2020-12-15

Dynamic partitioning of tropical Indian Ocean surface waters using ocean colour data management and modelling applications

Schwarz, Jill

http://hdl.handle.net/10026.1/16413

10.1016/j.jenvman.2020.111308
Journal of Environmental Management
Elsevier BV

Dynamic partitioning of tropical Indian Ocean surface waters using ocean colour data - management and modelling applications

Jill N. Schwarz

School of Biological & Marine Sciences, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, UK

Abstract

Over the past few decades, partitioning of the surface ocean into ecologically-meaningful spatial domains has been approached using a range of data types, with the aim of improving our understanding of open ocean processes, supporting marine management decisions and constraining coupled ocean-biogeochemical models. The simplest partitioning method, which could provide low-latency information for managers at low cost, remains a purely optical classification based on ocean colour remote sensing. The question is whether such a simple approach has value. Here, the efficacy of optical classifications in constraining physical variables that modulate the epipelagic environment is tested for the tropical Indian Ocean, with a focus on the Chagos marine protected area (MPA). Using remote sensing data, it was found that optical classes corresponded to distinctive ranges of wind speed, wind stress curl, sea surface temperature, sea surface slope, sea surface height anomaly and geostrophic currents (Kruskal-Wallis and post-hoc Tukey honestly significantly different

Email address: jill.schwarz@plymouth.ac.uk (Jill N. Schwarz)

tests, $\alpha = 0.01$). Between-class differences were significant for a set of subdomains that resolved zonal and meridional gradients across the MPA and Seychelles-Chagos Thermocline Ridge, whereas between-domain differences were only significant for the north-south gradient (PERMANOVA, $\alpha = 0.01$). A preliminary test of between-class differences in surface CO_2 concentrations from the Orbiting Carbon Observatory-2 demonstrated a small decrease in mean pCO_2 with increasing chl, from 418 to 398 ppm. Simple optical class maps therefore provide an overview of growth conditions, the spatial distribution of resources – from which habitat fragmentation metrics can be calculated, and carbon sequestration potential. Within the 18 year study period, biotic variables were found to have decreased at up to $0.025\%\mathrm{a}^{-1}$ for all optical classes, which is slower than reported elsewhere (Mann-Kendall-Sen regression, $\alpha = 0.01$). Within the MPA, positive Indian Ocean Dipole conditions and negative Southern Oscillation Indices were weakly associated with decreasing chl, fluorescence line height (FLH), eddy kinetic energy, easterly wind stress and wind stress curl, and with increasing FLH/chl, sea surface temperature, SSH gradients and northerly wind stress, consistent with reduced surface mixing and increased stratification. The optical partitioning scheme described here can be applied in Google Earth Engine to support management decisions at daily or monthly scales, and potential applications are discussed.
Keywords: Remote sensing, biogeography, habitat fragmentation, Orbiting Carbon Observatory-2 (OCO-2), Marine Protected Area, epipelagic

1. Introduction

Physical drivers such as light, temperature and dynamical mixing shape the epipelagic ecosystem, and the biota of deeper water layers is determined, at least in part, by the productivity of upper layers (Jerlov, 1976; Longhurst, 2007; Spalding et al., 2012; Kavanaugh et al., 2016; Proud et al., 2017). Dividing the oceans into geographical areas with common physical conditions has been approached using a range of methods and suites of data (reviewed by Krug et al., 2017; Kavanaugh et al., 2016), and referred to variously as ecological geography, partitioning, biogeography, biohydrography, biogeographical provinces and seascapes. Most schemes include nested spatial scales with slightly different nomenclature for different elements of heirarchical structure. Partitioning the oceans is similar to habitat mapping and species distribution modelling in the sense that a geographical representation of resources is produced. However, it does not relate to specific organisms, and no model of the relationships between predictor and response variables is produced (c.f. Blanco et al., 2015; Coelho et al., 2013; Scales et al., 2014; Zydelis et al., 2011).Partitioning schemes have found application to two key challenges. Firstly, they provide static and dynamical geographical boundaries to guide management planning (over the long term) and intervention (in the short term). Management tasks include monitoring ecosystem health, assessing risk and implementing control measures such as fishery closure, and information is required at multiple spatial scales and depth ranges to support these actions (Rice et al, 2011; Spalding et al., 2012; Caldow et al., 2015; Roberson et al, 2017). Short-term responses can only be supported when the frequency at which new data is available exceeds the rate of critical fluctuations occurring within the ecosystem. The second use is to provide spatial context for the evaluation of climate model reliability (Vichi et al., 2011; Kavanaugh et al., 2014; Fay & McKinley, 2014). The focus here is on the exchange of climate-relevant gases across the air-sea interface and there is no requirement for low-latency information. Partitions serve as a proxy for ecosystem function with the inference that they constrain rates of CO_2 diffusion, biotic carbon uptake and the efficiency with which carbon is removed from the surface layer. Earth system models (ESM), in which the representation of biogeochemical cycles remains quite simplistic (Hense et al., 2017; Jung et al., 2019; Sreeush et al., 2018) and which are sensitive to feedback between biotic and abiotic components (Lim & Kug, 2017; Park & Kug, 2014; Romanou et al., 2014) can then be evaluated in the context of static or dynamic seascapes (Landschützer et al., 2019; Lovenduski et al., 2019).

Satellite remote sensing data, which currently provides the best compromise between area coverage and temporal resolution for surface ocean studies, has increasingly been used for partitioning, as the disparate worlds of marine management and oceanography converge on how this rich data source can best be used (Kachelriess et al., 2014; Maxwell et al., 2015; Miloslavich et al., 2018). Figure 1 illustrates some of the pelagic abiotic and biotic factors that can be derived from remote sensing data, and characteristics relevant for modelling and management decisions.

[FIGURE 1 HERE: SINGLE-COLUMN; COLOUR ONLY ONLINE]

Of the variables amenable to remote sensing, only ocean colour is directly affected by the pelagic ecosystem at short time-scales. Partitioning schemes mostly use an ocean-colour-derived variable, chlorophyll-a concentration (hereafter chl), as a measure of biomass and combine it with other key physical drivers such as temperature. However, being a function of the concentrations of dissolved and particulate substances in the surface ocean, ocean colour itself represents a response to the physical drivers. Using ocean colour to derive chl incurs spatially-varying errors (Jackson et al., 2017) and ignores other coloured variables of interest that may, or may not, covary with chl (O'Reilly et al., 1998; Werdell et al., 2018).

Classifications based on water colour alone have shown that, whilst chl is generally constrained by these classifications, information about other water mass tracers is also present (Alvain et al., 2005, 2008; Vantrepotte et al., 2012; Trochta et al., 2015; Krug et al., 2018; Monolisha et al., 2018; Dierssen, 2010). Taken together with the knowledge that physical dynamics at all scales combine to control the growth environment, this suggests an opportunity to use pure optical classes as the smallest scale in a seascape heirarchy, with the advantage that it is a low-latency product which could feed into decision-making flows on a daily basis (e.g. using GoogleEarthEngine; Gorelick et al., 2017) where coverage allows. This possibility is explored here, with a focus on the Chagos marine protected area (MPA). If optical classes are found to constrain abiotic drivers as well as biotic response, then a further question arises of whether they can also be used to estimate carbon flows (extending Kavanaugh et al., 2014) without the need to identify individual elements, such as phytoplankton function type, as an intermediate step.

1.1. The study area

The Chagos marine protected area occupies 640,000 km^2 in the tropical Indian Ocean, with the Chagos Archipelago system of islands and atolls at it centre (Figure 2; UNEP-WCMC, 2016). The coastline of the Indian Ocean is shared amongst over twenty countries, representing a substantial fraction of the human population with variable socio-economic status and strong reliance on coastal and open-ocean fisheries (Hermes et al., 2019). The MPA is of particular value because of its coral health, resilience and diversity, extensive seagrass beds, potential support for the wider Indian Ocean fisheries and related benefits (Koldeway et al., 2010; Ateweberhan et al., 2018; Gravestock & Sheppard, 2015; Esteban et al., 2018). In common with shallow tropical corals around the world, reefs in the Chagos MPA are vulnerable to temperature increases associated with climate change as well as to increases in extreme high energy dynamics. Their relative resilience compared with other reef systems is associated with protection from human disturbance as well as to geographical location. Although extensive coral bleaching has occurred (Sheppard et al., 2008), interactions between dynamical processes at a range of scales and topographic diversity may alleviate temperature stress (Sheppard, 2012; Hosegood et al., 2019). Understanding whether this natural protection will continue under ongoing climate change is important in terms of economic as well as natural resource value.

The Chagos Archipelago lies at the edge of the South Indian tropical gyre within the influence of the Indian Ocean monsoon. In austral summer, northerly winds drive an anticyclonic cell and the gyre contracts, so that the archipelago is at the northern edge. For the rest of the year, the MPA lies at or near the southern edge of the gyre. The location and large-scale circulation features are summarised in Figure 2. In the west, divergent winds and negative wind stress curl lift the thermocline along 5-12°S creating the

Seychelles-Chagos Thermocline Ridge (SCTR), with a thermocline depth around of 50 m and no surface signature, which extends around the MPA throughout the year (Hermes & Reason, 2008, 2009; Aguiar-Gonzalez et al., 2016; Xie et al., 2002).

[FIGURE 2 HERE: SINGLE or 1.5 COLUMN; COLOUR ONLY ON-LINE]

The South Equatorial Current (SEC), flowing to the south of the MPA, denotes the boundary between relatively nutrient replete but O_2-poor surface waters to the north and southern sub-tropical gyre waters to the south which are nutrient-depleted throughout the water column but represent a CO_2 sink (Garcia et al., 2018; Landschützer et al., 2016). *In situ* biogeochemical data are sparse across the tropical Indian Ocean and considerable deviations from the mean conditions in the World Ocean Atlas have been reported (Subha Anand et al., 2017; Chinni et al., 2019). Whilst the tropics are generally considered to be oligotrophic, year-round elevated phytoplankton biomass is observed close to the archipelago as well as over the Mascarene Plateau to the west and broadly over the SCTR (Wilson & Qiu, 2008; Levy et al., 2007). *In situ* measurements of net primary production in this region range from close to zero up to 20 $\mathrm{mgCm^{-2}d^{-1}}$ and can be exceeded by bacterial production (Subha Anand et al., 2017; Fernandes et al., 2008; Veldhuis et al., 1997). A few high temporal resolution datasets from moored fluorometers have shown high frequency, high magnitude fluctuations in phytoplankton biomass (Hosegood et al., 2019; Strutton et al., 2015). Phytoplankton assemblages have been found to be dominated by Prochlorococcus and Synecococcus as expected in the oligotrophic gyres, but substantial fractions of diatoms, dinoflagellates and prymnesiophytes have also been reported in the TIO (Thorrington-Smith, 1971; Veldhuis et al., 1997; Soares et al., 2015). To my knowledge, there are no long-term biogeochemical monitoring efforts in the pelagic SCTR, despite detailed repeat monitoring in the shallow reef waters of the MPA (e.g. Sheppard, 2012). Many studies have used coupled ocean-biogeochemical models, together with available *in situ* or remote sensing data, to elucidate biophysical coupling in the tropical Indian Ocean (TIO) (e.g. Wiggert et al., 2006; Jin et al., 2012; Liu et al., 2013; Resplandy et al., 2009; George et al., 2018). Of particular interest here are the results of George et al. (2018), Dilmahamod et al. (2016) and Wiggert et al. (2006), who explore meridional and zonal gradients in the SCTR.

The epipelagic growth environment is directly modulated by entrainment and advection of nutrients and plankton, fluctuations in mixed layer temperature and depth, the relative euphotic to mixed layer depths, turbulence and varying illumination conditions. Conversely, feedback effects have been demonstrated between chl and shortwave heating, SST, surface convergence and basin-scale dynamical features (Back & Bretherton, 2009; Park & Kug, 2014). At the seasonal scale, the eastward extent of the SCTR and westward extent of Indonesian Throughflow (ITF) respond to monsoon wind weakening and reversal (Aguiar-Gonzalez et al., 2016). Two of the eight Madden-Julian Oscillation (MJO) phases are centred in the TIO (Hendon & Salby, 2994), with a westward-progagating Rossby wave (Seiki et al., 2013) impacting SST, evaporation, precipitation, cloud cover, rainfall, salinity gradients (Guan et al., 2014; Jin et al., 2013; McPhaden & Foltz, 2013) and wind-driven entrainment of nutrients into the mixed layer (Jin et al., 2012b). At the interannual scale, El Nino and the Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) events have been reported to affect surface conditions and surface chlorophyll concentrations. Nino conditions affect mixed layer temperatures through precipitation, downwelling (anticyclonic) winds and westward propagating Rossby waves (Santoso et al., 2010; Dilmahamod et al., 2016; Ma et al., 2014; Racault et al., 2017). Positive IOD phases coincide with cooler surface temperatures in the eastern TIO and warmer conditions to the west (SST anomalies of 0.1 to 0.3 °C or more; Currie et al., 2013; Vialard et al., 2009). The impact of ENSO and IOD events is amplified when they coincide, and both are expected to increase in frequency (IPCC, 2013; Sheppard et al., 2008; Currie et al., 2013; Cai et al., 2014).

Despite the scarcity of data in the epipelagic in and around the Chagos MPA, the importance of conditions in these waters has recently been highlighted by electronic tagging of seabirds and high trophic level pelagic feeders, which has been used to document foraging at considerable distances (Pecoraro et al., 2017; Danckwerts et al., 2014; Le Corre et al., 2012), with measurable positive impacts on reef health through nutrient redistribution (Graham et al., 2018). The near-shore pelagic biome is also a critical food resource for corals (Houlbréque & Ferrier-Pagés, 2009) and the strength of this relationship has now been demonstrated using remotely sensed ocean colour data (Fox et al., 2019). Elevated chl related to the wind-driven circulation around the Chagos MPA is therefore a potential resource both for the reef and for pelagic organisms.

The underlying hypothesis for this study is that water colour represents the evolving trophic status of the upper ocean and is characteristic of the suite of physical drivers of ecosystem function as illustrated in Figure 1 (e.g. Jerlov, 1976), making it a candidate as a low-latency, fine-scale level in a seascape heirarchy. This paper addresses how methodological choices affect optical classifications, whether pure optical classes provide useful information about the physical environment and potential applications of optical class maps.

2. Methods

2.1. Remote sensing data

This study spans August 2002 to October 2018. Remote sensing reflectance (Rrs), chlorophyll-a concentration (chl), normalised fluorescence line height (nFLH) and sea surface temperature (SST) data from the NASA Moderate Resolution Imaging Spectrometer aboard the Aqua satellite (MODIS-Aqua), at Level 2 and Level 3, were acquired from the Ocean Biology Processing Group (oceancolor.gsfc.nasa.gov). Rrs is the ocean colour product with the least degree of processing and therefore the lowest uncertainty, with errors on the order of 0.001 sr^{-1} but varying with waveband and water type (Franz et al., 2007; IOCCG, 2019). Reflectances from the seven 1 km resolution, 10 nm wavebands in the visible domain were augmented by band 1, with 250 m resolution (50 nm waveband) and bands 3 and 4, with 500 m resolution (20 nm wavebands), and these three bands were spatially averaged to match the 1 km wavebands. Globally, chl is the best-validated ocean colour product, with mean errors of ca. ± 33 mgm^{-3} (Hu, et al., 2012; O'Reilly et al., 1998). Little product validation data is available for nFLH in the tropical Indian Ocean, but the MODIS-Aqua and MODIS-Terra products perform well against *in situ* data in the Southern and Atlantic Oceans (Erickson et al., 2019; Hoge et al., 2003).

Daily sea surface height (SSH), height anomalies (SSHA) and geostrophic current velocities (denoted eastwards u and northwards v), from merged altimeter datasets, were acquired at ca. 30 km resolution from the Copernicus Marine Environment Monitoring Service (CMEMS). Reported errors on these products range from < 1 cm to 30 cm, with higher uncertainties under more dynamic conditions (CMEMS, 2020). Eddy kinetic energy was calculated as EKE = $1/2(u^2 + v^2)$. Sea surface slope was calculated pixelwise as $\nabla SSH = \partial SSH/\partial x + \partial SSH/\partial y$, with no smoothing. Each product was subsampled to 4 km resolution before applying the optical class masks (section 2.7).

Daily surface wind fields from the SeaWinds and ASCAT scatterometer sensors were acquired at 25 km resolution from the Jet Propulsion Laboratory Physical Oceanography Distributed Archive Center (SeaPAC, 2006; EUMETSAT/OSI SAF, 2018). Errors in these products are of order 0.1 ms^{-1} (Verhoef et al., 2017). Wind stress curl was calculated pixelwise as $\nabla \times \tau = -\partial\tau_N/\partial x - \partial\tau_E/\partial y$, where τ_N and τ_E are the north-south and east-west components of the wind stress and ∂x and ∂y are the pixel dimensions. Monthly averages were calculated before subsampling to 4 km resolution and applying the optical class masks (section 2.7)

Profiles of the partial pressure of CO_2 derived from Orbiting Carbon Observatory 2 (OCO-2) data using the full physics model version 7.3 (O'Dell et al., 2018) were acquired at native resolution (ca. 3 km^2) from the Goddard Earth Sciences Data and Information Services Center (OCO-2 Science Team, 2016, GES DISC). Only estimates of pCO_2 from the lowest model altitude were used and the ramifications of this are discussed in section 3.4. No spatial or temporal compositing was applied.

Bathymetry data were acquired at 1 arc-second resolution from the National Centers for Environmental Information (ETOPO1, 2019) and regridded using nearest-neighbour gridding to match the Level 3 Rrs data.

Data products, sources and abbreviations are summarised in table 1.

[TABLE 1 HERE]

2.2. Dynamic partitioning based on ocean colour

Fuzzy classification was applied to the Indian Ocean domain surrounding the BIOT MPA (after Moore et al., 2001; Jackson et al., 2017). This method was chosen for its potential to allow a single pixel to have multiple class memberships, which is likely in natural phytoplankton populations, particularly at the relatively coarse spatial scales of remote sensing data (1 to >30 km) in waters where mesoscale and submesoscale processes may be at play. The study bounds were 40^o to 100^o E, -20^o to 15^o N, spanning the central tropical Indian Ocean with the BIOT MPA roughly central to the domain (Figure 3; UNEP-WCMC, 2016). In the absence of *in situ* data with which to verify class memberships or interpret class composition, only the dominant class assigned to each pixel at any given time was retained (multiple class memberships were removed, to be considered in future work when validation data are available). Biovolumes calculated from miscroscopy analysis on five stations within the MPA were used as a preliminary test of whether different classes represented different phytoplankton biomass (Kruskal-Wallis test; Schwarz, 2020).

