical Properties (IOPs) are those which are dependent on the concentrations and characteristics of the constituents within ocean waters, and are independent of the incident light field. These properties include absorption and scattering which can then be further deconvolved into

the various contributions attributable to different constituents (i.e. pure water, phytoplankton, dissolved matter, detritus, non-algal particulates etc). Apparent Optical Properties (AOPs) are influenced by both the IOPs of the constituents present and the incident light field. These properties are typically derived from those measured by radiometers (i.e. upwelling radiance (L_u) and downwelling irradiance (E_d) and the derived remote sensing reflectance (R_{rs})).

To understand signals from ocean colour sensors, forward methods such as particle population modelling can be used to predict the inherent optical properties which could be expected from theoretical phytoplankton assemblages. AOPs associated with these IOPs can then be calculated using radiative transfer models such as reflectance approximation techniques (Zaneveld, 1995) or more advanced models such as Hydrolight (Mobley, 2011). In turn, this can inform the development of inversion algorithms to retrieve information about phytoplankton communities and other biogeochemical parameters from radiometric measurements (i.e satellite or *in situ* ocean colour).

From figure 1.3 the importance of IOPs as principle carriers of constituent signal can be seen. Therefore an understanding of IOP variability and suitable parameterisations of this variability, are of considerable importance when forming semi-analytical ocean colour inversion algorithms. The derivation of IOPs has allowed for the development of many approaches investigating phytoplankton characteristics beyond [Chl *a*] based biomass estimates. Phytoplantkon absorption (a_ϕ) is related to the concentration of [Chl *a*] present with a non-linear dependence on secondary assemblage characteristics - in particular, the variable presence of accessory pigments influences a_ϕ (Dierssen et al., 2006; Kyewalyanga et al., 2002; Sathyendranath et al., 2001, 2004). In addition, a_ϕ is also intimately related to cell size as a result of the effect of the cellular packaging of pigments (Ciotti et al., 2002; Morel, 1991, 2009; Morel and Bricaud, 1981, 1986). The spectral shape of a_ϕ has also been related to dominant cell size and therefore incorporates more information about the ecological and physiological processes at work within the phytoplankton community (Ciotti et al., 2002; Green et al., 2003a; Mouw and Yoder, 2005). Variability in backscattering coefficients have also been linked to phytoplankton characteristics. Kostadinov et al. (2009) showed that the slope of the oceanic particle size distribution was related to the particle backscattering coefficient. Recent modelling studies have suggested that variability relating to phytoplankton characteristics such as size and intracellular structure may be present in phytoplankton backscattering coefficients (Bernard et al., 2009; Matthews and Bernard, 2013; Robertson Lain et al., 2014) and thus may represent both an important contribution to AOP variability and one that can be exploited through semi-analytical ocean colour inversion algorithms.

Despite these encouraging observations, some doubt in the accuracy of certain IOPs determined by various analytical inversion algorithms has been revealed (Wang et al., 2005), calling for greater validation and theoretical understanding of the derivation of such products. Relationships observed between AOPs, IOPs and phytoplankton characteristics have formed the basis of many methods seeking to derive information about phytoplankton communities from ocean

colour. These are discussed further in the following section.

1.2.2 Phytoplankton functional types and cell size from space: bio-optical approaches and assimilation of environmental data.

A number of methods have been developed to assess the various aspects of phytoplankton community characteristics on a global scale from satellite ocean colour data. Brewin et al. (2011b) divided these methods into a number of broad categories. These are summarised below with examples covering both the common application in the global open ocean, and more specific examples in the context of Harmful Algal Blooms.

1. Spectral-response based approaches.

Spectral-response methods depend upon the detection of specific optical signatures, representative of specific phytoplankton characteristics or functional types (PFTs) in water leaving radiance/remote sensing reflectance or derived IOPs. Empirical approaches are based upon correlating features in spectral reflectance/radiance or IOPs with concentrations of particular species or functional type groups. The PHYSAT method developed by Alvain et al. (2005, 2008) correlates spectral characteristics of [Chl *a*] normalised water leaving radiance (L_w) with pigment data derived from High Pressure Liquid Chromatography (HPLC) to determine whether pixels are dominated by five broad types of phytoplankton including *Phaeocystis*, coccolithophores, *Prochlorococcus*, cyanobacteria and diatoms. Theoretical explanation for these empirical relationships was suggested in (Alvain et al., 2012). A number of spectral response approaches have been focussed on the development of algorithms which take advantage of the functional characteristics of individual species. For example, Holligan et al. (1983) identified coccolithophores through the high reflectance values caused by their highly reflective calcite coccoliths. Similar approaches have been developed for specific cyanobacteria (e.g. *Trichodesmium spp.* (Subramaniam et al., 1999) and *Microcystis aeruginosa* (Matthews et al., 2012)), which can show high backscattering (and thus elevated reflectance) due to vacuoles; and for toxic HAB species such as *Karenia brevis*, which shows distinctive high fluorescence and low backscattering characteristics (Amin et al., 2009).

A variety of semi-analytical approaches have also been used to derive IOPs which have subsequently been used to infer the presence of species, functional types or phytoplankton characteristics such as cell size. These approaches can allow for explanation and further understanding of features which can be observed in water leaving radiance such as in the case of *Karenia brevis* and its particular fluorescence and backscattering characteristics (Cannizzaro et al., 2008; Tomlinson et al., 2009). Sathyendranath et al. (2004) used the differential absorption related signatures of waters dominated by diatoms and mixed assemblages to discriminate between them. This study highlighted that this approach could be ambiguous as changes in the absorption coefficient could result from either changes in cell size or changes in the intracellular pigment concentrations associated with the dominant species (*ibid*). Ciotti *et al.,* also used the

influence of community structure on phytoplankton absorption in several studies (Ciotti and Bricaud, 2006; Ciotti et al., 1999, 2002). In particular, Ciotti et al. (2002) found that 80% of the variability in the spectral shape of the phytoplankton absorption coefficient could be explained by changes in cell size. This finding was exploited to form a spectral mixing model where cell size was parameterised through the effect on absorption resulting from small (< 2 μm) and large (> 20 μm) celled assemblages (Ciotti and Bricaud, 2006). Whilst acknowledging that the cell size of a given phytoplankton community is quite intimately linked with trophic status, this approach avoided explicit relationships with [Chl a] which are more typically associated with abundance approaches (see following section). Other approaches have used the impact of the particle size distribution on backscattering to elicit information about phytoplankton communities. Kostadinov et al. (2009, 2010) used a power law particle size distribution, combined with Mie modelling to derive size estimates from remotely sensed particle backscattering (b_{bp}). These techniques are well founded in particle modelling theory, however a number of studies (including (Kostadinov et al., 2009)) have found that some assumptions associated with using power law distributions and Mie theory to model phytoplankton (and their subsequent interaction with the light field) as spherical and homogenous particles can be inappropriate in high biomass blooms and turbid waters (Bernard et al., 2009, 2007; Chami et al., 2006; Quirantes and Bernard, 2004). Alternative models of phytoplankton as layered spheres and non-spherical particles with various distributions exist (Bernard et al., 2009; Clavano et al., 2007; Kitchen and Zaneveld, 1992; Matthews and Bernard, 2013; Quirantes and Bernard, 2004; Robertson Lain et al., 2014), but as yet have not been used in broad application with semi-analytical reflectance inversion algorithms.

These IOP based approaches (Ciotti and Bricaud, 2006; Kostadinov et al., 2009, 2010; Sathyendranath et al., 2004) often rely on semi-analytical algorithms for derivation of IOPs before the application of empirical or theoretical relationships to derive information on size or community structure. To an extent these approaches can therefore dependent on the suitability of the specific IOPs underlying these algorithms as representative of the variability they are seeking to exploit. Whilst some algorithms allow for variability in underlying IOPs (e.g. by including variable spectral phytoplankton absorption), no current methods use coupled IOPs (i.e. absorption and scattering/backscattering) explicitly parameterised by influences from phytoplankton characteristics e.g. size or accessory pigments. Further concerns about spectral response methods relate to the inability to detect small differences in signals (Morel, 1997), particularly when there exists a lot of noise and when groups or species have similar optical signals (Defoin-Platel and Chami, 2007; Dierssen et al., 2006; IOCCG, 2000; Nair et al., 2008). In particular, signals may be dependent on biomass concentrations (Amin et al., 2009; Subramaniam et al., 1999). Additional problems may be encountered when there is unaccounted for variability of optical characteristics within species/functional type e.g. as a result of physiology or life cycle state (Nair et al., 2008).

2. Abundance based approaches.

Abundance methods use observed links between biomass and phytoplankton community structure. These methods assume that biomass (usually represented by [Chl a] or a_ϕ) covaries with the dominance of specific species, size classes or PFTs. These assumptions are often based on empirical relationships and the allometric scaling laws associated with size which form the basis of many relationships describing biological processes (Agustí et al., 1987; Finkel et al., 2010; Finkel and Irwin, 2000; Marañón, 2008).

In additional to using spectral approaches, the previously mentioned HAB species *Karenia brevis* can also be detected in the Gulf of Mexico through methods using [Chl a] anomalies as a result of the distinctively high biomass range it is typically associated with (Stumpf et al., 2003).

Uitz et al. (2006) used the ratio of euphotic to mixed layer depth to determine stratified and mixed environments. These were then associated with [Chl a] intervals, phytoplankton size classes and specific marker pigments (Fig. 1.4).

Diagnostic Pigments	Abbreviations	Taxonomic Significance	Phytoplankton Size Class	$\langle Chla \rangle_{Zeu}$: $\langle P \rangle_{Zeu}$	Significance Level
Fucoxanthin	Fuco	diatoms	microplankton	1.41 ± 0.02	$p < 0.001$
Peridinin	Perid	dinoflagellates	microplankton	1.41 ± 0.10	$p < 0.001$
19′-hexanoyloxyfucoxanthin	Hex-fuco	chromophytes and nanoflagellates	nanoplankton	1.27 ± 0.02	$p < 0.001$
19′-butanoyloxyfucoxanthin	But-fuco	chromophytes and nanoflagellates	nanoplankton	0.35 ± 0.15	$p = 0.02$
Alloxanthin	Allo	cryptophytes	nanoplankton	0.60 ± 0.16	$p < 0.001$
chlorophyll *b* + divinyl chlorophyll *b*	TChl*b*	green flagellates and prochlorophytes	picoplankton	1.01 ± 0.10	$p < 0.001$
Zeaxanthin	Zea	cyanobacteria and prochlorophytes	picoplankton	0.86 ± 0.09	$p < 0.001$

Figure 1.4: Bio-marker pigments used by Uitz *et al.*, their taxonomic significance and relationship to chlorophyll *a*:pigment ratio at the euphotic zone depth. Reproduced from Uitz et al. (2006).

Similarly, Hirata et al. (2008) derived an abundance based approach using co-variance observed between phytoplankton absorption at 443 nm ($a_\phi(443)$) and the spectral shape between 443 nm and 510 nm. Figure 1.5 shows this relationship for the 3 phytoplankton size classes used in the study. Validation was conducted using NOMAD (NASA bio-optical Marine Algorithm Dataset, (Werdell and Bailey, 2005)) and similar bio-marker pigments to Uitz et al. (2006).

Abundance based approaches often perform well in the context of isolated [Chl a]/a_ϕ ranges for individual species or functional types, however in broader ranges e.g. where multiple PFT groups can cause similar optical variability, it may be difficult to distinguish between groups (Brewin et al., 2011b). For example, in the case of biomass based abundance methods, if multiple PFT groups occur at the same biomass, it would be difficult to distinguish them (*ibid*). Also, as with spectral response approaches, any within-group variability in parameters used (typically [Chl a] or a_ϕ) can result in erroneous retrievals (Brewin et al., 2011b; Dierssen et al., 2006). Similarly, abundance approaches can also be limited by the accuracy of the derived parameters needed for their use. For example, inaccuracies in derived [Chl a] or a_ϕ could result in variability being misattributed to a PFT or size class. The accuracy of the products used (i.e. [Chl a] or a_ϕ) is dependent on the empirical or analytical approaches used to derive them.

3. Ecological-based approaches.

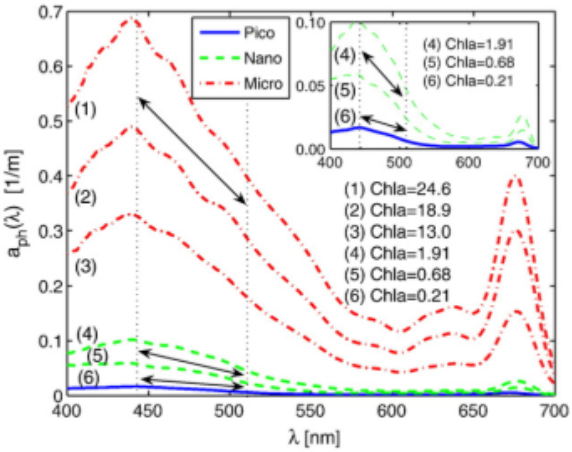

Figure 1.5: Relationship between $a_\phi(443)$ and the spectral slope of the absorption coefficient between 443 nm and 510 nm for different phytoplankton size classes. Reproduced from Hirata et al. (2008).

Ecological approaches incorporate ancillary data to identify and distinguish ecological niches where PFTs may occur. These methods are similar to the abundance approaches but frequently use various types of spatio-temporal and physical data in additional to bio-optical data. This style of approach offers a more holistic understanding of the multiple factors which have been observed to govern phytoplankton community dynamics. For example, see early understanding on the connections between cell size and nutrients in Yentsch and Phinney (1989). Raitsos et al. (2008)) used an ecological approach to distinguish dominance of PFT's in the North Atlantic. This was achieved using probabilistic neural networks to match *in situ* Continuous Plankton Recorder (CPR) data with SeaWiFS [Chl *a*], normalised water-leaving radiance (nLw) and photosynthetically active radiation (PAR), Advanced Very High Resolution Radiometer (AVHRR) sea surface temperature (SST) data and NASA QuikSCAT and European remote sensing satellites (ERS-2) wind stress data. The relative impact of these factors in the dominance of four functional types were investigated; a comparison for diatoms and dinoflagellates is shown in figure 1.6.

Ecological approaches potentially provide many ways of identifying niches where different PFTs may proliferate. This may make distinguishing individual groups easier than in abundance approaches which typically only use one variable. However, these approaches do require substantial amounts of coincident data, which may not always be available. Equally, evaluating these data can involve detailed statistical approaches, which can make interpretation difficult (Brewin et al., 2011b).

1.2.3 Algorithm validation and comparing different approaches for investigation of phytoplankton community structure.

An assessment of algorithm performance and estimates of errors and uncertainty are necessary to effectively use any algorithm and associated products. A number of efforts have been made to validate ocean colour data and products, particularly [Chl *a*] (Bernard et al., 1998; Doerffer and

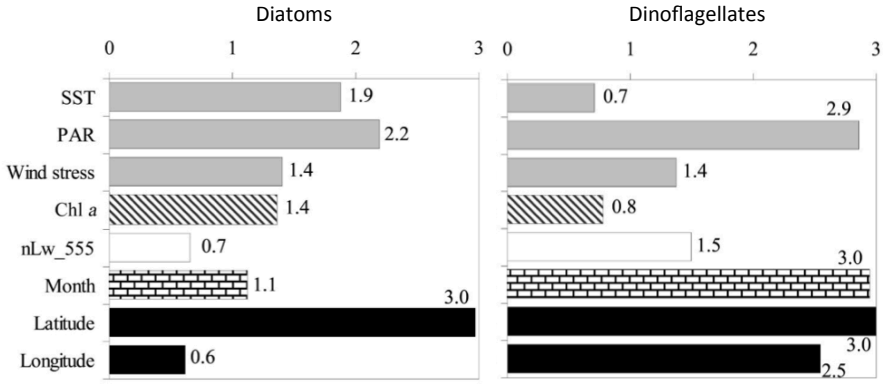

Figure 1.6: The relative impact (smoothing factor) of variables determining dominance of diatoms vs dinoflagellates. From top to bottom, the different shadings refer to physical, optical, temporal and spatial variables respectively. This figure is a simplified version of figure 3 in Raitsos et al. (2008), which covers additional groups.

Schiller, 2007; Mélin et al., 2013; Mueller et al., 2003a; Pinkerton and Aiken, 1999; Smith et al., 2013; Werdell and Bailey, 2005; Zibordi et al., 2009b, 2006a). These approaches have provided substantial insight on best practice for collecting suitable *in situ* data for validation purposes. Currently, this level of validation has not yet been applied to the evaluation of algorithms deriving information on phytoplankton community structure. Assessing and comparing performance of the range of algorithms currently available for determining phytoplankton community structure is made difficult by a number of factors. Firstly, there is a paucity of appropriate and diverse *in situ* measurements needed to characterise variability in phytoplankton characteristics and PFTs and to ultimately validate algorithms. Whilst many *in situ* measurements methods have been developed, all present advantages and disadvantages which can make them more or less suitable for algorithm development and validation (Sosik et al., 2014). Secondly, not all algorithms define phytoplankton community structure or PFTs in the same way and the data required to develop and validate them will vary accordingly. Some algorithms deduce a dominant group (either species or size), whilst others use a variety of methods to determine different suites of PFTs. The choice of community metric used largely (and appropriately) depends on the application. Detecting a specific species rather than cell size or vice versa may not accurately reflect potential impacts on the biological processes of interest. For example, in the study by Boyd and Newton (1995), both of their study sites were dominated by diatoms, however the size of the diatoms present was suggested as the factor determining observed differences in export. In contrast, the ability to detect particular species may be more useful when detecting toxicity in Harmful Algal Blooms (Bernard et al., 2006) and help inform model representations of other key cycles that specific species are involved in (i.e. through silicon or nitrogen cycling). From an optical perspective, cell size can have considerable impact on absorption and scattering, and

thus represents a direct link between phytoplankton characteristics and optical variability (Morel, 2009; Svensen et al., 2007). Similarly the effect of particular pigments (besides chlorophyll *a*) (Subramaniam et al., 1999), calcite structures (e.g. associated with coccolithophores (Moore et al., 2012)) and intracellular structures (Matthews and Bernard, 2013; Matthews et al., 2012) may impact IOPs, providing a causal link with reflectance that can be parameterised and used in algorithms. However, not all species may have sufficiently distinct optical characteristics to be distinguished from each other (Dierssen et al., 2006).

Brewin et al. (2011b) yielded no conclusive result with regards which algorithm type reproduces *in situ* observations most accurately. To make comparisons possible, results from each study were converted to a dominant phytoplankton type and several methods were used to score and compare results from 9 different algorithms from the 3 defined types. Sources of variability in the results from the different algorithms were suggested as including the size of the phytoplankton (i.e. some models better predicted certain size classes/functional types over others), input satellite data and availability of appropriate *in situ* validation data.

Whilst the diversity of approaches can make acquisition of comprehensive *in situ* validation data challenging, there exists some significant data sources and proxies which can be used to develop and assess the performance of the algorithms described above (Sosik et al., 2014). Throughout the literature, high pressure liquid chromatography (HPLC) measurements are commonly used as a frequent source of validation data. Despite the limited spatio-temporal coverage of many *in situ* validation data sources, several large, comprehensive optical and HPLC databases exist including NOMAD (National Aeronautics and Space Administration (NASA) bio-optical Marine Algorithm Dataset) (Werdell and Bailey, 2005), the NASA SeaWiFS Bio-optical Archive and Storage System (SeaBASS) (Hooker et al., 1994) and the Geochemistry, Phytoplankton and Color of the Ocean (GeP&Co) pigment database (http://www.lodyc.jussieu.fr/gep- co). The growing global coverage means these data represent an attractive source to validate algorithm results in different ocean regions. Dominant size classes or PFTs can be classified using dominant pigment analysis (DPA) to assess which pigment represents $> 45\%$ of the chlorophyll biomass (Hirata et al., 2008); this was used as a key validation source by Brewin et al. (2011b). However, it should be noted that the use of marker pigments can be ambiguous (i.e. pigments can occur similarly across broad functional types or size ranges) and may require careful interpretation appropriate to the system of interest and the potential for exceptions to classification schemes (Havskum et al., 2004; Schlüter and Møhlenberg, 2003; Sosik et al., 2014). Another commonly used data source is the Continuous Plankton Recorder (CPR). Towed behind ships across the North Atlantic, the CPR filters planktonic species, primarily microplankton (>10 μm) for subsequent species identification and counts. As a source of validation data for size/PFT algorithms, the CPR has been shown to reliably reflect changes in phytoplankton abundance (Batten et al., 2003). Though limited by its inability to detect small plankton, methods have been developed to statistically determine the remaining fraction of the plankton community from abundance of microplankton species (Brewin et al., 2011b; Raitsos et al., 2008). If de-

tecting cell size as opposed to pigment related characteristics or specific taxonomy, the most suitable data for algorithm validation may come from instruments which derive a particle size distribution. Commonly used instruments for this purpose include Coulter Counters and Laser In Situ Scattering and Transmissometry (LISST) which come with a number of advantages and disadvantages depending on assumptions about how size relates to the instrument measurement method (Reynolds et al., 2010). Further consideration of *in situ* validation data, with a focus on particle size distributions can be found in chapter four and a comprehensive review of a variety of optical methods for investigating PFTs can be found in Sosik et al. (2014).

The diversity of methods with which to determine phytoplankton community characteristics, both from *in situ* and satellite measurements, hints at the substantial regional diversity in phytoplankton communities. Open ocean phytoplankton community characteristics are known to be very different to those in coastal regions (Chisholm, 1992a; Margalef, 1978; Raven, 2009), and whilst a greater understanding of the underlying causes of optical variability should provide insight across trophic regimes, it would seem unlikely that there will be a single method which can perform adequately throughout the variety of water types which make up the global ocean. It should also be noted that although the satellite ocean colour record has provided an unprecedented view of global phytoplankton dynamics, it is still limited to a surface picture, which, in isolation, may not reveal key information about phytoplankton community responses to environmental variability. As such, the maximum value of ocean colour data is likely to be found through a combination of well characterised products derived from ocean colour ([Chl *a*] and some indication of community structure) with ancillary data and hydrodynamic and biogeochemical ecosystem models.

1.3 The southern Benguela: Challenges and opportunities.

The southern Benguela is a highly productive upwelling system situated off the west coast of South Africa. High phytoplankton biomass in the region has fuelled large fisheries and aquaculture industries (Pitcher and Calder, 2000). However, these industries are often at risk from Harmful Algal Bloom events, which can introduce toxins and/or cause anoxia upon their decay (*ibid*). To gain an understanding of how these blooms form, are advected alongshore and subsequently decay firstly requires highly spatially and temporally resolved data on biomass and ideally some level of functional or taxonomic information with appropriate understanding of likely errors (Bernard et al., 2006). These data can then be used to develop predictive capacity either through the use of statistical models (indicating spatial and temporal likelihood) or hydrodynamic and ecosystem models that appropriately account for influential forcing factors (e.g. wind, nutrients or higher trophic level dynamics) (*ibid*). To achieve these goals, there are a number of specific requirements:

- As a first step towards prediction of potential low oxygen conditions estimates of biomass must be made which cover the full range of biomass (often with [Chl *a*] in excess of 100 mg

m^{-3}) associated with Harmful Algal Blooms. Standard ocean colour [Chl *a*] algorithms frequently do not cover this range.

- Products should ideally provide indications of cell size and/or species/functional type to track succession and bloom development. This should ultimately allow for a level of succession to be observed, potentially modelled and connections drawn between species and environmental forcing to create statistical envelope or fuzzy logic type models indicating risk of harm from various sources (e.g. anoxia or toxicity) (Bernard et al., 2006).

- An understanding of sensitivity, errors, uncertainties and any constraints in terms of application (e.g. limits of detection) is needed to make products suitable for modelling and management purposes.

Whilst there is clearly a need for the types of information that ocean colour data can provide in the southern Benguela, the system also offers a number of opportunities to further an understanding of the relationships between phytoplankton and optical signals themselves. High biomass blooms are by nature greatly reduced in diversity and often dominated by one or a few species (Raven, 2009) and as such the high biomass blooms of the southern Benguela represent as close to a cultured sample as one may find in the ocean. A suite of biogeochemical, bio-optical and radiometric data has been collected from diverse bloom types, where the large majority of the signal is phytoplankton associated, over the last ten years. This allows for investigation of traits such as phytoplankton cell size or functional types in the context of a natural light environment, with both *in situ* and satellite derived ocean colour measurements available.

1.4 Conclusions and thesis content

From the literature reviewed here it is apparent that there is substantial potential to use ocean colour data to provide information about phytoplankton communities that is sorely needed for further research and management in systems like the southern Benguela. Several key points arise which directed the work which follows in this thesis:

- Use of a forward model, where phytoplankton characteristics such as cell size are captured from the IOP level through particle modelling, represents the most advanced way of parameterising the impacts of cell size on optics.

- A coupled forward and inverse modelling framework allows for a traceable understanding of how size based influence on IOPs is propagated to reflectance and how easily this can be retrieved using inversion methods.

- Application in the southern Benguela should enable a further understanding of how errors inherent in the inverse problem are added to by errors introduced in real world application i.e. due to inadequacies in the forward model and as a result of *in situ* instrumentation.

- To enhance understanding of Harmful Algal Blooms in the southern Benguela, a multi-year time series of highly temporally and spatially resolved data are needed. The satellite ocean colour archive can provide this, on the condition that the sensor in question has suitable specifications for application of semi-analytical inversion methods appropriate to a coastal and high biomass optical regime.

- This time series should then be related to ancillary environmental data to extend any conclusions made using simplified parameterisations (i.e. the detection of cell size towards an understanding of species succession) and relate them to the potential physical, chemical and biological forcing factors which may determine Harmful Algal Bloom development and impacts.

This thesis covers the development and application of a forward model and semi-analytical inversion algorithm that use coupled, size and functional type dependent IOPs derived from a phytoplankton particle population model. These are henceforth referred to as the Equivalent Algal Population (EAP) forward model and inversion algorithm. The EAP inversion algorithm is then validated with an *in situ* dataset and used to process MERIS satellite imagery from the last ten years in the southern Benguela. Examples are then shown to illustrate how the algorithm and new products can be used to enhance the investigation of interannual variability of Harmful Algal Blooms in the southern Benguela system.

Chapter two presents a sensitivity analysis of the Equivalent Algal Population (EAP) forward model. A simulated data set, representative of a broad range of optical water types, is generated and used to assess relationships between reflectance and cell size through the use of size (parameterised as an effective diameter) and functional type dependent IOPs. Assumptions with regards to the particular size distribution, intracellular [Chl *a*] concentration, and the presence of other optically significant constituents are evaluated. This simulated dataset is used in chapter three to investigate the minimum errors likely when the forward model is applied as an inversion algorithm using non-linear optimisation. Different non-linear optimisation techniques are tested and a number of assumptions assessed including the selection of initial conditions and the sensitivity of the inversion method to the number of unknown parameters to be solved. In chapter four, the EAP inversion algorithm is applied to an *in situ* dataset from the southern Benguela. The likely errors associated with the *in situ* data are assessed and together with the analysis from chapters two and three, provide a summary of the signal related to phytoplankton cell size which can be exploited through remote sensing products. Chapter five discusses the application of the unconstrained inversion approach to the MERIS satellite ocean colour archive for the southern Benguela. A time series of chlorophyll *a* and effective diameter is presented and a number of cases studies from key blooms are discussed. These time series are compared to those derived from other algorithms and used to assess the inter-annual variability of Harmful Algal Blooms in St Helena Bay within the southern Benguela. Metrics are developed to quantify bloom occurrence and persistence and results are related to wind forcing and current

understanding of physical dynamics of St Helena Bay. Chapter six summarises the major conclusions of this thesis and advises on further work which would enhance this new insight and address remaining questions.

Chapter 2

Sensitivity of an inherent optical property model for use in a semi-analytical inversion algorithm

Parts of this chapter formed content included in the paper: Evers-King et al. (2014). The forward modelling techniques within this paper were conducted by the lead author and have been expanded for inclusion in this thesis chapter. The original two-layered sphere model is described in Bernard et al. (2009) and the output was expanded for use here by the thesis author, with assistance from Lisl Robertson Lain. This work also formed part of a comparison to other Inherent Optical Property models, detailed in Robertson Lain et al. (2014).

2.1 Introduction

With the availability of remotely sensed optical data from satellites, and the development of numerous, but often regionally specific, empirically derived geophysical relationships, the need has arisen for a greater understanding and development of underlying theory linking optical signals to ocean and aquatic biogeochemistry. Studies in this pursuit frequently revolve around forward and inverse reflectance modelling. These modelling processes aim to link inherent and apparent optical properties of the ocean surface with biogeochemical properties. Inherent optical properties (IOPs) are measurable parameters, characteristic of the present biogeochemical properties, which are independent of the incident light field. Apparent optical properties (AOPs), whilst also being measurable and characteristic of the underlying biogeochemistry, are influenced by the incident light field and provide regular, stable features suitable for describing a water body (Mobley, 1994). Linking IOPs to AOPs with a known incident light field is termed the "forward" or "direct" problem. Accordingly, linking AOPs to IOPs and then with biogeochemical parameters, is known as the "inverse" problem (see chapter one, figure 1.3). Radiance (the total energy of photons in a given direction per time, solid angle, wavelength and area) and irradiance (the total energy of photons as before but without the angular component) are commonly measured.

These parameters are not themselves classed as AOPs, as they can vary very rapidly in time i.e. in the case of irradiance when a cloud passes over (Mobley, 1994). Remote sensing reflectance (R_{rs}) is a commonly derived AOP, representing the upwelling water-leaving radiance (L_u) divided by the downwelling plane irradiance (E_d). Calculation of this varies depending on the radiometric sensor used, but in all instances, forms the basis for ocean colour observation and *in situ* validation. A further AOP is the diffuse attenuation coefficient (K_d). This parameter describes the attenuation of light with depth in the water column. Inversion of K_d provides a key parameter for understanding remotely sensed bio-optical data - the "optical depth" or the depth to which satellites can "see" in the ocean. This value will vary over the ocean depending on the presence and concentration of various constituents (i.e. phytoplankton, dissolved organic matter, inorganic particles such as sediments etc).

There are two fundamental inherent optical properties: Absorption (a) and the volume scattering function (β). Additional IOPs are derived from these, including the total scattering coefficient (b), backscattering coefficient (b_b) and the beam attenuation coefficient (c, where c=b+a). Incident photons can be scattered in two ways, either elastically or inelastically. Elastically scattered photons have the same energy (frequency) and wavelength as incident photons. In the case of inelastically scattered photons, their energy is not conserved and consequently a change in wavelength is observed. Inelastic scattering of incident light in the oceans occurs through two main processes. Firstly, water molecules cause "Raman" scattering, where a portion of the incident light is absorbed and re-emitted at a shifted wavelength. This results in substantial impacts on in the 500-700 nm wavelength region (Bartlett et al., 1998). Secondly, incident photons can be scattered through the process of fluorescence by gelbstoff (also known as coloured, dissolved, organic matter (CDOM)) and phytoplankton. Water molecules, phytoplankton, gelbstoff and suspended particulate matter are the major components of ocean water responsible for absorption and scattering. Minor influences on absorption and scattering result from bacteria, viruses and bubbles in the water column (Balch et al., 2007; Stramski et al., 2004; Stramski and Kiefer, 1991).

Models of inherent optical properties form the first step in creating a semi-analytical inversion algorithm. In the simplest IOP models, total absorption and backscattering are derived additively from a number of absorbing and backscattering components - typically: phytoplankton (often scaled using chlorophyll a concentration ([Chl a]) as a proxy for biomass), gelbstoff/coloured dissolved organic matter and non-algal (typically mineral) particulates. Numerous expansions of this approach have sought to expand this construct - particularly the use of [Chl a] to determine the influence of phytoplankton on the optical signal. Measured IOPs and AOPs compared to those derived from simple, [Chl a] based methodologies have been found to differ by up to hundreds of percent, particularly for waters where dominant sources of optical variability do not co-vary with phytoplankton concentration (i.e. waters with high sediment loads) (Stramski et al., 2001). As a greater understanding of optically significant components has emerged, IOP models have been expanded to include them.

To derive information on phytoplankton community structure from bio-optical data, it is necessary to understand the influence that different assemblage types will have on these inherent and apparent optical properties. The optical properties of a phytoplankton assemblage are dependent on several factors, primarily [Chl a] but also including cell size, other pigments and intracellular structures (Morel, 1991; Morel and Bricaud, 1986). Considerable impacts on light interaction with phytoplankton cells also result from what is known as the "package effect" (Morel and Bricaud, 1986). This effect results from the various absorbing components of cells being contained in particles of varying sizes rather than dispersed in solution. When the absorption by a phytoplankton assemblage is measured, the variability observed can be a result of the biomass present, pigment density, the size distribution and/or the presence of pigments with distinctive absorption characteristics (Dierssen et al., 2006; Kirk, 1975). The effect of size appears dominant amongst these variables (after [Chl a]) (Morel and Bricaud, 1986) and consequently parameterisations of the relationship between size and absorption have been developed. Recently a number of studies have sought to link optical signatures with phytoplankton cell sizes or functional types. These approaches can be considered as "spectral based" (Brewin et al., 2011b). A significant number of these methods exploit the relationship between absorption and phytoplankton cell size. Simply put, a community dominated by small cells will show a higher absorption efficiency whilst a large cell dominated community will have a comparatively lower absorption efficiency (Ciotti et al., 2002; Morel and Bricaud, 1986; Sathyendranath et al., 1987; Uitz et al., 2008). This relationship has been further investigated to derive both empirical and semi-analytical relationships between absorption and cell size.

To parameterise the interactions between phytoplankton and the light field from a causal perspective, models have been developed from a single particle level. These models are based on optical efficiency factors for absorption, scattering and backscattering. At the level of a single particle these optical efficiency factors are controlled by particle size, the relative refractive index and particle shape. Mie modelling has formed a core of this work and particle size is described optically by the Mie size parameter (Bricaud and Morel, 1983; Morel and Bricaud, 1986; Stramski et al., 2001). Mie theory comprehensively describes absorption, scattering and attenuation (Ahn et al., 1992; Morel and Bricaud, 1986; Stramski et al., 2001), however the limited numbers of measurements of angular scattering suggest it is less well simulated through this method (Quinby-Hunt et al., 1989; Vaillancourt and Brown, 2004; Volten et al., 1998). These studies suggest that internal structures and non sphericity were linked to this inadequacy. Phytoplankton, due to their complex internal structure, have heterogenous refractive indices within their cells. This heterogeneity is associated with intracellular components such as membrane bound chloroplasts which are highly absorbant, other membrane bound organelles (e.g. mitochondria and the nucleus) and other intracellular constructions such as vacuoles, which can significantly impact scattering characteristics (Matthews and Bernard, 2013; Svensen et al., 2007; Whitmire et al., 2010).

2.1.1 Uncertainty in backscattering measurements and modelling techniques

Backscattering of light within the ocean is of significant importance to optical remote sensing techniques. However, understanding the variability and sources of backscattering in ocean waters, has been, and to some extent remains, a significant challenge. Whilst a significant portion of backscattering from any ocean waters (particularly open ocean waters), comes from backscattering by pure water itself, in other regions, other constituents can impart substantial effects. Historically, it has been suggested that phytoplankton contribute less to total backscattering than non-living particles in open ocean waters and understanding the interplay between various particle pools remains the "greatest challenge" in understanding backscattering in the ocean (Stramski et al., 2004). Living particles in the ocean can include: phytoplankton, zooplankton, viruses, bacteria; whilst non-living particles can include organic detritus and inorganic particles such as minerals. Bubbles have also been suggested to influence backscattering (Stramski et al., 2004). In ocean waters, particles of both living and non-living types can have very different sizes, shapes and refractive indices which may be heterogeneous due to internal structures; all of which can influence scattering (see Ahn et al. (1992); Stramski et al. (2004); Stramski and Kiefer (1991); Svensen et al. (2007); Volten et al. (1998) and references within). To understand these influences, models have been employed, however the necessary reductionist approach has often yielded conflicting results about how backscattering from each of these highly variable components contributes to total backscattering. To develop a causal, combined approach to simulate absorption and backscattering, parameterisations have been developed using particle models. As mentioned, Mie theory has been used to simulate absorption, attenuation and total scattering properties of algal cultures. However, when compared to measurements, it seems backscattering is not well reproduced through this theory (Chami et al., 2006; Vaillancourt and Brown, 2004; Volten et al., 1998). To complicate matters further, great uncertainty surrounds backscattering measurements themselves. Conflicting results have led to a number of unresolved questions and without consensus measurements, modelling the influence of highly variable phytoplankton communities and other particle assemblages remains difficult. Fully resolving the backscattering signal in ocean waters requires a complete measurement of the volume scattering function at all angles and wavelengths. This is time consuming, expensive and physically difficult to conduct in a fully accurate way. Studies often make use of fixed angle scattering meters which may not accurately reflect the full variability of backscattering in already highly variable ocean regions. A study by Chami et al. (2006) evaluated the assumptions inherent with the use of fixed angle scattering meters and found inconsistencies in the use of the conversion factor (χ_p) needed to extrapolate and convert fixed angle scattering measurements to particulate backscattering (b_{bp}; including phytoplankton and non-living particles) measurement. Spectral measurements of the full volume scattering function (VSF) for both *in situ* conditions and laboratory cultures found that χ_p varied significantly at 555 nm (associated with peak reflectance and low phytoplankton absorption) at certain angles. Thus an assumption that fixed angles can be used to assume fully angular and spectral measurements may be flawed, particularly under bloom-like conditions

(Chami et al., 2006).

Relationships, in particular between the particulate backscattering coefficient (b_{bp}) and other parameters, (e.g. [Chl a]) have displayed substantial variability; the exact causes of which are difficult to establish (Antoine et al., 2011). Some studies have evidenced an ability to correlate scattering (beam attenuation and backscattering) at specific wavelengths with [Chl a] and thus derive empirical relationships (Antoine et al., 2011; Behrenfeld et al., 2005). However these relationships are often regionally specific and significant scatter exists around the general relationships observed. This variability insinuates that coincident variability in the nature of the particulate pool across ocean regions/over time must influence backscattering. Although not the only, and perhaps not the most significant contribution to naturally occurring particulate assemblages, differences in phytoplankton community composition have been very clearly linked to differences in the b_{bp}:[Chl a] relationships (Whitmire et al., 2010). Although significant scatter was observed in measured b_{bp}, strong relationships between Chl and b_{bp} were observed (*ibid*). A number of studies suggest no spectral variability in observations of the particulate backscattering coefficient (b_{bp}) (Huot et al., 2008; Whitmire et al., 2007). However, the combined effects of uncertainties in measurement and theoretical assumptions make these conclusions uncertain (Antoine et al., 2011). Additionally, studies in more extreme circumstances (e.g. turbid waters, high biomass, monospecific blooms, phytoplankton communities with particularly obscure intracellular structures/pigmentation) suggest significant spectral variability can be observed in these cases (Chami et al., 2006). Modelling studies seeking to incorporate this complexity find agreement with these studies (Bernard et al., 2009; Matthews and Bernard, 2013).

2.1.2 Representing size - Phytoplankton absorption and backscattering from a causal perspective

In addition to the great need for phytoplankton size/species level information from remote sensing methods (see chapter one for discussion), current thinking also suggests that size and phytoplankton morphology could have a large part to play in explaining uncertainties in both measurement techniques and optical models. It has been shown to be possible to correlate optical signatures with dominant single/multiple size classes or phytoplankton functional types (Alvain et al., 2008; Brewin et al., 2011a; Uitz et al., 2006). However, establishing single parameters which can adequately describe, in causative terms, the combined or average optical effects of a complex particle size distribution (PSD) has proven difficult. Studies have employed log-normal and gamma distributions to approximate the PSD of monospecific phytoplankton cultures (Bricaud and Morel, 1986; Risović, 1993) and further adapted for non-monospecific cases (Jonasz and Fournier, 1996; Risović, 1993). However, the Junge or power law distributions has been perhaps the most extensively used across both cultures and *in situ* studies (Andrews et al., 2011; Babin et al., 2003; Bricaud et al., 1995b; Chami and Defoin-Platel, 2007; Chang and Whitmire, 2009; Dall'Olmo et al., 2009; Guidi et al., 2009; Junge, 1963; Kostadinov et al., 2009; Park and Ruddick, 2005; Stramski et al., 2001). Atmospheric physicists have used equivalent size distri-

butions and the "effective diameter" of a particle size distribution to adequately recreate the optical properties of clouds and aerosols. The effective diameter represents the mean volume to surface area ratio of the particle size distribution (Hansen and Travis, 1974). This method has been shown viable for characterising particle distributions of different shapes in a variety of circumstances including over various phytoplankton size distributions (Bernard et al., 2009, 2007).

Measurements of absorption have been used extensively to investigate phytoplankton properties and numerous studies have observed relationships between the particle size distribution and absorption coefficients. Ciotti et al. (2002) found that more than 80% of the variability in spectral shape in phytoplankton absorption measurements could be explained by coincident variability in size. The size parameter used represented that of the dominant organism (within filtered fractions) and was suggested to account for a number of size related, optically significant effects, such as pigment packaging and the concentration of accessory pigments (*ibid*). Similarly Kostadinov et al. (2010) were able to use relationships derived between the spectral backscattering coefficient and the parameters of a power-law particle size distribution to evaluate the presence of phytoplankton functional types associated with different power-law PSD slopes.

A number of modelling approaches have sought to simulate these relationships to greater understand how various phytoplankton characteristics influence IOPs. Stramski et al. (2001) divided the algal component in their model into 18 different size and species based classes. Estimates of phase function were made using Mie theory, assuming sphericity and cellular heterogeneity. The extensive dataset used to construct this model consisted of laboratory cultured phytoplankton, which allowed for explicit analysis of effects of variability in algal size on inherent optical properties. The study offered a methodology that allowed for detailed investigation of the effects of algal size on optical signals; indicating that phytoplankton community composition (at identical [Chl *a*]) could be responsible for substantial variation in IOPs. For the level of detail achieved by Stramski et al. (2001) cultures were used to to inform the IOPs used for the different groups. These may not representative of natural assemblages (and their associated optical properties), which are often more complex and (Stramski et al., 2001) indicated that their approach would benefit from sufficiently comprehensive *in situ* measurements of marine particles and their IOPs. This recommendation was taken up in work by Green et al. (2003a,b) which used a combination of Mie modelling and flow cytometry measurements to distinguish the contributions of different particle types (both phytoplankton and non-phytoplankton) to IOPs.

Stramski et al. (2001) also highlighted that the inherent assumptions of Mie theory have been criticised as too simplistic to account for the variety of shapes and heterogenous structures present amongst different phytoplankton species; particularly with regards to their impact on algal backscattering (*ibid*). More complex models have been developed to include non-spherical and layered spheres with multiple refractive indices (Bernard et al., 2009; Kitchen and Zaneveld, 1992; Matthews and Bernard, 2013; Quirantes and Bernard, 2004). Other assumptions with regards to the shape of the PSD have been criticised. As mentioned previously, the Junge distribu-

tion has been commonly applied to studies of phytoplankton size distributions. Reynolds et al. (2010) found that whilst for open ocean waters, dominated by small particles, the assumption of power-law type distributions (e.g. Junge) is suitable to characterise the PSD, these methods fail to characterise the PSDs associated with high biomass, monospecific blooms. Cases of monospecific blooms offer an opportunity to advance these methods towards a better understanding of bio-optical principles in high biomass, natural waters. In addition, deriving equivalent size distribution schemes for a variety of mixed assemblages has shown potential to provide an application of additive models for a wider range of natural waters (Bernard et al., 2007). As with all modelling approaches, more numerous and complex parameterisations increase computational requirements. The example approaches referred to here all contain simplifications including: performing computations over reduced spectral ranges, for small size ranges, for spheres/spheroid objects and/or for a spheres with varying numbers of layers. Deciding which are important parameterisations to maintain requires knowledge of the sensitivity of optical signals, in the context of regional optical properties and remote sensing requirements.

As can be seen from the literature reviewed in this and the previous chapter, linking the causes of optical variability with biogeochemical parameters is complex and a work in progress for biological oceanographers and remote sensing scientists. With size and taxonomy of utmost interest, defining the causality between these variables and optical properties observed from a variety of platforms has become a priority. As with any form of modelling in natural systems, parameterisations are required to simplify complexity and reduce computational demands. With regards to size there are a number of parameterisations which must be made. Firstly the particle size distribution (PSD) in natural ocean systems is often complex and highly variable, necessitating the derivation of a single parameter which can adequately describe both the inherent variability in PSDs and their interactions with the light field. The use of equivalent size schemes and the effective diameter (D_{eff}) as a singular descriptive variable, are parameterisations for PSDs that have been frequently used in the field of atmospheric physics (Alexandrov and Lacis, 2000; Hu and Stamnes, 1993; McGraw et al., 1998). Using the Standard size distribution (Hansen and Travis, 1974) has been shown to accurately represent typical PSDs and optical properties attributable to phytoplankton communities in upwelling regions (Bernard et al., 2009, 2007; Robertson Lain et al., 2014).

Although the use of an effective diameter (D_{eff}) to parameterise particle size distributions can be used to incorporate the causal effect of size on optical variables, factors other than size have been shown to significantly influence optical variability. As mentioned previous, Mie modelling techniques have been used to adequately simulate phytoplankton absorption, however with respect to scattering, these models perform less favourably. Consequently understanding and modelling variability of backscattering resulting from differences in phytoplankton communities remains difficult. From a single particle level, how can the abundance of variability found across the range of phytoplankton morphology be simply parameterised? Studies have shown non-sphericity to represent a relatively minor influence on absorption and backscattering (Aas,

1987; Quirantes and Bernard, 2004), whilst intercellular structures have been shown to have substantial effects on optical variability (Bernard et al., 2009; Matthews and Bernard, 2013; Quirantes and Bernard, 2004; Svensen et al., 2007).

2.1.3 Modelling other constituents: gelbstoff, detritus and non-algal particles

From an ecological perspective, the aim of most ocean colour studies has been the derivation of information about the algal biomass. However, there is substantial second order variability in ocean colour data, which can be attributed to absorption and scattering by non-algal components (Brown et al., 2008). Accordingly, an understanding of the optical characteristics of these components throughout the global ocean is necessary to extract the most accurate information about algal assemblages. In addition, information about non-algal components themselves is vital for many coastal and inland water quality monitoring approaches using ocean colour remote sensing. As a result many of the additive optical models discussed above include at least one, if not several distinct, non-algal absorbing and/or scattering components.

Gelbstoff, otherwise known as yellow substance or coloured dissolved organic matter (CDOM) derives from the degradation of organic matter and is primarily composed of humid and fulvic acids (Carder et al., 1989). As could be expected due to its origins, the highest concentrations of gelbstoff are typically found in coastal regions where there is significant land drainage and/or resuspension of sediments. Where terrestrial sources are absent, gelbstoff concentrations have been found to vary with [Chl a], as a by product of phytoplankton degradation (Bricaud et al., 1981), leading to the classic definition of Case 1 (where [Chl a] and phytoplankton degradation products co-vary) and Case 2 (where gelbstoff/detritus do not co-vary with phytoplankton biomass (Morel and Prieur, 1977)). Non-algal particles, such as mineral sediments, have been modelled simply using power law functions (Carder et al., 1999; Lee et al., 2002; Roesler and Perry, 1995). However studies have shown that these components, similarly to phytoplankton, can have variable refractive indices, resulting in different optical properties. The presence of both gelbstoff and non-algal particles, have been shown to cause significant ambiguity in the ocean colour inverse problem, as a result of covarying spectral features imparted by combinations of these constituents with algal biomass (Defoin-Platel and Chami (2007), see chapter three for further discussion on this).

2.1.4 Radiative transfer considerations for modelling reflectance from inherent optical property models

The development of radiative transfer models has vastly enhanced the ability to understand both measurement techniques and the interaction between inherent optical properties and the variable incident light field. However, solving the full, unpolarized, radiative transfer equation (RTE) can prove too computationally intensive for coupled hydrodynamic-biogeochemical modelling and semi-analytical inversion approaches. Reflectance approximation techniques (Zaneveld, 1995) have commonly been used in semi-analytical inversions to avoid this (Garver and Siegel, 1997;

Morel and Prieur, 1977; Roesler and Perry, 1995), however significant assumptions have to be made to use these approximations. Notably, assumptions must be made with regard to the bidirectionality of the light field. These are often represented by an "f/Q" parameter where f is a dimensionless coefficient regulating the upwelling irradiance magnitude and Q is the bidirectional function - the ratio of upward irradiance (just below the surface, (0-) to in-water upward radiance (0-) (Morel et al., 2002). These parameters are dependent on the IOPs and environmental conditions e.g. wind speed and solar zenith angle (*ibid*). Further, the propagation of upwelling irradiance/radiance to the surface when measurements are made or modelled at depth must be considered and is often represented by attenuation coefficients K_u or K_{Lu} (Albert and Mobley, 2003). The details of these parameterisations can have significant impacts on the radiometric values (particularly the remote sensing reflectance, R_{rs}) ultimately calculated. In particular, the use of f/Q parameterisations based on Case 1, single scattering, spectral invariance and [Chl *a*] related assumptions, can introduce significant errors when applied to turbid waters (Aurin and Dierssen, 2012; Hlaing et al., 2012). Radiative transfer models where accuracy of the RTE is compromised for the sake of computational efficiency have been developed in recent years (Mobley, 2011; Mobley et al., 2009). Models to solve the RTE typically require inputs of in-water IOPs, incident sky radiance, bottom reflectance (or an assumption of an "optically deep" ocean) and sea surface boundary conditions (often parameterised through windspeed). Efficiency can be maximised by using an azimuthally integrated version of the RTE and through performing calculations at limited wavelengths, with interpolation between and limiting the depth dependency of the RTE at different wavelengths, where appropriate (Mobley, 2011).

2.1.5 Forward model and experimental approach

A schematic summarising the EAP forward model and experimental approach used in this chapter is shown in figure 2.1.

2.1.6 Summary and chapter aims

Bio-optical models provide a framework within which causes of optical variability can be investigated and used to retrieve biogeochemical parameters from *in situ* and remotely sensed ocean colour radiometry. In this chapter, a dataset of simulated optical data is produced using the Equivalent Algal Population (EAP) forward model run over a range of biogeochemical parameter combinations. The model employs IOPs generated from a phytoplankton population model using equivalent size distributions of two-layered spheres and combined with a reflectance approximation or the radiative transfer model, EcoLight-S. The sensitivity of the simulated optical data to changes in these parameters and the choice of radiative transfer formulation is investigated, with a focus on the role of phytoplankton cell size in this sensitivity. The simulated dataset is then evaluated both in the context of naturally occurring biogeochemical parameter combinations in the southern Benguela, and with a view to spectral sensitivities used in both empirical and semi-analytical inversion algorithms.

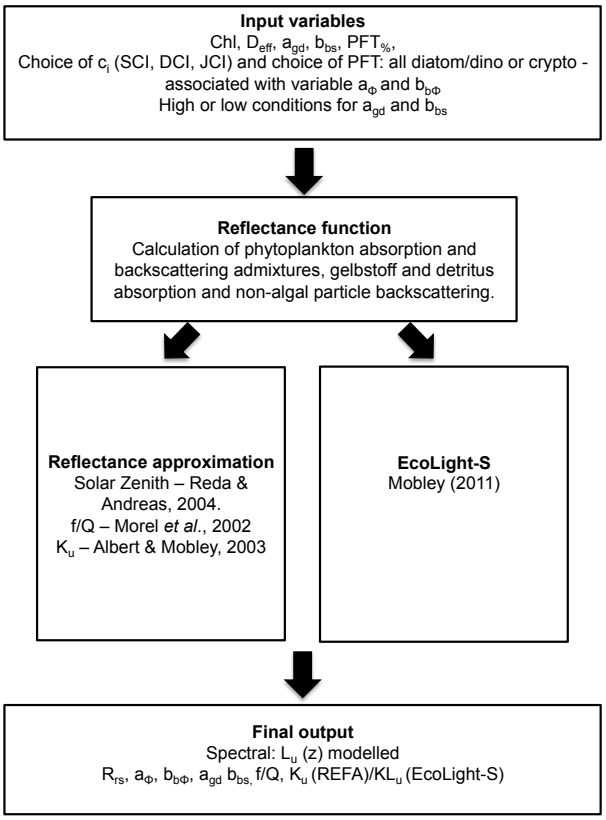

Figure 2.1: Schematic of the Equivalent Algal Population forward model indicating input parameters and options involved with regards to variable c_i, PFT group, a_{gd} and b_{bs} and radiative transfer technique.

2.2 Methods

The dataset created in this chapter has been simulated such that it is as analogous to data collected by in water *in situ* radiometers as possible. This was done to enable comparison to *in situ* data and an assessment of radiometric parameterisations used in future chapters. As such L_u(z), rather than R_{rs} is simulated. In this case, (z) is -0.66m, the fixed depth where measurements are made using the Satlantic Hyperspectral-Tethered surface radiometer buoy (see chapter four for further details of this instrument and measurement methodology).

2.2.1 Forward model development and set up

Bernard et al. (2009) presented a two layered model for investigating the impact of a spherical, two-layered cell geometry on optical interactions. In this model, phytoplankton are reduced

to spheres with two layers. The outer layer represents the chloroplast (set as 20% of the cell volume) and the inner layer, the cytoplasm - based on available morphometric data (Bernard et al., 2009). To couple the optical influence of second order variability such as cell size and pigment related features to estimation of R_{rs} through IOPs, the two-layered sphere model of (Bernard et al., 2009), is used here to generate [Chl a] and size (effective diameter, D_{eff}) related spectral phytoplankton absorption and backscattering basis vectors. The model makes two assumptions. Firstly, that complex phytoplankton communities can be feasibly represented using Standard size distributions, providing the effective diameter and total particle surface area are equivalent (Bernard et al., 2007). Secondly, that a two-layered spherical geometry can reasonably represent the absorbing and backscattering properties of diverse phytoplankton populations (Bernard et al., 2009; Matthews and Bernard, 2013). Basis vectors were generated based on imaginary refractive indexes representative of the intracellular absorption characteristics of several phytoplankton (Bernard et al., 2009). Full details of both the real and imaginary refractive indices, their derivation and use in the two-layered sphere model can be found in Bernard et al. (2009). Of importance for the sensitivity study conducted in this chapter, the imaginary refractive index derived for the chloroplast is dependent on the choice of intracellular chlorophyll concentration (c_i) The use of the heterogenous rather than homogenous geometry results in a considerable increase in backscattering. The location of the chloroplast, and its refractive indices were revealed to be the single most important intracellular variable influencing backscattering. Absorption was found to be fairly similar between both geometries (*ibid*).

$$a_\phi(\lambda) = [Chla].a^*_\phi(\lambda, F^*(d)); \tag{2.1}$$

$$b_{b\phi}(\lambda) = [Chla].b^*_{b\phi}(\lambda, F^*(d)); \tag{2.2}$$

The [Chl a] specific absorption (a^*_ϕ) and backscattering ($b^*_{b\phi}$) derived from the two-layered sphere model are used to calculate phytoplankton absorption (a_ϕ) and backscattering ($b_{b\phi}$) as shown in equations 2.1 and 2.2 respectively. In both equations, F*(d) is the [Chl a] specific size distribution (equation 2.5). The particle size distribution (F(d)) associated with F*(d) in equations 2.1 and 2.2 is given by a Standard distribution (equation 2.3) (Bernard et al., 2009, 2007; Hansen and Travis, 1974)

$$F(d) = ASF\frac{d}{2}^{[(1-3\nu_{eff})/\nu_{eff}]} exp\left[\frac{d}{2}/\left(\frac{D_{eff}2}{\nu}_{eff}\right)\right] \tag{2.3}$$

Where ASF is an area scaling factor (Bernard et al., 2007), which adjusts the magnitude of the equivalent distribution to the total projected surface area in the particle size distribution and ν_{eff} is the effective variance, which describes the width of the Standard distribution, for which a constant of 0.6 is used (Quirantes and Bernard, 2004). To scale this to a [Chl a] specific size distribution as required for generation of the basis vectors, the total relative particle volume of

the Standard distribution ($\bar{V}$) is calculated:

$$\bar{V} = \frac{\pi}{6} \int F(d) d^3 d(d) \tag{2.4}$$

And the [Chl *a*] specific size distribution is calculated from this:

$$F^*(d) = \frac{F(d)}{\bar{V} c_i} \tag{2.5}$$

where c_i is the intracellular chlorophyll *a* concentration (kg m^{-3}).

Future application to *in situ* data prompted a sensitivity investigation of several key assumptions using different variants and implementations of output from the two-layered model (Bernard et al., 2009; Robertson Lain et al., 2014). Firstly, the use of multiple groups of phytoplankton basis vectors was assessed. The two-layered model allows for the optical properties of different groups of phytoplankton species to be simulated through the application of different spectral refractive index data (see Bernard et al. (2009) for details). Two sets of basis vectors were generated to simulate known differences in the broad absorption characteristics of diatoms/dinoflagellate and cryptophytes/cryptophyte endosymbionts (e.g. *Mesodinium rubrum*). Values of 3.5 kg m^{-3} and 2.5 kg m^{-3} were used for the intracellular c_i of each group respectively. Secondly, in the original variant of this model the intracellular concentration of chlorophyll (c_i) is assumed constant with size (henceforth referred to as the static c_i, or SCI basis vectors). However, there is evidence that this is not representative of reality in that generally, with increasing size, c_i is reduced (Agustí, 1991; Marañón et al., 2007; Roy et al., 2011; Zhang et al., 2011). The implication of this from an optical perspective is that, given constant c_i, phytoplankton absorption maybe overestimated for communities where large cells dominate the biomass. To investigate this, a variable c_i, hereafter referred to as the dynamic c_i (DCI) was introduced. Variability in c_i with cell size was parameterised according to the relationship shown in equation 2.6, formulated from the results described in Zhang et al. (2011). These c_i values were applied on a per size bin level (from D_{eff} of 2 to 50 μm, at 1 μm intervals) in the two-layered sphere model to generate the [Chl *a*] and D_{eff} specific IOP basis vectors. It should be noted that even this dynamic c_i parameterisation still represents a simplification of reality, as c_i has also been shown to vary on a species level and as a result of the light environment (Agustí, 1991; Kirk, 1975; Marañón et al., 2007; Morel and Bricaud, 1981; Raven, 1984; Sathyendranath et al., 1987; Sosik and Mitchell, 1991).

$$D_{c_i} = 7.94.D_{eff}^{-0.3} \tag{2.6}$$

A final assumption, relating to the use of Standard size distributions was also tested. Given that in most open ocean, low [Chl *a*] environments, small celled phytoplankton dominate the biomass, a Jungian distribution can be used to suitably summarise the particle size distribution,

whilst a Standard size distribution may be most appropriate for characterisation of PSD's associated with mono-specific blooms. A new set of phytoplankton absorption and backscattering basis vectors were generated from the two-layered sphere model assuming a Jungian distribution. This Jungian distribution employed a slope value of 4.02, corresponding to an effective diameter of 3 μm. To simulate the impact of a presence of smaller cells at low [Chl a], these basis vectors were used to form a [Chl a] dependent admixture with those generated from the Standard size distributions. Admixture proportions were decided dependent on [Chl a] such that at concentrations up to and including 1 mg m^{-3}, full basis vectors were taken from the Jungian distribution model, and beyond this basis vectors from the Standard size distribution model were added in increments of 10% per 1 mg m^{-3} increase, up to 10 mg m^{-3}. These IOP basis vectors are henceforth referred to as "JCI".

Basis vectors for a given [Chl a] and D_{eff} are used in combination with seawater absorption and backscattering estimates, gelbstoff/detritus absorption (a_{gd}) and backscattering by non-algal particles (b_{bs}) to calculate total absorption and scattering and backscattering. Absorption and backscattering of seawater were accounted for using the data from a number of sources which is offered as default within Hydrolight 5.0 (Mobley and Sundman, 2008). The spectral absorption of gelbstoff and detritus (a_{gd} are represented using an exponential shape function (Roesler and Perry, 1995), where a scaling factor is explicitly solved for and the exponential slope factor "S" was determined as a constant of 0.012 for typical waters in the Benguela (Bernard et al., 1998). Small, non-algal, particle backscattering is represented by a $\lambda(^{-1.2})$ power law, and a scale factor is explicitly solved for (*ibid*).

$$a_{gd}(\lambda)(m^{-1})) = a_{gd}(400)exp[-S(\lambda - 400)] \tag{2.7}$$

2.2.2 Radiative transfer considerations

In preparation for future applications to *in situ* and satellite data, the forward model configuration used here was formulated to simulate these type of measurements and the uncertainties associated with their use. Two different techniques were used to address radiative transfer requirements. Firstly, a reflectance approximation (REFA) based method was used. The original equation from Zaneveld (1995) (equation 3.1) was reformulated to allow for calculation of L_u(z) (equation 2.8).

$$L_u(z,\lambda) = \frac{f}{Q}\frac{b_{\mathrm{b}}(\lambda)}{a(\lambda)+b_b(\lambda)}\tau\frac{E_d(\lambda,0+)}{exp(K_u(\lambda)z)} \tag{2.8}$$

Where the upwelling radiance at depth z is explicitly solved for, allowing for comparison to (and input from) the Satlantic Hyperspectral Tethered Surface Radiometer Buoy at z = -0.66 m (μW cm^{-2} nm^{-1} sr^{-1}). To calculate R_{rs}, a $\eta^2\tau$ is used as an assumed constant for the transmittance of upwelling radiance across the air-sea interface.

The bidirectionally of the upwelling radiance is described by the f/Q parameter in equation (2.8). This parameter is dependent on the surface roughness, solar zenith angle and the, often

highly variable, IOPs of seawater, particularly the backscattering efficiency and phase function (Morel et al., 2002). In the absence of full radiative transfer simulation, Morel et al. (2002) offer look-up tables where an f/Q value can be selected based on [Chl a] and solar zenith angle (θ). These look up tables were linearly interpolated to a 5nm resolution for use with the hyper spectral Satlantic H-TSRB data and a constant f/Q value associated with chlorophyll concentration of 10 mg m^{-3} was used when input values were greater than 10.

The upward diffuse attenuation coefficient (K_u) is parameterised as per the formulation of Albert and Mobley (2003) (equation 2.9), using total absorption and backscattering and the solar zenith angle. A constant, seasonally averaged downwelling irradiance (E_d) generated from EcoLight-S was used for all forward model runs and the solar zenith angle (θ) assumed to be a constant of 0. As a number of key components of the forward model described here are dependent on the solar zenith angle, the assumption of a constant value is a simplification, to increase the applicability of the simulated data, further simulations should be conducted over a range of solar zenith angles.

$$K_u(\lambda) = (a + b_b \left[1 + (\frac{b_b}{a + b_b})\right]^{3.452} \left(1 - \frac{0.2786}{cos\theta_s}\right) \tag{2.9}$$

A second approach was constructed using EcoLight-S (ES) (Mobley, 2011) to replace the reflectance approximation and associated parameterisations (f/Q , K_u etc). f/Q values for comparison with the Morel et al. (2002) look up tables were calculated from EcoLight-S using equation 2.10 (Morel et al., 2002), where E_u is the upwelling irradiance. The upward diffuse attenuation coefficient as derived from EcoLight-S is referred to as K_{Lu} such as to distinguish it from the Albert and Mobley (2003) formulation.

$$f/Q(\lambda) = \left[\frac{E_u * (a + b_b)/E_d * b_b}{E_u/L_u}\right] \tag{2.10}$$

A spectrally variable Fournier-Forand phase function based on the total backscattering probability was selected as the most suitable for the broad range of waters to be simulated in this study (Mobley et al. (2002) and see discussion with reference to the forward model developed in this chapter in (Robertson Lain et al., 2014)). It should be noted that the angular variation of the Fournier Forand phase functions is fairly flat in the backwards direction, whilst various experimental approaches have shown that directional effects of phytoplankton may be significant in the backwards direction (Chami et al., 2014) and this may particularly effect the derivation of f/Q factors. Further consideration may well be needed to provide more suitable phase functions (Robertson Lain et al., 2014), however this is outside of the scope of the study here.

The output from the two-layered sphere model (Bernard et al., 2009), combined with the above parameterisations is employed here to derive a set of simulated optical data suitable for testing the various assumptions about phytoplankton cell size and other characteristics inherent in the model and estimate radiometric sensitivity to these. In this context the model is known as the Equivalent Algal Population (EAP) forward model.

2.2.3 Sensitivity assessment

A simulated dataset of L_u (z) (μW cm^{-2} nm^{-1} sr^{-1}), R$_{rs}$ (sr^{-1}) and associated IOPs were generated from the forward model described above over a range of the 5 solvable unknowns in this model configuration (table 2.1). This model dataset is henceforth referenced to as FWD$_A$.

Table 2.1: Summary of parameter ranges used in initial forward model data simulation

Parameter	*Range*
Chlorophyll	0.1, 0.5, 1, 3, 5, 10, 15, 20, 30 50, 100, 300 mg m^{-3}
Effective diameter	2, 5, 10, 20, 30, 40 μm
Combined absorption (m^{-1}) of gelbstoff + detritus (a$_{gd}$) at 400nm	Low (0.02) and High(2)
Non-algal particle backscattering (m^{-1}) (b$_{bs}$) at 550nm	Low(0.0005) and High (0.5)
Diatom/dinoflagellate vs cryptophyte basis vectors (PFT$_{\%}$)	0 and 100 %

Values for the 5 solvable unknowns were selected to encompass both open ocean conditions and the extremes which may be experienced in highly turbid waters and Harmful Algal Blooms, and can be seen in table 2.1. Values for [Chl a] were selected between the lower concentrations typically observed offshore in the southern Benguela (0.1-30 mg m^{-3}) and the higher concentrations observed during bloom conditions (50, 100 and 300 mg m^{-3}). Effective diameters were chosen to cover the range of cell sizes observed in both bloom and non-bloom conditions. Low values for combined gelbstoff and detritus (a$_{gd}$) and non algal particulate backscattering (b$_{bs}$) were selected as generally representative of Case 1 waters and the conditions in the southern Benguela where there are relatively few terrigenous sources (Bernard et al., 2001, 1998). High values for both a_{gd} and b_{bs} were also selected to simulate Case 2 conditions. Although these values are likely more extreme than those which occur in the southern Benguela, this should provide some wider context for the sensitivity of reflectance to phytoplankton characteristics such as cell size, under these conditions. A sensitivity study was conducted on this forward model output. Spectral sensitivity in the calculated radiance and reflectance was estimated in response to variation of the five parameters. This analysis was then repeated to assess the impact of (1) the use of a static or dynamic c_i assumption, (2) the effect of using Standard versus jungian size distributions at low [Chl a] and, (3) the resultant effects on size sensitivity in radiance and reflectance spectra manifested through each of the radiative transfer techniques (REFA and ES).

2.3 Results

2.3.1 Sensitivity in reflectance spectra as a result of varying biogeochemical parameters

Sensitivity in R_{rs} spectra is summarised in figures 2.2 to 2.4 for a range of biogeochemical parameter combinations generated using the diatom/dinoflagellate IOPs, three c_i assumptions, and reflectance approximation (REFA) and EcoLight-S (ES). Examples of the spectral range of reflectance associated with changes in D_{eff} at different [Chl a] are also shown (Fig. 2.2(c, f, i) to 2.4(c, f, i)). Changes in R_{rs} as a result of changes in D_{eff} with [Chl a] are indicated by the thickness of the curves. Reflectance sensitivity to size is both spectrally variant and influenced by [Chl a] indicated by the different R_{rs} magnitudes at different wavelengths and [Chl a]. However this sensitivity is also highly influenced by the presence of other constituents (figures 2.3 and 2.4). In the high a_{gd}/low b_{bs} case (Fig. 2.3), the R_{rs} and associated size sensitivity are suppressed, particularly at shorter wavelengths (445 nm and 520 nm) and under lower [Chl a]. Substantial size related sensitivity in this instance, only manifests in the green (550 nm) and red (710 nm) beyond [Chl a] of above 5 mg m^{-3}. Similarly for the high a_{gd}/high b_{bs} case (Fig. 2.4), size sensitivity is again suppressed, though R_{rs} magnitudes are much higher as a result of the increased backscattering. The trends in size related sensitivity are different under this scenario as well, with smaller sizes generally resulting in lower R_{rs} values. Overall, maximum size related sensitivity occurs at low b_{bs} and low a_{gd} in the green (here at 550 nm) for both the REFA (Fig. 2.2(a, d, g)) and ES (Fig. 2.2 (b, e, h)) above [Chl a] of 1 mg m^{-3}. Substantial size related variability can be observed under all scenarios in the red (710 nm). Non-linear responses in size sensitivity are observed, most clearly in the low a_{gd}/low b_{bs} scenario (Fig. 2.2). Particularly at the shorter wavelengths, as [Chl a] increases, R_{rs} associated with smaller sizes increases, resulting in the non-linear sensitivity observed. Looking at individual spectra associated with these changes, in some spectral regions peaks shift towards/away from the bands chosen for these figures, creating this effect.

From figures 2.2 to 2.4, it is apparent that significant differences in model response can occur between the REFA and ES approaches. Under the low a_{gd}/low b_{bs} scenario (Fig. 2.2), the size related responses are broadly similar between the REFA (a, d, g) and ES (b, e, h) approaches, with a couple of exceptions. Firstly, R_{rs} magnitudes are generally lower from the ES method, particularly at low [Chl a]. Secondly the REFA approach, exhibits a tailing off of size related sensitivity at 550 nm with [Chl a], such that peak sensitivity is seen between 15 and 20 mg m^{-3}, followed by a decline. In contrast, using the ES approach, size related sensitivity in both the green and red continues to increase with [Chl a] although the sensitivity does begin to plateau at higher biomass. Similar behaviour is observed at higher biomass at 710 nm, although the size related R_{rs} sensitivity in the ES method does not appear to tail off at higher biomass for this wavelength. As the phytoplankton absorption and backscattering used for both the REFA and ES simulations are identical, these differences must result from assumptions made with regards

to radiative transfer and the suitability of these, particularly at high biomass.

2.3.2 Sensitivity in radiance and reflectance spectra as a result of IOP model structure

In general, applying a dynamic c_i reduces the size related sensitivity in R_{rs}, however the extent of this reduction is spectrally variant, given the spectral shape associated with the absorption/backscattering of phytoplankton in the IOPs. Comparing figure 2.2(a, b, c) (SCI) with figure 2.2(d, e, f) (DCI), the greatest differences can be seen at 550 nm, i.e. in closest proximity to peak chlorophyll *a* absorption. Adding IOPs generated using a Jungian distribution, as a [Chl *a*] dependent admixture, creates the spectral response shown in figure 2.2. At lower [Chl *a*] (below 10 mg m^{-3}) where the admixture is implemented, the sensitivity to size is reduced particularly in the blue.

Using the cryptophyte basis vectors does not produce substantially different spectral responses (not shown) at the selected wavelengths used in figures 2.2 to 2.4. Although a slightly decreased sensitivity to size is seen in general across the four wavelengths, as a result of the lower c_i used to generate the cryptophyte basis vectors. Significant differences in R_{rs} and associated size sensitivity can however be seen at other wavelengths as a result of the simulated absorption effects of accessory pigments. Figure 2.5 shows the differences between R_{rs} generated using diatom/dinoflagellate ($R_{rs(d)}$) and cryptophyte ($R_{rs(c)}$) basis vectors. At 520 nm $R_{rs(c)}$ is greater than $R_{rs(d)}$, whilst at 600 nm $R_{rs(d)}$ is greater than $R_{rs(c)}$. For the diatom/dinoflagellate dominated assemblages modelled here, peak reflectance at higher biomass (above 3-10 mg m^{-3}) is seen between 550 and 570nm, whilst for the cryptophyte dominated assemblages, this peak is bifurcated as a result of absorption by accessory pigments (including phycoerythrin). As above, the extent of these effects is biomass and size dependent and significantly impacted by the presence of other absorbing and scattering components (Fig. 2.5(d, e, f) and 2.5(g, h, i)). In the high a_{gd}/low b_{bs} case (Fig. 2.5(d, e, f)), differences in R_{rs} as a result of the different basis vectors only become apparent at [Chl *a*] above 10 mg m^{-3}. Responses related to the different radiative transfer techniques are similar to those observed in the diatom/dinoflagellate only R_{rs} examples (figures 2.2 to 2.4, including the apparent "tail off" of size related variability when using the REFA method (Fig. 2.5(a, d, e)) and greatly increased R_{rs} in the high b_{bs} and high a_{gd} case (Fig. 2.5(h and i)).

Similar trends are seen when the cryptophyte data are generated using the DCI parameterisation (not shown). R_{rs} magnitude is reduced as observed using the diatom/dinoflagellate basis vectors (see Fig. 2.2). The R_{rs} differences between the two groups at other wavelengths are also reduced, as would be expected given the reduced R_{rs} magnitudes seen for both groups under the DCI parameterisation.

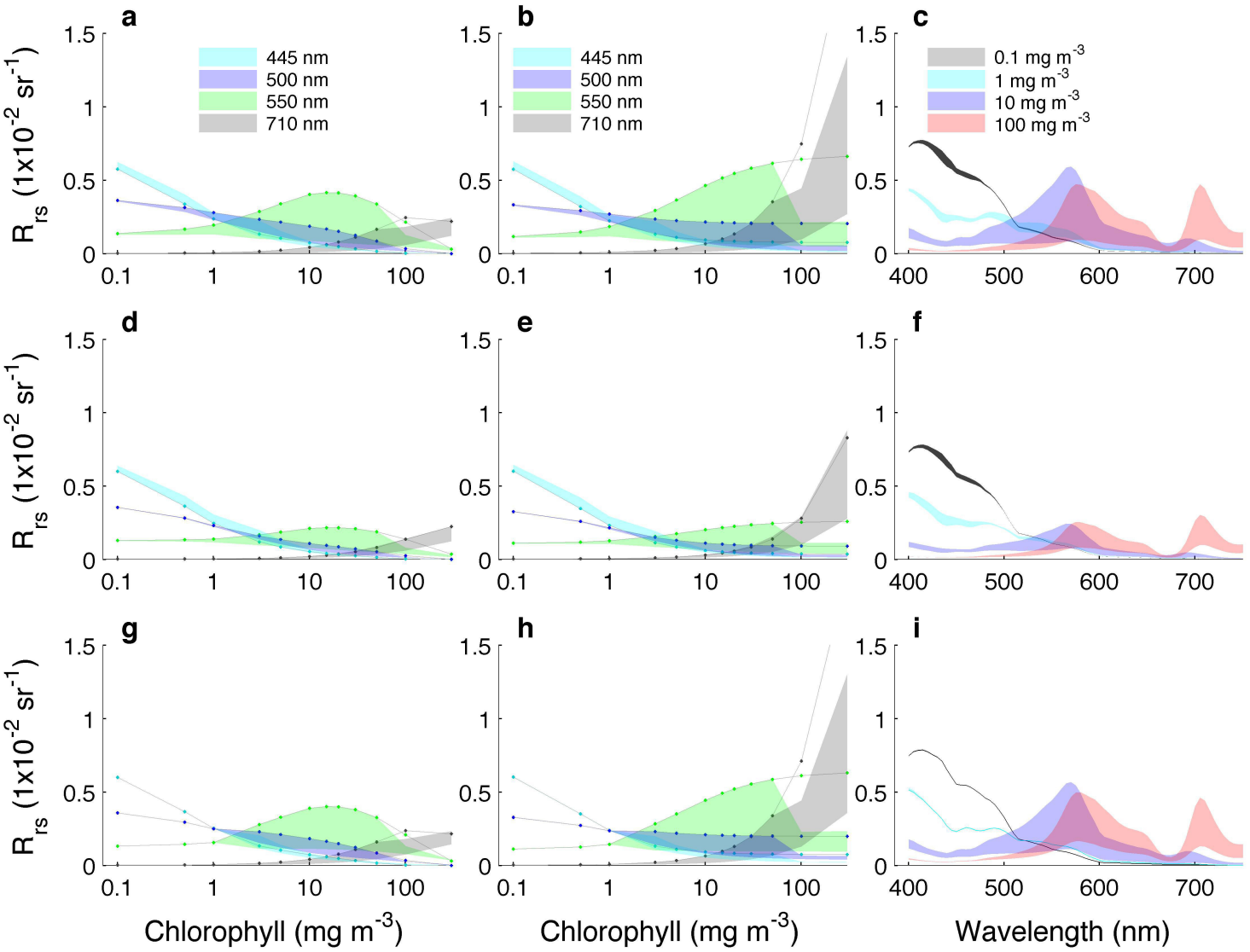

Figure 2.2: Spectral sensitivity of modelled R_{rs} to variations in size (effective diameter) and biomass ([Chl a]), under low b_{bs} and low a_{gd} conditions for the (a, d, g) REFA and (b, e, h) EcoLight-S approaches using SCI (a, b, c), DCI (d, e, f) and JCI (g, h, i). Dots indicate R_{rs} associated with the smallest cells. Panels c, f and i show example ranges of spectral R_{rs} at selected [Chl a] across the modelled size range using ES. The shaded range of the curves shows the range of R_{rs} associated with changes in size (from a minimum of 2 μm to maximum of 40 μm). A filter has been applied to remove ecologically unrealistic spectra i.e. dominance of large cells (+20 μm) at low [Chl a] (<10 mg m^{-3}) and small cells (<10 μm) at high [Chl a] (>30 mg m^{-3}), see section 2.4.2 for discussion.

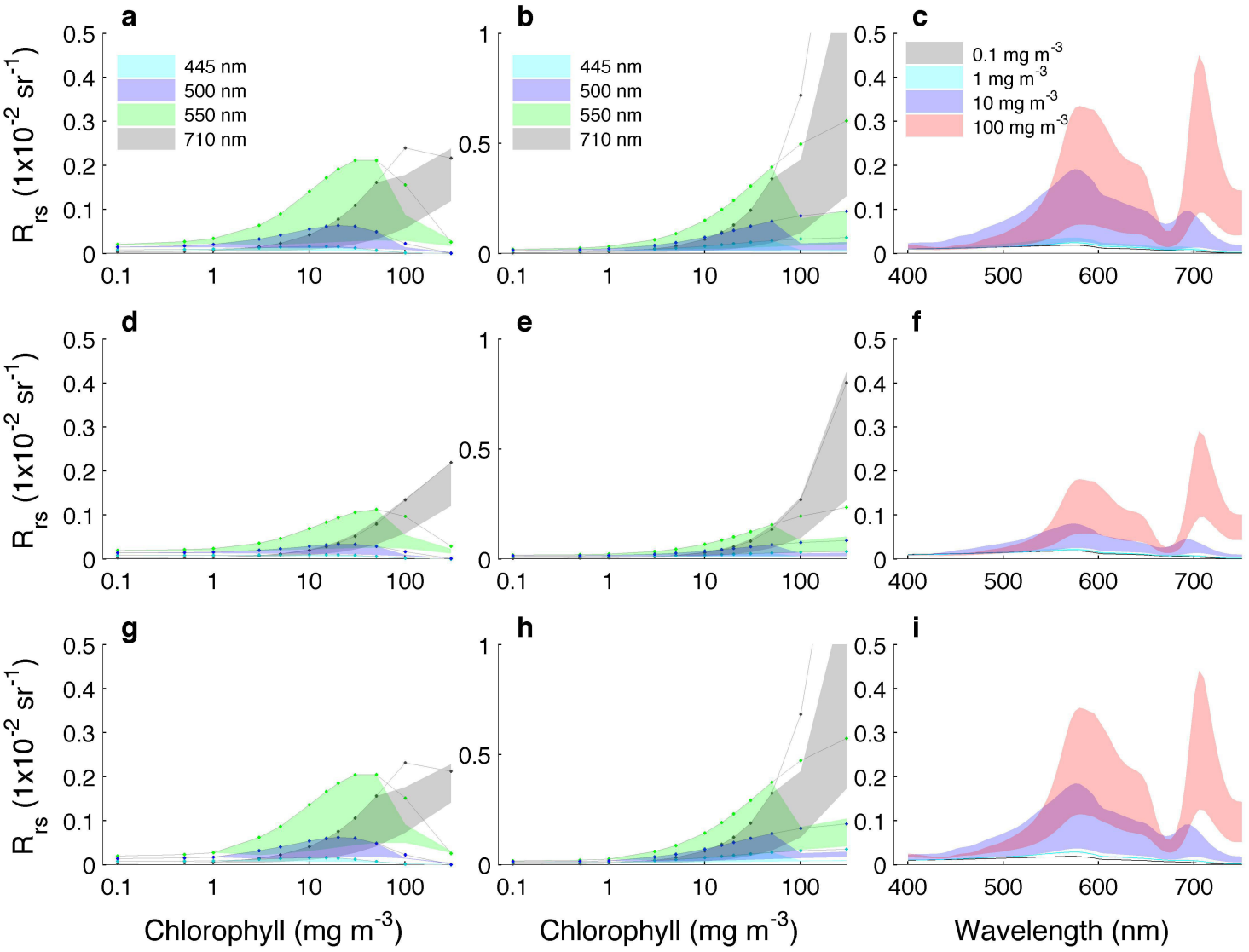

Figure 2.3: Spectral sensitivity of modelled R_{rs} to variations in size (effective diameter) and biomass ([Chl a]), under low b_{bs} and high a_{gd} conditions for the (a, d, g) REFA and (b, e, h) EcoLight-S approaches using SCI (a, b, c), DCI (d, e, f) and JCI (g, h, i). Dots indicate R_{rs} associated with the smallest cells. Panels c, f and i show example ranges of spectral R_{rs} at selected [Chl a] across the modelled size range using ES. The shaded range of the curves shows the range of R_{rs} associated with changes in size (from a minimum of 2 μm to maximum of 40 μm). A filter has been applied to remove ecologically unrealistic spectra i.e. dominance of large cells (+20 μm) at low [Chl a] (<10 mg m^{-3}) and small cells (<10 μm) at high [Chl a] (>30 mg m^{-3}), see section 2.4.2 for discussion.

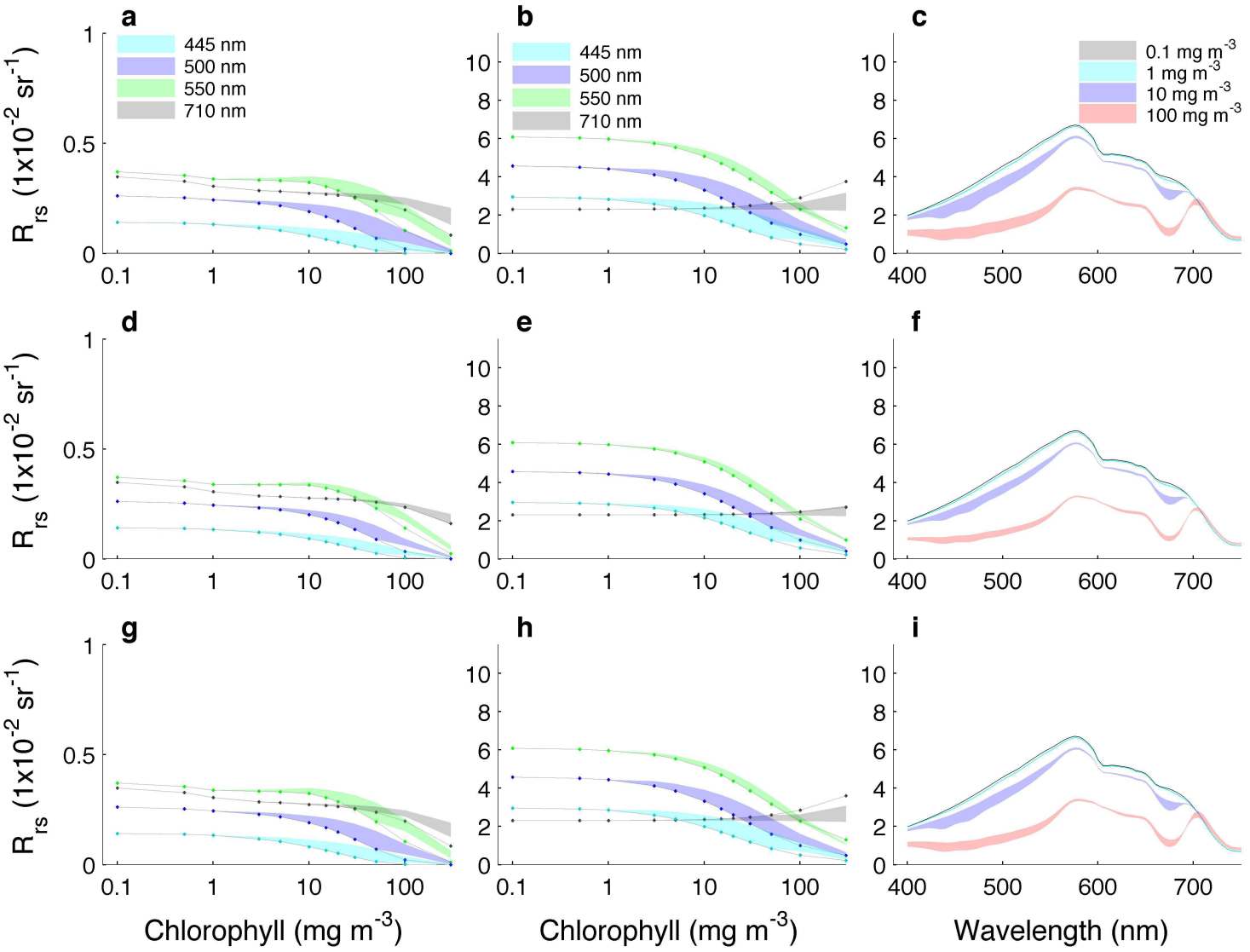

Figure 2.4: Spectral sensitivity of modelled R_{rs} to variations in size (effective diameter) and biomass ([Chl a]), under high b_{bs} and high a_{gd} conditions for the (a, d, g) REFA and (b, e, h) EcoLight-S approaches using SCI (a, b, c), DCI (d, e, f) and JCI (g, h, i). Dots indicate R_{rs} associated with the smallest cells. Panels c, f and i show example ranges of spectral R_{rs} at selected [Chl a] across the modelled size range using ES. The shaded range of the curves shows the range of R_{rs} associated with changes in size (from a minimum of 2 μm to maximum of 40 μm). A filter has been applied to remove ecologically unrealistic spectra i.e. dominance of large cells (+20 μm) at low [Chl a] (<10 mg m^{-3}) and small cells (<10 μm) at high [Chl a] (>30 mg m^{-3}), see section 2.4.2 for discussion.

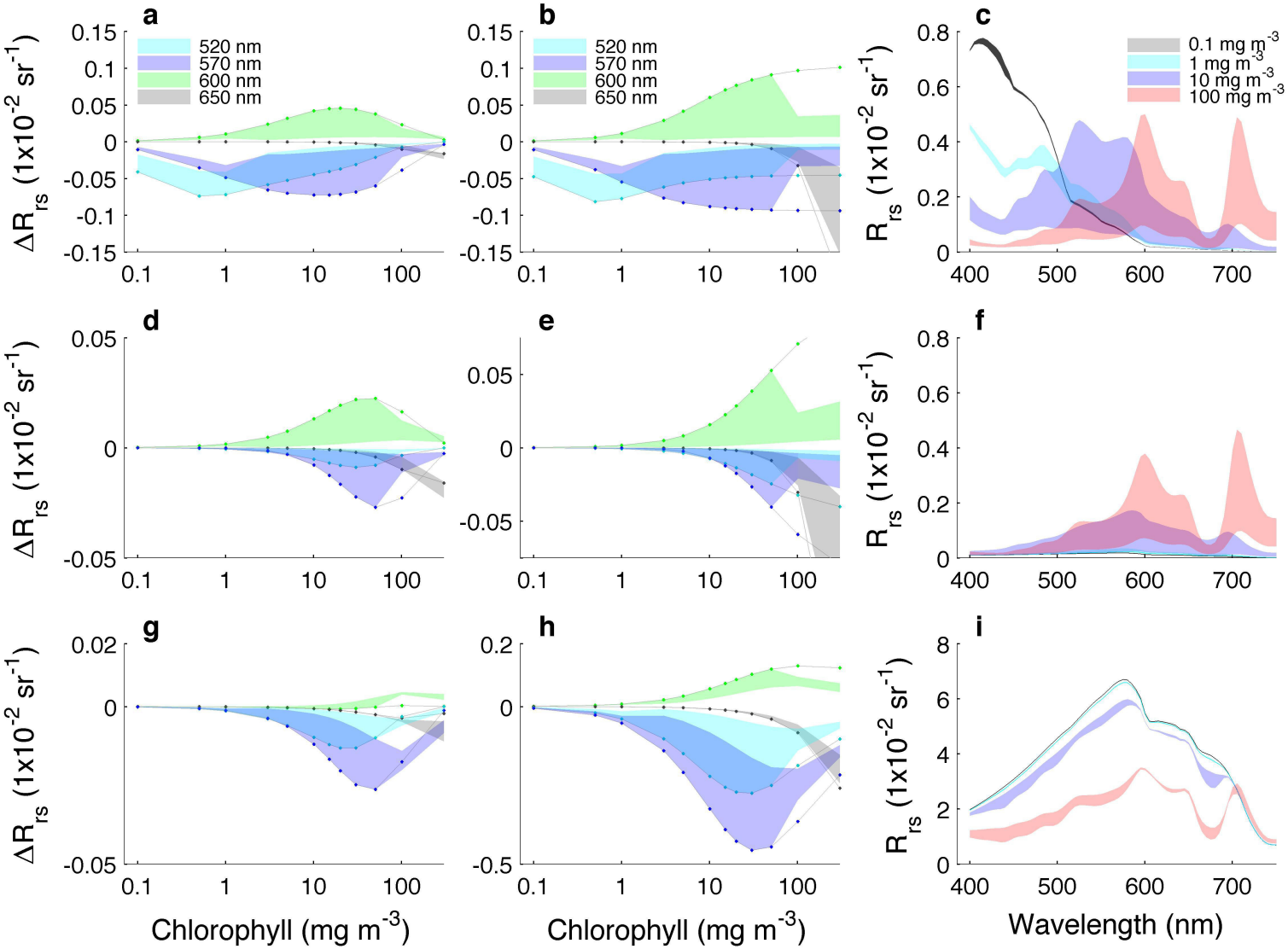

Figure 2.5: Differences in modelled R_{rs} for diatom/dinoflagellate ($R_{rs(d)}$) and cryptophyte ($R_{rs(c)}$) assemblages at selected wavelengths with variations in size (effective diameter) and biomass ([Chl *a*]) using the (a, d, g) REFA and (b, e, h) EcoLight-S approaches under low b_{bs} and low a_{gd} conditions (a, b, c), low b_{bs} and high a_{gd} conditions (d, e, f) and high b_{bs} and high a_{gd} conditions (g, h, i). Dots indicate R_{rs} associated with the smallest cells. Panels c, f and i show example ranges of spectral $R_{rs(c)}$ at selected [Chl *a*] across the modelled size range using ES. The shaded range shows the range of R_{rs} associated with changes in size (from a minimum of 2 μm to maximum of 40 μm). A filter has been applied to remove ecologically unrealistic spectra i.e. dominance of large cells (+20 μm) at low [Chl *a*] (<10 mg m^{-3}) and small cells (<10 μm) at high [Chl *a*] (>30 mg m^{-3}), see section 2.4.2 for discussion.

2.4 Discussion

2.4.1 Inherent optical property budgets

As the modelled reflectance spectra from the forward model are the combined result of several absorbing and scattering components, influenced by both the spectral magnitude and shape of the absorption and scattering characteristics of these constituents and the subsequent effect on radiative transfer calculations, it is useful to establish how the relative contributions of each component vary spectrally under different parameter combinations. Using this budgetary approach, the observations made in the sensitivity study above can be explained.

The budgets for total absorption under the low a_{gd} and high a_{gd} scenarios summarised in figures 2.2 to 2.4, are shown in figures 2.6 to 2.7 respectively. These figures use data from the forward model runs with static c_i assumption and basis vectors from the diatom/dinoflagellate group.

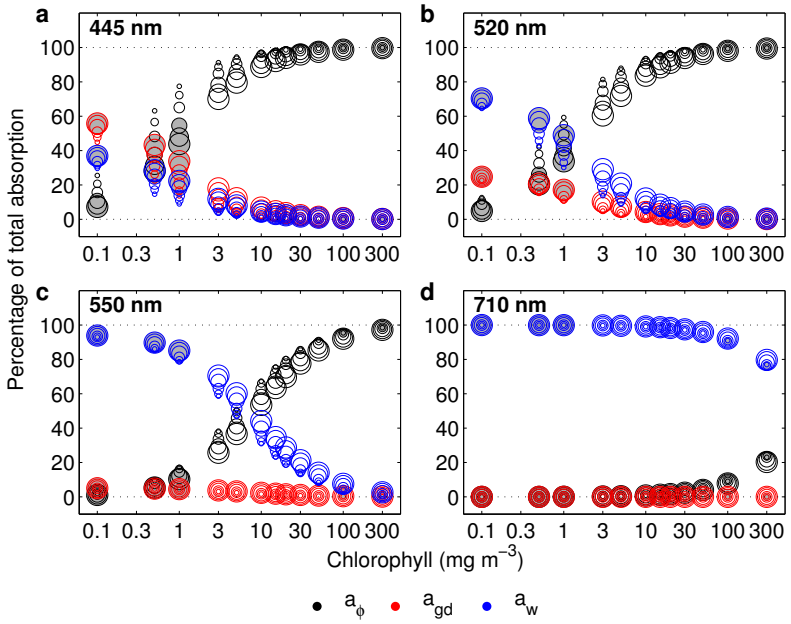

Figure 2.6: Proportions of total absorption attributed to different components of the additive model at different wavelengths, for the low a_{gd} scenarios, over a range of sizes and [Chl a]. Size is indicated by increasing size of bubble between 2 and 40 μm. Grey filled bubbles signify scenarios which are ecologically unlikely, see section 2.4.2 for discussion.

In the low a_{gd} case, at 445 nm, contributions to total absorption from phytoplankton begin to overtake contributions from other components at low [Chl a] (i.e. above 0.5 mg m^{-3}). Assemblages with smaller D_{eff} absorb more and so absorption from these assemblages becomes a dominant contributor at lower [Chl a] values than for assemblages dominated by larger cells. Similar trends are seen at 520 nm and 550 nm, although the intersection where the phytoplankton absorption becomes dominant versus the other components occurs at progressively higher [Chl a] (i.e. above 3 mg m^{-3} at 520 nm and above 10 mg m^{-3} at 550 nm. At 710 nm, the absorption budget is overwhelmingly dominated by water absorption. When the a_{gd} is increased

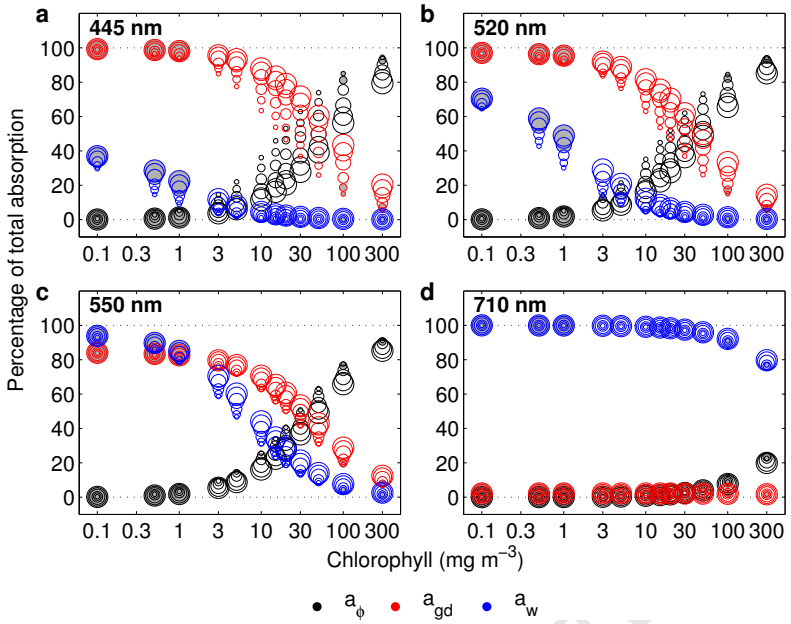

Figure 2.7: Proportions of total absorption attributed to different components of the additive model at different wavelengths, for the high a_{gd} scenarios, over a range of sizes and [Chl a]. Size is indicated by increasing size of bubble between 2 and 40 μm. Grey filled bubbles signify scenarios which are ecologically unlikely, see section 2.4.2 for discussion.

to the high value, the contribution to a from a_ϕ does not exceed that from a_{gd} until much greater [Chl a] across most wavelengths shown in figure 2.7.

In reality it is unlikely that a_{gd} values would be constant across the ranges of [Chl a] modelled here, except in the case of high input from terrestrial sources etc. Hence it is likely (especially under the low a_{gd} scenarios) that the size related sensitivity would be further suppressed in relation to absorption from a_{gd}, as [Chl a] increased. This effect would be seen most strongly in the blue where spectral absorption from a_{gd} is highest, as is seen in the high a_{gd} case (Fig. 2.7). Similarly, in waters with very low [Chl a], where a Case 1 assumption is valid, a_{gd} may be less than that presumed in the low scenario here, resulting in a greater size related sensitivity, especially at blue wavelengths.

In general, phytoplankton contribute substantially less to the total backscattering than other components, across most values of [Chl a] and across different wavelengths (figures 2.8 and 2.9). However, size related sensitivity in the backscattering budget can be substantial, particularly under low contributions from non-algal components (Fig. 2.8). At 445nm, under low b_{bs} conditions, the smallest cells make more than a 30% contribution to backscattering at [Chl a] as low as 3 mg m^{-3} and $b_{b\phi}$ becomes the dominant source of backscattering at this wavelength for the smallest D_{eff} at [Chl a] of 100 mg m^{-3}.

The budgets above can be used to understand the size related sensitivity in R_{rs} shown in figures 2.2 to 2.5. As biomass levels increase (i.e. above 1-3 mg m^{-3}), the effects of changes in size become most significant in the green (550 nm) when the effects of a minima in chlorophyll a and size related absorption are combined with a decrease in backscattering from pure water at these

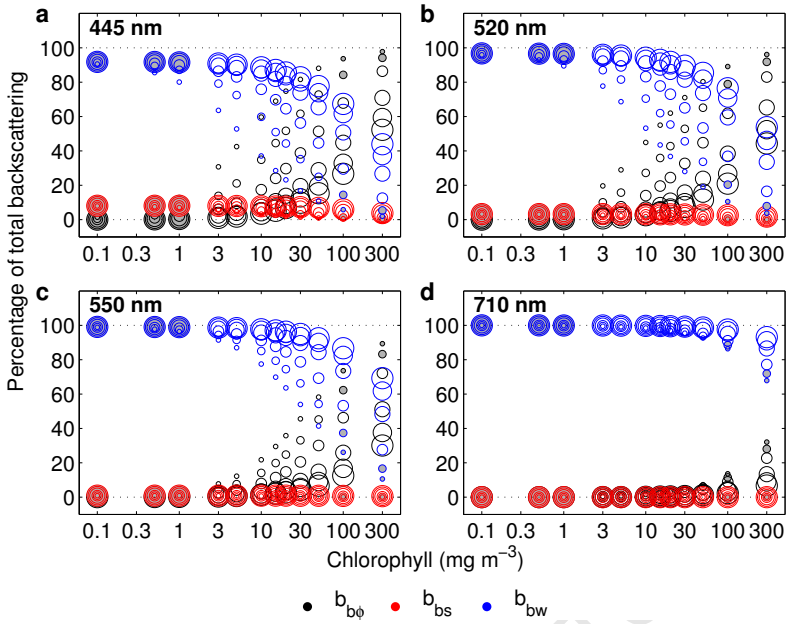

Figure 2.8: Proportions of total backscattering attributed to different components of the additive model at different wavelengths, for the low b_{bs} scenarios, over a range of sizes and [Chl a]. Size is indicated by increasing size of bubble between 2 and 40 μm. Grey filled bubbles signify scenarios which are ecologically unlikely, see section 2.4.2 for discussion.

wavelengths. Smallest cells do not always produce the highest R_{rs} as spectral shape imparted by the combination of changes in biomass and size in both the absorption and backscattering of phytoplankton create a non-linear response at specific [Chl a] and D_{eff} e.g. as seen at the selected wavelengths, particularly in figure 2.2. As [Chl a] increases, the position of the phytoplankton absorption minima with respect to that of pure water absorption shifts from shorter to longer wavelengths (Dierssen et al., 2006), resulting in varying locations of peak R_{rs}.

The changes in IOP contributions are determined by the forward model (and limited selection of parameter values) used here, and there would likely be considerably more variability in natural waters. However, the discussion here should indicate the value of using an IOP budget approach to consider the likely sensitivities and errors associated with simulated and *in situ* optical data, and their use e.g. in inversion algorithms. Limitations of the simulated dataset produced in this chapter and ways in which it could be expanded, are discussed in the following sections.

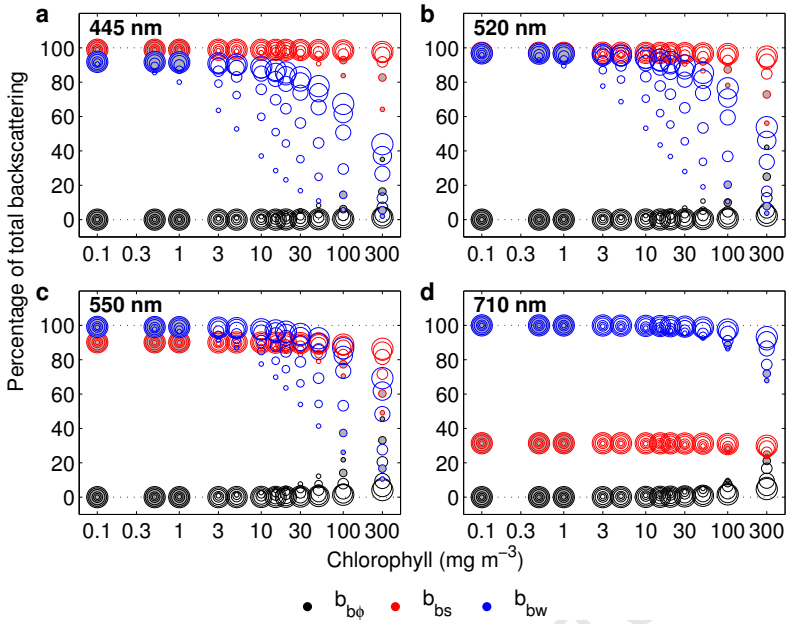

Figure 2.9: Proportions of total backscattering attributed to different components of the additive model at different wavelengths, for the high b_{bs} scenarios, over a range of sizes and [Chl a]. Size is indicated by increasing size of bubble between 2 and 40 μm. Grey filled bubbles signify scenarios which are ecologically unlikely, see section 2.4.2 for discussion.

2.4.2 Suitability of simulated data for real world application

For coherency and ease of implementation the forward model was run over all possible parameter combinations within the ranges specified. However, whilst these ranges might be suitable in isolation, the full range of each of these parameters may not occur coincidently over the full range of each of the others. Relationships between [Chl a] and cell size have been observed (Chisholm, 1992b; Yentsch and Phinney, 1989) and, as discussed in the introduction to this chapter, have formed the basis of abundance methods for deriving information on phytoplankton community structure. These observations indicate some ecologically unlikely (if not impossible) scenarios. An example would be the occurrence of an extremely high [Chl a], coincidently with a very small effective diameter. Phytoplankton of such small size, even in bloom conditions do not contain enough chlorophyll a on a per cell level to reach these concentrations due to allometric scaling (Agustí et al., 1987; Irwin et al., 2006). Similarly, it is unlikely that effective diameters of 20-40 μm would occur at lower biomass concentrations, given the established paradigm of dominance of smaller cells at lower [Chl a] in oligotrophic conditions (Aiken et al., 2008; Marañón, 2008; Margalef, 1978). This likelihood is further supported in terms of optical studies by the suitability and consistent application of a Jungian distribution in size studies in low biomass waters, where to achieve such D_{eff} values would require a Jungian slope far smaller than the range between 2.5 and 5.0 typically used (Andrews et al., 2011; Chami and Defoin-Platel, 2007; Chami and Robilliard, 2002; Guidi et al., 2009).

To contextualise the sensitivity to size seen here, with respect to current ocean colour algorithms and a more realistic co-variance of IOPs, a second set of forward model runs (henceforth

referred to as FWD$_N$) were conducted to approximate the NASA bio-Optical Marine Algorithm Dataset (NOMAD) (Werdell and Bailey, 2005). NOMAD has been used in the generation and evaluation of many ocean colour algorithms. Hence, creating a similar dataset here allows for a preliminary investigation of the role of size in second order variability in the context of both global optical measurements and empirical band ratio algorithms. This variant of the forward model was run with just the diatom/dinoflagellate basis vectors and with static c_i and Standard size distribution assumptions. This is justified as [Chl a] within NOMAD are relatively low (compared to the ranges used in the original forward model range), and the results above show relatively little sensitivity to these assumptions at lower biomass ranges. More conservative values of a_{gd} and b_{bs} were used (compared to FWD$_A$), where a_{gd} = 50% (N-) and 400% (N+) of $a_\phi(445)$ and $b_{bs}(550)$ = 0.005 (N-) and 0.0103 (N+)

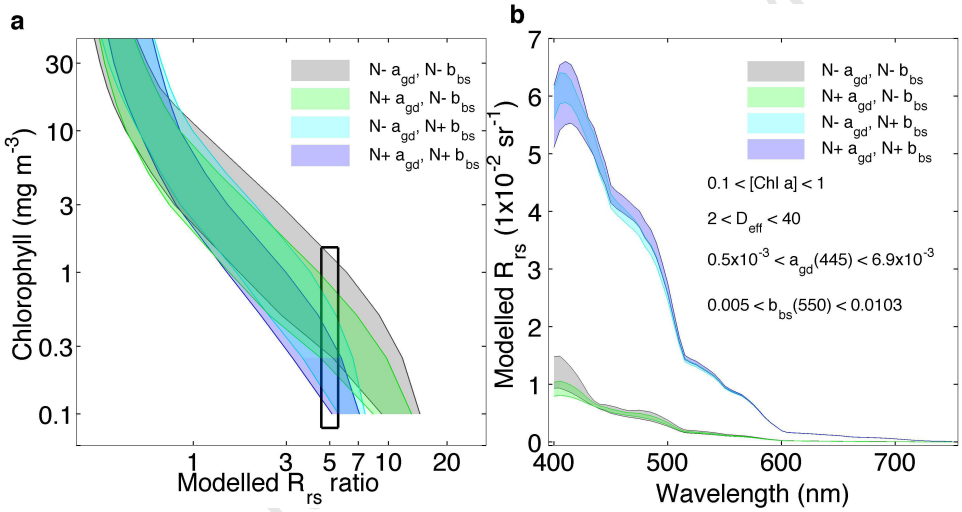

Figure 2.10: (a) OC4 maximum band reflectance ratios, versus [Chl a] from the forward model using EcoLight-S and a_{gd} and b_{bs} values to approximately simulate the range of data covered by the NOMAD dataset (b) Example spectra from samples with similar (within 10%) reflectance ratios to highlight significant ambiguity within a reasonable error that could be associated with reflectance measurements. Thickness of the curves plotted refers to the effect of cell size.

Figure 2.10 shows an approximation of the maximum band ratio (MBR) approach used in the OC4 algorithm (O'Reilly et al., 1998), using output from the FWD$_N$ model. Figure 2.10 suggests two key points. Firstly, changes in size (especially at higher biomass) could explain some of the scatter associated with these ratios, however the extent of this is subject to influence of coincident optical variability in gelbstoff/detritus and non-algal particles (Fig. 2.10(a)). Secondly, similar changes in ratio can result from changes in [Chl a], size (as parameterised here through D_{eff}), a_{gd} and b_{bs} independently or coincidently, hence the retrieval of variables as distinct from each other is highly ambiguous (Fig. 2.10(a and b)). These results support those of Sauer et al. (2012), suggesting that variability in a^*_ϕ (in the EAP case, coincident with changes in size) may be obscured by a_{gd}, particularly at lower biomass, where the majority of the size related signal occurs in the blue and MBR approaches are typically applied (Fig. 2.2). Sauer et al. (2012) indicate little influence of total particle backscattering (b_{bp}) on MBR approaches in the

context of NOMAD. The data presented here support this, and closer examination of the IOP budgets under both this and the previous forward model scenarios, suggests that phytoplankton backscattering (and the associated size influence inferred by the two-layered sphere model) only becomes a dominant contributing factor (versus scattering by water) in the absence of significant contribution from non-algal particles and at [Chl a] greater than 3-100 mg m^{-3} depending on the wavelength and size; whilst a_ϕ at all sizes can represent the dominant contribution to total absorption at [Chl a] as low as 1, at some wavelengths and when there is minimal influence from a_{gd} (as shown in figure 2.6).

2.4.3 Differences in forward model response attributable to radiative transfer technique

As can be seen from the results in figures 2.2 to 2.5, the use of the two radiative transfer techniques influences both the absolute magnitudes of reflectance generated in the forward models and the sensitivity in R_{rs} to size and to assumptions made within the two-layered sphere model. Additionally, these differences between the results of the two different radiative transfer techniques become more/less substantial, dependant upon the IOP budgets of the modelled water type. The two key differences between the REFA and ES approaches are the propagation of the light at depth to the surface (as parameterised through the K_u approximation in the REFA approach (Albert and Mobley, 2003) and K_{Lu} in EcoLight-S) and the method through which the bidirectionality of light is accounted for (where REFA uses a [Chl a] dependent f/Q parameterisation (Morel et al., 2002), whilst EcoLight-S can be used to accurately calculate this parameter as dependent on the IOPs used).

Figure 2.11 shows the spectral K_u (REFA) and K_{Lu} (ES) values for forward model runs using the diatom/dinoflagellate basis vectors with SCI c_i assumption and under the low a_{gd}/low b_{bs} (Fig. 2.11(a, b, c)), high a_{gd}/low b_{bs} (Fig. 2.11(d, e, f)) and high a_{gd}/high b_{bs} (Fig. 2.11(g, h, i)) scenarios. Under the low a_{gd}/low b_{bs} scenario (Fig. 2.11(a, b, c)), relatively minimal differences are seen between the two radiative transfer approaches. K_{Lu} values from ES are slightly higher in the green, compared to those estimated from the parameterisation of Albert and Mobley (2003) and in general higher values of K_{Lu} (ES) are seen at higher [Chl a] at all wavelengths, with the exception of the 709 nm values. Similarly, there is relatively minimal differences in K_u/K_{Lu} when the DCI basis vectors are employed, with only a slight reduction in size sensitivity in the DCI case (not shown). Although the two attenuation coefficients (K_u and K_{Lu}) are not identically derived, they are almost identical in the low a_{gd}/low b_{bs} case, with only some small differences seen at higher biomass.

At high a_{gd} but low b_{bs}, both the K_u and K_{Lu} increase in the blue and green relative to those in the previous scenario (Fig. 2.11 (d, e, f)), as could be expected due to increased absorption by gelbstoff/detritus. Whilst in the high b_{bs} case, both K_u and K_{Lu} values are substantially increased across all wavelengths, particularly at low [Chl a] in the REFA approach (Fig. 2.11(g)) and at higher biomass in the ES approach (Fig. 2.11(h and i)). The latter in particular shows

substantial spectral shape and size sensitivity at [Chl a] 10 and above. Mobley et al. (2002) looked at the effects of phase functions on ocean optics, finding that errors in K_{Lu} were highly dependent on the phase function selected, with the spectrally variable Fournier-Forand used here in the ES simulations proving the most suitable for a variety of waters (Mobley et al., 2002). The decay of L_u with was found to vary little with changes in backscatter fractions in high absorption cases due to the dominance of single scattering, whilst under high scattering, L_u decay substantially varied under different backscattering fractions (Mobley et al., 2002). These results suggest that, despite the differences in derivation, the K_u parameterisation is suitable for application in the model here whilst in more scattering waters, EcoLight-S should either be used in full, or a new parameterisation derived from a wide range of simulated EcoLight-S K_{Lu}.

The corresponding spectral f/Q ratios used for the REFA approach (Fig. 2.12(a)) and calculated from EcoLight-S according to equation 2.10 are shown in figure 2.12 (b-h) for forward model runs using the diatom/dinoflagellate basis vectors with static c_i. The f/Q values used in the Morel et al. (2002) parameterisation in the REFA approach show no size related variation in the forward model data, due to the parameterisation being exclusively [Chl a] dependent in this case. In EcoLight-S, greater spectral shape is introduced, with f/Q values in the green and red showing substantially lower values than those in the Morel et al. (2002) formulation. However at low biomass (i.e. below 10 mg m^{-3}), very little size related variability is observed in f/Q values When the IOP budget becomes dominated by phytoplankton (see figures 2.6 to 2.9) at higher [Chl a], size related sensitivity at green and red wavelengths becomes more substantial. The assumption of constant f/Q above 10 mg m^{-3} made in the REFA approach becomes increasingly limited as the size related effects on absorption and backscattering by phytoplankton become significant and affect the f/Q ratio derived from EcoLight-S (Fig. 2.12(b)). The inclusion of these effects likely accounts for the elevated R_{rs} in the red generated with EcoLight-S under these conditions (figure 2.2. Lower levels of size related sensitivity, are seen in the high a_{gd} (Fig. 2.12(c)) and especially the high b_{bs} (Fig. 2.12(d)) cases, as in these cases the IOP budget is dominated by high a_{gd} (Fig. 2.7) and high b_{bs} (Fig. 2.9) respectively. In both these Case 2 type scenarios, the Morel et al. (2002) f/Q parameterisation becomes inadequate compared to those generated as IOP dependent by EcoLight-S, as expected, given that this parameterisation was developed for Case 1 waters where IOPs are largely [Chl a] dependent. Again, the differences here between the REFA K_u and the ES K_{Lu} likely accounts for a significant portion of the differences in R_{rs} between the two methods (figures 2.2 to 2.4).

The highest and most variable f/Q values are associated with the cases where biomass is high, with the exception of the highly scattering case (Fig. 2.12(d)). This is consistent with observations in other studies, which found high values in bloom conditions and sensitivity to the slope of the PSD (Chang et al., 2007b; Kostadinov et al., 2009, 2010). However, in contradiction to the results above, these studies also suggested that highest f/Q values were associated with dominance of large particles, which is not the case in the idealised modelling here, except in the highly scattering case at blue/green wavelengths. Thus, these observations may reflect the

tendency for blooms to be dominated by larger particles, rather than reflecting a specific size related effect, and/or that there is significant non-algal particle scattering occurring coincidentally (*ibid*).

The two-layered sphere model imparts significant shape in the spectral backscattering probability associated with phytoplankton (Bernard et al., 2009; Robertson Lain et al., 2014). The f/Q values computed from the EcoLight-S forward model variant, show significant spectral shape (figure 2.12(b)), beyond those typically modelled in the literature and including those simulated by (Morel et al., 2002) and used in the REFA approach (figure 2.12(a)), even at low biomass. This is a result of the inclusion of phytoplankton [Chl *a*] and size dependent basis vectors in the EcoLight-S derived f/Q estimates. The volume scattering function (VSF) (and variability in it) have been shown to significantly impact bidirectionality (Chang et al., 2007a; Morel et al., 2002), hence the development of parameterisations such as that by Morel et al. (2002), which incorporate a [Chl *a*] and wavelength dependent particulate phase function. Although the use of size dependent b/b_b from the two-layered sphere model goes some way to incorporating variability in backscattering, and the spectrally variable Fournier-Forand phase function employed is able to respond to changes in backscattering fraction, there is still a need for discretisation of phase functions relative to the phytoplankton IOPs generated from the two-layered sphere model (Robertson Lain et al., 2014). The sensitivity of the f/Q ratio and final R_{rs} solutions to phase function assumptions could then be established. This may be particularly important in these high biomass cases where the phytoplankton related backscattering is highly variable and already shows considerable influence on f/Q and where single scattering is not dominant (i.e. in highly scattering waters) (Mobley et al., 2002). An appropriate phase function (and effects of this through f/Q and K_{Lu}), has been shown to result in an order of magnitude less error in L_u just below the surface in other studies (*ibid*). It should be noted that the f/Q analysis conducted here is preliminary, and further understanding and development of new parameterisations would require additional consideration of other factors which influence f/Q e.g. solar zenith angle, viewing geometry and wind speed (Morel et al., 2002).

2.4.4 Potential consequences for application of the Equivalent Algal Population model in an inversion algorithm

Whilst there are additional errors to consider when applying this model for retrieval of biogeochemical parameters to *in situ* or satellite ocean colour data (e.g. instrument error, patchiness, atmospheric correction and inversion methodology (see chapter three and four)), the forward model sensitivity can be used to formulate tentative hypotheses about likely limitations. The analysis above indicates optical water types where this model (and indeed any others with similar assumptions) may struggle to detect or differentiate between biogeochemical parameters and changes in them. This "ambiguity" has been discussed as a major issue relating to ocean colour inversion (Defoin-Platel and Chami, 2007) and manifests as multiple parameter combinations which result in similar reflectance spectra. Figure 2.4 exemplifies this situation, where, at several

wavelengths, there is no discernible differences between R_{rs} at different [Chl a] and/or different sizes, until biomass levels reach a point that absorption and backscattering by phytoplankton outweigh the absorption and backscattering associated with the high a_{gd} and high b_{bs}. Although this is an extreme condition, a similar situation is seen using the simulated NOMAD data, where a_{gd} and b_{bs} vary within ranges commonly found throughout the global oceans (figure 2.10).

It is also clear that there is an abundance of spectrally variant effects which need to be accounted for when attempting to fully capture variability associated with co-incident changes in ocean constituents, and in particular, phytoplankton characteristics. This spectral sensitivity insinuates that there are likely to be limitations when applying this approach to satellite data, especially when bands are not available within close spectral proximity of important features e.g. at sufficient resolution to resolve the bifurcated peaks associated with the presence of cryptophytes such as in *Mesodinium rubrum* dominated blooms. Similarly, the results from the FWD_N suggest that the use of band ratio approaches can result in significant ambiguity, in which size can play a role in obscuring [Chl a] estimations. This is consistent with the findings of Sauer et al. (2012). Whilst the model presented here allows characterisation of size related and some pigment related effects on optical properties, other factors may need to be considered. Particularly in the case of monospecific blooms, species with unique intracellular structure or particularly distinct pigment suites, could have substantial impacts on algal IOPs and ultimately on R_{rs} (e.g. Matthews and Bernard (2013)). Additionally, the optical properties of living algae and the extent of their degradation products are known to be influenced by algal physiology, life stage, and response to growth conditions, including the light environment (see Dierssen et al. (2006) and references within). Incorporating all these effects into an unconstrained model may be computationally intensive and ultimately unnecessary if sensitivity is not substantial. Flags based on specific optical signals or auxiliary data could be used to indicate when specific basis vectors should be used (e.g. in the case of cryptophytes/cryptophyte endosymbionts), or when the optical conditions are beyond those accounted for in the forward model.

Finally, the choice of radiative transfer technique appears to have a range of impacts on the final R_{rs} values and sensitivity to size/functional type. The extent of this impact is dependent on the underlying IOP budgets associated with the modelled water types. Differences between the two approaches, largely as a result of their different f/Q and K_{Lu} parameterisations, are minimal under low a_{gd}, low b_{bs} and higher levels of biomass. Whilst it would be optimal to always choose the most accurate radiative transfer methods, the computational efficiency of an approach needs to be considered. In the forward model case, run time using EcoLight-S is approximately 300 times longer per sample. Whilst for the relatively small number of samples here (288), this is not a limiting factor, when moving towards an inversion method, which may need to conduct a large (+1000) number of iterations of this model per sample/over a large number of satellite image pixels, run time could increase substantially. Parallel and high performance computing methods may be able to address this problem, however within some of the cases seen here, the extra computational expense may not be necessary, given the relatively

small differences between the radiative transfer solutions. The increased accuracy gained from using a more advanced radiative transfer method will need to be balanced against the accuracy of parameter retrieval in an inversion method. Alternatively, the results here could be expanded using EcoLight-S (or full Hydrolight) to develop more appropriate parameterisations of f/Q in particular, under various viewing geometries, wind speed and IOP budgets.

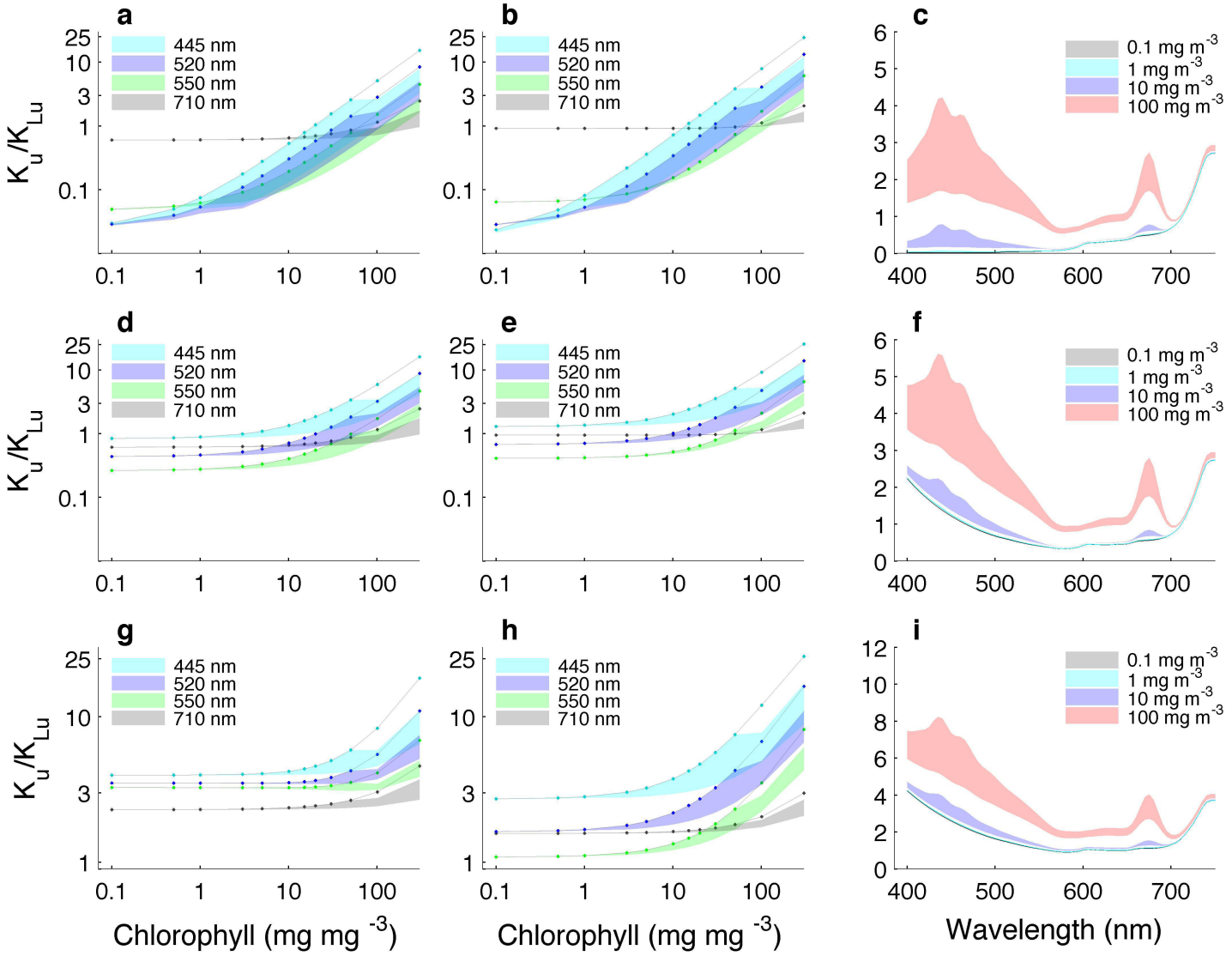

Figure 2.11: K_u/K_{Lu} values from the REFA approach using the K_u parameterisation of Albert and Mobley (2003) (a, d, g) and K_{Lu} from EcoLight-S (b, e, h), for the low a_{gd}/low b_{bs} scenario (a, b, c), high a_{gd}/low b_{bs} (d, e, f) and high a_{gd}/high b_{bs} (g, h, i). Dots indicate K_u/K_{Lu} associated with the smallest cells. Panels c, f and i show example ranges of spectral K_u/K_{Lu} at selected [Chl a] across the modelled size range using ES. Ecologically unlikely parameter combinations are excluded, see section 2.4.2 for discussion.

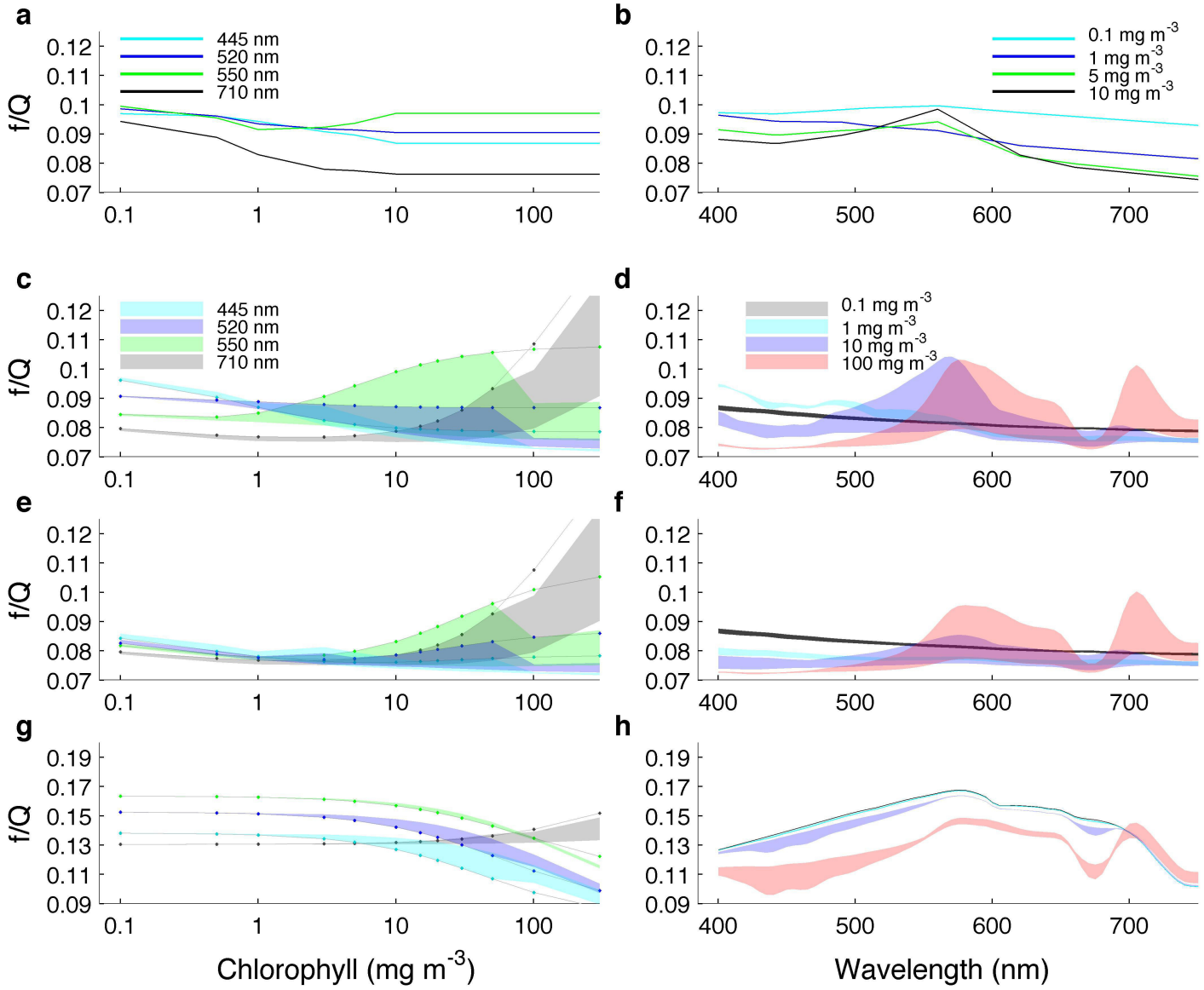

Figure 2.12: Ranges and spectral f/Q values across the simulated dataset using the diatom/dinoflagellate basis vectors. The REFA approach uses the f/Q parameterisation of Morel et al. (2002) (a, b). These values will be constant across the different a_{gd} and b_{bs} scenarios and with size due to the solely [Chl a] dependent parameterisation. The rest of the panels use the EcoLight-S simulated data (with SCI assumption) under (c, d) low a_{gd} and low b_{bs}, (e, f) high a_{gd} and low b_{bs} and (g, h) high a_{gd} and high b_{bs}. Dots indicate f/Q associated with the smallest cells. Ecologically unlikely parameter combinations are excluded, see section 2.4.2 for discussion

2.5 Conclusions and recommendations

The EAP model has substantial utility for investigating the effects of phytoplankton cell size on optical signals. Several conclusions can be made from the preliminary sensitivity analysis presented here:

- Size related sensitivity in R_{rs} is highly dependent on algal biomass, as determined by the relative algal contribution to the IOP budget.
- When the IOP budget of a water type is dominated by other constituents, this size related variability is reduced, particularly at low biomass.
- Radiative transfer techniques able to quantitatively account for phytoplankton spectral scattering and bidirectionality are needed to propagate assemblage related effects on IOPs through to reflectance, especially in optically complex waters.
- Results here suggest that the occurrence of high biomass blooms, combined with minimal gelbstoff/detritus and non-algal particulate input in the southern Benguela may make it an ideal location to exploit sensitivity in optical signals related to size and other characteristics of phytoplankton (e.g. pigments such as those associated with cryptophytes).

The results of this study indicate that whilst there is likely significant phytoplankton size related sensitivity in R_{rs}, the extent of this sensitivity is dependent on several things. Firstly, the accuracy of the underlying IOP model (in this case the two layered sphere model) and its assumptions for accurately representing the absorption and backscattering of diverse phytoplankton communities. Secondly, the accuracy of the radiative transfer method or any parameterisations employed and finally, the underlying IOP budget/the physical part of the ocean colour problem, which can lead to high levels of non-uniqueness amongst parameters. For application to *in situ* and satellite derived ocean colour data, exploiting these sensitivities will rely on appropriately resolved data (hyperspectral or multispectral), sufficient sensitivity in the context of the underlying ambiguity and how these two factors interact with the spectral shape and magnitude of the parameters in the additive model. To develop this model for application as an inversion algorithm, and to further investigate the role of ambiguity in parameter estimate errors, it is recommended that inversion methods be tested using the forward model datasets produced here (or similar) to understand ambiguity in size related signals before application of the model to *in situ* data.

Chapter 3

Impact of inversion mathematics on retrievals from a semi-analytical algorithm

Parts of this chapter formed content included in the paper Evers-King et al. (2014). The inverse modelling and discussion of ambiguity within this paper were conducted by the lead author and have been expanded for inclusion in this thesis chapter.

3.1 Introduction

3.1.1 Semi-analytical inversion algorithms

Semi-analytical algorithms use radiative transfer theory and relationships between radiance/reflectance and inherent optical properties (IOPs). These relationships can be inverted to derive IOPs and biogeochemical parameters with the inclusion of empirical relationships and assumptions (see chapter one (Fig. 1.3) and chapter two for discussion of forward modelling techniques). Many semi-analytical bio-optical algorithms are designed around a similar framework and the major differences arise from the parameterisations and assumptions used to represent IOPs and light field shape factors (see discussion and example of the Generalised Inherent Optical Property (GIOP) framework in Franz and Werdell (2010)). The basic structure of most semi-analytical bio-optical algorithms centres around the reflectance approximation (equation 3.1), where R_{rs} is remote sensing reflectance, $L_u(0^+, \lambda)$ is the wavelength dependent upwelling radiance just above the sea surface, $E_d(0^+, \lambda)$ is the wavelength dependent downwelling irradiance just above the sea surface, f/Q describes the angular structure/bidirectionality of the light field (see Morel et al. (2002) for discussion), $\eta^2\tau$ represents the transmission parameters associated with the air/sea interface, $a(\lambda)$ is the wavelength dependent total absorption coefficient and $b_b(\lambda)$ is the wavelength dependent total backscattering coefficient. However, it is becoming more common to use full radiometric modelling with programmes that solve either the full or

an azimuthally integrated version of the radiative transfer equation (RTE) (Mobley, 2011).

$$R_{rs} = \frac{L_u(0^+, \lambda)}{E_d(0^+, \lambda)} = \frac{f}{Q}\eta^2\tau\frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \tag{3.1}$$

Semi-analytical algorithms require the characterisation of the spectral IOPs associated with the various constituents which affect the ocean colour (Robinson, 2010). Differences in semi-analytical algorithms can arise from their choice of Inherent Optical Properties. Specific inherent optical properties (SIOPs) frequently form the basis for calculation of full IOPs (i.e. spectral absorption and backscattering) and can be concentration specific in relation to different components (e.g. phytoplankton/sediments/dissolved matter) (Lee et al., 2002; Smyth et al., 2006; Tilstone et al., 2012). SIOPs vary regionally and are often selected based on regional observations, which can improve regional algorithm performance, but may limit wide scale application (Ambarwulan et al., 2010; D'Sa et al., 2002; Franz and Werdell, 2010; Kahru and Mitchell, 2001; Reynolds et al., 2001; Tilstone et al., 2012; Tzortziou et al., 2006; Zibordi et al., 2011). Whilst large datasets exist for the purpose of algorithm development (e.g. the NASA bio-Optical Marine Algorithm Development (NOMAD) dataset (Werdell and Bailey, 2005)), these datasets may fail to provide enough coverage to characterise the extremes found in coastal regions and the extensive *in situ* sampling campaigns needed to adequately capture this variability can be costly and time consuming. Alternatively, theoretical models can be used to generate the IOPs, allowing investigation of the optical characteristics of different components from a first principles approach (e.g. using particle models to characterise complex phytoplankton assemblages, as in chapter two) (Bernard et al., 2009; Robertson Lain et al., 2014). However ultimately these models require a degree of *in situ* validation to quantify their suitability. Use of semi-analytical algorithms allows the composition of ocean waters to be broken down into any number of potentially optically significant components. However, due consideration must be given to the sensitivity of the available signal to different constituents and the possibility of non-uniqueness amongst R_{rs} spectra for different IOP combinations (Defoin-Platel and Chami, 2007).

In addition to an understanding of radiative transfer, closure of the forward and inverse problem of ocean optics requires characterisation (either through models or *in situ* measurements) of absorption and backscattering. Absorption has been relatively well measured and modelled over a range of biogeochemically variable regions and most models are robust to assumptions which simplify the complexity of the phytoplankton community, at least in the Case 1 waters of the global ocean (see chapter two, section one for discussion and examples in Ciotti et al. (2002); Morel and Bricaud (1986); Sathyendranath et al. (1987); Uitz et al. (2008)). Backscattering however proves difficult to measure and where measurements and comparable models exist, many gaps in understanding still remain (Chami et al., 2006; Vaillancourt and Brown, 2004; Volten et al., 1998). When comparing differences in studies, several key factors emerge which may explain these discontinuities. These factors arise mostly from the assumptions brought with the models which work well for modelling phytoplankton absorption (i.e. Mie modelling Bricaud and Morel

(1983); Morel and Bricaud (1986); Stramski et al. (2001)) but show conflicting findings about the impacts of size (Ahn et al., 1992; Stramski and Kiefer, 1991; Svensen et al., 2007; Zhou et al., 2012), cellular chlorophyll (Ahn et al., 1992) and carbon content (Vaillancourt and Brown, 2004; Zhou et al., 2012), cell shape (Quinby-Hunt et al., 1989) and cellular heterogeneity arising from complex intracellular structures with differing refractive indices (Matthews and Bernard, 2013; Stramski et al., 2004; Svensen et al., 2007; Vaillancourt and Brown, 2004; Volten et al., 1998) on scattering.

3.1.2 Errors in semi-analytical algorithms

Errors and uncertainties with regards to *in situ* and satellite derived ocean colour data and products can come from a number of sources. Instrument error (frequently minimised through calibration (Pinkerton and Aiken, 1999)) and atmospheric correction (Antoine and Morel, 1999; IOCCG, 2010; Zhu et al., 2012) are two substantial error sources (see further discussion of errors associated with *in situ* and satellite radiometry in chapters four and five respectively). Calibration, atmospheric correction and/or correction for other sources of instrument error are most frequently dealt with before the application of semi-analytical algorithms. Additional errors in resultant products can be introduced from semi-analytical algorithms, including the following main sources, which are discussed further below:

- Use of underlying Inherent Optical Properties which are not representative of the range of optical variability associated with different constituents.
- Loss of signal information resulting from use of multispectral versus hyperspectral radiometric data (Aurin and Dierssen, 2012).
- Inappropriate simplifications and assumptions with regards to radiative transfer i.e. with regards to the shape of the light field, bidirectionality etc (Aurin and Dierssen, 2012; Chami and Defoin-Platel, 2007; Mobley et al., 2002).
- Ambiguity, where multiple combinations of IOPs can result in indistinguishable reflectance spectra and ultimately in inaccurate retrieval of biogeochemical parameters (Defoin-Platel and Chami, 2007).
- The choice of inversion mathematics which can find local rather than globally optimal solutions (as in the case of non-linear optimisation) or be over constrained by limited data (as in the case of trained Artificial Neural Networks (ANN)) (Defoin-Platel and Chami, 2007; Loyola and Diego, 2006; Robinson, 2010).

Through necessity, semi-analytical algorithms must make assumptions about the relationships between the biogeochemical parameters of interest and the IOPs of ocean waters. Errors introduced by semi-analytical algorithms can result from inappropriate parameterisation of these relationships. Additive models require groupings of these biogeochemical components and the

suitability of these groups will often play a significant role in determining the accuracy of retrievals when a generic set of IOPs is specified for each group. As mentioned above, the suitability of these groupings may vary per region, depending on what the causes of the underlying variability are. For example, in coastal regions where sources of particulate matter are often prominent, there may be justification (and need) for having multiple particulate components, each represented with a different set of IOPs. Similarly, a range of IOPs representing distinguishable phytoplankton functional groups would seem necessary for an algorithm seeking to detect this variability.

Conversely, to design an algorithm approach which distinguishes different components, it must be ensured that the IOPs of each component are sufficiently different and resultant effects on reflectance sufficiently sensitive, that the inverse problem does not become too ambiguous (Defoin-Platel and Chami, 2007). Whilst empirical datasets can help to establish the most realistic groupings and associated IOPs for an inversion approach, there are limitations to this. To sufficiently parameterise IOPs based on observational data requires large amounts of high quality data, in order for it to be assumed that a significant amount of the variability associated with the IOPs has been captured. In practice this can be very difficult due to time and financial constraints and the inherent errors associated with many data collection methods. Whilst retrieving accurate results in comparison to *in situ* data is the ultimate goal for an algorithm product, it can be difficult to establish sources and scales of error when errors associated with measurement methodology cannot be separated from those introduced through the inversion approach.

The use of simulated data provides several benefits to help the algorithm developer. Simulated datasets can be generated from known biogeochemical parameters and associated IOPs. In this way, the impact of inversion choice and IOP parameterisation scheme can be isolated, varied, sensitivity established and approximate error estimates made. To make conclusions based on reality, a simulated dataset should be informed by ranges of biogeochemical parameters and IOPs observed *in situ*, in addition to any theoretical knowledge of IOP covariance (Defoin-Platel and Chami, 2007; IOCCG, 2006). These datasets can then be used to expand knowledge of potential variability in systems where measurements are sparse.

The ocean colour problem is well-posed in that a solution can always be found through the radiative transfer equation and weak variations in IOPs typically result in weak variations in calculated R_{rs} spectra (Defoin-Platel and Chami, 2007). However, it is inherently ambiguous as multiple combinations of IOPs can produce the same reflectance spectra (*ibid*). The selection of an inversion method can enhance or help reduce this ambiguity. Common approaches have included non-linear optimisation (Roesler and Perry, 1995), genetic programming algorithms (Chami and Robilliard, 2002) and Artificial Neural Network approaches (Defoin-Platel and Chami, 2007; Loyola and Diego, 2006; Raitsos et al., 2008; Schiller and Doerffer, 2005)). Artificial Neural Networks have been particularly successful in coastal waters (Schiller and Doerffer, 2005), however they require a substantial amount of training data to set up, and

can fail in cases where spectra and constituent concentrations fall outside of the range of this training data (Robinson, 2010). Non-linear inversion methods offer a less constrained approach, however there can be difficulties when methods find local, as opposed to global optima; a result of the ambiguity inherent in the ocean colour problem and the types of mathematical solvers used. Predefinition of a suitable range, or probability density function for reasonable solutions can be included in inversion methodology to improve the likelihood of reaching a globally optimum solution (Defoin-Platel and Chami, 2007). However, these approaches can require substantial *in situ* datasets and/or well constrained simulated datasets.

In addition to choice of inversion technique, the incorporation of ancillary data to inversion methods has been suggested to improve retrievals (Defoin-Platel and Chami, 2007)). In these cases, ancillary data can be used to predefine the regions in parameter space where likely and unlikely solutions are expected to exist. Similarly, increased spectral resolution of radiometric instruments or regionally and mechanistically appropriate selection of bands, can provide information that makes inversion less ambiguous. Several studies have supported the use of hyperspectral data for inversion approaches in coastal waters from *in situ* (Rehm and Mobley, 2013; Ruddick et al., 2001), airborne (Dierssen et al., 2003; Lee et al., 2001) and satellite mounted sensors (Brando and Dekker, 2003). The availability of data at higher spectral resolution may allow for further use of techniques developed to detect specific spectral features. For example, decomposition and derivative analysis have been used with *in vivo* absorption spectra to investigate the presence of particular species and accessory pigments (Bidigare et al., 1989; Hoepffner and Sathyendranath, 1993; Millie and Moline, 2000). Attempts have been made to extend these techniques to remotely sensed radiometric data and derived products for the detection of phytoplankton species, suspended and bottom sediments (Chen et al., 1992; Craig et al., 2006; Hochberg et al., 2003; IOCCG, 2014; Louchard et al., 2002; Roelke et al., 1999). It should be noted that the forward data simulated in chapter two and inverted in this chapter, is hyperspectral at 5 nm resolution. Discussion of R_{rs} sensitivity and inversion errors when the EAP forward and inverse approaches are applied to multispectral resolution data (i.e. from the MERIS instrument) can be found in chapter five.

3.1.3 Summary and chapter aims

This chapter investigates the sensitivity of the Equivalent Algal Population (EAP) model, used as a semi-analytical inversion algorithm, to choice of inversion mathematics. The simulated datasets from chapter two are inverted and returned biogeochemical parameter values compared to those used to generate the dataset. Several inversion methods are employed and the sensitivity of various assumptions within these methods evaluated. The error associated with inversion choice can then be estimated and the variability introduced by different IOP and radiative transfer assumptions can be explored. Errors are evaluated in the context of different water types with different IOPs with a particular focus on application to high biomass blooms in the southern Benguela system.