Variable	Abbreviation	Source
Remote sensing reflectance, Level 2	L2 Rrs	Ocean Biology Processing Group Level 2 data oceancolor.gsfc.nasa.gov
Remote sensing reflectance, Level 3	L3 Rrs	MODIS-Aqua Level 3 monthly data, Giovanni data portal giovanni.gsfc.nasa.gov
Surface chlorophyll-a concentration	chl	
Surface normalised fluorescence line height	nflh	
Ratio of fluorescence line height to chlorophyll-a	flh:chl	
Sea surface temperature	SST	
Metric for within-class spectral variability (chlorophyll-like pigments): ratio of residual reflectances (Rrs - class mean) at 555 nm to 443 nm	R555:443	
Metric for within-class spectral variability (accessory pigments): ratio of residual reflectances (Rrs - class mean) at 555 nm to 488 nm	R555:488	
Sea surface height (absolute dynamic topography)	SSH	AVISO Level 4 reprocessed gridded sea surface heights and derived variables (product suite SEALEVEL GLO_PHY_L4 REP_OBSERVATIONS_008_047), marine.copernicus.eu, 0.25^o x 0.25^o regridded to 4 x 4 km
Sea surface slope	∇ SSH	
Sea surface height anomaly	SLA	
Eastward component of the geostrophic current	u	
Northward component of the geostrophic current	v	
Eddy kinetic energy	EKE	
Eastward component of the surface wind field	τ_E	Quikscat and ASCAT Level 3 gridded wind fields, podaac.jpl.nasa.gov, 25 x 25 km regridded to 4 x 4 km
Northward compoent of the surface wind field	τ_N	
Wind stress curl	$\nabla \times \tau$	
Water depth	z	ETOPO 1 arc-minute bathymetry regridded to 4 x 4 km (Amante & Eakins, 2009)
Indian Ocean Dipole index	IOD	Dipole mode index (Saji & Yamagata, 2003) esrl.noaa.gov/psd/gcos_wgsp/Timeseries/DMI
Madden Julian Oscillation index	MJO	Kilidas et al. (2014) esrl.noaa.gov/pas/mjo/mjoindex
Southern Oscillation Index	SOI	Ropelewski & Jones (1987) esrl.noaa.gov/psd/data/20thC_Rean/ timeseries/monthly/SOI
Surface partial pressure of CO_2	pCO_2	Level 2 OCO-2 physical model surface pCO_2, release 9, GES DISC (Boesch et al., 2019)

Table 1: Optical, biological and physical remote sensing products used to explore optical classes

[FIGURE 3 HERE: SINGLE-COLUMN or 1.5 COLUMN; PLEASE QUOTE FOR COLOUR PRINTING COSTS]

2.3. Impact of spatial and temporal compositing on optical classification

As noted by Jackson et al. (2017, hereafter JSM), averaging of the remotely sensed reflectance spectra from instantaneous, Level 2 (ca. 1 x 1 km resolution) to Level 3 (ca. 4 x 4 km resolution, weekly or monthly) data incurs a risk of smoothing out phytoplankton dynamics associated with growth/decay/advection events that are short-lived, and it increases the difference between any available *in situ* reflectance data used in characterising or evaluating optical classes from the already spatially-averaged remote sensing pixel values. Daily data may also be most appropriate for some management applications. However, cloud-cover and the reduced overpass rates at low latitudes make the use of daily data for capturing spatial patterns challenging in this region. The impact of spectral aliasing (from averaging) on the classification was tested by creating classes firstly from all the Level 2 (L2) data for 2003, with no averaging or regridding. The resulting classes were compared against the same suite of classification procedures applied to the Level 3 (L3) monthly, 4 km gridded data for 2003 and against the hybrid classification scheme produced by JSM and Moore et al. (2001) who used *in situ* data.

Both the L2 and L3 data were classified in three forms: Remote sensing reflectance (Rrs), Rrs with the mean 2003 Rrs subtracted (Rrs-$\overline{Rrs2003}$) and Rrs normalized to 488 nm (Rrs/Rrs488). Training data were selected randomly in space and time from the study domain. For comparability with previous studies, three sizes of training dataset were used. The smallest dataset contained 42,000 pixels, corresponding approximately to the volume of data used by JSM for two of their study sites - they used two central Indian Ocean sites referenced to Longhurst (2007) provinces. Additional datasets of double and fifty times the original size were added to test for sensitivity of the classification scheme to dataset size. Each dataset was classified using the Matlab fcm function (Bezdek, 1981) with between 2 and 9 classes and the weighting exponent m was varied between 1.05 and 2.0. Class separability and compactness were assessed using the partition coefficient (F) and compactness and separation index (S) as in Moore et al. (2001); Windham (1982); Xie & Beni (1991). In contrast to previous studies, all ten available MODIS visible wavebands were used in the fuzzy classification procedure.

2.4. Interpretation of within-class spectral variability

The use of ten wavebands for an optical classification allows limited exploration of within-class spectral variability, which may be related to pigmentation and size differences caused by change in phytoplankton community composition or physiology, differences in backscatter related to the viral, bacterial and phytoplankton communities, variability in the relative quantities of coloured, dissolved organic matter or inorganic particulate matter, variability in the depth distributions of coloured materials and noise in the satellite signal (Kirk, 1994; Brown et al., 2008; Defoin-Platel & Chami, 2007; Alvain et al., 2005; Lain & Bernard, 2018; Brewin et al., 2011b). Having excluded water depths shallower than 200 m and in the absence of *in situ* validation data, the main focus here is on testing whether Rrs spectra varied uniformly with optical class. Residual reflectance ratios Rrs(555)/Rrs(443) and Rrs(555)/Rrs(488) were calculated after subtracting the dominant class mean reflectance spectrum (Table 4) at each pixel, as indicators of spectral variability. In a full application of the fuzzy classification scheme (Moore et al., 2001), this step could be pre-empted by allowing multiple class memberships at each pixel.

2.5. Comparison against other classifications

Class spectra produced in this study and by JSM were mapped to other interpretations for comparison. Firstly, the standard NASA OC3M-CI algorithms were applied to each central class spectrum to produce chl concentration (O'Reilly et al., 1998; Hu, et al., 2012). This was passed to abundance-based algorithms for phytoplankton size class published by Brewin et al. (2010, Atlantic Ocean), Brewin et al. (2011a, global *in situ* data), Brewin et al. (2012, eastern Indian Ocean) and Devred et al. (2011) (North Atlantic and global *in situ* data; IOCCG, 2014). Chl was also used to select the closest stratified water trophic class from Uitz et al. (2006). For comparison against Alvain et al. (2005, 2008), every 10th Level 2 MODIS Rrs file from 2003 was used to generate local PHYSAT-equivalent mean-chl spectra (39 chl divisions from 0.01 to 4.00 mgm^{-3} in intervals of ln(0.15); between 67016 and 5488110 pixels per chl interval with a total of 64733448 pixels). These were subtracted from each mean optical class spectrum and the residual compared against the criteria provided by Alvain et al. (2008).

Name	Weighting coefficient m	Partition coefficient F	Separation index S	Source	
N5	1.05	10	0.986	0.211	This study
N8	1.05	10	0.982	0.342	This study
JSM	Unknown	6	Unknown	Unknown	Jackson et al. (2017)

Table 2: Classification schemes used for further analysis. See supplementary online material, Table S1, for full classification evaluation metrics.

2.6. Classification of the 17 year Level 3 dataset

Following evaluation of the L2 and L3 classifications, a suite of three classification schemes was chosen and applied to 17 full years of MODIS-Aqua Level 3 Rrs version 2018.0 (Table 2). Water depths shallower than 200 m were excluded from analysis. This provided the framework for testing whether other remotely sensed parameters were well constrained by the optical classes.

[TABLE 2 HERE]

2.7. Relationship between colour class and biophysical parameters

Summary statistics of remotely sensed biological and physical variables were produced to characterize each class in each of seven geographical domains (Figure 3): SCTR-W (10^o to 5^o S, 50^o to 62^o E), SCTR-E (10^o to 5^o S, 63^o to 75^o E), SCTR-C (10^o to 5^o S, 76^o to 88^o E), Wiggert-N (7^o to 2^o S, 60^o to 85^o E), Wiggert-S (12^o to 7^o S, 60^o to 85^o E), the MPA, z200 (the whole study domain at depths greater than 200 m) and z1000 (the whole study domain at depths greater than 1000 m).

The potential of purely optical classifications as indicators of the pelagic growth environment was explored using within- and between-class statistics. Non-parametric analysis of variance was used to test whether the optical classifications reflected differences in each of the the biological and physical variables (Kruskal-Wallis, applied to one variable at a time, assuming that underlying distributions of each variable within each class subset of each geographical domain are the same, followed by post-hoc Tukey's honestly significantly difference, HSD, tests between pairs of classes; Ruxton & Beauchamp, 2008) for each of the geographical domains. The distinctiveness of smaller domains for which contrasting processes across zonal or meridional gradients have been discussed in the literature (SCTR-W,E,C; Wiggert-N, S Wiggert et al., 2006; George et al., 2018; Dilmahamod et al., 2016) was tested using permuted multivariate analysis of variance (PERMANOVA; Anderson, 2001, 2017) on standardised variables (z-scores), with all fifteen biological and physical variables. Anomaly time-series were used in addition to simple z-scores where clear seasonal cycles were present. Because of the large dataset sizes, 100 subsets of 1000 pixels were selected randomly through time within each geographical domain for bootstrapped testing; pixels with missing data were excluded at each iteration. Domain was taken as the first, fixed, factor and class as the second, nested factor. Euclidean distance, correlation distance and squared correlation distance gave similar PERMANOVA hypothesis test results, whilst χ and χ^2 results varied; euclidean distance and 1000 permutations were used for all reported results (McCune & Grace, 2002; Anderson, 2017; Pillar, 2013). The fathom Matlab toolbox was used for PERMANOVA tests (Jones, 2012). Mann-Kendall-Sen correlation coefficients (Sen, 1968; Hamed & Rao, 1998) were used to identify which biological and physical variables covaried within each class. Correlations were tested firstly using all data and secondly using bootstrapping to insure against spatial and temporal autocorrelation effects (100 random subsamples of 1000 pixels).

A full comparison of optical classes with OCO-2 pCO_2 data was beyond the scope of this study - OCO-2 was launched in July, 2014, and the data therefore do not span the study period considered here. OCO-2 data are also lower in spatial and temporal coverage, although the recent launch of OCO-3 will mitigate this. As a proof of concept, a single month of OCO-2 surface pCO_2 data were matched to MODIS Level 2 optical classes for January 2015 (same-day match-ups only), and between-class differences evaluated using Kruskal-Wallis and post-hoc HSD tests, for the z1000 study domain.

Assuming that optical classifications do partition pixels dynamically in space and time according to the physical and biological variables that can be derived from remote sensing data, two additional tests were applied to establish whether conditions within pixels assigned to each class changed during the study period, and whether they are correlated with basin- and global-scale circulation patterns. Mann-Kendall-Sen trend tests were applied in regressions of the class-averaged time-series of each remote sensing variable against climate indices that characterize the Indian Ocean Dipole (Dipole Mode Index, referred to hereafter as IOD; Saji & Yamagata, 2003), Madden-Julian Oscillation (MJO; Kilidas et al., 2014) and El Nino-Southern Oscillation, Southern Oscillation index (SOI; Ropelewski & Jones, 1987). The SOI was chosen because it represents variability in the Walker circulation, rather than directly in SST or combinations of variables. Habitat fragmentation metrics were used to characterise the distribution of lower trophic level resources in the MPA, Wiggert-S and Wiggert-N domains to investigate whether changes related to climate patterns can be detected in the pelagic growth environment using available remote sensing data between 2002 and

2018. Patches of the same optical class were created using each month's class map as a binary image, grouping adjacent pixels of like-class (Matlab bwconncomp) and finding the perimeter, patch centre and number of pixels contained within each class patch (Matlab regionprops). Average patch area, distance between patches and patch density were calculated (Wang et al., 2014) for each month for temporal regression and regression against climate indices (Mann-Kendall-Sen test). Finally, correlations between the total area occupied by each optical class within each subdomain and climate indices were tested (Mann-Kendall-Sen trend test).

To test whether optical classifications could form a useful basis for fisheries management and enforcement, fishing catch and effort data from the Indian Ocean Tuna Commission were acquired at monthly temporal resolution, in 1^o and 5^o grids (IOTC, 2020, 2014). The most common optical class was assigned to each of these coarse fisheries grid cells, matched by year and month, and between-class differences in catch, effort and catch per unit effort (CPUE) were evaluated (Kruskal-Wallis and post-hoc HSD). This comparison was applied across the z1000 domain and within the IOTC data gridcells that contain the MPA as well as 15^o x 15^o and 25^o x 25^o domains centred on the MPA and the full z1000 area. Long-line fishing effort was reported in hooks, whereas surface fishery effort was reported in hours; the two datasets were analysed separately.

All data analysis and visualisation was carried out using Matlab 2018a running under MacOS10.12.6.

3. Results & Discussion

3.1. Comparison of classification results with Level 2 and Level 3 input data

Classification results were similar for L2 and L3 data for the two smaller datasets, but classes were less compact when the largest dataset was used. Partition coefficients (F) remained above 0.9 for all three Rrs treatments with weighting coefficients 1.05, 1.1 and 1.2, but separation coefficient S increased from O0.15 to O0.3 as the number of classes, N, was increased from two to nine, and the best separability was obtained for 2 classes with all data treatments. Beyond N=8 classes, processing time increased significantly and F and S values were unstable between repeat runs, therefore no classification with 9 or more classes was pursued. Classification performance metrics (F and S) for all L2 and L3 classifications are provided in full in the Supplementary Online Material, Table S1, sheet 'L2 L3 Classification F S'. Rrs treatment made little difference to F and S for Rrs and Rrs-$\overline{(Rrs(2003))}$ (<1% variability in performance metrics for m $<$ 1.3), but separation index increased by a factor of 20 for Rrs/Rrs(488).

Figure 4 shows the JSM classes and the L2 and L3 class spectra produced using Rrs with m = 1.05 and two to eight classes, and Table 3 shows the proportions of pixels mapped to each class, with N=5 and N=8 class schemes mapped to JSM classes using Euclidean distance between Rrs values at the six common wavebands. Fewer of the JSM classes were reproduced in L3 classifications than in L2 classifications, with the extreme blue-water spectra (highest and lowest Rrs(412) values) lost in an N=8-class scheme (Table 3). Class spectra fell within or at the edges of the JSM classes, with most deviation in spectral shape (relative to the closest JSM class) in the blue:green wavebands. Only one of the three JSM coastal water classes was found in the N=8 classifications produced here, and none with N=5.

[FIGURE 4 HERE, SINGLE-COLUMN; COLOUR ONLINE ONLY]

[TABLE 3 HERE]

As a compromise between F and S metrics and representation of the JSM classes, and to enable direct comparison of results against similar studies, the 5- and 8-class schemes (henceforth N5 and N8) produced using m=1.05, 42,000 training pixels and absolute Rrs spectra were used for further analysis, together with the 6-waveband JSM classification. These Level 2 and 3 classifications for 2003 are summarized in Table 3 and the class spectra are given in Table 4.

[TABLE 4 HERE]

Mapping of the N5 and N8 classifications from Level 2 and Level 3 data to abundance-based and reflectance-based PFT algorithms is shown in Figure 5. PFT algorithms consistently interpreted the lower-OC3-CI chl classes as being dominated by picoplankton, with the contribution of nano- and microplankton fractions increasing with increasing chl. Eutrophic-type distributions, dominated by microplankton, were only produced in the N8 classification, but this was the only difference between 5- and 8-class schemes. Since the proportion of pixels assigned to the 8th class was 0.3% (Table 3), the abundance-based PFT algorithms generally classify these optical classes as dominated by small cells. In contrast, switching from Level 2 to Level 3 data produced more PHYSAT-type spectra that fell into the pseudo-diatom class (similar spectral shape to PHYSAT-diatom, but higher Rrs values). The N5 and N8 schemes correspond to JSM classes between 1 and 9, and

Class	Level 2 % pixels assigned to each class			Level 3 % pixels assigned to each class		
	N5	N8	JSM	N5	N8	JSM
1		5.7	1.9			0.5
2	10.7		6.7	11.9	8.1	5.6
3		14.8	14.6		14.0	13.2
4	24.7	19.7	18.0	20.9	16.3	17.2
5	28.0	20.9	16.6	24.7	17.7	17.2
6		18.5	13.4		14.8	12.9
7	23.9		9.7	19.7	10.9	8.3
8		14.0	7.7			6.0
9	12.8		4.6	12.0	6.0	3.5
10		6.3	3.6			2.3
11			2.6			0.8
12			0.3			0.3
13		0.3	0.2		0.3	0.2
14			0.1			0.1

Table 3: Total number of Level 2 and Level 3 dominant classes mapped to each fuzzy cluster in the 10-waveband classifications (5 classes and 8 classes, N5, N8) and in the Jackson-Moore 6-waveband classification. The 10-waveband classes are mapped to the JSM classification by Euclidean distance (Figure 4).

Classification / class no.	Rrs(412) x1000	Rrs(443) x1000	Rrs(469) x1000	Rrs(488) x1000	Rrs(531) x1000	Rrs(547) x1000	Rrs(555) x1000	Rrs(645) x1000	Rrs(667) x1000	Rrs(678) x1000
N5: 1	15.032	11.065	9.366	7.298	2.621	1.929	1.535	0.009	0.204	0.256
N5:2	12.186	9.035	7.988	6.455	2.523	1.857	1.489	0.076	0.197	0.248
N5:3	9.620	7.332	6.782	5.672	2.447	1.814	1.458	0.059	0.191	0.247
N5:4	6.974	5.602	5.482	4.787	2.388	1.807	1.468	0.058	0.196	0.262
N5:5	3.961	3.476	3.748	3.480	2.215	1.768	1.474	0.085	0.224	0.324
N8:1	15.530	11.400	9.556	7.381	2.558	1.861	1.467	0.077	0.183	0.236
N8:2	13.131	9.683	8.426	6.718	2.519	1.842	1.469	0.070	0.189	0.241
N8:3	11.193	8.355	7.507	6.138	2.462	1.808	1.445	0.057	0.184	0.236
N8:4	9.375	7.166	6.656	5.583	2.418	1.787	1.433	0.048	0.183	0.241
N8:5	7.500	5.948	5.743	4.965	2.375	1.778	1.435	0.045	0.186	0.250
N8:6	5.514	4.593	4.680	4.196	2.313	1.779	1.457	0.054	0.195	0.271
N8:7	3.184	2.913	3.268	3.103	2.146	1.746	1.467	0.096	0.236	0.348
N8:8	7.662	9.253	11.578	12.541	14.101	14.027	13.220	4.781	3.947	3.880

Table 4: Class Rrs spectra for the N5 and N8 classifications produced in this study using Level 3 data.

with a finer gradation of Rrs spectra corresponding to 0.03 to 0.5 mgm^{-3}, the JSM classes include more scope for identifying mixed size-class waters with the PFT algorithms included here. The PHYSAT classification was designed to identify cases in which a single PFT dominates water colour and, correspondingly, not all classes could be mapped to a PHYSAT class (Figure 5). Of those that did, Synechococcus-like cyanobacteria was the most common designation (3 classes at Level 2, N5; 4 classes at Level 2, N8; 1 class at Level 3, N5 and JSM and 2 classes at Level 3, N8).