Table 3.1: Summary of parameter ranges used in initial forward model data simulation

Parameter	*Range*
Chlorophyll	0.1, 0.5, 1, 3, 5, 10, 15, 20, 30 50, 100, 300 mg m^{-3}
Effective diameter	2, 5, 10, 20, 30, 40 μm
Combined absorption of gelbstoff + detritus (a_{gd}) at 400nm in m^{-1}	Low (0.02) and High(2)
Non-algal particle backscattering (b_{bs}) at 550nm in m^{-1}	Low(0.0005) and High (0.5)
Diatom/dinoflagellate vs cryptophyte basis vectors	0 and 100 %

3.2 Methods

3.2.1 Forward model and simulated data

The simulated dataset (FWD_A) from chapter two is used here as the input data for a number of experiments designed to test the sensitivity of different inversion schemes to the various IOP and radiative transfer assumptions used. The forward model used to generate FWD_A is detailed in chapter two, however a summary of the parameter ranges covered are shown in table 3.1.

Output L_u(z) and R_{rs} are calculated from the forward model for all combinations of the parameters detailed in table 3.1 with each of the assumptions listed below systematically varied. As discussed in chapter two, the L_u is calculated at a z of -0.66 m such that results can be compared to *in situ* data in chapter four.

1. Different c_i parameterisations where the SCI set uses a static c_i per group and the DCI set uses an effective diameter (D_{eff}) dependent parameterisation based on Zhang et al. (2011).

2. The use of a Jungian distribution to characterise the size dependent phytoplankton IOPs at low [Chl a] as opposed to the Standard size distribution. These IOPs were used to create [Chl a] dependent admixtures of phytoplankton absorption and backscattering between 0 and 10 mg m^{-1}. These results are henceforth referred to as the JCI set.

3. Different radiative transfer methods including a reflectance approximation (REFA) and EcoLight-S (ES).

The EAP forward model used in the inversion process is identical in construct to that described fully in chapter two. In each case, the assumptions used to generate the forward model simulated data, are used in the inverse approach i.e. to invert the SCI data, IOPs with an SCI assumption are used. The exception to this is the use of the basis vectors from the two functional type groups. In this case, both constrained (i.e. only diatom/dinoflagellate or cryptophyte basis vectors for the case where the simulated data uses these) and unconstrained (i.e. both sets of

IOPs were made available to the inversion method) basis vectors were used. The impact of using a reflectance approximation (REFA) versus an EcoLight-S (ES) approach to deal with radiative transfer considerations is also assessed. However, due to the additional computational expense required to run EcoLight-S in the iterative way required for the inversion techniques, the ES technique is only applied to a selection of the inversion experiments.

3.2.2 Inversion techniques

Beyond the selection of specific inherent optical properties for the components of this additive model, the inversion methodology used to solve for these from measured radiance (L_u(z)) or reflectance (R_{rs}) is a major consideration. Non-linear optimisation methods seek to resolve a parameterised function by varying unknowns to minimise a modelled solution towards a given, measured solution. In this instance, the simulated upwelling radiance (L_u (z), μW cm^{-2} nm^{-1} sr^{-1}) is the assumed solution to the parameterisations set in the forward model, given a set of four or five unknowns: Chlorophyll-a concentration ([Chl a], mg m^{-3}), effective diameter (D_{eff}, μm), combined absorption of gelbstoff and detritus (a_{gd}, m^{-1}), backscattering of non-algal particles (b_{bs}, m^{-1}) and an optional percentage admixture between two different functional groups of phytoplankton (represented by different absorption and backscattering basis vectors). For this study, sensitivity assessment of the inversion was conducted using two versions of the EAP algorithm with basis vectors available to the inversion process from the diatom/dinoflagellate or cryptophyte groups with an *a priori* dependence on those used in the simulated data (1v) or both dinoflagellate/diatom and cryptophyte groups available for selection (2v).

A number of inversion techniques were tested to solve for the 4/5 parameters described (Table 3.2).

Table 3.2: Summary of inversion techniques used

Inversion method	*Reference*	*Main characteristics*
Nelder-mead simplex	Nelder and Mead (1965) Matlab R12, The Mathworks	Initial conditions (3 variants) Parameter variation windows given
Nelder-mead fminsearch	Nelder and Mead (1965) Matlab R12, The Mathworks	Initial conditions (3 variants) Unconstrained parameter variation
Levenberg-Marquardt	Levenberg (1944); Marquardt (1963) Matlab R12, The Mathworks	Initial conditions (3 variants)
Evolutionary algorithm	Based upon the optimisation technique described in Rajpaul (2012)	Initial conditions (3 variants)

Where appropriate (i.e. offered within the mathematical framework of the inversion technique), the impact of different initial conditions on the inversion result was tested. Proximity of initial conditions to the global optimal solution for an inversion, and setting of appropriate parameter variation constraints, can increase the likelihood of finding this solution (see discussion of this in Rehm and Mobley (2013)). Testing of a number of initial conditions also allows for an assessment of the sensitivity and general robustness of the solution derived through the inversion technique to the choice of initial conditions. Firstly, a baseline set of static initial conditions were used. These values for the four or five iterated parameters represent reasonable averages for the range of Benguela waters (both offshore, Case 1 waters and inshore blooms). For the standard set, initial conditions (p_{init}) were 10, 5, 0.02, 0.005 for respectively: Chlorophyll a concentration ([Chl *a*], mg m^{-3}), effective diameter (D_{eff}, μ m), absorption of combined gelbstoff and detritus ($a_{gd}(400)$, m^{-1}) and non algal particle backscattering ($b_{bs}(550)$, m^{-1}). In the 2v variant an additional initial condition of 50% was used for the proportion of the functional groups ($PFT_{\%}$) where 0 and 100 mean that entirely diatom/dinoflagellate and cryptophyte basis vectors are selected respectively, the intermediate values formulate a spectral admixture. Secondly, a simple grid search technique was employed to search a reasonable parameter space for a closest match to each spectra to use as a start point for the inversion technique. A 4x4 (1v)/5x5 (2v) grid was set up using five logarithmically distributed values of [Chl *a*] between 0.1 and 100, and five linearly distributed values of the following ranges of the other four/five variables: D_{eff} = 2 μm to 50 μm, a_{gd} (400) = 0.2:2 m^{-1}, $b_{bs}(550)$ = 0.0005: 0.5 m $^{-1}$ and $PFT_{\%}$ = 0 (all diatoms/dinoflagellate basis vectors used): 100 (all cryptophyte basis vectors used). Finally, as the main source of optical variability in phytoplankton dominated waters, the use of case specific initial estimates for [Chl *a*] was tested. To this end, an empirical estimate of [Chl *a*] was used as an initial condition based on a red band ratio (equation 3.2).

$$[Chl\ a] = 237.5.exp(-2.13.(L_u(z, 665nm)/L_u(z, 709nm)) \tag{3.2}$$

Where L_u is the simulated upwelling radiance at 0.66m depth (μ W cm^{-2} nm^{-1} sr^{-1}), at 665 and 709 nm wavelengths respectively. This equation is based on a 665/709 nm ratio commonly used for estimation of high biomass [Chl *a*] in the Benguela (Bernard et al., 2005). Similar relationships have also been used in other regions (Gilerson et al., 2010; Gons, 2002, 2005; Ruddick et al., 2001).

The simulated radiance (L_u(z)) calculated from the forward model and the seasonally averaged upwelling irradiance (E_d) were used as input to the various variants of the inversion approach described above to obtain modelled radiance (L_u(z)) and remote sensing reflectance (R_{rs}), the four/five iterated parameters ([Chl *a*], D_{eff}, $a_{gd}(400)$, $b_{bs}(550)$ and $PFT_{\%}$), spectral absorption and backscattering of phytoplankton (calculated from the selected [Chl *a*], D_{eff} and $sp_{\%}$ parameters), spectral a_{gd} (from the a_{gd} scale factor) and non algal particulate backscattering (from the b_{bs} scale factor). Where appropriate, the maximum number of iterations and function tolerance were set to 1000 iterations and 1e-6 respectively.

Figure 3.1 presents a simplified schematic of this inversion approach, henceforth referred to as the Equivalent Algal Population (EAP) inversion algorithm.

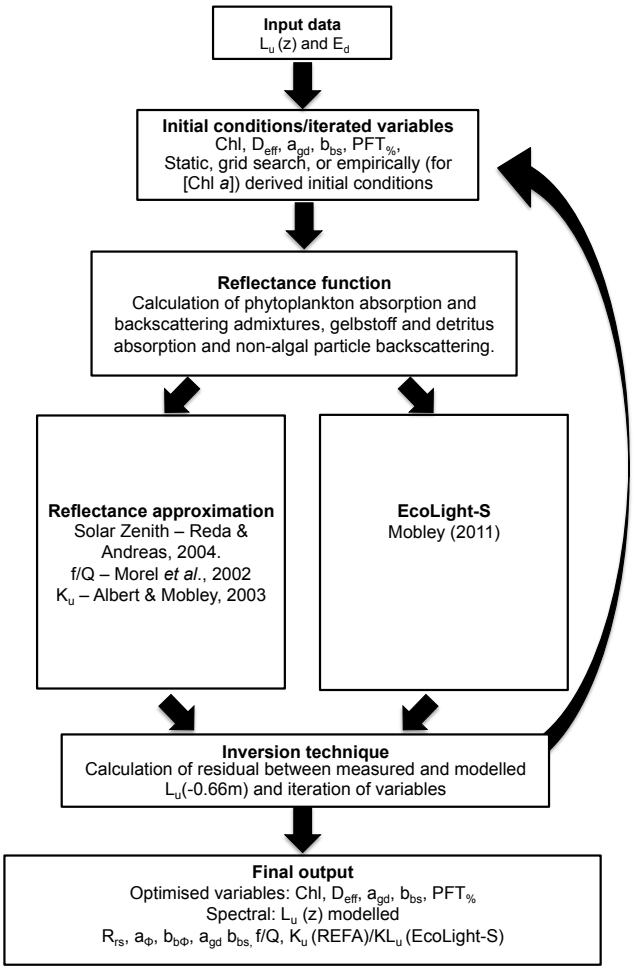

Figure 3.1: Schematic of the Equivalent Algal Population algorithm. Input in this case consists of $L_u(-0.66m)$ spectra generated using the forward model and seasonally averaged E_d. Inversion techniques are as detailed in table 3.2. The objective function to be minimised is the reflectance function for $L_u(-0.66m)$.

3.3 Results

3.3.1 Convergence of modelled solutions to input radiance

The inversion methods employed seek to match a modelled L_u(z) spectrum with that given as input (in this case from the simulated datasets in chapter two) through iterative variation of the four/five unknowns in the forward model. It thus follows that, to expect reasonable estimates of these parameters, the inversion method should closely converge upon the input solution.

Examples of this convergence spectrally across the simulated dataset are shown in figure 3.2(a, b, c), where the r^2 value is plotted for each wavelength across the dataset and colour of each point represents the corresponding root mean squared error (RMSE) in units of μW cm^{-2} nm^{-1} sr^{-1}.

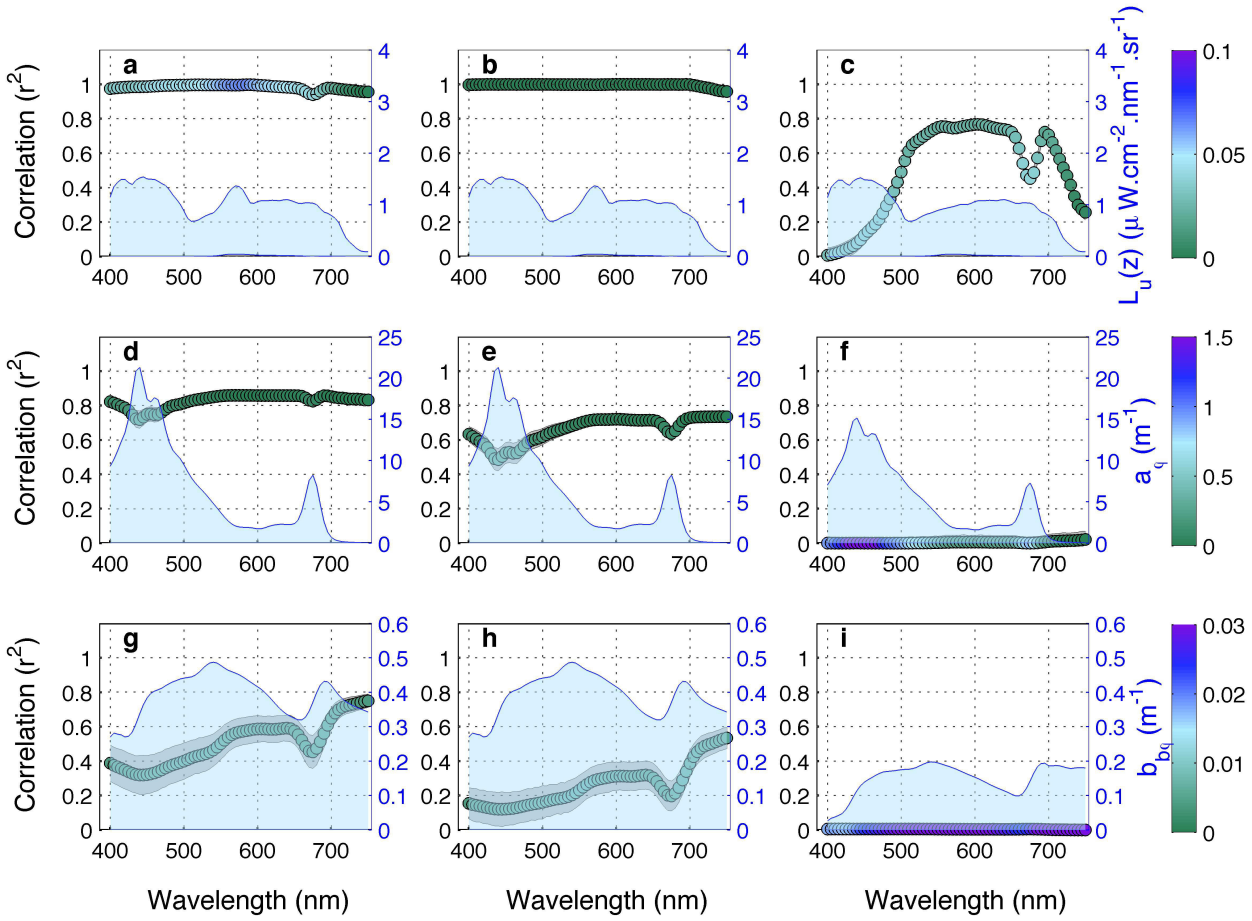

Figure 3.2: Spectral correlation for L_u (z) (a, b, c), a_ϕ (d, e, f) and $b_{b\phi}$ (g, h, i) estimated by the inversion process versus that input from the simulated dataset (n = 288, $p < 0.01$ in all cases). Dot fill colour represents the Root Mean Squared Error between the modelled output and simulated input, whilst the grey fill indicates standard error. The blue fill shows the range of each parameter associated with the simulated dataset. These figures are examples from inversion experiments using the 1v variant with simplex, SCI diatom/dinoflagellate basis vectors and static initial conditions (a, d, g), 1v variant with Levenberg-Marquardt, SCI diatom/dinoflagellate basis vectors and grid searched initial conditions (b, e, h) and Evolutionary Algorithm with DCI diatom/dinoflagellate basis vectors and initial conditions with an empirically derived estimate of [Chl a] (c, f, i.

Figure 3.2(a) shows an example where the inversion approach converges well across the dataset (indicated by the high r^2 across the spectral range). Overall, the simplex technique performs well, showing high levels of correlation and little spectral variation in performance across the various IOP model variants/initial condition combinations (e.g. Fig. 3.2(a)). The fminsearch technique, which is also based on a Nelder-Mead optimisation, performs well, showing high levels of correlation although there is slightly greater spectral variation in performance (not shown). The Levenberg-Marquardt technique shows varying levels of performance, where under

some conditions, the modelled solution converges well (e.g Fig. 3.2(b)). Lowest convergence is seen when using the Evolutionary algorithm (e.g. Fig. 3.2(c)), which shows much the same response over the various IOP model/initial condition experiments. In general, when lower levels of convergence are seen, these occur in the blue and red spectral regions, whilst highest levels (in terms of r^2) are seen in the green. There does not seem to be a consistent response across the inversion techniques, using the different basis vectors, although in several individual instances the use of JCI basis vectors does appear to produce better convergence (not shown). Similarly, there does not seem to be significantly different convergence between the experiments where the diatom/dinoflagellate and cryptophyte basis vectors are used, other than a small amount more spectral variation in the correlation coefficients in the cryptophyte cases. Additionally, in the 2v cases (where the inversion technique has an additional parameter allowing choice of admixture between the diatom/dinoflagellate and cryptophyte IOPs), substantial differences are not apparent, although retrievals appear broadly similar to those observed in the solely diatom/dinoflagellate cases.

3.3.2 Inherent optical property retrieval

Under a number of the experiments, suitably proximate solutions for modelled L_u(z) are converged upon across the simulated dataset. However, considering that the inverse ocean colour problem is highly ambiguous, this could have occurred without the selection of IOPs which are analogous to those used to create the input simulated data. To investigate this, the IOPs used in the simulated dataset are compared to those returned from the various inversion experiments (Fig. 3.2(d-i)) .

Figure 3.2(d) shows an example where spectral phytoplankton absorption (a_ϕ) from the inversion approach closely matches that from the simulated data. However, figure 3.2(e) shows that, even under cases where L_u(z) convergence is high (in this instance for the experiment using Levenberg-Marquardt inversion in Fig. 3.2(b)), estimation of a_ϕ can be poor. These data can be further broken down by looking at the performance in the four broad water types simulated. Figure 3.3(a, c, e) reveals that the ability to retrieve a_ϕ (at 445 nm) varies by water type, even when both L_u convergence and spectral a_ϕ are high across the dataset. Highest performance is seen under the Case 1/ low a_{gd}/low b_{bs}, whilst the other water types show greater spread in performance. Likewise, low correlation in spectral a_ϕ (e.g. Fig 3.3(c)), hides substantially different performances seen in the different water types. In the case of the Levenberg-Marquardt example here (Fig. 3.2(b)), poor performance is seen at low a_ϕ for all the water types except the low a_{gd}/low b_{bs}, the performance increases for these water types at higher levels of a_ϕ, though this again begins to tail off under extremely high a_ϕ. Performance under low a_{gd}/low b_{bs} is good throughout the dataset (black circles in Fig. 3.3(a)).

Overall, spectral $b_{b\phi}$ correlation across the different inversion experiments is lower than seen for a_ϕ (for example, see Fig. 3.3(b, d, f) for the water type breakdown of the spectral correlation achieved in the simplex example shown in figure 3.2(a, d and g)). However, the differences

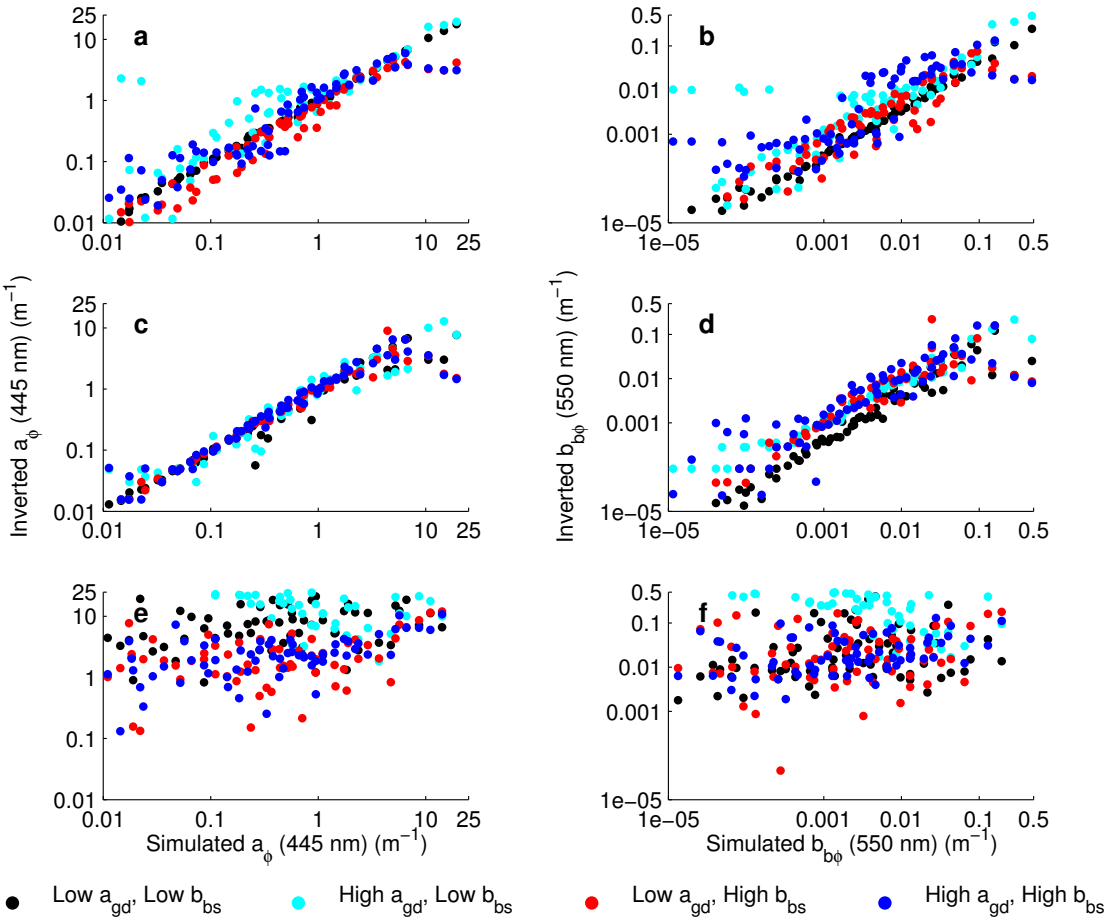

Figure 3.3: Correlation for a_ϕ (445 nm) (a, c, e) and $b_{b\phi}$ (550 nm) (b, d, f) estimated by the inversion process versus that input from the simulated dataset. Results are divided into the four broad water types simulated in the input data. These figures are examples from inversion experiments using the Reflectance approximation and a 1v variant with simplex, SCI diatom/dinoflagellate basis vectors and static initial conditions (a, b), 1v variant with Levenberg-Marquardt, SCI diatom/dinoflagellate basis vectors and grid searched initial conditions (c, d) and Evolutionary Algorithm with DCI diatom/dinoflagellate basis vectors and initial conditions with an empirically derived estimate of [Chl a] (e, f).

between the simulated water types, are even more substantial than seen for a_ϕ (Fig. 3.3(a)). Again, high levels of correlation are seen for the Case 1/ low a_{gd} and low b_{bs} data (black circles).

Although the primary focus of this work is on algal characteristics and their inherent optical properties in typically Case 1 waters (and as such the simulation of a_{gd} and b_{bs} here is simplified and limited), when attempting to understand ambiguity using a holistic reflectance model it is useful to see how these parameters are retrieved. This can also help explain how some of the experiments may show high levels of convergence and yet poor retrieval of algal IOPs/biogeochemical parameters. To examine whether variability in L_u(z) has been incorrectly attributed to a_{gd} and/or b_{bs}, the mean and standard deviations of the estimated values at 400 nm and 550 nm respectively, can be compared to the static values used under the four water types simulated. Table 3.3 shows the values calculated for the simplex example used above, which showed high convergence in modelled L_u(z) (Fig. 3.2(a)) and high capacity for estimating a_ϕ (Fig. 3.3(a)) and $b_{b\phi}$ (Fig. 3.3(b)) and for the Levenberg-Marquardt example, which showed

high convergence (Fig. 3.2(b)) but lower capacity in terms of estimating the IOPs (Fig. 3.3(c) and 3.3(c)). Values retrieved for a_{gd} (400 nm) and b_{bs} (550 nm) were generally good when the simplex was applied across the different water types. Comparatively, substantial over- and underestimation of of a_{gd} and b_{bs} occurred using the Levenberg-Marquardt technique, particularly in the highly scattering (high b_{bs}) cases. This suggests that the high levels of convergence seen in the Levenberg-Marquardt example result from non-unique solutions to L_u(z), and corresponding inaccurate estimates of the input a_{gd} and b_{bs} and algal IOPs. It therefore seems likely that incorrect [Chl a] and D_{eff} parameters will have been selected in this case.

Table 3.3: Summary statistics (mean and standard deviation (SD)) for a_{gd} (400 nm) and b_{bs} (550 nm). Low a_{gd} (400 nm) = 0.0197 m^{-1}, high a_{gd} (400 nm) = 1.97 m^{-1}, low b_{bs} (550 nm) = 0.0005 m^{-1}, high b_{bs} (550 nm) = 0.5147 m^{-1}. Examples here from (column 2 and 3) simplex experiment using diatom/dinoflagellate SCI basis vectors and static initial conditions and (column 4 and 5) Levenberg-Marquardt (LM) example using diatom/dinoflagellate SCI basis vectors and grid-search derived initial conditions.

Water type	*Simulated value for* $a_{gd}(400)$ *and* $b_{bs}(500)$	*Simplex* a_{gd}	*Simplex* b_{bs}	*LM* a_{gd}	*LM* b_{bs}
Low a_{gd} and Low b_{bs}	0.02 & 0.0005	0.02 (0.03)	0.003 (0.007)	0.03 (0.05)	0.002 (0.007)
High a_{gd} and Low b_{bs}	2 & 0.0005	1.4 (0.83)	0.007 (0.01)	0.25 (0.6)	0.004 (0.01)
Low a_{gd} and high b_{bs}	0.02 & 0.51	0.47 (1.0)	0.71 (0.2)	3.3 (1.1)	0.11 (0.05)
High a_{gd} and High b_{bs}	2 & 0.51	1.1 (1.5)	0.71 (0.18)	3.8 (1.3)	0.12 (0.04)
Low a_{gd} and Low b_{bs}	0.02 & 0.0005	0.02 (0.03)	0.003 (0.007)	0.03 (0.05)	0.002 (0.007)
High a_{gd} and Low b_{bs}	2 & 0.0005	1.4 (0.83)	0.007 (0.01)	0.25 (0.6)	0.004 (0.01)
Low a_{gd} and high b_{bs}	0.02 & 0.51	0.47 (1.0)	0.71 (0.2)	3.3 (1.1)	0.11 (0.05)
High a_{gd} and High b_{bs}	2 & 0.51	1.1 (1.5)	0.71 (0.18)	3.8 (1.3)	0.12 (0.04)

3.3.3 Performance statistics

All performance statistics discussed in this section relate to the experiments performed using the reflectance approximation (REFA) technique.

Simulated data using diatom/dinoflagellate basis vectors

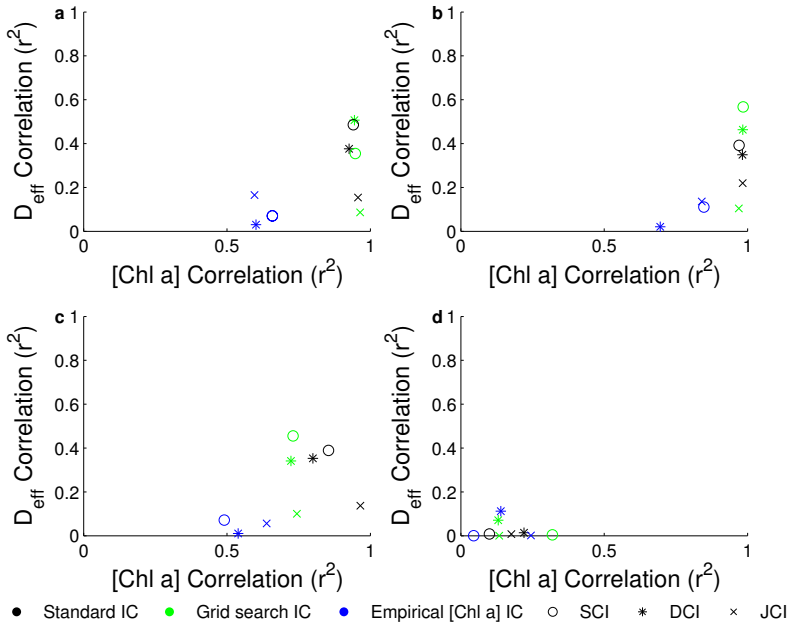

Figure 3.4: Correlation coefficients for [Chl a] and D_{eff} estimated by the algorithm versus those used to generate the simulated dataset using (a) Nelder-Mead simplex, (b) Nelder-Mead fminsearch (c) Levenberg-Marquardt and (d) Evolutionary algorithm non linear optimisation techniques. Colours indicate the different initial conditions used and different shapes refer to different IOP assumptions. Data here from the simulated data generated using low a_{gd} and low b_{bs} and diatom/dinoflagellate basis vectors.

To evaluate the ability of the inversion approach to return values consistent with those used to generate the simulated dataset, correlation coefficients (r^2) were derived between the input data from the simulated dataset and those estimated by the inversion approach. Figure 3.4 shows r^2 values for D_{eff} and [Chl a] estimation for the four inversion techniques (Fig. 3.4(a-d)) using different initial conditions (indicated by different colours) and using different basis vectors (indicated by different shapes) under low a_{gd} and low b_{bs} conditions. High levels of correlation for [Chl a] estimation are seen using the Nelder-Mead simplex with all 3 basis vectors and using both the standard and grid-selected initial conditions (Fig. 3.4(a)). The fminsearch optimisation (Fig. 3.4(b)) performs similarly to the simplex in terms of [Chl a] estimation, which may be expected given the similarity of the underlying mathematics (i.e. both are based on the Nelder-Mead technique (Nelder and Mead, 1965)). The Levenberg-Marquardt optimisation produces lower r^2 for estimated [Chl a] with only the standard initial condition/JCI basis vector combination resulting in an r^2 greater than 0.9. This is consistent with the suggestion made above that this approach has misattributed variance in L_u to other parameters, resulting in poor selection

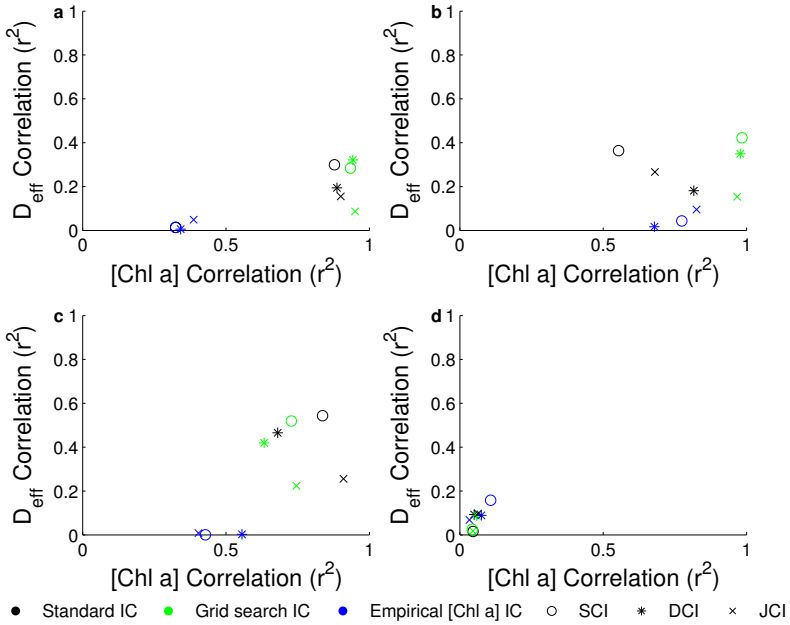

Figure 3.5: Correlation coefficients for [Chl a] and D_{eff} estimated by the algorithm versus those used to generate the simulated dataset using (a) Nelder-Mead simplex, (b) Nelder-Mead fminsearch, (c) Levenberg-Marquardt and (d) Evolutionary algorithm non linear optimisation techniques. Colours indicate the different initial conditions used and different shapes refer to different IOP assumptions. Data here from the simulated data generated using high a_{gd} and low b_{bs} and diatom/dinoflagellate basis vectors.

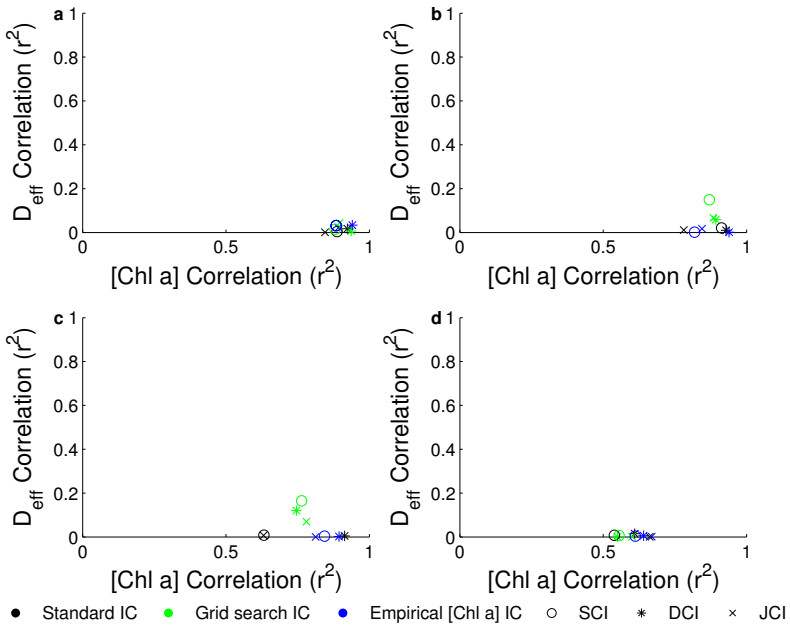

Figure 3.6: Correlation coefficients for [Chl a] and D_{eff} estimated by the algorithm versus those used to generate the simulated dataset using (a) Nelder-Mead simplex, (b) Nelder-Mead fminsearch, (c) Levenberg-Marquardt and (d) Evolutionary algorithm non linear optimisation techniques. Colours indicate the different initial conditions used and different shapes refer to different IOP assumptions. Data here from the simulated data generated using high a_{gd} and high b_{bs} and diatom/dinoflagellate basis vectors.

of [Chl a] and D_{eff}. The evolutionary algorithm produces generally poor returns for [Chl a] with r^2 values across all basis vector and initial condition combinations between 0.05 and 0.25. This could be expected given the relatively poor convergence seen using this technique to model the reflectance spectra. In terms of the different initial conditions, lowest [Chl a] correlation occurs across the inversion techniques (excepting the evolutionary algorithm) when using the empirically derived [Chl a] as an initial estimate. The use of standard and grid-search selected initial conditions results in broadly similar performance in [Chl a] estimation. Responses using the different basis vectors are not consistent across the inversion techniques.

Under low a_{gd}/low b_{bs} conditions, highest r^2 for D_{eff} estimation (above 0.8) is again found using the Nelder-Mead simplex inversion technique and SCI basis vectors with initial conditions determined using a grid search of the parameter space (Fig. 3.4(a), green circle). A general trend across the inversion techniques is that r^2 values for [Chl a] estimation are generally higher than those for D_{eff}; consistent with their typical role as first and second order causal variability. Additionally, high r^2 associated with D_{eff} and [Chl a] estimation co-occur, excepting the evolutionary algorithm, where both [Chl a] and D_{eff} r^2 values are low across the various IOP/initial condition combinations.

Compared to the performance results under low a_{gd}/low b_{bs} (Fig. 3.4), under high a_{gd} and low b_{bs} lower r^2 values are found for [Chl a] estimation (Fig. 3.5) whilst under high a_{gd} and high b_{bs} conditions, the r^2 values for [Chl a] estimation are slightly higher (Fig. 3.6). Estimation of D_{eff}, as indicated by the r^2 values become increasingly poor in the high a_{gd}/low b_{bs} (Fig. 3.5) and high a_{gd}/high b_{bs} (Fig. 3.6) conditions.

From the figures above it is clear that the performance of the inversion algorithm, in so far as its ability to retrieve the biogeochemical parameters prescribed in the generation of the simulated dataset, varies dependent on the inversion technique used, the choice of initial conditions and the water type. Generally, weaker performance in terms of both [Chl a] and D_{eff} is seen outside of the low a_{gd} and low b_{bs} data. Importantly, the higher levels of convergence associated with the use of the JCI basis vectors in some instances do not seem to translate into an increased capacity for accurate estimation of the [Chl a] and D_{eff}. The lack of consensus response when different IOP assumptions are used, suggests that the differences in reflectance spectra, imparted by these IOPs, are not substantial enough to significantly reduce the levels of ambiguity.

Simulated data using crytophyte basis vectors

Retrieval of [Chl a] and D_{eff} is generally weaker in the cryptophyte case than achieved when the diatom/dinoflagellate basis vectors are used in the inversion experiments. Similar trends are seen in terms of responses from inversion method and initial conditions used (i.e. the simplex and standard/grid searched initial conditions provide the best returns) under low a_{gd}/low b_{bs} conditions (Fig. 3.7). Performance decreases under higher a_{gd} and b_{bs} as seen previously.

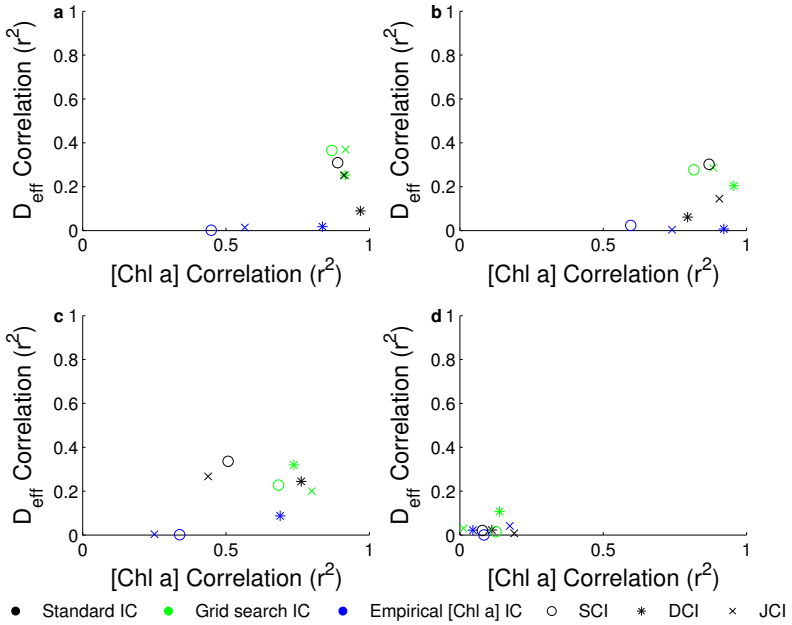

Figure 3.7: Correlation coefficients for [Chl a] and D_{eff} estimated by the algorithm versus those used to generate the simulated dataset using (a) Nelder-Mead simplex, (b) Nelder-Mead fminsearch, (c) Levenberg-Marquardt and (d) Evolutionary algorithm non linear optimisation techniques. Colours indicate the different initial conditions used and different shapes refer to different IOP assumptions. Data here from the simulated data generated using low a_{gd} and low b_{bs} and cryptophyte basis vectors.

Simulated data using combined basis vectors

In reality, phytoplankton assemblages are often mixed in that they may contain a number of functional groups with similar or unique characteristics, where dominance may not be easily attributable to one species in terms of numerical, volumetric or optical proxy abundance. As such, it may not be possible to preselect the most suitable IOP suite for application in blooms dominated by different assemblage types. However it also remains to be seen whether these different IOP suites ultimately impart substantial and unambiguous differences to the reflectance in such a way that they are not excessively ambiguous. To investigate this further, a 2v version of the EAP approach was used, where the inversion method was allowed to choose between the two suites of IOP basis vectors as an additional parameterisation. Although mixed samples were not simulated, giving the inversion technique the freedom to select basis vectors, allows (to first order) a determination of the extent to which second order variability (other than size) is detectable in conditions where species are dominant and in the context of varying competing sources of optical variability (i.e. in Case 2 type conditions).

Performance in terms of [Chl a] and D_{eff} (Fig. 3.8) is broadly similar to that seen in the 1v version using diatom/dinoflagellate (Fig. 3.4) and cryptophyte (Fig. 3.7) basis vectors with a slight reduction in the r^2 values for D_{eff} retrieval in the optimal cases (e.g. simplex and fminsearch optimisations) (Fig. 3.8(a and b)).

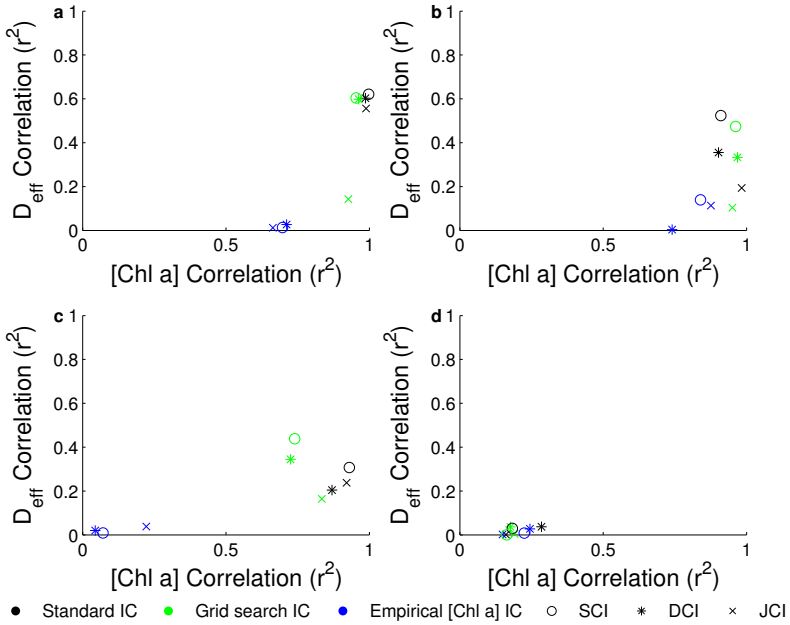

Figure 3.8: Correlation coefficients for [Chl a] and D_{eff} estimated by the algorithm versus those used to generate the simulated dataset using (a) Nelder-Mead simplex, (b) Nelder-Mead fminsearch, (c) Levenberg-Marquardt and (d) Evolutionary algorithm non linear optimisation techniques. Colours indicate the different initial conditions used and different shapes refer to different IOP assumptions. Data here from the simulated data generated using low a_{gd} and low b_{bs} for both sets of basis vectors and inverted with the 2v algorithm variant.

Comparing the % of each basis vector selected by the inversion approach to that used to generate the simulated data (either 0% or 100% diatoms/dinoflagellate or cryptophyte basis vectors) suggests that the correct IOPs can be selected by some of the inversion approaches. Most accurate selection of basis vectors (i.e. diatom/dinoflagellate or cryptophyte as appropriate) occurs consistently across all the inversion experiments under the low a_{gd}/low b_{bs} scenario. As could be expected from the performance seen so far, the simplex and fminsearch techniques perform best, frequently returning r^2 values above 0.7 for basis vector selection under the low a_{gd}/low b_{bs} conditions. The Levenberg-Marquardt and evolutionary algorithm approach by comparison, show r^2 of around 0.4-0.5 under these conditions. Lower r^2 are seen under the other water types, indicating the likelihood of greater difficulty retrieving species/functional type related IOPs under these conditions.

3.3.4 Impact of radiative transfer model choice

A selection of the inversion experiments above were run using the EcoLight-S radiative transfer code (Mobley (2011)) Figure 3.9 below show the EcoLight-S equivalents for figure 3.2(a, d, g) above, using the simplex technique, diatom/dinoflagellate SCI basis vectors and static initial conditions.

The modelled L_u(z) from EcoLight-S converges well on the measured solution (Fig. 3.9(a)). Note the range of EcoLight-S L_u(z) is substantially higher than that produced by the REFA approach. This was discussed in chapter two, and is largely the result of EcoLight-S more

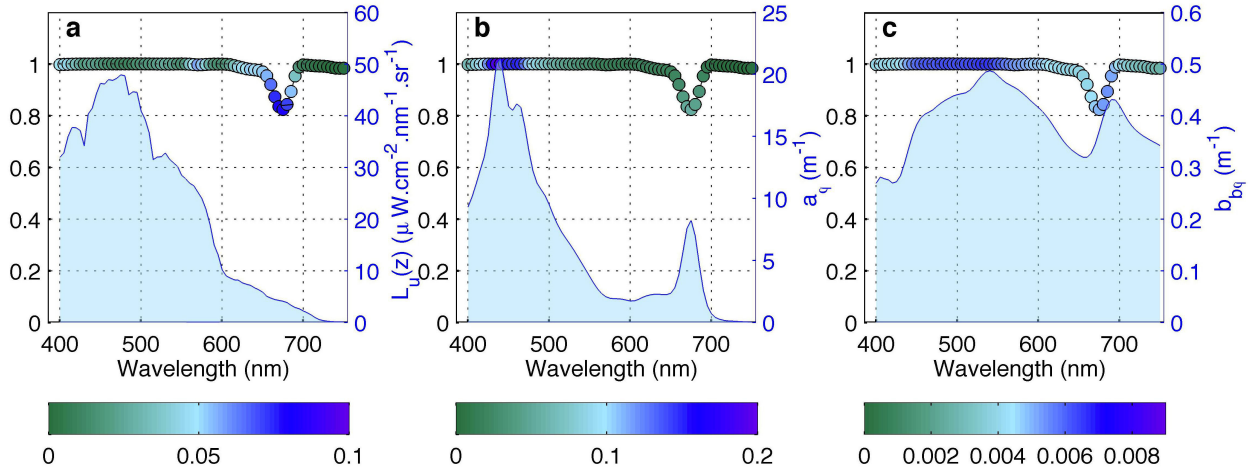

Figure 3.9: Spectral correlation for L_u(z) (a), a_ϕ (b) and $b_{b\phi}$ (c) estimated by the inversion process versus that input from the simulated dataset (n = 288, $p < 0.01$ in all cases). Dot fill colour represents the Root Mean Squared Error between the modelled output and simulated input, whilst the grey fill indicates standard error. The blue fill shows the range of each parameter associated with the simulated dataset. These figures are examples from inversion experiments using EcoLight-S and the 1v variant with simplex, SCI diatom/dinoflagellate basis vectors and static initial conditions.

appropriately simulating highly scattering waters where the single scattering approximation does not hold. Despite this much greater range, RMSE for EcoLight-S modelled L_u (z) versus that simulated is still relatively low.

a_ϕ is retrieved well, with low RMSE and high correlation observed when compared to the input simulated data (Fig. 3.9(b)). Retrieval of $b_{b\phi}$ is not as consistent as a_ϕ, however the results are still favourable (Fig. 3.9(c)) and comparable with those achieved through the REFA approach (Fig. 3.2(g)). $a_\phi(445)$ and $b_{b\phi}(445)$ retrieval varies by water type as before (Fig. 3.10(a) and (b)). In the EcoLight-S case, accurate retrievals are again seen under low a_{gd}/low b_{bs} conditions. Under the low a_{gd}/high b_{bs} conditions, improved retrieval of both $a_\phi(445)$ and $b_{b\phi}$ are seen using EcoLight-S compared to using the REFA approach (Fig. 3.2(d) and (g)). Both high a_{gd} cases show poor retrieval of $a_\phi(445)$ and $b_{b\phi}(445)$ when simulated $a_\phi(445)$ and $b_{b\phi}(445)$ are low, likely as a result of increased absorption ambiguity in the blue. Retrieval success increases under these water types as $a_\phi(445)$ and $b_{b\phi}(445)$ from the simulated dataset increase. [Chl a] are retrieved well, with r^2 values ranging from 0.93 (in the high a_{gd}/b_{bs} case) to 0.99 (in the low a_{gd}/b_{bs} case). D_{eff} retrievals are also favourable, although not as accurate as those retrieved through the REFA method, with a maximum r^2 value of 0.54 in the low a_{gd}/high b_{bs} case, and a minimum of 0.27 in the high a_{gd}/high b_{bs} case. Under several of the broad water type cases, the EcoLight-S inversion variant here shows substantial errors in the a_{gd}/b_{bs} estimation. It is possible that errors resulting from the mis-selection of size related absorption/backscattering are compensated for in terms of the modelled radiance by the selection of inaccurate a_{gd}/b_{bs} e.g.there is a substantial underestimation of a_{gd} and significant overestimation of b_{bs} across the retrieved estimates, which may result in accurate modelled L_u(z), despite inaccurate selection of the D_{eff}.

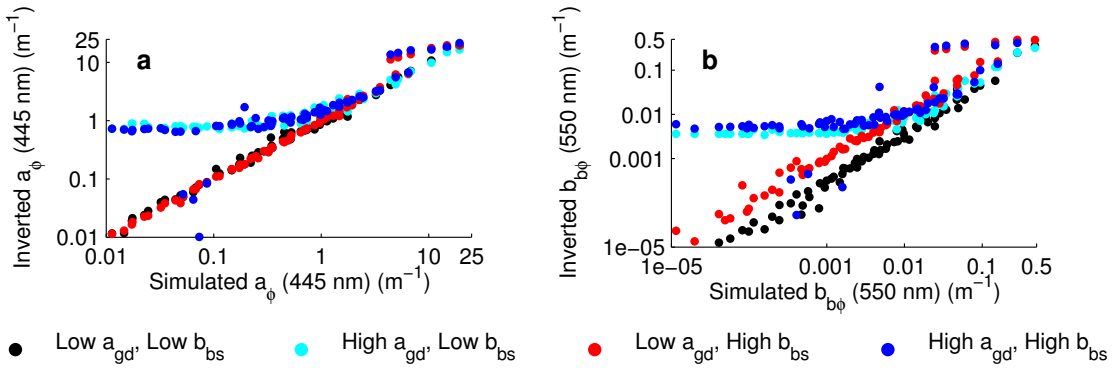

Figure 3.10: Correlation for a_ϕ (445 nm) (a) and $b_{b\phi}$ (550 nm) (b) estimated by the inversion process versus that input from the simulated dataset. Results are divided into the four broad water types simulated in the input data. These figures are examples from inversion experiments using EcoLight-S and the 1v variant with simplex, SCI diatom/dinoflagellate basis vectors and static initial conditions.

3.4 Discussion

3.4.1 Variability of estimation errors with regards to IOP budgets

The results above indicate that the likelihood of the inversion technique finding a "correct" solution (in terms of the [Chl a] and D_{eff} used in the simulated dataset) varies by broad water type. Greatest capacity (in terms of both estimation of IOPs and biogeochemical parameters/unknowns) is consistently found under low a_{gd} and b_{bs}. However, improvement for the highly scattering waters is seen when using EcoLight-S. This can be understood in the context of the sensitivity study and IOP budgets summarised in chapter two. Under the higher a_{gd} and higher b_{bs} scenarios, there is less size related spectral sensitivity in reflectance particularly at lower [Chl a] (below 10 mg m^{-3}) (see figures 2.3 and 2.4 in chapter two). This suggests that at least half of the samples in the simulated dataset have little to no size related sensitivity for the non-linear optimisation process to exploit during the fitting process. Under these conditions, derivation of size is likely highly ambiguous as spectra may be practically indistinguishable. This is in line with the results of (Sauer et al., 2012) who found that under relatively low biomass, where maximum band reflectance ratio algorithms are typically used, differences resulting from variable a_ϕ could easily be obscured by differences in a_{gd}.

To examine the extent to which this ambiguity is [Chl a] dependent in the inversion approach, figure 3.11 shows the root mean squared errors (RMSE) in D_{eff} estimation across the simulated range of [Chl a] for the four possible water types modelled in the simulated dataset and subsequently inverted: low a_{gd} and low b_{bs} (Case 1/the "Benguela type" waters above), high a_{gd} and low b_{bs}, low a_{gd} and high b_{bs}, and high a_{gd} and high b_{bs} (Case 2/gelbstoff and sediment influenced waters). Lowest errors occur in the context of Benguela type waters, with low a_{gd} and low b_{bs}. High error and substantial scatter in RMSE values across biomass levels for the high b_{bs} scenarios suggests significant ambiguity may be introduced under highly scattering conditions. As a result of this presumably more accurate handling of bidirectionality in Case 2 type waters, using the ES approach results in more consistent trends in error reduction with

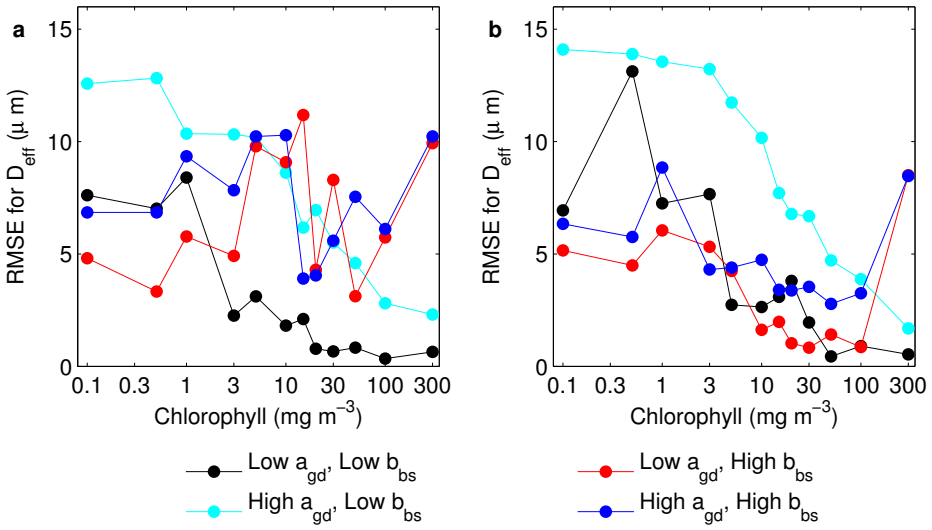

Figure 3.11: Root mean squared errors (RMSE) for D_{eff} retrievals at different [Chl a] under various parameter combinations for (a) the REFA and (b) ES methods. These data are from the simplex example above with SCI diatom/dinoflagellate basis vectors and static initial conditions.

increased biomass under high b_{bs} scenarios (Fig. 3.11(b)).

It was also indicated by the results in figures 3.4 to 3.8 that the successful estimation of D_{eff} and [Chl a] (in terms of highest r^2 values) generally occurred coincidently with successful estimation of [Chl a] or D_{eff} respectively. To further explore this, absolute errors in estimated D_{eff} are plotted against absolute percentage errors in estimated [Chl a] (Figure 3.12). Figure 3.12 suggests that obtaining an accurate size estimate is less likely when the [Chl a] is estimated incorrectly and/or [Chl a] is low (<10 mg m^{-3}, red dots).

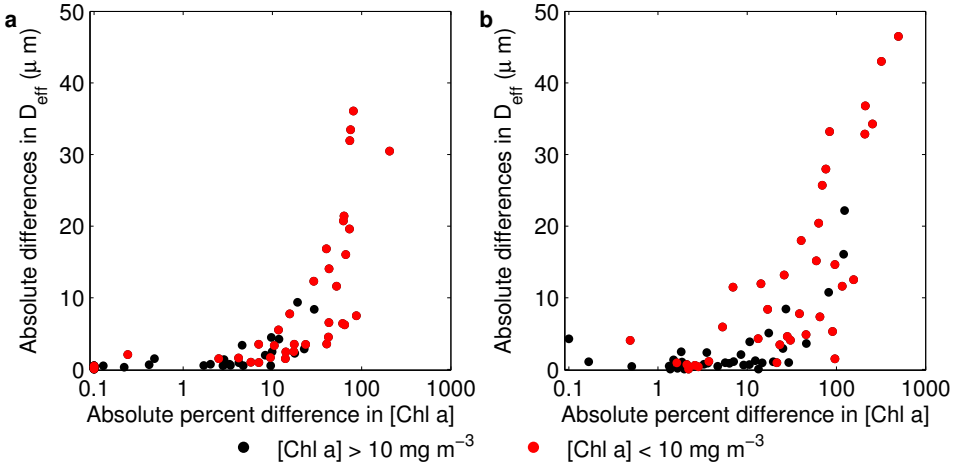

Figure 3.12: Errors in D_{eff} and [Chl a] estimation when inverting simulated data from the forward EAP model for low a_{gd} and low b_{bs} conditions for a) REFA and b) ES methods. These data are from the simplex example above with SCI diatom/dinoflagellate basis vectors and static initial conditions.

As mentioned above there is no clear distinction in performance (particularly for [Chl a] and D_{eff}) between the different IOP assumptions with regards to c_i. This would suggest that any differences in IOPs ultimately imparted to the reflectance spectra used as input here, do

not significantly alter the level of ambiguity present. In context, this presents potential difficulties in using optical inversion techniques such as this one to detect characteristics of particular phytoplankton species, which may manifest in a similar way. Encouragingly however, it seems the inversion approach is able to largely correctly select a set of basis vectors (between diatoms/dinoflagellates and cryptophyte sets) when presented with a spectrum generated using solely those absorption/backscattering characteristics. This supports the discussion in chapter two, that hyperspectral data should be most suited to detection of these features, given that significant differences in reflectance are only apparent at certain wavelengths, and in the hyperspectral inversion here, these differences can be identified. This is further supported by Sauer et al. (2012), whose modelled data and application of maximum band ratios suggest that approaches at specific wavelengths will be unlikely to retrieve information about second order variability of particles, due to insensitivity to particulate backscattering at specific wavelengths.

3.4.2 Role of cell size in ocean colour ambiguity

The extent of ambiguity in the ocean colour problem, has been the focus of a number of studies in recent years. Defoin-Platel and Chami (2007) conducted an extensive analysis of the extent of non-uniqueness (and hence ambiguity) in the ocean colour problem in coastal waters. Whilst ambiguity is essentially a simple concept (i.e. that multiple parameter combinations (of absorption and scattering) may yield indistinguishable R_{rs}), quantifying the extent of this in a meaningful way has not been done routinely, likely due to a lack of data covering coincident R_{rs} and a full suite of IOPs and no specific formulations to quantify ambiguity. To address the first issue, Defoin-Platel and Chami (2007) created a large synthetic dataset, based on radiative transfer simulations, constrained by observations and statistical relationships to encompass co-variations in IOPs typically seen in the coastal region. Of particular relevance to the results in this chapter, the simulated data by Defoin-Platel and Chami (2007) incorporated variable refractive indices for phytoplankton (1.05-1.10) and a variable exponent for a Junge power law (3-5) as an approach to incorporate variable particle phase functions into this simulated dataset. However the sensitivity of inversion response to this variability was not explicitly quantified. The lack of formulations was addressed through creation of several parameters to quantify uniqueness, including spectral distance (between R_{rs} spectra) and spectral neighbourhood (incorporating the likely uncertainty in R_{rs} measurements). This was then further deconstructed to assess the extent of ambiguity i.e. looking at the maximum difference in IOPs in ambiguous samples. Using their formulations, Defoin-Platel and Chami (2007) suggested that ambiguity in R_{rs} may exceed 90%, with variance dependent on spectral characteristics e.g. that minimum errors in absorption were highest when R_{rs} is low in the blue and highest for backscattering when spectral R_{rs} values were high. Similarly, Sauer et al. (2012) employed a model with variable a_{ϕ} (which could indirectly represent changes in size) and found that this variability could be obscured by a_{gd}, particularly at low biomass, consistent with the results here.

Although the simulated dataset created by Defoin-Platel and Chami (2007) incorporated

some variability which may be attributable to changes in phytoplankton community structure, the formulations used (e.g. Jungian) may be less suitable than those used in the two-layered sphere model, particularly for application in coastal, high biomass blooms (Bernard et al., 2009, 2007; Robertson Lain et al., 2014). Similarly, Sauer et al. (2012) used a variable a_ϕ, which may represent changes in size but not explicitly. Additionally, they did not account for size related variability in $b_{b\phi}$. Given the additional backscattering magnitude and shape imparted by the two-layered sphere model (Bernard et al., 2009; Robertson Lain et al., 2014), it is thus possible that they did not account for all potential size related sensitivity in R_{rs}. As such, until now no studies have explicitly quantified the ambiguity introduced by variation in coupled absorption and scattering specifically attributable to realistic variability in cell size. Whilst creation of a larger simulated dataset to explicitly quantify spectral errors attributable to ambiguity under different IOP scenarios is beyond the scope of this study, insight can be gained into the role of cell size in ocean colour ambiguity from the results above. In the context of the spectral sensitivities shown in chapter two, the errors estimated from the inversion experiments determined here, suggest both spectral and IOP parameter space regions where size related changes in optics may impact ambiguity. The impact of cell size on ambiguity in the ocean colour problem can be conceptualised in two main ways. Firstly, in terms of broader aims to retrieve information on phytoplankton community structure from optical data, there are likely regions of parameter space where these changes maybe insignificant in the total absorption and/or backscattering budgets, such that detection of accurate size from their impact on reflectance may be unlikely. Secondly, inclusion in an inversion approach may provide an additional degree of freedom which an unconstrained inversion approach may exploit at the expense of other parameters. The results in this chapter can show how this ambiguity manifests. In the first case, high errors are associated with estimation of size in water types where chapter two indicated that there is little size related variability in the associated IOP budget (e.g. low biomass, high a_{gd} and/or high b_{bs} conditions). Secondly, the non-constrained inversion techniques used here can find a proximate modelled reflectance solution and show incorrect estimations of parameters. These errors are also sensitive to the choice of radiative transfer technique as improvements in retrievals can be seen using EcoLight-S under the highly scattering scenarios.

3.4.3 Role of radiative transfer assumptions in ocean colour ambiguity

From the inversion experiments conducted in this chapter, it is apparent that the choice of radiative transfer technique used can substantially influence inversion potential. Although it may initially seem counter intuitive that the use of a more accurate radiative transfer scheme (EcoLight-S) induces more varied response in inversion retrieval of D_{eff}, this result is consistent with other studies on ambiguity. Defoin-Platel and Chami (2007) suggested that most (+80%) of the ambiguity in IOPs was related to uncertainties in particle phase function. They further suggested that variations in the directional effects of particles increase the number of combinations leading to non-unique R_{rs} between IOPs, and a resultant increase in ambiguity.

Whilst the models used in this chapter and in the study by Defoin-Platel and Chami (2007) are not that similar, the EcoLight-S variant of the EAP inversion approach using a radiative transfer model and spectrally variable Fournier-Forand phase functions allows for unconstrained f/Q. Hence, the observation by Defoin-Platel and Chami (2007) is consistent with the number of erroneous retrievals increasing when EcoLight-S is used, compared to those from the REFA approach, where directional effects are constrained by the f/Q parameterisation of Morel et al. (2002). The effect of this when results are compared to the REFA approach are two fold. Firstly, using EcoLight-S under high b_{bs} conditions, a wider range of radiance is simulated and can subsequently be more easily distinguished from others upon retrieval (i.e. explaining the reduced ambiguity in size retrievals under highly scattering scenarios (Fig. 3.11)). Secondly, the inclusion of more variable directional effects likely leads to more non-unique solutions from which the inversion process may select as suggested by Defoin-Platel and Chami (2007), such that under conditions where the various REFA approximations are more appropriate (i.e. low a_{gd} and low b_{bs}), greater variability in D_{eff} retrieval is seen using EcoLight-S.

3.4.4 Further considerations for optimisation techniques

The extent of ambiguity in the ocean colour problem apparent from both this and other studies, suggests that serious consideration must be made towards how inversion techniques can be optimised and constrained (Defoin-Platel and Chami, 2007; Huang et al., 2013; Rehm and Mobley, 2013). The experiments here do not suggest a clear way of predetermining optimum initial conditions to ensure retrieval of global rather than local minima from the inversion techniques tested. Although a substantial number of experiments were conducted here, other options for optimising the inversion approach could be explored. Rehm and Mobley (2013) constrained an inverse algorithm using the measured attenuation coefficient at a single wavelength, finding that this substantially reduced errors in total IOP determination. Termed the "enrichment approach" by Defoin-Platel and Chami (2007), the use of auxiliary measurements/information may be possible when applying this algorithm to *in situ* data. Given the [Chl *a*] dependence of the D_{eff} errors for example, it may be advisable to constrain the [Chl *a*] using either empirical estimates or *in situ* data. With regards to *in situ* application; although the range of phytoplankton IOPs simulated and modelled here does occur within the context of the high biomass, inshore and low biomass, offshore of the southern Benguela, further effort could be made to constrain their covariance, referred to as a "divide and conquer strategy" (Defoin-Platel and Chami, 2007; Rehm and Mobley, 2013).

3.4.5 Consideration of computational expense for scaling to satellite application

Conducting the inversion experiments above has highlighted differences in terms of computational efficiency in application of the different inversion methods. In terms of the optimisation scheme chosen, the Levenberg-Marquardt technique, offers the fastest convergence of all the methods tested, however, the results are particularly poor across the inversion experiments (e.g. figures 3.4 to 3.8). This is unlikely to improve on application to *in situ* data, where additional, measurement related errors are likely to arise. The simplex and fminsearch methods, offer the most accurate overall retrievals of the biogeochemical parameters, however both of these methods take substantially longer to converge than the Levenberg-Marquardt technique (approximately 8 times, dependent on the individual spectrum). The results in terms of relative computational expense versus performance, are broadly similar to those observed using similar mathematical schemes applied to the Garver-Siegel-Maritorena semi-analytical model in the recent study by Huang et al. (2013), though the addition of several degrees of freedom as a result of including size and species/type basis vectors could potentially increased the number of iterations needed to search parameter space optimally.

Although the best performance (in terms of [Chl a]/D_{eff}) was seen using the simplex technique with grid-search retrieved initial conditions, this technique does substantially increase algorithm run time. As the grid-search technique itself does not necessarily avoid the problems associated with ambiguity and does not produce substantially better results using the static initial conditions under the low a_{gd}/low b_{bs} conditions modelled which are most similar to the Benguela, eventual satellite application of this technique may be inefficient.

Similarly, although there are myriad benefits to using the EcoLight-S radiative transfer code in terms of accuracy in reflectance calculations, and an ability to investigate fully coupled scattering and bi-directional effects from custom IOP suites, inclusion substantially increases run time. In the case of the simplex and fminsearch optimisation variants, including EcoLight-S increases run time by around 300 times, although for some samples this can be up to 700 times. The exact run time is dependent on the number of iterations used and this could be optimised to a degree by setting a lower maximum number of iterations, or a more lenient function end tolerance, though this may potentially compromise finding of an appropriate solution. Additionally, it may be possible to further optimise the EcoLight-S run time using the built in options for this e.g. through wavelength skipping, and limiting the number of optical depths to which results are computed (Mobley, 2011). However, as can be seen from the results here and in chapter two, differences between the REFA approach and EcoLight-S are least substantial under low a_{gd}/low b_{bs} conditions, such as those typically found in the Benguela, and the inclusion of EcoLight-S in the inversion approach does not necessarily increase accuracy of parameter retrievals/decrease ambiguity. As such, it may be prudent in terms of computational efficiency, to use the REFA approach in Case 1 water types and/or to use the results from EcoLight-S runs to re-parameterise parts of the REFA approach, e.g. the f/Q or K_u assumptions.

3.5 Conclusions and recommendations

A number of conclusions and recommendations can be made about the EAP inversion approach used here, which, given the fully coupled nature of the modelled IOPs, are likely broadly applicable to other ocean colour studies:

- The ocean colour problem is highly ambiguous, as indicated by the variance in biogeochemical parameter estimation, despite close convergence of input simulated data and modelled spectral solutions in many cases.
- Success of the inversion technique (in terms of retrieving accurate D_{eff}) is highly [Chl a] dependent, and this dependency is further increased (i.e. higher biomass is required) when distinguishing cell size under more Case 2 conditions (i.e. higher a_{gd} and b_{bs}).
- The simplex inversion technique, using a static set of initial conditions, produces the best results for both [Chl a] and D_{eff} retrieval.
- As may be intuitive, given the history of ocean colour algorithm development, experiments here suggest lower errors in [Chl a] estimation will occur under all water types where biomass is highest and are lowest under Case 1 type conditions (low a_{gd} and low b_{bs}). Minimum errors range from 0.1 - 100% with increasing biomass under these conditions and increase substantially for high a_{gd}, high b_{bs} and combined Case 2 type conditions.
- Errors in D_{eff} estimation are also biomass dependent and follow similar trends to those seen in [Chl a] estimation under different simulated water types. Minimum errors range between 1 μm (under high biomass) and 8 μm (under low biomass) for low a_{gd}/b_{bs}. The error can be increased under high a_{gd}, b_{bs}, dependent on the radiative transfer assumptions made to account for high scattering in particular.
- The results here suggest that even with a sophisticated IOP and radiative transfer model, and under optimal conditions for detection of phytoplankton characteristics (i.e. when they dominate the IOP budget), it may not be possible to reliably escape the ambiguity inherent in the ocean colour problem. Hence statistically supported constraint/regionally specific training of inversion techniques is likely necessary to further increase accuracy of retrievals.
- Further work is needed to fully understand the role of cell size in ocean colour ambiguity, more broadly and with relevance to the EAP model. A more comprehensive set of simulated data, taking in to account IOP covariance, and a systematic analysis such as that conducted by Defoin-Platel and Chami (2007) would allow for further quantification of this.
- Encouragingly, most accurate performance for all parameters occurs under the high biomass, low a_{gd}/b_{bs} conditions which are common in the southern Benguela. This bodes well for future *in situ* and satellite application of these methods in this region.

Chapter 4

Application of the Equivalent Algal Population (EAP) inversion algorithm to *in situ* data from the southern Benguela.

Parts of this chapter formed content included in the paper Evers-King et al. (2014). The *in situ* validation and discussion of error and ambiguity within this paper were conducted by the lead author and have been expanded further for inclusion in this thesis chapter.

4.1 Introduction

The simulated datasets and forward modelling approaches shown in chapters two and three provide the opportunity to isolate causal variability in optical signals. However, to determine whether any radiometric inversion approach is suitable for systematic application to *in situ* and satellite derived ocean colour data, validation against an *in situ* dataset is required. Ideally, data that form a validation set should be sufficiently comprehensive in spatial and temporal distribution to capture the optical properties of different water types. Global datasets now exist for the purpose of open ocean product validation (e.g. NOMAD (Werdell and Bailey, 2005), MERMAID (Barker et al., 2007), SeaBASS (Hooker et al., 1994) and GeP&Co (http://www.lodyc.jussieu.fr/gep- co) - see chapter five for further discussion). However regionally specific algorithm validation efforts can suffer from a lack of data coverage, both geographically and with respect to water type i.e. measurements in high biomass conditions are particularly sparse. In addition, due to a paucity of data on suitable temporal, spatial and bio-optical scales, it can often be difficult to ensure consistency of data collected with different instruments, by different research groups, using different processing methods. *In situ* radiometry and biogeochemical data can provide detailed information for application to research and management problems (Bernard et al., 2006; Erickson et al., 2012; Pitcher et al., 2008a; Roesler and Boss,

2008; Schofield et al., 2002; Zibordi et al., 2006b) and are also vital for satellite ocean colour validation (Aiken et al., 2007; Mueller et al., 2003a; Pinkerton and Aiken, 1999; Zibordi et al., 2006a,b, 2004b) and even calibration (Franz et al., 2007; Mélin and Zibordi, 2010; Mueller et al., 2003b; Pinkerton and Aiken, 1999). However, making accurate *in situ* measurements of the light field requires consideration of sensitive measurement protocols and many sources of error can be introduced depending on the measurement techniques used (Zibordi and Voss, 2010).

In the following sections of this introduction, the various errors associated with *in situ* radiometric measurements are discussed, with a focus on in-water as opposed to above-water radiometers. Following this, an assessment is made of the different ways in which *in situ* measurements of phytoplankton cell size can be made, and how this relates to the various assumptions made in optical modelling.

4.1.1 Radiometric instruments, deployment, and measurement protocols

Observations of the marine light field began in the 1920s with designs for spectral radiometers, measurement methods and calibration following between the 1940s and 1970s (Zibordi and Voss, 2010). Quantitative measurements were initially made with custom built, in-water instruments and subsequently instruments have been developed by a number of companies for commercial sale. In the current market, *in situ* radiometers can take a variety of forms and are associated with different deployment protocols and data processing methods. All radiometers measure radiant flux at various geometries, which is then either separated through filters at specific spectral bands (multispectral), or passed through gratings or prisms over a desired range of wavelengths (hyperspectral). Many *in situ* radiometry instruments comprise both radiance and irradiance instruments, collecting the radiant flux over a specific solid angle and diffusively through a cosine collector, respectively.

Differences in radiometer type, deployment method and the bio-optical complexity of the water necessitate a number of protocols for data processing. In-water systems, both profiling and fixed depth, measure radiance distributions below the surface. The in-water part of the instrument can take several forms. Fixed depth system are typically either attached to a buoy with above and below water sensors at set height/depth e.g. the Buoy for the acquisition of long term optical time series (BOUSSOLE) (Antoine et al., 2008) and Marine Optical Buoy (MOBY) platforms (Franz et al., 2007) and the commonly used Satlantic Hyperspectral Tethered Surface Radiometer Buoy (H-TSRB) (Chang et al., 2003; Garcia et al., 2005; Leathers et al., 2004, 2001; Matthews et al., 2010; Schallenberg et al., 2008; Wang et al., 2005; Zhang et al., 2010). Profiling instruments can be attached to a winch or free falling, usually with a ship mounted above water sensor. In-water measurements are usually coupled to an above water irradiance measurement to correct for variation in the incident light field during the measurement period i.e. from changes in cloud cover. Measurement below the water surface is necessary for practical deployment of in-water radiometers and can help limit errors introduced by rapid fluctuations of the light field associated with wave action (Hooker et al., 2002; Zaneveld et al., 2001; Zibordi and Voss,

2010). However, an extrapolation scheme is then necessary to compare these measurements to the water leaving radiance (L_w) or Remote Sensing Reflectance (R_{rs}) products derived from satellite measurements. Examples of different in-water systems are shown in figure 4.1 and compared in Barlow et al. (2003). Above water systems have not been as thoroughly assessed as in-water methods (Zibordi et al., 2006b). However, recent years have seen an increase in the use of these systems for calibration and validation activities (Hooker et al., 2002; Hooker and Morel, 2003; Zibordi et al., 2006b) and new methods have emerged to measure L_w more directly from above water radiometry through a skylight-blocked approach (Lee et al., 2013). Both above and in-water systems offer a number of advantages and disadvantages which will vary in severity depending on the water type and deployment platform (Lee et al., 2013).

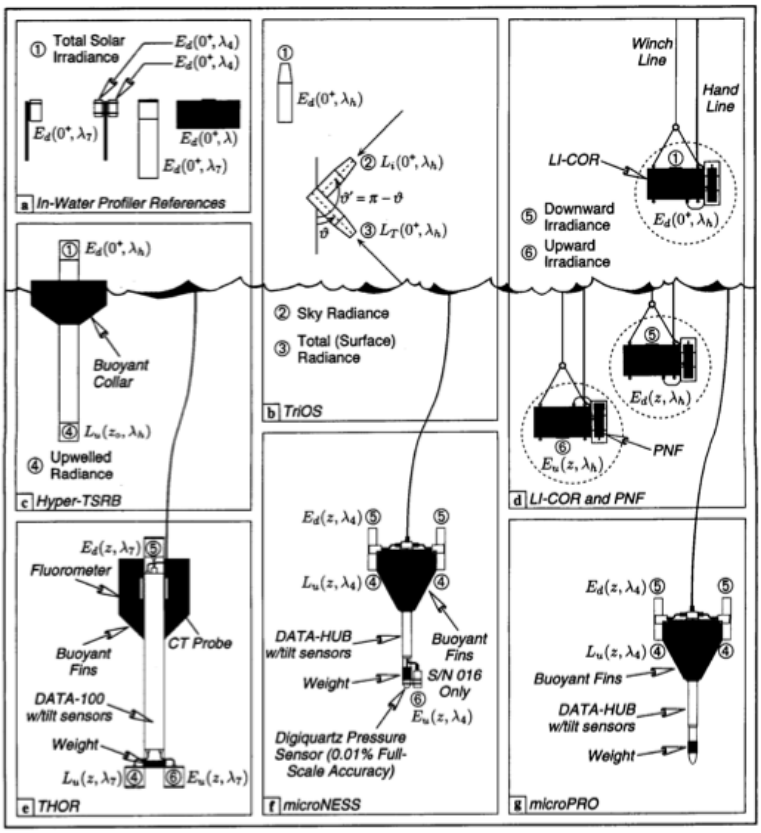

Figure 4.1: The range of in-water radiometer systems deployed during the BENCAL cruise, including above- and below- water sensors. Further details of the instruments, abbreviations and deployment methodology can be found in Barlow et al. (2003) from where this figure was reproduced.

A number of initiatives have sought to produce standard protocols for deployment of radiometric instruments and subsequent processing of data, to minimise the uncertainties introduced. The launch of the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) prompted development of protocols for calibration and validation activities, which were subsequently revised and updated during the SIMBIOS project (Mueller et al., 2003a). Despite the protocols introduced, differences in deployment and processing can still result in significant uncertainties in the Apparent Optical Properties (AOPs) measured, which can often be substantially more than required by mission uncertainty budgets (Antoine et al., 2008; Hooker et al., 2002; Hooker and

Maritorena, 2000; Hooker et al., 2001; Siegel et al., 1994). Major sources of error in radiometric measurements are summarised in table 4.1, with percentage error estimates from the literature.

Table 4.1: Summary of error sources and range estimates associated with in-water radiometric methods

Error Source	*Percentage error*	*Instrument type*	*References*
Calibration (Radiance)	1.5-6.3%	All	Hooker et al. (2002); Zibordi and Voss (2010)
Calibration (Irradiance)	1.1-3.4%	All	Hooker et al. (2002); Zibordi and Voss (2010)
Cosine response (Irradiance)	1.5-15%	All	(Zibordi and Bulgarelli, 2007)
Immersion factor (Radiance)	1-3%	In-water	Zibordi and Darecki (2006)
Immersion factor (Irradiance)	0.5%	In-water	Zibordi et al. (2004a)
Deployment structure perturbations	2-20%	All	Helliwell et al. (1990); Voss et al. (1986) Suraya et al. (1996); Weir et al. (1994) Hooker and Morel (2003); Piskozub (2004)
Wave focussing	0.1-30%	All	Zibordi et al. (2009a, 2004a, 1999)
Self Shading	3-36%	In-water	Gordon and Ding (1992); Kahru and Mitchell (1998) Doyle and Voss (2000); Piskozub et al. (2000) Leathers et al. (2004, 2001); Matthews et al. (2010)
Subsurface extrapolation	0.9-2.4%	In-water	Zibordi et al. (2004a); Zibordi and Voss (2010)

The errors listed in table 4.1 are not always spectrally homogeneous in their effects and can affect the various radiometric quantities (i.e. L_u/E_d) to different extents (Zibordi et al., 2004a). Additionally, many of these errors are highly dependent on the water type and environmental conditions (e.g. solar zenith angle, wind speed or cloud cover) (Zibordi and Voss, 2010). For self shading, wave focussing, correction for variability in incident irradiance and sub-surface propagation, errors are significantly larger for red than for blue wavelengths (Zibordi et al., 2004a; Zibordi and Voss, 2010). It is not always possible to conduct the extensive measurements and modelling techniques used in the studies cited in table 4.1 alongside those collected for scientific research and management purposes (Leathers et al., 2001). As such, instrument development has largely been focussed on reducing these errors through instrument design, deployment methodology and processing protocols and so a number of these errors can be managed and reduced substantially by most users (Leathers et al., 2004; Mueller et al., 2003b; Zibordi and Voss, 2010). Errors from deployment structure for example, can be minimised using free floating, profiling instruments which can take measurements a significant distance away from the boat/platform being used (Hooker and Morel, 2003; Zibordi and Voss, 2010). Wave focussing and shading effects can also be minimised through instrument design, deployment method and correction schemes (Leathers et al., 2004; Zaneveld et al., 2001; Zibordi and Voss, 2010). For static systems, increasing acquisition rate for measurements can average wave focussing effects over time (Zibordi and Voss, 2010). For profilers, in addition to increasing acquisition rate, deployment speeds can be decreased to average effects over depth and time (Zibordi et al., 2004a). Self shading effects have been reduced by the design of smaller in-water systems (Voss and Chapin, 2005) and corrections have been proposed and can be applied given knowledge of instrument specifications (Gordon and Ding, 1992; Leathers et al., 2004, 2001). Gordon and Ding (1992) found self shading errors to be a function of the absorption coefficient and hence under cases of high pigment concentration/beyond wavelengths of 600 nm, errors are increased. A correction scheme for this was proposed by Leathers et al. (2004, 2001). Their findings, using both Monte Carlo and analytical models to assess self shading, indicated that errors were highest when the ratio of scattering to absorption was smallest and under small solar zenith angles.

Characterisation and reduction of the errors associated with subsurface extrapolation is particularly crucial as *in situ*, in-water data are often used to validate and even calibrate ocean colour sensors. For this purpose remote sensing reflectance (R_{rs}) i.e. the water leaving signal, is required. Chapters two and three have shown that there can be substantial variance in attenuation coefficients (K_u/K_{Lu}), and the choice of parameterisation or radiative transfer model can influence the R_{rs} ultimately derived through the EAP forward model and inversion approach. Application of the EAP inversion approach to *in situ* data from subsurface instrumentation will thus require some consideration of this error source.

In principle, a layer where optically significant constituents are relatively homogenous should be selected, allowing for a linear decay scheme to be implemented. The type of instrument used will influence how this extrapolation interval is selected (Zibordi et al., 2004a). For fixed meas-

urements (i.e. those taken at a constant depth, from a buoy or mooring system), extrapolation usually takes place over the depth of the sensor. For data from a profiling instrument, an extrapolation interval can be assigned either manually or through statistical determination to a relatively homogenous layer in the upper part of the profile (Wright and Hooker, 2009; Zibordi et al., 2004a; Zibordi and Voss, 2010). In different water types, the depth of this homogenous layer can vary, with steep gradients potentially present in optically significant constituents, particularly in coastal areas (*ibid*). In particular, in waters that are highly attenuating, the homogenous layer is more likely to coincide with the upper few metres of the water column, where effects such as wave focussing are highest (Zaneveld et al., 2001; Zibordi et al., 2004a; Zibordi and Voss, 2010). Extrapolation intervals are typically set using a red wavelength as a reference (Wright and Hooker, 2009), where attenuation is highest, further limiting the depth of the homogenous layer in highly absorbent water i.e. in coastal regions (Zibordi et al., 2004a). Hence, the use of long sampling periods and fast acquisition rates over this part of the water column becomes even more crucial in error reduction (Hooker et al., 2002; Zibordi et al., 2004a; Zibordi and Voss, 2010).

4.1.2 Characterisation of bulk phytoplankton assemblage properties

Chlorophyll *a* concentration ([Chl *a*]) has historically been the starting point for linking optical signals with phytoplankton. This has largely been as a result of its ubiquity as a phytoplankton pigment (and thus as a proxy for biomass) and the subsequent collection of extensive *in situ* [Chl *a*] measurements (IOCCG, 2006; McClain, 2009). Fundamentally, empirically derived [Chl *a*] is dependent on IOPs (particularly phytoplankton absorption) (IOCCG, 2006). This, combined with greater understanding of how [Chl *a*] and phytoplankton absorption vary within species and as a result of physiology, has led to greater use of IOPs for deriving information on [Chl *a*] (Roesler and Barnard, 2013) and other phytoplankton assemblage characteristics (Ciotti et al., 2002; Kostadinov et al., 2009).

From an ecological modelling perspective, the size of the phytoplankton community has long been of interest to biological oceanographers (Agustí et al., 1987; Anderson, 2005; Chisholm, 1992a; Finkel et al., 2010; Finkel and Irwin, 2000; Irwin et al., 2006; Marañón, 2008; Raven, 1986; Raven and Kübler, 2002; Sieburth et al., 1978). However, quantifying and then simply representing the diversity of sizes and shapes present in any given algal assemblage is not without difficulty. In terms of size, measuring *in situ* particle size distributions in the ocean, and separating this into different organic and inorganic components produces a number of complications. In an idealised sense, measurements should cover the entire range of optically significant particle sizes - from the submicron to the millimeter scale (Reynolds et al., 2010). To link these size distributions with optical variability, measurements should also be able to capture this range appropriately in terms of its spatial/temporal variability influential at the scale of *in situ*/satellite radiometry (Stemmann and Boss, 2012). A number of instruments and methods exist for measuring phytoplankton cell size and PSDs, each with their own assumptions, advantages and

disadvantages (see recent reviews by Erickson et al. (2012); Koike and Fukuda (2007); Reynolds et al. (2010); Sosik et al. (2014); Stemmann and Boss (2012)). These methods are summarised in the following list and discussed in more detail below:

- Microscopy counts e.g. the Utermöhl and related methods e.g Havskum et al. (2004); Utermöhl (1931).
- Filtration and size fractionation e.g. Acevedo-Trejos et al. (2013); Ciotti et al. (2002); Dall'Olmo et al. (2009).
- Inferred size from chemo-taxonomic techniques i.e. High Performance Liquid Chromotography (HPLC) e.g. Antoine et al. (2011); Brewin et al. (2011a, 2010); Havskum et al. (2004); Vidussi et al. (2001).
- Particle imaging e.g. with FlowCAM, Imaging FlowCytobot and ZooSCAN and similar imaging instruments e.g. Álvarez et al. (2011); Erickson et al. (2012); Guidi et al. (2009); Jakobsen and Carstensen (2011); Olson and Sosik (2007); Reynolds et al. (2010).
- Optical sizing techniques e.g. Flow cytometry and Laser *In Situ* Scattering and Transmissometry (LISST) e.g. Andrews et al. (2011); Arin et al. (2002); Jonker et al. (1995); Karp-Boss et al. (2007); Koike and Fukuda (2007); Neukermans et al. (2012); Olson et al. (2003).
- Electronic particle sizing e.g. using a Coulter Counter e.g. Jackson et al. (1997); Li and Logan (1995); Mulligan and Kingsbury (1968); Olivieri (1985); Sheldon and Parsons (1967); Sheldon et al. (1972).

Microscopy counts arguably allow for the most in depth, qualitative understanding of the nature of the phytoplankton assemblage. The Utermöhl method has been used extensively throughout the literature for quantifying phytoplankton community cell size (Havskum et al., 2004; Hu et al., 2005; Kamykowski and Zentara, 2003; Marañón, 2008; Marañón et al., 2007; Pitcher et al., 2011). However, particles can be affected by sample handling and preservation methods mont(Montagnes et al., 1999), there is potential for operator error (Lund et al., 1958) and the time and taxonomic expertise required to process samples can be limiting (Sheldon and Parsons, 1967). Despite methodological improvements, time constraints can still make this method difficult for extensive sampling programs, i.e. those needed to validate satellite scale data.

Several options exist for processing samples using more automated approaches. Size fractionation through successive filtering and inference of size from High Performance Liquid Chromotography (HPLC) can provide bulk inferences about dominant cell sizes (Dall'Olmo et al., 2009; Uitz et al., 2006). However these approaches do not provide quantitative particle size distributions of the kind that may be required for development of particle based optical models

e.g. (Bernard et al., 2009, 2007; Kostadinov et al., 2010; Zhang et al., 2011). A number of instruments can provide quantitative PSDs of phytoplankton communities including flow cytometers and related instruments (Erickson et al., 2012), electronic particle sizers such as the Coulter Counter (Olivieri, 1985; Sheldon and Parsons, 1967; Sheldon et al., 1972) and particle imaging techniques (Álvarez et al., 2011; Erickson et al., 2012; Jakobsen and Carstensen, 2011; Neukermans et al., 2012; Reynolds et al., 2010). Flow cytometry can characterise broad communities much faster than microscopy, however sample volume can be small and it can be difficult to assess the full range of phytoplankton size (Erickson et al., 2012; Sosik et al., 2014). These limitations are being addressed with the development of autonomous *in situ* flow cytometry instruments specifically for oceanographic applications, such as FlowCytobot (Olson et al., 2003; Olson and Sosik, 2007) and others (Erickson et al., 2012). Limitations with conventional flow cytometry are also common to electronic particle sizes; including uncertainties with regards to how clusters or chains and non-spherical particles are quantified (Erickson et al., 2012). Optical methods which infer size based upon assumptions relating to particle scattering, may be less suitable in cases where those underlying assumptions are less valid (see chapter two for a discussion of the limitations of Mie theory). Electronic particle sizers such as the Coulter Counter avoid these assumptions, although as mentioned, some assumptions with regards to size and shape still have to be made. Further, electronic particle size measurements, like microscopy and conventional flow cytometry, require discrete water samples and cannot take place *in situ* which is a benefit of some new imaging and flow cytometry approaches such as FlowCAM and FlowCytobot (Álvarez et al., 2011; Reynolds et al., 2010). The impact of making discrete measurements rather than using in-water measurements must be balanced against the assumptions inherent in the method and whether these are valid for the system of interest. In the coastal setting of the southern Benguela, electronic particle sizing methods have historically proven useful for quantifying phytoplankton assemblages (Bernard et al., 2007; Crichton et al., 2013; Olivieri, 1985).

A comparison of the current, major techniques - electrical particle sizing (using a Coulter Counter), particle imaging (using FlowCam and microscopy) and optical approaches (flow cytometry and Laser *In Situ* Scattering and Transmissometry (LISST) was conducted by Reynolds et al. (2010). All approaches produced similar results for known suspensions and for polydisperse particle assemblages between 1 and 50 μm. Whilst optical approaches can provide a multitude of opportunities to either measure or infer particle size distributions, assumptions made with respect to the optical characteristics of particles can be limiting (Reynolds et al., 2010). In particular, the various assumptions made in the application of Mie theory (homogenous, spherical particles) and the shape of the distribution may be problematic, particularly for algal populations characterised by peaks in the PSD e.g. in coastal blooms (*ibid*). A further point of consideration is that many techniques discussed above cannot be done autonomously *in situ*. This can considerably limit application at a variety of scales and either preservation or delays in processing samples can result in degradation of samples which may be biased across taxa.

As discussed in previous chapters, the impact of phytoplankton cell size on optical signals

has been variously quantified through *in situ* observations and particle models. However, although cell size has been widely used as parameterisation for biogeochemical modelling purposes (Litchman and Klausmeier, 2008; Silió-Calzada et al., 2008; Ward et al., 2012), cases exist where aspects of ecosystem biogeochemical functionality may be better represented by phytoplankton functional types or even individual species (Anderson, 2005). The paradigm of Phytoplankton Functional Types (PFTs) and it's application to remote sensing is discussed in depth by a recent International Ocean Colour Coordinating Group (IOCCG) report (IOCCG, 2014) and a number of reviews (Nair et al., 2008; Rudorff and Kampel, 2012). The techniques used for the assessment of PFTs *in situ* have much in common with those used to characterise the phytoplankton particle size distribution. Cell size has been used as a metric in a number of PFT approaches, however others use additional metrics to classify phytoplankton groups. Sosik et al. (2014) give an overview of how various PFT information can be retrieved from microscopy, HPLC, molecular methods, flow cytometry and other optically based methods. PFT derivation from optical measurements and application to satellite ocean colour is discussed further in chapter five.

Remotely sensed ocean colour data, which have already been used extensively for validation of [Chl *a*] from biogeochemical model (e.g. Doney et al. (2009); Gregg et al. (2003); Kone et al. (2005); Lacroix et al. (2007); Siddorn et al. (2007), may be able to provide information about both phytoplankton size and phytoplankton functional types/individual species. However, as highlighted in the review by Anderson (2005) the reliability of these approaches must be assessed, and a lack of data and common metrics for PFT use in both satellite algorithms and biogeochemical modelling currently limits validation potential (*ibid*). The forward model and inversion scheme applied in chapters two and three respectively indicate that both cell size and functional type information (related to varying refractive indices) cause sensitivity in reflectance signals using the EAP forward model and can be retrieved to varying degrees of error using an inversion approach. However, as with any reductionist modelling approach, comparison to independent *in situ* data is required to assess real world applicability.

4.1.3 Summary and chapter aims

In summary the following points should be considered when selecting the most appropriate *in situ* phytoplankton assemblage data for investigating relationships with optical variability and for algorithm validation:

- Spatial and temporal coverage of the data needs to be sufficiently high that scales of variability in the phytoplankton community (biologically and optically) are suitably captured.
- As such, the volume of sample and processing time for each method should be considered and approached either through autonomous application or ensuring the integrity of manually acquired discrete samples
- Care must be taken with regards to how the method characterises the phytoplankton assemblage i.e. with regards to particle shape, volume and abundance.

- Similarly, for optical techniques, the suitability of any underlying assumptions should be considered e.g. for particle sizing techniques with regards to Mie theory.

This chapter assesses the extent to the which the EAP inversion approach can retrieve in-water IOPs and biogeochemical parameters from radiance measurements made with *in situ* radiometers. A validation dataset is compiled from field campaigns spanning the last 10 years, over a range of ecological and bio-optical conditions. The errors associated with the validation dataset are estimated from the existing literature. The idealised sensitivity study in chapters two and three was specifically configured for comparison to the *in situ* application of the EAP inversion algorithm to the data from the in-water Satlantic H-TSRB. Comparison between the simulated data and *in situ* validation gives context for how ambiguity in the inversion process may limit retrievals under different water types given *in situ* error budgets. Measurements from two in-water radiometers are cross-compared to investigate likely sources of error; particularly under high biomass conditions. Recommendations are then made with regards to further work required to establish suitable protocols for making measurements suitable towards bio-optical validation in high biomass and eutrophic waters.

4.2 Methods

4.2.1 Field methods and validation dataset

Due to its distinctive characteristics from a physical, bio-optical and ecological perspective, the southern Benguela has been the focus of numerous field studies over the last 12 years. As a consequence there is considerable detailed physical, biogeochemical and optical data coverage for a wide range of upwelling events and consequent bloom succession involving both mixed and mono-specific assemblages. A dataset was compiled where radiometric, absorption, [Chl *a*] and measurements of the particle size distribution were available. During most years, field campaigns have been undertaken during the high productivity season which occurs between February and April (Pitcher and Calder, 2000) from land based stations in the St Helena Bay area at Lambert's Bay and Elands Bay by a team of scientists from the University of Cape Town (UCT); Department for Agriculture, Forestry, and Fisheries (DAFF); and the Council for Scientific and Industrial Research (CSIR) for which methodological details are provided below. In addition to the data collected specifically during peak bloom activity, data were collected during the Benguela Calibration (BENCAL) cruise as part of a multi-institution effort for ocean colour satellite validation. Full details of the data collected can be found in the cruise report (Barlow et al., 2003). All historical data have been reprocessed from raw format to ensure consistency.

Radiometric data

Radiometric measurements used in this study came from several sources. Over most of the

sampling campaigns, a Satlantic H-TSRB was deployed. This instrument measures upwelling radiance (L_u(z)) at z = -0.66m, μW cm^{-2} nm^{-1} sr^{-1}) and above surface downwelling irradiance (E_d(0+), μW cm^{-2} nm^{-1}). Two 256 channel spectrographs are linked to an upward looking cosine corrected irradiance sensor and a downward looking 8.5^o field of view radiance sensor. Acquisition rates vary between 0.7 and 1.6 Hz in response to the light field. Spectral range covered by both spectrographs is 400 to 800 nm, with a sampling distance of 3.3 nm to an accuracy of 0.3 nm. The instrument was released to float a significant distance from the boat to prevent any interference/shadowing and radiometric measurements were typically recorded for between 2 and 5 minutes. Raw data were processed with Prosoft 6.3d (Satlantic: Halifax, Canada) and median values over the deployment were calculated and resampled to a 5 nm spectral resolution.

During 2011 and 2012, the Compact Optical Profiling System (C-OPS, Biospherical instruments) was also deployed. These data were used as auxiliary information here for comparison to the H-TSRB and discussion of likely sources of error. At each sampling location, three profiles were made to a depth of between 15 and 25 m. This instrument has two sets of multispectral (19 wavelength) radiometers on the profiler, for radiance (upwelling and downwelling) and irradiance (upwelling and downwelling). Coincident measurements are also recorded through the instrument deck box and μprofile software from a surface irradiance sensor to correct for variations in incoming solar radiation from cloud cover. Data collected by the C-OPS was processed using the PROSIT software package (S. Hooker and J. Brown (*personal communication*) also see meeting minutes in Wright and Hooker (2009)). This trial software is largely based upon the bio-optical processing protocols for profiling radiometry (Mueller et al., 2003b). The data were extrapolated to just below the surface using extrapolation intervals chosen on a cast by cast basis to ensure the selection of a relatively homogenous layer.

To measure total particulate absorption, the quantitative filter pad technique was used (Yentsch, 1962). Varying volumes from 100 ml to 1 l of seawater from the surface at each sampling location, were filtered through GF/F 0.7μm Whatmann filters at 10 mm mercury pressure. In some cases where excessive biomass was present, the amount filtered was reduced to prevent clogging. The absorbance of these filters were measured using a spectrophotometer (Shimadzu UV-2501PC, ISR-2200 integrating sphere) over a wavelength range of 350 to 750 nm with an air reference and baseline After analysis, each filter was placed in a petridish and methanol extraction was used to prepare the filters for detrital measurement (Kishino et al., 1985). These samples were left for 24 hours to extract pigments. After this each filter pad was returned to the filtering apparatus and washed with methanol before re-analysis in the spectrophotometer to determine detrital absorbance. Blanks were prepared by filtering several hundred ml of Milli-Q water through a GF/F filter and analysed using the spectrophotometer in the same way as each sample. Absorption coefficients were calculated from the raw absorbance data using equation 4.1 below where λ is wavelength, denoting spectral dependence, 2.303 is the coefficient representing transformation from $\log_{10}$ to $\log_e$, D'_f is the baseline corrected optical

density of the filter sample, D'_{750} is the null correction point, set at 750 nm, β' is the path-length amplification factor for the filter pads used, which is assumed to be 2 (Roesler, 1998) and l is the ratio of volume filtered (varies per sample) to the effective area of the filter pad (0.021 m^2).

$$a(\lambda) = 2.303 \frac{D'_f(\lambda) - D'_{750}}{\beta' l} \tag{4.1}$$

Spectral phytoplankton absorption ($a_\phi(\lambda)$) was calculated by subtracting detrital values ($a_d(\lambda)$) from the total particulate absorption ($a_p(\lambda)$) measured using the filter technique described above. Typical errors in this methodology include the amplification of path-length introduced by scattering caused by the fibres of the filter pads, and variability of the optical density of the filter pads due to moisture, contamination and uniformity of the pads. Path-length errors are addressed using the filter pad amplification factor (Roesler, 1998).

Measurements of particle size were made using two particle sizers. For most samples, a 128 channel Coulter Multisizer II with either a 50 μm or 140 μm aperture was used. In addition to this, samples from 2005 were analysed using an analogous method for a Beckmann Coulter Z2 cell and particle counter. In instances where biomass was extremely high, samples were diluted up to 1 in 100 times to enable counting without blocking the apertures. Dilution was made with 0.2 μm membrane filtered sea water. This methodology allows particles in the range 1 - 70 μm to be detected, an adequate range within the ecological context of the southern Benguela, where larger cells typically dominate. A daily blank was treated in the same way and results were removed from each sample. The effective diameters (D_{eff}) of the algal particle size distributions were calculated as shown in equation 4.2 (Ahn et al., 1992; Hansen and Travis, 1974) where d is the particle diameter and $F(d)d(d)$ is the number of particles per unit volume in the size range $d \pm 1/2d(d)$.

$$D_{eff} = \frac{\int_1^{100} \frac{\pi}{6} d^3 F(d) d(d)}{\int_1^{100} \frac{\pi}{4} d^2 F(d) d(d)} \tag{4.2}$$

Pigment analysis

Fluorometric measurements of chlorophyll were made using a Turner Designs 10-AU fluorometer (calibrated with commercial Chl *a* (Sigma)) based on the protocols of Mueller et al. (2003a). 100 ml of each sample was filtered through GF/F filters at 10 mm mercury pressure. The filter pads were ground in a polypropylene test tube using a glass rod with 9 ml of 90% acetone. The samples were then frozen for 24 hours to allow for pigments to be extracted. After 24 hours, the samples were thawed and centrifuged at 2500 rpm for 10 minutes. The supernatant was carefully poured i to new 13 mm x 100 mm disposable glass culture tubes. The fluorometer was zeroed with 90% acetone before each set of readings. Each sample was read in the fluorometer before being acidified to correct for phaeophyton pigments with a 0.15 ml 0.2N HCl solution and allowed to stand for three minutes before re-reading. Chlorophyll

concentrations were calculated using equation 4.3 where a_c is the calibration coefficient for the fluorometer, R_b is the measured value before acidification and R_a is the measured value after acidification, DF is the dilution factor. The extraction volume was 9 ml in all cases and the sample volume varied slightly per sample.

$$[\text{Chl } a](\text{mg m}^{-3}) = a_c * (R_b - R_a * \frac{\text{extraction volume}}{\text{sample volume}} * DF) \tag{4.3}$$

Species identification

Dominant species for samples were determined using microscopy counts according to methods described in Pitcher et al. (1991).

4.2.2 Inversion algorithm

The Equivalent Algal Population (EAP) inversion algorithm presented in chapter three and based upon the EAP forward model described in chapter two is applied here. The Satlantic H-TSRB derived L_u(z = -0.66m, λ) and E_d(0+, λ) are used as input.

The EAP inversion algorithm can be summarised as follows: Inputs are hyperspectral, subsurface radiance (L_u(-0.66 m),(λ), μW cm^{-2} nm^{-1} sr^{-1}) and hyperspectral, above surface downwelling irradiance (E_d(0+) (λ), μW cm^{-2} nm^{-1}). There are four or five solvable unknowns dependent on the algorithm variant (1v or 2v): Chlorophyll *a* concentration ([Chl *a*], (mg m^{-3})), algal effective diameter (D_{eff}, μm), combined gelbstoff and detrital absorption (a_{gd}(400 nm), m^{-1}), small particle backscattering (b_{bs}(550 nm), m^{-1}) and (in the case of the 2v version) a percentage of the two species groups. Algal optical properties are derived from several variants of the two-layered sphere model (Bernard et al., 2009) described below. For gelbstoff/detrital absorption and small particle backscattering an exponential shape function and power law respectively, with variable magnitude are assumed (see explanation in chapter two and Roesler and Perry (1995)).

The EAP inversion algorithm was applied to the *in situ* radiometric data using all the permutations described in chapter three. To summarise this included:

- Applying two different radiative transfer techniques: 1. a reflectance approximation with associated f/Q (Morel et al., 2002) and K_u (Albert and Mobley, 2003) parameterisations and 2. the EcoLight-S radiative transfer code.
- Running the algorithm in 1v and 2v modes, such that the inversion method was either constrained to one set of functional type IOPs (diatom/dinoflagellate or cryptophyte) or could choose a percentage of each to form spectral phytoplankton absorption (a_ϕ) and backscattering ($b_{b\phi}$) admixtures.
- Using 3 different sets of intracellular chlorophyll *a* concentration (c_i) assumptions to generate the size specific IOPs: 1. static c_i (SCI) (2.5/3.5 kg m^{-3} for diatoms/dinoflagellates

and cryptophyte groups respectively), 2. dynamic c_i (DCI) (parameterised with respect to size according to Zhang et al. (2011) and 3. IOPs generated using a Jungian size distribution as an admixture with the SCI IOPs below 10 mg m^{-3} [Chl *a*] (JCI).

- Using four different mathematical inversion techniques: 1. Nelder-Mead simplex, 2. Nelder-Mead fminsearch, 3. Levenberg-Marquardt optimisation and 4. Evolutionary algorithm. See table 3.2 and chapter three for details.

- Initialising the algorithm with different sets of initial conditions: 1. a static set of values appropriate as averages for those observed in the southern Benguela. 2. Sample specific values selected from a simple grid search over the range associated with each of the four/five parameters and 3. An empirically derived estimate for initial [Chl *a*] and static values for the other parameters.

The algorithm was applied to 75 casts from the Satlantic H-TSRB. These casts include those from the BENCAL cruise in October 2002 (mostly representing more offshore, Case 1 conditions) and from nearshore studies of Harmful Algal Blooms (HABs) near Lambert's and Elands Bay (dominated by high phytoplankton biomass) during October 2002 and March-April 2003, 2004 and 2005. Species present in the various diverse and monospecific phytoplankton assemblages ranged from the small ($\approx$ 2 μm) *Aureococcus anophagefferens* to the large *Ceratium furca* and included the autotrophic ciliate *Mesodinium rubrum*, containing cryptophyte endosymbionts. Casts were mostly taken between 9 and 10 am GMT to coincide with ENVISAT overpasses for the St Helena Bay region.

4.3 Results

4.3.1 Convergence of modelled versus measured upwelling radiance

A first order estimate of the efficacy of any inversion process is the ability of the technique to accurately converge towards the input solution, in this case reflectance (as derived from L_u(z) and $E_d(z)$) from the Satlantic H-TSRB. E_d(z) is not independently modelled by the inversion approach, so L_u(z) is used here to evaluate convergence. Figure 4.2(a, b, c) show the spectral convergence of the modelled versus measured L_u(z) from the Satlantic H-TSRB for three inversion experiments over the validation dataset in terms of r^2, with associated errors. Similar trends in terms of broad performance to those seen in chapter three are seen for the different inversion techniques applied to the *in situ* data. The simplex method produces the closest convergence across the different IOP suite/initial condition combinations (e.g. Fig. 4.2(a)), although convergence in terms of r^2 and RMSE is slightly poorer than observed in the previous idealised experiments using the simulated data (chapter three, Fig. 3.2(a)). This is not surprising, given the additional uncertainties involved in the *in situ* measurement techniques and likelihood of unaccounted for variability in IOPs etc. (see discussion). Convergence is also achieved with high r^2 and low RMSE using the fminsearch technique (not shown). As noted

in the previous chapter, the similarity between the simplex and fminsearch techniques is not unexpected as both are based on Nelder-Mead optimisation. Convergence is also relatively high using the Levenberg-Marquardt technique (e.g Fig. 4.2(b)), though poorer than using either of the Nelder-Mead techniques and less consistent across the various IOP/initial condition combinations. The poorest performance occurred using the evolutionary algorithm (e.g. Fig. 4.2(c)). Across all the techniques there was spectral variance in convergence in several key spectral regions, not obvious in the simulated experiments in the previous chapter, which may begin to shed some light on some of the model inadequacies and errors in the *in situ* data. Firstly, lower convergence and highest RMSE are seen in the blue, between 400 and 430 nm. As the southern Benguela is typically dominated by high biomass, $L_u(z)$ values in the blue are very small for many samples (e.g. < 0.0012 μW cm^{-2} nm^{-1} sr^{-1}). The sensitivity study by Defoin-Platel and Chami (2007) suggested that highest ambiguities in ocean colour inversion for total absorption may be found when this is the case. The EAP inversion algorithm also applies a relatively simple, constant parameterisation for the absorption of gelbstoff and detritus (a_{gd}). Although a_{gd} is likely a relatively minor constituent in St Helena Bay, any uncertainty in accounting for this would likely manifest in this region of the spectrum. Secondly, convergence decreases in the red, beyond 710 nm. As $L_u(z)$ measurements are small (< 0.0006 μW cm^{-2} nm^{-1} sr^{-1}), leading to low signal-to-noise ratios in the red region of the spectrum (Bergmann et al., 2004; Louchard et al., 2003), this may not be unexpected. Finally, a lower level of convergence is seen around 675-700 nm. This is due to the lack of fluorescence term in the model and low convergence weighting set in this range for the optimisation techniques. The trends in convergence across the inversion approaches are broadly similar when the 1v (diatom/dinoflagellate or cryptophyte) and 2v (admixture) are applied and there is no consistent difference in convergence as a result of the different initial conditions. This may indicate that, for this *in situ* dataset, there is insufficient sensitivity in reflectance attributable to changes in functional type, relative to the inherent ambiguity in the inverse problem.

4.3.2 Comparison between measured and modelled phytoplankton absorption

Given the forward and inverse approach used (see chapter one, figure 1.3), the ability of the EAP algorithm to determine intermediate IOPs is of importance for the ultimate derivation of accurate biogeochemcial parameters. For appropriate validation, algorithm output must be compared to independently derived *in situ* data. Figure 4.2(d, e, f) compares the a_ϕ measured from *in situ* samples with that estimated in the three inversion experiments. When the algorithm estimated a_ϕ is compared to that derived from the spectrophotometer, results vary substantially depending on the inversion method, basis vectors and initial conditions selected. The simplex technique produces highest spectral correlation (Fig. 4.2(d)) of all the inversion techniques, as observed in the chapter three inversion experiments. Similar spectral performance is seen using the fminsearch technique, whilst poorer responses are seen from the Levenberg-Marquardt and

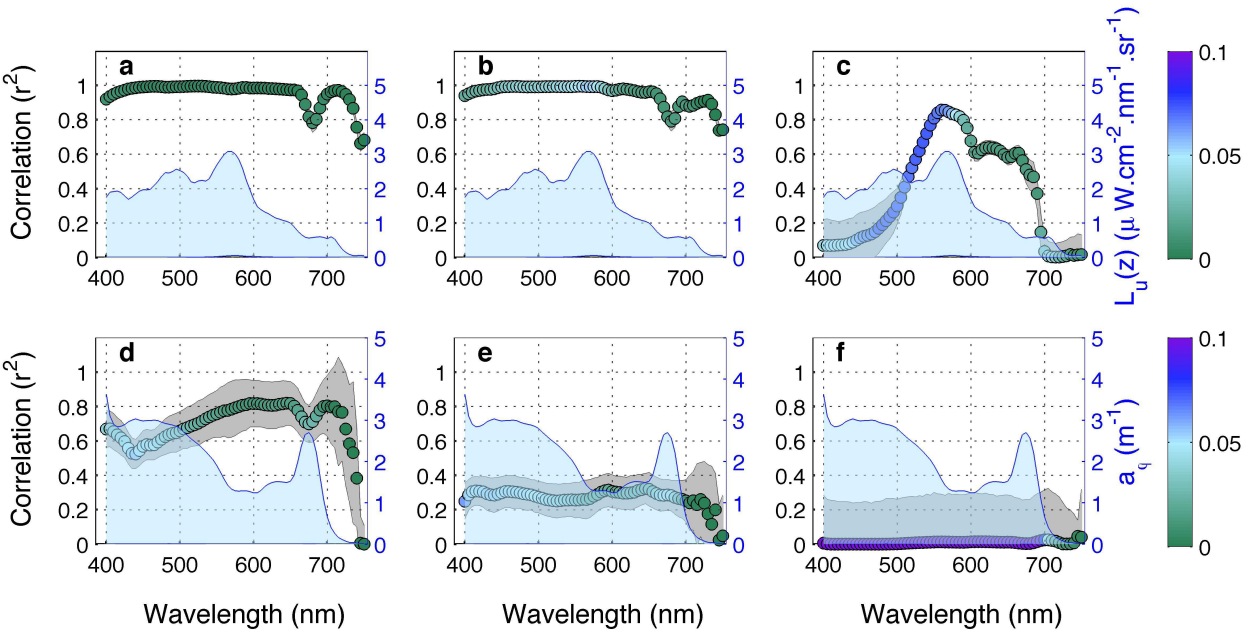

Figure 4.2: Examples from inversion experiments using the 1v variant with simplex, SCI diatom/dinoflagellate basis vectors and static initial conditions (a, d), Levenberg-Marquardt, SCI diatom/dinoflagellate basis vectors and grid searched initial conditions (b, e) and Evolutionary Algorithm with DCI diatom/dinoflagellate basis vectors and initial conditions with an empirically derived estimate of [Chl *a*] (c, e). Spectral correlation coefficients for L_u(z) (a, b, c) and a_ϕ (d, e, f) estimated by the inversion process versus that input from the H-TSRB dataset (a, b, c) and derived from *in situ* absorption data (d, e, f) ($p < 0.01$ in all cases). Dot fill colour represents the Root Mean Squared Error (RMSE) between the modelled output and H-TSRB/*in situ* a_ϕ, whilst the grey fill indicates standard error. The blue fill shows the range of (a) L_u(z) associated with the H-TSRB dataset and (b) a_{phi} derived from *in situ* samples.

evolutionary algorithm methods (Figures 4.2(e and f)).

In terms of the use of different IOP basis vectors (with respect to the c_i parameterisation), substantial sensitivity in response was seen across all the inversion techniques. In particular, the use of JCI vectors tended to coincide with a reduced ability to accurately retrieve a_ϕ, with higher RMSE seen in the blue in many cases. This may be as a result of the reduced sensitivity imparted by these vectors, as identified in chapters two and three, resulting in greater ambiguity. There is also a likelihood that these vectors (and the use of a Jungian distribution) may not be suitable in the context of the southern Benguela (Bernard et al., 2007). Besides this, there is little consistent trend in inversion response across the different inversion techniques when different initial conditions are used. No substantial differences are noted when both sets of

diatom/dinoflagellate or cryptophyte basis vectors are available for selection by the inversion process. However, slightly poorer returns are seen when only the cryptophyte basis vectors are used.

4.3.3 Estimation of algal descriptors; chlorophyll a concentration ([Chl a]) and effective diameter (D_{eff})

Chlorophyll a values in the *in situ* dataset range from 0.2 to 309 mg m^{-3} whilst effective diameters derived from the Coulter Counter particle size distributions range between 2.5 and 31.8 μm. Figure 4.3 to 4.5 summarise the algorithm estimation of [Chl a] and D_{eff} for each inversion method and the various basis vector/initial condition combinations for the three versions of the EAP algorithm; i.e. only diatom/dinoflagellate basis vectors (Fig. 4.3)); only cryptophyte basis vectors (Fig. 4.4) and both basis vectors available as an admixture dependent on a variable parameter iterated through the optimisation approaches (Fig. 4.5).

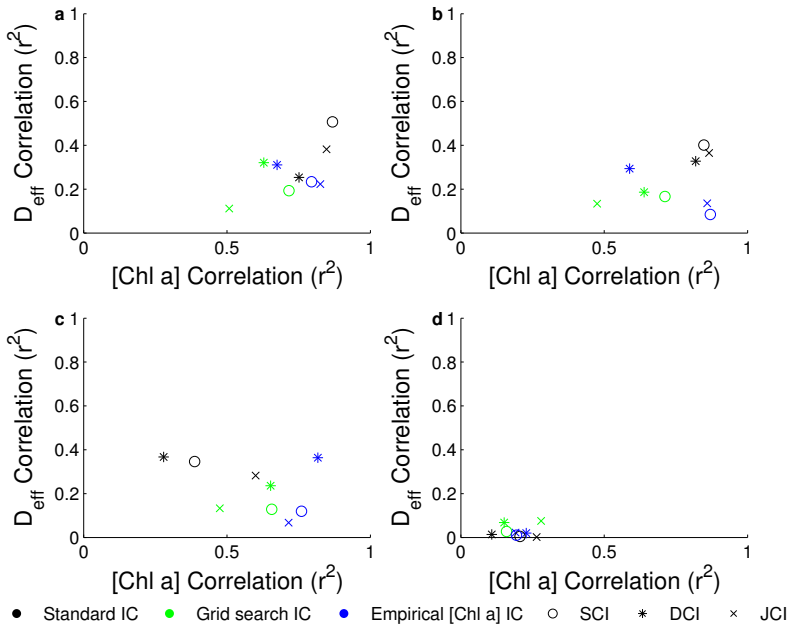

Figure 4.3: Correlation coefficients for [Chl a] and D_{eff} estimated by the algorithm versus those from the *in situ* dataset using (a) Nelder-Mead simplex, (b) Nelder-Mead fminsearch (c) Levenberg-Marquardt and (d) Evolutionary algorithm non linear optimisation techniques. Colours indicate the different initial conditions (IC) used and different shapes refer to different IOP assumptions. Results here used the diatom/dinoflagellate basis vectors. All results are statistically significant at $p < 0.01$, where n = 72 for [Chl a] and n = 44 for D_{eff}

Best performance across the validation set (in terms of r^2 for [Chl a] and D_{eff}) is seen using the Nelder-Mead simplex, SCI, diatom/dinoflagellate basis vectors and standard static initial conditions (Fig. 4.3(a)). Performance decreases slightly when using only the cryptophyte basis vectors (Fig. 4.4), but is similar when the admixture approach is available to the optimisation (Fig. 4.5). Generally, the fminsearch approach performs similarly to the simplex across the various basis vector/initial condition combinations (Figs. 4.3b, 4.4b and 4.5(b)). The Levenberg-Marquardt technique is slightly poorer than the simplex and fminsearch techniques, particularly

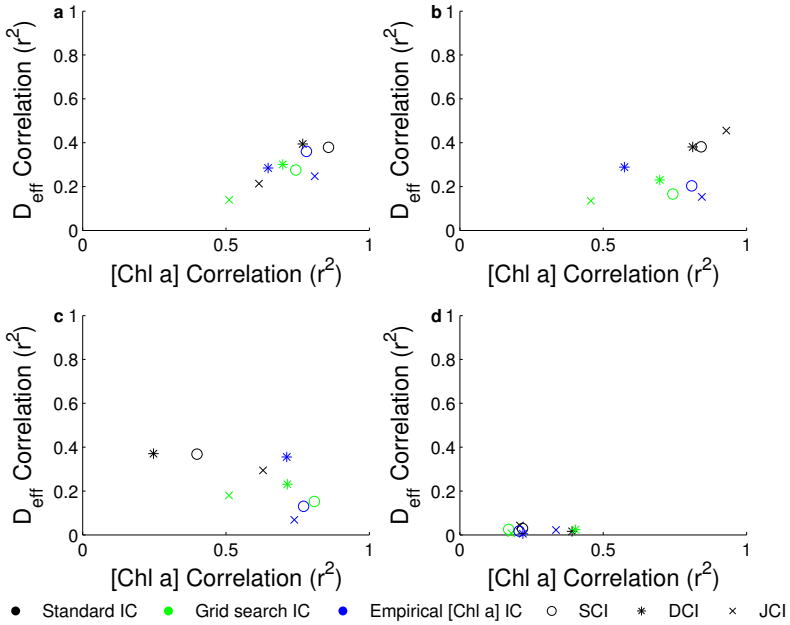

Figure 4.4: Correlation coefficients for [Chl a] and D_{eff} estimated by the algorithm versus those from the *in situ* dataset using (a) Nelder-Mead simplex, (b) Nelder-Mead fminsearch (c) Levenberg-Marquardt and (d) Evolutionary algorithm non linear optimisation techniques. Colours indicate the different initial conditions (IC) used and different shapes refer to different IOP assumptions. Results here used the cryptophyte basis vectors. All results are statistically significant at $p < 0.01$, where n = 72 for [Chl a] and n = 44 for D_{eff}

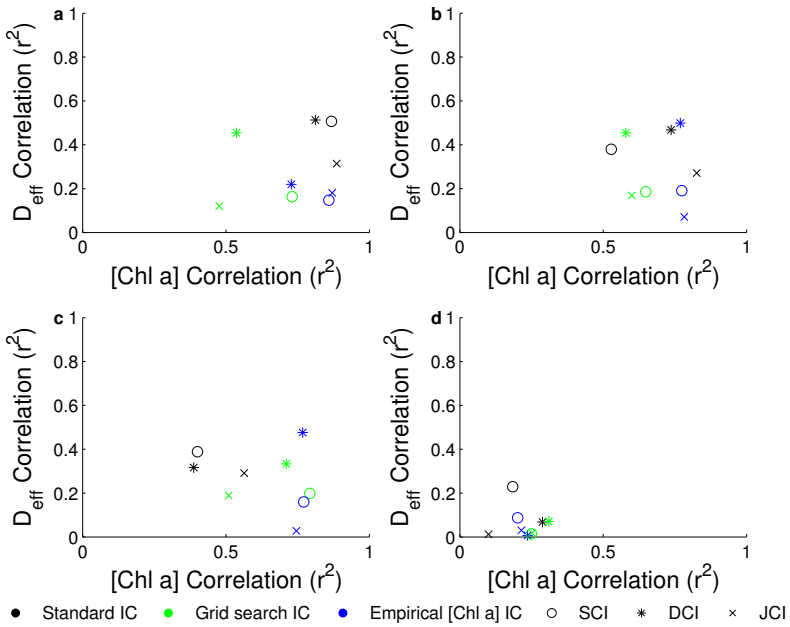

Figure 4.5: Correlation coefficients for [Chl a] and D_{eff} estimated by the algorithm versus those from the *in situ* dataset using (a) Nelder-Mead simplex, (b) Nelder-Mead fminsearch (c) Levenberg-Marquardt and (d) Evolutionary algorithm non linear optimisation techniques. Colours indicate the different initial conditions (IC) used and different shapes refer to different IOP assumptions. Results here used both sets of basis vectors with a variable admixture applied through the optimisation techniques. All results are statistically significant at $p < 0.01$, where n = 72 for [Chl a] and n = 44 for D_{eff}

for [Chl a] estimation (Figs. 4.3(c), 4.4(c) and 4.5(c)). Performance is similarly poor across all combinations for the evolutionary algorithm technique (Figs. 4.3(d), 4.4(d) and 4.5(d)).

Most notably, the use of an empirical estimate for the initial value for [Chl a], appears to improve results, contradictory to the results from the simulated data inversion in chapter three. In contrast, results using a grid search to determine the initial conditions are poorer when applied to the *in situ* data compared to the simulated data inversion in chapter three. Although in some instances there is a general trend that high r^2 for [Chl a] and D_{eff} coincide, this trend is not as clear as that seen in chapter three.

4.3.4 Performance using EcoLight-S

As in the previous chapter, EcoLight-S was used instead of the reflectance approximation to examine the effects of the assumptions made (i.e. with regards for f/Q and K_u) on both radiometric and biogeochemical algorithm performance.

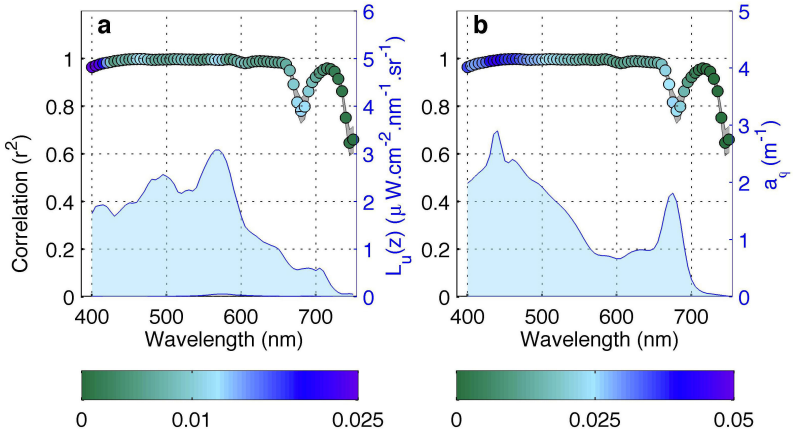

Figure 4.6: Spectral correlation for (a) L_u(z) and (b) a_ϕ estimated by the inversion process versus that from the H-TSRB and *in situ* absorption data respectively ($p < 0.01$ in all cases). Dot fill colour represents the RMSE between the modelled output and H-TSRB/*in situ* a_ϕ, whilst the grey fill indicates standard error. The blue fill shows the range of (a) L_u(z) associated with the H-TSRB dataset and (b) a_{phi} derived from *in situ* samples. This example using EcoLight-S, the simplex technique, SCI basis vectors for the diatom/dinoflagellate group and static initial conditions.

Whilst the response in terms of L_u convergence using EcoLight-S (Fig. 4.6(a)) is almost identical to that seen in the equivalent case using the reflectance approximation (Fig. 4.2(a)), and the estimated a_ϕ compares favourably to the *in situ* data (Fig. 4.6(b)), the performance statistics for [Chl a] and D_{eff} are poorer using the EcoLight-S case (Fig. 4.7(b)) than seen in the optimum REFA case (Fig. 4.7(a)). Despite seeming somewhat counter intuitive (as one may expect more accurate radiative transfer assumptions to lead to improved results) this is consistent with the results seen in the previous chapters, where greater ambiguity was seen when EcoLight-S was used, leading to more potential solutions for the inversion approach to select. This is discussed further below with regards to the *in situ* dataset in particular.

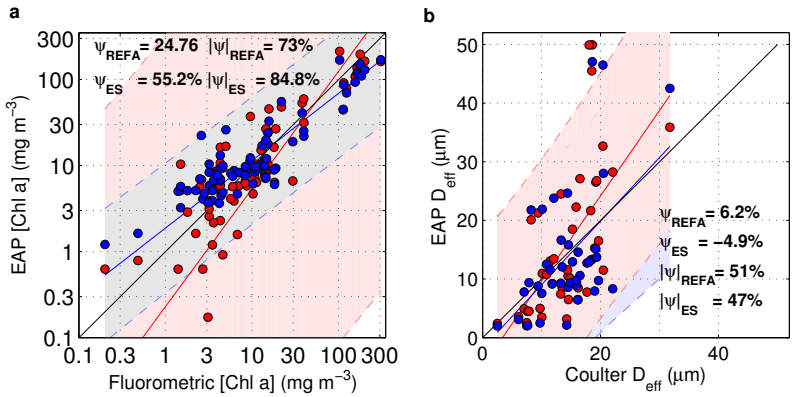

Figure 4.7: Correlations between measured and algorithm estimated (a) [Chl *a*] (n= 73) and (b) D_{eff} (n=44). Red and blue dots represent values derived from the REFA and ES approaches respectively. r^2 values for [Chl *a*] estimation were 0.81 (REFA) and 0.86 (ES), and 0.4 (REFA) and 0.36 (ES) for D_{eff} estimates, with $p < 0.001$ in all cases. Estimates of absolute percentage error ($|\psi|$) and bias (ψ) are given as per the method used in Zibordi et al. (2004b). Shaded areas show 95% confidence intervals based on linear regression for each dataset. These examples both use the SCI, diatom/dinoflagellate basis vector with a static set of initial conditions.

4.3.5 Comparison between the Satlantic and C-OPS radiometers, evaluation of self-shading errors and consequences for processing radiometric data in high biomass waters

During 2011 and 2012 the Satlantic and C-OPS radiometers were deployed in close coincidence, allowing for a preliminary intercomparison of their radiometric performance and the utility of their individual characteristics for determining both radiometric variables and biogeochemical information through the inversion algorithm. During 2011, conditions were generally lower biomass than observed in previous years, with no reported incidence of obvious high biomass bloom activity in the area over the typical peak bloom season (March-April). Comparatively, 2012 was a year of exceptionally high bloom activity with high biomass conditions reported *in situ* and observed by satellite from January through the sampling period in March and into April.

Figure 4.8 below shows some example L_u(z) profiles from the C-OPS radiometer deployed in Elands Bay on the 6th March 2011 (4.8a and 4.8c) and Lambert's Bay on the 7th March 2012 (Fig. 4.8(b and d)). L_u profiles are shown for the 443 nm (Fig. 4.8 (a and b)) and 665 nm (Fig. 4.8(c and d)) wavelengths as these spectral regions are used by the PROSIT software to determine suitable extrapolation intervals for propagation of the signal to above water. Figure 4.9 shows the equivalent R_{rs} spectra processed from these L_u profiles using PROSIT and the R_{rs} estimated by the EAP inversion algorithm applied to coincident Satlantic H-TSRB casts. *In situ* [Chl *a*] values of 3 mg m^{-3} and 88.7 mg m^{-3} were associated with the 2011 and 2012 blooms respectively.

From figures 4.8 and 4.9 several observations can be made. Firstly, as would be expected, near surface L_u values are elevated at 665nm in the high biomass conditions (Fig. 4.8(d)), relative to the low biomass conditions (Fig. 4.8(c)) and vice versa at 443 nm. Secondly, the rate

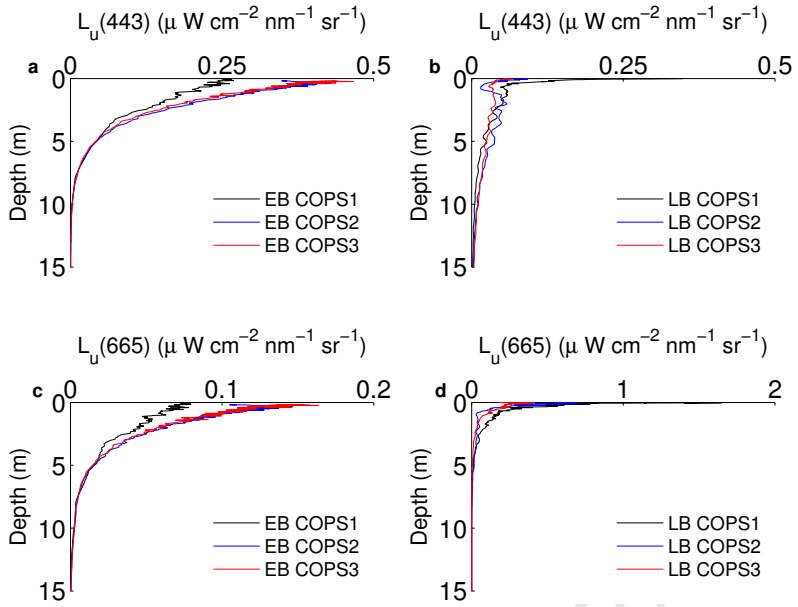

Figure 4.8: L_u (z) profiles from the C-OPS radiometer from the 6th March 2011 at (a) 443nm and (c) 665nm and from the 7th March 2012 at (b) 445 nm and (d) 665 nm. C-OPS casts were processed using PROSIT to correct for self shading, tilt and roll effects etc. Note the scale differences for panel (c) and (d). Numbers 1-3 refer to consecutive casts made at the same sampling station.

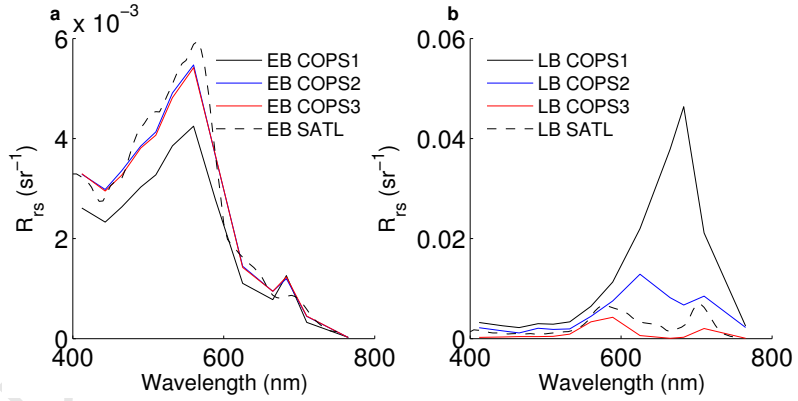

Figure 4.9: R_{rs} spectra derived from the C-OPS (solid lines) profiles and Satlantic H-TSRB (dashed line) casts on (a) 6th March 2011(EB - Elands Bay) and (b) 7th March 2012 (LB - Lambert's Bay). C-OPS casts were processed to R_{rs} using PROSIT, whilst the H-TSRB casts were processed using the inversion algorithm above with SCI, diatom/dinoflagellate basis vectors, static initial conditions and reflectance approximation.

of attenuation with depth is substantially higher in the high biomass case, particularly at 665 nm. Finally, based on figure 4.9 and secondary experiments using the PROSIT processing software in a non-automatic mode, substantial sensitivity in R_{rs} can result from the choice of subsurface extrapolation intervals. For PROSIT, and in the NASA bio-optical protocols (Mueller et al., 2003b), it is recommended to first use a red wavelength to assign the extrapolation interval. In the case of the 2012 bloom it can be seen that the red wavelengths are rapidly attenuated with depth. This is likely a result of the high levels of biomass that were present in the surface waters during this sampling period. As such, resultant R_{rs} can be highly variable depending on the extrapolation interval assigned (Fig. 4.9(d)).

The use of free floating, fixed depth or profiling radiometers is beneficial to avoid deployment platform shading and other forms of error associated with *in situ* radiometry (Hooker et al., 2002; Hooker and Maritorena, 2000). However this can can limit the amount of time spent sampling at any single point which can exacerbate a number of errors (Zibordi and Voss, 2010). The C-OPS is specifically designed to maximise sampling in the upper water column as its sinking rate is density (and therefore depth-) dependent. The Satlantic H-TSRB by comparison samples for (in this case) three minutes at a continuous depth (-0.66 m). Whilst the profiling action of the C-OPS should allow for better depth resolution of attenuation (and hence a more accurate derivation of K_{Lu} with depth), it does mean that any single measurement at a single depth may be more susceptible to errors from wave focussing which are typically dealt with through increased sampling time (Hooker et al., 2002; Zibordi et al., 2004a; Zibordi and Voss, 2010). In the case of high biomass waters, where much of the signal is confined to the upper metre of the water column, this may mean discarding large amounts of crucial data using tilt/roll filtering or that these measurements may be severely affected by wave focussing.

Additionally, this surface drift can potentially result in mismatches between *in situ* point measurements of [Chl *a*] and those sampled by the radiometer. This may be particularly important in high biomass situations, where blooms can be very patchy on horizontal scales of less than 50 m (Lucas et al., 2014; Matthews et al., 2012; Pitcher et al., 1992a, 2010). This is also particularly important for the C-OPS radiometer, as it is free to drift both horizontally and vertically. It is possible that at least some of the differences in the C-OPS L_u spectra in figure 4.8 can be attributed to this, especially under the high biomass case, where patches were visible in the colour of the surface waters during sampling (Fig. 4.9). Whilst this is not in and of itself a problem, if the 3D location is taken into account, it is likely to introduce error in any algorithm response when compared to point measurements of [Chl *a*] or D_{eff}, which may not be a representative average of the waters sampled with the radiometers. Matthews et al. (2012), used the relative standard error in the 710 nm band across the three minute sampling burst for the H-TSRB in high biomass water samples in the Benguela as an approximation for patchiness, given the correlation between this band and high [Chl *a*]. Results indicated a 10.8% error over sampling time. This may be more crucial when considering the greater spatial averaging that occurs when comparing to satellite measurements made at 1km or greater scales.

The differences between the R_{rs} measurements derived from the C-OPS and Satlantic H-TSRB can give an indication of the severity of the errors involved in the measurements. Under the low biomass conditions in 2011, the R_{rs} spectra from the C-OPS and Satlantic H-TSRB are similar (Fig. 4.9(a)). In fact two of the C-OPS derived spectra are more similar to the Satlantic derived spectrum than they are to the third C-OPS spectrum. This indicates that several potential forms of error are likely relatively minimal in these conditions. Firstly, the C-OPS results are corrected for self shading using PROSIT, whilst the H-TSRB results are not. However, the similarity of the spectra suggests that this form of error is likely relatively minor, although it is possible that it is compensated for by another form of error. Patchiness on both

horizontal and vertical scales could potentially compensate for differences, however as the results from the individual C-OPS casts are relatively similar and also similar to the results from the Satlantic, it seems that this is unlikely. In the high biomass case (Fig. 4.9(b)), the differences between the R_{rs} derived from the C-OPS casts are substantial, and all but one are substantially different from the H-TSRB result. The results from these casts do not follow a typical spectral shape that would be expected from a high biomass bloom (see examples simulated in chapter two and further examples of high biomass spectra in chapter five). This suggests that the longer, fixed depth sampling time used by the Satlantic may be more beneficial for sampling high biomass water. Additionally, this shows the final R_{rs} from the C-OPS may be very sensitive to the choice of extrapolation scheme which is assigned on a cast by cast basis in PROSIT. Setting a shallow extrapolation interval was necessary to achieve the result from the C-OPS cast that appears most similar to those from the H-TSRB.

Self shading errors were investigated, with a particular focus on the Satlantic H-TSRB by Leathers et al. (2004, 2001). Using Monte Carlo simulations, they showed that the magnitude of errors from instrument self shading were dependent on the sun position (with small solar zenith angles (θ) causing greater shading errors) and in-water IOPs. Errors were estimated as ranging between $>$ 20% and 50 % where absorption coefficients were $>$ 0.2 m $^{-1}$ and $>$ 0.65 m $^{-1}$ for wavelengths $>$ 600 nm and $>$700 nm respectively. Increased scattering (b) relative to absorption, was found to reduce errors (Leathers et al., 2001). To prevent the need for Monte Carlo simulations to derive error for instruments, a model was presented which deviated from the simulations at high absorption, small sun position angle and shallow depths (*ibid*). Applying the model presented in Leathers et al. (2001) to the Satlantic H-TSRB measurements in this chapter presents a number of difficulties. Firstly, application of the model requires a determination of total absorption (a) and, if a is large and θ is small, an estimation of total scattering (b). For the dataset here, *in situ* a is frequently large, as a result of the high [Chl a] biomass associated with many of the samples. As such, application of the Leathers et al. (2001) model to the data in this chapter would likely benefit from accurate estimation of a and b. Although some measurements of a were available for the samples in the validation dataset, coverage was not complete and no *in situ* scattering data were available. When ancillary data are lacking Leathers et al. (2001) recommend the use of an analytical approach for deriving a and b (if a is large). The EAP approach, as an analytical framework, allows for the derivation of a and b (if using EcoLight-S). To extract these from the inversion experiments and subsequently correct the initial input L_u is admittedly circuitous, given that the initial L_u(z) is used to estimate a, b_b and, in the case of EcoLight-S, b. Nonetheless, the ranges of a and b associated with the *in situ* and simulated dataset can be used as a first attempt at a self shading correction; and to explore the consequences on the EAP inversion approach.

Correction of L_u(z) according to the methods of Leathers et al. (2001) requires several stages and these were applied as follows to create a corrected set of L_u(z):

1. Estimation of $a(\lambda)$. For the purposes of this error analysis, a from the simulated data-

set (using SCI, diatom/dinoflagellate basis vectors and EcoLight-S) is used, with context provided by the *in situ* absorption values in the validation dataset.

2. Estimation of $b(\lambda)$. Leathers et al. (2001) state that this is only important where $a(\lambda)$ is large and solar zenith angle θ is small ($< 30^o$). $a(\lambda)$ is large throughout the *in situ* validation dataset, exceeding 7 m^{-1}, far higher than the maximum of 1 m^{-1} used by Leathers et al. (2001). Solar zenith angle (θ) ranged between 21 and 57^o.

3. Use of $a(\lambda)$ and $b(\lambda)$ to interpolate table one in Leathers et al. (2001) to estimate the direct sunlight ($\epsilon_{sun}(\lambda)$) and diffuse skylight ($\epsilon_{sky}(\lambda)$) errors. Table one in Leathers et al. (2001) shows variations in $\epsilon_{sky}(\lambda)$ values at different b/a, a, and for $\epsilon_{sun}(\lambda)$, θ values. To cover the variability of these parameters in the *in situ* validation dataset, data from Leathers et al. (2001) table one were extracted for $\epsilon_{sun}(\lambda)$ at θ of 20^0 and $b/a = 1$, 2 and 4. For θ of 30, 40, 50 and 60^0, $\epsilon_{sun}(\lambda)$) was only estimated by Leathers et al. (2001) at $b/a = 2$; beyond this the effect of increasing b/a ratios was deemed insignificant. Diffuse skylight errors ($\epsilon_{sky}(\lambda)$) were extracted for $b/a = 1$, 2 and 4. Linear regression was then used to determine the equation of a best fit line through the extracted data to relate $\epsilon_{sun}(\lambda)$ and $\epsilon_{sky}(\lambda)$ to a at the various θ and b/a combinations; a total of seven parameterisations of $\epsilon_{sun}(\lambda)$ and 3 parameterisations for $\epsilon_{sky}(\lambda)$.

4. Calculation of total spectral errors (ϵ). Total spectral error ($\epsilon(\lambda)$) was calculated by Leathers et al. (2001) using equation 4.4 below where $\epsilon_{sky}(\lambda)$ is the direct sunlight error, $\epsilon_{sky}(\lambda)$ is the diffuse skylight error and f is the ratio of downwelling irradiance from skylight to that from direct sunlight. Values for f on the west coast of South Africa were obtained from Walters et al. (1985).

5. Correction of spectral L_u(z, λ) using total spectral error ($\epsilon(\lambda)$): Corrected radiance was calculated using equation 4.5 below where $L_{u(corr)}$(z, λ) is the corrected radiance, L_u (z, λ) is the original measured radiance from the Satlantic H-TSRB and ϵ is the total spectra error derived from equation 4.4.

$$\epsilon(\lambda) = \frac{\epsilon_{sun}(\lambda) + \epsilon_{sky}(\lambda)}{1 + f} \tag{4.4}$$

$$L_{u(corr)}(z, \lambda) = \frac{L_u(z, \lambda)}{1 - \epsilon} \tag{4.5}$$

Two examples of corrected L_u(z) spectra are shown, along with the components of the ϵ calculation (equation 4.4) in figures 4.10 and 4.11. Due to the discretised nature of the parameterisations used to calculate the error (ϵ), some spectral artifacts occurred when calculating and applying the total spectral errors (ϵ). (see Fig. 4.11 for example). Similarly, high errors in the blue resulted in some negative values of the 1- f term in equation 4.5. These artifacts occurred where L_u(z) values were very low, as a result of the high absorption which was likely responsible

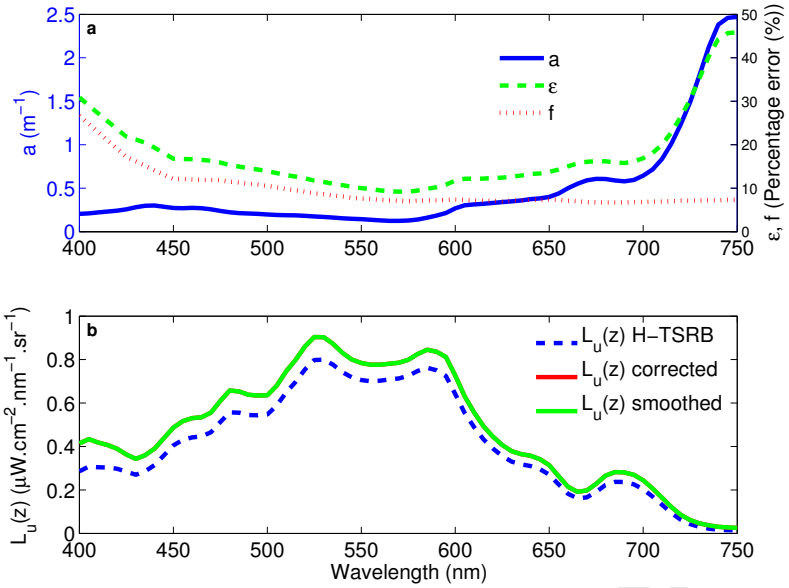

Figure 4.10: Application of the Leathers et al. (2001) correction scheme to L_u(z) from a Satlantic H-TSRB cast taken in a *Mesodinium rubrum* bloom ([Chl *a*] = 18.2 mg m^{-3}) on 29th February 2004. Panel (a) shows total absorption associated with the sample, the total error (ϵ) and the diffuse sky error (f) from Walters et al. (1985). Panel (b) shows the original L_u(z) from the Satlantic H-TSRB, the corrected L_u(z) and final smoothed L_u(z). In this case, no smoothing was necessary, thus the smoothed and corrected L_u(z) are identical.