[FIGURE 5 HERE: SINGLE-COLUMN; ONLINE COLOUR ONLY]

Class maps for every second month of 2003 (Figure 6) show more fragmented spatial distribution of classes in the L2 data, corresponding to the higher spatial resolution of the data and higher spectral separability. Broader spatial patterns (O1000 km) are consistent between all classifications and include a limited seasonal north to south shift in class, distinction of coastal, Somali Current and Arabian Sea waters from the central domain and coherent and elongated patches of higher class (4-5 in the N5 classification) waters in the central Indian Ocean, often enveloping the Chagos MPA. These zonal bands are more evident in Figure 7, which summarises class diversity at each pixel, as corresponding to the seasonally-varying SCTR, SEC and SECC domains. Figure 8 shows a subset of the remote sensing climatologies, for comparison.

[FIGURE 6, 7, 8 CLOSE TO HERE: SINGLE-COLUMN or 1.5 COLUMN; PLEASE QUOTE FOR COLOUR PRINTING]

Whether or not the smaller features correspond to ecologically meaningful variations in the microbial biome or carbon cycling can only be answered definitively with *in situ* data. The microscopy stations lay within classes 3 and 4, and a significant difference in phytoplankton biovolume was confirmed (class 3 mean biovolume = 8.1x10^5 $\pm$ 2.6x10^5 $\mu m^3 l^{-1}$, N=3; class 4 mean biovolume = 2.3x10^6 $\pm$ 8.4x10^5 $\mu m^3 l^{-1}$, N=9; Kruskal Wallis p = 0.0126). There was also an order of magnitude difference in the ratio between phytoplankton and zooplankton biovolume (Schwarz, 2020). In the absence of a larger *in situ* dataset with which to evaluate the full classification, comparison between these classifications and previous studies is helpful. In spatial and temporal variability, these optical classifications are most similar to previous studies that use Rrs or radiance, as expected, and to some of the mesopelagic biogeographies. The degree of patchiness is consistent with examples given by JSM, and the classes assigned to the SCTR in their example of July 2004 are in direct agreement with the N5 classification produced here (JSM classes 7/8, based on 6 wavebands, correspond to N5 class 4, based on 10 wavebands; Table 3). A similar degree of patchiness is reported by Oliver & Irwin (2008) using nLw412, nLw551 and SST, and their approach, allowing the number of classes to emerge from the data, assigned up to ten classes over the TIO, which supports the richer N8 or JSM (N=14) classifications tested here in terms of spectral separability. George et al. (2013) reported multiple patches of elevated chl extending some 200 km along 67^o E in the SCTR, as well as suppression of surface chl by eddies further south. In contrast, classifications that used chl as the only ocean colour variable, together with other physical drivers, have been less spatially diverse (e.g. Longhurst, 2007; Spalding et al., 2012; Reygondeau et al., 2013; Fay & McKinley, 2014; Sayre et al., 2017, surface zone) and do not distinguish the SEC or SCTR domains clearly.

Mesopelagic classifications using a range of approaches including derived ocean colour variables, acoustic data, World Ocean Atlas data and species abundance mostly do distinguish the SCTR and SEC zones (Proud et al., 2017; Sutton et al., 2017; Sutton & Beckley, 2017; Sayre et al., 2017, 200 to 800 m zones) although in some cases the distinction between coastal influences and SEC/SCTR features is unclear (Costello et al., 2017; Reygondeau et al., 2018). Differences in the spatial richness of mixed-input epipelagic classifications relate partly to the scales and methods used, but may also reflect subtle changes in the growth environment that are related to phytoplankton community composition that are not detected in the chl algorithms, or to chl variability being outweighed in a classification by the contribution of SST, producing spatially coarser structures because of dynamics that have no surface signature in SST (e.g. Santoso et al., 2010; Drushka et al., 2012; Strutton et al., 2015), or both. Previous analysis of phytoplankton bloom dynamics in the TIO, based on satellite-derived chl and biogeochemical modelling, suggested a summer bloom spanning the full breadth of the basin at SEC latitudes, but no winter bloom (Levy et al., 2007), and the modelled emergent biogeography of Follows & Dutkiewicz (2011) predicted a band of Prochlorococcus analogs and low species richness in this region which is consistent with the PFT algorithm interpretation of classes 1 to 3 with N5. However, Wiggert et al. (2006) predict larger phytoplankton cells between January and May in the deep chlorophyll maximum in the SCTR and Jeffries et al. (2015) found eukaryotes contributed >10% to relative cell abundance at a deep water site within the Chagos MPA. Similarly, Thorrington-Smith (1971) found diatom and dinoflagellite communities in water samples from 100 m depth across the western TIO - a signal that is consistent with the higher-chl PFT interpretations of classes 5 (N5) and 6-8 (N8) which could be expected to be detected in satellite data in zones of strong vertical mixing, such as the tropical gyre boundaries.

The loss, at L3, of classes representing high and very low chl values, may be important for monitoring carbon export and ecosystem resources and Duarte et al. (2013) suggested that a chlorophyll concentration of 0.44 mgm^{-3} represents a transition between heterotrophic and autotrophic communities. Application of the Duarte et al. (2013) threshold to L3 chl values in this area is consistent with Fernandes et al. (2008), who reported net heterotrophic production between 1^o N and 5^o S at 83^o E, but the appearance of higher ranges in L2 data, the prevalence of higher classes for most of the year and the paucity of *in situ* data for the pelagic MPA domain area renders this use of the classification results uncertain. Level 2 data classifications are therefore potentially valuable for modelling and monitoring tasks (Tweddle et al., 2018), despite the low daily spatial coverage. For the purposes of exploring spatiotemporal variability over the MPA and wider TIO domain, Level 3 classifications are explored further.

3.2. Are biological and physical variables distinct for each optical class ?

The class-specific ranges of remote sensing variables are summarised for the MPA and z1000 domains in Figure 9 and Appendix 1. The N5 optical classification constrained all variables most effectively. For the MPA, only sea surface height (classes 4 and 5), eddy kinetic energy (classes 3 and 5) and the R443:R555 reflectance ratio (classes 1 and 5) were ambiguous (Kruskal-Wallis with post-hoc HSD tests, $p<0.01$, $N>10000$). Increasing class number in the MPA was associated with increasing chl, FLH, ∇SSH, u, τ_N and depth, and with decreasing FLH:chl ratio, SST, SSH, v and τ_E. Westerly and northerly currents, and westerly wind stress, associated with SWM wind reversals, were associated with class 1 only (lowest chl).

[FIGURE 9 HERE; SINGLE-COLUMN; COLOUR ONLINE ONLY]

Greater between-class ambiguity in optical and physical variables was found with N8 and JSM, both of which included sparse classes ($N<10000$). For N8, class 8, representing OC3M-CI chl = 2.39 mgm^{-3}, was always sparse and class 7 (chl = 0.46 mgm^{-3}) was sparse in the SCTR-E and -C domains. Similarly, JSM classes 10 to 14, representing chl > 0.62 mgm^{-3}, were always sparse or empty, and class 9 (chl = 0.47 mgm^{-3}) was sparse in all the sub-domains except z1000 and SCTR-W. Between-class variability was not significant in the MPA for the majority of physical variables in N8 and JSM (SSH, EKE, u, v, τ, $\nabla \times \tau$) or for FLH, but chl and the residual reflectance ratios were significantly different in all 8 classes. This could be interpreted as a smaller number of physically-distinct conditions hosting a larger number of optically distinct conditions, consistent with growth, decay and succession occuring within each physical 'province' over the averaging period of 1 month. In JSM, only 8 classes were well-populated (No. pixels > 10000) and all variables except $\nabla \times \tau$ were ambiguous for two or more classes. In the wider z1000 domain, between-class variability was significant for most variables in each of the classifications (exceptions were N5: SST for classes 2 and 3; N8: ∇SSH for classes 4 and 7, u for classes 6 and 7; JSM: SST for classes 1 and 8, v for classes 7 and 8 and τ_E for classes 1 and 2), reflecting the much greater size of this dataset. In this wider domain, the tendency of windstress variables with increasing class number was reversed so that increasing class and chl were associated with increasing τ_E and decreasing τ_N.

All between-class test results are given in Supplementary Online Material Table S1, sheet 'Variable Ranges by Class'. The optical classifications do correspond to distinct ranges of biotic and abiotic variables, suggesting their potential value in providing a useful diagnostic for management and modelling applications. Five optical classes produces least ambiguity in physical variables, although significant residual ocean colour differences are detected in up to 7 classes. However, the biophysical relationships vary within the wider domain, as may be expected from the known oceanographic processes in the region, suggesting that the use of optical classes may be most appropriate within a heirarchical scheme (c.f. Kavanaugh et al., 2014; Oliver et al., 2004; Oliver & Irwin, 2008). Exploration of between-class variability within different sub-domains is addressed in the next section.

To test whether seasonal variability in winds and associated mixing and entrainment (Halkides & Lee, 2011; Wiggert et al., 2006) produces between-class ambiguity, Kruskal-Wallis tests were applied to the MPA data for each month (Online Supplementary Figures S2, S3; Tables S2, S3). Seasonal wind stress reversals were detected with the annual average conditions (negative τ_E decreasing with increasing class number) found during austral winter and the opposite trend (positive τ_E increasing with increasing class number) during austral summer (Figure S2).The winter months were significantly different at α=0.01 for classes 3, 4 and 5 (higher chl), whereas austral summer months were distinct in classes 2 and 4, and τ_N trends generally mirrored τ_E trends. Less pronounced seasonal reversals were found for u, v and ∇SSH. One heirarchical partitioning approach could therefore be to use the optical classifications with a monthly or seasonal interpretation to constrain variability in the epipelagic growth environment, but a more objective approach using the physical variables at higher levels of the heirarchy avoids the need to assume a regular seasonal cycle. In either case, the correspondence between optical class and biotic environmental conditions needs to be explored using *in situ* biogeochemical data if the optical classification is to be used to deduce ecological function.

3.3. Do optical classifications capture zonal or meridional differences around the Chagos MPA?

Between-class variability in remotely sensed variables for each of the subdomains is shown in Figure 10. Most remotely sensed variables tended to increase or decrease monotonically with increasing optical class number, and spatially segmenting the dataset had little effect on these tendencies. Of the three SCTR domains, the westernmost area had the lowest SST values in all classes, as well as lower SST and higher τ_N in the highest-chl class 5. The central TIO domain, SCTR-C, which is the easternmost of the three, had the highest SST and EKE values and much lower nFLH:chl values in all classes. Stronger contrasts were evident between the Wiggert-N and -S domains. nFLH, v, τ_N and nFLH:chl were lower in the northern than the southern domain in some or all classes, whilst SST, ∇SSH, EKE, u and τ_E were higher. Whereas the depth-resolved modelling studies of George et al. (2018) and Dilmahamod et al. (2016) suggested east-west gradients in biophysical mechanisms operating across the SCTR, sub-domains SCTR-E,-W and -C could not be distinguished in the surface remote sensing variables studied here (PERMANOVA, $p > 0.1$, N $>$ 691x100; Table **??**), although the optical classes were significantly different in all domains and classifications ($p < 0.005$, N $>$ 692x100). In contrast, differences were detected between classes and domains for the north-south division discussed by Wiggert et al. (2006) ($p < 0.037$, N $>$ 692; Table 5). Excluding the shelf slope depths between 200 and 1000 m had no impact on class ambiguities in any variable (Figure 10).

[FIGURE 10 HERE; SINGLE-COLUMN; ONLINE COLOUR ONLY]

[TABLE 5 HERE]

Correlations between physical and biological variables within each class and domain are shown in Figure 8. Chl was negatively correlated with SST in all but the z200/z1000 domains. In the MPA, increasing SST was associated with lower EKE and τ_N and with higher v, τ_E and $\nabla \times \tau$ in the lowest-chl class (class 1) only, with the reverse relationships found in classes 2 to 5. The lower chl class 1 is therefore associated with conditions in which τ_E and $\nabla \times \tau$

Table 5: Results of PERMANOVA tests for differences between regions. a) Fixed factor 1 is location in the western or eastern SCTR or adjacent central Indian Ocean domain; nested factor 2 is optical class. All 15 remotely sensed variables were included in the test as z-scores. Average values are from 100 random draws of 1000 data points from each domain, time-matched. Pixels with any missing data were excluded, yielding residual degrees of freedom between 600 and 800 (denoted 'Resid'). p-values were calculated using 999 permutations. b) Fixed factor 1 is location in the northern or southern central Indian Ocean sectors; other details as for a).

a) Seychelles-Chagos Thermocline Ridge W / E / C (George et al., 2018)					
	Factor	df	$\overline{F}$	$\overline{p}$	p range
N5	F1: Domain	2	1.096	0.406	0.104 to 0.847
	F2: Class	10.9 (10-12)	15.066	0.0012	0.001 to 0.006
	Resid:	[691-766]			
N8	F1: Domain	2	1.256	0.271	0.048 to 0.700
	F2: Class	15.8 (14-18)	13.785	0.001	0.001 to 0.003
	Resid:	[703-765]			
JSM	F1: Domain	2	1.515	0.142	0.015 to 0.424
	F2: Class	18.3 (16-22)	12.825	0.001	0.001 to 0.001
	Resid:	[688-771]			

b) Central tropical Indian Ocean N / S (Wiggert et al., 2006)					
	Factor	df	$\overline{F}$	$\overline{p}$	p range
N5	F1: Domain	1	3.420	0.0093	0.001 to 0.037
	F2: Class	7.4 (6-8)	20.871	0.0011	0.001 to 0.005
	Resid:	[692-755]			
N8	F1: Domain	1	4.030	0.003	0.001 to 0.014
	F2: Class	10.2 (9-11)	18.127	0.001	0.001 to 0.001
	Resid:	[708-775]			
JSM	F1: Domain	1	4.415	0.0023	0.001 to 0.013
	F2: Class	12.1 (10-15)	16.044	0.001	0.001 to 0.001
	Resid:	[692-755]			

do not correspond to surface cooling and the northward geostrophic current component is positive. Chl and nFLH were positively correlated everywhere except in class 1 (lowest chl) in the MPA, SCTR-C and Wiggert-N,S areas. Higher nFLH:chl ratios were associated with lower τ_E, $\nabla \times \tau$, u, EKE, ∇SSH, SSH and SST in the MPA class 5, and to a lesser extent for classes 3 and 4, suggesting the importance of the tropical gyre strength for surface chl. These relationships are similar, but less pronounced, in the SCTR-W, Wiggert-N and Wiggert-S domains, and in z1000 classes 3 and 4. Assuming relative homogeneity of the phytoplankton community within a given optical class, nFLH:chl can be interpreted as a proxy for relatively high iron limitation (as opposed to other limiting factors, Behrenfeld et al., 2009). Although Chinni et al. (2019) and Wiggert et al. (2006) suggest Fe-limitation in some seasons within and north of the SCTR, George et al. (2013) reported that the deep chlorophyll maximum (DCM) followed the nitricline and did not measure iron concentrations, so the interpretation of nFLH:chl requires more *in situ* data in this region. SLA, which is related to westward-propagating, downwelling Rossby waves in the SCTR (George et al., 2018), decreased with increasing optical class and was negatively correlated with chl in at least 4 classes in all domains, including z1000 (Figure 11). The use of class-specific correlations across different sub-domains captured other contrasts in physical relationships, such as a switch from positive to negative coupling between u and EKE in MPA class 1, SCTR-W classes 1,2,3 and 5; SCTR-C classes 1 and 2; Wiggert-N class 1, whilst the relationship was negative for all classes in Wiggert-S, suggesting dominance of the westward SEC in the southern domain and more varied interactions in the northern SCTR (e.g.

mesoscale and fine-scale processes, George et al., 2013; Hosegood et al., 2019). For all domains, the greatest between-class differences were found between classes [1,2] and classes [3,4,5], suggesting that the different physical processes driving the growth environment can be distinguished by the chl 0.08 mgm^{3} isoline (Figure 5). Biophysical coupling in the Chagos MPA was most similar to that in the Wiggert-S domain (Figure 10; Supplementary Material Table 1, sheet 'Variable Ranges by Class').

[FIGURE 11 HERE; SINGLE-COLUMN; COLOUR ONLINE ONLY]

3.4. Modelling applications of optical partitioning

Assessment of the global carbon cycle and sequestration of anthropogenic CO_2 emissions underpins the Paris Agreement (UN, 2015). The oceanic biogeochemical models used for global carbon cycle assessment remain fairly simplistic and uncertainties are high (Lim & Kug, 2017; Le Quéré et al., 2013, 2018; Gruber et al., 2019); increasing the complexity of ecosystem dynamics in models without rendering them unstable is challenging (Anderson, 2005) and different approaches are still being developed (e.g. Hense et al., 2017; Wanninkhof et al., 2013). Optical classes offer an empirical contstraint on ecosystem models and provide a dynamic framework for aggregating model outputs and assessing model skill, for example in the prediction of CO_2 uptake or sequestration rates. The preliminary comparison of pCO_2 between optical classes supports both of these applications (Figure 12). Between-class differences in surface partial pressure of CO_2 were significant ($p < 0.01$, $N > 13{,}000$) for all N5 classes. pCO_2 distributions were mostly bi-modal, reflecting a background latitudinal gradient in class (increasing class number, reflecting increasing chl, to the north) with patches of higher classes to the south. The average pCO_2 value decreased slightly with increasing class number (416±49, 418±48, 404±44, 406±48, 398±48 ppm), in contrast with Nagelkerken et al. (2015) who reported no simple relationship between primary production and CO_2 uptake in this area. Between-class differences in this study were within the version 8 OCO-2 model error for X_{CO_2}, but uncertainties in the profile retrieval are not specified (O'Dell et al., 2018). Because of these uncertainties, the scope of this comparison and because too few data were available to test between-area differences, a more complete comparison is reserved for a future study.

[FIGURE 12 HERE; TWO-COLUMN; COLOUR ONLINE ONLY]

In a fully-realised fuzzy optical classification, multiple colour class memberships enable mixed phytoplankton communities (or mixed water types) to be represented and this has been shown to be useful in enhancing the interpretation of ocean colour data as well as constraining chl algorithm errors (e.g. Moore et al., 2001, 2009). Better remote sensing information about the ocean surface microbial community feeds into the Conservation on Biodiversity as well as monitoring and understanding the ecosystem services they provide (CBD, 2010; Tweddle et al., 2018; Roberts et al., 2017). In this study, the Rrs class residuals did not show promise for identifying details of phytoplankton community composition, but ensemble class biophysical agreement suggests that ecosystem function may be constrained by optical class.