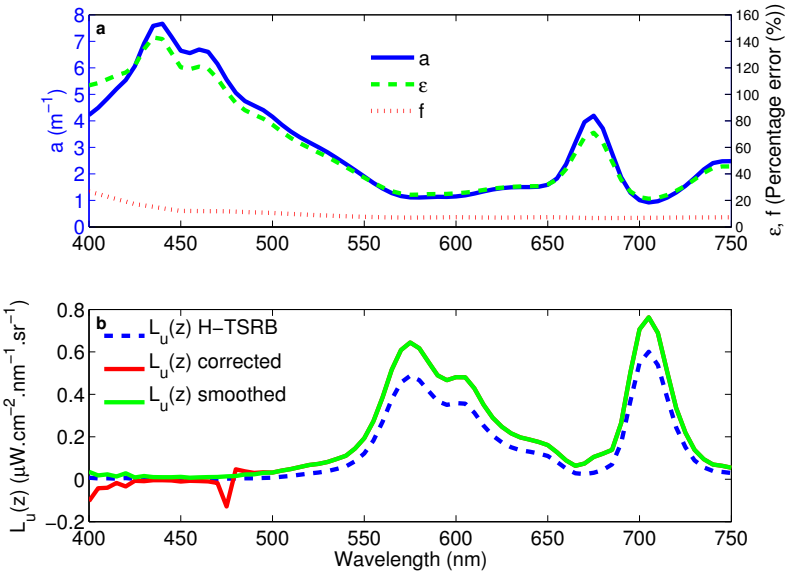

Figure 4.11: Application of the Leathers et al. (2001) correction scheme to L_u(z) from a Satlantic H-TSRB cast taken in a *Prorocentrum triestinum* bloom ([Chl *a*] = 103 mg m^{-3}) on the 29th March 2005. Panel (a) shows total absorption associated with the sample, the total error (ϵ) and the diffuse sky error (f) from Walters et al. (1985). Panel (b) shows the original L_u(z) from the Satlantic H-TSRB, the corrected L_u(z) and final smoothed L_u(z). In this case, smoothing was necessary, thus the smoothed and corrected L_u(z) are different.

for the high errors. In these cases a filter was introduced to smooth these spectral artifacts (as in Fig. 4.11). Matthews et al. (2010) report similar difficulties applying the correction scheme of Leathers et al. (2001) to Satlantic H-TSRB casts from the hypertrophic Zeekoevlei lake, opting

to apply a smoothed error function at blue wavelengths.

Mean spectral self shading errors estimated using the Leathers et al. (2001) method applied to the Satlantic H-TSRB L_u(z) dataset are shown in figure 4.12. Errors show significant spectral range (12 - 102 %) with maximum errors in the blue and red where absorption is highest. As reflectance values are typically low in these regions, the absolute magnitude of the corrected L_u(z) is still small.

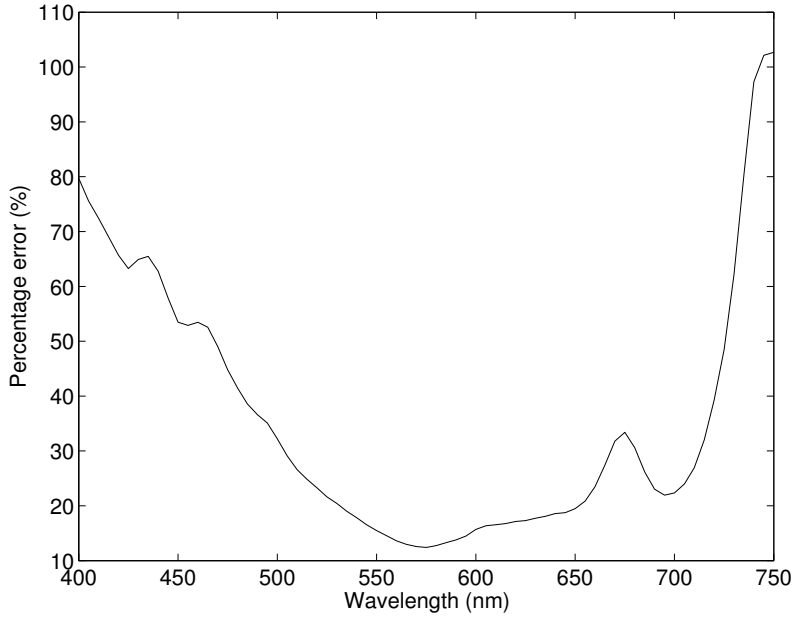

Figure 4.12: Mean percentage difference between the original Satlantic H-TSRB L_u(z) and the corrected L_u(z) using the Leathers et al. (2001) method.

Values of a from the *in situ* validation dataset and those simulated in chapter two, greatly exceed those used in the study by Leathers et al. (2001). Ultimately this can result in high estimates of ϵ_{sun}, ϵ_{sky} and total error (ϵ). However, b/a are also higher than those accounted for by Leathers et al. (2001) and show substantial spectral variance. Mean b/a estimated from the EcoLight-S *in situ* output shows a spectral range between 1.2 (lowest in the red) and 12.2 (highest in the green). The use of the two-layered sphere, size based IOPs, may impart substantially more phytoplankton backscattering in the EAP approach than in other analytical models (Robertson Lain et al., 2014), particularly in the case of relatively high biomass blooms dominated by small cells (for example the *Aureococcus anophagefferens* bloom sampled in the validation dataset). However, it should also be considered that the b_{bs} component of the EAP model is simple and unconstrained which may account for additional scattering. Regardless, although Leathers et al. (2004) suggest that the effect of b becomes insignificant at higher θ (above 30^o), this may warrant further investigation for water types similar to those dealt with in this work.

To assess the effect of self shading errors on the retrieved biogeochemical parameters, the EAP inversion approach, using SCI, diatom/dinoflagellate basis vectors, simplex method, static initial conditions and reflectance approximation was applied to the corrected L_u(z) data. Small

decreases in the r^2 values for [Chl a] and D_{eff} (0.82 and 0.35 respectively) were found compared to the uncorrected data (0.87 and 0.51 respectively).

Beyond self shading, as a floating buoy the Satlantic H-TSRB is subject to tilt and roll, which can influence the accuracy of radiometric measurements. In the absence of accurate tilt and roll measurements, the data cannot be filtered, as is done for the C-OPS. However, the substantially longer sampling time at consistent depth should contribute to reducing this error (Zibordi et al., 2009a, 2012).

Comparison of the Satlantic and C-OPS performance to above-water radiometric instrumentation should be made to investigate the most accurate method of making radiometric measures in high biomass and eutrophic waters. Errors introduced by the in-water method of the deployment will need to be considered against errors inherent in above water methods i.e. correction for fluctuations in skylight reflectance introduced by surface waves (Hooker et al., 2002). An additional consideration is the logistics involved in collecting in-water versus above water measurements. Fixed depth, in-water systems have formed a major part of ocean colour calibration and validation efforts as they can be deployed for substantial amount of time e.g. in the case of the MOBY and BOUSSOLE buoys (Antoine et al., 2008; Franz et al., 2007) or easily moved and deployed from boats, as in the case of the Satlantic H-TSRB system used here and in other studies (Chang et al., 2003; Garcia et al., 2005; Leathers et al., 2004, 2001; Matthews et al., 2010; Schallenberg et al., 2008; Wang et al., 2005; Zhang et al., 2010)). However long term deployment can be limited in high biomass waters due to increased biofouling (Pitcher et al., 2011; Pitcher and Weeks, 2006; Zibordi et al., 2004b). This can be avoided with above water systems. However, for long term deployment of above water sensors, stability of platform becomes a concern, which may require the availability of, or a large investment in infrastructure e.g a stable, offshore tower (Zibordi et al., 2004b).

A three dimensional sea surface model, coincident spectral IOPs and Monte Carlo simulations are likely needed to fully characterise the spectral errors introduced by interactions between instrument sampling specifications, self shading and wave action under different IOP conditions (see discussion in Zaneveld et al. (2001)). Whilst application of these approaches is not possible in the scope of the current study and with the data currently available, they are highly recommended for future validation of new remote sensing platforms and to refine measurement and algorithm uncertainties for management purposes, especially given the paucity of such work in the literature including high biomass waters similar to the southern Benguela.

4.4 Discussion

4.4.1 Consideration of EAP performance compared to that of other algorithms

The results above show that the EAP inversion approach has substantial utility for the estimation of [Chl a] associated with both lower biomass waters and high biomass blooms in the southern

Benguela. However the success of the algorithm is highly dependent on the optimal choice of inversion technique and underlying basis vectors. Most algorithms for derivation of [Chl a] are applied to R_{rs}. The EAP algorithm (at least as applied here) uses L_u(z), and subsequent calculation of R_{rs} is thus dependent on the IOP choice and the effect of this on radiative transfer methods (i.e. through selection of f/Q or K_u parameterisation, or more fully in EcoLight-S). This prevents application of other algorithms to the *in situ* data in a fully independent way. However, the range successfully retrieved is substantially larger than any offered by standard satellite ocean colour products (see Matthews et al. (2012) for further discussion of these products in the southern Benguela). The EAP [Chl a] results compare favourably with those of the Maximum Peak Height (MPH) algorithm applied in the southern Benguela (Matthews et al., 2012). Though direct comparison is not possible in this instance as the MPH is applied to Rayleigh and gaseous corrected, Level 1 satellite data, it appears that the EAP is less prone to overestimation, particularly in high biomass conditions. The EAP method can also give an indication of the broad size structure of a community through the effective diameter (D_{eff}) estimates, which are likely most useful for discriminating between assemblages dominated by smaller or larger cells under bloom scenarios. In the EAP method, a_ϕ and $b_{b\phi}$ are coupled through D_{eff} as representative of the underlying phytoplankton particle population model. This is a novel approach for which there are no direct comparisons in the literature. Chapter five will further assess the performance of the EAP approach compared to other [Chl a] algorithms through application to Level 2 MERIS R_{rs} data, where the very important consideration of atmospheric correction must be accounted for. Further, chapter five will consider the validity of this approach compared to other methods for determining information about phytoplankton community structure.

4.4.2 Errors attributable to IOP model and inversion ambiguity

The [Chl a] and D_{eff} performance results, combined with those for a_ϕ suggest that, despite converging to a modelled reflectance solution that is proximate to that used as input, some of the inversion techniques may fail to find a realistic solution in terms of the [Chl a] and/or D_{eff}. This indicates that there is indeed substantial ambiguity in the range of reflectance signals measured in the Benguela, at least with regards to the current set of parameterisations in the model used here. The results are most similar to those exhibited in the previous chapter when low a_{gd} and low b_{bs} conditions were simulated. However, slightly poorer results are seen in the *in situ* inversion here versus those from the simulated data in chapter three. This suggests that the ambiguity errors are likely compounded by the additional error inherent in *in situ* measurements; the likely greater range of natural variability in IOPs not accounted for by the IOP model; variability in solar zenith angle; and other unaccounted for factors i.e. surface roughness.

Whilst the behaviour of the EAP forward model and inversion approach have shown that some second order variability in radiometric measurements can be attributed to size, other

sources of variability not currently included in the model may cause additional scatter in retrieval success. The inclusion of any parameterisations (represented by unknowns to be solved for) must be balanced against the potential for increased ambiguity (particularly for inverse modelling). The EAP approach used here has existing capacity for using multiple algal IOP functional groups, although additional basis vectors could also be included e.g. to capture optical signatures associated with unique morphometric features such as vacuoles (Matthews and Bernard, 2013). However, an assessment is needed as to whether this merely increases ambiguity without any subsequent increase in retrieval accuracy. The *in situ* application of the algorithm in this chapter, informed by the sensitivity study in chapters two and three can offer some insight on this. Chapter two indicated that substantial changes in R_{rs} could be induced using the cryptophyte basis vectors versus the diatom/dinoflagellate basis vectors. However the extent of these differences was wavelength dependent and influenced by [Chl a] and the presence of competing sources of optical variability (in this case represented by increased a_{gd} or b_{bs}). In chapter three, it was shown that the two algal groups could be distinguished to a reasonable degree (up to an r^2 of 0.7) by the simplex technique under idealised conditions where each group was used to create simulated data in isolation. However, performance statistics (in terms of [Chl a] and D_{eff} retrieval in particular) were poorer in these cases, suggesting that this additional degree of freedom in the inversion technique possibly causes an increased likelihood of finding a close (radiometrically) solution which may not be associated with accurate [Chl a] or D_{eff}. In the *in situ* application here, the use of only the cryptophyte basis vectors appears to limit inversion potential. As there are only relatively few samples (≈ 10) containing cryptophyte endosymbionts (e.g. *Mesodinium rubrum*) at substantial biomass, it is likely that the cryptophyte basis vectors are not appropriate for the majority of the samples in the data, which were mostly dominated by diatoms/dinoflagellates. Additionally, chapters two and three showed that the impact of functional group differences in IOPs (at least according to this model) are biomass dependent, meaning that in some cases, they may not be distinguishable. This is consistent with other studies which have suggested that many algal species may have generally indistinguishable absorption spectra (Dierssen et al., 2006). Including basis vectors representative of spectrally distinctive pigments, such as those associated with cryptophytes and cryptophyte endosymbionts may improve retrieval for specific samples, especially at elevated [Chl a] when these spectral signatures can be distinguished from those associated with other species. However, in terms of ambiguity, it may be prudent to constrain the use of these basis vectors. For example, basis vectors could be preselected through empirical techniques to identify spectrally distinctive pigments or features associated with highly scattering intracellular structures e.g. vacuoles (Matthews and Bernard, 2013).

Several assumptions in the two-layered model used to derive the IOP basis vectors for the EAP approach may also introduce scatter and explain some sources of ambiguity. The use of a constant or variable c_i may result in under/over expression of size related, optical characteristics. Although no explicit analysis is possible here, given the limited data on intracellular c_i for the

field samples, the use of a static c_i assumption does produce the most accurate results for [Chl a] and D_{eff}. The assumption of constant Standard size distributions, whilst appropriate for mono-specific blooms, may be less relevant for samples where a background of smaller cells dominate (Bernard et al., 2007; Reynolds et al., 2010). The use of the JCI basis vectors represents an attempt to incorporate alternative size distributions on the basis of known ecological context e.g. the dominance of smaller cells at lower [Chl a] (Agustí et al., 1987; Raven, 2009; Raven and Kübler, 2002). However, this does not improve performance across the dataset. As a large proportion of the data used was obtained under bloom conditions it is likely that a Jungian distribution would be inappropriate. The approach may become more suitable as part of a low [Chl a] switching algorithm for application to mixed low and high biomass conditions. The two-layered sphere model used here has been shown to be appropriate for broad application to high biomass blooms in the southern Benguela (Bernard et al., 2009; Robertson Lain et al., 2014).

Finally, the optical properties of living algae and the extent of their degradation products are known to be influenced by physiology, life stage, and response to growth conditions, including the light environment (Dierssen et al., 2006). Though there is currently no explicit accommodation for this in the two-layered model, other bio-optical measurements (e.g. fluorescence products linked to physiology) and ancillary environmental data (e.g. bloom phenology, nutrient levels) could indicate when these factors may play a role in inversion error and ambiguity (Defoin-Platel and Chami, 2007). Further constraint of the inversion method with regards to the above factors and IOP co-variance may reduce the likelihood of finding local as opposed to global minima (Defoin-Platel and Chami, 2007; Rehm and Mobley, 2013).

4.4.3 Radiative transfer considerations and application to satellite data

Assumptions made when solving or approximating solutions to the radiative transfer equation (e.g. f/Q or K_u parameterisations) can introduce significant errors in the radiometric quantities estimated (Mobley, 2011). However, for coupled modelling and inversion approaches, compromises in accuracy must be optimised against gains in computational efficiency. Here, using EcoLight-S to invert *in situ* measurements, results in a substantial increase in computational time ($\approx$ 300 fold), versus using the REFA method. Whilst for the limited number of measurements in the *in situ* validation dataset, this is not unpractical, scaling up the inversion approach for eventual application to satellite-derived time series of ocean colour data requires further consideration of whether any accuracy is gained. The sensitivity study conducted with the forward model (chapter two), revealed several differences between the two methods, which may become increasingly important in other water types (e.g. Case 2/highly scattering waters). Chapter three revealed that the introduction of EcoLight-S for inversion did not necessarily increase accuracy of results under low a_{gd}/low b_{bs} conditions. This is likely as a result of increased ambiguity from the increased number of non-unique solutions available as a result of the less constrained, IOP dependent, bidirectionality and attenuation in the EcoLight-S radiometric calculations (Defoin-Platel and Chami, 2007). These factors are more constrained in the

reflectance approximation approach through the f/Q and K_u parameterisations. However, for highly scattering waters, using EcoLight-S improved inversion retrievals from the simulated data (chapter three). The *in situ* analysis allows a consideration of where the Benguela fits within the simplified water types simulated in chapter three and how this influences the assumptions made with regards to radiative transfer.

Figure 4.13 below compares the f/Q values selected by the inversion method over the *in situ* set for (a) the reflectance approximation method using the parameterisation of Morel et al. (2002) and (b) f/Q derived from EcoLight-S. The most obvious differences are the greater spectral variability and range of values shown in the EcoLight-S results (Fig. 4.13(b)), compared to those available through the Morel et al. (2002) parameterisation (Fig. 4.13(a)). Whilst the Morel et al. (2002) f/Q parameterisation is dependent on [Chl *a*] and solar zenith angle, the EcoLight-S derived f/Q is further dependent on the total absorption and scattering determined by the algorithm including any size related influences, presence of gelbstoff/non algal particles etc. Further, the model simulations used to generate the f/Q parameterisation of Morel et al. (2002) have a number of differences to the EAP approach, which may influence the suitability of the f/Q parameterisation in the context used here. For example, Morel et al. (2002) use a two-component model of phase function to account for small and large particles (2 - 14 μm) with assumed Jungian distributions of homogenous spheroidal particles. However, due to the nature of the southern Benguela and St Helena Bay region (i.e. phytoplankton are the dominant optical constituent), it is likely that the differences in magnitude between the f/Q used in the reflectance approximation and those calculated from EcoLight-S are largely due to the far greater range of [Chl *a*] associated with the *in situ* dataset, than was used in the model runs (0.03 to 10 mg m^{-3}) which form the Morel et al. (2002) parameterisation. As may be expected due to the limited spectral resolution of the Morel et al. (2002) parameterisation scheme (interpolated for hyperspectral application here and in chapter three), the f/Q values from Morel et al. (2002) (Fig. 4.13a) show substantially less spectral shape than those calculated from EcoLight-S (Fig. 4.13b). In particular, elevated values in the spectral region of the fluorescence and backscattering induced peaks (685-709 nm) are not captured in the Morel et al. (2002) parameterisation as a result of the wavelengths selected for the modelling employed to generate the parameterisation, and the low range of [Chl *a*] simulated. Clearly there is significant signal at these wavelengths in high biomass waters such as the southern Benguela. Additionally, whilst the modelling undertaken by Morel et al. (2002) to generate the f/Q tables used in the REFA approach involved a [Chl *a*] dependent phase function with allowances for small and large particles, the study admits that this necessitates assumptions about particle size and shape. In the EcoLight-S variant of the EAP forward model and inversion, the spectrally variant Fournier-Forand phase function is dependent on the variability in backscattering probability inferred by the basis vectors. However, sensitivity of the approach to discretised phase functions has not been conducted and is necessary to determine the extent to which the model is sensitive to the detailed spectral backscattering component (Robertson Lain et al., 2014).

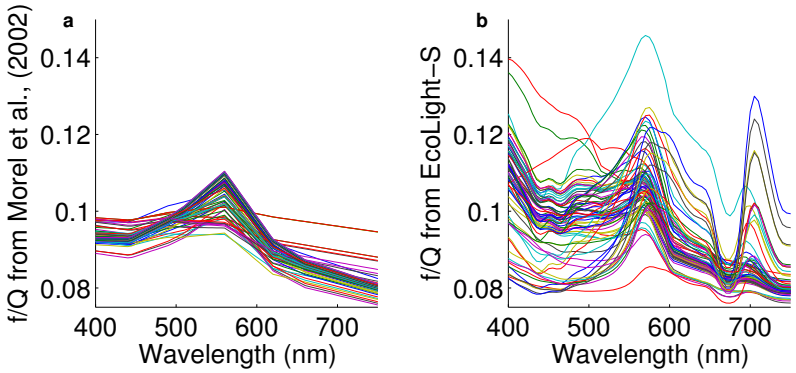

Figure 4.13: Spectral f/Q values over the validation dataset derived from (a) the Morel et al. (2002) parameterisation used in the REFA approach and (b) EcoLight-S. Results here using the inversion algorithm with SCI, diatom/dinoflagellate basis vectors and static initial conditions.

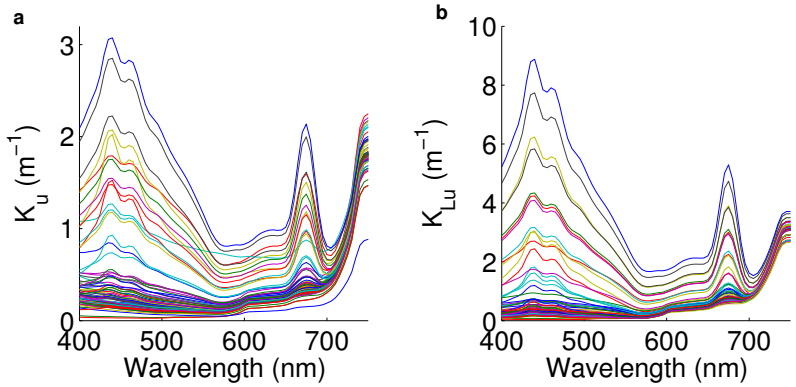

Figure 4.14: Spectral K_{Lu} values over the validation dataset derived from (a) the parameterisation of Albert and Mobley (2003) used in the REFA approach and (b) EcoLight-S. Results here using the inversion algorithm with SCI, diatom/dinoflagellate basis vectors and static initial conditions. Note the different scales on the y axes.

Whilst the spectral shapes are broadly similar, the specific K_{Lu} values from EcoLight-S (Fig. 4.14(b)) are substantially higher than those derived from the K_u parameterisation of Albert and Mobley (2003)(Fig. 4.14(b)). In particular, K_{Lu} values from EcoLight-S are substantially greater in the 450 nm and 685 nm centred peaks than the K_u derived from Albert and Mobley (2003). It is worth noting however that the highest K_{Lu} values are associated with high biomass examples ([Chl a] >100 mg m^{-3}), where the unconstrained inversion algorithm has selected unrealistically small cell size (< 5 μm) and as such the accompanying K_{Lu} values maybe also be unrealistic. Although the K_u parameterisation is dependent on sample specific absorption and backscattering, provided in the EAP approach by the [Chl a] and size dependent basis vectors, the radiative transfer model used by Albert and Mobley (2003) to parameterise the relationship between these IOPs and Ku may not be optimal to fully represent the impact of these on factors such as K_u or K_{Lu}. The model employed by Albert and Mobley (2003) was based upon typical Case 2 IOPs and employed a spectrally flat specific scattering to backscattering coefficient of 0.019 as a phase function. By comparison the variant of EcoLight-S used in this chapter (and all variables derived from it, such as the f/Q and K_{Lu} shown in figures 4.13 and 4.14) are dependent

on the [Chl a] and size dependent IOPs and employs a spectrally variant Fournier-Forand phase function to account for the spectrally variably backscattering probability apparent from the two-layered sphere model (Bernard et al., 2009). This has been found most suitable for generating suitable modelled R_{rs} for high biomass conditions in the Benguela (Robertson Lain et al., 2014).

To first order, the disparities between the EcoLight-S derived f/Q and K_{Lu} and those from the parameterisations employed in the reflectance approximation may seem substantial. However, when the associated R_{rs} measurements are compared (Fig. 4.15) the differences between the REFA and ES derived R_{rs} are relatively minimal with a few notable exceptions. On the whole the magnitude of the R_{rs} from the two approaches is comparable although the main reflectance peak (550 - 570 nm) is in general associated with slightly higher values when EcoLight-S is used. Similarly, R_{rs} values associated with the red peak centred around 709 nm are elevated in EcoLight-S compared to the REFA approach. Further, the EcoLight-S R_{rs} shows higher values and more spectral shape at blue wavelengths (400-470 nm).

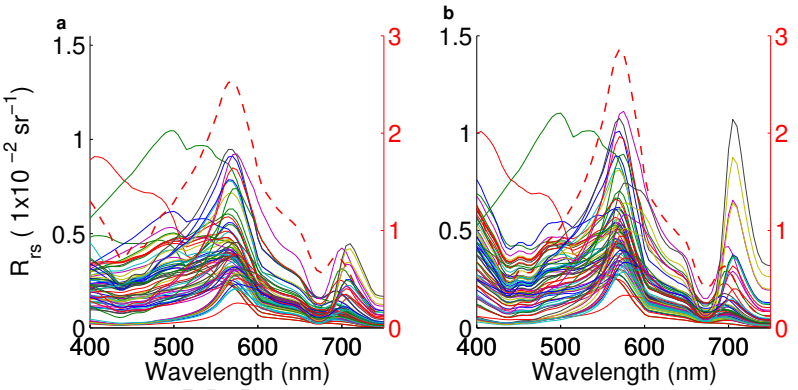

Figure 4.15: R_{rs} spectra output derived from the EAP inversion algorithm applied to the validation dataset using (a) the REFA approach and (b) EcoLight-S. The second axis in red provides scaling for the sample taken in an *Aureococcus anophagefferens* bloom (red dashed line). Results here using the inversion algorithm with SCI, diatom/dinoflagellate basis vectors and static initial conditions.

To further isolate the influence of using the K_u parameterisation from other assumptions in the process of radiative transfer, the K_u values from the Albert and Mobley (2003) parameterisation and K_{Lu} from EcoLight-S were applied to the measured L_u (z) and E_d from the Satlantic to generate R_{rs} spectra (Fig. 4.16). The results further suggest that the use of the K_u parameterisation has relatively little impact on the eventual R_{rs} spectral magnitude with the majority of examples showing a ±20% variability between the K_{Lu} from EcoLight-S and the Albert and Mobley (2003) K_u parameterisation, resulting in a ±20% variability in R_{rs}. Several examples show spectrally variant differences in K_{Lu} of up to ± 100% between those from EcoLight-S and the Albert and Mobley (2003) K_u parameterisation with resultant differences in R_{rs} of up to 100%. In the case where the measured L_u and E_d are directly used (Fig. 4.15), the elevated peaks seen in figure 4.15 in the red remain in the case where an EcoLight-S K_{Lu} is applied directly to the Satlantic $L_u(z)$ (Fig. 4.16) and are thus likely caused by the differences in the REFA vs ES approaches attributable to the handling of K_u. The higher values in the

blue are reduced (Fig. 4.16(b)) and are therefore likely dependent on the bi-directional effects characterised by the f/Q (Fig. 4.13). A number of strange features are apparent in the blue and red when the EcoLight-S derived K_{Lu} values are applied. Whilst the EAP model has no fluorescence term included, there is a fluorescence related signal present in the Satlantic H-TSRB L_u measurements, as such, these features interact with the application of the EcoLight-S K_{Lu} in the exponential format used to calculate R_{rs} from L_u and E_d. Similarly, jagged peaks are seen at blue wavelengths between 400 and 500 nm in figure 4.16(b). Small variations in the low values of L_u between different wavelengths are magnified in the R_{rs} by the exponential K_{Lu} function. These features are absent from the EcoLight-S R_{rs} as EcoLight-S does not directly use these K_{Lu} values in its radiative transfer calculations (Fig.4.15(b)). The results here suggest that although it may seem wise to create new K_{Lu} specific parameterisations which are more suitable to the IOP suite used in the EAP approach and that are present in waters such as the southern Benguela, care must be taken to avoid introducing unrealistic features such as those seen in figure 4.16(b). However given the relatively minimal differences in R_{rs} and the success of the EAP approach in returning suitable biogeochemical and IOP values compared to the validation dataset, it seems that the inversion approach is relatively insensitive to errors introduced by the use of the K_u parameterisation to deal with subsurface extrapolation. Indeed, in the case of the REFA applications, it seems that these parameterisations may offer a constraint that reduces ambiguity, compared to full radiative transfer calculations, although this is unintentional.

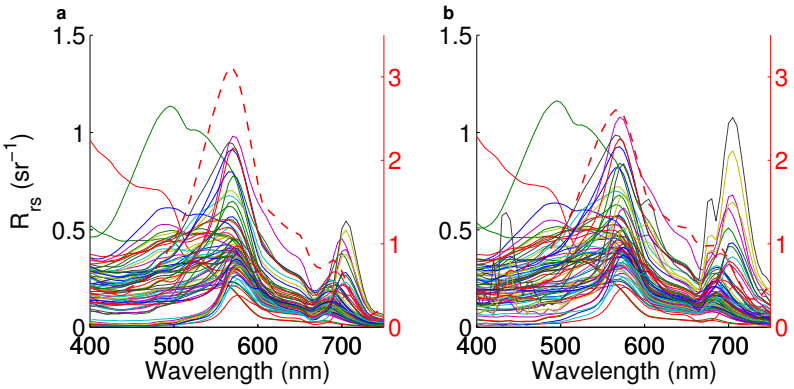

Figure 4.16: R_{rs} spectra from the validation dataset using (a) Albert and Mobley (2003) derived K_u and (b) EcoLight-S derived K_{Lu} applied to the measured L_u (-0.66m) from the Satlantic. The second axis in red provides scaling for the sample taken in an *Aureococcus anophagefferens* bloom (red dashed line). Results here using the inversion algorithm with SCI, diatom/dinoflagellate basis vectors and static initial conditions.

It has been suggested that errors in radiative calculations can arise when assumptions made associated with single scattering are invalid and/or the single scattering albedo (ratio of scattering to total attenuation) is greater than 0.6 (Aurin and Dierssen, 2012). Comparison of f/Q values from EcoLight-S to those from Morel et al. (2002) in chapter two suggested that as scattering increases, the f/Q factors used in the REFA approach, become increasingly inadequate. Estimates of the single scattering albedo derived from the EcoLight-S variant of the EAP inversion algorithm suggest that a number of samples exhibit spectrally variant values above the value

of 0.6 cited by Aurin and Dierssen (2012). This appears somewhat contrary to the assumption that (in terms of IOPs) the southern Benguela is typically dominated by highly absorbing phytoplankton biomass, rather than sediments or other optically significant constituents. However, as the non-algal particle scattering estimated by the inversion approach has not been validated, the formulation used is relatively simple and any errors would propagate to the full scattering output from EcoLight-S, these values may not be reliable. Assessing the validity of the single scattering assumption (and subsequent consequences for measurement errors and radiative transfer) in the southern Benguela will require further measurements of angular scattering.

4.5 Summary of errors and uncertainties and impact on size related sensitivity in the context of global ocean colour measurements

Table 4.2 provides potential values for ranges of key sources of error associated with the Satlantic data used in this validation dataset. These values were selected from the literature as most appropriate where indication was given with regards to higher biomass/absorbing waters and informed by the discussion above e.g. the upper bound for wave focussing error was selected. Where errors were suggested to be spectrally variable either a range or the maximum reported value is given. See Zibordi and Voss (2010) and Zibordi et al. (2012) for a detailed overview of error sources and estimated magnitudes. It must however be further emphasised that all studies reviewed within this chapter were performed either under measured or modelled [Chl *a*] far lower than the samples from the Benguela, and often in Case 2 waters with high concentrations of gelbstoff and sediments.

Important in the context of ocean colour algorithms, is the extent to which size related variability in reflectance falls within the error associated with measurements, either *in situ* or satellite derived. The error estimates here (and from the literature) can be compared to the sensitivity derived in chapter two to give an estimate of the size related signal in reflectance that exists outside of these error ranges. Figure 4.17 shows the range of potential error associated with low biomass and high biomass R_{rs} spectra (expressed by error bars), if total, spectrally constant errors of 5% in reflectance are assumed for low biomass, optically clear waters (as is the common standard for ocean colour radiometry missions), and of 30% for higher biomass cases based on an average of the errors sources described above. This is admittedly an oversimplified approach given the likely spectral variance of errors and difficulty in meeting these error targets even with high quality validation data (Zibordi et al., 2006a,b, 2004b). However, this exercise should serve to illuminate a "best case" scenario and can be easily expanded with the EAP forward model in future, given more detailed *in situ* error characterisation. The modelled data here suggests that, at the 5% error level, it is possible that the majority of size related variability at low biomass (0.1 - 1 mg m^{-3}) is contained within the potential errors and uncertainties associated with R_{rs} measurements. At biomass levels $>$ 3 mg m^{-3} size related changes in R_{rs}

Table 4.2: Summary of error sources and range estimates associated with Satlantic H-TSRB measurements

Error Source	*Percentage Error*
Calibration and general instrument error	2.8%
Wave focussing	up to 20%
Propagation error through K_u	20%
Bidirectionality assumptions (f/Q parameterisations)	3%
Self shading	up to 30%

begin to exceed the errors associated with R_{rs}, however at this point the assumption of 5% error may be becoming inappropriate (Zibordi et al., 2004b). By contrast, even withstanding the higher errors likely associated with more optically complex, high biomass waters, size related variability in R_{rs} exceeds that which may reasonably occur as a result of error in R_{rs} for most of the spectral range modelled. However, small changes in size (i.e. $\pm\ 5 - 10\mu$m) could still potentially be contained within these ranges, particularly if errors are larger than estimated here. Hence the EAP approach is likely to be most effective when assemblage characteristics (in this case, size) shift from one extreme to another (i.e. dominated by small ($\approx$ 2 μm e.g. *Aureococcus anophagefferens* or large (+ 30 μm e.g. *Ceratium furca* cells). It should be noted that these modelled scenarios additionally represent a hypothetical best case with regards to cell size detection for both low biomass and high biomass waters, as the modelled spectra here exclude substantial variability of other phytoplankton characteristics (i.e. accessory pigments) and of non-phytoplankton constituents. Presence of these may further obfuscate detection of size related signals in R_{rs}, dependent on the impact of these characteristics on the underlying IOP budget.

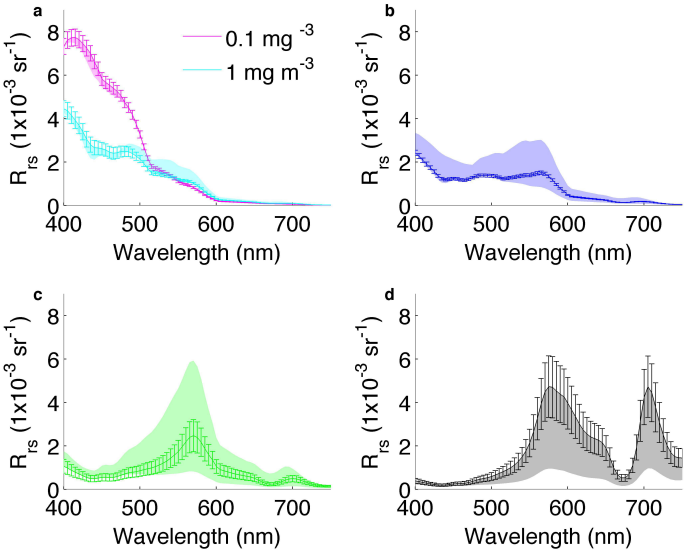

Figure 4.17: R_{rs} spectrum from the forward model with associated error estimates for [Chl a] of (a) 0.1 mg m^{-3}, (b) 3 mg m^{-3}, (c) 10 mg m^{-3} and (d) 100 mg m^{-3}. The solid line with error bars represents the spectra generated with a D_{eff} of 10 μm. Error bars were applied on the basis of a 5% spectral error for low biomass (<10 mg m^{-3}) and 30% for high biomass (>10 mg m^{-3}). The shaded range shows the range of R_{rs} associated with changes in size (from a minimum of 2 μm to maximum of 40 μm). A filter has been applied to remove unrealistic spectra i.e. dominance of large cells (+20 μm) at low [Chl a] (<10 mg m^{-3}) and small cells (<10 μm) at high [Chl a] (>30 mg m^{-3}). See further discussion of this in chapter two. Results here using the forward model with SCI, diatom/dinoflagellate basis vectors, low a_{gd} and low b_{bs} and EcoLight-S.

4.6 Conclusions

Application of the EAP inversion approach to an *in situ* validation set from the southern Benguela yields encouraging results for the estimation of the full range of chlorophyll a concentrations associated with high biomass blooms in this region. The approach also shows potential for distinguishing changes in cell size in phytoplankton communities using the effective diameter parameterisation, coupled to size dependent inherent optical properties. There are however a number of conditions to the success of the approach and several recommendations can be made:

- As seen in the inversion experiments using the simulated data (chapter three), the Nelder-Mead simplex, with static c_i assumption and a static set of initial conditions, provides the best returns for optical and biogeochemical parameters across the dataset. These results indicate a high capacity for the estimation of a_ϕ and [Chl a] and an ability to detect size which will likely be most useful for distinguishing between blooms where cells have considerably different size characteristics, under high biomass conditions.

- The use of EcoLight-S does not improve the result from the reflectance approximation (REFA) approach. This is likely because of increased ambiguity, due to comparatively unconstrained variability in bidirectionality and attenuation with depth in Ecolight-S. The REFA approx likely presents fewer non-unique solutions due to the relatively more

constrained f/Q and K_u parameterisations.

- EcoLight-S substantially increases computational expense, and given the lack of improvement of algorithm performance for the independent validation dataset used here, it would seem unnecessary to apply it to inversion of satellite imagery for the southern Benguela. However it may be argued that a new parameterisation of the f/Q and K_{Lu} factors would be appropriate and EcoLight-S should certainly be used for any further investigation of impact of IOPs on f/Q, K_{Lu}, phase functions etc in a more systematic manner than the preliminary analysis presented here.

- Additional work is required to investigate and minimise the errors for in-water radiometric sampling in high biomass waters, particularly for satellite validation efforts. This may further improve the capacity of the EAP approach for deriving IOPs, [Chl a] and D_{eff}. Comparison between the Satlantic and C-OPS radiometers indicates that care must be given towards length of sampling period at depth in these waters to minimise effects of wave focussing and other errors which are more prevalent in the upper water column, where biomass is typically concentrated in these cases. The use of above-water radiometry for high biomass waters should also be considered for observations in high biomass and eutrophic waters, to avoid errors associated with in-water systems which are exacerbated in high biomass waters, such as self shading.

Chapter 5

Phytoplankton biomass and cell size in the southern Benguela from remotely sensed satellite ocean colour data: Application to the study of interannual variability of Harmful Algal Blooms.

Output images from this chapter, produced by the thesis author, were contributed to Bernard et al. (2014). An adapted version of the pixel extraction and the processed Maximum Peak Height algorithm data were also contributed to Pitcher et al. (2014).

5.1 Introduction

Whilst *in situ* observations allow for detailed understanding of bio-optics and phytoplankton dynamics, satellite data can allow for routine and cost effective production of synoptic scale information about phytoplankton communities. If suitable algorithms can be developed, satellite ocean colour imagery allows for the observation of biomass variability over scales that are inconceivable with even the most extensive *in situ* monitoring scheme. Over the last few decades, as the archive of satellite ocean colour data has expanded, much research has attempted to use these data to extract more detailed information about phytoplankton populations, beyond the use of chlorophyll *a* concentration ([Chl *a*]) as a proxy for biomass.

5.1.1 Regional and global algorithms for satellite ocean colour data

Determination of chlorophyll pigment concentration ([Chl *a*]) from water leaving radiance has been the primary objective of most satellite ocean colour missions to date. This has been achieved to some degree of accuracy (± 35%) over the global ocean using empirical band ratios (McClain, 2009). However, substantial regional variability in performance has been observed (Bailey and Werdell, 2006; Gregg and Casey, 2004; McClain, 2009). The most commonly used empirical algorithms rely on the bio-optical assumption that "ocean optical properties should co-vary with chlorophyll concentration" (Morel and Maritorena, 2001; Siegel et al., 2005; Smith and Baker, 1978). This holds in Case 1 (usually open ocean) waters, however in Case 2 (often coastal) waters, high concentrations of coloured dissolved organic matter (CDOM or Gelbstoff), particulate organic matter, and suspended sediments may represent a significant component of the optical signal (Morel, 2009; Morel and Prieur, 1977). Empirical algorithms based on the bio-optical assumption are typically derived from regression analysis of satellite and *in situ* variables. These relationships form the basis for most global [Chl *a*] algorithms in use today (e.g. the OC3 (MODIS) and OC4 (SeaWiFS) algorithms (O'Reilly et al., 1998), and the Algal 1 pigment index (MERIS) (Morel and Antoine, 2000)). However, even in Case 1 waters, large regional variations in algorithm performance have been observed. This has been attributed to various influences, primarily involving variable success in atmospheric correction and natural variations in Inherent Optical Properties (IOPs) (McClain, 2009; Morel and Maritorena, 2001; Prieur and Sathyendranath, 1981; Sauer et al., 2012).

Regional algorithm development typically begins with collection of *in situ* data for that specific region. Existing algorithms can then be validated and re-parameterised if necessary. Whilst this may provide the most accurate empirical parameterisations for the region in question, it can limit application of these algorithms in systems with different optical characteristics. Several large databases have emerged to provide data for algorithm development and validation on a global scale e.g. the NASA bio-Optical Marine Algorithm Dataset (NOMAD) (Werdell and Bailey, 2005), the SeaWiFS Bio-optical Archive and Storage System (SeaBASS) (Hooker et al., 1994), MERis MAtchup In-situ Database (MERMAID) (Barker et al., 2008) and the Geochemistry, Phytoplankton and Color of the Ocean (GeP&Co) pigment database (http://www.lodyc.jussieu.fr/gep- co). From these, a variety empirical algorithms can be developed, relating features in reflectance to biogeochemical parameters.

In some cases these databases also contain *in situ* IOPs. This allows for both parameterisation and validation of semi-analytical algorithms (see chapter one, figure 1.3). Semi-analytical algorithms can incorporate some of the variability in IOPs, which has been identified as a source of second order variability in empirical relationships (Sauer et al., 2012). Increased theoretical understanding of IOP variability has lead to extensive development of semi-analytical algorithms (e.g the Generalized Inherent Optical Property (GIOP) framework (Franz and Werdell, 2010)), particularly for application in Case 2 waters (e.g. Doerffer and Schiller (2007)). However, the extent to which IOPs used in semi-analytical algorithms capture natural variability is still

dependent on how well resolved the underlying biogeochemical causes of this variability are. Chapter two discussed how phytoplankton characteristics such as cell size and absorption by specific pigments influence IOPs. Databases such as those listed above are rarely accompanied by routine information on phytoplankton cell size and other characteristics, so investigation of global relationships between IOPs and phytoplankton characteristics is currently limited.

Coherent regional and global scale datasets of IOPs and AOPs calculated from a variety of models have also been produced to aid algorithm development (Chami and Defoin-Platel, 2007; Chami and Robilliard, 2002; Chang et al., 2007b; Chang and Whitmire, 2009; Defoin-Platel and Chami, 2007; IOCCG, 2006). Informed by *in situ* measurements, simulated datasets have been generated using sophisticated radiative transfer models such as Hydrolight (Mobley and Sundman, 2008) and others (Chami and Robilliard, 2002). The extent to which variability in AOPs associated with phytoplankton characteristics is captured through these models varies and is dependent on the choice of input IOPs and often governed by availability of *in situ* data. The International Ocean Colour Coordinating group (IOCCG) generated a simulated dataset with extensive coverage of the natural variation observed in the specific absorption coefficient $a_\phi(\lambda)$ for phytoplankton. The absorption coefficients were normalised to $a_\phi(440)$ in nine groups separated by values at $a_\phi(440)$ and [Chl a] range (Maritorena et al., 2006). The simulated data were then used to investigate the performance of a number of different algorithms, detailed in the various chapters of IOCCG report number 5 (IOCCG, 2006) and has since been used in a number of other studies (Chang et al., 2007a; Defoin-Platel and Chami, 2007). However, currently no simulated dataset (with the exception of that produced in chapter two of this thesis) explicitly characterises the impact of phytoplankton characteristics such as size or accessory pigments on coupled IOPs and ultimately R_{rs} across different optical regimes. The use of IOP models which capture this variability, can offer substantial advantages where size and other phytoplankton characteristics influence the IOP budget i.e. in eutrophic conditions (Robertson Lain et al., 2014). The future use and expansion of IOP models such as these would benefit substantially from the availability of coincident information on IOPs and phytoplankton characteristics including particle size distributions and functional type classifications.

5.1.2 Sources of uncertainty in satellite derived ocean colour products

When using satellite ocean colour products for analysis of ecological variability, understanding and quantifying the associated errors is of substantial importance. Without estimates of errors in the products used, it can become difficult to resolve the differences between biogeochemical variability of the target system and random or systematic observational error created by instrument, processing and methodological errors. In particular the signal-to-noise ratio should be considered as trends may be obscured within error and ambiguity as discussed in the preceding chapters. Ideally estimates of error should be derived at each stage of the processing chain to include, for example, those from atmospheric and other corrections. For applications of remotely sensed products in ecosystem modelling (e.g. to calculate primary production/carbon export),

errors associated with reflectance and IOP/AOP/[Chl *a*] or other derived parameters need to be known before errors associated with modelling applications can be added.

Validation represents the first step necessary to quantify regional uncertainties in remote sensing data. This process itself however, is not without error due to the vastly different temporal and spatial scales over which *in situ* validation data and ocean colour remote sensing data are collected. Satellite ocean colour measurements represent a surface water column average, dependent on optical depth, over spatial scales of between 250m and 1km, whereas *in situ* data are often taken over discrete depth samples at a single point which may or may not be representative of the average pixel captured by the satellite. The errors introduced by this mismatch in spatial and temporal variability are inherently difficult to quantify. Useful measures of "patchiness" may help provide some indication of the level of uncertainty associated with spatial heterogeneity. For the southern Benguela, previous studies have made estimates of patchiness associated with using *in situ* sampling and estimate that this may be in the order of 11% (Matthews et al., 2012). Quantification of errors in satellite radiometry, can only be as accurate as the errors associated with the *in situ* radiometry provided as validation data. Chapter four revealed the differences that choice of radiometer can have on spectral water leaving radiance and derived products such as remote sensing reflectance (R_{rs}).

Satellite radiometry measurements are subject to additional sources of uncertainty. Calibration is made more complex due to the remoteness of the instruments themselves, although over the course of the last decade's worth of missions, newer, more advanced calibration procedures have been implemented. Measurements of top of atmosphere radiance require correction to remove the atmospheric component of the signal to derive upwelling, water-leaving radiance (L_w). Many methods exist for this and developments have been made with the increased and varied wavebands available on newer satellites (IOCCG, 2010; McClain, 2009). Retrieval of accurate water-leaving radiance and quantification of error can be complicated by influences such as white cap waves, sun glint and sun angle (Yoder and Kennelly, 2006) and, in shallow waters, by bottom reflectance. For global missions, error estimates for open water (Case 1) L_w are $\pm$ 5% (McClain, 2009). However, validation exercises have found that this can vary substantially (Antoine et al., 2008; Mélin et al., 2007; Zibordi et al., 2011, 2004b), particularly in coastal regions. Case 2 and coastal waters represent additional complications for the process of atmospheric correction. Due to the masking effect of high particulate/CDOM waters and close proximity to the land (adjacency effect) and associated aerosol input, new methods have been sought to address atmospheric correction in coastal waters (Matthews et al., 2012; Moore et al., 1999; Schiller and Doerffer, 2005; Zhu et al., 2012). Appropriate flagging schemes can help identify pixels likely to be influenced by errors with processing or algorithm application. Analysis and comparison of flags and underlying reflectances can hint at processing errors that may be corrected to obtain more accurate products and prevent compounding of errors if these data are used to derive further products.

The ways in which bio-optical data and their derived products are used also has the potential

to add error to results. Derived products are available averaged over multiple days and can be composed of data from multiple sensors. Level 3 and merged products are often used to avoid issues with cloud cover and to consider seasonal biogeochemical inventories or climatologies (IOCCG, 2007). Errors can be introduced in this averaging process, and the sources of error potentially missed (*ibid*). For example, depending on the percentage of cloud cover over the pixels of interest, the averages may not represent real averages, rather just the short timescale variability that was captured during sparse clear days. Ultimately, any gaps in data, or the processes of averaging, may impact the scientific conclusions made, particularly with regards to long term trends or seasonal variability (Cole et al., 2012; Henson et al., 2010).

5.1.3 Approaches and limitations of deriving phytoplankton size from satellite ocean colour data

Ocean colour data, including reflectance and [Chl a] products have been used in a growing body of literature to derive information about phytoplankton communities and their behaviour. Chapter one briefly summarised some of the major methodologies used in recent times to examine phytoplankton cell size or functional type (PFT) including: 1. Spectral-response, 2. Abundance, and 3. Ecological approaches (Brewin et al., 2011b). These methods have been the subject of the most recent International Ocean Colour Coordinating Group report (IOCCG, 2014). Whilst many cell size and PFT approaches have achieved significant results, there exist a number of limitations in their use which can begin to be addressed through application of semi-analytical models and inversion algorithms of the structure presented in chapters two to four, at a satellite scale.

Some spectral approaches for deriving PFT information have empirically related features in normalised water leaving radiance (nL_w) to concentrations of pigments associated with different phytoplankton species (Alvain et al., 2005). Radiative transfer simulations have subsequently been used to provide some theoretical explanation for the anomalies in nL_w and how these relate to phytoplankton absorption, absorption from coloured dissolved organic matter (gelbstoff) and scattering from non algal particles (Alvain et al., 2012). Further empirical relationships have used the principles of abundance to link satellite derived [Chl a] with broad size classes derived from databases using diagnostic pigment analysis (Uitz et al., 2006). These relationships have a theoretical understanding in principles of abundance (Agustí et al., 1987; Irwin et al., 2006), where low biomass tends to be dominated by small cells and only large cells contain enough [Chl a] to occur at maximum density and account for high biomass.

Other studies have used the output of semi-analytical inversion algorithms, applied to satellite data, to link features in IOPs to different PFTs and size classes. Hirata et al. (2008) developed a model relating size classes determined on the basis of diagnostic pigment analysis to the slope of the phytoplankton absorption coefficient at 443 - 510 nm. This relationship was then used to associate classes with phytoplankton absorption coefficient slopes derived from Sea viewing Wide Field of view Sensor (SeaWiFS) data using the semi-analytical algorithm of Smyth

et al. (2006). Ciotti et al. (1999) and Ciotti et al. (2002) developed a semi-analytical approach to estimating cell size which also drew on the principles of abundance used by Uitz et al. (2006). The shape of the phytoplankton absorption coefficient was ultimately used to distinguish small from large assemblages through application to SeaWiFs data in Ciotti and Bricaud (2006).

Although there has been much focus on absorption, as the dominant IOP where phytoplankton biomass has an influence; backscattering has also been used to infer information about the size of phytoplankton assemblages. Kostadinov et al. (2009) related the particle backscattering coefficient derived from SeaWiFS data using the semi-analytical algorithm of Loisel et al. (2006), to the parameters of a power law particle size spectrum to summarise the particle size distribution. Results were then used to infer information about phytoplankton size classes, in regions where the particle size distribution is dominated by phytoplankton (Kostadinov et al., 2009, 2010). Outside of the open ocean, lower biomass trophic paradigm (i.e. in eutrophic and hypertrophic conditions), phytoplankton backscattering may provide further information about phytoplankton cell size and other characteristics (Matthews and Bernard, 2013; Robertson Lain et al., 2014).

Semi-analytical inversion algorithms have been designed with variable phytoplankton absorption coefficients to take into account the natural variability observed in this parameter (Bricaud et al., 1995a, 1988; Chang et al., 2007a; Lee et al., 2002; Sathyendranath et al., 2001; Tilstone et al., 2012). Cell size is well known to influence a_ϕ and a number of approaches have sought to exploit this (e.g. (Ciotti et al., 2002)). However, none of the current suite of semi-analytical inversion algorithms incorporates size structure as a factor influencing the underlying basis vectors/Specific Inherent Optical Properties (SIOPs) in an explicit and coupled way i.e. so that [Chl *a*] and cell size are related to both phytoplankton absorption and scattering simultaneously.

5.1.4 Scales of variability in the Benguela

As an upwelling system, characterised by frequent, high biomass blooms, dominated by a variety of functional types, the Benguela represents an exciting region to test and develop ocean colour methods. The Benguela is a highly dynamic system with phytoplankton community response to physical and biological forcing occurring across a broad range of temporal and spatial scales (Hutchings et al., 2009; Pitcher et al., 1992b). Routine observations from remote sensing platforms allow these responses to be investigated. Besides the potential for developing new techniques, the Benguela suffers from Harmful Algal Blooms (HABs) which threaten the extensive fishery and aquaculture industry within the region, making monitoring a necessity. The occurrence of large blooms of both toxic and non-toxic species of phytoplankton has had extensive impacts on fisheries in the St Helena Bay region of the southern Benguela - particularly the rock lobster (*Jasus lalandii*) (Cockcroft, 2001) and mussel farms (Pitcher et al., 2011). Two primary impacts can arise from HABs in the Benguela. Firstly, HABs can contain phytoplankton species which produce toxic compounds, which can cause mortality of marine organisms and secondary impacts through human consumption of affected organisms, including various types of shellfish

poisoning and death (Pitcher and Calder, 2000; Probyn et al., 2000). Secondly, the presence of vast volumes of phytoplankton biomass can cause oxygen depletion and mass mortality of marine organisms, when the bloom decays (Pitcher and Calder, 2000). Predicting whether anoxic conditions will arise and/or whether blooms contain toxic species or not is consequently of high interest to fisheries managers in this region (Bernard et al., 2006).

Monitoring the Benguela to understand ecosystem productivity for fisheries purposes and to predict and manage impacts from HABs, requires an understanding of the scales of variability in this complex ecosystem. Variability in phytoplankton biomass in the Benguela exists over a large range of temporal and spatial scales in response to variable physical and biological forcing mechanisms. Fully observing this variability requires observational platforms which measure appropriate parameters at these scales. Whilst wind-driven upwelling is the primary influence on phytoplankton communities in the Benguela, variability in this wind forcing can be observed at seasonal, event (3-6 days) and inter-annual scales, eliciting a complex array of responses in phytoplankton bloom extent and community structure when acting in combination with local coastal morphological and bathymetric features and biological succession (Figueiras et al., 2006; Lucas et al., 2014; Pitcher et al., 2010; Smayda, 2000).

Phytoplankton blooms in the Benguela are not evenly distributed synoptically within the region. Studies of satellite derived chlorophyll have suggested that the fundamental frequencies of variability differ significantly between the northern (15 S - 27.5 S) and southern Benguela with the northern Benguela exhibiting an almost opposite seasonal response (Demarcq et al., 2007). However this trend of opposition did not hold for all years included in the study mentioned and the potential for long term interannual variability of these spatial and seasonal trends warrants further investigation incorporating newer satellite data. At the seasonal scale, phytoplankton biomass in the southern Benguela follows a pattern associated with periods of stratification during austral summer-autumn, mixing during winter and upwelling during spring (Weeks et al., 2006). Chlorophyll *a* concentrations broadly follow this trend, showing generally higher values and frequent "red tides" and HABs (often in excess of 100mg m^{-3}) following spring upwelling and persisting through the stratified summer conditions until the onset of winter mixing abruptly reduces values (Pitcher et al., 1992a). Between upwelling cycles (typically 3-6 days) chemical and biological factors (e.g. nutrient use and grazing controls) can determine bloom extent and species succession (Pitcher et al., 1992a; Smayda, 2008).

Within this general trend there exists high frequency variability associated with wind variability and species succession (Fawcett et al., 2007; Pitcher et al., 1991). Accordingly seasonal trends in oxygen depletion and toxicity from species have been observed (Pitcher and Calder, 2000; Pitcher and Probyn, 2011). Anoxia is driven by both the physical process of advection of oxygen depleted water into in-shelf regions and oxygen demand driven by bloom decay linked to prolonged quiescence in the upwelling cycle (*ibid*). Cold water with low oxygen is generally observed during summer/autumn, whilst in winter, bottom water in the Benguela has been observed to be much warmer and well ventilated. Significant variability exists around these

general trends, characterised by event scale upwelling of different deep water sources advected onto the shelf and the extent of subsequent phytoplankton bloom events. The observed seasonality of phytoplankton blooms is accompanied by a change in the dominant species present. The transition from winter, well mixed conditions to summer/autumn upwelling conditions sees dinoflagellate species generally dominate over diatoms as the seasons progress and the increasing stratification favours bloom conditions (Pitcher et al., 2002; Pitcher and Nelson, 2006).

Superimposed upon the general seasonal trends described above are sub-regional and sub-seasonal trends and events. Maximum biomass in the Benguela is typically found in the southern Benguela, within the greater St Helena Bay region (Pitcher and Probyn, 2011; Weeks et al., 2006) (Fig. 5.1). Increased stratification as a result of the broader shelf, and the retentive circulation in the bay, promotes this high biomass which is frequently dominated by species of dinoflagellates (Pitcher and Nelson, 2006).

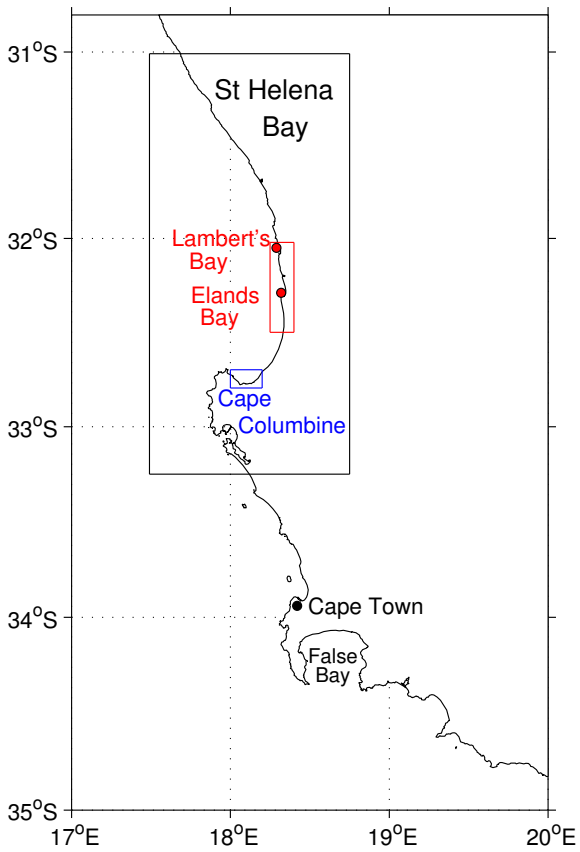

Figure 5.1: Schematic map of the study region within the southern Benguela, showing the greater St Helena Bay area, with Elands Bay, Lambert's Bay and Cape Columbine highlighted. The red box shows the extent of the region used for determining averages for the Elands/Lambert's Bay subregion and the blue box shows the corresponding subregion defined as Cape Columbine.

Within St Helena bay, blooms are often transported south by trapped coastal waves or wind stress driven flows (Fawcett et al., 2008; Pitcher and Nelson, 2006; Probyn et al., 2000). These flows result in a southerly net bloom transport, even against South Easterly winds (Pitcher and Nelson, 2006). This transport can cause bloom waters to become trapped in inshore regions around Cape Columbine, causing anoxic events (Pitcher and Probyn, 2011). These flows can also result in "flood events" where blooms are advected around Cape Columbine and into small bays further down the coast (e.g. Saldahna Bay (Probyn et al., 2000)).

In addition to affecting bloom transport, the physical dynamics of the St Helena Bay region have been linked to phytoplankton community structure dynamics (Fawcett et al., 2007). Inshore, more turbulent waters are typically dominated by diatoms during upwelling. However it has been hypothesised that during quiescent or downwelling periods, dinoflagellates are advected from the offshore region as a result of upwelling front collapse. Further, a typical successional pattern over the upwelling season from diatom to dinoflagellate dominance has been observed. This is supported by the prevalence of high biomass, dinoflagellate dominated HABs later in the upwelling season (March-May) over the full extent of the wide shelf around St Helena Bay when stratification is most prevalent (Pitcher et al., 1998; Pitcher and Calder, 2000). However, it must be noted that this dominance and abundance cannot be explained by local growth alone and thus communities are likely concentrated by physical processes (Pitcher and Nelson, 2006).

Observation of the particle size distributions between the inshore, frontal and offshore regions of St Helena Bay suggest a shift in volume dominance from large cells to small cells is seen across these regions progressively (Armstrong et al., 1987; Pitcher et al., 1992a). However, it must be noted that the selection and proliferation of individual species and functional types is stochastic and reliant on other factors beyond the physical environment; including the presence of a seed population, nutrient availability and top down, biological controls including grazing and viral infection (Smayda, 2008; Smayda and Reynolds, 2003).

5.1.5 Application of ocean colour to understanding Harmful Algal Blooms in the southern Benguela

5.1.6 What is a "bloom"?

Though discussions of phytoplankton blooms abound throughout the literature, defining what constitutes a "bloom" in quantitative terms is not simple. If considering the occurrence of phenomena which have been called "blooms", events are encountered which can vary extensively in their temporal and spatial characteristics (Blondeau-Patissier et al., 2014; Smayda, 1997b). On a seasonal time scale, the annual mid-high latitude growth of phytoplankton in the spring is arguably the canonical example of a bloom. However, in many coastal regions, intra-seasonal events known as Harmful Algal Blooms (HABs) or red tides are also defined as blooms. Definition of an event as a phytoplankton bloom therefore likely depends on the system of interest. A primary interest in mid-latitude and non-coastal regions has been the phenology of the bloom in relation to physical and chemical dynamics and grazing pressure. A number of methods have

been proposed to quantify its spatial and temporal extent (see reviews by (Cole et al., 2012) and Brody et al. (2013)). In these mid-high latitude regions, the timing of bloom initiation, peak biomass and decline have potential consequences for higher trophic levels and carbon export dynamics. As such, metrics have been designed, evaluated and applied to satellite data to evaluate these parameters in both space and time (Brody et al., 2013; Cole et al., 2012).

Phytoplankton blooms have long been characterised by the presence of elevated biomass over a background threshold. In the seasonal bloom case, a "bloom initiation date" has often been defined as the first occurrence of a chlorophyll *a* concentration or growth rate which exceeds a given threshold above an annual average (Brody et al., 2013; Henson et al., 2009; Henson and Thomas, 2007; Platt et al., 2009; Siegel et al., 2002). A number of studies have discussed the exact threshold/growth rate implemented, whether this must be exceeded for a minimum amount of time and how the average seasonal cycle is defined e.g as a median/mean. Variations of this metric haveonto been successfully applied to a number of regions throughout the global oceans (Brody et al., 2013; Cole et al., 2012; Henson et al., 2009; Henson and Thomas, 2007; Racault et al., 2012; Thomalla et al., 2011). Harmful Algal Blooms have also been defined with respect to presence of species above a background concentration (Ryan et al., 2008; Smayda and Reynolds, 2001a), however this approach can be limiting for evaluating HABs depending on how these techniques are applied. Though a first occurrence of high biomass within a seasonal scale, may be useful for understanding interannual seasonal dynamics, harmful bloom events may occur multiple times within a season. Application of an initiation date in this instance may understate the importance of large events happening much later in the season. In addition, depending on the species involved, large growth in a community - arguably a bloom, may not manifest as high [Chl *a*] biomass. This is particularly the case when considering small celled species, which may grow rapidly whilst [Chl *a*] values may not rise above a seasonal average, leaving them undetectable by threshold based quantification methods. Equally, if one is interested in detection of toxic blooms, elevated biomass may be similarly unsuitable for an identification metric due to presence of toxins at relatively low biomass (Pitcher et al., 2011).

Though elevated biomass can indicate high phytoplankton growth and productivity, from a HABs management perspective, this is, in itself, not a definitive indicator of harmful consequences. As mentioned, toxic effects can occur at low biomass and impacts from oxygen related dynamics (anoxia and hypoxia), though requiring elevated biomass levels to occur, do not follow every high biomass bloom to the same extent. Oxygen depletion (seasonal hypoxia and episodic anoxia) in the southern Benguela is fundamentally dependent on the timing and duration of wind and coastal upwelling which not only influences the extent of phytoplankton productivity and formation of HABs but also the advection of oxygen depleted bottom waters and control of stratification and mixing (Pitcher et al., 2014).

From a system perspective, regions that are affected by HABs, may differ from the canonical, annual diatom bloom in terms of species succession (Smayda, 1997b). Harmful Algal Blooms can be dominated by species across different functional groups (Smayda and Reynolds, 2001a;

Smayda, 1997a). Although there is a general trend of bloom progression following wind driven nutrient upwelling in coastal upwelling systems, the spatial and temporal distribution of these patterns can be highly heterogenous (Kudela et al., 2005). Succession can be interrupted by wind patterns that can be related to both local (Fawcett et al., 2007; Pitcher and Nelson, 2006) and global forcing e.g El Nino Southern Oscillation, North Atlantic Oscillation, Pacific Decadal Oscillation (*ibid*). In addition to this, the occurrence of HABs in coastal bays means continental run off, bay bathymetry and coastal circulations can influence bloom progression (Kudela et al., 2005; Pitcher and Nelson, 2006; Probyn et al., 2000). An ideal metric for examining HAB development would incorporate ways of identifying general patterns of succession to inform model parameterisations and explain physical, chemical and biological forcing dynamics.

From the literature, several characteristics of phytoplankton blooms that may warrant quantification arise and a consideration can be made of how to quantify these in a HABs setting:

1. Elevated biomass

Most definitions of a phytoplankton bloom would include elevated biomass levels as a descriptive factor. However, it is not simple to globally characterise what constitutes elevated biomass. Outside of the techniques used to characterise spring bloom initiation, which typically involve either a growth rate, or threshold above an annual average to define a bloom, absolute thresholds have been used to indicate HABs (Ryan et al., 2008). A further complication is that HABs typically occur in coastal regions, often affected by other optically significant constituents where band ratio algorithms may fail (McClain, 2009). Even in waters where phytoplankton are still the dominant constituent, the range of [Chl *a*] associated with HABs can be beyond that from which empirical band ratio (and many semi-analytical) relationships are derived (O'Reilly et al., 1998). In high biomass, the relationship between blue and green band ratios and [Chl *a*] diminishes and greater signal is found in the red (See figure 6, (Robertson Lain et al., 2014)).

2. Species identification and succession

Appropriate observation and management of HABs requires some level of taxonomic identification for estimation of toxic and noxious effects (Bernard et al., 2006; Kudela et al., 2005). To predict bloom occurrence, progression and response to physical, chemical and biological forcing, a conceptual model of species or functional type succession, can be useful (e.g. Margalef's Mandala and subsequent variations (Margalef, 1961; Smayda and Reynolds, 2001a)). However observation of taxonomy and broad functional types at suitable temporal and spatial scales is currently very difficult either due to physical (i.e. coverage of detailed *in situ* data required) or methodological limitations (e.g. many species exhibit similar optical characteristics, which make distinguishing them extremely difficult from optical based measurements such as satellite ocean colour). However, impacts of cell size on optics have been clearly observed and may represent a useful metric for HABs modelling and monitoring, at least in terms of further characterising the

relationships between size and physical/chemical/biological controls on production. In addition to toxic effects induced by the presence of particular species in a HAB assemblage, further impacts may arise from physical characteristics of species present, which may be characterised by size, e.g. size mismatch in terms of predation, gill clogging and other forms of physical damage to higher trophic level species (Smayda, 1997b). Additionally, larger cells with a higher nitrate dependence (Irwin et al., 2006; Probyn, 1985) maybe at greater risk from nutrient depletion and thus are more likely to be associated with anoxia.

3. Timing and persistence

To understand what drives elevated biomass and the presence of certain species, cell sizes or phytoplankton functional types, a time based metric must be developed to test relationships with suspected physical/chemical or biological forcing factors. As previously mentioned, mismatches with predators (in terms of nutrient provision, size and prey density) can be impacts of HABs (Smayda, 1997b) and to quantify these, a time based metric in terms of occurrence and persistence is required to assess impacts.

5.1.7 Harmful Algal Bloom monitoring - what information and understanding is needed?

The development of models of HABs has been somewhat hampered by the lack of consistent and quantitive definitions of what constitutes a bloom and in particular those considered "harmful", "exceptional" or "rare" (Smayda, 1997b). However several general models have been proposed to summarise some typical features and driving forces, particularly in eastern boundary upwelling systems (EBUs). In these systems, the upwelling of nutrient rich water is driven by Ekman transport from alongshore winds and the development of HABs is principally linked to this (Pitcher et al., 1998, 2010). As such, interannual variability in HAB formation has also been linked to interannual variability in wind forcing, often caused by climatic scale oscillation of high and low pressure systems (Hickey, 1998). Wind driven upwelling, whilst providing nutrients to fuel phytoplankton growth, also affects the stability of the water column and consequently levels of light experienced by phytoplankton. In addition to this classical view, the movement of water in upwelling regions is influenced in multiple dimensions by advective and turbulent processes, defined by local topography and bathymetry (Smayda, 2000). Nutrient dynamics are similarly non-linearly dependent on broad-scale upwelling dynamics. Changes in nutrient ratios, speciation and the extent of remineralisation produce spatial gradients and heterogeneous distributions when combined with physical ocean variability and plankton community structure dynamics (Barlow, 1982; Shannon and Pillar, 1986).

This continuum in physical and chemical conditions, between high-low nutrient and turbulent and stratified waters, creates an abundance of ecological niches for which certain phytoplankton species are more or less suited (Margalef, 1961, 1978; Smayda and Reynolds, 2001a). Con-

sequently, species succession is often observed after wind events and can be influenced and even reset by further wind activity (Fawcett et al., 2007). Typically, in EBUs, diatom dominance occurs during winter and spring due to their ability to succeed in highly turbulent and both high and low nutrient environments. As water becomes more stratified, dinoflagellates are able to compete for high levels of nutrients with small dinoflagellates remaining as nutrient levels are reduced (Kudela et al., 2005; Margalef, 1978). Within this general seasonal trend, other species can succeed and bloom in areas with specific hydrodynamic features, or as the result of competitive advantages against other species and reaction to grazing pressures. Swimming and chain forming species fare better in areas of high turbulence, often concentrating around convergence zones (Kudela et al., 2005). These succession dynamics have been summarised by several schematic models. Margalef initially developed a mandala to describe the succession of phytoplankton species in relation to turbulence and nutrient concentrations in 1978. However, Margalef himself admitted that "red tide" species were problematic within this model and Smayda and Reynolds have since expanded the work by Margalef to account for the independence of nutrient input and turbulence and shared morphological characteristics (Smayda and Reynolds, 2001b). Margalef originally segregated phytoplankton species into two groups: "r" and "K" based upon their success under high and low nutrient and turbulence environments respectively (Margalef, 1978). Attempts to attribute HABs species to these groups proved complicated and the original mandala was suggested to be too simple to reflect the range of environmental conditions in which HABs develop, particularly in upwelling systems (Smayda, 2000). The two groups proposed by Margalef (1978) have been expanded to three (C, R and S) based on morphological traits and adaptive strategies including growth rate, size and light and nutrient stress tolerance (Smayda and Reynolds, 2001a). Within this ecological space, Smayda and Reynolds (2001a) described 9 subgroup "types" of HAB species following a continuum of decreasing nutrients and mixing and increasing euphotic zone depth. A simple schematic depicting this "life form" model is shown in figure 5.2 below with the addition of some typical HAB forming species found in the southern Benguela. General trends observed along the nutrient/turbulence and light availability gradient include increasing size (from small to large to colonies to even larger forms) and change from simple, spherical shape to longer, chain forming and ornamented species (Smayda and Reynolds, 2001a). This trend is accompanied by increased maximum abundance, increased generation time and increased prevalence for mixotrophy and endosymbiosis *in situ* (*ibid*). Two bloom sequences are proposed by Margalef's mandala and subsequent expansions - a "main sequence" and an "alternative main sequence" associated with red tides (and HABs) in upwelling systems (Fig. 5.2) (Smayda, 2000). Pitcher and Nelson (2006) discussed this mandala and the expansion by Smayda and Reynolds (2001a) in the context of the southern Benguela, suggesting that a number of species occurred across overlapping types. It was suggested that it would therefore be difficult to make distinction beyond diatoms vs dinoflagellates based on the physical oceanographic environment (Pitcher and Nelson, 2006).

Obtaining taxonomic level information can be a problem for effective HABs management.

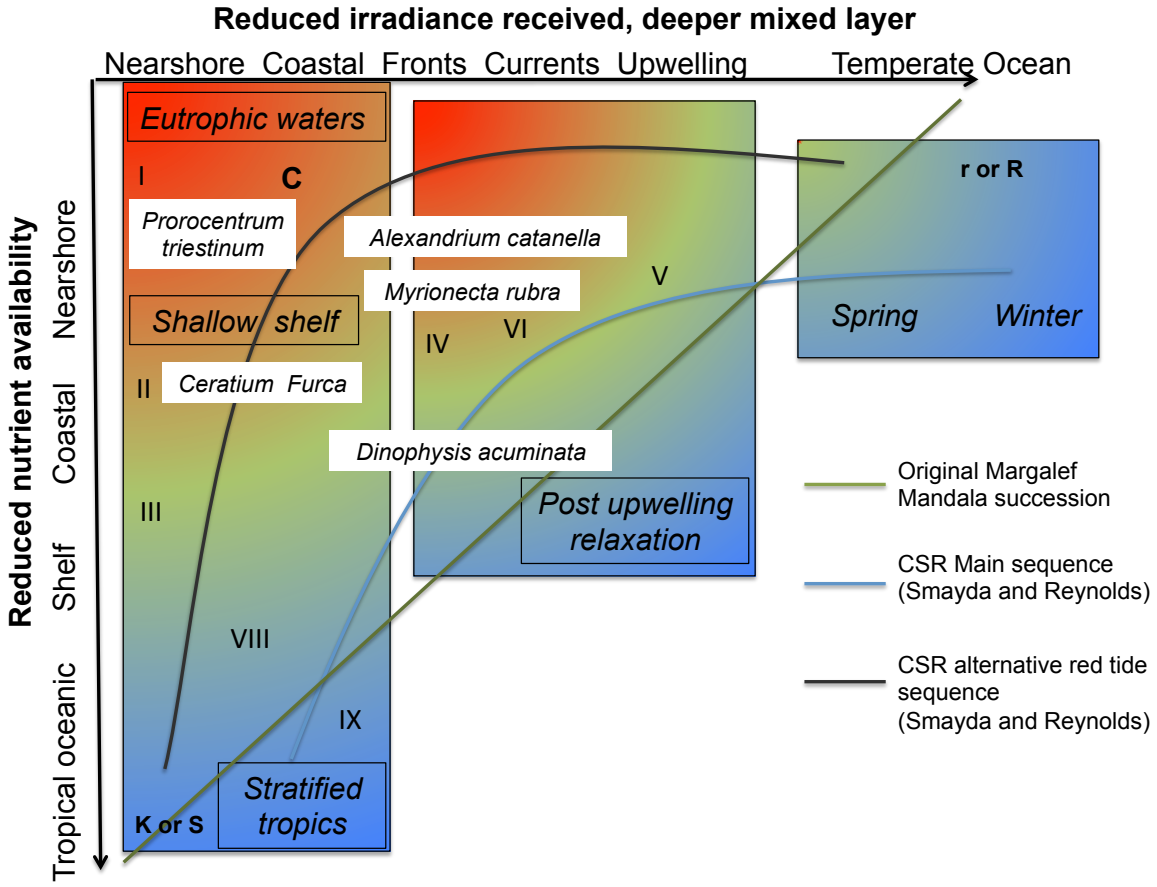

Figure 5.2: Species succession in relation to physical forcing factors showing typical regions where different types of species occur and typical succession sequences for typical phytoplankton blooms and HABs. See(Smayda and Reynolds, 2001a) upon which this figure is based for more details on species types (I-IX). Examples of species which commonly occur in the southern Benguela are given, however as noted by Pitcher and Nelson (2006), these species may occupy more broad, overlapping ranges than can be indicated in this schematic.

Typically, succession schemes for understanding HAB development have been developed at a class level, generalising broadly over the diversity of species within diatoms, dinoflagellate, ciliates etc. This appears to be a quite valid simplification, as common morphological traits within species can be linked with phytoplankton processes influenced by physical dynamics (e.g. growth rate, swimming efficacy, predation potential etc) (Smayda and Reynolds, 2001a). However, impacts resulting from toxic effects can vary widely between and within species which may be considered within the same higher taxonomic or functional type group. Predicting occurrence of harmful impacts from individual species, likely needs observations at both autecological and synecological levels. Whilst prediction of impacts from species containing toxins may require species level information, studies have suggested that overall production is unaffected by bloom species, and rather is governed by the length and intensity of upwelling during the upwelling season (Smayda, 2000). Anoxia, due to its relation to biomass decomposition, is one impact associated with Harmful Algal Blooms that may therefore result from blooms of a variety of single or mixed species assemblages. In this instance, rather than focussing on methods for

detecting individual species, the key antecedent and "in bloom" conditions that lead to anoxia must be established.

HABs are highly spatially and temporally variable phenomena. Several key questions for HABs research focus were presented by Kudela et al. (2005), covering the impact of small scale physics on HAB growth and dispersion; seeding strategies; nutrient supply and acquisition; effects of coastal morphology and bathymetry; the role of genetics and environmental conditions in toxin production; impacts of cross and along shelf advection; and use of climate indicators to predict HABs. Interannual variability in forcing from wind patterns has been well observed over the global oceans and high levels of interannual variability in upwelling communities is evident (Fawcett et al., 2007; Kudela et al., 2005; Ryan et al., 2009). However determining the extent of year to year variability in upwelling indices and the effect of this on bulk properties such as total primary production is not without difficulty. Historically, assessing the nature of phytoplankton variability in upwelling systems has proven difficult due to the inconsistent nature of ship based measurements (Thomas et al., 2004). Long term records are required to establish the existence of any interannual periodicities and these data must be of a suitable temporal and spatial scale to capture long term effects from the events scale upwards. Satellite data represents perhaps the best option to meet these conditions. Though depth limited, remotely sensed ocean colour data offers synoptic scale measurements on an almost daily basis at 1km spatial resolution - suitable for capturing the many levels of variability associated with phytoplankton blooms in the Benguela and other upwelling systems.

5.1.8 Summary and chapter aims

The development of semi-analytical algorithms represents a step towards more holistic modelling of water-light-biological interactions for the purpose of retrieving biogeochemical information about the oceans. Many methods exist for deriving phytoplankton community structure from satellite ocean colour data (see chapter one, Brewin et al. (2011b) and IOCCG (2014) for discussion). However major limitations exist in coherent application at the regional and global scales and through limited understanding of the variable optical characteristics of different regions. To begin to address these limitations, a variant of the *in situ* Equivalent Algal Population (EAP) algorithm is adapted for application to Level 2 data from the Medium Resolution Imaging Spectrometer (MERIS) and compared to other algorithms and *in situ* radiometric, [Chl *a*] and Coulter Counter derived effective diameter (D_{eff}) data for a selection of case studies. A time series is produced covering the full MERIS archive for the southern Benguela. This time series is then used to assess long term trends in phytoplankton biomass and cell size in the St Helena Bay region. Several experimental metrics are then designed to assess the event scale, seasonal and interannual variability of HABs in this region. Finally, a discussion is presented to examine ways in which these data and the metrics can be developed and used to enhance understanding of the environmental forcing of these phenomena and how the EAP algorithm approach could be extended for use in the southern Benguela and other systems.

5.2 Methodology

5.2.1 Algorithm design and application to MERIS data

For application to satellite data, the Equivalent Algal Population (EAP) algorithm design is essentially the same as for *in situ* application (see chapter four). The major difference being the lack of depth dependency and therefore need for subsurface propagation of radiance measurements. For satellite application Remote sensing reflectance (R_{rs}) is the input parameter for the minimisation procedure. To generate a time series, the entire available archive of 1km (reduced resolution) MERIS Level 2 data were searched using custom search procedures to extract all swaths containing any pixels within latitudes and longitudes predefined as the southern Benguela region (-30 to-36 ^{o} N, 15: to 20 ^{o} E). Subsets were extracted to include per pixel latitude, longitude and solar zenith values and flags using the Graph Processing Tool (GPT) available through the Basic ENVISAT Toolbox for (A)ATSR and MERIS (BEAM). Reflectance from the Level 2 data was converted to R_{rs} as per equation 13 in ESA et al. (2011). The resultant time series begins on the 25th October 2002 and ends on the 8th April 2012 shortly before the ENVISAT mission failed.

As R_{rs} is used, the forward model is reworked to remove the depth dependency associated with the Satlantic H-TSRB and associated parameterisations (i.e. the K_u model from Albert and Mobley (2003)). To account for bidirectionality, the f/Q parameterisation of Morel et al. (2002) is applied as used previously, with resampling to MERIS wavelengths (412.5 nm, 442.4 nm, 490 nm, 510 nm, 560 nm, 620 nm, 665 nm, 681 nm and 709 nm). Given the previous experience with the *in situ* application, the diatom/dinoflagellate phytoplankton absorption (a_ϕ) basis vectors were selected and the 4 component model applied without any ability to differentiate between the species/functional type basis vectors. In a few case studies, where species showing substantially different optical properties are present (e.g. *Mesodinium rubrum*), the use of this assumption is assessed by applying the 5 component model, with an admixture of cryptophyte basis vectors available to the optimisation approach. Occurrence of these species were determined through *in situ* microscopy and/or through the presence of distinctive bifurcated reflectance peaks associated with the presence of phycoerythrin pigments.

Based on the results from the *in situ* application (chapter four), a set of initial conditions including empirically derived starting values were selected. For [Chl a] an "Algal 1" to 665/709 switching algorithm is applied (Bernard et al., 2005). Algal 1 values are extracted from the original Level 2 files and used as part of a switching algorithm which reverts to an empirical relationship between 665 and 709 nm for conditions where Algal 1 suggests [Chl a] is greater than 10 mg m^{-3}. It became apparent that this approach presented a problem involving fill values in Algal 1. This was traced back to pixels where some wavelengths presented negative reflectances. These features occurred predominantly over known bloom conditions and negative reflectances were confined to the blue (as could be expected given low R_{rs} under highly absorbing bloom conditions). Whilst these negative reflectances will impact the use of empirical band ratio

approaches such as Algal 1, semi-analytical algorithms using multi-spectral R_{rs} should be able to circumvent this problem. In these cases, the 665/709 empirical [Chl a] derivation was used as initial conditions and the negative values replaced with very small, near zero R_{rs} values. An effective diameter of 10 μm was used as a starting value, as in all previous applications of the algorithm. Initial conditions for gelbstoff and detrital absorption (a_{gd}) and small particle backscattering (b_{bs}) were calculated proportional to the empirically estimated [Chl a], given that the Benguela is likely a Case 1 type system. Solar zenith values were extracted from within the subset Level 2 data files.

The Nelder-Mead simplex optimisation was selected given its consistent performance across the work conducted in chapters two to four. The optimisation was applied on a per pixel basis, with pixels skipped whenever a land, cloud, or high glint flag was raised. These flags were maintained as Not a Number (NaN) values in the resulting [Chl a], D_{eff} and other products.

The processing code was translated from Matlab to Python (2.6.5) to allow for parallelised image processing on the Centre for High Performance Computing (CHPC) cluster (CSIR, Rosebank).

5.2.2 Case study and time series extractions

Several case studies were selected, based upon examples of Harmful Algal Bloom events reported in the literature. For these cases, images shown are from the subset Level 2 data output from GPT (in the case of R_{rs}, Algal 1, flags and any subsidiary products discussed) and the raw output of the EAP algorithm for any EAP products. Matchups for R_{rs}, [Chl a] and D_{eff} were extracted from each raw image output using a spheric radius method to extract all pixels within 1km of the *in situ* sampling point (see details of this process in (Pitcher et al., 2014)). For time series analysis, the St Helena Bay region was the focus of this study due to its previously acknowledged high biomass and frequent occurrence of HABs (Pitcher and Calder, 2000). Figure 5.1 shows the study region covered by the applications within this chapter. Imagery was subset over the entire domain shown in figure 5.1, which covers most of the southern Benguela. Areas which are used to determine subregional averages are highlighted in black (St Helena Bay) red (Elands/Lambert's Bay) and blue (Cape Columbine).

5.2.3 Data from other algorithms

In addition to the EAP [Chl a] output, the full time series of Algal 1 data was extracted from the Level 2 subsets. Algal 2 was not extracted at this time, as Algal 1 and similar band ratio approaches have been most commonly used in the Benguela in previous studies (Demarcq et al., 2003; Machu et al., 1999; Silió-Calzada et al., 2008; Weeks et al., 2006) and the dynamic range of Algal 2 does not confer any substantial advantages over Algal 1 (Doerffer and Schiller, 2007). For comparison over a range more suitable for the EAP and high biomass detection, the Maximum Peak Height (MPH) algorithm of Matthews et al. (2012) was applied to the Level 1 MERIS archive. This algorithm uses bottom-of-Rayleigh reflectances to calculate a maximum peak

height from peaks at the 681nm, 709nm and 753nm bands, using a baseline between 664nm and 885nm. The peak height is then empirically related to ranges of chlorophyll concentrations observed *in situ* in South Africa's coastal and inland, high biomass waters. Full details on the simplified atmospheric correction, the MPH algorithm and its validation can be found in Matthews et al. (2012).

5.2.4 Regridding and Metric design

To assess the spatial variability of [Chl *a*] and D_{eff} in the St Helena Bay region over inter annual timescales, the data output from the EAP algorithm was regridded on to a common (1/100 ^{o}) grid using linear interpolation. This allowed for a series of metrics to be developed to investigate bloom variability over both spatial and temporal domains. The dataset was rearranged according to upwelling seasons, covering 2 years from September to September i.e. year 2004/2005 covers data from September 2004 until September 2005.

Firstly, a threshold method was developed to provide a way of classifying blooms. Threshold techniques have been a common way of defining the start of a bloom used in multiple phenology metrics (Blondeau-Patissier et al., 2014; Brody et al., 2013; Cole et al., 2012). Thresholds have been defined in different ways, but often in relation to a seasonal average value. In the case of HABs however, where high levels of biomass and their persistence in time may be more relevant than seasonally scaled values, the use of absolute thresholds may be justified. Extreme blooms in coastal settings have been defined using absolute thresholds in the literature. Ryan et al. (2008) used a threshold applied to the fluorescence based Maximum Chlorophyll Index (MCI) to indicate when biomass had exceed a concentration of 75 mg m^{-3}. *In situ* samples from St Helena Bay indicate that blooms in the area can exceed 100 mg m^{-3} (Pitcher et al., 2011), however impacts from HABs have been recorded at a range of lower concentrations, and may not correspond clearly with presence of toxins (*ibid*). A comparison was carried out to compare the extent of bloom occurrence detected when thresholds of 20, 30, 50, 75 and 100 mg m^{-3} were used with the various [Chl *a*] products. A threshold of 75 mg m^{-3} was decided for further development of metrics and assessments of variability (discussed below).

From the binary bloom flag maps where 0 represents no bloom and 1 indicates that the threshold was exceeded and the pixel classified as a bloom, estimates of bloom persistence (in terms of days) were generated on a per pixel basis over the time series. To do this, once the binary bloom flag was raised a count was initiated which was advanced with every subsequent image where the bloom flag remained raised and was returned to zero if the flag was not raised. From this, the number of days of persistence was calculated from the image metadata, given the uneven temporal distribution of the images (i.e. a count of 6 may cover more than 6 days if a clear overpass is only retrieved every 2 to 3 days).

5.2.5 *In situ* and auxiliary data

To allow for validation with *in situ data* and synoptic scale qualitative understanding of algorithm performance, a number of distinct case studies were selected. *In situ* data used for validation for these case studies is detailed in chapter four. Results used here comparing *in situ* algorithm application (i.e. R_{rs} from the Satlantic H-TSRB) correspond to those derived in chapter four using the static intracellular chlorophyll a concentration (c_i) assumption, diatom/dinoflagellate IOP basis vectors, empirically derived initial conditions and the reflectance approximation method.

Where appropriate, pixels were extracted from the MERIS imagery and associated standard, MPH and EAP products for comparison of R_{rs} to that derived from the Satlantic H-TSRB dataset and *in situ* [Chl a] and D_{eff} described in chapter four. The R_{rs} derived from the Satlantic H-TSRB is dependent on the *in situ* EAP inversion algorithm where L_u (-0.66m) is used for convergence of the forward model and the ultimate calculation of R_{rs} is dependent on the IOPs selected and the consequential parameterisations used for K_u and f/Q. However, as the R_{rs} extracted from the MERIS imagery and used in the satellite EAP inversion algorithm is initially independent of these assumptions, an intercomparison of the results should yield insight in to firstly, the performance of MERIS (with regards to atmospheric correction etc) and secondly the performance of the EAP inversion approach as a tool for processing the H-TSRB data to R_{rs}.

Images were predominantly chosen where coincident *in situ* matchup data were available. In some cases, full *in situ* match up, including radiometric data were not available coincidently but case studies were selected on the basis of ecological context and previous studies conducted in the region with regard to coastal circulations, occurrence of particular species and species succession. All images presented are from the third reprocessing of the MERIS data using the MERIS Ground Segment (MEGS) v8.0.

Wind data were obtained from the South African Weather Service (SAWS) for two *in situ* stations, one at Cape Columbine and the other above Lambert's Bay (the Nortier station - approximately 8.5 km inland). In addition to these data, the merged "blended SeaWinds" wind speed product at daily, 0.25 ^{o} resolution was downloaded for the St Helena Bay region (Zhang et al. (2006), available at: http://www.ncdc.noaa.gov/oa/rsad/air-sea/SeaWinds.html). These winds were then converted to alongshore winds to derive upwelling favourable and relaxed/downwelling favourable conditions.

5.3 Results

5.3.1 Bloom case studies

Case study 1: 25th October 2002 - Alexandrium catenella

On the 25th October 2002, *in situ* sampling was conducted in a large bloom off Lambert's Bay. The dominant species for this bloom was determined as *Alexandrium catenella* with *in situ* Coulter Counter derived D_{eff} and [Chl a] of 25.1 μm and 309 mg m^{-3} respectively. Figure 5.3 shows the Algal 1 [Chl a] (a), EAP [Chl a] (b) and EAP derived D_{eff} (c) from the MERIS image associated with this day. Radiometric match-ups were extracted from the MERIS imagery and are shown in comparison to R_{rs} derived from the EAP approach applied to the Satlantic H-TSRB data (Fig. 5.3(d)).

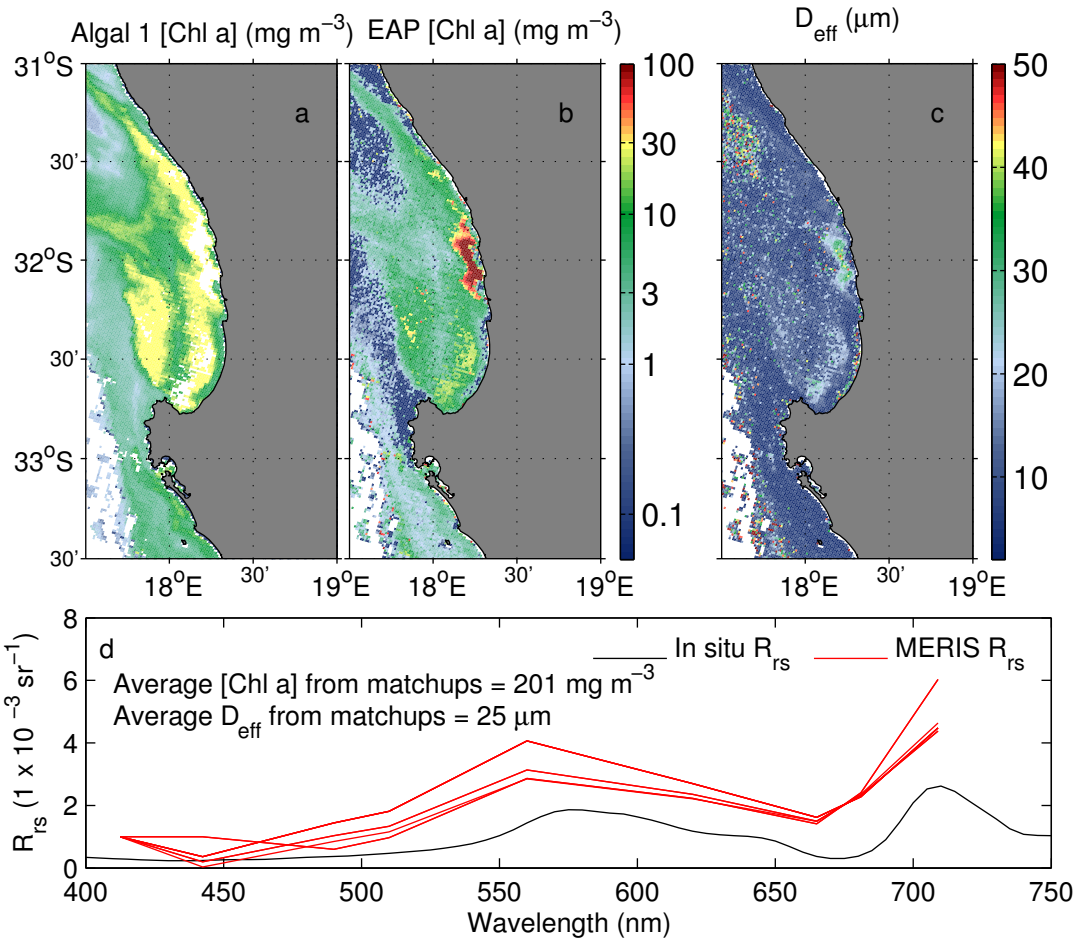

Figure 5.3: Products derived from MERIS Level 2 imagery of St Helena Bay on the 25th October 2002. The Algal 1 [Chl a] product (a) is the standard [Chl a] product for Case 1 waters. Panel (b) shows the [Chl a] product for the same image derived through the EAP algorithm. Panel (c) shows the associated D_{eff} product from the EAP algorithm. R_{rs} data from the Satlantic H-TSRB (an output from the *in situ* application of the EAP algorithm) is compared to MERIS R_{rs} extracted from pixels within 1km of the *in situ* sampling station at 32.08750^{o}S 18.2680^{o}E. Average [Chl a] and D_{eff} from these match up pixels are given in the legend of panel (d).

Figure 5.3(d) shows the R_{rs} from the MERIS sensor matches well with *in situ* R_{rs} in terms of broad spectral shape, however the magnitude of the MERIS R_{rs} is greater than that estimated from the *in situ* radiometry. Closer inspection of the flags raised for this image indicates that the Bright Pixel Atmospheric Correction (BPAC) (Aiken and Moore, 2000) was used over most of the image, and ubiquitously over the bloom region. The matchup between satellite and *in situ* R_{rs} suggests that this atmospheric correction method is largely appropriate for this region, with spectral features maintained. The results here also emphasise the importance of a 709 nm band for capturing signal in this region associated with high phytoplankton biomass. Low R_{rs} in the blue can be seen from the pixels extracted, particularly at 443 nm. These are examples of the pixels discussed in the methods section above, where negative reflectances in the blue cause the failure of the Algal 1 product. The extent of this problem can be seen by comparing the Algal 1 [Chl *a*] product (Fig. 5.3(a)) with the EAP [Chl *a*] product (Fig. 5.3(b)). In the Algal 1 product NaN values (white) are present over most of the bloom region which can be identified in the EAP [Chl *a*] product, showing that the EAP algorithm provides a substantial advantage over the Algal 1 product for identifying and quantifying the levels of biomass associated with high biomass dinoflagellate blooms of this nature. From the match up pixels extracted and detailed in (Fig. 5.3(d)) the EAP [Chl *a*] (mean of 201 mg m^{-3}, with maximum of 271 mg m^{-3}) compares favourably to the *in situ* value (309 mg m^{-3}), especially given the extremely high biomass associated with this bloom and the discrepancies between the magnitude of satellite derived and *in situ* R_{rs}. In general, outside of the bloom region, the EAP product underestimates [Chl *a*] compared to the Algal 1 product, however the synoptic scale distribution of high and low biomass are mostly similar.

The EAP D_{eff} product also compares favourably to *in situ* values, with match up pixels indicating an average D_{eff} of 25.7 surrounding the sampling location where the Coulter Counter returned a D_{eff} of 25.1 μm. High D_{eff} values of between 20 and 30 μm occur coherently across the high biomass bloom feature identified in the EAP [Chl *a*] (Fig. 5.3(a)), consistent with a bloom dominated by a large celled organism such as *A. catenella* (Equivalent Spherical Diameter (ESD) $\approx$ 32.6 μm (Yoo et al., 2009)). D_{eff} values of around 20 μm also occur coincidently with a high biomass feature (around 30 mg m^{-3}) towards the south of the bay. Some pixels with high D_{eff} values are also apparent in the "speckled" patch in the top left of the image, however as these are associated with low biomass, it is likely that these are a result of the ambiguity identified at low biomass as a feature of the EAP approach in previous chapters.

Case study 2: Changes in species dominance between March and April 2005

A high biomass bloom was observed during a shore based field campaign from Lambert's Bay in March and April 2005. Figure 5.4(d) shows a matchup between Satlantic H-TSRB and MERIS R_{rs} from valid pixels within a 1 km radius of the sampling location. As seen in figure 5.3(d), broad spectral shape is similar between the satellite and *in situ* derived R_{rs}, however the MERIS R_{rs} appears to overestimate in comparison to the Satlantic H-TSRB derived R_{rs}. As in

the previous case study, the Algal 1 [Chl a] product fails over most of the inshore region where the bloom occurred. The bloom can be more readily identified in the EAP [Chl a] product, which gives [Chl a] values (from 147 to 163 mg m^{-3}, with an average of 155 mg m^{-3}) close to those measured *in situ* (between 115 and 173 mg m^{-3}, with an average of 148 mg m^{-3}). Again, outside of the bloom conditions, the EAP [Chl a] generally underestimates compared to Algal 1, though features are similar between the two products. The D_{eff} product over the high biomass feature suggests that most of the bloom was dominated by smaller cells (between 10 and 20 μm) (Fig. 5.4(b)). This is consistent with microscopy samples during this period which suggest a mixed dinoflagellate assemblage including toxic species *Dinophysis acuminata*, *Dinophysis fortii* and *Protoceratium reticulatum* with a dominance of the non-toxic, *Prorocentrum triestinum* (ESD $\approx$ 12.6 μm (Jeong et al., 2005)) in the nearshore region off Lambert's Bay (Fawcett et al., 2007). In contrast, the southern part of the bloom and a high biomass feature offshore of Cape Columbine, appears to be dominated by larger cells (D_{eff} between 30 and 50 μm).

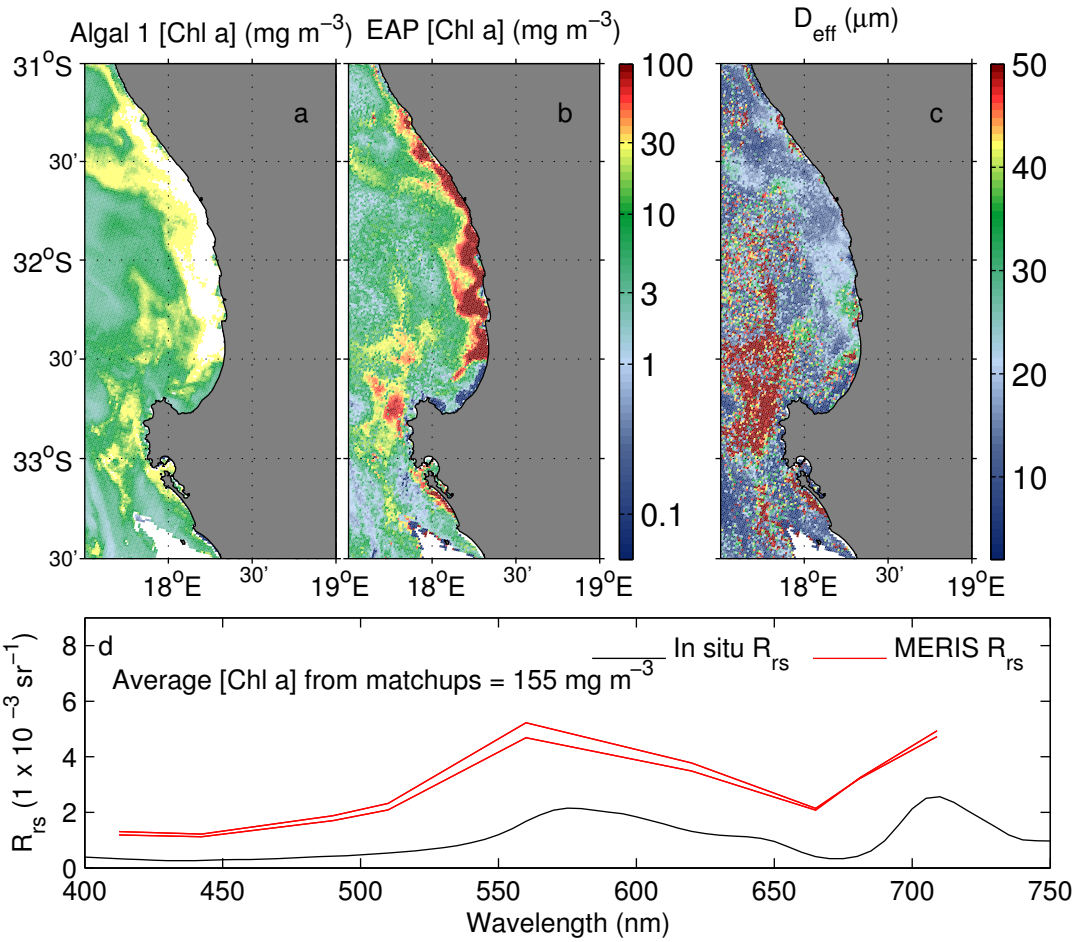

Figure 5.4: Products derived from MERIS Level 2 imagery of St Helena Bay on the 30th March 2005. The Algal 1 [Chl a] product (a) is the standard [Chl a] product for Case 1 waters. Panel (b) shows the [Chl a] product for the same image derived through the EAP algorithm. Panel (c) shows the associated D_{eff} product from the EAP algorithm. R_{rs} data from the Satlantic H-TSRB (an output from the *in situ* application of the EAP algorithm) is compared to MERIS R_{rs} extracted from pixels within 1km of the *in situ* sampling station at 32.08750^{o}S 18.2678^{o}E. Average [Chl a] from these match up pixels is given in the legend of panel (d).

Figure 5.5 shows results from St Helena Bay a week later on the 5th April 2005. The EAP [Chl *a*] product suggests that high biomass was present throughout the bay at this time, with extracted values from near the sampling location off Lambert's Bay between 64 and 94 mg m^{-3}, showing an over prediction compared to those measured *in situ* (40 mg m^{-3}).

Effective diameter estimates from the EAP algorithm for both the 30th March and 5th April suggest a significant change in the estimated size of the assemblage between the two dates (Fig. 5.4(c) and Fig. 5.5(c)). The inshore high biomass bloom patch on the 30th March identified in figure 5.4(b), appears to be dominated by relatively smaller cells (10 - 20 μm) in the north of the bay, with a shift to relatively larger cells (30 -50 μm) in the south and offshore bloom patches. Microscopy estimates suggest that the predominance of smaller cells may be attributable to the dominance of *Prorocentrum triestinum* in a mixed dinoflagellate assemblage (Fawcett et al., 2007). By the 5th of April, the estimated D_{eff} (Fig. 5.5(c)) over the high biomass patches (Fig. 5.5(b)) has shifted further towards larger cells, particularly in the south of the bay, which microscopy data attribute to changes in species dominance favouring *Ceratium furca* ($\approx$ 28 μm (Smalley et al., 2003)). A period of upwelling favourable wind between the 30th March and 5th April (Fawcett et al., 2007) may have provided additional nitrate, favouring the proliferation of the larger cells. As in figure 5.3(c), speckling is seen in the D_{eff} product associated with lower biomass estimates, which may be the result of ambiguity.

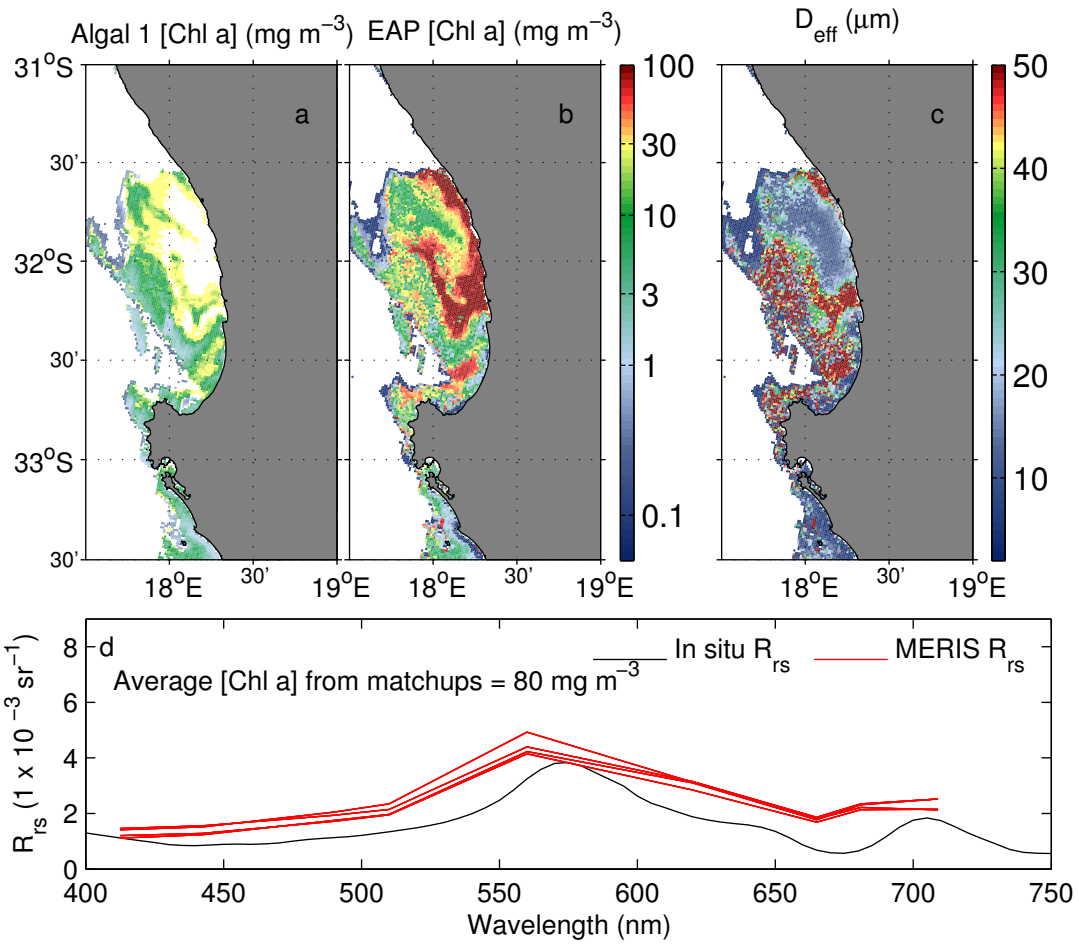

Figure 5.5: Products derived from MERIS Level 2 imagery of St Helena Bay on the 5th April 2005. The Algal 1 [Chl *a*] product (a) is the standard [Chl *a*] product for Case 1 waters. Panel (b) shows the [Chl *a*] product for the same image derived through the EAP algorithm. Panel (c) shows the associated D_{eff} product from the EAP algorithm. R_{rs} data from the Satlantic H-TSRB (an output from the *in situ* application of the EAP algorithm) is compared to MERIS R_{rs} extracted from pixels within 1km of the *in situ* sampling station at 32.08750^{o}S 18.2678^{o}E. Average [Chl *a*] from these match up pixels is given in the legend of panel (d).

Case study 3: Different blooms in different bays during 2007

Pitcher et al. (2008b) reported on two concurrent Harmful Algal Bloom events in two different bays along South Africa's south west coast in 2007. A bloom dominated by *Alexandrium catenella* occurred off Elands Bay within the greater St Helena Bay region, whilst a bloom of *Gonyaulax polygramma* dominated the inshore region of False Bay during March (Pitcher et al., 2008b). The *A. catenella* bloom was associated with episodes of toxicity in mussels whilst the *G. polygramma* bloom was linked to the mortality of marine fish and invertebrates, likely through inshore anoxia caused by the eventual decay of the bloom. Chlorophyll *a* concentration for the southern Benguela region between the 15th February and 13th March 2007 is shown in figure 5.6. The high biomass associated with both blooms is captured well by the EAP [Chl *a*] product.

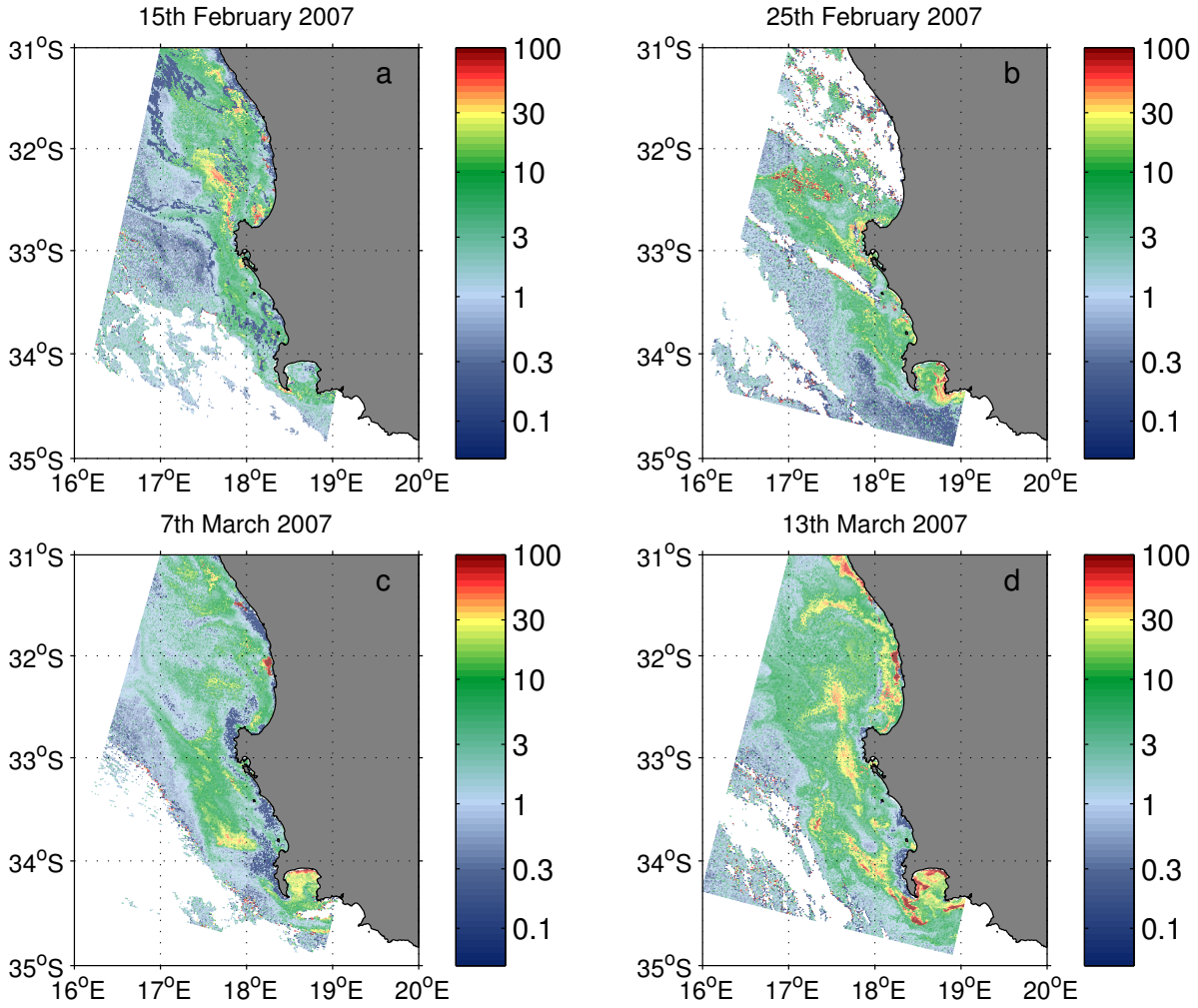

Figure 5.6: EAP [Chl *a*] products derived from MERIS Level 2 imagery for the (a) 15th February 2007, (b) 25th February 2007, (c) 7th March 2007 and (d) 13th March 2007.

Evidence for the large nature of the cells associated with the *A. catenella* bloom in the inshore region of St Helena Bay can be seen in the EAP D_{eff} estimates for the February-March period (Fig. 5.7). Despite also being a large cell (ESD $\approx$ 32.6 μm (Yoo et al., 2009)), the D_{eff} response is not as strong over the high biomass patches of the *G. polygramma* bloom in False Bay, with only slightly elevated D_{eff} values ($\approx$ 20 μ m). However, higher D_{eff} values are seen offshore, associated with the Cape Columbine frontal system and jet, where *in situ* measurements by Pitcher et al. (2008b) indicated the dominance of *G. polygramma*, presumably transported to this region from the south coast bloom, whilst the *A. catenella* bloom was confined to a narrow band near the coast in St Helena Bay. As with the previous case studies, speckling in the D_{eff} product is seen. The additional spatial coverage associated with the 2007 case study (Fig. 5.7) lends support to the previous inference that this is related to ambiguity under low biomass conditions, as the speckling is most pronounced in the most offshore areas of the image.

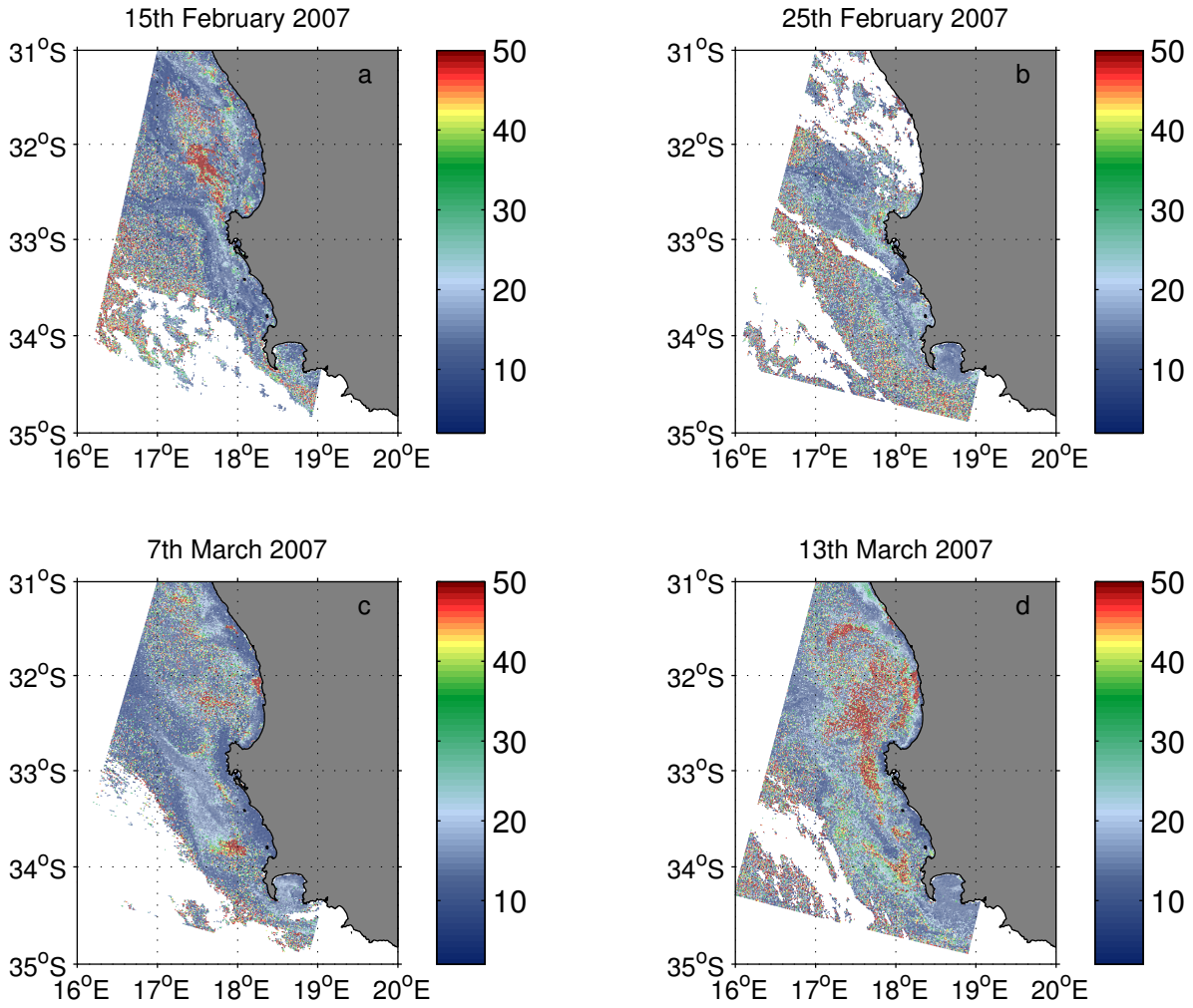

Figure 5.7: EAP D_{eff} products derived from MERIS Level 2 imagery for the (a) 15th February 2007, (b) 25th February 2007, (c) 7th March 2007 and (d) 13th March 2007.

Case study 4: Bloom transport in St Helena Bay between February and May 2009

High biomass was observed through an *in situ* water quality mooring and field sampling campaigns off Elands Bay between late November 2008 and mid October 2009 (Pitcher and Probyn, 2011). Sampling between March and May 2009 during extremely high biomass, where [Chl *a*] values frequently exceeded 100 mg m^{-3}, indicated high cell concentrations of the large dinoflagellate *Ceratium balechii* (*ibid*). According to *in situ* sampling, high biomass persisted through the austral summer and autumn throughout the St Helena Bay region and ultimately led to severely low oxygen concentrations and mortalities in the south of St Helena Bay between Dwarskersbos and the Berg River mouth (Pitcher and Probyn, 2011).

The full spatial extent of this high biomass over St Helena Bay region during the February to May 2009 period can be assessed using the EAP [Chl *a*] product. Figure 5.8 suggests that the bloom initially formed in the north of the bay in late February, extending south throughout the bay by mid April before becoming confined to the nearshore in the most southerly part of St Helena Bay behind Cape Columbine and eventually around the headland by early May.

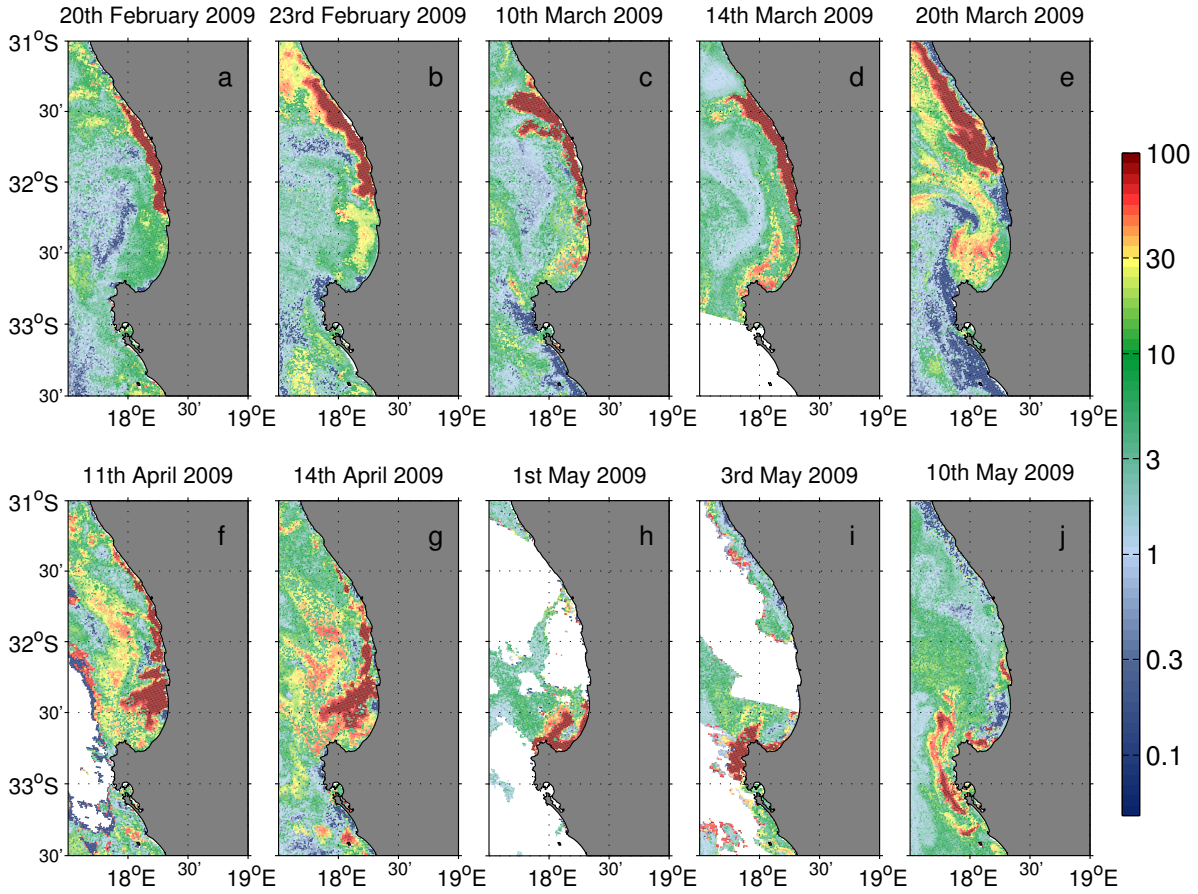

Figure 5.8: EAP [Chl *a*] products derived from MERIS Level 2 imagery between the 20th February 2009 and 10th May 2009.

The coincident D_{eff} product for the images in figure 5.8 is shown in figure 5.9. These results suggest that the bloom may initially have been dominated by a smaller species or a mixed assemblage with a smaller overall D_{eff}. However, as the time series of images progresses, the D_{eff} estimates suggest dominance by a very large (up to 50 μm) cell, coincident with the *in situ* measurements of *Ceratium balechii* cell concentrations (Pitcher and Probyn, 2011).

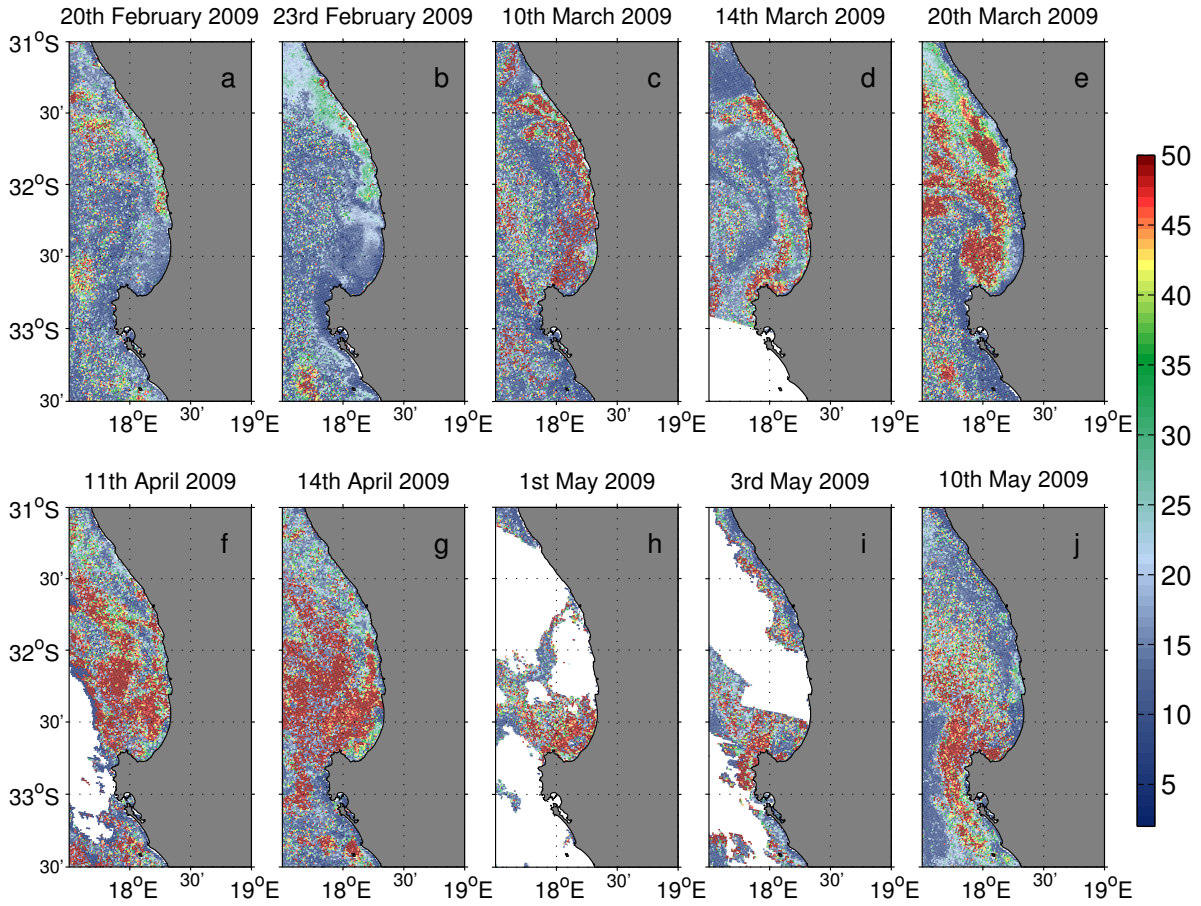

Figure 5.9: EAP D_{eff} products derived from MERIS Level 2 imagery between the 20th February 2009 and 10th May 2009.

Case study 5: Distinguishing Mesodinium rubrum blooms

The distinction of different phytoplankton species through satellite remote sensing techniques is of benefit to HABs monitoring (Bernard et al., 2006; Kudela et al., 2005) and ecosystem studies in general (Anderson, 2005; Nair et al., 2008; Rudorff and Kampel, 2012). The examples from the previous case studies show that the EAP approach can be used to derive information on bulk community cell size, however the use of the additional, pigment based IOP basis vectors representing cryptophytes and endosymbionts available from the two-layered sphere model (Bernard et al., 2009) has not yet been assessed in the satellite application of the EAP algorithm. At the satellite scale, the amount of information contained in R_{rs} spectra is likely diminished due to the reduced wavelength resolution compared to *in situ* hyperspectral radiometry, making the distinction of features relating to absorption of particular pigments for example, more difficult. In the case of *Mesodinium rubrum*, a frequent component of HAB assemblages in the southern Benguela (Pitcher et al., 2010), the MERIS wavebands are not ideally placed to resolve the bifurcated reflectance peaks associated with absorption by phycoerythrin pigments. A few test images were processed where the EAP algorithm was allowed to select between the diatom/dinoflagellate and cryptophyte basis vectors. These images were selected from 2004,

where *in situ* samples showed the presence of *Mesodinium rubrum* at varying concentrations. Satlantic H-TSRB derived R_{rs} were also investigated to identify changes in shape of spectra which may be attributed to the presence of *Mesodinium rubrum* at different concentrations.

Figure 5.10 shows the output [Chl *a*] and D_{eff} from the extended version of the EAP where the optimisation is allowed to select between the diatom/dinoflagellate and cryptophyte IOP basis vectors as an admixture. The broad patterns in [Chl *a*] (Fig. 5.10(a) and (d)) and D_{eff} (Fig. 5.10(b and e)) are not greatly altered from those generated using the diatom/dinoflagellate version (not shown). Effective diameter estimates are broadly consistent with blooms being dominated by a relatively large cell (*M. rubrum* cell size ≈ 30-50 μm (Smayda, 2000)). Further examination of the admixture selected shows no particularly coherent patterns in basis vectors selected, with the exception of an apparent relationship between lower biomass, smaller D_{eff} and the selection of cryptophyte basis vectors.

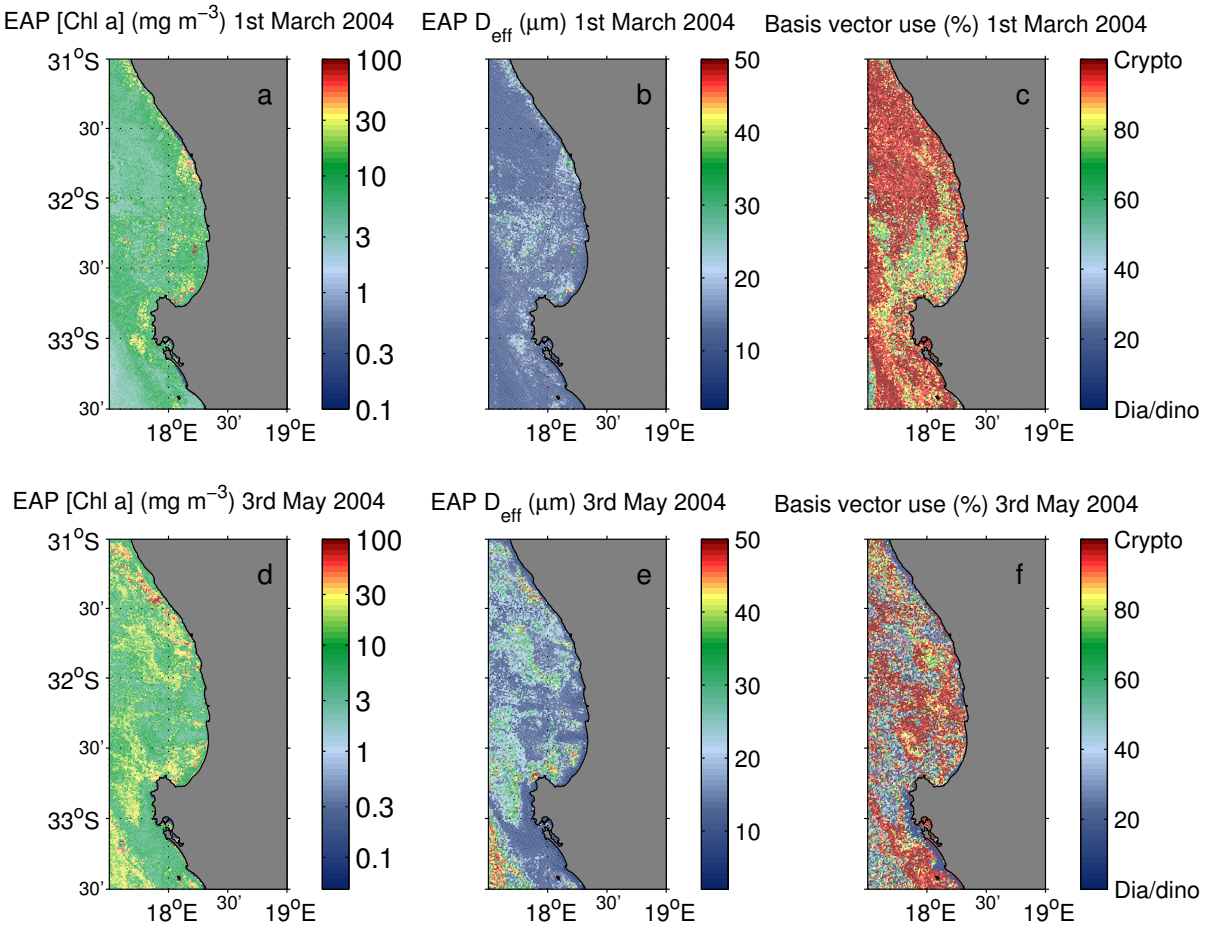

Figure 5.10: Extended EAP products ([Chl *a*], D_{eff} and basis vector admixture) derived from MERIS Level 2 imagery for the 1st March (a, b and c) and 3rd May 2004 (d, e and f).

The characteristic reflectance peaks associated with the presence of *Mesodinium rubrum* are apparent in the example spectra from various samples in 2004 shown in figure 5.11(a). Looking at these spectra, it is apparent that at higher biomass levels, samples containing *Mesodinium rubrum* exhibit elevated R_{rs} in the 600 to 650 nm region compared to samples

dominated by diatoms. A ratio between reflectance at bands 620 and 550 nm can therefore be used to distinguish the presence of *M. rubrum* from diatoms or dinoflagellates. With increasing biomass the 620/550 nm ratio becomes increasingly positive when *M. rubrum* is present. Although this is a largely qualitative application at present, figure 5.11(c) shows how patches of potential *M. rubrum* dominance can be distinguished from background high biomass using this ratio. Additional data i.e. R_{rs} from a variety of mixed assemblages with coincident cell counts could allow this ratio to be developed further for detection of this species.

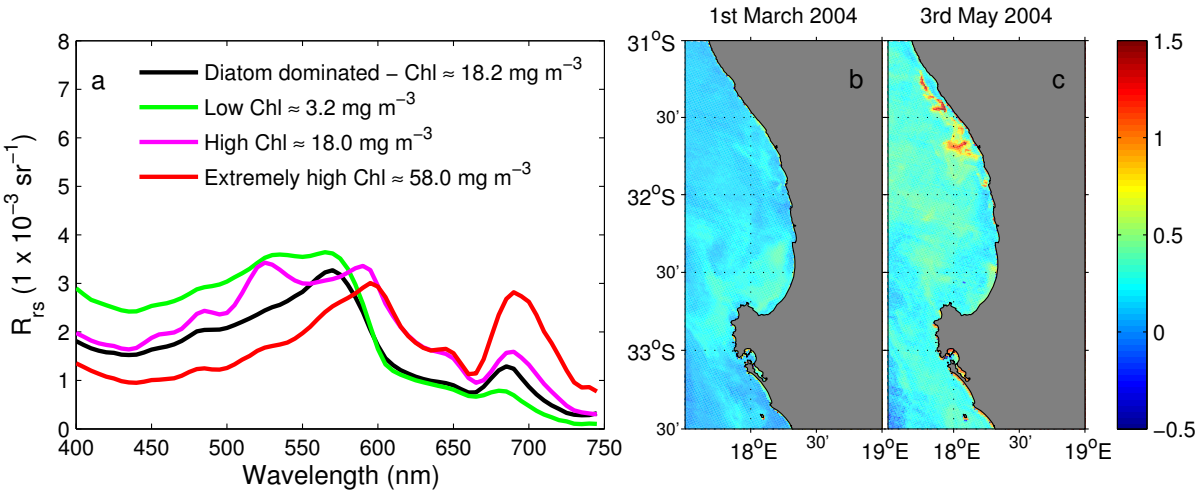

Figure 5.11: Spectral features associated with the presence of *Mesodinium rubrum* (a) can be exploited at satellite resolution using a reflectance ratio between 620 and 550 nm, however this appears strongly biomass dependent (b) and (c).

5.3.2 Time series of [Chl *a*] and effective diameter for St Helena Bay

Spatially averaged time series for [Chl *a*] in the St Helena Bay region are shown in figure 5.12. Time series are derived from the standard Level 2 Algal 1 product, the EAP algorithm and the Maximum Peak Height Algorithm of Matthews et al. (2012) applied to Level 1 data.

The case studies in the previous section of this chapter highlighted that the Algal 1 algorithm, and associated flagging may result in missing data, particularly over the inshore region of St Helena Bay during high biomass blooms. The effect of this, combined with the smaller range of [Chl *a*] from the Algal 1 algorithm is apparent in the spatial averaging conducted for figure 5.12. Both the EAP and MPH approaches capture the larger range of [Chl *a*] values associated with high biomass blooms in the southern Benguela. Further specific validation is required to assess the amount of over/underestimation resulting from these approaches, however for the purposes of identifying high biomass (above 75 mg m^{-3}) most frequently associated with HABs and anoxia, the broad range captured by these approaches is sufficient.

All three [Chl *a*] products capture a seasonal cycle, associated with the increase in upwelling wind conditions in September of each year, through to peak conditions in March/April of the following year (Fig. 5.12(a)). Further variability is apparent at the subregional scale. The Elands/Lambert's Bay averages (Fig. 5.12(b)) show greater interannual variability in the mag-

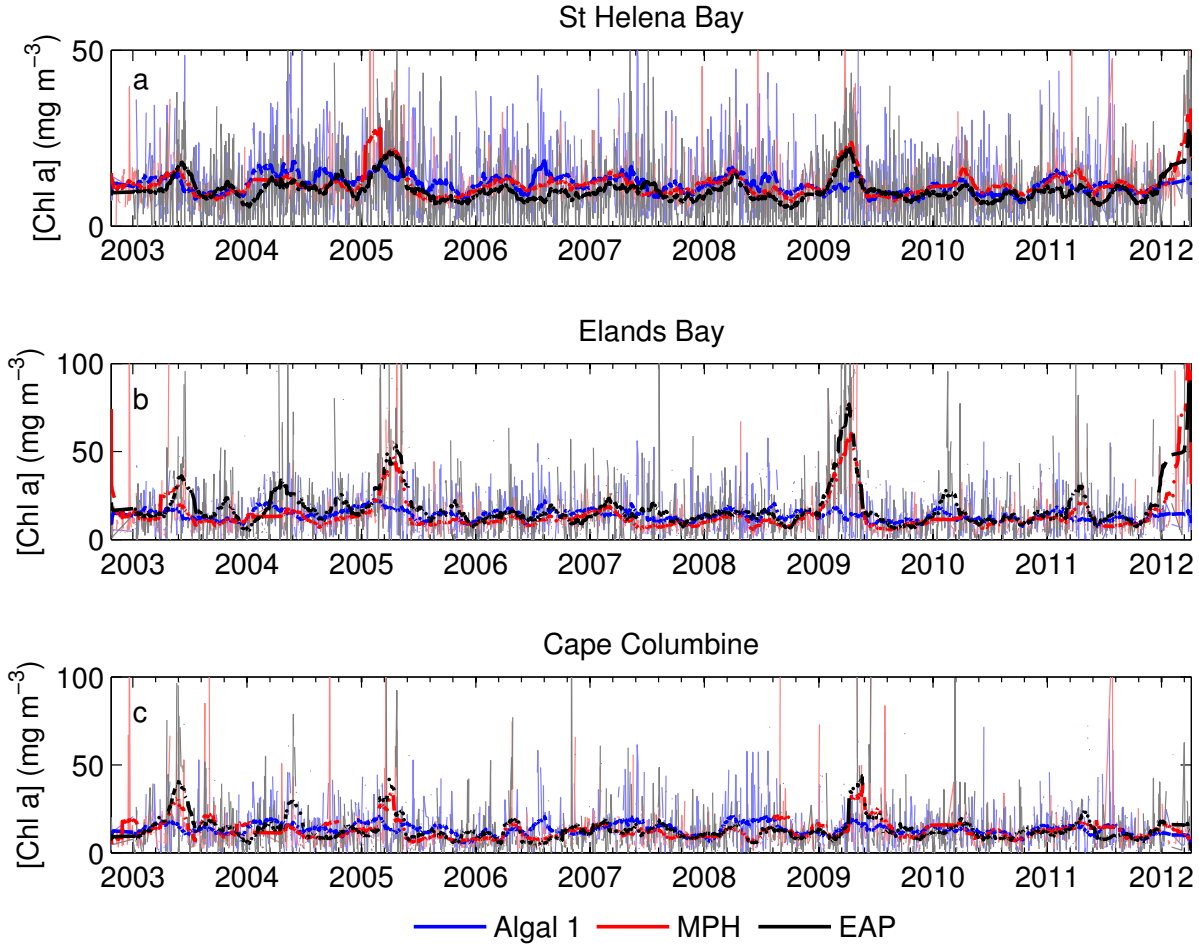

Figure 5.12: Time series of [Chl a] (mg m^{-3}) for (a) The whole St Helena Bay domain, (b) the Elands/Lambert's Bay subregion and (c) Cape Columbine subregion as denoted in Figure 5.1 using Algal 1 (blue), MPH (red) and EAP (black) algorithms. A 30 day running mean is shown by the bold lines, with the full time series shown in the corresponding background fill.

nitude of peak regional average [Chl a], likely associated with the occurrence of high biomass blooms (such as those shown in the case studies) in this inshore region. However, this is not consistent interannually and figure 5.12(b) suggests that higher than average high biomass may have occurred in 2004/2005, 2008/2009 and 2011/2012 upwelling seasons. The Cape Columbine averages (Fig. 5.12(c)), show higher subseasonal variability with less clear 12 monthly peaks, although high values are still apparent in the years identified in the St Helena Bay (Fig. 5.12(a)) and Elands/Lambert's Bay (Fig. 5.12(b)) time series, with an additional peak in 2003/2004.

A time series of bay averaged D_{eff} is presented in figure 5.13. There appears to be a strong seasonal cycle present in average D_{eff} values in both the St Helena Bay (Fig. 5.13(a)), although the peak average values appear to be centred mid year, after the peak [Chl a].

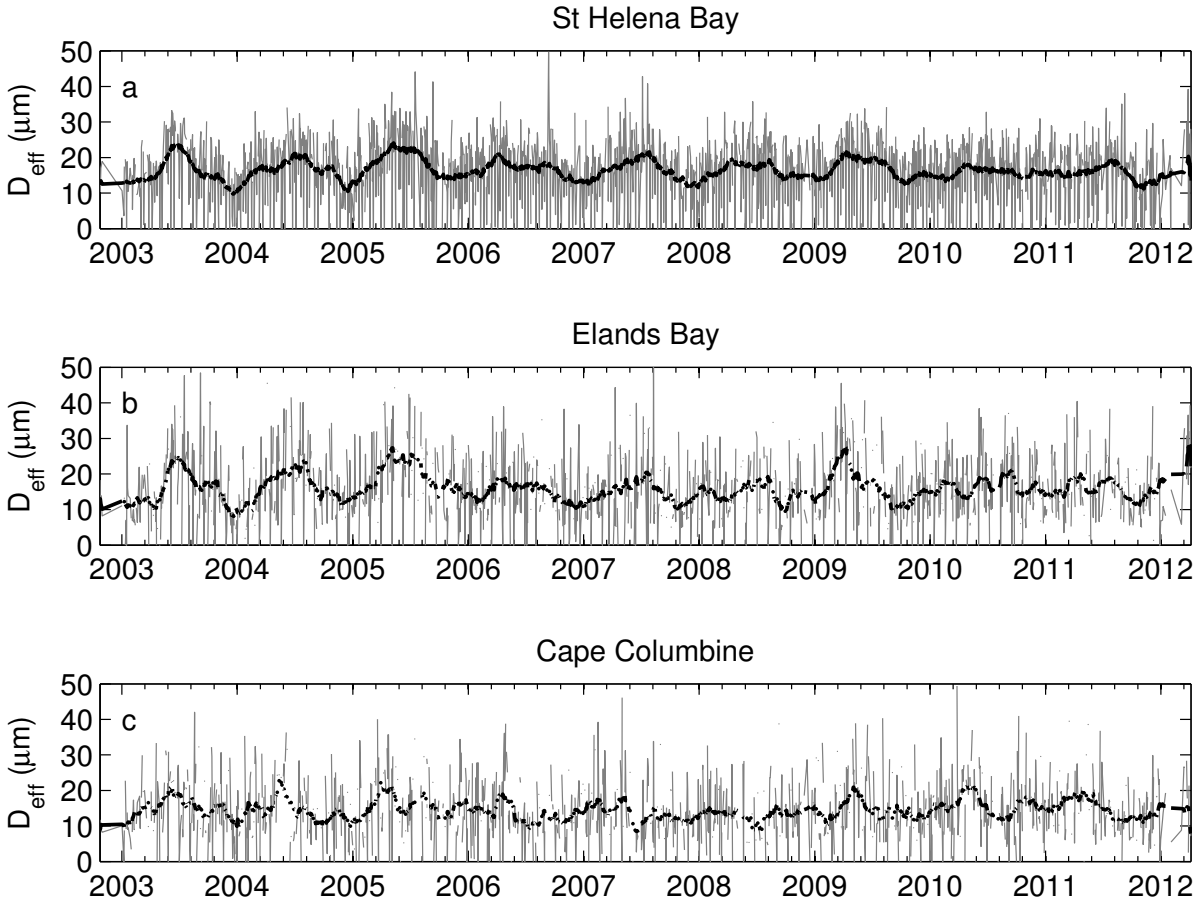

Figure 5.13: Time series of effective diameter (D_{eff}, μm) for (a) The whole St Helena Bay domain, (b) the Elands/Lambert's Bay subregion and (c) Cape Columbine subregion as denoted in Figure 5.1 using the EAP algorithm. A 30 day running mean is shown by the bold lines, with the full time series shown in the corresponding background fill.