3.5. Management and conservation applications of optical partitioning

Information is lost when Rrs spectra are partitioned into discrete classes, rather than applying an algorithm to produce a continuous biological variable such as chl, but a distinct advantage of this is the possibility of using habitat fragmentation metrics on the class maps. Fragmentation metrics applied to the N5 class maps are shown in Figure 13 for the MPA and the Wiggert-N and -S domains. The lower-chl classes 1 and 2 behaved differently in the three domains: Class 1 patches were of order 5000 km^2 and generally largest to the south, where they were separated by ca. 200 km. In the MPA, between-patch distances were shorter and stable through time, whereas they fluctuated between 20 and 100 km to the south. The 2011 La Nina period coincided with particularly high separation distances in Wiggert-N (separation ca. 500 km) and with high patch sizes in Wiggert-S (areas up to ca. 15 000 km^2; Figure 13). In contrast, class 2 patches were larger overall (O20 000 km^2) with higher values to the north, and a maximum coinciding with the 2015 El Nino (mean patch size ca. 50 000 km^2). Class 2 patch separation distances fluctuated in Wiggert-N as for class 1 but with a lower range (20 to 70 km). The 2015 El Nino coincided with the highest class 2 patch sizes in Wiggert-N (up to ca. 50 000 km^2). Classes 1 and 2 represent clear, warm water, which is a foraging habitat used by seabirds preying on flying fish and squid, often in association with subsurface predators (e.g. Weimerskirch et al., 2005; Catry et al., 2009b; Le Corre et al., 2012). The absence of seasonal cycles in the fragmentation metrics for these classes is marked: Prey occurence, driven by cetaceans and tuna, is stochastic, but patches of similar foraging conditions are predictable at the monthly scales used here, with patch separations that are within the known range of some seabirds (Weimerskirch et al., 2007; Nel et al., 2001; Pinaud & Weimerskirch, 2007).

[FIGURE 13 HERE: SIGNLE-COLUMN; BLACK AND WHITE]

Patch sizes were smaller in the higher-chl classes 3 and 4 (mostly within 5000 km^2), with separation distances between 20 and 200 km and marked seasonality. For classes 1 to 4 the MPA clearly straddles the Wiggert-N and Wiggert-S conditions, potentially providing stability of resource within foraging range in the event of extreme (Nino/Nina/IOD) conditions. The highest-chl class 5 was not always present (Figure 13e, j, o) and represents the smallest but most intense resource patches (10 to 300 km^2 in size), separated by 20 to 100 km in the MPA, 40 to 300 km in Wiggert-S and 20 to 1000 km in Wiggert-N. Scott et al. (2010) and Trevail et al. (2019) highlight the importance of fine spatial and temporal scales in prey resource; Level 2 (ca. daily, 1 km) or higher spatial resolution data are therefore also of interest.

Class 5 patches were largest within the MPA domain up to 2014, after which increasingly large class 5 patches appear in Wiggert-N (Figure 13j). However, the time-series is too short to confirm whether this is a robust trend. Significant temporal trends in fragmentation metrics were only found for fragmentation distance in the MPA, where the average distance between class 2 patches increased over the study period, whilst the distance between class 3 patches decreased ($p < 0.05$, $N = 204$). There is evidence of a shift in the spatial distribution of resources over the study period, but without loss of any of the colour classes, suggesting that the range of niches that seabirds exploit has been maintained across the Wiggert-N and -S domains (Waugh & Weimerskirch, 2003; Catry et al., 2009a; Le Corre et al., 2012). In this study, water depths shallower than 200 m were excluded to avoid land adjacency and bottom reflectance effects, so the MPA domain metrics do not include the near-shore and lagoon waters which may augment class 5.

Although mesopelagic biomes have been shown to reflect the spatial distribution of primary production (Proud et al., 2017), the only direct link between optical classes and large, commercially-fished species that could be expected is through water clarity for foraging. To test whether optical classes could provide useful fisheries management information, class maps were compared against fisheries records. The Chagos MPA has been a no-take zone since it was established in 2010, and IOTC fishing records amalgamated over the study period (2002 to 2018) are correspondingly lower in waters immediately adjacent to the MPA and increase further away (Table 6). However, 284 surface fishery records and 43 longline records were reported after 2010 in the IOTC gridcells that contain the MPA (-2.3 to -10.8^{o}S, 67.9 to 79.4^{o}E). Figure 14 shows the distribution of average effort, catch and CPUE at increasing distances from the MPA for the surface and longline fisheries. In the wider domain (z1000), surface fishery catch and CPUE mostly increased with increasing N5 class, as expected (e.g. Solanki et al., 2015, 2017; Mohamed et al., 2018), with significant between-class differences between low- and high-chl waters (Kruskal-Wallis, $p < 0.01$), and effort focussed in class 5 waters. Longline fishery effort and catch were highest in class 3, and although a trend for increasing CPUE with increasing class was apparent, it was not statistically significant. Few significant between-class differences were found in the MPA, but CPUE was highest in class 4, whereas effort was decreased from class 2 to 4 (Figure 14d, f). These patterns reflect reported catches, but may be used to infer behaviours of illegal fisheries and so target monitoring resources. This management application of the optical classification is easy to apply using L2 data, but is limited by cloud cover and ca. 1 day data

Table 6: Number of fishing records submitted to the IOTC within the 1^o and 5^o IOTC gridcells that contain the MPA, then incrementally as boxes of 3 x 3 and 5 x 5 of the 5^o IOTC gridcells are added around the MPA area, and finally for entire z1000 domain. LL=long-line fisheries; Surf=surface fisheries; + denotes an increment from the MPA number and Total is the number of records reported within the full z1000 domain.

N5 class	MPA		3x3 x5^o		5x5 x5^o		z1000		Total	
	LL	Surf	LL	Surf	LL	Surf	LL	Surf	LL	Surf
1	0	73	+15	+131	+84	+254	+792	+948	891	1406
2	14	588	+527	+1375	+650	+2821	+1691	+12078	3152	16862
3	23	668	+648	+1381	+1681	+4674	+3264	+20870	5616	27593
4	6	82	+63	+178	+701	+1665	+2614	+20746	3384	22671
5	0	0	0	0	+25	+100	+1266	+11457	1291	11557

latency.

[TABLE 6 ROUGHLY HERE]

Knowledge of fishing behaviours is also relevant to MPA design and, potentially, to the use of dynamic MPA designations. Dynamic protection boundaries serve as a compromise between static MPA boundaries, which protect relatively small areas of rare or valuable habitat (e.g. Oppel et al., 2018; Heerah et al., 2019; Handley et al., 2020; Williamson et al., 2019), and the very large foraging domains that are intractable to static protection but where seabirds, cetaceans and large predatory fishes such as tuna are known to colocate in the tropical Indian Ocean (Anderson, 2014; Letessier et al., 2017, 2019; Pinheiro et al., 2019; Hobday et al., 2010).

[FIGURE 14 ROUGHLY HERE: 1-COLUMN or 1.5-COLUMN; COLOUR ONLINE ONLY]

3.6. Are there class-specific temporal trends in biological/physical variables?

All conservation and management efforts must consider temporal variability and the possible impacts of climate change (IPCC, 2013). The 17-year study period used here is sufficient to characterise temporal variability but not to detect long-term change in this region (Henson et al., 2016). However, where biophysical conditions are coupled to large-scale climate indices for which robust, long-term time-series are available, the patterns observed can be extrapolated backwards and, with modelling, forward in time.

Time-series of selected variables are shown in Figure 15 for the MPA and Figure S4 for the Wiggert-S domain. A weak tendency for increased chl values in the austral winter was observed, but with no clear seasonal cycle either for the MPA as a whole or within a given class. This is consistent with removal of seasonality by class-switching (c.f. Figure 6). The class 4 and 5 chl values overlapped (Figure 15a), with class 4 representing an elevated background level of chl compared with classes 1 to 3, superimposed with stochastic, higher chl events in class 5, which often coincided with increases in nFLH. The strongest chl peak, in 2011, coincided with higher ∇SSH and a protracted period of positive SOI index (Nina conditions). Higher nFLH:Chl ratios and SSH values were evident for class 1 in both the MPA and Wiggert-S domains, as were lower SST values for class 5. Class 5 chl peaks in the MPA were not synchronised (or time-lagged, judging by visual inspection) with those in the Wiggert-S domain, suggesting small-scale, rather than basin-wide processes are being captured, despite the use of composited Level 3 data. Significant trends are not shown on Figures 15 and S4, for clarity, but are summarised in Figure 16a. Chl, nFLH and nFLH:Chl decreased over the study period in most classes and most sub-domains (when appraised using both absolute values and with anomaly time-series). For chl, the rate of change was between -7.5 x10^{-6} and -1.2 x10^{-4} mgm^{-3}a^{-1} (up to 0.025%a^{-1}

in the MPA and 0.055%a^{-1} in z1000, compared with decreases of 0.7%a^{-1} reported by Gregg et al., 2017, for the Indian Ocean).

Of the abiotic variables, significant trends in all five N5 classes were only found for the MPA, z1000 and Wiggert-N domains, with weaker positive trends in some areas (domain-integrated) and classes for SST, ∇SSH, u and v, and a positive trend in $\nabla \times \tau$ in Wiggert-S only. Full temporal trend results, including 95% confidence intervals on the rates of change, are given in Table S1, sheet 'Trends Correlations'.

[FIGURE 15 ROUGHLY HERE: SINGLE-COLUMN; PLEASE QUOTE FOR COLOUR PRINTING]

[FIGURE 16 ROUGHLY HERE: 1.5 OR SINGLE-COLUMN; PLEASE QUOTE FOR COLOUR PRINTING]

No significant correlations were found between the MJO and remotely sensed variables, consistent with the brief residence of MJO events over the TIO as well as the short time-scales of response of chl to MJO events which precludes detection of MJO effects in this Level 3 data analysis (order of days to weeks; Vialard et al., 2009; Jin et al., 2012; Wheeler & Hendon, 2004). Surface cooling and enhanced surface primary production have been documented in response to the MJO (Vialard et al., 2009; Resplandy et al., 2009, Supplementary material Figure S5), making the use of higher temporal resolution data desirable where coverage allows. The monthly-averaged MJO index was not correlated with the IOD or SOI indices over the study period. However, a weak, negative correlation was found between the IOD and SOI (SOI = -0.39 IOD + 0.27; n=202, p = 0.084), in contrast to the decoupling of these cycles found using EOF analysis by Saji et al. (1999). Fragmentation metrics for N5 class 3 were related to the SOI (negative relationship for patch density; positive relationship for patch area) and IOD (positive relationship for patch density only). No relationships between fragmentation metrics and any climate index were found for the Wiggert-S domain, whereas for Wiggert-N, patch density increased with increasing IOD for classes 1 and 3, and decreased with increasing IOD for classes 4 and 5. Similarly, patch area increased with increasing IOD for classes 1 and 2 and decreased for classes 3 and 4. Fewer significant relationships were found with the SOI, but they mirrored the IOD relationships, consistent with a negative relationship between the two climate indices. These results suggest that within the MPA, if the frequency of Nino events increases as predicted, incurring more negative SOI conditions, the higher chl N5 classes 3 and 4 will yield to larger and more closely spaced patches of lower N5 class 1. If positive IOD events increase in frequency, fewer, smaller N5 class 4 and 5 patches separated by greater distances are predicted. The observed North-South diversity in epipelagic conditions and temporal trends may contribute to resilience of reefs and mobile species in the MPA. However, only surface effects are characterised and, in this study, processes lasting days to weeks, such as MJO events and cyclones that have an impact on vertical mixing (Jin et al., 2012; Webster et al., 2005), may be averaged out.

Positive IOD events are associated with elevated SST in the western TIO as the easterly wind driving the SEC converges further west (Saji et al., 1999). Domain-specific correlations between IOD and SST (Table 16) were in agreement, with positive correlations in all but the SCTR-C (eastern-most) domain. Positive relationships were also found for SSH, τ_N and nFLH:chl, whilst negative relationships were found with Chl, FLH, R555:488, EKE and τ_E. Mostly positive, class-specific relationships only were found for ∇SST, suggesting stronger gradients at the sub-domain spatial scale which can not be explained by the regressions undertaken here as the spatially coarse altimetry products used to calculate ∇SSH and EKE, which might indicate small-scale processes, tended to decrease with increasing IOD. The SST product (4 x 4 km resolution) is more sensitive to mesoscale and perhaps submesoscale processes.

Negative impacts of Nino and (positive) IOD conditions on biotic remotely sensed variables is consistent with surface warming and deepening of the mixed layer in the western TIO, corresponding to down-mixing of phytoplankton within a strong, westward SEC current extending several hundreds of metres below the surface (Vialard et al., 2009). The response of elevated nFLH:chl ratios (Figure 15) under such a deeply-mixed layer could indicate nutrient stress (e.g. Fe, NO_3) or possibly a thin, surface freshening related to precipitation that is isolating a surface, light-stressed population (Behrenfeld et al., 2009; Chinni et al., 2019; George et al., 2013). The satellite data used for this study can not distinguish between these possibilities and are further limited by their short time-span (Dilmahamod et al., 2016; Landschützer et al., 2019), coarse spatial and temporal resolution (Hosegood et al., 2019; Vialard et al., 2009) and lack of information about depth variability. *In situ* data are needed to interpret many of the possible biophysical interactions in the MPA domain in terms of management application. For example, deepening of the mixed layer in the SCTR has been found to be associated with deepening of the DCM with an increase in chl owing to nutrient entrainment

(George et al., 2013); or with no net impact on water-column productivity owing to redistribution of light (Resplandy et al., 2009) or with a decrease in chl owing to reduced nutrient entrainment (Ma et al., 2014). Mesopelagic biogeographies suggest a strong positive relationship between satellite-derived primary production and zooplankton scattering layers (Proud et al., 2017), which is consistent with two of those scenarios. Depending on the response of grazers, a deepened DCM may have a protective effect on pelagic fish that are forced to forage further from the surface (Vialard et al., 2009) but a negative effect for surface-foraging seabirds, so that any future increase in the frequency of these events may have unexpected ramifications at higher trophic levels and for the Chagos MPA reefs (Graham et al., 2018; Fox et al., 2019).

4. Conclusions

Pure, optical classifications of unnormalised satellite remotely sensed reflectance data have been shown here to constrain physical variables that shape the epipelagic growth environment, making them a potentially useful source of management-relevant information at low- to medium latency. Constraint of most remotely sensed variables was best when used within spatial sub-domains such as the MPA area, suggesting their use within a seascape heirarchy. Potential applications discussed here include monitoring ecosystem services, including CO_2 uptake, and resource distribution, but in all cases *in situ* validation data are needed to elucidate optical biome composition and function.

Optical classification provides a snapshot at monthly (or better) time-scales of spatial variability of epipelagic resources that are amenable to habitat fragmentation analysis, which suggested a change in the spacing and size of richer surface food resources in response to Nino and IOD events in this study. However, ecological interpretation of the trends and interactions between remotely sensed variables requires knowledge of higher trophic level responses. Inclusion of fragmentation metrics in species distribution models might help to address this, where target species observations are sparse. Although ocean colour remote sensing is among the most finely resolved satellite products in space and time, it is limited by cloud cover and does not capture the full range of dynamical interactions that are relevant to habitat structure and use. Increased spatiotemporal coverage is needed and may be provided by combining information from all available sensors (though this is problematic) and by increased *in situ* monitoring using moorings and perhaps unmanned devices. The improved spectral resolution of NASAs forthcoming PACES mission may provide better discrimination of microbial community composition.

The suggested applications of optical classifications are globally applicable, but the need for more *in situ* data is not restricted to the tropical Indian Ocean. Data requirements include repeat vertical profiles (reflectance, temperature, salinity, vertical mixing, nutrient concentrations and carbon cycle parameters) and spatial fields of surface bio-optical and oceanographic conditions that resolve sub-pixel variability, so that appropriate spatial scales can be identified for a given question. This scope of fieldwork has been attempted in a few international, inter-disciplinary projects and it is to be hoped that more will take shape under the biogeoscapes programme (**?**).

Acknowledgements

I would like to thank Gillian Glegg, Ken Kingston and two anonymous reviewers for helpful comments on the manuscript.

References

Aguiar-Gonzalez, B., L. Ponsoni, H. Ridderinkhof, H. M. van Aken, W.P.M. de Ruijter, L.R.M. Maas, 2016. Seasonal variation of the South Indian tropical gyre. Deep-Sea Research I, 110, 123-140, doi:101016/j.dsr.2016.02.004.

Alvain, S., C. Moulin, Y. Dandanneau, F.M. Bréon, 2005. Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. Deep-Sea Research Part I 52, 1989-2004, doi:10.1016/j.dsr.2005.06.015.

Alvain, S., C. Moulin, Y. Dandanneau, H. Loisel, 2008. Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: A satellite view. Global Biogeochemical Cycles 22, GB3001, doi:10.1029/2007GB003154.

Amante, C. and B.W. Eakins, 2009. ETOPO1 1 arc-minute global relief model: Procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Centre, NOAA, doi:10.7289/V5C8276M.

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecology 26(1), 32-46, doi:10.1111/j.1442-9993.2001.01070.pp.x.

Anderson, T.R., 2005. Plankton functional type modelling: running before we can walk? Journal of Plankton Research 27(11), 1073-1081, doi:10.1093/plankt/fbi076.

Anderson, R. C., 2014. Cetaceans and Tuna Fisheries in the Western and Central Indian Ocean. IPNLF Technical Report 2, International Pole and Line Foundation, London, 133 pp.

Anderson, M.J., 2017. Permutational Multivariate Analysis of Variance (PERMANOVA). Wiley StatsRef: Statistics Reference Online, 2014-17, doi:10.1002/9781118445112.stats07841.

Ateweberhan, M., T.R. McClanahan, J. Maina, C. Sheppard, 2018. Thermal energy and stress properties as the main drivers of regional distribution of coral species richness in the Indian Ocean. Journal of Biogeography 45, 1355-1366, doi:10.1111/jbi.13224.

Back, L.E., C.S. Bretherton, 2009. On the relationship between SST gradients, boundary layer winds and convergence over the tropical oceans. Journal of Climate 22, 4182-4196, doi:10.1175/2009JCLI2392.1.

Behrenfeld, M.J., T.K. Westberry, E.S. Boss, R.T. O'Malley, D.A. Siegel et al., 2009. Satellite-detected fluorescence reveals global physiology of ocean phytoplankton. Biogeosciences 6, 779-794, doi:10.5194/bg-6-779-2009.

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, NY, doi:10.1007/978-1-4757-0450-1.

Blanco, G.S., J.P. Pisoni, F. Quintana, 2015. Characterization of the seascape used by juvenile and wintering adult Southern Giant Petrels from

Patagonia Argentina. Estuarine, Coastal and Shelf Science 153, 135–144, doi:10.1016/j.ecss.2014.12.007.

Boesch, H., L. Brown, R. Castano, M. Christi, B. Connor et al., 2019. Orbiting Carbon Observatory (OCO)-2 Level 2 Full Physics Algorithm Theoretical Basis Document, version 2.0, Rev. 3. National Aeronautics and Space Administration, Jet Propulsion Laboratory, OCO D-55207, 92pp.

Brewin, R.J.W., S. Sathyendranath, T. Hirata, S.J. Lavender, R. Barciela, N.J. Hardman-Mountford, 2010. A three-component model of phytoplankton size class for the Atlantic Ocean. Ecological Modelling 221, 1472-1483, doi:10.1016/j.ecolmodel.2010.02.014.