5.3.3 Development of metrics useful for quantifying bloom behaviour

As discussed in the introduction and methods section of this chapter, various metrics have been used to help characterise interannual variability in phytoplankton blooms in the global oceans and investigate response to various physical and biological forcing factors (see examples discussed in Cole et al. (2012)). To extend the insight from the time series of bay scale [Chl *a*] averages, several metrics were designed to focus on inter annual variability of event scale HABs.

Identification of blooms using threshold techniques

Blooms were identified on a per pixel basis when [Chl *a*] exceeded a threshold of 75 mg m^{-3}, as used in (Ryan et al., 2008). For each image, each pixel was assessed with regards to the threshold criteria, raising a flag value of one if the [Chl *a*] value was greater than the threshold, or a zero if not. Whilst a threshold technique alone will not quantify all blooms that could be classified as "harmful", due to the presence of toxins at much lower [Chl *a*] than

used as the threshold (e.g. (Fawcett et al., 2007)), this value should be sufficient to distinguish blooms which may lead to anoxia from those associated with general high productivity in the upwelling season. Figure 5.14 below shows an example bloom threshold map associated with the *Prorocentrum triestinum* bloom from case study 1 (Fig. 5.4).

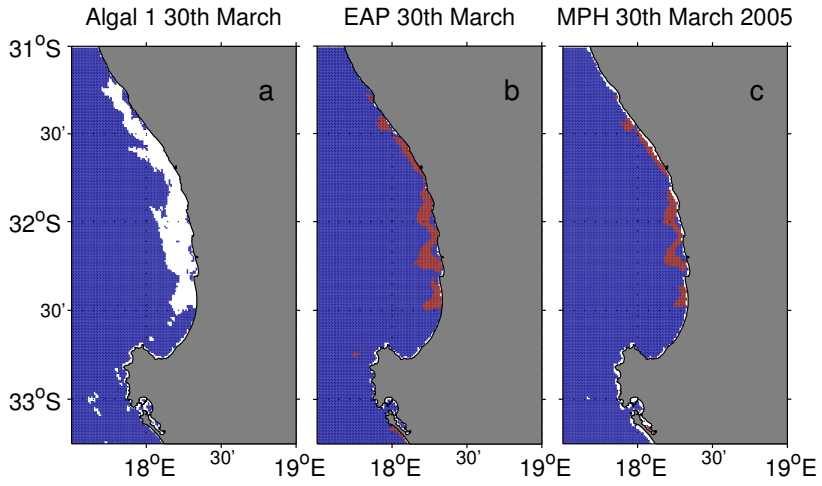

Figure 5.14: Pixels classified as "bloom" (red) and "non-bloom" (blue) for the 30th March 2005 using the (a) Algal 1 (b) EAP and (c) MPH [Chl *a*] products.

Figure 5.15 shows the percentage of valid pixels classified as in bloom for each upwelling season of the 2002 - 2012 time series. Highest years of bloom activity are as identified in the bay average time series (Fig. 5.12) including 2004/2005, 2008/2009 and 2011/2012 upwelling seasons. As might be expected from the performance over bloom conditions in the case studies, using the Algal 1 product returns very low counts of bloom pixels across most upwelling seasons (Fig. 5.15(a)). In addition to this, inspection of the threshold maps generated using Algal 1 indicates that these pixels are typically associated with features around the edges of clouds or along the edges of the swaths, and not in the inshore region typically associated with blooms and may therefore be erroneous high estimates of [Chl *a*]. Identifiable in the thresholds determined from the EAP and MPH [Chl *a*] products, are the years of extremely high biomass identified in the time series of bay scale and subregional averages i.e. 2004/2005, 2008/2009 and (in the case of the MPH), 2011/2012. Whilst the patterns of high/low bloom activity years are similar for the EAP and MPH problems, they are not identical. The EAP and MPH were not applied to identical sets of imagery, as not all imagery was available at the required level with appropriate processing at the time of analysis. So, some discrepancy in bloom detection is likely attributable to this, particularly the 2011/2012 bloom for example, where only 27 images were suitable at Level 2 for the application of the EAP algorithm (due to high glint), whilst 46 Level 1 images were suitable for the application of the MPH. However, broadly it seems that the MPH identifies less bloom pixels than the EAP. This could be for a number of reasons, including that the EAP generally overestimates high biomass compared to the MPH, however this seems unlikely, given the similarity of the time series of bay scale and subregional averages (Fig. 5.12).

Examples of the spatial distribution of bloom classification are shown in figure 5.16 for the

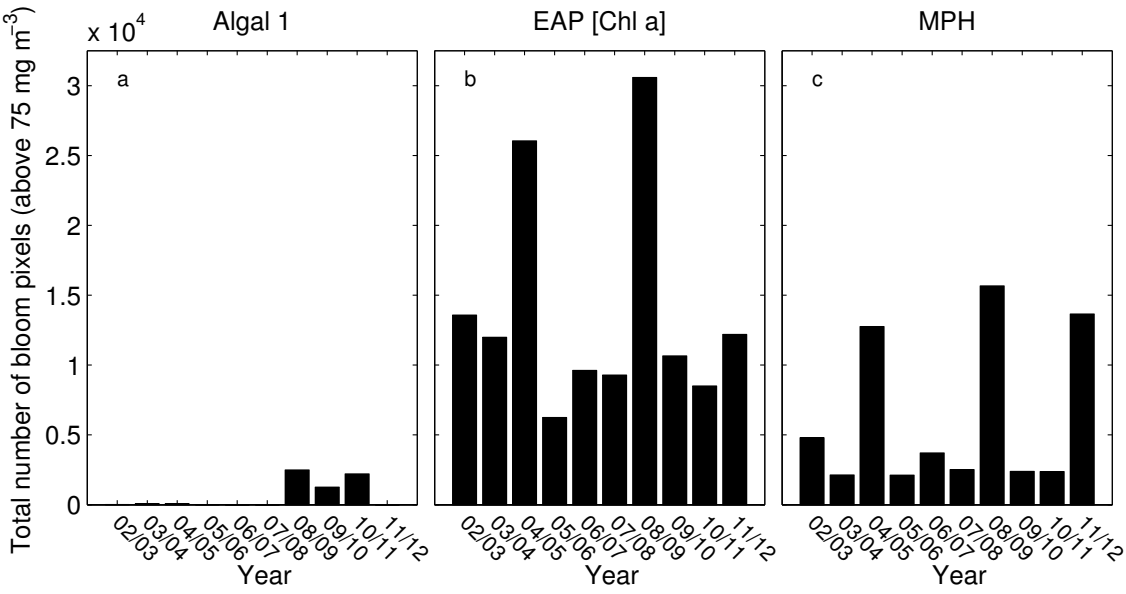

Figure 5.15: Total number of pixels classified as "bloom" for each upwelling season using the (a) Algal 1, (b) EAP and (c) MPH [Chl *a*] products.

high bloom activity years identified in figures 5.12 and 5.15. The percentage of pixels classified as bloom is highest over the inshore region where high biomass dinoflagellate blooms typically occur and consistent with the higher average [Chl *a*] observed, particularly for the Elands/Lambert's Bay subregion. Generally, the EAP classifies more pixels as bloom than the MPH in the inshore region. However the EAP also classifies far more pixels offshore as bloom compared to the MPH. This may explain why the magnitude of pixels classified as bloom is twice that when using the EAP [Chl *a*] over the MPH [Chl *a*]. This may be attributable to erroneous high [Chl *a*] resulting from the ambiguity of the semi-analytical inversion approach of the EAP, which is particularly high at low biomass (see chapter two and three). As expected from figure 5.15, the EAP underestimates the number of bloom pixels in 2011/2012 compared to the MPH (Fig. 5.16(c and f)).

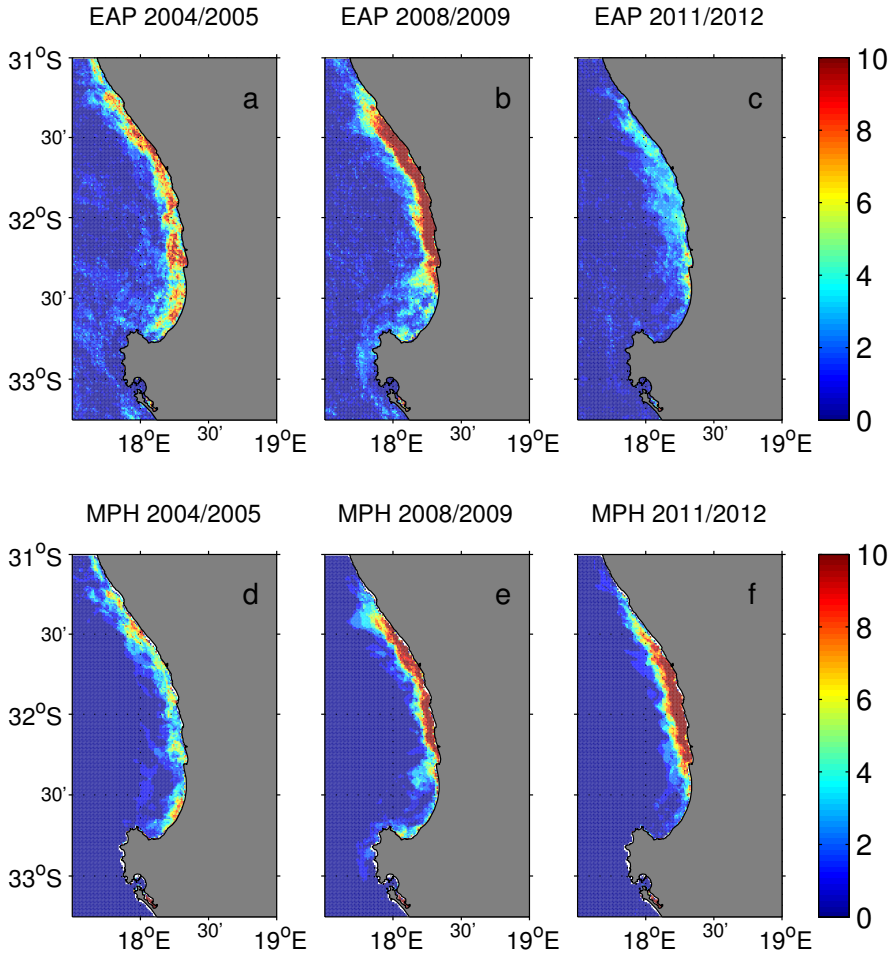

Figure 5.16: Percantage of valid pixels classified as "bloom" over the St Helena Bay region for the (a, d) 2004/2005, (b, e) 2008/2009 and (c, f) 2011/2012 upwelling seasons using the EAP (a, b, c) and MPH (d, e, f) [Chl *a*] products.

Bloom persistence in time

From the bloom threshold metric results, a time series of per pixel bloom persistence was generated. Figure 5.17 and 5.18 shows the frequency of persistent bloom conditions using threshold masks derived from the MPH and EAP [Chl *a*] products respectively. The longest periods of bloom persistence were associated with the years identified as high bloom years in the threshold, showing that not only was there high bloom activity in these years but that this high activity occurred more consistently in time and space. However, there is a flaw in using this metric of persistence to infer likelihood of anoxia for example, which may be caused by persistent blooms. Given the physical circulation of St Helena Bay, it is highly likely that blooms may be advected over relatively short (weekly) time scales. As such, the bloom persistence metric, if it records a single day of bloom activity, does not explicitly account for the fact that this bloom may have moved in to/from an adjoining pixel. Similarly, a bloom may drop in [Chl *a*] below the threshold and then rise above it again. This would be recorded as two separate blooms.

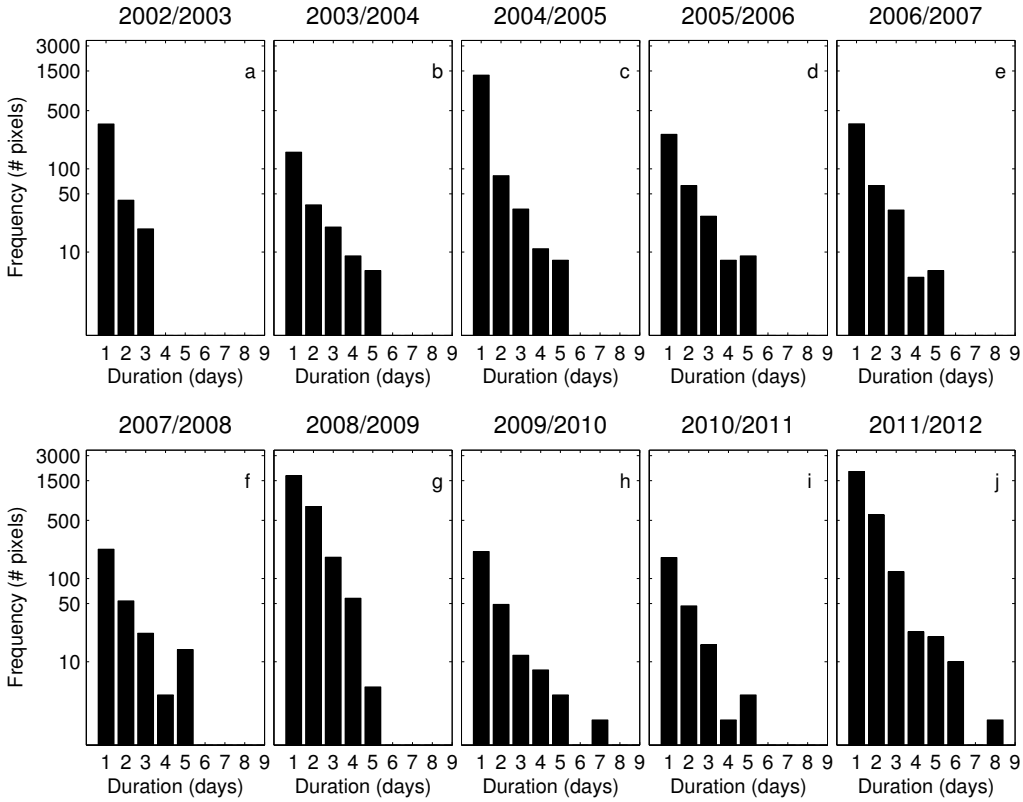

Figure 5.17: Persistence of pixels classified as "bloom" over the St Helena Bay region using threshold masks derived from the MPH [Chl a] product using a threshold of 75 mg m^{-3} [Chl a]

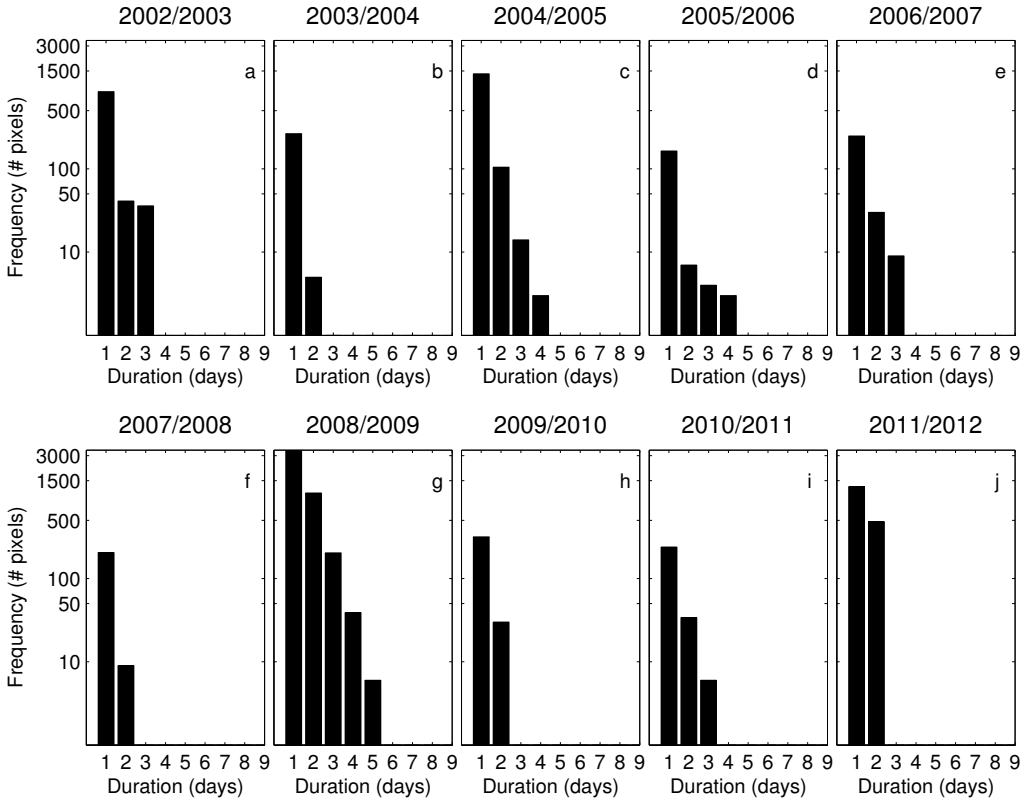

Figure 5.18: Persistence of pixels classified as "bloom" over the St Helena Bay region using threshold masks derived from the EAP [Chl a] product using a threshold of 75 mg m^{-3} [Chl a]

Effective diameters associated with bloom metrics

Given that the previous chapters have shown that accurate D_{eff} retrievals are most likely when [Chl a] is higher, the bloom metrics described should provide a means for selecting the most accurate D_{eff} data for consideration of long-term trends and event scale species succession. Effective diameter estimates were extracted over the regions classified as "bloom" using the threshold metric and averaged per image (Fig. 5.19). As would be expected given general relationships between phytoplankton abundance and [Chl a], the average D_{eff} associated with the blooms points towards dominance of large cells and shows substantially higher average values than those seen in the bay scale averages (Fig. 5.13). However, as in figure 5.13, both seasonal and interannual variability can be observed in the average size values.

Of particular note, is the comparatively low effective diameter associated with blooms in March 2003. Between March and May 2003, a bloom dominated by the coccolithophore *Syrachosphaera pulchra* (ESD $\approx$ 19.9 μm (Napp et al., 1988)) in a mixed assemblage with *Prorocentrum triestinunum* (ESD $\approx$ 12.6 μm (Jeong et al., 2005)) and *Scrippsiella trochoidea* (ESD $\approx$ 22.8 μm (Jeong et al., 2005)) occurred in St Helena Bay (Weeks et al., 2003). Although the optical characteristics of coccolithophores are not explicitly characterised in the two-layered sphere model and consequently in the EAP IOP basis vectors, their highly scattering properties may well be interpreted by the EAP inversion algorithm as an indication of smaller cells.

Closer inspection of the March-April 2005 time period associated with case study 2, supports the inferences made that the bloom assemblage shifted from the dominance of smaller (*Prorocentrum triestinum*) to larger (*Ceratium furca*) celled species. Bloom D_{eff} ranged from 22 to 37 μm between the 31st March and 5th April.

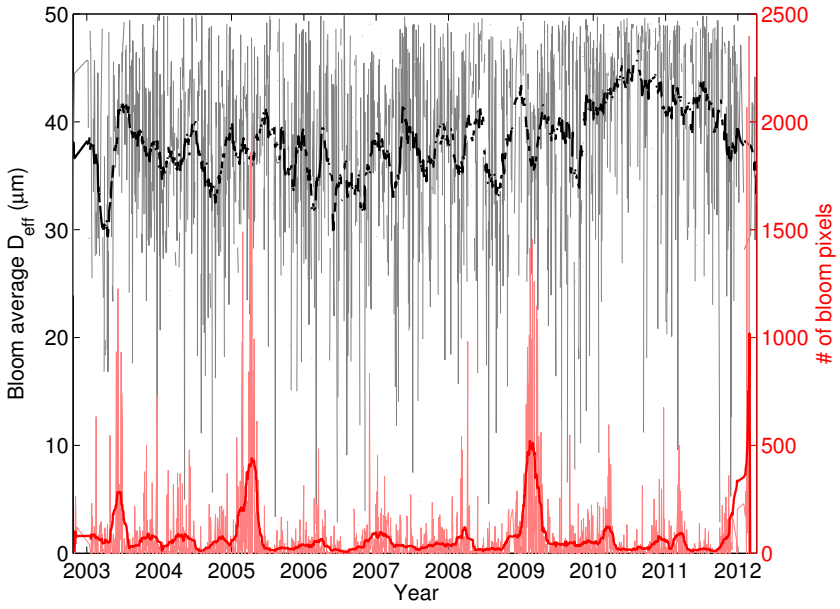

Figure 5.19: Average D_{eff} (per image) associated with pixels classified as "bloom" over the MERIS archive (grey fill), with a 30 day running mean (solid black line). Numbers of bloom pixels per day are given for context (red fill), again with a 30 day running mean (solid red line).

5.4 Discussion

5.4.1 Ambiguity, error and future constraint of the approach for application to satellite data.

The sensitivity analysis of the forward model (chapter two), the inversion of the simulated (chapter three) and *in situ* (chapter four) data assumed the use of hyperspectral data (i.e. at 5 nm resolution). For the satellite resolution, it is likely that the available information in the radiometry is reduced, due to the multispectral resolution (i.e. specific bands on the MERIS satellite). In particular, this is likely to effect the detection of specific R_{rs} features. Figure 5.20 shows the sensitivity of the simulated R_{rs} signal to size, under different [Chl a] for the MERIS wavelengths.

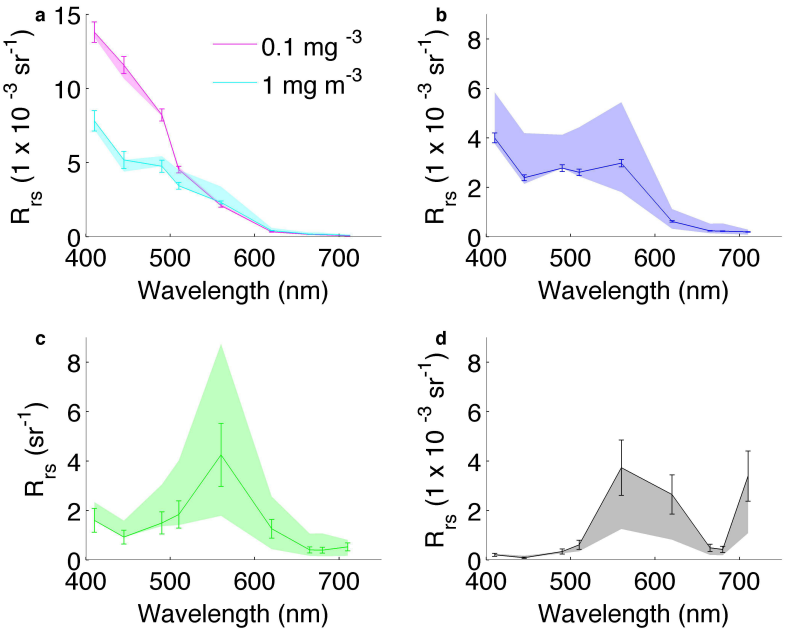

Figure 5.20: R_{rs} spectra from the forward model with associated error estimates for [Chl a] of (a) 0.1 mg m^{-3}, (b) 3 mg m^{-3}, (c) 10 mg m^{-3} and (d) 100 mg m^{-3}. The solid line represents the spectra generated with a D_{eff} of 10 μm. The shaded range shows the range of R_{rs} associated with changes in size. Results here use the forward model with SCI, diatom/dinoflagellate basis vectors, low a_{gd} and low b_{bs} and EcoLight-S. A filter has been applied to remove unrealistic spectra i.e. dominance of large cells (+20 μm) at low [Chl a] (<10 mg m^{-3}) and small cells (<10 μm) at high [Chl a] (>30 mg m^{-3}). See further discussion of this in chapter two. Error bars were applied on the basis of a 5% spectral error for low biomass (<10 mg m^{-3}) and 30% for high biomass (>10 mg m^{-3}).

Changes in spectral shape as a result of instrument error/processing are likely to have significant impacts, particularly on the detection of size related variability. In particular, reduced resolution of the peaks in the red/red edge region of the spectrum could substantially limit information on size related variability, particularly at high biomass. Therefore much care and attention should be given to the use of atmospheric correction when the EAP is applied to ensure that these features and others such as those imparted by accessory pigments are maintained. Although the *in situ* and satellite derived R_{rs} in the case studies above show largely coherent

trends in spectral shape and magnitude, further work on the suitability of atmospheric correction schemes will ensure that size related variability, which is already relatively minimal, is not lost through inappropriate atmospheric correction.

The variant of the EAP algorithm applied here may benefit from additional developments as indicated by the performance here and in previous chapters. The inverse ocean colour problem has been termed as highly ambiguous (Defoin-Platel and Chami, 2007) and options to reduce ambiguity centre around suitable constraints applied to maintain naturally observed covariance in parameters (e.g. IOPs or biogeochemical parameters) (Defoin-Platel and Chami, 2007; Rehm and Mobley, 2013). This should help to prevent mathematical methods such as non-linear optimisation from selecting parameters outside the realm of possibility even though a suitable modelled convergence solution is found, whilst maintaining the freedom to select from large ranges of parameter combinations which can be limited under *in situ* and simulated data trained neural networks. Given the knowledge of abundance scaling laws (Agustí et al., 1987; Irwin et al., 2006) and as can be seen from the success of abundance approaches applied to satellite ocean colour data (Uitz et al., 2010, 2006, 2008), it would seem appropriate to constrain the EAP approach to incorporate these laws. In practice, this would mean limiting the range of effective diameters (and associated IOPs) that could be accessed by the optimisation approach when estimated [Chl *a*] is within a set range. This could limit the "speckling" artifacts seen under low biomass conditions in the case studies above. Determination of exact ranges over which to set these constraints requires further investigation of the relationship between [Chl *a*] and D_{eff} particularly under low biomass conditions, where the *in situ* validation set used in chapter four is particularly lacking. Similarly, it would be appropriate to constrain the estimates of a_{gd} and b_{bs} either through a minimum and maximum range of values or through a relationship to [Chl *a*]. Again this would require further *in situ* measurements to determine the extent of covariance.

Figure 5.10 indicated that the 2v variant of the EAP algorithm (where the inversion method is allowed to select an admixture of diatom/dinoflagellate and cryptophyte IOP basic vectors) was unable to differentiate and apply suitable basis vectors to blooms dominated by *Mesodinium rubrum*. However figure 5.11 suggested that a band ratio could be used to detect *Mesodinium rubrum* at high biomass. The failure of the EAP inversion approach and the band ratio technique (at low biomass) to distinguish *Mesodinium rubrum* is apparent when figure 5.20 is compared to the corresponding figure using cryptophyte basis vectors (Fig. 5.21). At the MERIS wavelengths, the differences between R_{rs} generated using the diatom/dinoflagellate and cryptophyte is relatively minimal. The only substantial difference is a shift towards higher R_{rs} around 520 nm under the cryptophyte basis vectors, as exploited by the ratio approach used in figure 5.11. The shift from hyperspectral (chapters two and three) to multispectral, results in less of the variability in spectral shape associated with phycoerythrin being captured (see chapter two, Fig. 2.5).

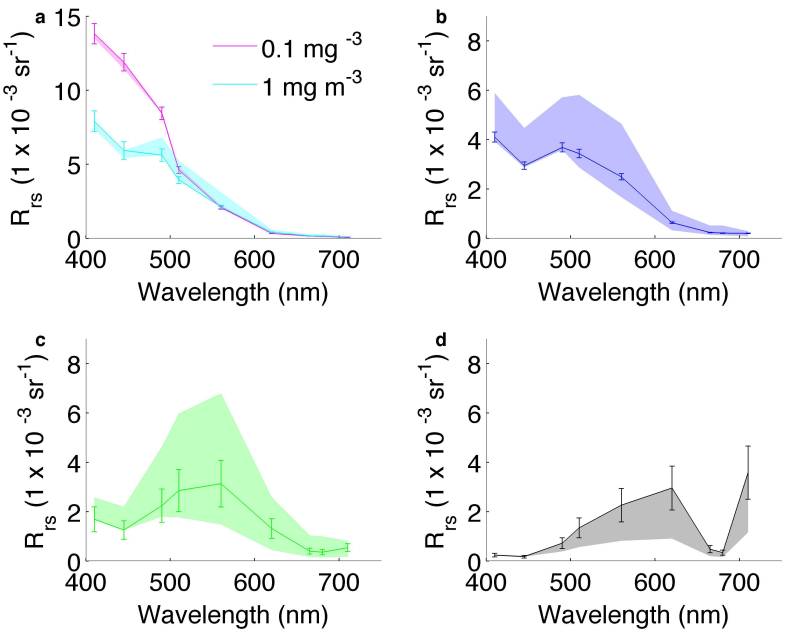

Figure 5.21: R_{rs} spectra from the forward model with associated error estimates for [Chl a] of (a) 0.1 mg m^{-3}, (b) 3 mg m^{-3}, (c) 10 mg m^{-3} and (d) 100 mg m^{-3}. The solid line represents the spectra generated with a D_{eff} of 10 μm. The shaded range shows the range of R_{rs} associated with changes in size. Results here using the forward model with SCI, cryptophyte basis vectors, low a_{gd} and low b_{bs} and Reflectance approximation. A filter has been applied to remove unrealistic spectra i.e. dominance of large cells (+20 μm) at low [Chl a] (<10 mg m^{-3}) and small cells (<10 μm) at high [Chl a] (>30 mg m^{-3}). See further discussion of this in chapter two. Error bars were applied on the basis of a 5% spectral error for low biomass (<10 mg m^{-3}) and 30% for high biomass (>10 mg m^{-3}).

5.4.2 Validation and comparison to other approaches for deriving information about phytoplankton community cell size

As mentioned in the introduction to this chapter, the EAP algorithm represents a fundamentally different approach to empirical and other semi-analytical techniques for determining cell size/PFT. As such, comparison to other algorithms is particularly difficult as already highlighted by reviews of existing approaches (Brewin et al., 2011b). Additionally, most applications of cell size/PFT detection algorithms developed thus far have been applied to lower resolution satellite data in the open ocean, compared to the application of the EAP in the southern Benguela which necessitates higher temporal and spatial resolution data due to patchiness and high variable physical dynamics (i.e. 3-6 day upwelling cycles). Further, the range of application in terms of biogeochemical (and consequently IOP) variability in the southern Benguela is substantially wider than other studies, particularly in terms of [Chl a]. For some key examples from the literature, the following ranges of [Chl a] were covered: 0.004 to 3 mg m^{-3} (Alvain et al., 2005), 0.01 to 10 mg m^{-3} (Uitz et al., 2006), > 0.01 to < 10 mg m^{-3} (Ciotti and Bricaud, 2006). Finally, the ways in which techniques are validated varies, such that even broad performance statistics provided are not entirely comparable due to varying levels of independence. Nevertheless, the performance of the EAP can be contextualised through a qualitative understanding of

its performance versus other approaches.

The abundance method of Uitz et al. (2006) was adapted for investigation of primary production in the Benguela by Silió-Calzada et al. (2008). Contributions of 3 size classes to total biomass ([Chl a]) was derived and primary production closest to the shore was attributed to microphytoplankton (>20 μm) with nanoplankton (2-20 μm) showing a stronger contribution offshore. This is consistent with the broad trends observed in the EAP D_{eff} output across the case studies and through the differences between the inshore (Fig. 5.13(b)) and bay scale (Fig. 5.13(a)) averages.

Ciotti and Bricaud (2006) applied a semi-analytical approach based on variable phytoplankton absorption to the continental shelf region off Brazil. Using both spectral decomposition and non-linear optimisation, Ciotti and Bricaud (2006) found that estimates of a phytoplankton cell size parameter (0 - 1 where 1 is $<$ 2 μm and 0 is $>$ 20 μm) derived from satellite retrieved absorption compared favourably to those derived from *in situ* absorption. Root Mean Squared Errors for the size parameter were lowest using a non-linear optimisation method applied to total absorption to derive the size parameter. Results suggested that application to multi-spectral rather than hyperspectral data did not significantly affect the retrieval of the size parameter, however the use of non-linear optimisation did help include more of the variability in spectral shape. Although based on IOP measurements, this application differs from the EAP approach in that it only used absorption, and that the cell size parameter is not coupled to the derivation of total absorption or reflectance and is instead dependent on accurate retrieval of absorption parameters from a suitable semi-analytical inversion method. It is worth noting here that the validation of returned size parameters in Ciotti and Bricaud (2006) was based upon comparison to those derived from *in situ* absorption spectra, rather than to direct measurements of particle size through pigment or other particle sizing techniques.

Kostadinov et al. (2009) developed a method based on the slope of backscattering coefficients derived from a semi-analytical algorithm to infer information about the particle size distribution which was subsequently applied to global SeaWiFS data (Kostadinov et al., 2010). Highest uncertainties in their method were found in regions of high productivity, which is somewhat counter to the findings presented in this and previous chapters about increased availability of size based information in high biomass conditions. However (as discussed by (Kostadinov et al., 2009)), it is possible that their method of parameterising the particle size distribution (PSD) is less suitable for high biomass waters (Bernard et al., 2007; Chami et al., 2006) and could result in inaccurate relationships with backscattering in these waters as a result of the assumptions made with regards to Jungian distributions and Mie theory. The EAP approach, with its inherent assumptions to counter the disadvantages of Mie theory, should provide more appropriate phytoplankton backscattering estimates and therefore, relationships with size at high biomass. Indeed the backscattering from the two-layered sphere would violate the assumptions used by Kostadinov et al. (2009) of constancy of constituent IOP spectral shapes. A further key difference in the two approaches, is that Kostadinov et al. (2009) characterises all particle

backscattering as one combined parameter, whilst the EAP separates these two components and applies a simple power law for non algal particle backscattering. Despite these differences, and the fact that Kostadinov et al. (2009) applied their method to much lower resolution satellite data, the broad conclusions made indicating that eastern boundary upwelling systems are dominated by microplankton (20 - 50 μm) are consistent with the results shown here for St Helena Bay (Fig. 5.13).

Kostadinov et al. (2010) also highlighted the difficulty of intercomparison of PFT approaches and the need for consistent and varied validation data across the global ocean to include pigment analysis and measures of the particle size distribution. In particular, Kostadinov et al. (2010) discussed the fundamental differences between approaches using pigment rather than explicit size based parameterisations. Use of pigments to characterise size variability is somewhat limited due to the influence of physiology on pigments and the inconsistency of chemotaxonomy across very different trophic regimes, and thus may not represent true validation of size (*ibid*). The use of the effective diameter parameterisation in the EAP algorithm may also be affected by this variability, if variability in pigments and ultimate effects on absorption are misinterpreted by the optimisation as variability in size. This is entirely possible in the context of a highly ambiguous problem with multiple covarying parameters.

5.4.3 New insights into phytoplankton variability in the southern Benguela and relationships between bloom occurrence and physical forcing factors.

Analysis of the time series in figure 5.12 indicates no significant trend in average [Chl *a*] over the St Helena Bay region from 2002 to 2012 using the MPH product and a small negative trend for Algal 1, ($r^2 = 0.009$, $p < 0.01$) and a negative but only weakly significant negative trend for the EAP [Chl *a*] ($r^2 = 0.003$, $p = 0.03$) respectively). The use of the EAP product in both the case studies and metrics shows that there is much greater variability between the inshore and offshore region than is possible to quantify with the use of standard [Chl *a*] products such as Algal 1. This variability has not previously been observed in other studies using satellite data in the St Helena Bay region. [Chl*a*] from the EAP is strongly correlated with the D_{eff} product ($r^2 = 0.53$ $p < 0.01$). This is expected given typical abundance scaling laws which govern cell size (Agustí et al., 1987; Irwin et al., 2006) but also confirms that the unconstrained inversion approach is in general retrieving coincident estimates of [Chl *a*] and D_{eff} which are in line with reality.

Evidence based on reported incidence has suggested an increase in both occurrence and severity of HABs in the southern Benguela region (Stephen and Hockey, 2007). This is not corroborated in the study here. It might not be expected to see an increase in HABs in the average regional [Chl *a*] trends. Further, the current time series of 11 years is likely much too short to make firm conclusions about long term trends (Henson et al., 2010), however, applied to future sensors (i.e. OLCI aboard Sentinel-3 and other sensors with suitable band selections),

the EAP algorithm should allow long term time series of the full range of [Chl *a*] and D_{eff} to be generated for the southern Benguela and other regions with suitable parameterisation and constraint.

The case studies and average effective diameters confirm the spatial variability of particle size distributions described by Armstrong et al. (1987) and further discussed by Pitcher et al. (1992a) and Crichton et al. (2013), where larger cells are typically more dominant inshore and at frontal features, with smaller cells offshore. Size progression from smaller to larger cells over event scale upwelling cycle and upwelling season is also seen in the examples from 2005 and 2009 (Figs. 5.4, 5.5 and 5.9 respectively).

The basic principle of upwelling favourable winds driving enhanced phytoplankton growth is confirmed through comparison of the EAP [Chl *a*] times series in figure 5.12 with wind data derived from 2 *in situ* stations and SeaWinds archive shown in figure 5.22(a) which show similar annual cycles. The SeaWinds data for the bay correspond well to those collected at the *in situ* station at Cape Columbine. Wind data collected at the Lambert's Bay *in situ* station shows lower peak alongshore wind speeds, possibly the result of the more inland situation of the station. On closer inspection, it can be seen that years with particularly high bloom activity, identifiable through the averages in figure 5.12 and corroborated by the threshold method of identifying blooms (Fig. 5.15), do not appear to correspond with exceptionally high average wind conditions. Pitcher et al. (2014) showed that several years of high bloom activity (and resultant anoxia) occurred coincidently with years of high cumulative upwelling winds. Figure 5.22(b-d) extends the analysis of Pitcher et al. (2014) over the full MERIS archive. Cumulative upwelling winds (shown as total northerly displacement) in figure 5.22(b-d) suggest that years of high bloom activity occupy the mid range of cumulative wind conditions on the interannual scale. This is particularly clear from the Cape Columbine *in situ* data (Fig. 5.22(c)). It seems likely that the wind conditions associated with the years of highest bloom activity represent a balance between upwelling favourable winds that provide nutrients and suitably long periods of quiescent conditions which allow the largest cells to proliferate and use these nutrients under more stratified conditions.

A general model of the physical circulation associated with HAB formation in St Helena Bay was proposed by Pitcher and Nelson (2006) on the basis of *in situ* physical data and the distribution of different phytoplankton assemblages. This conceptual model is summarised in schematic form in figure 5.23, which shows the circulation (solid arrows) and concentration of HAB forming dinoflagellates (dashed lines) under (a) upwelling conditions and (b) relaxation/downwelling conditions. A coastal, equatorward jet is present offshore of Cape Columbine, whilst within the bay a retentive circulation develops consisting of an eddy in the lee of Cape Columbine and (under relaxation/downwelling conditions) a poleward current. The extent of the circulation, and this retention time within the bay are dependent on wind stress (Penven et al., 2000). Whilst increased retention under upwelling wind conditions may work to keep blooms within the bay, distribute them from the north to the south and introduce populations from south

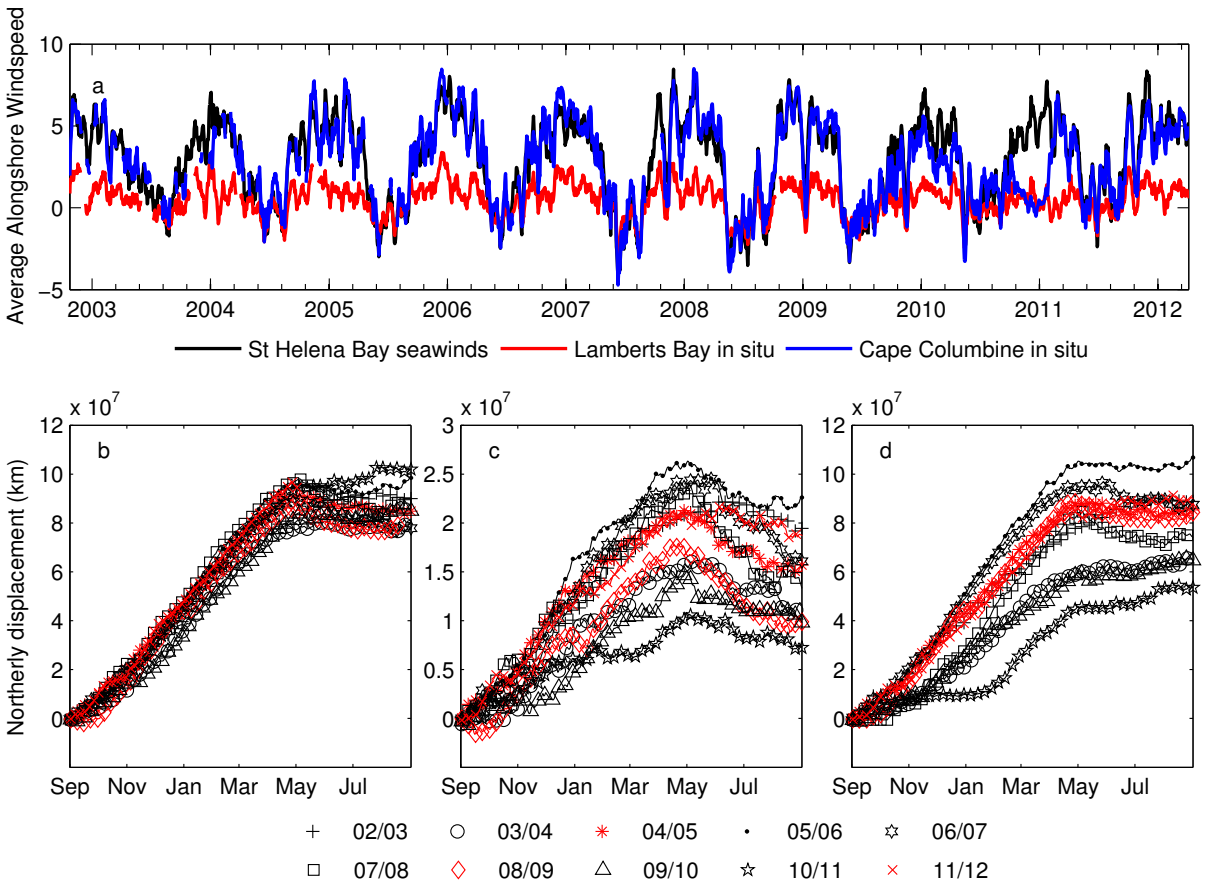

Figure 5.22: (a) Alongshore winds derived using a 30 day running mean. St Helena Bay averages were calculated from the SeaWinds product over the region detailed in figure 5.1 (black), whilst the Lambert's Bay (red) and Cape Columbine (blue) data were calculated from hourly measurements of wind speed at two *in situ* stations. Cumulative upwelling wind (in the form of northerly displacement) was calculated from the daily averages of wind from the three sources in panels (c), (d) and (e) respectively. Years with high bloom activity are highlighted in red in panels (c), (d) and (e) for clarity.

coast via the coastal jet (Pitcher et al., 2008b), increased wind also promotes mixing and can reduce stratification (Penven et al., 2000). Stratified conditions are most closely associated with the proliferation of the large dinoflagellates where as an increase in upwelling favourable winds can disperse a bloom and inhibit persistence (Pitcher and Calder, 2000; Pitcher and Nelson, 2006; Probyn et al., 2000). A balance between the two conditions of upwelling and subsequent relaxation would seem to provide the ideal environmental envelope for HABs dominated by large dinoflagellates to develop and be maintained and concentrated within the bay. This set of conditions may be exemplified by years associated with high bloom activity occupying the mid range of cumulative upwelling winds in figure 5.22(b-d).

Both the [Chl *a*] and D_{eff} products derived from the EAP algorithm applied to the MERIS time series provide substantial evidence in support of the conceptual model described in (Pitcher and Nelson, 2006) and figure 5.23. The years of high biomass detected in the time series (Fig. 5.12) and bloom occurrence and persistence metrics (Figs. 5.15, 5.17 and 5.18) all show evidence of the inshore poleward current (baroclinic instability (Penven et al., 2000)) predicted by Pitcher and Nelson (2006) and shown in figure 5.23(b). Figure 5.24 shows examples from the 2004/2005,

2008/2009 and 2011/2012 upwelling seasons. Identification of these features is only possible as a result of using algorithms (such as the EAP and MPH) which are optimised for high biomass conditions.

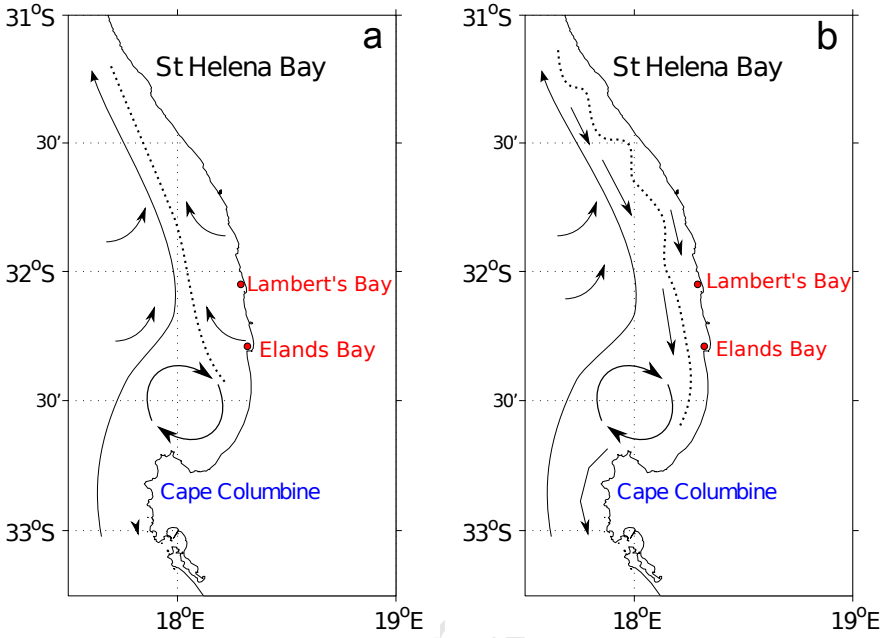

Figure 5.23: Schematic describing the physical circulation in St Helena Bay under (a) upwelling and (b) relaxation/downwelling with poleward current. An equatorward coastal jet and retentive eddy in the lee of Cape Columbine are present under both circumstances. Dashed lines indicate concentration of dinoflagellates. This figure is based on figure 15 in Pitcher and Nelson (2006).

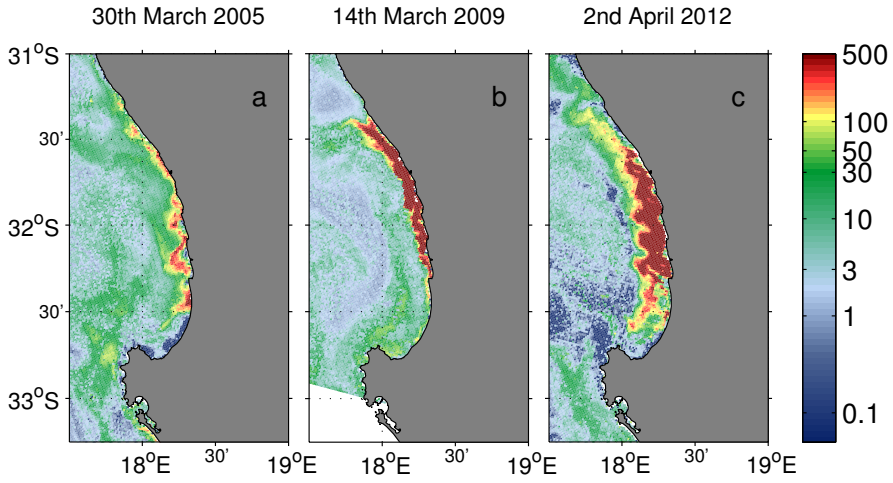

Figure 5.24: Examples of the EAP [Chl *a*] product for (a) 30th March 2005, (b) 14th March 2009 and (c) 2nd April 2012 showing distinctive wave features associated with years of high biomass (See Figs. 5.12 (a and b) and 5.15). Note the extended colour bar to further distinguish wave features.

Evidence of the poleward transport and retentive circulation in the Bay can also be found using a start date metric over several of the events highlighted in the case studies. The first date within the defined time period where each pixel exceeds the bloom threshold was extracted and recorded. Figure 5.25 shows the results of this bloom start date metric over the period of

February - May 2009 covered in case study 4 (Figs. 5.8 and 5.9). Southerly progression of the bloom can be seen in the escalating bloom start dates between the north and south of the bay. The formation of a retentive eddie circulation, entraining water from the offshore coastal jet can be seen during the middle part of this period in figure 5.8(e). Also evidenced in the bloom start dates (particularly Fig. 5.8(j) and Fig. 5.25(b)) is an example of the "flood" events described by Probyn et al. (2000), which result in HABs being transported south out of St Helena Bay and into bays south of Cape Columbine e.g. Saldanha Bay.

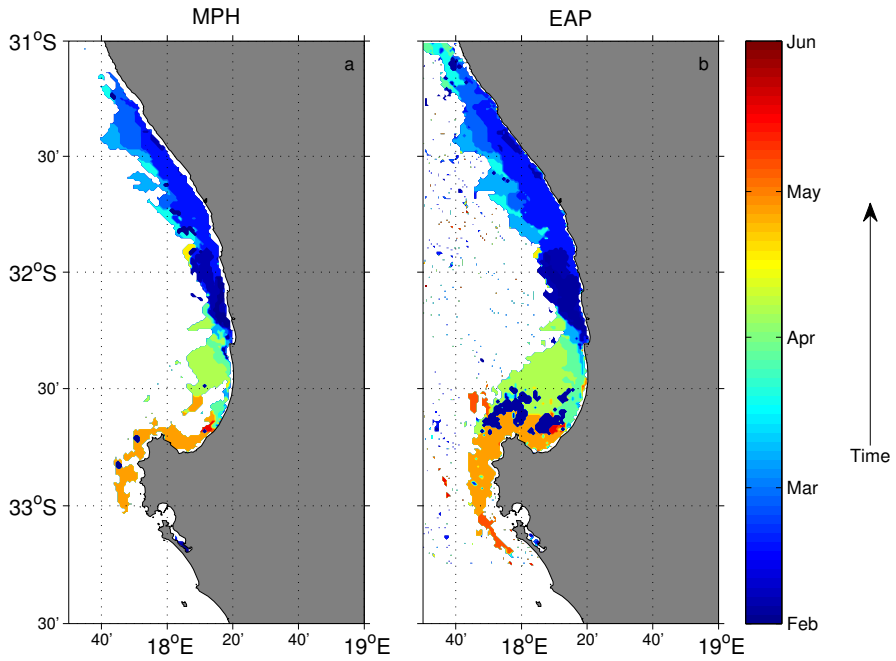

Figure 5.25: Bloom start dates between February and May 2009 derived using a 75 mg m^{-3} [Chl *a*] threshold applied to (a) the MPH and (b) the EAP [Chl *a*] products.

Associated with this physical circulation, transects conducted across St Helena Bay have found the onshore and offshore waters to be dominated by different species, thought to be a result of the differing sources of these waters (i.e. the coastal jet versus an inshore assemblage) (Pitcher et al., 2008b; Pitcher and Nelson, 2006). For example, Pitcher et al. (2008b) found that inshore, near Lambert's Bay, *A. catenella* dominated an algal bloom whilst offshore the assemblage became dominated by *G. polygramma*, presumably transported north in the coastal jet from the south coast where a bloom dominated by *G. polygramma* persisted in False Bay. This is consistent with the inshore and offshore trends in D_{eff} seen in figure 5.7

5.4.4 Further work required on phytoplankton seasonality and Harmful Algal Blooms in the southern Benguela

Whilst the ranges of [Chl *a*] and D_{eff} derived in this application of the EAP algorithm to the MERIS Level 2 archive, are broadly consistent with expected trends and *in situ* data in St Helena Bay several aspects can be identified that would improve the utility of this time series for assessing variability throughout the southern Benguela. Appropriate constraint and flagging

as discussed above should improve the coherency between inshore and offshore data and reduce the speckling of high values that are likely the result of ambiguity. For time series analysis and application of metrics, particularly those tracking bloom progression and persistence, it may be justifiable to implement filling procedures as has been done in previous studies (Weeks et al., 2006).

Further work is needed to analyse the suitability of wind products (satellite and *in situ*) for capturing the scales of variability which influence phytoplankton production. Although an envelope of bloom favourable conditions can be gleaned from the results here and supporting studies (Fawcett et al., 2008; Pitcher and Nelson, 2006; Pitcher et al., 2014) on a yearly basis, the event scale processes which lead to the formation of extreme blooms over the period of short term wind driven upwelling cycles needs further work. Metrics, which summarise the occurrence and persistence of wind conditions which are/are not favourable to phytoplankton production, will need to be developed and then related to phytoplankton community characteristics (biomass, size, physiology etc) and other forcing factors not considered (seeding, nutrient dynamics, grazing, viral control etc (see Anderson et al. (2008) and Smayda (2008) and references within.)

Whilst the seasonal variability in phytoplankton production in the southern Benguela can be easily attributed to wind forcing, there is substantial event scale variability associated with HABs which can not (at least according to the metrics employed here) be explicitly connected to wind variability, other than through a broad envelope of wind conditions which appear to lead to certain physical circulation features (Pitcher and Nelson, 2006). Hydrodynamic models have already been applied to understand the generation of such features in a theoretical way (Penven et al., 2000). Further development of these models for the St Helena Bay region could produce realistic response to wind forcing and allow for greater comparison to ocean colour data and the products developed in this chapter. The use of Lagrangian particle tracking to analyse potential scenarios for bloom transport and dispersion could help indicate the likelihood of risk for various parts of St Helena Bay (e.g. as applied to *Karenia brevis* bloom in Tampa Bay by Havens et al. (2010)) and potentially form part of a forecast system for blooms once initiated. Full forecasting would require a far greater understanding of other factors (not all easily observable from satellite) that can inhibit bloom initiation, proliferation and persistence. Relationships with nutrients based upon relationships with regularly observable parameters such as temperature (e.g. Goes et al. (2000)), understanding of phenology of grazers (Smayda, 2008), indication of bloom health/viral infection (e.g. Balch et al. (2007)) and species competition in response to these factors may aid the development of an ecological niche approach which could indicate the likelihood of success for different phytoplankton species (Bernard et al., 2006).

5.5 Conclusions and recommendations

The generation of a time series of inherent optical properties, chlorophyll *a* concentration ([Chl *a*]) and estimates of phytoplankton community cell size (effective diameter, D_{eff}) represents a substantial opportunity to develop insight into both the use of optical data and ecosystem dynamics in the southern Benguela. A number of conclusions can be made from the results shown in this chapter and several recommendations will allow this new resource to be fully exploited.

- Applications shown in this chapter indicate that the EAP algorithm can be used to provide suitable estimates of [Chl *a*] and D_{eff} associated with a variety of high biomass events. The [Chl *a*] product offers a substantial advantage over [Chl *a*] products from traditional blue/green empirical band ratio approaches (e.g. Algal 1) and is comparable to other empirical band ratio approaches developed for high biomass waters i.e. the MPH.

- Time series of [Chl *a*] from the EAP and MPH, offer new evidence of interannual variability in these phenomena.

- A bloom metric using a threshold technique supports this observation of interannual variability and the application of a persistence metric suggests that years of high bloom activity are also accompanied by spatially more persistent blooms.

- The wavebands available on the MERIS sensor, appear suitable for broad application of the EAP algorithm. As such the algorithm should also be applicable to the expected OLCI sensor aboard Sentinel 3 and possibly to Sentinel 2.

- However, the success of the algorithm will be reliant on appropriate atmospheric correction. Further investigation of atmospheric correction techniques in the southern Benguela is necessary to ensure full spectral data relating to algal biomass and characteristics such as size or functional type is available.

- With regards to further validation activities, the following recommendations are made: Estimates of patchiness over the typical scales of satellite data used in the region (i.e. 350m - 1km); matchup analysis using multiple radiometers (given the difficulties associated with *in situ* in water radiometry (see chapter four)); collection of both inshore and offshore/bloom and non-bloom data for validation; collection of a full suite of IOPs and biogeochemical parameters to assess the extent of covariance and suitability of parameterisations (e.g. for a_{gd} and b_{bs}).

- Whilst the EAP algorithm has shown success in application to inshore waters and HAB events, it is recommended that the MERIS archive be reprocessed with a constrained version of the EAP algorithm, to reduce ambiguity outside of these regions and ensure the provision of realistic cell size estimates in low biomass waters and prevent overestimation of a_{gd} and b_{bs}.

- For the purpose of time series analysis, a new gridded product with appropriate fill values may allow the further development of bloom metrics.
- Use of a suitable scale hydrodynamic model should expand on the extent to which advection and retentive circulations are controlled by wind variability and ultimately control environmental conditions and the distribution and persistence of HABs.

Chapter 6

Conclusions and further work

From the work contained within this thesis, the following conclusions and recommendations for further work can be made.

6.1 Conclusions

For the first time, the sensitivity in R_{rs} attributable to changes in phytoplankton community cell size has been quantified through the use of a fully coupled IOP and reflectance model - the Equivalent Algal Population (EAP) forward model:

- It was found that this sensitivity is spectrally variant and highly biomass dependent. Sensitivity increases substantially in the range of 3 -10 mg m^{-3} [Chl a] with maximum signal available when phytoplankton absorption and scattering become a dominant component in the Inherent Optical Property budget.
- As such, this sensitivity is also dependent on the presence and concentration of optically significant, non-phytoplankton components i.e. gelbstoff or small, highly scattering particles such as sediments. It is proposed that spectral IOP budgets be used in future to further consider the levels of ambiguity associated with second order variability (e.g. cell size or functional type) in different water types.
- At low concentrations of gelbstoff and non-algal particles, the reflectance approximation and more sophisticated radiative transfer techniques (in this case, EcoLight-S), produce similar simulated reflectance spectra across a range of biomass concentrations. The assumptions associated with the reflectance approximation begin to break down at higher biomass and with an increase in scattering. The use of EcoLight-S is more suitable in these cases (i.e. when the single scattering assumption is less valid).

The EAP forward model was then developed into a semi-analytical inversion algorithm and tested using both simulated and *in situ* data:

- A wide range of chlorophyll *a* concentrations (0.1 to 300 mg m^{-3}) can be retrieved using the EAP inversion algorithm under Case 1 type conditions applied to both simulated and *in situ* radiometric data from the southern Benguela upwelling system.

- Estimates of phytoplankton community cell size, as parameterised through the use of standard size distributions and the calculation of an effective diameter, can be retrieved using the EAP inversion algorithm. However, the accuracy of these retrievals is biomass dependent. Minimum likely errors for cell size prediction ranged between 1 μm and 8 μm under high and low biomass, Case 1 conditions respectively.

- Similarly, it is possible to distinguish between two groups of IOPs, representative of different functional types (diatoms/dinoflagellates vs cryptophytes and associated endosymbionts) when inverting hyperspectral simulated data. However applying this unconstrained approach to *in situ* data from a variety of mixed and mono-specific assemblages did not improve bulk performance statistics over using a single set of IOPs. This is likely the result of increased ambiguity from the additional degree of freedom and relatively smaller sensitivity to functional type specific IOPs in mixed assemblages *in situ*. Additionally the application to *in situ* data is likely further compounded by *in situ* measurement error.

- Dependent on the inversion technique used and the underlying IOP budget, close matches in measured and modelled reflectance can be achieved despite inaccurate biogeochemical parameters selected by the non linear optimisation. This indicates that the inverse problem (at least as parameterised here) can be highly ambiguous.

- Accurate estimates of phytoplankton community cell size were found to be dependent on coincident estimation of chlorophyll *a* i.e. if chlorophyll *a* concentration was estimated inaccurately by the inversion approach, it was likely that the size would be incorrectly estimated also, as a result of the inherent ambiguity in the inverse ocean colour problem.

- Use of more sophisticated radiative transfer models (EcoLight-S), as opposed to reflectance approximation techniques, did not necessarily decrease the ambiguity present in the inverse problem. An inversion approach using the reflectance approximation is suitable for deriving the necessary range of chlorophyll *a* concentrations and for distinguishing between assemblages dominated by small or large cells in simulated Case 1 type waters, and for *in situ* application in the southern Benguela.

To demonstrate potential applications, the EAP inversion algorithm was adapted for application to satellite derived radiometry at Level 2 (water leaving radiance/remote sensing reflectance):

- Results show substantial improvements over standard products (Algal 1) for the detection of the high levels of biomass associated with Harmful Algal Blooms.
- At the event scale, application to a number of case studies suggests that results from the EAP algorithm, in terms of chlorophyll *a* concentrations and community cell size are broadly consistent with available information from a range of case studies i.e. the ranges of predicted [Chl *a*] associated with bloom features are consistent with *in situ* estimates, cell size is consistent with those observed *in situ*.
- Application of the EAP inversion algorithm to a 10 year (2002 to 2012) archive of MERIS Level 2 data allows for substantial new insight into the spatial and temporal variability of Harmful Algal Blooms, particularly within the St Helena Bay region. For the first time, the full range of chlorophyll *a* concentrations associated with this highly dynamic upwelling region are available through satellite remote sensing.
- Using this new data, substantial interannual variability in the occurrence of Harmful Algal Blooms was discovered through the development of threshold and persistence metrics for quantifying high biomass blooms.
- Years of high bloom activity were not associated with extremely high or low cumulative wind conditions, but rather occurred in years of intermediate cumulative wind conditions. Wind dynamics and the phytoplankton community response are complex in both time and space. As such, these intermediate cumulative wind conditions presumably represent a suitable balance of upwelling and quiescent periods, providing suitable nutrient supply and stratification respectively, such that proliferation of bloom forming phytoplankton types is favoured.
- Years of high bloom activity were commonly associated with physical features indicative of poleward surface currents and the formation of a strongly retentive circulation in St Helena Bay.

6.2 Recommendations for further work

- A new simulated dataset, accounting for community structure across a broad range of trophic regimes from oligotrophic to eutrophic waters, should be developed. This new simulated dataset should be informed by *in situ* data and suitably constrained with regards to IOP covariance. This would extend the current understanding of EAP performance and ocean colour ambiguity, and potentially help extend the application to other regions.

- For Case 2 and eutrophic waters, it is recommended that models which solve the radiative transfer equation, such as those available through EcoLight-S are used where possible with respect to computational demands. In cases where this is prohibitive (i.e. under inversion techniques which demand many iterations), it is recommended that a comprehensive simulated dataset, provided with size dependent IOPs, be used to create new f/Q and K_{Lu} parameterisations.

- Further consideration is needed with regard to other sources of second order variability e.g. physiology and the additional uncertainty that this may impart on IOPs and derivation of biogeochemical parameters. Limitations of the IOP model with regards to physiology and species with unusual optical characteristics should also be considered and potentially incorporated into future algorithm variants. For example, the IOPs used in the current study are likely insufficient to account for elevated scattering from coccolithophores or vacuolate cyanobacteria.

- Algorithm performance would likely benefit from being suitably constrained, in particular to limit the selection of very large or very small effective diameter values under high and low biomass conditions respectively. Similar benefit would likely be achieved through better understanding of regional IOP budgets for gelbstoff/detritus absorption and small particle scattering.

- Further work is needed to establish suitable radiometric measurement protocols for sampling in high biomass and eutrophic waters. Preliminary comparisons conducted in this thesis suggest that fixed depth radiometry, with longer, depth specific sampling time, may help reduce errors which can be exacerbated by high frequency variability in the upper water column. In-water approaches should also be compared to those used in above-water radiometry. Further work is also needed to evaluate correction schemes for eutrophic waters e.g. Monte Carlo methods may assist in the development of appropriate self shading corrections for eutrophic and highly scattering waters.

- Most accurate results over a wide range of optical conditions may come from using the EAP as part of a suite of regional algorithms i.e. using regionally specific empirical band ratios to provide [Chl *a*] under very low concentrations, to constrain ranges of parameters and to select different IOPs as a prerequisite (i.e. using crytophyte basis vectors when empirical band ratios indicate the likelihood of significant phycoerythrin absorption).

- The components of the EAP forward model offer a powerful means of parameterising phytoplankton community variability. Use of similar size based metrics in a hydrodynamically driven ecosystem model, coupled to EcoLight-S could substantially extend the satellite ocean colour perspective. In the southern Benguela this could be used to investigate the response of physical circulation to wind forcing and subsequent bloom formation and transport within St Helena Bay. This should allow progress towards developing statistical models and predictive physical models indicating the risk of bloom occurrence and persistence.

References

Aas, E. (1987). Two-stream irradiance model for deep waters. *Applied Optics*, 26(11):2095–2101.

Acevedo-Trejos, E., Brandt, G., Merico, A., and Smith, S. L. (2013). Biogeographical patterns of phytoplankton community size structure in the oceans. *Global Ecology and Biogeography*, 22(9):1060–1070.

Agustí, S. (1991). Light environment within dense algal populations: cell size influences on self-shading. *Journal of Plankton Research*, 13(4):863–871.

Agustí, S., Duarte, C. M., and Kalff, J. (1987). Algal cell size and the maximum density and biomass of phytoplankton. *Limnology and Oceanography*, 32(4):983–986.

Ahn, Y.-H., Bricaud, A., and Morel, A. (1992). Light backscattering efficiency and related properties of some phytoplankters. *Deep Sea Research*, 39(11/12):1835–1855.

Aiken, J., Fishwick, J. R., Lavender, S., Barlow, R., Moore, G. F., Sessions, H., Bernard, S., Ras, J., and Hardman-Mountford, N. J. (2007). Validation of MERIS reflectance and chlorophyll during the BENCAL cruise October 2002: preliminary validation of new demonstration products for phytoplankton functional types and photosynthetic parameters. *International Journal of Remote Sensing*, 28(3-4):497–516.

Aiken, J., Hardman-Mountford, N. J., Barlow, R., Fishwick, J., Hirata, T., and Smyth, T. (2008). Functional links between bioenergetics and bio-optical traits of phytoplankton taxonomic groups: an overarching hypothesis with applications for ocean colour remote sensing. *Journal of Plankton Research*, 30(2):165–181.

Aiken, J. and Moore, G. (2000). Algorithm Theoretical Basis Document. ATBD 2.6. Case 2 (S) Bright Pixel Atmospheric Correction. Technical Report PO-TN-MEL-GS-0005, Plymouth, UK.

Albert, A. and Mobley, C. D. (2003). An analytical model for subsurface irradiance and remote sensing reflectance in deep and shallow case-2 waters. *Optics Express*, 11(22):2873–2890.

Alexandrov, M. D. and Lacis, A. A. (2000). A new three-parameter cloud/aerosol particle size distribution based on the generalized inverse Gaussian density function. *Applied Mathematics and Computation*, 116:153–165.

Allen, J., Brown, L., Sanders, R., Moore, C. M., Mustard, A., Fielding, S., Lucas, M., Rixen, M., Savidge, G., Henson, S., and Mayor, D. (2005). Diatom carbon export enhanced by silicate upwelling in the northeast Atlantic. *Nature*, 437:728–732.

Alvain, S., Loisel, H., and Dessailly, D. (2012). Theoretical analysis of ocean color radiances anomalies and implications for phytoplankton groups detection in case 1 waters. *Optics Express*, 20(2):1070–1083.

Alvain, S., Moulin, C., Dandonneau, Y., and Bréon, F. M. (2005). Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep Sea Research I*, 52:1989–2004.

Alvain, S., Moulin, C., Dandonneau, Y., and Loisel, H. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: A satellite view. *Global Biogeochemical Cycles*, 22:GB3001.

Álvarez, E., Lopez-Urrutia, A., Nogueira, E., and Fraga, S. (2011). How to effectively sample the plankton size spectrum? A case study using FlowCAM. *Journal of Plankton Research*, 33(7):1119–1133.

Ambarwulan, W., Salama, M. S., Mannaerts, C. M., and Verhoef, W. (2010). Estimating specific inherent optical properties of tropical coastal waters using bio-optical model inversion and in situ measurements: case of the Berau estuary, East Kalimantan, Indonesia. *Hydrobiologia*, 658(1):197–211.

Amin, R., Zhou, J., Gilerson, A., Gross, B., Moshary, F., and Ahmed, S. (2009). Novel optical techniques for detecting and classifying toxic dinoflagellate Karenia brevis blooms using satellite imagery. *Optics Express*, 17(11):9126–9144.

Anderson, D. M., Burkholder, J. M., Cochlan, W. P., Glibert, P. M., Gobler, C. J., Heil, C. A., Kudela, R. M., Parsons, M. L., Rensel, J. E. J., Townsend, D. W., Trainer, V. L., and Vargo, G. A. (2008). Harmful algal blooms and eutrophication: Examining linkages from selected coastal regions of the United States. *Harmful Algae*, 8(1):39–53.

Anderson, T. (2005). Plankton functional type modelling: running before we can walk? *Journal of Plankton Research*, 27(11):1073–1081.

Andrews, S. W., Nover, D. M., Reuter, J. E., and Schladow, S. G. (2011). Limitations of laser diffraction for measuring fine particles in oligotrophic systems: Pitfalls and potential solutions. *Water Resources Research*, 47:W05523.

Antoine, D., d'Ortenzio, F., Hooker, S. B., Bécu, G., Gentili, B., Tailliez, D., and Scott, A. J. (2008). Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color sensors (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean Sea (BOUSSOLE project). *Journal of Geophysical Research*, 113:C07013.

Antoine, D. and Morel, A. (1999). A multiple scattering algorithm for atmospheric correction of remotely sensed ocean colour(MERIS instrument): principle and implementation for atmospheres carrying various aerosols including absorbing ones. *International Journal of Remote Sensing*, 20(9):1875–1916.

Antoine, D., Siegel, D. A., Kostadinov, T. S., Maritorena, S., Nelson, N. B., Gentili, B., Vellucci, V., and Guillocheau, N. (2011). Variability in optical particle backscattering in contrasting bio-optical oceanic regimes. *Limnology and Oceanography*, 56(3):955–973.

Arin, L., Morán, X. A. G., and Estrada, M. (2002). Phytoplankton size distribution and growth rates in the Alboran Sea (SW Mediterranean): short term variability related to mesoscale hydrodynamics. *Journal of Plankton Research*, 24(10):1019–1033.

Armstrong, D. A., Mitchell-Innes, B. A., Verheye-Dua, F., Waldron, H., and Hutchings, L. (1987). Physical and biological features across an upwelling front in the southern Benguela. *South African Journal of Marine Science*, 5(1):171–190.

Aurin, D. A. and Dierssen, H. M. (2012). Advantages and limitations of ocean color remote sensing in CDOM-dominated, mineral-rich coastal and estuarine waters. *Remote Sensing of Environment*, 125:181–197.

Babin, M., Morel, A., Fournier-Sicre, V., Fell, F., and Stramski, D. (2003). Light scattering properties of marine particles in coastal and open ocean waters as related to the particle mass concentration. *Limnology and Oceanography*, 48(2):843–859.

Bailey, S. W. and Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment*, 102:12–23.

Balch, W. M., Vaughn, J. M., Goes, J. I., Novotny, J. F., Drapeau, D. T., Booth, E. S., and Vining, C. L. (2007). Bio-optical consequences of viral infection of phytoplankton: I. Experiments with the cyanobacterium, Synechococcus sp. *Limnology and Oceanography*, 52(2):727–738.

Banse, K. (1976). Rates of growth, respiration and photosynthesis of unicellular algae as related to cell size - a review. . *Journal of Phycology*, 12:135–140.

Barker, K., Mazeran, C., Lerebourg, C., Bouvet, M., Antoine, D., Ondrusek, M., Zibordi, G., and Lavender, S. J. (2008). MERMAID: The MEris MAtchup In-situ Database. *2nd MERIS (A)ATSR Users Workshop*.

Barker, K., Smyth, T., Lavender, S. J., and Aiken, J. (2007). A Novel Technique to Estimate Primary Production Directly From Earth Observation Data: An Inherent Optical Property Approach. *Proceedings "Envisat symposium" 2007*.

Barlow, R. G. (1982). Phytoplankton ecology in the southern Benguela current. I. Biochemical composition. *Journal of Experimental Marine Biology and Ecology*, 63(3):209–227.

Barlow, R. G., Sessions, H., Siliulwane, N., Engel, H., Hooker, S., Aiken, J., Fishwick, J., Vicente, V., Morel, A., Chami, M., Ras, J., Bernard, S., Pfaff, M., Brown, J. W., and Fawcett, A. (2003). 2003: BENCAL Cruise Report, NASA/TM 2003-206892. Technical report, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

Bartlett, J. S., Voss, K. J., Sathyendranath, S., and Vodacek, A. (1998). Raman scattering by pure water and seawater. *Applied Optics*, 37(15):3324–3332.

Batten, S. D., Clark, R., Flinkman, J., Hays, G., John, E., John, A. W. G., Jonas, T., Lindley, J. A., Stevens, D. P., and Walne, A. (2003). CPR sampling: the technical background, materials and methods, consistency and comparability. *Progress in Oceanography*, 58:193–215.

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

Bergmann, T., Fahnenstiel, G., Lohrenz, S., Millie, D., and Schofield, O. (2004). Impacts of a recurrent resuspension event and variable phytoplankton community composition on remote sensing reflectance. *Journal of Geophysical Research*, 109:C10S15.

Bergquist, A., Carpenter, S., and Latino, J. (1985). Shifts in phytoplankton size structure and community composition during grazing by contrasting zooplankton assemblages. *Limnology and Oceanography*, 30(5):1037–1045.

Bernard, S., Balt, C., Pitcher, G. C., Probyn, T. A., Fawcett, A., and Du Randt, A. (2005). The use of MERIS for harmful algal bloom monitoring in the Southern Benguela. In Lacoste, H., editor, *Proceedings of the MERIS (A)ATSR Workshop 2005 (ESA SP-597)*, Frascati, Italy.

Bernard, S., Kudela, R. M., Franks, P. J. S., Fennel, W., Kemp, A., Fawcett, A., and Pitcher, G. C. (2006). The Requirements for Forecasting Harmful Algal Blooms in the Benguela. *Large Marine Ecosystems*, 14:281–302.

Bernard, S., Pitcher, G. C., Evers-King, H., Robertson, L., Matthews, M. W., Rabagliati, A., and Balt, C. (2014). Ocean Colour Remote Sensing of Harmful Algal Blooms in the Benguela System. In Barale, V. and Gade, M., editors, *Remote Sensing of the African Seas*, pages 185–203. Springer.

Bernard, S., Probyn, T. A., and Barlow, R. G. (2001). Measured and modelled optical properties of particulate matter in the southern Benguela. *South African Journal of Science*, 97(9):410–420.

Bernard, S., Probyn, T. A., and Quirantes, A. (2009). Simulating the optical properties of phytoplankton cells using a two-layered spherical geometry. *Biogeosciences Discussion*, 6:1–67.

Bernard, S., Probyn, T. A., and Shillington, F. A. (1998). Towards the validation of SeaWiFS in southern African waters: the effects of gelbstoff. *South African Journal of Marine Science*, 19(1):15–25.

Bernard, S., Shillington, F. A., and Probyn, T. A. (2007). The use of equivalent size distributions of natural phytoplankton assemblages for optical modeling. *Optics Express*, 15(5):1995–2007.

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

Blondeau-Patissier, D., Gower, J. F. R., Dekker, A. G., Phinn, S. R., and Brando, V. E. (2014). A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans. *Progress in Oceanography*, 123:123–144.

Boyd, P. and Newton, P. (1995). Evidence of the potential influence of planktonic community structure on the interannual variability of particulate organic carbon flux. *Deep Sea Research I*, 42(5):619–639.

Brando, V. E. and Dekker, A. G. (2003). Satellite hyperspectral remote sensing for estimating estuarine and coastal water quality. *IEEE Transactions on Geoscience and Remote Sensing*, 41:1378–1387.

Brewin, R. J. W., Devred, E., Sathyendranath, S., Lavender, S. J., and Hardman-Mountford, N. J. (2011a). Model of phytoplankton absorption based on three size classes. *Applied Optics*, 50(22):4535–4549.

Brewin, R. J. W., Hardman-Mountford, N. J., Lavender, S. J., Raitsos, D., Hirata, T., Uitz, J., Devred, E., Bricaud, A., and Gentili, B. (2011b). An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sensing of Environment*, 115:325–339.

Brewin, R. J. W., Lavender, S. J., and Hardman-Mountford, N. J. (2010). Mapping size-specific phytoplankton primary production on a global scale. *Journal of Maps Student Edition*, pages 12–25.

Bricaud, A., Babin, M., Morel, A., and Claustre, H. (1995a). Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization. *Journal of Geophysical Research*, 100:13321–13332.

Bricaud, A., Bédhomme, A.-L., and Morel, A. (1988). Optical properties of diverse phytoplanktonic species: experimental results and theoretical interpretation. *Journal of Plankton Research*, 10(5):851–873.

Bricaud, A. and Morel, A. (1983). Optical efficiency factors of some phytoplankters. *Limnology and Oceanography*, 28(5):816–832.

Bricaud, A. and Morel, A. (1986). Light attenuation and scattering by phytoplanktonic cells: a theoretical modeling. *Applied Optics*, 25(4):571–580.

Bricaud, A., Morel, A., and Prieur, L. (1981). Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains. *Limnology and Oceanography*, 26(1):43–53.

Bricaud, A., Roesler, C., and Zaneveld, J. R. V. (1995b). In situ methods for measuring the inherent optical properties of ocean waters. *Limnology and Oceanography*, 40(2):393–410.

Brody, S. R., Susan Lozier, M., and Dunne, J. P. (2013). A comparison of methods to determine phytoplankton bloom initiation. *Journal of Geophysical Research*, 118:1–13.

Brown, C. A., Huot, Y., Werdell, P. J., Gentili, B., and Claustre, H. (2008). The origin and global distribution of second order variability in satellite ocean color and its potential applications to algorithm development. *Remote Sensing of Environment*, 112:4186–4203.

Cannizzaro, J. P., Carder, K. L., Chen, F. R., Heil, C. A., and Vargo, G. A. (2008). A novel technique for detection of the toxic dinoflagellate, Karenia brevis, in the Gulf of Mexico from remotely sensed ocean color data. *Continental Shelf Research*, 28(1):137–158.

Carder, K. L., Chen, F. R., Lee, Z. P., and Hawes, S. K. (1999). Semianalytic Moderate-Resolution Imaging Spectrometer algorithms for chlorophyll a and absorption with bio-optical domains based on nitrate-depletion temperatures. *Journal of Geophysical Research*, 104:5403–5421.

Carder, K. L., Steward, R. G., Harvey, G., and Ortner, P. B. (1989). Marine Humic and Fulvic-Acids - Their Effects on Remote-Sensing of Ocean Chlorophyll. *Limnology and Oceanography*, 34(1):68–81.

Chami, M. and Defoin-Platel, M. (2007). Sensitivity of the retrieval of the inherent optical properties of marine particles in coastal waters to the directional variations and the polarization of the reflectance. *Journal of Geophysical Research*, 112:C05037.

Chami, M., Marken, E., Stamnes, J., Khomenko, G. A., and Korotaev, G. K. (2006). Variability of the relationship between the particulate backscattering coefficient and the volume scattering function measured at fixed angles. *Journal of Geophysical Research*, 111:C05013.

Chami, M. and Robilliard, D. (2002). Inversion of oceanic constituents in case I and II waters with genetic programming algorithms. *Applied Optics*, 41(30):6260–6275.

Chami, M., Thirouard, A., and Harmel, T. (2014). POLVSM (Polarized Volume Scattering Meter) instrument: an innovative device to measure the directional and polarized scattering properties of hydrosols. *Optics Express*, 22(21):26403.

Chang, C.-H., Liu, C.-C., and Wen, C.-G. (2007a). Integrating semianalytical and genetic algorithms to retrieve the constituents of water bodies from remote sensing of ocean color. *Optics Express*, 15(2):252–265.

Chang, G., Barnard, A., and Zaneveld, J. R. V. (2007b). Optical closure in a complex coastal environment: particle effects. *Applied Optics*, 46(31):7679–7692.

Chang, G. and Whitmire, A. L. (2009). Effects of bulk particle characteristics on backscattering and optical closure. *Optics Express*, 17(4):2132–2142.

Chang, G. C., Dickey, T. D., Mobley, C. D., Boss, E., and Pegau, W. S. (2003). Toward Closure of Upwelling Radiance in Coastal Waters. *Applied Optics*, 42(9):1574–1582.

Chen, Z., Curran, P. J., and Hansom, J. D. (1992). Derivative reflectance spectroscopy to estimate suspended sediment concentration. *Remote Sensing of Environment*, 40:67–77.

Chisholm, S. W. (1992a). Phytoplankton Size. In *Primary productivity and biogeochemical cycles in the ...*, pages 213–237. Springer US, Boston, MA.

Chisholm, S. W. (1992b). Phytoplankton Size. In Falkowski, P. G. and Woodhead, A. D., editors, *Primary Productivity and Biogeochemical Cycles in the sea*, pages 213–237. Plenum Press, New York.

Chiswell, S. M. (2011). Annual cycles and spring blooms in phytoplankton: don't abandon Sverdrup completely. *Marine Ecology Progress Series*, 443:39–50.

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

Ciotti, A. M., Cullen, J. J., and Lewis, M. R. (1999). A semi-analytical model of the influence of phytoplankton community structure on the relationship between light attenuation and ocean color. *Journal of Geophysical Research*, 104:1559–1578.

Ciotti, Á. M., Lewis, M. R., and Cullen, J. J. (2002). Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. *Limnology and Oceanography*, 47(2):404–417.

Clavano, W. R., Boss, E., and Karp-Boss, L. (2007). Inherent optical properties of non-spherical marine-like particles - From theory to observation. *Oceanography and Marine Biology*, 45:1–38.

Cockcroft, A. C. (2001). Jasus lalandii'walkouts' or mass strandings in South Africa during the 1990s: an overview. *Marine and Freshwater Research*, 52(8):1085–1093.

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

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

Crichton, M., Hutchings, L., Lamont, T., and Jarre, A. (2013). From physics to phytoplankton: prediction of dominant cell size in St Helena Bay in the Southern Benguela. *Journal of Plankton Research*, 35(3):526–541.

Dall'Olmo, G., Westberry, T. K., Behrenfeld, M. J., Boss, E., and Slade, W. H. (2009). Direct contribution of phytoplankton-sized particles to optical backscattering in the open ocean. *Biogeosciences Discussions*, 6:291–340.

Defoin-Platel, M. and Chami, M. (2007). How ambiguous is the inverse problem of ocean color in coastal waters? *Journal of Geophysical Research*, 112:C03004.

Demarcq, H., Barlow, R., and Hutchings, L. (2007). Application of a chlorophyll index derived from satellite data to investigate the variability of phytoplankton in the Benguela ecosystem. *African Journal of Marine Science*, 29(2):271–282.

Demarcq, H., Barlow, R., and Shillington, F. A. (2003). Climatology and variability of sea surface temperature and surface chlorophyll in the Benguela and Agulhas ecosystems as observed by satellite imagery. *African Journal of Marine Science*, 25(1):363–372.

Dierssen, H. M., Kudela, R. M., Ryan, J. P., and Zimmerman, R. C. (2006). Red and black tides: Quantitative analysis of water-leaving radiance and perceived color for phytoplankton, colored dissolved organic matter, and suspended sediments. *Limnology and Oceanography*, 51(6):2646–2659.

Dierssen, H. M., Zimmerman, R. C., Leathers, R. A., Downes, T. V., and Davis, C. O. (2003). Ocean Color Remote Sensing of Seagrass and Bathymetry in the Bahamas Banks by High-Resolution Airborne Imagery. *Limnology and Oceanography*, 48(1, part 2):444–455.

Doerffer, R. and Schiller, H. (2007). The MERIS Case 2 water algorithm. *International Journal of Remote Sensing*, 28(3-4):517–535.

Doerffer, R., Schiller, H., and Krasemann, H. (2002). MERIS Case 2 water validation early results North Sea/Helgoland/German Bight. In *Proceedings of the ....*

Doney, S. C., Lima, I., Moore, J. K., Lindsay, K., Behrenfeld, M. J., Westberry, T. K., Mahowald, N., Glover, D. M., and Takahashi, T. (2009). Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data. *Journal of Marine Systems*, 76:95–112.

Doyle, J. P. and Voss, K. J. (2000). 3D instrument self-shading effects on in-water multi-directional radiance measurements. In *Ocean Optics XV*.

D'Sa, E. J., Hu, C., Muller-Karger, F. E., and Carder, K. L. (2002). Estimation of colored dissolved organic matter and salinity fields in case 2 waters using SeaWiFS: Examples from Florida Bay and Florida Shelf. *Journal of Earth System Science*, 111(3):197–207.

Erickson, J. S., Hashemi, N., Sullivan, J. M., Weidemann, A. D., and Ligler, F. S. (2012). In situ phytoplankton analysis: there's plenty of room at the bottom. *Analytical chemistry*, 84(2):839–850.

ESA, ARGANS, and ACRI-ST (2011). MERIS Optical Measurement Protocols. Part A: In-situ water reflectance measurements. Technical Report CO-SCI-ARG-TN-008.

Evers-King, H., Bernard, S., Robertson Lain, L., and Probyn, T. A. (2014). Sensitivity in reflectance attributed to phytoplankton cell size: forward and inverse modelling approaches. *Optics Express*, 22(10):11536–11551.

Fawcett, A., Pitcher, G. C., Bernard, S., Cembella, A., and Kudela, R. M. (2007). Contrasting wind patterns and toxigenic phytoplankton in the southern Benguela upwelling system. *Marine Ecology Progress Series*, 348:19–31.

Fawcett, A., Pitcher, G. C., and Shillington, F. A. (2008). Nearshore currents on the southern Namaqua shelf of the Benguela upwelling system. *Continental Shelf Research*, 28(8):1026–1039.

Figueiras, F. G., Pitcher, G. C., and Estrada, M. (2006). Harmful Algal Bloom Dynamics in Relation to Physical Processes. In Graneli, E. and Turner, T. J., editors, *Ecology of Harmful Algae*, pages 127–138. Springer-Verlag, Berlin.

Finkel, Z., Beardall, J., Flynn, K., Quigg, A., Rees, T., and Raven, J. (2010). Phytoplankton in a changing world: cell size and elemental stoichiometry. *Journal of Plankton Research*, 32(1):119–137.

Finkel, Z. and Irwin, A. J. (2000). Modeling Size-dependent Photosynthesis: Light Absorption and the Allometric Rule. *Journal of Theoretical Biology*, 204(3):361–369.

Franz, B. and Werdell, P. (2010). A Generalized Framework for Modeling of Inherent Optical Properties in Ocean Remote Sensing Applications. In *Ocean Optics XX*.

Franz, B. A., Bailey, S. W., Werdell, P. J., and McClain, C. R. (2007). Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry. *Applied Optics*, 46(22):5068–5082.

Garcia, C., Garcia, V., and McClain, C. (2005). Evaluation of SeaWiFS chlorophyll algorithms in the Southwestern Atlantic and Southern Oceans. *Remote Sensing of Environment*, 95:125–137.

Garver, S. A. and Siegel, D. A. (1997). Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation 1. Time series from the Sargasso Sea. *Journal of Geophysical Research*, 102:18607–18625.

Gilerson, A. A., Gitelson, A. A., Zhou, J., Gurlin, D., and Moses, W. (2010). Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. *Optics Express*, 18(23):109–124.

Goes, J. I., Saino, T., Oaku, H., Ishizaka, J., Wong, C., and Nojiri, Y. (2000). Basin scale estimate of Sea Surface Nitrate and New Production from remotely sensed Sea Surface Temperature and Chlorophyll. *Geophysical Research Letters*, 27(9):1263–1266.

Gons, H. J. (2002). A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters. *Journal of Plankton Research*, 24(9):947–951.

Gons, H. J. (2005). Effect of a waveband shift on chlorophyll retrieval from MERIS imagery of inland and coastal waters. *Journal of Plankton Research*, 27(1):125–127.

Gordon, H. R. and Ding, K. (1992). Self-shading of in-water optical instruments. *Limnology and Oceanography*, 37(3):491–500.

Green, R. E., Sosik, H. M., and Olson, R. J. (2003a). Contributions of phytoplankton and other particles to inherent optical properties in New England continental shelf waters. *Limnology and Oceanography*, 48(6):2377–2391.

Green, R. E., Sosik, H. M., Olson, R. J., and DuRand, M. D. (2003b). Flow cytometric determination of size and complex refractive index for marine particles: comparison with independent and bulk estimates. *Applied Optics*, 42(3):526–541.

Gregg, W. W. and Casey, N. W. (2004). Global and regional evaluation of the SeaWiFS chlorophyll data set. *Remote Sensing of Environment*, 93:463–479.

Gregg, W. W., Ginoux, P., Schopf, P. S., and Casey, N. W. (2003). Phytoplankton and iron: validation of a global three-dimensional ocean biogeochemical model. *Deep Sea Research II*, 50:3143–3169.

Guidi, L., Stemmann, L., Jackson, G., Ibanez, F., Claustre, H., Picheral, M., and Gorsky, G. (2009). Effects of phytoplankton community on production, size, and export of large aggregates: A world-ocean analysis. *Limnology and Oceanography*, 54(6):1951–1963.

Hansen, J. E. and Travis, L. D. (1974). Light scattering in planetary atmospheres. *Space Science Review*, 16:527–610.

Havens, H., Luther, M. E., Meyers, S. D., and Heil, C. A. (2010). Lagrangian particle tracking of a toxic dinoflagellate bloom within the Tampa Bay estuary. *Marine pollution bulletin*, 60(12):2233–2241.

Havskum, H., Schlüter, L., Scharek, R., Berdalet, E., and Jacquet, S. (2004). Routine quantification of phytoplankton groups-microscopy or pigment analyses? *Marine Ecology Progress Series*, 273:31–42.

Helliwell, W. S., Sullivan, G. N., MacDonald, B., and Voss, K. J. (1990). Ship shadowing: Model and data comparisons. Ocean . In *Ocean Optics X*.

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

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

Henson, S. A. and Thomas, A. C. (2007). Interannual variability in timing of bloom initiation in the California Current System. *Journal of Geophysical Research*, 112:C08007.

Hickey, B. M. (1998). Coastal oceanography of western North America from the tip of Baja California to Vancouver Island. In Brink, K. H. and Robinson, A. R., editors, *The Sea*, pages 345–393. John Wiley and Sons Inc., New York.

Hirata, T., Aiken, J., Hardman-Mountford, N. J., Smyth, T., and Barlow, R. (2008). An absorption model to determine phytoplankton size classes from satellite ocean colour. *Remote Sensing of Environment*, 112:3153–3159.

Hlaing, S., Gilerson, A., Harmel, T., Tonizzo, A., Weidemann, A. D., Arnone, R., and Ahmed, S. (2012). Assessment of a bidirectional reflectance distribution correction of above-water and satellite water-leaving radiance in coastal waters. *Applied Optics*, 51(2):220–237.

Hochberg, E., Atkinson, M. J., and Andréfouët, S. (2003). Spectral reflectance of coral reef bottom-types worldwide and implications for coral reef remote sensing. *Remote Sensing of Environment*, 85:159–173.

Hoepffner, N. and Sathyendranath, S. (1993). Determination of the major groups of phytoplankton pigments from the absorption spectra of total particulate matter. *Journal of Geophysical Research*, 98(C12):22789–22803.

Holligan, P. M., Viollier, M., Harbour, D. S., Camus, P., and Champagne-Philippe, M. (1983). Satellite and ship studies of coccolithophore production along a continental shelf edge. *Nature*, 304:339–342.

Hooker, S. B., Lazin, G., and Zibordi, G. (2002). An evaluation of above-and in-water methods for determining water-leaving radiances. *Journal of Atmospheric and Oceanic Technology*, 19:486–515.

Hooker, S. B. and Maritorena, S. (2000). An evaluation of oceanographic radiometers and deployment methodologies. *Journal of Atmospheric and Oceanic Technology*, 17:811–830.

Hooker, S. B., McClain, C. R., Firestone, J. K., Westphal, T. L., Yeh, E.-N., and Ge, Y. (1994). The SeaWiFS Bio-Optical Archive and Storage System (SeaBASS). Part 1. Technical Report 104566, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

Hooker, S. B. and Morel, A. (2003). Platform and Environmental Effects on Above-Water Determinations of Water-Leaving Radiances. *Journal of Atmospheric and Oceanic Technology*, 20:187–205.

Hooker, S. B., Zibordi, G., Berthon, J.-F., D'Alimonte, D., Maritorena, S., McLean, S. D., and Sildam, J. (2001). Results of the Second SeaWiFS Data Analysis Round Robin. Technical report, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

Hu, C., Muller-Karger, F., Taylor, C., and Carder, K. L. (2005). Red tide detection and tracing using MODIS fluorescence data: A regional example in SW Florida coastal waters. *Remote Sensing of Environment*, 97:311–321.

Hu, Y. X. and Stamnes, K. (1993). An Accurate Parameterization of the Radiative Properties of Water Clouds Suitable for Use in Climate Models. *Journal of Climate*, 6(4):728–742.

Huang, S., Li, Y., Shang, S., and Shang, S. (2013). Impact of computational methods and spectral models on the retrieval of optical properties via spectral optimization. *Optics Express*, 21(5):6257–6273.

Huot, Y., Morel, A., Twardowski, M. S., Stramski, D., and Reynolds, R. A. (2008). Particle optical backscattering along a chlorophyll gradient in the upper layer of the eastern South Pacific Ocean. *Biogeosciences*, 5:495–507.

Hutchings, L., Van der Lingen, C., Shannon, L. V., Crawford, R., Verheye, H., Bartholomae, C., Van der Plas, A., Louw, D., Kreiner, A., Ostrowski, M., Fidel, Q., Barlow, R., Lamont, T., Coetzee, J., Shillington, F. A., Veitch, J., Currie, J., and Monteiro, P. M. S. (2009). The Benguela Current: An ecosystem of four components. *Progress in Oceanography*, 83:15–32.

IOCCG (2000). Remote Sensing of Ocean Colour in Coastal and Other Optically-Complex, Waters. Technical Report 3, Dartmouth, Canada.

IOCCG (2006). Remote Sensing of Inherent Optical Properties: Fundamental Tests of Algorithms and Applications. Technical Report 5, Dartmouth, Canada.

IOCCG (2007). Ocean-Color Data Merging. Technical Report 6, Dartmouth, Canada.

IOCCG (2009). Remote Sensing in Fisheries and Aquaculture. Technical Report 8, Dartmouth, Canada.

IOCCG (2010). Atmospheric Correction for Remotely-Sensed Ocean- Colour Products. Technical Report 10, Dartmouth, Canada.

IOCCG (2014). Phytoplankton Functional Types from Space . Technical Report 15, Dartmouth, Canada.

Irwin, A. J., Finkel, Z., Schofield, O. M. E., and Falkowski, P. G. (2006). Scaling-up from nutrient physiology to the size-structure of phytoplankton communities. *Journal of Plankton Research*, 28(5):459–471.

Jackson, G. A., Maffione, R., Costello, D. K., Alldredge, A. L., Logan, B. E., and Dam, H. G. (1997). Particle size spectra between 1 μm and 1 cm at Monterey Bay determined using multiple instruments. *Deep Sea Research I*, 44(11):1739–1767.

Jakobsen, H. H. and Carstensen, J. (2011). FlowCAM: Sizing cells and understanding the impact of size distributions on biovolume of -planktonic community structure. *Aquatic Microbial Ecology*, 65:75–87.

Jeong, H. J., Yoo, Y. D., Park, J. Y., Song, J. Y., Kim, S. T., Lee, S. H., Kim, K. Y., and Yih, W. H. (2005). Feeding by phototrophic red-tide dinoflagellates: five species newly revealed and six species previously known to be mixotrophic. *Aquatic Microbial Ecology*, 40:133–150.

Joint, I. and Groom, S. (2000). Estimation of phytoplankton production from space: current status and future potential of satellite remote sensing. *Journal of Experimental Marine Biology and Ecology*, 250:233–255.

Jonasz, M. and Fournier, G. (1996). Approximation of the size distribution of marine particles by a sum of log-normal functions. *Limnology and Oceanography*, 41(4):744–754.

Jonker, R., Meulemans, J., Dubelaar, G., Wilkins, M., and Ringelberg, J. (1995). Flow cytometry: A powerful tool in analysis of biomass distributions in phytoplankton. *Water Science and Technology*, 32(4):177–182.

Junge, C. E. (1963). *Air Chemistry and Radioactivity.* Academic Press, New York.

Kahru, M. and Mitchell, B. G. (1998). Evaluation of instrument self-shading and environmental errors on ocean color algorithms. In *Ocean Optics XIV*.

Kahru, M. and Mitchell, B. G. (2001). Seasonal and nonseasonal variability of satellite-derived chlorophyll and colored dissolved organic matter concentration in the California Current. *Journal of Geophysical Research*, 106:2517–2529.

Kamykowski, D. and Zentara, S.-J. (2003). Can phytoplankton community structure be inferred from satellite-derived sea surface temperature anomalies calculated relative to nitrate depletion temperatures? *Remote Sensing of Environment*, 86:444–457.

Karp-Boss, L., Azevedo, L., and Boss, E. (2007). LISST-100 measurements of phytoplankton size distribution: evaluation of the effects of cell shape. *Limnology and Oceanography: Methods*, 5:396–406.

Kirk, J. T. O. (1975). A Theoretical Analysis of the Contribution of Algal Cells to the Attenuation of Light Within Natural Waters. I. General Treatment of Suspension of Pigmented Cells. *New Phytologist*, 75(1):11–20.

Kishino, M., Takahashi, M., Okami, N., and Ichimura, S. (1985). Estimation of the spectral absorption coefficients of phytoplankton in the sea. *Bulletin of Marine Science*, 37:634–642.

Kitchen, J. C. and Zaneveld, J. R. (1992). A three-layered sphere model of the optical properties of phytoplankton. *Limnology and Oceanography*, 37(8):1680–1690.

Koike, I. and Fukuda, H. (2007). Mini-review: Assessment of size distribution of suspended particles in coastal and estuarine environments usin in situ instruments. *Coastal Marine Science*, 31(1):1–8.

Kone, V., Machu, E., Penven, P., Andersen, V., Garcon, V., Freon, P., and Demarcq, H. (2005). Modeling the primary and secondary productions of the southern Benguela upwelling system: A comparative study through two biogeochemical models. *Global Biogeochemical Cycles*, 19(GB4021).

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *Journal of Geophysical Research*, 114:C09015.

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2010). Global variability of phytoplankton functional types from space: assessment via the particle size distribution. *Biogeosciences Discussions*, 7:4295–4340.

Kudela, R., Pitcher, G. C., Probyn, T., Figueiras, F., Moita, T., and Trainer, V. (2005). Harmful algal blooms in coastal upwelling systems. *Oceanography*, 18(2):184–197.

Kyewalyanga, M., Sathyendranath, S., and Platt, T. (2002). Effect of Mesodinium rubrum (= Myrionecta rubra) on the action and absorption spectra of phytoplankton in a coastal marine inlet. *Journal of Plankton Research*, 24(7):687–702.

Lacroix, G., Ruddick, K., Park, Y., Gypens, N., and Lancelot, C. (2007). Validation of the 3D biogeochemical model MIRO&CO with field nutrient and phytoplankton data and MERIS-derived surface chlorophyll a images. *Journal of Marine Systems*, 64:66–88.

Leathers, R., Downes, T., and Mobley, C. (2004). Self-shading correction for oceanographic upwelling radiometers. *Optics Express*, 12(20):4709–4718.

Leathers, R., Downes, T. V., and Mobley, C. (2001). Self-shading correction for upwelling sea-surface radiance measurements made with buoyed instruments. *Optics Express*, 8(10):561–570.

Lee, Z., Carder, K. L., and Arnone, R. (2002). Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, 41(27):5755–5772.

Lee, Z., Carder, K. L., Chen, R. F., and Peacock, T. G. (2001). Properties of the water column and bottom derived from Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data. *Journal of Geophysical Research*, 106:11639–11651.

Lee, Z., Pahlevan, N., Ahn, Y.-H., Greb, S., and O'Donnell, D. (2013). Robust approach to directly measuring water-leaving radiance in the field. *Applied Optics*, 52(8):1693.

Levenberg, K. (1944). A Method for the Solution of Certain Non-Linear Problems in Least Squares . *Quarterly of Applied Mathematics*, pages 164–168.

Li, X. and Logan, B. E. (1995). Size distributions and fractal properties of particles during a simulated phytoplankton bloom in a mesocosm. *Deep Sea Research II*, 42(1):125–138.

Litchman, E. and Klausmeier, C. A. (2008). Trait-Based Community Ecology of Phytoplankton. *Annual Review of Ecology, Evolution, and Systematics*, 39(1):615–639.

Loisel, H., Nicolas, J., Sciandra, A., Stramski, D., and Poteau, A. (2006). Spectral dependency of optical backscattering by marine particles from satellite remote sensing of the global ocean. *Journal of Geophysical Research*, 111:C99024.

Louchard, E., Reid, R., Stephens, C., Davis, C., Leathers, R., Downes, T., and Maffione, R. (2002). Derivative analysis of absorption features in hyperspectral remote sensing data of carbonate sediments. *Optics Express*, 10(26):1573–1584.

Louchard, E. M., Reid, R. P., Stephens, F. C., Davis, C. O., Leathers, R. A., and Downes, T. V. (2003). Optical remote sensing of benthic habitats and bathymetry in coastal environments at Lee Stocking Island, Bahamas: A comparative spectral classification approach. *Limnology and Oceanography*, 48(1, part 2):511–521.

Loyola, R. and Diego, G. (2006). Applications of neural network methods to the processing of earth observation satellite data. *Neural Networks*, 19(2):168–177.

Lucas, A. J., Pitcher, G. C., Probyn, T. A., and Kudela, R. M. (2014). The influence of diurnal winds on phytoplankton dynamics in a coastal upwelling system off southwestern Africa. *Deep-Sea Res Part II: Topical Studies in Oceanography*, 101:50–62.

Lund, J. W. G., Kipling, C., and Cren, E. D. (1958). The inverted microscope method of estimating algal numbers and the statistical basis of estimations by counting. *Hydrobiologia*, 11(2):143–170.

Machu, E., Ferret, B., and Garcon, V. (1999). Phytoplankton pigment distribution from SeaWiFS data in the subtropical convergence zone south of Africa: A wavelet analysis. *Geophysical Research Letters*, 26(10):1469–1472.

Marañón, E. (2008). Inter-specific scaling of phytoplankton production and cell size in the field. *Journal of Plankton Research*, 30(2):157–163.

Marañón, E., Cermeno, P., Rodriguez, J., Zubkov, M. V., and Harris, R. P. (2007). Scaling of phytoplankton photosynthesis and cell size in the ocean. *Limnology and Oceanography*, 52(5):2190–2198.

Margalef, R. (1961). Communication of structure in planktonic populations. *Limnology and Oceanography*, 6(2):124–128.

Margalef, R. (1978). Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanologica Acta*, 134:493–509.

Margalef, R., Estrada, M., and Blasco, D. (1979). Functional morphology of organisms involved in red tides, as adapted to decaying turbulence . In Taylor, D. and Seliger, H., editors, *Toxic Dinoflagellate Blooms*, pages 89–94. Elsevier, New York.

Maritorena, S., Lee, Z. P., Du, K., Loisel, H., Doerffer, R., Roesler, C., Lyon, P., Tanaka, A., Babin, M., and Kopelevich, O. V. (2006). Synthetic and In Situ Data Sets for Algorithm Testing . In Lee, Z. P., editor, *Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms and Applications*. IOCCG, Dartmouth, Canada.

Marquardt, D. W. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441.

Matthews, M. and Bernard, S. (2013). Using a two-layered sphere model to investigate the impact of gas vacuoles on the inherent optical properties of M. aeruginosa. *Biogeosciences*, 10:10531–10579.

Matthews, M., Bernard, S., and Robertson, L. (2012). An algorithm for detecting trophic status (chlorophyll- a), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. *Remote Sensing of Environment*, 124:637–652.

Matthews, M. W., Bernard, S., and Winter, K. (2010). Remote sensing of cyanobacteria-dominant algal blooms and water quality parameters in Zeekoevlei, a small hypertrophic lake, using MERIS. *Remote Sensing of Environment*, 114:2070–2087.

McClain, C. (2009). A Decade of Satellite Ocean Color Observations. *Annual Review of Marine Science*, 1:19–42.

McGraw, R., Nemesure, S., and Schwartz, S. E. (1998). Properties and evolution of aerosols with size distributions having identical moments. *Journal of Aerosol Science*, 29(7):761–722.

Mélin, F., Holben, B. N., and Courcoux, Y. (2013). Validation of aerosol products derived from ocean colour in East African coastal waters. *African Journal of Marine Science*, 35(3):351–356.

Mélin, F. and Zibordi, G. (2010). Vicarious calibration of satellite ocean color sensors at two coastal sites. *Applied Optics*, 49(5):798–810.

Mélin, F., Zibordi, G., and Berthon, J. F. (2007). Assessment of satellite ocean color products at a coastal site. *Remote Sensing of Environment*, 110:192–215.

Millie, D. F. and Moline, M. A. (2000). Optical discrimination of a phytoplankton species in natural mixed populations. *Limnology and Oceanography*, 45(2):467–471.

Mobley, C. D. (1994). *Light and Water*. Radiative Transfer in Natural Waters. Academic Press.

Mobley, C. D. (2011). Fast light calculations for ocean ecosystem and inverse models. *Optics Express*, 19(20):18927–18944.

Mobley, C. D. and Sundman, L. K. (2008). HydroLight 5.0, Technical Documentation. Technical report.

Mobley, C. D., Sundman, L. K., Bissett, W. P., and Cahill, B. (2009). Fast and accurate irradiance calculations for ecosystem models. *Biogeosciences Discussions*, 6:10625–10662.

Mobley, C. D. C., Sundman, L. K. L., and Boss, E. E. (2002). Phase function effects on oceanic light fields. *Applied Optics*, 41(6):1035–1050.

Montagnes, D., Berges, J., Harrison, P., and Taylor, F. (1999). Estimating carbon, nitrogen, protein, and chlorophyll a from volume in marine phytoplankton. *Limnology and Oceanography*, 39(5):1044–1060.

Moore, G., Aiken, J., and Lavender, S. (1999). The atmospheric correction of water colour and the quantitative retrieval of suspended particulate matter in Case II waters: application to MERIS. *International Journal of Remote Sensing*, 20(9):1713–1733.

Moore, T. S., Dowell, M. D., and Franz, B. A. (2012). Detection of coccolithophore blooms in ocean color satellite imagery: A generalized approach for use with multiple sensors. *Remote Sensing of Environment*, 117:249–263.

Morel, A. (1991). Optics of marine particles and marine optics. *Optical Aspects of Oceanography, NATO ASI Series*, G27:141–188.

Morel, A. (1997). Consequences of a Synechococcus bloom upon the optical properties of oceanic (case 1) waters. *Limnology and Oceanography*, 42(8):1746–1754.

Morel, A. (2009). Are the empirical relationships describing the bio-optical properties of case 1 waters consistent and internally compatible. *Journal of Geophysical Research*, 114:C01016.

Morel, A. and Antoine, D. (2000). Pigment Index Retrieval in Case 1 Waters. MERIS ATBD 2.9. Technical report, France.

Morel, A., Antoine, D., and Gentili, B. (2002). Bidirectional Reflectance of Oceanic Waters: Accounting for Raman Emission and Varying Particle Scattering Phase Function. *Applied Optics*, 41(30):6289–6306.

Morel, A. and Bricaud, A. (1981). Theoretical results concerning the optics of phytoplankton, with special reference to remote sensing applications. In *Oceanography from Space*, pages 313–327. Plenum Press.

Morel, A. and Bricaud, A. (1986). Inherent optical properties of algal cells including picoplankton: Theoretical and experimental results. In Platt, T. and Li, W., editors, *Photosynthetic picoplankton*, pages 521–559. Canadian Bulletin Fishery and Aquatic Science.

Morel, A. and Maritorena, S. (2001). Bio-optical properties of oceanic waters: A reappraisal. *Journal of Geophysical Research*, 106(C4):7163–7180.

Morel, A. and Prieur, L. (1977). Analysis of variations in ocean color. *Limnology and Oceanography*, 22(4):709–722.

Moses, W. J., Gitelson, A. A., Berdnikov, S., Saprygin, V., and Povazhnyi, V. (2012). Operational MERIS-based NIR-red algorithms for estimating chlorophyll-a concentrations in coastal waters — The Azov Sea case study. *Remote Sensing of Environment*, 121:118–124.

Mouw, C. B. and Yoder, J. A. (2005). Primary production calculations in the Mid-Atlantic Bight, including effects of phytoplankton community size structure. *Limnology and Oceanography*, 50(4):1232–1243.

Mueller, J., Bidigare, R. R., Trees, C., Balch, W. M., Dore, J., Drapeau, D. T., Karl, D. M., Van Heukelem, L., and Perl, J. (2003a). Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 5, Volume V: Biogeochemical and Bio-Optical Measurements

and Data Analysis Protocols. Technical report, NASA Goddard Space Flight Space Centre, Greenbelt, Maryland 20771.

Mueller, J. L., Morel, A., Frouin, R., Davis, C., Arnone, R., Carder, K. L., Lee, Z. P., Steward, R. G., Hooker, S. B., Mobley, C. D., Mclean, S., Holben, B., Miller, M., Pietras, C., Knobspiesse, K. D., Fargion, G. S., Porter, J., and Voss, K. (2003b). Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 4, Volume III: Radiometric Measurements and Data Analysis Protocols. Technical report, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

Mulligan, H. F. and Kingsbury, J. M. (1968). Application of an electronic particle counter in analyzing natural populations of phytoplankton. *Limnology and Oceanography*, (13):499–506.

Munk, W. H. and Riley, G. A. (1952). Absorption of nutrients by aquatic plants. *Journal of Marine Research*, 11:215–240.

Nair, A., Sathyendranath, S., Platt, T., Morales, J., Stuart, V., Forget, M.-H., Devred, E., and Bouman, H. (2008). Remote sensing of phytoplankton functional types. *Remote Sensing of Environment*, 112:3366–3375.

Napp, J. M., Brooks, E. R., Reid, F. M. H., Matrai, P., and M, M. M. (1988). Vertical distribution of marine particles and grazers. I. Vertical distribution of food quality and quantity . *Marine Ecology Progress Series*, 50:45–58.

Nelder, J. A. and Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313.

Neukermans, G., Loisel, H., Mériaux, X., Astoreca, R., and McKee, D. (2012). In situ variability of mass-specific beam attenuation and backscattering of marine particles with respect to particle size, density, and composition. *Limnology and Oceanography*, 57(1):124–144.

Olivieri, E. T. (1985). Feasibility of estimating phytoplankton size and biomass in fresh and preserved samples from the Benguela Current with a Coulter counter. *South African Journal of Marine Science*, 3(1):99–110.

Olson, R. J., Shalapyonok, A., and Sosik, H. M. (2003). An automated submersible flow cytometer for analyzing pico- and nanophytoplankton: FlowCytobot. *Deep Sea Research Part 1*, 50(2):301–315.

Olson, R. J. and Sosik, H. M. (2007). A submersible imaging-in-flow instrument to analyze nano- and microplankton: Imaging FlowCytobot. *Limnology and Oceanography: Methods*, 5:195–203.

O'Reilly, J., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., Kahru, M., and McCain, C. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research*, 103:24937–24953.

Park, Y.-J. and Ruddick, K. (2005). Model of remote-sensing reflectance including bidirectional effects for case 1 and case 2 waters. *Applied Optics*, 44(7):1236–1249.

Parsons, T. R. and Takahashi, M. (1973). Environmental control of phytoplankton cell size. *Limnology and Oceanography*, 18(4):511–515.

Penven, P., Roy, C., Colin de Verdière, A., and Largier, J. (2000). Simulation of a Coastal Jet Retention Process Using a Barotropic Model. *Oceanologica Acta*, 23(5):615–634.

Pinkerton, M. and Aiken, J. (1999). Calibration and validation of remotely sensed observations of ocean color from a moored data buoy. *Journal of Atmospheric and Oceanic Technology*, 16:915–923.

Piskozub, J. (2004). Effect of ship shadow on in-water irradiance measurements. *Oceanologia*, 46:103–112.

Piskozub, J., Weeks, A. R., Schwarz, J. N., and Robinson, I. S. (2000). Self-shading of upwelling irradiance for an instrument with sensors on a sidearm. *Applied Optics*, 39(12):1872–1878.

Pitcher, G. C., Bernard, S., and Fawcett, A. (2008a). Real-time coastal observing systems for ecosystem dynamics and harmful algal blooms: the needs and expectations of the user. In *Real-time Coastal Observing Systems for Marine Ecosystem Dynamics and Harmful Algal Blooms: Theory, Instrumentation and Modelling*, pages 765–798. UNESCO, Paris, France.

Pitcher, G. C., Bernard, S., and Ntu, J. (2008b). Contrasting Bays and Red Tides in the Southern Benguela Upwelling System. *Oceanography*, 21(3):82–91.

Pitcher, G. C., Boyd, A., Horstman, D. A., and Mitchell-Innes, B. A. (1998). Subsurface dinoflagellate populations, frontal blooms and the formation of red tide in the southern Benguela upwelling system. *Marine Ecology Progress Series*, 172:253–264.

Pitcher, G. C., Brown, P., and Mitchell-Innes, B. A. (1992a). Spatio-temporal variability in phytoplankton in the southern Benguela upwelling system. *South African Journal of Marine Science*, 12(1):439–456.

Pitcher, G. C., Brown, P. C., and Mitchell-Innes, B. A. (1992b). Spatio-temporal variability of phytoplankton in the southern Benguela upwelling system. *South African Journal of Marine Science*, 12(1):439–456.

Pitcher, G. C. and Calder, D. (2000). Harmful algal blooms of the southern Benguela Current: a review and appraisal of monitoring from 1989 to 1997. *South African Journal of Marine Science*, 22(1):255–271.

Pitcher, G. C., Figueiras, F. G., Hickey, B. M., and Moita, M. T. (2010). The physical oceanography of upwelling systems and the development of harmful algal blooms. *Progress in Oceanography*, 85(1-2):5–32.

Pitcher, G. C., Krock, B., and Cembella, A. D. (2011). Accumulation of diarrhetic shellfish poisoning toxins in the oyster Crassostrea gigasand the mussel Choromytilus meridionalisin the southern Benguela ecosystem. *African Journal of Marine Science*, 33(2):273–281.

Pitcher, G. C., Monteiro, P., and Kemp, A. (2002). The potential use of a hydrodynamic model in the prediction of harmful algal blooms in the southern Benguela. In Steidinger, K. A., Lansberg, J. H., Tomas, C. R., and Vargo, G. A., editors, *Harmful algae 2002*, pages 11–13. Florida Fish and Wildlife Conservation Commission, Florida Institute of Oceanography, and Intergovernmental Oceanographic Commission of UNESCO.

Pitcher, G. C. and Nelson, G. (2006). Characteristics of the surface boundary layer important to the development of red tide on the southern Namaqua shelf of the Benguela upwelling system. *Limnology and Oceanography*, 51(6):2660–2674.

Pitcher, G. C. and Probyn, T. A. (2011). Anoxia in southern Benguela during the autumn of 2009 and its linkage to a bloom of the dinoflagellate Ceratium balechii. *Harmful Algae*, 11:23–32.

Pitcher, G. C., Probyn, T. A., du Randt, A., Lucas, A. J., Bernard, S., Evers-King, H., Lamont, T., and Hutchings, L. (2014). Dynamics of oxygen depletion in the nearshore of a coastal embayment of the southern Benguela upwelling system. *Journal of Geophysical Research*, 119:2183–2200.

Pitcher, G. C., Walker, D. R., Mitchell-Innes, B. A., and Moloney, C. L. (1991). Short-term variability during an anchor station study in the southern Benguela upwelling system: Phytoplankton dynamics. *Progress in Oceanography*, 28(1-2):39–64.

Pitcher, G. C. and Weeks, S. J. (2006). The variability and potential for prediction of harmful algal blooms in the southern Benguela ecosystem. . In Shannon, V., Hempel, G., Malanotte-Rizzoli, P., Moloney, C., and Woods, J., editors, *The Benguela: Predicting a Large Marine Ecosystem*. Elsevier.

Platt, T., White, III, G. N., Zhai, L., Sathyendranath, S., and Roy, S. (2009). The phenology of phytoplankton blooms: Ecosystem indicators from remote sensing. *Ecological Modelling*, 220(21):3057–3069.

Prieur, L. and Sathyendranath, S. (1981). An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments, dissolved organic matter, and other particulate materials. *Limnology and Oceanography*, 26(4):671–689.

Probyn, T. A. (1985). Nitrogen uptake by size-fractionated phytoplankton populations in the southern Benguela upwelling system. *Marine Ecology Progress Series*, 22:249–258.

Probyn, T. A., Pitcher, G. C., Monteiro, P. M. S., Boyd, A. J., and Nelson, G. (2000). Physical processes contributing to harmful algal blooms in Saldanha Bay, South Africa. *South African Journal of Marine Science*, 22(1):285–297.

Quinby-Hunt, M. S., J, H. A., Lofftus, K., and Shapiro, D. (1989). Polarized-light scattering studies of marine Chlorella. *Limnology and Oceanography*, 34(8):1587–1600.

Quirantes, A. and Bernard, S. (2004). Light scattering by marine algae: two-layer spherical and nonspherical models. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 89(1-4):311–321.

Racault, M. F., Le Quéré, C., and Buitenhuis, E. (2012). Phytoplankton phenology in the global ocean. *Ecological Indicators*, 14:152–163.

Raitsos, D., Lavender, S. J., and Maravelias, C. (2008). Identifying four phytoplankton functional types from space: An ecological approach. *Limnology and Oceanography*, 53(2):605–613.

Rajpaul, V. (2012). Evolutionary-algorithm-based analysis of gravitational microlensing light curves. *Monthly Notices of the Royal Astronomical Society*, 427:1755–1768.

Raven, J. A. (1984). A Cost-Benefit Analysis of Photon Absorption by Photosynthetic Unicells. *New Phytologist*, 98(4):593–625.

Raven, J. A. (1986). Physiological consequences of extremely small size for autotrophic organisms in the sea. In Platt, T. and Li, W. K. W., editors, *Photosynthetic Picoplankton*. Canadian Bulletin Fishery and Aquatic Science.

Raven, J. A. (2009). The Bigger The Fewer: Size, Taxonomic Diversity and The Range of Chlorophyll(Ide) Pigments in Oxygen-Evolving Marine Photolithotrophs. *Journal of the Marine Biological Association of the United Kingdom*, 76(01):211–217.

Raven, J. A. and Kübler, J. E. (2002). New light on the scaling of metabolic rate with the size of algae. *Journal of Phycology*, 38(1):11–16.

Rehm, E. and Mobley, C. D. (2013). Estimation of hyperspectral inherent optical properties from in-water radiometry: error analysis and application to in situ data. *Applied Optics*, 52(4):795–817.

Reynolds, R. A., Stramski, D., and Mitchell, B. G. (2001). A chlorophyll-dependent semianalytical reflectance model derived from field measurements of absorption and backscattering coefficients within the Southern Ocean. *Journal of Geophysical Research*, 106(C4):7125–7138.

Reynolds, R. A., Stramski, D., Wright, V. M., and Wozniak, S. B. (2010). Measurements and characterization of particle size distributions in coastal waters. *Journal of Geophysical Research*, 115:C08024.

Richardson, T. and Jackson, G. (2007). Small phytoplankton and carbon export from the surface ocean. *Science*, 315:838–840.

Risović, D. (1993). Two-component model of sea particle size distribution. *Deep Sea Research Part I: Oceanographic Research Papers*, 40(7):1459–1473.

Robertson Lain, L., Bernard, S., and Evers-King, H. (2014). Biophysical modelling of phytoplankton communities from first principles using two-layered spheres: Equivalent Algal Populations (EAP) model. *Optics Express*, 22(14):16745–16758.

Robinson, I. S. (2010). *Discovering the Ocean from Space.* The unique applications of satellite oceanography. Springer Science & Business Media.

Roelke, D. L., Kennedy, C. D., and Weidemann, A. D. (1999). Use of discriminant and fourth-derivative analyses with high-resolution absorption spectra for phytoplankton research: Limitations at varied signal-to-noise ratio and spectral resolition. *Gulf of Mexico Science*, 2:75–86.

Roesler, C. S. (1998). Theoretical and experimental approaches to improve the accuracy of particulate absorption coefficients derived from the quantitative filter technique. *Limnology and Oceanography*, 43(7):1649–1660.

Roesler, C. S. and Barnard, A. H. (2013). Optical proxy for phytoplankton biomass in the absence of photophysiology: Rethinking the absorption line height. *Methods in Oceanography*, 7:79–94.

Roesler, C. S. and Boss, E. (2008). In situ measurement of the inherent optical properties (IOPs) and potential for harmful algal bloom detection and coastal ecosystem observations. In *Real-time Coastal Observing Systems for Marine Ecosystem Dynamics and Harmful Algal Blooms: Theory, Instrumentation and Modelling*, pages 153–206. UNESCO, Paris, France.

Roesler, C. S. and Perry, M. J. (1995). In situ phytoplankton absorption, fluorescence emission, and particulate backscattering spectra determined from reflectance. *Journal of Geophysical Research*, 100(C7):13279–13294.

Roy, S., Sathyendranath, S., and Platt, T. (2011). Retrieval of phytoplankton size from bio-optical measurements: theory and applications. *Journal of The Royal Society Interface*, 8(58):650–660.

Ruddick, K. G., Gons, H. J., Rijkeboer, M., and Tilstone, G. (2001). Optical Remote Sensing of Chlorophyll a in Case 2 Waters by Use of an Adaptive Two-Band Algorithm with Optimal Error Properties. *Applied Optics*, 40(21):3575–3585.

Rudorff, N. D. M. and Kampel, M. (2012). Orbital remote sensing of phytoplankton functional types: a new review. *International Journal of Remote Sensing*, 33(6):1967–1990.

Ryan, J. P., Fischer, A. M., Kudela, R. M., Gower, J. F. R., King, S. A., Marin, III, R., and Chavez, F. P. (2009). Influences of upwelling and downwelling winds on red tide bloom dynamics in Monterey Bay, California. *Continental Shelf Research*, 29(5-6):785–795.

Ryan, J. P., Gower, J. F. R., King, S. A., Bissett, W. P., Fischer, A. M., Kudela, R. M., Kolber, Z., Mazzillo, F., Rienecker, E. V., and Chavez, F. P. (2008). A coastal ocean extreme bloom incubator. *Geophysical Research Letters*, 35(12):L12602.

Sathyendranath, S., Cota, G., Stuart, V., Maass, H., and Platt, T. (2001). Remote sensing of phytoplankton pigments: a comparison on empirical and theoretical approaches. *International Journal of Remote Sensing*, 22(2&3):249–273.

Sathyendranath, S., Lazzara, L., and Prieur, L. (1987). Variations in the spectral values of specific absorption of phytoplankton. *Limnology and Oceanography*, 32(2):403–415.

Sathyendranath, S., Watts, L., Devred, E., Platt, T., Caverhill, C., and Maass, H. (2004). Discrimination of diatoms from other phytoplankton using ocean-colour data. *Marine Ecology Progress Series*, 272:59–68.

Sauer, M. J., Roesler, C. S., Werdell, P. J., and Barnard, A. (2012). Under the hood of satellite empirical chlorophyll a algorithms: revealing the dependencies of maximum band ratio algorithms on inherent optical properties. *Optics Express*, 20(19):20920–20933.

Schallenberg, C., Lewis, M. R., Kelley, D., and Cullen, J. J. (2008). Inferred influence of nutrient availability on the relationship between Sun-induced chlorophyll fluorescence and incident irradiance in the Bering Sea. *Journal of Geophysical Research*, 113:C07046.

Schiller, H. and Doerffer, R. (2005). Improved determination of coastal water constituent concentrations from MERIS data. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(7):1585–1591.

Schlüter, L. and Møhlenberg, F. (2003). Detecting presence of phytoplankton groups with non-specific pigment signatures - Springer. *Journal of applied phycology.*

Schofield, O., Grzymski, J., Bissett, W. P., Kirkpatrick, G. J., Millie, D. F., Moline, M., and Roesler, C. S. (2002). Optical Monitoring and Forecasting Systems for Harmful Algal Blooms: Possibility or Pipe Dream? *Journal of Phycology*, 35(6):1477–1496.

Semina, H. J. (1972). The Size of Phytoplankton Cells in the Pacific Ocean. *Internationale Revue der gesamten Hydrobiologie und Hydrographie*, 57(2):177–205.

Shannon, L. V. and Pillar, S. C. (1986). The Benguela Ecosystem part 3. Plankton. *Oceanography and Marine Biology: An annual review*, 24:65–170.

Sheldon, R. W. and Parsons, T. R. (1967). A Continuous Size Spectrum for Particulate Matter in the Sea. *Journal of the Fisheries Research Board of Canada*, 24:909–915.

Sheldon, R. W., Prakash, A., and Sutcliffe, W. H. (1972). The size distribution of particles in the ocean. *Limnology and Oceanography*, 17(3):327–340.

Siddorn, J. R., Allen, J. I., Blackford, J. C., Gilbert, F. J., Holt, J. T., Holt, M. W., Osborne, J. P., Proctor, R., and Mills, D. K. (2007). Modelling the hydrodynamics and ecosystem of the North-West European continental shelf for operational oceanography. *Journal of Marine Systems*, 65(1-4):417–429.

Sieburth, J., Smetacek, V., and Lenz, J. (1978). Pelagic ecosystem structure: heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnology and Oceanography*, 23(6):1256–1263.

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

Siegel, D. A., Maritorena, S., and Nelson, N. B. (2005). Colored dissolved organic matter and its influence on the satellite-based characterization of the ocean biosphere. *Geophysical Research Letters*, 32:L20605.

Siegel, D. A., OBrien, M. C., Sorensen, J. C., Konnoff, D. A., and Brody, E. A. (1994). SeaWiFS technical Report Series. Volume 26, Results of the SeaWiFS Data Analysis Round Robin, July 1994 (DARR-94). Technical report, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

Silió-Calzada, A., Bricaud, A., Uitz, J., and Gentili, B. (2008). Estimation of new primary production in the Benguela upwelling area, using ENVISAT satellite data and a model dependent on the phytoplankton community size structure. *Journal of Geophysical Research*, 113:C11023.

Smalley, G. W., Coats, D. W., and Stoecker, D. K. (2003). Feeding in the mixotrophic dinoflagellate Ceratium furca is influenced by intracellular nutrient concentrations. *Marine Ecology Progress Series*, 262:137–151.

Smayda, T. and Reynolds, C. (2001a). Community assembly in marine phytoplankton: application of recent models to harmful dinoflagellate blooms. *Journal of Plankton Research*, 23(5):447.

Smayda, T. J. (1997a). Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea. *Limnology and Oceanography*, 42(5, part 2):1137–1153.

Smayda, T. J. (1997b). What is a bloom? A commentary. *Limnology and Oceanography*, 42(5, part 2):1132–1136.

Smayda, T. J. (2000). Ecological features of harmful algal blooms in coastal upwelling ecosystems. *South African Journal of Marine Science*, 22(1):219–253.

Smayda, T. J. (2008). Complexity in the eutrophication–harmful algal bloom relationship, with comment on the importance of grazing. *Harmful Algae*, 8:140–151.

Smayda, T. J. and Reynolds, C. S. (2001b). Community assembly in marine phytoplankton: application of recent models to harmful dinoflagellate blooms. *Journal of Plankton Research*, 23(5):447–461.

Smayda, T. J. and Reynolds, C. S. (2003). Strategies of marine dinoflagellate survival and some rules of assembly. *Journal of Sea Research*, 49(2):95–106.

Smith, M. E., Bernard, S., and O'Donoghue, S. (2013). The assessment of optimal MERIS ocean colour products in the shelf waters of the KwaZulu-Natal Bight, South Africa. *Remote Sensing of Environment*, 137:124–138.

Smith, R. C. and Baker, K. S. (1978). Optical classification of natural waters. *Limnology and Oceanography*, 23(2):260–267.

Smyth, T. J., Moore, G. F., Hirata, T., and Aiken, J. (2006). Semianalytical model for the derivation of ocean color inherent optical properties: description, implementation, and performance assessment. *Applied Optics*, 45(31):8116–8131.

Sosik, H. M. and Mitchell, B. G. (1991). Absorption, fluorescence, and quantum yield for growth in nitrogen-limited Dunaliella tertiolecta. *Limnology and Oceanography*, 36(5):910–921.

Sosik, H. M., Sathyendranath, S., Uitz, J., Bouman, H., and Nair, A. (2014). *In situ* Methods of Measuring Phytoplankton Functional Types . In Sathyendranath, S., editor, *Phytoplankton Functional Types from Space*, pages 21–38. IOCCG, Dartmouth, Canada.

Stemmann, L. and Boss, E. (2012). Plankton and Particle Size and Packaging: From Determining Optical Properties to Driving the Biological Pump. *Annual Review of Marine Science*, 4:263–290.

Stephen, V. C. and Hockey, P. A. (2007). Evidence for an increasing incidence and severity of Harmful Algal Blooms in the southern Benguela region. *South African Journal of Science*, 103:223–231.

Stramski, D., Boss, E., Bogucki, D., and Voss, K. J. (2004). The role of seawater constituents in light backscattering in the ocean. *Progress in Oceanography*, 61(1):27–56.

Stramski, D., Bricaud, A., and Morel, A. (2001). Modeling the inherent optical properties of the ocean based on the detailed composition of the planktonic community. *Applied Optics*, 40(18):2929–2945.

Stramski, D. and Kiefer, D. A. (1991). Light-Scattering by Microorganisms in the Open Ocean. *Progress in Oceanography*, 28(4):343–383.

Stumpf, R. P., Culver, M. E., Tester, P. A., Tomlinson, M., Kirkpatrick, G. J., Pederson, B. A., Truby, E., Ransibrahmanakul, V., and Soracco, M. (2003). Monitoring Karenia brevis blooms in the Gulf of Mexico using satellite ocean color imagery and other data. *Harmful Algae*, 2(2):147–160.

Subramaniam, A., Carpenter, E., and Falkowski, P. G. (1999). Bio-optical properties of the marine diazotrophic cyanobacteria Trichodesmium spp. II. A reflectance model for remote sensing. *Limnology and Oceanography*, 44(3):618–627.

Suraya, Y., Oishi, T., Kishino, K., Jodai, Y., and Tanaka, A. (1996). Influence of ship shadow on underwater irradiance fields . In *Ocean Optics XIII*.

Svensen, Ø., Frette, Ø., and Erga, S. R. (2007). Scattering properties of microalgae: the effect of cell size and cell wall. *Applied Optics*, 46(23):5762–5769.

Thomalla, S. J., Fauchereau, N., Swart, S., and Monteiro, P. M. S. (2011). Regional scale characteristics of the seasonal cycle of chlorophyll in the Southern Ocean. *Biogeosciences*, 8:2849–2866.

Thomas, A., Strub, P., Carr, M.-E., and Weatherbee, R. (2004). Comparisons of chlorophyll variability between the four major global eastern boundary currents. *International Journal of Remote Sensing*, 25(7-8):1443–1447.

Tilstone, G. H., Peters, S. W. M., van der Woerd, H. J., Eleveld, M. A., Ruddick, K., Schönfeld, W., Krasemann, H., Martinez-Vicente, V., Blondeau-Patissier, D., Röttgers, R., Sørensen, K., Jørgensen, P. V., and Shutler, J. D. (2012). Variability in specific-absorption properties and their use in a semi-analytical ocean colour algorithm for MERIS in North Sea and Western English Channel Coastal Waters. *Remote Sensing of Environment*, 118:320–338.

Tomlinson, M. C., Wynne, T. T., and Stumpf, R. P. (2009). An evaluation of remote sensing techniques for enhanced detection of the toxic dinoflagellate, Karenia brevis. *Remote Sensing of Environment*, 113(3):598–609.

Tzortziou, M., Herman, J. R., Gallegos, C. L., Neale, P. J., Subramaniam, A., Harding, Jr., L. W., and Ahmad, Z. (2006). Bio-optics of the Chesapeake Bay from measurements and radiative transfer closure. *Estuarine, Coastal and Shelf Science*, 68(1-2):348–362.

Uitz, J., Claustre, H., Gentili, B., and Stramski, D. (2010). Phytoplankton class-specific primary production in the world's oceans: seasonal and interannual variability from satellite observations. *Global Biogeochemical Cycles*, 24.

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *Journal of Geophysical Research*, 111:C08005.

Uitz, J., Huot, Y., Bruyant, F., Babin, M., and Claustre, H. (2008). Relating phytoplankton photophysiological properties to community structure on large scales. *Limnology and Oceanography*, 53(2):614–630.

Utermöhl, H. (1931). Neue Wege in der quantitativen Erfassung des Planktons. (Mit besonderer Beriicksichtigung des Ultraplanktons.). *Verhandlungen des Internationalen Verein Limnologie*, pages 567–595.

Vaillancourt, R. and Brown, C. (2004). Light backscattering properties of marine phytoplankton: relationships to cell size, chemical composition and taxonomy. *Journal of Plankton Research*, 26(2):191–212.

Vidussi, F., Claustre, H., Manca, B. B., Luchetta, A., and Marty, J.-C. (2001). Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. *Journal of Geophysical Research*, 106(C9):19939–19956.

Volten, H., de Haan, J., Hovenier, J., Schreurs, R., Vassen, W., Dekker, A., Hoogenboom, H., Charlton, F., and Wouts, R. (1998). Laboratory measurements of angular distributions of light scattered by phytoplankton and silt. *Limnology and Oceanography*, 43(6):1180–1197.

Voss, K. and Chapin, A. (2005). Upwelling radiance distribution camera system, NURADS. *Optics Express*, 13(11):4250–4262.

Voss, K. J., Nolten, J. W., and Edwards, G. D. (1986). Ship Shadow Effects On Apparent Optical Properties . In *Ocean Optics IIX*.

Walters, M. N., Kok, J. C., and Claase, C. (1985). Optical properties of the South African marine environment . In Shannon, L. V., editor, *South African ocean colour and upwelling experiment*, pages 157–174. Sea Fisheries Research Institute, Cape Town, South Africa.

Wang, P., Boss, E. S., and Roesler, C. (2005). Uncertainties of inherent optical properties obtained from semianalytical inversions of ocean color. *Applied Optics*, 44(19):4074–4085.

Ward, B. A., Dutkiewicz, S., Jahn, O., and Follows, M. J. (2012). A size-structured food-web model for the global ocean. *Limnology and Oceanography*, 57(6):1877.

Wasmund, N., Nausch, G., and Matthäus, W. (1998). Phytoplankton spring blooms in the southern Baltic Sea—spatio-temporal development and long-term trends. *Journal of Plankton Research*, 20(6):1099–1117.

Weber, T. and Deutsch, C. (2010). Ocean nutrient ratios governed by plankton biogeography. *Nature*, 467:550–554.

Weeks, S. J., Barlow, R., Roy, C., and Shillington, F. A. (2006). Remotely sensed variability of temperature and chlorophyll in the southern Benguela: upwelling frequency and phytoplankton response. *African Journal of Marine Science*, 28(3-4):493–509.

Weeks, S. J., Pitcher, G. C., and Bernard, S. (2003). Satellite monitoring of the evolution of a Coccolithophorid bloom in the Southern Benguela upwelling system. *Oceanography*, 17(1):83–89.

Weir, C. T., Siegel, D. A., Michaels, A. F., and Menzies, D. W. (1994). In-situ evaluation of a ship's shadow. In *Ocean Optics XII*. SPIE.

Werdell, P. and Bailey, S. (2005). An improved bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment*, 98:122–140.

Whitmire, A. L., Boss, E., and Cowles, T. J. (2007). Spectral variability of the particulate backscattering ratio. *Optics Express*, 15(11):7019–7031.

Whitmire, A. L., Pegau, W. S., Karp-Boss, L., Boss, E., and Cowles, T. J. (2010). Spectral backscattering properties of marine phytoplankton cultures. *Optics Express*, 18(14):15073–15093.

Wright, V. M. and Hooker, S. B. (2009). AOP Workshop: Development of a Web-based Community AOP Processor . In *AOP Workshop, University of California Santa Barbara*.

Yentsch, C. S. (1962). Measurement of visible light absorption by particulate matter in the ocean. *Limnology and Oceanography*, 7(2):207–217.

Yentsch, C. S. and Phinney, D. A. (1989). A bridge between ocean optics and microbial ecology. *Limnology and Oceanography*, 34(8):1694–1705.

Yoder, J. A. and Kennelly, M. (2006). What have we learned about ocean variability from satellite ocean color imagers? *Oceanography*, 19(1):152–171.

Yoo, Y. D., Jeong, H. J., Kim, M. S., Kang, N. S., Song, J. Y., Shin, W., Kim, K. Y., and Lee, K. (2009). Feeding by Phototrophic RedâĂŘTide Dinoflagellates on the Ubiquitous Marine Diatom Skeletonema costatum. *Journal of Eukaryotic Microbiology*, 56(5):413–420.

Zaneveld, J. R. V. (1995). A Theoretical Derivation of the Dependence of the Remotely-Sensed Reflectance of the Ocean on the Inherent Optical-Properties. *Journal of Geophysical Research*, 100(C7):13135–13142.

Zaneveld, J. R. V., Boss, E., and Hwang, P. (2001). The influence of coherent waves on the remotely sensed reflectance. *Optics Express*, 9(6):260–266.

Zhang, H.-M., Bates, J. J., and Reynolds, R. W. (2006). Assessment of composite global sampling: Sea surface wind speed. *Geophysical Research Letters*, 33(17):L17714.

Zhang, T. L., Liu, B., and Qiu, G. Q. (2010). An Empirical Method for Determining the Water Leaving Radiance from TSRB Radiometer Measurements. *Periodical of Ocean University of China*.

Zhang, X., Huot, Y., Gray, D. J., Weidemann, A., and Rhea, W. J. (2011). Biogeochemical origins of particles obtained from the inversion of the volume scattering function and spectral absorption in coastal waters. *Applied Optics*, 50(9):1240–1259.

Zhou, W., Wang, G., Sun, Z., Cao, W., Xu, Z., and Hu, S. (2012). Variations in the optical scattering properties of phytoplankton cultures. *Optics Express*, 20(10):11189–11206.

Zhu, J., Chen, J., Matsushita, B., Yang, W., and Fukushima, T. (2012). Atmospheric correction of ENVISAT/MERIS data over case II waters: the use of black pixel assumption in oxygen and water vapour absorption bands. *International Journal of Remote Sensing*, 33(12):3713–3732.

Zibordi, G., Berthon, J.-F., and D'Alimonte, D. (2009a). An Evaluation of Radiometric Products from Fixed-Depth and Continuous In-Water Profile Data from Moderately Complex Waters. *Journal of Atmospheric and Oceanic Technology*, 26:91–106.

Zibordi, G., Berthon, J. F., Mélin, F., and D'Alimonte, D. (2011). Cross-site consistent in situ measurements for satellite ocean color applications: BiOMaP radiometric data set. *Remote Sensing of Environment*, 115:2104–2115.

Zibordi, G., Berthon, J.-F., Mélin, F., D'Alimonte, D., and Kaitala, S. (2009b). Validation of satellite ocean color primary products at optically complex coastal sites: Northern Adriatic Sea, Northern Baltic Proper and Gulf of Finland. *Remote Sensing of Environment*, 113:2574–2591.

Zibordi, G. and Bulgarelli, B. (2007). Effects of cosine error in irradiance measurements from field ocean color radiometers. *Applied Optics*, 46(22):5529–5538.

Zibordi, G., D'Alimonte, D., and Berthon, J.-F. (2004a). An evaluation of depth resolution requirements for optical profiling in coastal waters. *Journal of Atmospheric and Oceanic Technology*, 21:1059–1073.

Zibordi, G. and Darecki, M. (2006). Immersion factors for the RAMSES series of hyper-spectral underwater radiometers. *Journal of Optics A: Pure and Applied Optics*, 8(3):252–258.

Zibordi, G., Doyle, J. P., and Hooker, S. B. (1999). Offshore tower shading effects on in-water optical measurements. *Journal of Atmospheric and Oceanic Technology*, 16:1767–1779.

Zibordi, G., Holben, B., Hooker, S. B., Mélin, F., and Berthon, J. (2006a). A network for standardized ocean color validation measurements. *EOS Transactions*, 87(30):293–304.

Zibordi, G., Mélin, F., and Berthon, J. (2006b). A time-series of above-water radiometric measurements for coastal water monitoring and remote sensing product validation. *IEEE Geoscience and Remote Sensing Letters*, 3:120–124.

Zibordi, G., Mélin, F., Hooker, S., D'Alimonte, D., and Holben, B. (2004b). An Autonomous Above-Water System for the Validation of Ocean Color Radiance Data. *IEEE Transactions on Geoscience and Remote Sensing*, 42:401–415.

Zibordi, G., Ruddick, K., Ansko, I., Moore, G., Kratzer, S., Icely, J., and Reinart, A. (2012). In situ determination of the remote sensing reflectance: an inter-comparison. *Ocean Science Discussions*, 9(2):787–833.

Zibordi, G. and Voss, K. J. (2010). Field Radiometry and Ocean Color Remote Sensing. In Barale, V., editor, *Oceanography from Space*, pages 1–28. Springer.