Brewin, R.J.W., E. Devred, S. Sathyendranath, S.J. Lavender, N.J. Hardman-Mountford, 2011. Model of phytoplankton absorption based on three size classes. Applied Optics 50(2), 4535-4549, doi:10.1364/AO.50.004535.

Brewin, R.J., N.J. Hardman-Mountford, S.J. Lavender, D.E. Raitsos, T. Hirata, J. Uitz, E. Devred, A. Bricaud, A. Ciotti, B. Gentili, 2011. An intercomparison of bio-optical techniques for detecting dominant phytoplankton size-class from satellite remote sensing. Remote Sensing of Environment 115(2), 325-339, doi:10.1016/j.rse.2010.09.004.

Brewin, R.J.W., T. Hirata, N.J. Hardman-Mountford, S. Lavender, S. Sathyendranath, R. Barlow, 2012. The influence of the Indian Ocean Dipole on interannual variations in phytoplankton size structure as revealed by Earth Observation. Deep-Sea Research Part II 77-80, 117-127, doi:10.1016/j.dsr2.2012.04.009.

Brown, C., Y. Huot, J. Werdell, B. Gentili, H. Claustre, 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(12), 4186-4203, doi:10.1016/j.rse.2008.06.008.

Cai, W., S. Agus, G. Wang, E. Weller, L. Wu et al., 2014. Increased frequency of extreme Indian Ocean Dipole events due to greenhouse warming. Nature 510, 254-258, doi:10.1038/nature13327.

Caldow, C., M.E. Monaco, S.J. Pittman, M.S. Kendall, T.L. Goedeke, C. Menza, B.P. Kinlan, B.M. Costa, 2015. Biogeographic assessments: A framework for information synthesis in marine spatial planning. Marine Policy 51, 423-432, doi:10.1016/j.marpol.2014.07.023.

Catry, T., J.A. Ramos, S. Jaquemet, L. Faulquier, M. Berlincourt, A. Hauselmann, P. Pinet, M. Le Corre, 2009. Comparative foraging ecology of a tropical seabird community of the Seychelles, western Indian Ocean. Marine Ecology Progress Series 374, 259 – 272, doi:10.3354/meps07713.

Catry, T., J.A. Ramos, M. Le Corre, R.A. Phillips, 2009. Movements, at-sea distribution and behaviour of a tropical pelagic seabird: the wedge-tailed shearwater in the western Indian Ocean. Marine Ecology Progress Series 391, 231 – 242, doi:10.3354/meps07717.

Chinni, V., S. Kumar Singh, R. Bhushan, V.V.S.S. Sarma, 2019. Spatial variability in dissolved iron concentrations in the marginal and open waters of the Indian Ocean. Marine Chemistry 208, 11-28, doi:10.1016/j.marchem.2018.11.007.

Convention on Biological Diversity 10, Decision X/2, 2010. Strategic plan for biodiversity 2011-2020 (Target 11, Strategic goal C).

Copernicus Marine Environment Monitoring Service (CMEMS), 2020. Quality information document for sea level TAC DU-ACS products. CMEMS-SL-QUID-008-032-062, 8th April, 2020, https://resources.marine.copernicus.eu/documents/QUID/CMEMS-SL-QUID-008-032-062.pdf.

Coelho, R., P. Infante, M.N. Santos, 2013. Application of generalized linear models and generalised estimation equations to model at-haulback mortality of blue sharks captured in a pelagic longline fisher in the Atlantic Ocean. Fisheries Research 145, 66–75, doi:10.1016/j.fishres.2013.02.010.

Costello, M.J., P. Tsai, P. Shan Wong, A. Kwok Lun Cheung, Z. Basher, C. Chaudhary, 2017. Marine biogeographic realms and species endemicity. Nature Communications 8:1057, doi:10.1038/s41467-017-01121-2.

Currie, J.C., M. Lengaigne, J. Vialard, D.M. Kaplan, O. Aumont et al., 2013. Indian Ocean Dipole and El Nino/Southern Oscillation impacts on regional chlorophyll anomalies in the Indian Ocean. Biogeosciences 10, 6677-6698, doi:10.5194/bg-1-/6677-2013.

Danckwerts, D.K., C.D. McQuaid, A. Jaeger, G.K. McGregor, R. Dwight, M. Le Corre, S. Jaquemet, 2014. Biomass consumption by breeding seabirds in the western Indian Ocean: Indirect interactions with fisheries and implications for management. ICES Journal of Marine Science 71(9), 2589-2598, doi:10.1093/icesjms/fsu093.

Defoin-Platel, M., M. Chami, 2007. How ambiguous is the inverse problem of ocean color in coastal waters? Journal of Geophysical Research: Oceans 112(C3), C03004, doi:10.1029/2006JC003847.

Devred, E., S. Sathyendranath, V. Stuart, T. Platt, 2011. A three component classification of phytoplankton absorption spectra: Applications to ocean colour data. Remote Sensing of Environment 115(9), 2255-2266, doi:10.1016/j.rse.2011.04.025.

Dierssen, H.M., 2010. Perspectives on empirical approaches for ocean color remote sensing of chlorophyll in a changing climate. Proceedings of the National Academy of Sciences 107(40), 17073-17078, doi:10.1073/pnas.0913800107.

Dilmahamod, A.F., J.C. Hermes, C.J.C. Reason, 2016. Chlorophyll-a variability in the Seychelles-Chagos Thermocline Ridge: Analysis of a coupled biophysical model. Journal of Marine Systems 154, 220-232, doi:10.1016/j.jmarsys.2015.10.011.

Drushka, K., J. Sprintall, S.T. Gille, S. Wijffels, 2012. In situ observations of Madden-Julian Oscillation mixed layer dynamics in the Indian and western Pacific Oceans. Journal of Climate 25, 2306-2328, doi:10.1175/JCLI-D-11-00203.1.

Duarte, C.M., A. Regaudie-de-Gioux, J.M. Arrieta, A. Delgado-Huertas, S. Agustí, 2013. The oligotrophic ocean is hetertrophic. Annual Review of Marine Science 5, 551-569, doi:10.1146/annurev-marine-121211-172337.

Erickson, Z. K., Frankenberg, C.,Thompson, D. R., Thompson, A. F., Gierach, M., 2019. Remote sensing ofchlorophyll fluorescence in the ocean using imaging spectrometry: Toward a vertical profile of fluorescence. Geophysical Research Letters 46, 1571–1579, doi:10.1029/2018GL081273.

Esteban, N., R.K.F. Unsworth, J.B.Q. Gourlay, G.C. Hays, 2018. The discovery of deep-water seagrass meadows in a pristine Indian Ocean wilderness revealed by tracking green turtles. Marine Pollution Bulletin 134, 99-105, doi:10.1016/j.marpolbul.2018.03.018.

ETOPO1 Global Relief Model, www.ngdc.noaa.gov/mgg/global doi:10.7289/V5C8276M.

EUMETSAT/OSI SAF. 2018. MetOp-A ASCAT Level 2 25-km Ocean Surface Wind Vector Climate Data Record. Ver. 1.0. PO.DAAC, CA, USA. Dataset accessed 2019-03 at https://doi.org/10.15770/EUM_SAF_OSI_0006.

Fay, A.R., G.A. McKinley, 2014. Global open-ocean biomes: mean and temporal variability. Earth System Science Data 6, 273-284, doi:10.5194/essd-6-273-2014.

Fernandes, V., V. Rodrigues, N. Ramaiah, J.T. Paul, 2008. Relevance of bacterioplankton abundance and production in the oligotrophic equatorial

Indian Ocean. Aquatic Ecology 42(4), 511-519, doi:10.1007/s10452-007-9142-y.

Follows, M.J., S. Dutkiewicz, 2011. Modeling diverse communities of marine microbes. Annual Review of Marine Science 3, 427-451, doi:10.1146/annurev- marine-120709-142848.

Fox, M.D., G.J. Williams, M.D. Johnson, V.Z. Radice, B.J. Zgliczynski et al., 2019. Gradients in primary production predic trophic strategies of mixotrophic corals across spatial scales. Current Biology (2018), doi:10.1016/j.cub.2018.08.057.

Franz, B.A., S.W. Bailey, P.J. Werdell, C.R. McClain, 2007. Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry. Applied Optics 46(22), 5068 - 5082, doi:10.1364/AO.46.005068.

Garcia, H.E., T. Boyer, O.K. Baranova, R.A. Locarnini, A.V. Mishonov et al., 2018. World Ocean Atlas 2018 (pre-release). Produce Documentation. A. Mishonov, Technical Editor. Available online at https://www.nodc.noaa.gov/OC5/woa18.

George, J.V., M. Nuncio, R. Chacko, N. Anilkumar, S.B. Noronha et al., 2013. Role of physical processes in chlorophyll distribution in the western tropical Indian Ocean. Journal of Marine Systems 113-114, 1-12, doi:10.1016/j.jmarsys.2012.12.001.

George, J.V., M. Nuncio, N. Anilkumar, R. Chacko, D. Rajashekhar, 2018.

Seasonal surface chlorophyll-a variability in the Seychelles-Chagos Thermocline Ridge. Current Science 114(4), 868-878.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, R. Moore, 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment 202, 18 - 27, doi:10.1016/j.rse.2017.06.031.

Graham, N.A.J., S.K. Wilson, P.Carr, A.S. Hoey, S. Jennings, M.A. MacNeil, 2018. Seabirds enhance coral reef productivity and functioning in the absence of invasive rats. Nature 559, 250-253, doi:10.1038/s41586-018-0202-03.

Gravestock, P., C. Sheppard, 2015. Valuing the ecosystem services of the Chagos: a review of challenges and estimates. Marine Ecology Progress Series 530, 255-270, doi:10.3354/meps11235.

Gregg, W.W., C.S. Rousseaux, B.A. Franz, 2017. Global trends in ocean phytoplankton: a new assessment using revised ocean colour data. Remote Sensing Letters 8(12), 1102-1111, doi:10.1080/2150704X.2017.1354263.

Gruber, N., D. Clement, B.R. Carter, R.A. Feely, S. van Heuven et al., 2019. The oceanic sink for anthropogenic CO2 from 1994 to 2007. Science 363:6432, 1193-1199, doi:10.1126/science.aau5153.

Guan, B., T. Lee, D.J. Halkides, D.E. Waliser, 2014. Aquarius surface salinity and the Madden-Julian Oscillation: The role of salinity in surface layer density and potential energy. Geophysical Research Letters 41, 2858-2869, doi:10.1002/2014GL059704.

Halkides, D.J., T. Lee, 2011. Mechanisms controlling seasonal mixed layer temperature and salinity in the Southwestern Tropical Indian Ocean. Dynamics of Atmospheres and Oceans 51, 77-93, doi:10.1016/j.dynatmoce.2011.03.002.

Hamed, K.H., A.R. Rao, 1998. A modified Mann-Kendall trend test for autocorrelated data. Journal of Hydrology 204, 182-196.

Handley, J.M., E.J. Pearmain, S. Oppel, A.P.B. Carneiro, C. Hazin et al., 2020. Evaluating the effectiveness of a large multi-use MPA in protecting key biodiversity areas for marine predators. Diversity and Distributions 26, 715 – 729, doi:10.1111/ddi.13041.

Heerah, K., M.P. Dias, K. Delord, S. Oppel, C. Barbraud, H. Weimerskirch, C.A. Bost, 2019. Important areas and conservation sites for a community of globally threatened marine predators of the Southern Indian Ocean. Biological Conservation 234, 192 – 201, doi:10.1016/j.biocon.2019.03.037.

Hendon, H.H., M.L. Salby, 1994. The life cycle of the Madden-Julian Oscillation. Journal of the Atmospheric Sciences 51(15), 2225-2237.

Hense, I., I. Stemmler, S. Sonntag, 2017. Ideas and perspectives: climate-relevant marine biologically driven mechanisms in Earth system models. Biogeosciences 14, 403-413, doi:10.5194/bg-14-403-2017.

Henson, S.A., C. Beaulieu, R. Lampitt, 2016. Observing climate change trends in ocean biogeochemistry: when and where. Global Change Biology 22, 1561-1571, doi:10.1111/gbc.13152.

Hermes, J.C., C.J.C. Reason, 2008. Annual cycle of the South Indian Ocean (Seychelles-Chagos) thermocline ridge in a regional ocean model. Journal of Geophysical Research 113, C04035, doi:10.1029/2007JC004363.

Hermes, J.C., C.J.C. Reason, 2009. The sensitivity of the Seychelles-Chagos thermocline ridge to large-scale wind anomalies. ICES Journal of Marine Science 66, 1455-1466, doi:10.1093/icesjms/fsp074.

Hermes, J.C., Y. Masumoto, L.M. Beal, M.K. Roxy, J. Vialard et al., 2019. A sustained ocean observing system in the Indian Ocean for climate related scientific knowledge and societal needs. Frontiers in Marine Science 6:355, doi:10.3389/fmars.2019.00355.

Hobday, A., J. Hartog, T. Timmiss, and J. Fielding 2010. Dynamic spatial zoning to manage southern bluefin tuna (Thunnus maccoyii) capture in a multispecies longline fishery. Fisheries Oceanography 19:243–253.

Hoge, F.E., P.E. Lyon, R.N. Swift, J.K. Yungel, M.R. Abbott et al., 2003. Validation of Terra-MODIS phytoplankton chlorophyll fluorescence line height. I. Initial airborne lidar results. Applied Optics 42(15), 2767-2771, doi:10.1364/AO.42.002767.

Hosegood, P.J., W.A.M. Nimmo-Smith, R. Proud, K. Adams, A.S. Brierley et al., 2019. Internal lee waves and baroclinic bores over a tropical seamount shark 'hot-spot'. Progress in Oceanography 172, 34-50, doi:10.1016/j.pocean.2019.01.010.

Houlbréque, F., C. Ferrier-Pagés, 2009. Heterotophy in tropical scleractinian corals. Biological Reviews 84, 1-17, doi:10.1111/j.1469-185X.2008.00058.x.

Hu, C., Z. Lee, B. Franz, 2012. Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. Journal of Geophysical Research 117(C1), doi:10.1029/2011jc007395.

Indian Ocean Tuna Commission, catch and effort data by month, species and gear, by vessel flag reporting country. Data accessed on 2020-06028 from http://iotc.org/datasets.

Indian Ocean Tuna Commision Secretariat, 2014.Guidelines for the reporting of fisheries statistics to the IOTC. IOTC Secretariat, Mahe, Seychelles, January 2014, 70pp.

IPCC, 2013. Climate change 2013: The physical science basis. Contribution of working group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschun, A. Nauels, Y. Xia, V. Bex, P.M. Midgley (eds)]. Cambridge University Press, Cambridge, U.K. and New York, NY, USA, 1535pp.

International Ocean Color Coordinating Group (IOCCG), 2014. Phytoplankton Functional Types from Space. Sathyendranath, S. (ed.), Reports of the International Ocean-Color Coordinating Group No. 15, IOCCG, Dartmouth, Canada.

International Ocean Color Coordinating Group (IOCCG), 2019. Uncertainties in Ocean Colour Remote Sensing. Mélin F. (ed.), IOCCG Report Series, No. 18, International Ocean Colour Coordinating Group, Dartmouth, Canada. http://dx.doi.org/10.25607/OBP696

Jackson, T., S. Sathyendranath, F. Melin, 2017. An improved optical classification scheme for the Ocean Colour Essential Climate Variable and its applications. Remote Sensing of Environment 203, 152-161.

Jeffries, T.C., M. Ostrowski, R.B. Williams, C. Xie, R.M. Jensen et al., 2015. Spatially extensive microbial biogeography of the Indian Ocean provides insights into the unique community structure of a pristine coral atoll. Scientific Reports 5:15383, doi:10.1038/srep15383.

Jerlov, N.G., 1976. Marine Optics. Elsevier Oceanography Series, 14. 2nd revised and enlarged edition of Optical Oceanography (Elsevier Oceanography Series, 5), Elsevier Scientific Publishing Company, Amsterdam, Oxford, New York.

Jin, D., R. Murtugudde, D.E. Waliser, 2012. Tropical Indo-Pacific Ocean chlorophyll response to MJO forcing. Journal of Geophysical Research: Oceans 117, C11008, doi:10.1029/2012JC008015.

Jin, D., D.E. Waliser, C. Jones, R. Murtugudde, 2012b. Modulation of tropical ocean surface chlorophyll by the Madden-Julian Oscillation. Climate Dynamics 41:1321, doi:10.1007/s00382-012-1321-4.

Jin, D., R.G. Murtugudde, D.E. Waliser, 2013. Intraseasonal atmospheric forcing effects on the mean state of ocean surface chlorophyll. Journal of Geophysical Research: Oceans 118, 184-196, doi:10.1029/2012JC008256.

Jones, D. L., 2012. The Fathom Toolbox for Matlab: multivariate ecological and oceanographic data analysis. College of Marine Science, University of South Florida, St. Petersburg, Florida, USA. Available from: https://github.com/stripathy/neuroelectro_matlab_analysis/tree/master/Fathom.

Jung, H.-C., B.-K. Moon, J. Wie, H.-S. Park, J. Lee, Y.-H. Byun, 2019. A single-column ocean biogeochemistry model (GOTM-TOPAZ) version 1.0. Geoscientific Model Development 12, 699-722, doi:10.5194/gmd-12/699-2019.

Kachelriess, D., M. Wegmann, M. Gollock, N. Pettorelli, 2014. The application of remote sensing for marine protected area management. Ecological Indicators 36, 169-177, doi:10.1016/j.ecolind.2013.07.003.

Kavanaugh, M.T., B. Hales, M. Saraceno, Y.H. Spitz, A.E. White, R.M. Letelier, 2014. Heirarchical and dynamic seascapes: A quantitative framework for scaling pelagic biogeochemistry and ecology. Progress in Oceanography 120, 291-304, doi:10.1016/j.pocean.2013.10.013.

Kavanaugh, M.T., M.J. Oliver, F.P. Chavez, R.M. Letelier, F.E. Muller-Karger, S.C. Doney, 2016. Seascapes as a new vernacular for pelagic ocean monitoring, management and conservation. ICES Journal of Marine Science 73(7), 1839-1850, doi:10.1093/icesjms/fsw086.

Kilidas, G.N., J. Dias, K.H. Straub, M.C. Wheeler, S.N. Tulich, K. Kikuchi, .M. Weickmann and M.J. Ventrice, 2014. A comparison of OLR and circulation based indices for tracking the MJO. Monthly Weather Review 142, May 2014, 1697-1715.

Kirk, J.T.O., 1994. Light and photosynthesis in aquatic ecosystems, 2nd Edition. Cambridge University Press, UK, ISBN: 0 521 45966 4, 509pp.

Koldeway, H.J., D. Curnick, S. Harding, L.R. Harrison, M. Gollock, 2010. Potential benefits to fisheries and biodiversity of the Chagos Archipelago/British Indian Ocean Territory as a no-take marine reserve. Marine Pollution Bulletin 60, 1906-1915, doi:10.1016/j.marpolbul.2010.10.002.

Krug, L.A., T. Platt, S. Sathyendranath, A.B. Barbosa, 2017. Ocean surface partitioning strategies using ocean colour remote sensing: A review. Progress in Oceanography 155, 41-53, doi:10.1016/j.pocean.2017.05.013.

Krug, L.A., T. Platt, A.B. Barbosa, 2018. Delineation of ocean surface provinces over a complex marine domain (off SW Iberia): An objective abiotic-based approach. Regional Studies in Marine Science 18, 80-96, doi:10.1016/j.rsma.2018.01.003.

Lain, L.R., S. Bernard, 2018. The fundamental contribution of phytoplankton spectral scattering to ocean colour: Implications for satellite detection of phytoplankton community structure. Applied Sciences 8, 2681, doi:10.3390/app8122681.

Landschützer, P., N. Gruber, D.C.E. Bakker, 2016. Decadal variations and trends of the global ocean carbon sink. Global Biogeochemical Cycles 30, 1396-1417, doi:10.1002/2015GB005359.

Landschützer, P., T. Ilyina, N.S. Lovenduski, 2019. Detecting regional modes of variability in observation-based surface pCO2. Geophysical Research Letters 46, 2670-2679, doi:10.1029/2018GL081756.

Le Corre, M., A. Jaeger, P. Pinet, M.A. Kappes, H. Weimerskirch, T. Catry et al., 2012. Tracking seabirds to identify potential Marine Protected Areas in the tropical western Indian Ocean. Biological Conservation 156, 83-93, doi:10.1016/j.biocon.2011.11.015.

Le Quere, C., R.J. Andres, T. Boden, T. Conway, R.A. Houghton et al., 2013. The global carbon budget 1959-2011. Earth System Science Data 5, 165-185, doi:10.5194/essd-5-165-2013.

Le Quere, C., R.M. Andrew, P. Friedlingstein, S. Sitch, J. Hauck et al., 2018. Global Carbon Budget 2018. Earth System Science Data 10, 2141-2194, doi:10.5194/essd-10-2141-2018.

Letessier, T.B., P.J. Bouchet, J.J. Meeuwig, 2017. Sampling mobile oceanic fishes and sharks: implications for fisheries and conservation planning. Biological Reviews 92, 627-646, doi:10.1111/brv.12246.

Letessier. T.B., D. Mouillot, P.J. Bouchet, L. Vigliola, M.C. Fernandes et al., 2019. Remote reefs and seamounts are the last refuges for marine predators across the Indo-Pacific. PLoS Biology 17(8):e3000366, doi:10.1371/journal.pbio.3000366.

Levy, M., D. Shankar, J.-M. Andre, S.S.C. Shenoi, F. Durand, C. Boyer de Montegut, 2007. Basin-wide seasonal evolution of the Indian Ocean's phytoplankton blooms. Journal of Geophysical Research: Oceans 112(C12), C12014, doi:10.1029/2007JC004090.

Lim, H.G., J.S. Kug, 2017. Impact of chlorophyll bias on the tropical Pacific mean climate in an earth system model. Climate Dynamics doi:10.1007/s00382-017-4036-8.

Longhurst, A.R., 2007. Ecological Geography of the Sea, 2nd Edition. Academic Press, Burlington, MA. ISBN 0-12-455521-7.

Lovenduski, N.S., S.G. Yeager, K. Lindsay, M.C. Long, 2019. Predicting near-term variability in ocean carbon uptake. Earth System Dynamics 10, 45-57, doi:10.5194/esd-10-45-2019.

Liu, L., L. Feng, W. Yu, H. Wang, Y. Liu, S. Sun, 2013. The distribution and variability of simulated chlorophyll concentration over the tropical Indian Ocean from five CMIP5 models. Journal of Ocean University of China 12(2), 253-259, doi:10.1007/s11802-013-2168-y.

Ma, J., Y. Du, H. Zhan, H. Liu, J. Wang, 2014. Influence of oceanic Rossby waves on phytoplankton production in the southern tropical Indian Ocean. Journal of Marine Systems 134, 12-19, doi:10.1016/j.marsys.2014.02.003.

Maldonaldo, M.T., A. Marchetti, M. Saito, A. Tagliabue, 2018. Biogeoscapes: Ocean Metabolism and Nutrient Cycles on a Changing Planet. Minutes of a working group at Woods Hole, November, 2018. Accessed online 2020-06 at https://drive.google.com/file/d/1EjIE4Fz2edd_hmusaaio-JHQQaDazGYn/view.

Maxwell, S.M., E.L. Hazen, R.L. Lewison, D.C. Dunn, H. Bailey et al., 2015. Dynamic ocean management: Definining and conceptualizing real-time management of the ocean. Marine Policy 58, 42-50, doi:10.1016/j.marpol.2015.03.014.

McCune, B., J.B. Grace, 2002. Analysis of Ecological Communities. MjM Software, Gleneden Beach, Oregon, USA. ISBN: 0-9721290-0-6.

McPhaden, M.J., G.R. Foltz, 2013. Intraseasonal variations in the surface layer heat balance of the central equatorial Indian Ocean: The importance of zonal advection and vertical mixing. Geophysical Research Letters 40, 2737-2741, doi:10.1002/grl.50536.

Miloslavich, P., N.J. Bax, S.E. Simmons, E. Klein, W. Appeltans et al., 2018. Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. Global Change Biology 24(6), 2416-2433, doi:10.1111/gcb.14108.

Mohamed, K.S., K.K. Sajikumar, N. Ragesh, T.V. Jayasankar, K.P. Said Koya, G. Sasikumar, 2018. Relating abundance of purpleback flying squid Sthenoteuthis oualaniensis (Cephalopoda: Ommastrephidae) to environmental parameters using GIS and GAM in southeastern Arabian Sea. Journal of Natural History 52(29-30), 1869 – 1882, doi:10.1080/00222933.2018.1497721.

Monolisha, S., T. Platt, S. Sathyendranath, J. Jayasankar, G. George, T. Jackson, 2018. Optical classification of the coastal waters of the northern Indian Ocean. Frontiers in Marine Science 5:87, doi:10.3389/fmars.2018.00087.

Moore, T.S., J.W. Campbell and H. Feng, 2001. A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. IEEE Transactions in Geoscience and Remote Sensing 39(8), 1764-1776.

Moore, T.S., J.W. Campbell, M.D. Dowell, 2009. A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. Remote Sensing of Environment 113(11), 2424-2430, doi:10.1016/j.rse.2009.07.016.

Nagelkerken, I., S.D. Connell, 2015. Global alteration of ocean ecosystem functioning due to increasing human CO2 emissions. Proceedings of the National Academy of Sciences 112(43), 13727 – 13277, doi:10.1073/pnas.151056112.

Nel, D.C., J.R.E. Lutjeharms, E.A. Pakhomov, I.J. Ansorge, P.G. Ryan, N.T.W. Klages, 2001. Exploitation of mesoscale oceanographic features by grey-headed albatross Thalassarche chrysostoma in the southern Indian Ocean. Marine Ecology Progress Series 217, 15-26, doi:10.3354/meps217015.

OCO-2 Science Team/M. Gunson, A. Eldering, 2016. ACOS GOSAT/TANSO-FTS Level 2 Full Physics Standard Product V7.3, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC). Dataset accessed 2019-04 at https://disc.gsfc.nasa.gov/datacollection/ACOS_L2S_7.3.html.

O'Dell, C.W., A. Eldering, P.O. Wennberg, D. Crisp, M.R. Gunson et al, 2018. Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm. Atmospheric Measurement Techniques 11, 6539-6576, doi:10.5194/amt-11-6539-2018.

Oliver, M.J., S. Glenn, J.T. Kohut, A.J. Irwin, O.M. Schofield, M.A. Moline, W.P. Bissett, 2004. Bioinformatic approaches for objective detection of water masses on continental shelves. Journal of Geophysical Research 109, C07S04, doi:10.1029/2003.JC002072.

Oliver, J.J., A.J. Irwin, 2008. Objective global ocean biogeographic provinces. Geophysical Research Letters 35, L15601, doi:10.1029/2008GL034238.

Oppel, S., M. Bolton, A.P.B. Carneiro, M.P. Dias, J.A. Green et al., 2018. Spatial scales of marine conservation management for breeding seabirds. Marine Policy 98, 37 – 46, doi:10.1016/j.marpol.2018.08.024.

O'Reilly, J.E., S. Maritorena, B.G. Mitchell, D.A. Siegel, K.L. Carder et al., 1998. Ocean color chlorophyll algorithms for SeaWiFS. Journal of Geophysical Research: Oceans 103(C11), 24937-24953, doi:10.1029/98JC02160.

Park, J.-Y., J.-S. Kug, 2014. Marine biological feedback associated with Indian Ocean Dipole in a coupled ocean/biogeochemical model. Climate Dynamics 42(1-2), 329-343, doi:10.1007/s00382-012-1640-5.

Pecoraro, C., I. Zudaire, N. Bodin, H. Murua, P. Taconet et al., 2017. Putting all the pieces together: integrating current knowledge of the biology, ecology, fisheries status, stock structure and management of yellowfin tuna (Thunnus albacares). Reviews in Fish Biology and Fisheries 27(4), 811-841, doi:10.1007/s11160-016-9460-z.

Pinaud, D., H. Weimerskirch, 2007. At-sea distribution and scale-dependent foraging behaviour of petrels and albatrosses: a comparative study. Journal of Animal Ecology 76, 9-19, doi:10.1111/j.1365-2656.2006.01186.x.

Pinheiro, H.T., J.B. Teixeira, R.B. Francini-Filho, A. Soares-Gomes, C.E.L. Ferreira, L.A. Rocha, 2019. Hope and doubt for the world's marine ecosystems. Perspectives in Ecology and Conservation 17, 19 – 25, doi:10.1016/j.pecon.2018.11.001.

Pillar, V.D., 2013. How accurate and powerful are randomization tests in multivariate analysis of variance? Community Ecology 14(2), 153-163, doi:10.1556/ComEc.14.2013.2.5.

Proud, R., M.J. Cox, A.S. Brierley, 2017. Biogeography of the global ocean's mesopelagic zone. Current Biology 27(1),113-119, doi:10.1016/j.cub.2016.11.003.

Racault, M.F., S. Sathyendranath, R.J. Brewin, D.E. Raitsos, T. Jackson, T. Platt, 2017. Impact of El Nino variability on oceanic phytoplankton. Frontiers in Marine Science 4:133, doi:10.3389/fmars.2017.00133.

Resplandy, L., J. Vialard, M. Levy, O. Aumont, Y. Dandonneau, 2009. Seasonal and intraseasonal biogeochemical variability in the thermocline ridge of the southern tropical Indian Ocean. Journal of Geophysical Research 114, C07024, doi:10.1029/2008JC005246.

Reygondeau, G., A. Longhurst, E. Martinez, G. Beaugrand, D. Antoine, O. Maury, 2013. Dynamic biogeographical provinces in the global ocean. Global Biogeochemical Cycles 27(1), 1046-1058, doi:10.1002/gbc.20089.

Reygondeau, G., L. Guidi, G. Beaugrand, S.A. Henson, P. Koubbi et al., 2018. Global biogeochemical provinces of the mesopelagic zone. Journal of Biogeography 45, 500-514, doi:10.1111/jbi.13149.

Rice, J., K.M. Gjerde, J. Ardron, S. Arico, I. Cresswell et al., 2011. Policy relevance of biogeographic classification for conservation and management of marine biodiversity beyond national jurisdiction, and the GOODS biogeographic classification. Ocean & Coastal Management 54, 110-122, doi:10.1016/j.ocecoaman.2010.10.010.

Roberson, L.A., E. Lagabrielle, A.T. Lombard, K. Sink, T. Livingstone et al., 2017. Pelagic bioregionalisation using open-access data for better planning of marine protected area networks. Ocean & Coastal Management 148, 214-230, doi:10.1016/j.ocecoaman.2017.08.017.

Roberts, C.M., B.C. O'Leary, D.J. McCauley, P.M. Cury, C.M. Duarte et al., 2017. Marine reserves can mitigate and promote adaptation to climate change. Proceedings of the National Academy of Sciences 114(24), 6167-6175, doi:10.1073/pnas.1701262114.

Romanou, A., J. Romanski, W.W. Gregg, 2014. Natural ocean carbon cycle sensitivity to parameterizations of the recycling in a climate model. Biogeosciences 11, 1137-1154, doi:10.5194/bg-11-1137-2014.

Ropelewski, C.F. and P.D. Jones, 1987. An extension of the Tahiti-Darwin Southern Oscillation Index. Monthly Weather Review 115, 2161-2165, doi:10.1175/1520-0493(1987)115¡2161:AEOTTS¿2.0.CO;2.

Ruxton, G.D., G. Beauchamp, 2008. Some suggestions about appropriate use of the Kruskal-Wallis test. Animal Behaviour 76, 1083-1087, doi:10.1016/j.anbehav.2008.04.01 1.

Saji, N.H., B.N. Goswami, P.N. Vinayachandran, T. Yamagata, 1999. A dipole mode in the tropical Indian Ocean. Nature 401, 360-363, 23 Sept. 1999.

Saji, N.H. and T. Yagamaga, 2003. Possible impacts of Indian Ocean Dipole mode events on global climate. Climate Research 25(2), 151-169.

Santoso, A., A. Sen Gupta, M.H. England, 2010. Genesis of Indian Ocean mixed layer temperature anomalies: A heat budget analysis. Journal of Climate 23(20), 5375-5403, doi:10.1175/2010JCLI2072.1.

Sayre, R.G., D.J. Wright, S.P. Breyer, K.A. Butler, K. Van Graafeiland et al., 2017. A three-dimensional mapping of the ocean based on environmental data. Oceanography 30(1), 90-103, doi:10.5670/oceanog.2017.116.

Scales, K.L., P.I. Miller, C.B. Embling, S.N. Ingram, E. Pirotta, S.C. Votier, 2014. Mesoscale fronts as foraging habitats: composite front mapping reveals oceanographic drivers of habitat use for a pelagic seabird. Journal of the Royal Society Interface 11, doi:10.1098/rsif.2014.0679.

Schwarz, J.N., 2020. Phytoplankton biovolumes from light microscopy; Chagos Archipelago, Indian Ocean, May 2018. Data submission 2020-06-30T16:25:05Z, pangaea.de.

Scott, B.E., J. Sharples, O.N. Ross, J. Wang, G.J. Pierce, C.J. Camphuysen, 2010. Foraging plasticity in a pelagic seabird species along a marine productivity gradient. Marine Ecology Progress Series 408, 207 – 226, doi:10.3354/meps08552.

SeaPAC. 2006. SeaWinds on QuikSCAT Level 3 Daily Gridded Ocean Wind Vectors (JPL Version 2). Ver. 2. PO.DAAC, CA, USA. Dataset accessed 2019-03 at https://doi.org/10.5067/QSXXX-L3002.

Seiki, A., M. Katsumata, T. Horii, T. Hasegawa, K.J. Richards et al., 2013. Abrupt cooling associated with the oceanic Rossby wave and lateral advection during CINDY2011. Journal of Geophysical Research: OCeans 118, 5523-5535, doi:10.1002/jgrc.20381.

Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's tau. Journal of the Americal Statistical Association 63(324), 1379-1389.

Sheppard, C.R.C., A. Harris, A.L.S. Sheppard, 2008. Archipelago-wide coral recovery patterns since 1998 in the Chagos Archipelago, central Indian Ocean. Marine Ecology Progress Series 362, 109-117, doi:10.3354/meps07436.

Sheppard, C.R.C., M. Ateweberhan, B.W. Bowen, P. Carr, C.A. Chen et al., 2012. Reefs and islands of the Chagos Archipelago, Indian Ocean: why it is the world's largest no-take marine protected area. Aquatic Conservation 22(2), 232-261, doi:10.1002/aqc.1248.

Soares, M.A., P.V. Bhaskar, R.K. Naik, D. Dessai, J. George et al., 2015. Latitudinal $\delta^{1}3$C and $\delta^{1}5$N variations in particulate organic matter (POM) in surface waters from the Indian ocean sector of Southern Ocean and the Tropical Indian Ocean in 2012. Deep-Sea Research 118, 186-196, doi:10.1016/j.dsr2.2015.06.009.

Solanki, H.U., D. Bhatpuria, P. Chauhan, 2015. Signature analysis of satellite derived SSHa, SST and chlorophyll concentration and their linkage with marine fishery resources. Journal of Marine Systems 150, 12 – 21, doi:10.1016/j.jmarsys.2015.05.004.

Solanki, H.U., D. Bhatpuria, P. Chauhan, 2017. Applications of generalized additive model (GAM) to satellite-derived variables and fishery data for prediction of fishery resources distributions in the Arabian Sea. Geocarto International 32(1), 30 – 43, doi:10.1080/10106049.2015.1120357.

Spalding, M.D., V.N. Agostini, J. Rice, S.M. Grant, 2012. Pelagic provinces of the world: A biogeographic classification of the world's surface pelagic waters. Ocean & Coastal Management 60, 19-30, doi:10.1016/j.ocecoaman.2011.12.016.

Sreeush, M.G., V. Valsala, S. Pentakota, K. Venkata Siva Rama Prasad, R. Murtugudde, 2018. Biological production in the Indian Ocean upwelling zones – Part 1: refined estimation via the use of a variable compensation depth in ocean carbon models. Biogeosciences 15, 1895-1918, doi:10.5194/bg-15-1895-2018.

Strutton, P.G., V.J. Coles, R.R. Hood, R.J. Matear, M.J. McPhaden, H.E. Phillips, 2015. Biogeochemical variability in the central equatorial Indian Ocean during the monsoon transition. Biogeosciences 12, 2367-2382, doi:10.5194/bg-12-2367-2015.

Sutton, T.T., M.R. Clark, D.C. Dunn, P.N. Halpin, A.D. Rogers et al., 2017.

A global biogeographic classification of the mesopelagic zone. Deep-Sea Research Part I 126, 85-102, doi:10.1016/j.dsr.2017.05.006.

Sutton, A.L., L.E. Beckley, 2017. Species richness, taxonomic distinctness and environmental influences on Euphausiid zoogeography in the Indian Ocean. Diversity 9, 23, doi:10.3390/d9020023.

Subha Anand, S., R. Rengarajan, V.V.S.S. Sarma, A.K. Sudheer, R. Bhushan, S.K. Singh, 2017. Spatial variability of upper ocean POC export in the Bay of Bengal and the Indian Ocean determined using particle-reactive 234Th. Journal of Geophysical Research – Oceans 122, 3753-3770, doi:10.1002/2016JC012639.

Talley, L.D., G.L. Pickard, W.J. Emery, J.H. Swift, 2011. Indian Ocean. Chapter 11 in Descriptive Physical Oceanography (6th Ed.), Academic Press, doi:10.1016/B978-0-7506-4552-2.10033-2.

Thorrington-Smith, M., 1971. West Indian Ocean phytoplankton: a numerical investigation of phytohydrographic regions and their characteristic phytoplankton associations. Marine Biology 9, 115-137.

Trevail, A.M., J.A. Green, J. Sharples, J.A. Poloton, J.P. Arnould, S.C. Patrick, 2019. Environmental heterogeneity amplifies behavioural response to a temporal cycle. Oikos 128, 517 – 528, doi:10.1111/oik.05579.

Trochta, J.T., C.B. Mouw, T.S. Moore, 2015. Remote sensing of physical cycles in Lake Superior using a spatio-temporal analysis of optical water typologies. Remote Sensing of Environment 171, 149-161, doi:10.1016/j.rse.2015.10.008.

Tweddle, J.F., M. Gubbins, B.E. Scott, 2018. Should phytoplankton be a key consideration for marine management? Marine Policy 97, 1-9, doi:10.1016/j.marpol.2018.08.026.

Uitz, J., H. Claustre, A. Morel, S.B. Hooker, 2006. Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. Journal of Geophysical Research 111, C08005, doi:10.1029/2005JC003207.

United Nations Framework Convention on Climate Change, 2015. Adoption of the Paris Agreement. Proposal by the President. UNFCCC. Conference of the Parties (COP), United Nations, Geneva. Available at: https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement.

UNEP-WCMC, 2016. World database on protected areas user manual 1.4. UNEP-WCMC: Cambridge, UK. Available at: http://wcmc.io/WDPA_Manual.

Vantrepotte, V., H. Loisel, D. Dessailly, X. Mériaux, 2012. Optical classification of contrasted coastal waters. Remote Sensing of Environment 123, 306-323, doi:10.1016/j.rse.2012.03.004.

Veldhuis, M.J.W., G.W. Kraay, J.D.L. van Bleijswijk, M.A. Baars, 1997. Seasonal and spatial variability in phytoplankton biomass, productivity and growth in the northwest Indian Ocean: the southwest and northeast monsoon, 1992-1993. Deep-Sea Research I 44(3), 425-449.

Verhoef, A., J. Vogelzang, J. Verspeek, A. Stoffelen, 2018. Long-term scatterometer wind climate data records. IEEE Journal of Selected Topics in Applied Earth Observation 10(5), 2186-2194, doi:10.1109/JSTARS.2016.2615873.

Vialard, J., J.P. Duvel, M.J. McPhaden, P. Bouruet-Aubertot, B. Ward et al., 2009. CIRENE: Air-sea interactions in the Seychelles-Chagos thermocline ridge region. Bulletin of the American Meteorological Society 90(1), 45-61, doi:10.1175/2008BAMS2499.1

Vichi, M., J. Icarus Allen, S. Masina, N.J. Hardman-Mountford, 2011. The emergence of ocean biogeochemical provinces: A quantitative assessment and a diagnostic for model evaluation. Global Biogeochemical Cycles 25, GB2005, doi:10.1029/2010GB003867.

Wang, X., F. Guillaume Blanchet, N. Koper, 2014. Measuring habitat fragmentation: An evaluation of landscape pattern metrics. Methods in Ecology & Evolution 5, 634-646, doi:10.1111/2041-210X.12198.

Wanningkhof, R., G.-H. Park, T. Takahashi, C. Sweeney, R. Feely et al., 2013. Global ocean carbon uptake: magnitude, variability and trends. Biogeosciences 10, 1983-2000, doi:10.5194/bg-10-1983-2013.

Waugh, S.M., H. Weimerskirch, 2003. Environmental heterogeneity and the evolution of foraging behaviour in long ranging greater albatrosses. Oikos 103, 374-384.

Webster, P.J., G.J. Holland, J.A. Curry, H.-R. Chang, 2005. Changes in tropical cyclone number, duration and intensity in a warming environment. Science 309(5742), 1844-1846, doi:10.1126/science.1116448.

Weimerskirch, H., M. Le Corre, Y. Ropert-Coudert, A. Kato, F. Marsac, 2005. The three-dimensional flight of red-footed boobies: adaptations to foraging in a tropical environment? Proceedings of the Royal Society B 272, 53 – 61, doi:10.1098/rspb.2004.2918.

Weimerskirch, H., 2007. Are seabirds foraging for unpredictable resources? Deep-Sea Research II 54, 211 – 223, doi:10.1016/j.dsr2.2006.11.013.

Werdell, P.J., L.I.W. McKinna, E. Boss, S.G. Ackleson, S.E. Craig et al., 2018. An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing. Progress in Oceanography 160, 186-212, doi:10.1016/j.pocean.2018.01.001.

Wheeler, M.C., H.H. Hendon, 2004. An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. Monthly Weather Review 132, 1917-132, doi:10.1175/1520-0493(2004)132¡1917:AARMMI¿2.0.CO;2.

Williamson, M.J., E.J. Tebbs, T.P. Dawson, D.M.P. Jacoby 2019. Satellite remote sensing in shark and ray ecology, conservation and management. Frontiers in Marine Science 6:135, doi:10.3389/fmars.2019.00135.

Wiggert, J.D., R.G. Murtugudde, J.R. Christian, 2006. Annual ecosystem variability in the tropical Indian Ocean: Results of a coupled biophysical ocean general circulation model. Deep-Sea Research II 53, 644-676, doi:10.1016/j.dsr2.2006.01.027.

Wilson, C., X. Qiu, 2008. Global distribution of summer chlorophyll blooms in the oligotrophic gyres. Progress in Oceanography 78(2), 107-134, doi:10.1016/j.pocean.2008.05.002.

Windham, M.P., 1982. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-4(4), 357-363.

Xie, X.L. and G. Beni, 1991. A validity measure for fuzzy clustering.IEEE Transactions on Pattern Analysis and Machine Intelligence 13(8), 841-847.

Xie, S.-P., H. Annamalai, F.A. Schott, J.P. McCreary, 2002. Structure and mechanisms of south Indian Ocean climate variability. Journal of Climate 15, 864-878.

Zydelis, R., R.L. Lewison, S.A. Schaffer, J.E. Moore, A.M. Boustany, J.J. Roberts, M. Sims, D.C. Duncan, B.D. Best, Y. Tremblay et al., 2011. Dynamic habitat models: using telemetry data to project fisheries bycatch. Proceedings of the Royal Society B – Biological Sciences 278(1722), 3191 – 3200, doi:10.1098/rspb.2011.0330.

5. Figure captions

Figure 1: Schematic representation of key biophysical linkages (not exhaustive). The flow of information begins with sunlight to the left. Physical variables that can be detected using remote sensing, followed by the oceanographic variables derived from them, are shown between the sun and the ocean surface processes. Oceanographic variables of interest that can be derived from remote sensing data are outlined in blue and abbreviations are explained in Table 1. The other variables shown are of interest to conservation, management or climate change applications but are not amenable to remote sensing.

Figure 2: Location of the study domain, adapted from Talley et al. (2011); Aguiar-Gonzalez et al. (2016). SCTR = Seychelles Chagos Thermocline Ridge; MPA = Chagos Marine Protected Area; SECC = South Equatorial Countercurrent

Figure 3: Bathymetry of the study domain. Shaded regions denote the sub-areas related to other published studies. Black line: coast; grey line: 200 m contour; black dashed line: 1000 m contour; thick black line: MPA boundary.

Figure 4: Central spectra for the 10-band classifications from Level 2 (upper) and Level 3 (lower) datasets with two to nine classes, compared with the 6-band JSM classification (shaded).

Figure 5: Mapping of the N5, N8 and JSM classes onto published PFT algorithms. SLC=Synechococcus-like cyanobacteria; xDiat was assigned where

the residual PHYSAT-like spectrum resembled the diatom criteria in Alvain et al. (2008) but with higher values; SynPro=spectrum matches SLC or Prochlorococcus except for one waveband, which fell in the other small-cell category. Algorithm acronyms are explained in section 3.1.

Figure 6: Comparison of class maps produced with the Level 2 and Level 3 Rrs data, classifications N5, N8 and JSM, in 2003. The Chagos MPA outline, 0 m and 1000 m isobaths are shown in black and the 2000 m isobath in grey.

Figure 7: Number of classes held at each pixel over the 17 year study period, by month, for Level 3 data, N=5. The Chagos MPA outline is shown in black.

Figure 8: Monthly 17-year climatologies of key remote sensing variables. The Chagos MPA outline and 0 m contour are shown in black, and the 3000 m isobath in grey.

Figure 9: Variability of remotely sensed parameters in the MPA and z1000 domains: Top row = N5; Middle row = N8; Lower row = JSM; black=MPA; cyan=whole domain with depth > 1000 m (z1000).

Figure 10: Variability of remotely sensed parameters for each sub-domain using the N5 classification.

Figure 11: Averaged Mann-Kendall-Sen correlation coefficients between variables within each class for each domain: a) MPA, b) SCTR-W, c) SCTR-

C, d) Wiggert-N, e) Wiggert-S, f) z1000. Only results with a p-value < 0.01, confirmed using bootstrapping to remove autocorrelation and subsampling effects, are shown.

Figure 12: OCO-2 lowest altitude pCO_2 distributions within the z1000 domain for the N5 classes applied to MODIS Level 2 data, January 2015.

Figure 13: a) to e) Density of patches of each N5 class (a=class 1, e=class 5); f) to j) Average patch area (f=class 1, j=class 5); k) to o) Average distance between patches (k=class 1, o=class 5) for the MPA, Wiggert-N and Wiggert-S domains. Note different y-axis limits are used to show detail.

Figure 14: IOTC fishing catch and catch per unit effort for the Indian Ocean surface (top row; a, b,c) and longline (lower row; d, e, f) fisheries.

Figure 15: Time-series for each N5 class within the MPA for a) Chl, b) FLH, c) FLH:Chl, d) SST, e) SSH and f) ∇SSH.

Figure 16: a) Summary of the temporal trends found within each domain ($p < 0.05$), including ranges of the Sen regression coefficient and rates of change. Trends for specific classes are denoted by class number; trends for the entire domain are indicated by block colour (red=positive trend; blue=negative trend). b) Summary of significant ($p < 0.05$) correlations between remotely sensed variables and the Indian Dipole Mode Index, including ranges of the Sen regression coefficient. Correlations for specific classes are denoted by class number; correlations for the entire domain are indicated

by block colour (red=positive trend; blue=negative trend). Full correlation results, including 95% confidence intervals on the regression slopes, are given in Table S1.

Appendix A. Appendix 1

Table A.7: Characterisation of variables within each class for the N5, N8 and JSM classification schemes in the whole study site where depth >200 m (total no. pixels 177,145,574)

Variable	Mean ±1σ												
	N5					N8							
Class	1	2	3	4	5	1	2	3	4	5	6	7	8
N (%)	13	25	30	22	10	10	15	21	21	17	11	5	<1
chl	0.053	0.086	0.122	0.181	0.456	0.049	0.074	0.097	0.126	0.166	0.242	0.612	1.914
	±.011	±.013	±.022	±.065	±.994	±.009	±.010	±.012	±.021	±.044	±.127	±1.336	±1.508
flh	0.040	0.045	0.050	0.064	0.107	0.040	0.044	0.046	0.051	0.061	0.076	0.128	0.248
	±.021	±.022	±.027	±.040	±.077	±.021	±.021	±.022	±.028	±.037	±.050	±.087	±.163
flh/chl	0.778	0.528	0.409	0.346	0.277	0.822	0.595	0.473	0.399	0.360	0.306	0.264	0.142
	±.433	±.258	±.208	±.197	±.142	±.457	±.294	±.228	±.205	±.203	±.170	±.127	±.094
SST	27.62	28.58	28.65	28.02	27.30	27.55	28.22	28.77	28.61	28.10	27.76	27.01	28.67
	±1.67	±1.99	±1.73	±1.56	±1.3	±1.57	±2.03	±1.88	±1.71	±1.62	±1.35	±1.36	±1.43
SLA	0.083	0.060	0.048	0.038	0.026	0.087	0.067	0.055	0.047	0.039	0.035	0.023	0.074
	±.080	±.076	±.072	±.075	±.090	±.078	±.080	±.073	±.073	±.074	±.079	±.094	±.085
∇SSH	-23.1	-5.8	-0.3	-1.7	-1.3	-23.8	-13.3	-1.0	-0.4	-3.1	2.8	-1.3	-22.4
	±62.0	±52.9	±49.1	±55.7	±72.5	±62.9	±56.5	±49.9	±49.0	±54.8	±59.2	±80.7	±67.5
u	-0.059	0.002	0.018	-0.010	0.020	-0.058	-0.031	0.023	0.016	-0.012	0.003	0.029	-0.000
	±.146	±.188	±.196	±.211	±.244	±.141	±.177	±.193	±.196	±.207	±.223	±.257	±.171
v	0.012	-0.006	-0.003	0.012	0.021	0.013	-0.000	-0.008	-0.002	0.010	0.016	0.027	-0.028
	±.089	±.111	±.120	±.147	±.222	±.087	±.104	±.115	±.121	±.138	±.178	±.245	±.201
EKE	0.017	0.024	0.027	0.034	0.056	0.016	0.022	0.026	0.027	0.032	0.042	0.066	0.037
	±.024	±.036	±.041	±.057	±.107	±.022	±.032	±.038	±.041	±.052	±.075	±.124	±.059
τ_E	-0.064	-0.033	-0.023	-0.021	-0.004	-0.066	-0.044	-0.026	-0.023	-0.024	-0.013	0.002	-0.012
	±.044	±.057	±.057	±.055	±.054	±.041	±.055	±.056	±.057	±.057	±.048	±.058	±.053
τ_N	0.031	0.024	0.023	0.022	0.008	0.031	0.028	0.023	0.024	0.025	0.011	0.013	0.003
	±.027	±.034	±.038	±.048	±.064	±.025	±.033	±.035	±.039	±.046	±.054	±.069	±.046
$\nabla \times \tau$	-1.43	-3.66	-3.80	-2.20	-2.14	-0.86	-3.38	-3.86	-3.76	-2.57	-1.77	-2.12	-0.76
	±31.8	±23.2	±20.8	±22.9	±27.0	±32.9	±25.7	±21.7	±20.7	±22.8	±21.8	±31.7	±43.4
Depth	4350	4275	3988	3599	3309	4320	4365	4184	3953	3666	3453	3192	677
	±1005	±980	±1047	±1222	±1347	±1002	±992	±978	±1057	±1197	±1276	±1378	±674
R555:443	0.006	0.022	0.101	-0.90	0.070	0.136	0.156	0.179	0.207	0.272	0.350	0.485	-0.893
	±8.05	±9.75	±23.38	±32.55	±30.69	±.028	±.058	±.082	±.093	±.103	±.126	±.152	±31.044
R555:488	-0.352	0.356	0.437	0.262	2.896	0.201	0.218	0.236	0.258	0.309	0.367	0.451	-0.003
	±88.63	±15.80	±50.55	±22.06	±482.27	±.095	±.054	±.077	±.089	±.098	±.115	±.127	±16.181
hline													

Table A.8: Characterisation of variables within each class for the N5, N8 and JSM classification schemes within the MPA (total no. pixels 5,343,309).

Variable	Mean ±1σ												
	N5					N8							
Class	1	2	3	4	5	1	2	3	4	5	6	7	8
N (%)	8	37	39	16	1	4	20	32	26	15	3	¡1	¡1
chl	0.063	0.088	0.123	0.180	0.253	0.058	0.077	0.099	0.128	0.171	0.227	0.244	0.404
	±.008	±.012	±.021	±.046	±.087	±.007	±.009	±.012	±.019	±.036	±.073	±.099	±.108
flh	0.049	0.049	0.060	0.094	0.117	0.050	0.048	0.050	0.062	0.089	0.114	0.081	0.092
	±.022	±.021	±.027	±.037	±.053	±.024	±.020	±.021	±.028	±.035	±.044	±.049	±.002
flh/chl	0.791	0.560	0.482	0.521	0.460	0.868	0.632	0.510	0.479	0.521	0.503	0.357	0.209
	±.376	±.240	±.197	±.190	±.198	±.427	±.269	±.214	±.195	±.191	±.187	±.205	±.070
SST	29.59	29.41	28.71	27.62	27.43	29.55	29.56	29.22	28.60	27.70	27.35	28.05	28.41
	±.77	±.89	±.99	±1.06	±1.04	±.72	±.87	±.89	±.99	±1.08	±.94	±1.20	±.41
SLA	0.147	0.072	0.027	0.008	0.015	0.163	0.100	0.051	0.023	0.009	0.008	0.025	0.110
	±.084	±.074	±.066	±.063	±.042	±.084	±.078	±.070	±.065	±.063	±.057	±.039	±.049
∇SSH	24.3	0.83	8.52	2.66	8.35	-29.4	-8.19	6.74	8.55	2.55	5.38	14.04	9.05
	±39.4	±34.3	±33.1	±34.6	±32.6	±38.7	±37.1	±32.2	±33.5	±34.9	±33.2	±29.3	±15.5
u	-0.118	0.018	0.071	0.040	0.048	-0.138	-0.038	0.058	0.071	0.041	0.046	0.048	0.110
	±.162	±.190	±.193	±.175	±.175	±.150	±.184	±.192	±.192	±.177	±.169	±.178	±.160
v	0.002	-0.014	-0.020	-0.010	-0.038	0.003	-0.007	-0.019	-0.020	-0.009	-0.019	-0.074	0.038
	±.089	±.090	±.088	±.093	±.111	±.086	±.090	±.090	±.088	±.092	±.102	±.110	±.051
EKE	0.024	0.023	0.026	0.021	0.024	0.025	0.022	0.025	0.025	0.021	0.021	0.026	0.020
	±.026	±.028	±.033	±.029	±.029	±.026	±.026	±.031	±.033	±.030	±.030	±.027	±.036
τ_E	0.003	-0.007	-0.027	-0.056	-0.047	0.002	-0.001	-0.014	-0.030	-0.055	-0.057	-0.016	-0.069
	±.034	±.040	±.047	± .050	±.059	±.034	±.037	±.042	±.048	±.049	±.051	±.068	±.0008
τ_N	0.000	0.011	0.027	0.047	0.040	-0.001	0.005	0.016	0.029	0.046	0.047	0.020	0.057
	±.017	±.025	±.032	±.034	±.039	±.017	±.022	±.027	±.032	±.034	±.034	±.046	±.003
$\nabla \times \tau$	-7.48	-7.04	-7.92	-7.85	-7.16	-7.86	-6.70	-7.52	-7.99	-7.77	-8.12	-3.01	0.00
	±17.7	±16.4	±16.5	±17.3	±26.6	±18.5	±16.5	±16.4	±16.4	±16.6	±22.0	±25.4	±8.2
Depth	3891	3730	3598	3396	3011	3938	3785	3688	3574	3441	3129	3113	368
	±1143	±1083	±1173	±1368	±1433	±1149	±1102	±1087	±1193	±1343	±1446	±1471	±140
R555:443	-0.113	0.048	0.216	0.250	0.139	0.117	0.147	0.173	0.203	0.260	0.300	0.620	6.934
	±8.70	±9.88	±21.98	±28.20	±22.04	±.028	±.048	±.080	±.099	±.110	±.084	±.123	±17.700
R555:488	-5.161	0.477	0.825	0.179	15.117	0.179	0.207	0.229	0.252	0.295	0.325	0.559	-12.526
	±136.2	±16.07	±51.55	±21.12	±617.4	±.538	±.046	±.076	±.094	±.105	±.075	±.084	±40.68

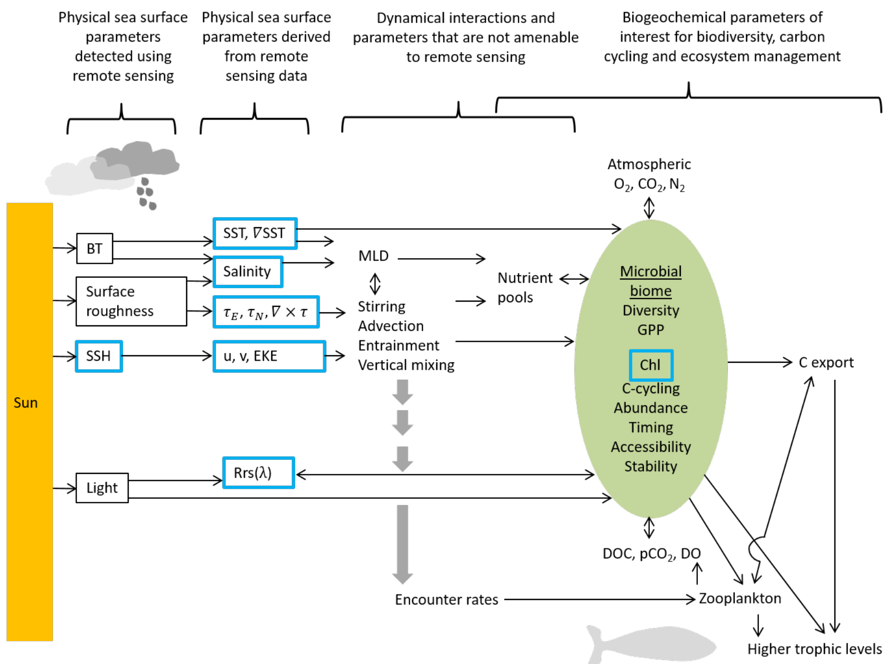

Figure 1: Schematic representation of key biophysical linkages (not exhaustive). The flow of information begins with sunlight to the left. Physical variables that can be detected using remote sensing, followed by the oceanographic variables derived from them, are shown between the sun and the ocean surface processes. Oceanographic variables of interest that can be derived from remote sensing data are outlined in blue and abbreviations are explained in Table 1. The other variables shown are of interest to conservation, management or climate change applications but are not amenable to remote sensing.

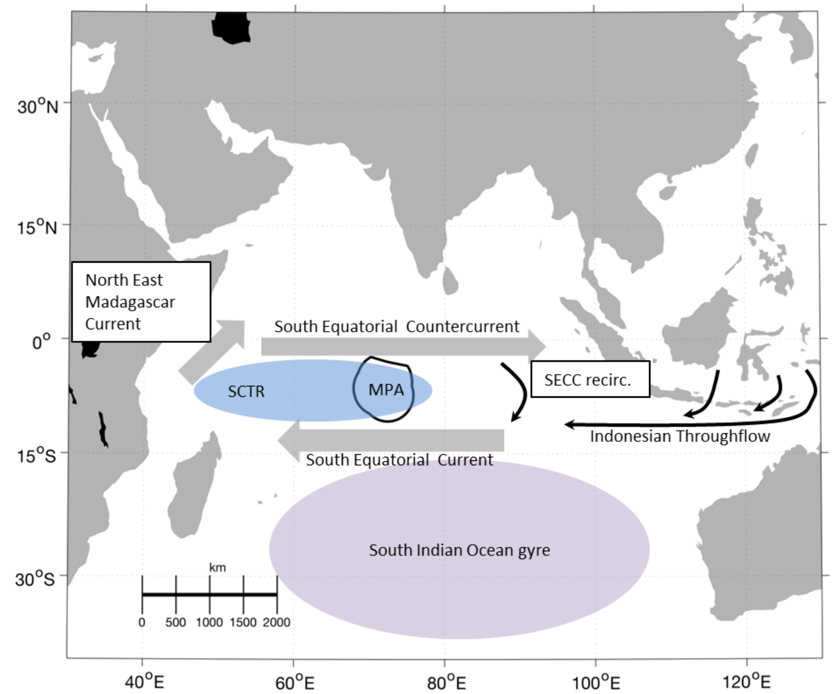

Figure 2: Location of the study domain, adapted from Talley et al. (2011); Aguiar-Gonzalez et al. (2016). SCTR = Seychelles Chagos Thermocline Ridge; MPA = Chagos Marine Protected Area; SECC = South Equatorial Countercurrent

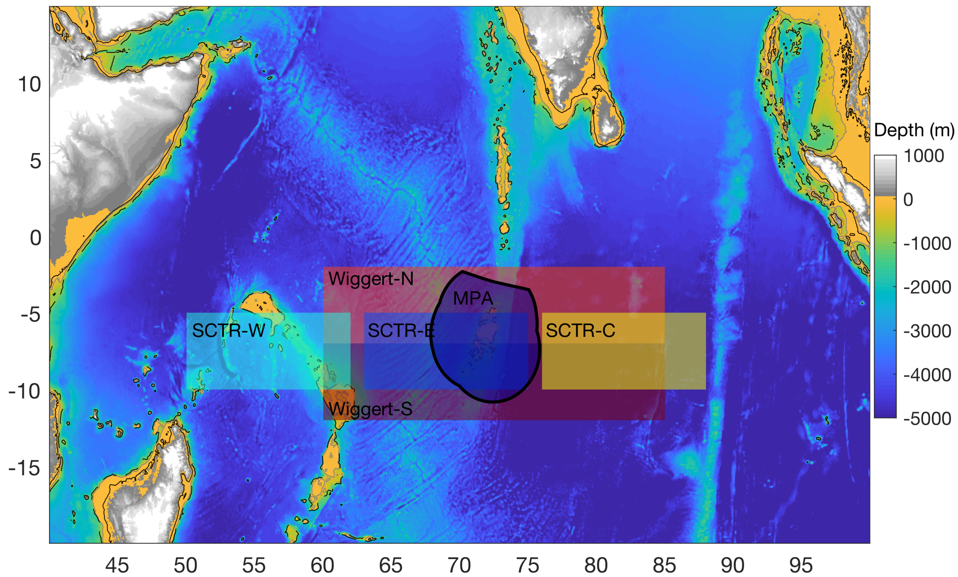

Figure 3: Bathymetry of the study domain. Shaded regions denote the sub-areas related to other published studies. Black line: coast; grey line: 200 m contour; black dashed line: 1000 m contour; thick black line: MPA boundary.

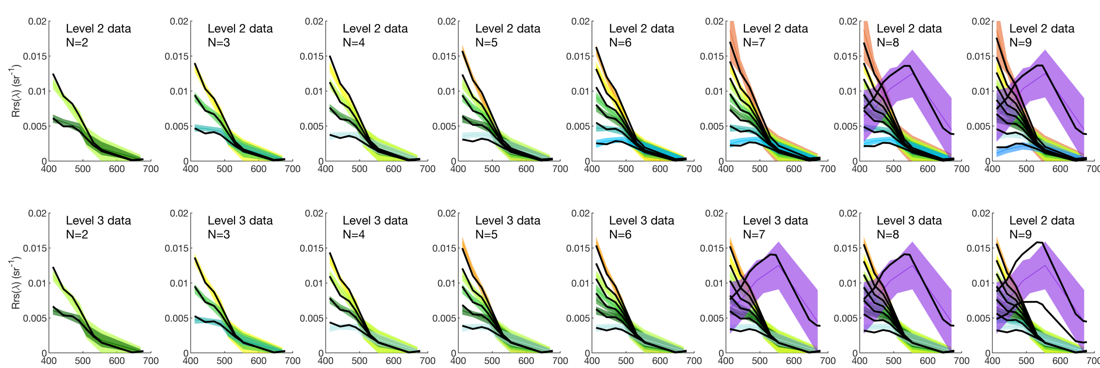

Figure 4: Central spectra for the 10-band classifications from Level 2 (upper) and Level 3 (lower) datasets with two to nine classes, compared with the 6-band JSM classification (shaded).

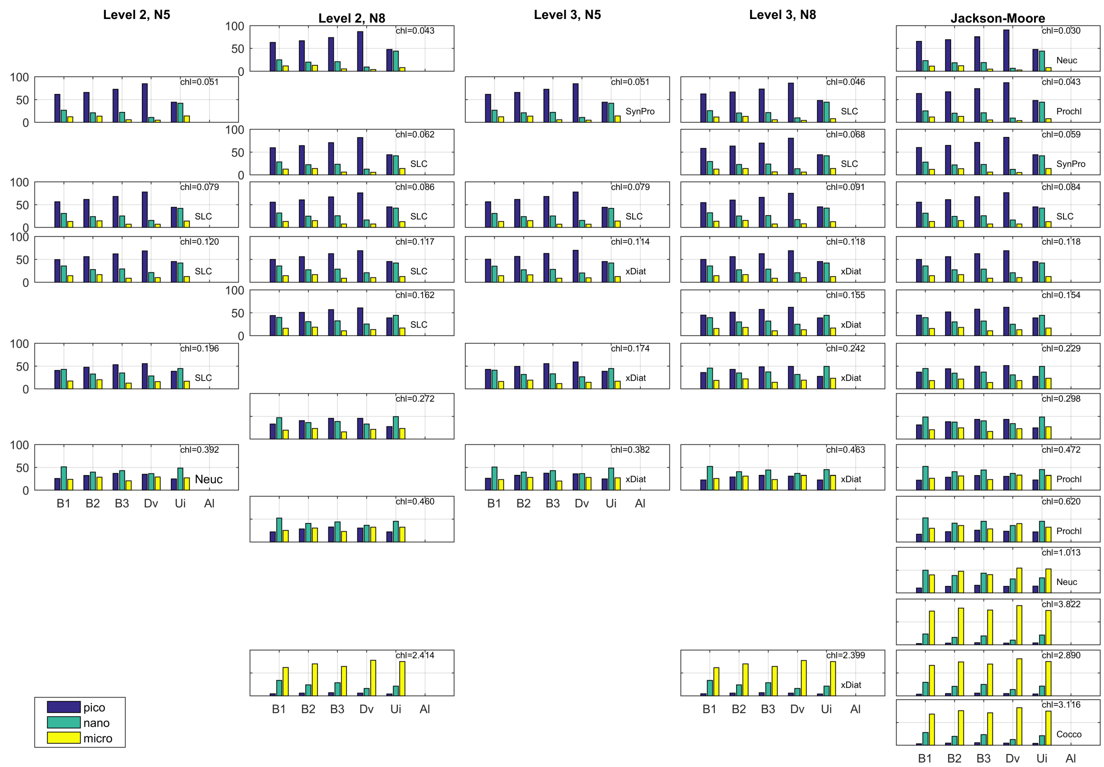

Figure 5: Mapping of the N5, N8 and JSM classes onto published PFT algorithms. SLC=Synechococcus-like cyanobacteria; xDiat was assigned where the residual PHYSAT-like spectrum resembled the diatom criteria in Alvain et al. (2008) but with higher values; SynPro=spectrum matches SLC or Prochlorococcus except for one waveband, which fell in the other small-cell category. Algorithm acronyms are explained in section 3.1.

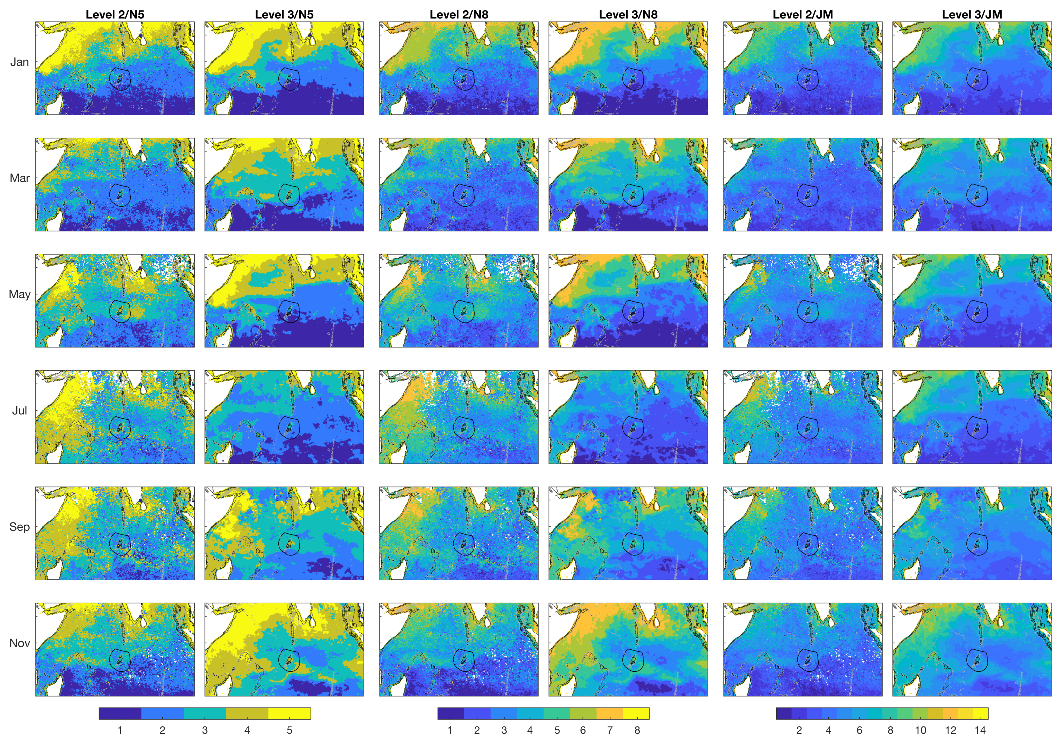

Figure 6: Comparison of class maps produced with the Level 2 and Level 3 Rrs data, classifications N5, N8 and JSM, in 2003. The Chagos MPA outline, 0 m and 1000 m isobaths are shown in black and the 2000 m isobath in grey.

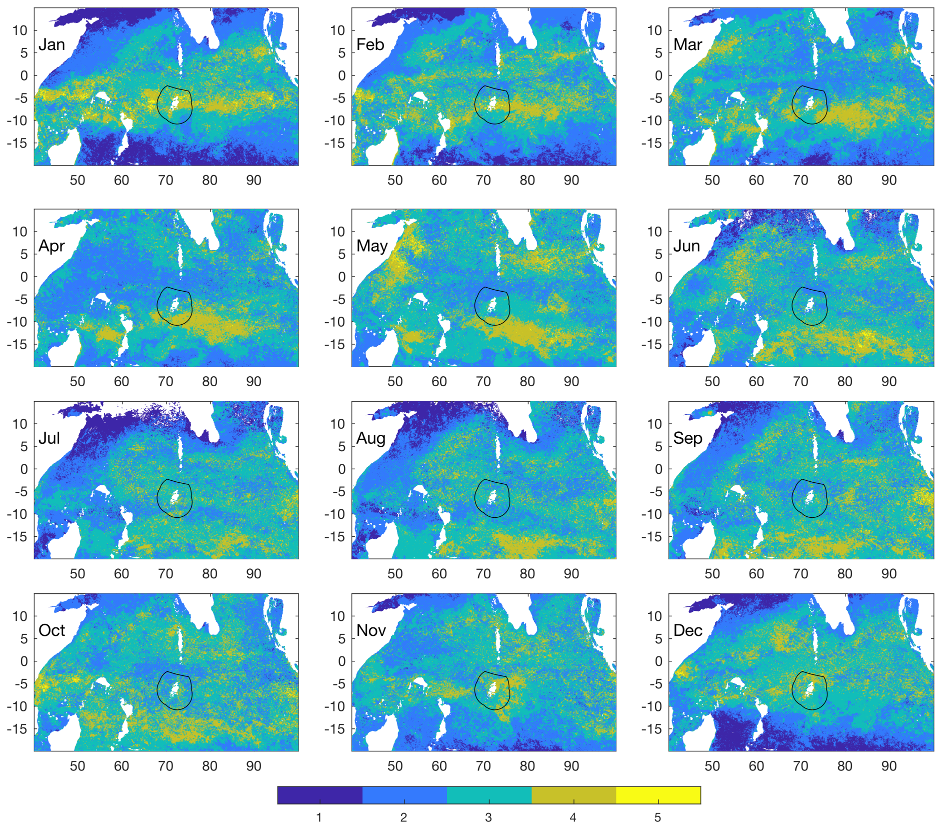

Figure 7: Number of classes held at each pixel over the 17 year study period, by month, for Level 3 data, N=5. The Chagos MPA outline is shown in black.

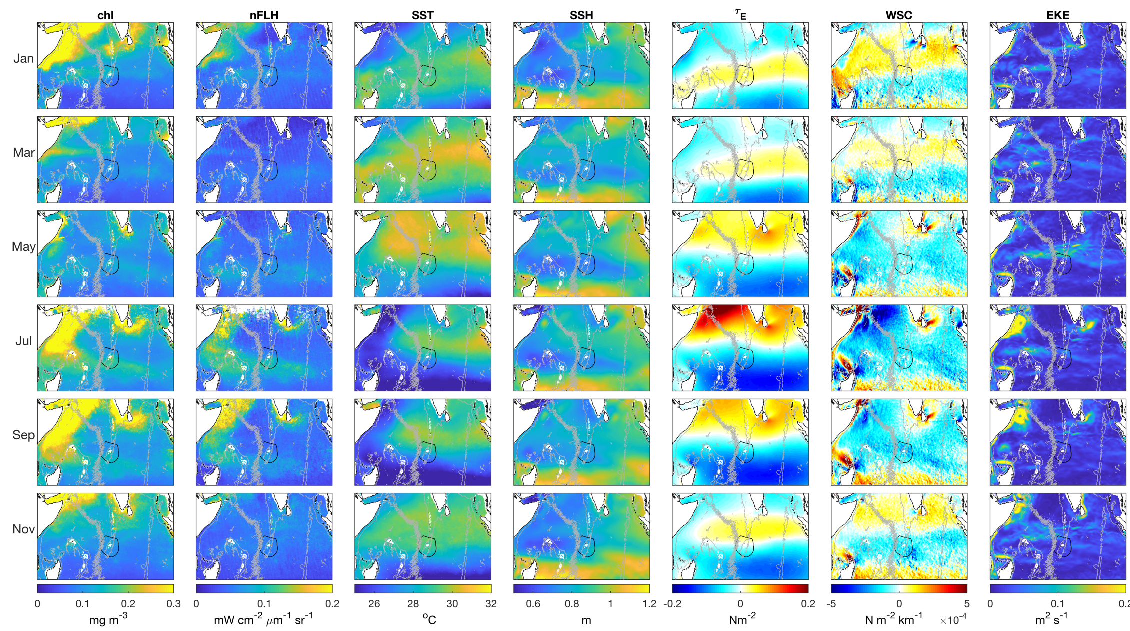

Figure 8: Monthly 17-year climatologies of key remote sensing variables. The Chagos MPA outline and 0 m contour are shown in black, and the 3000 m isobath in grey.

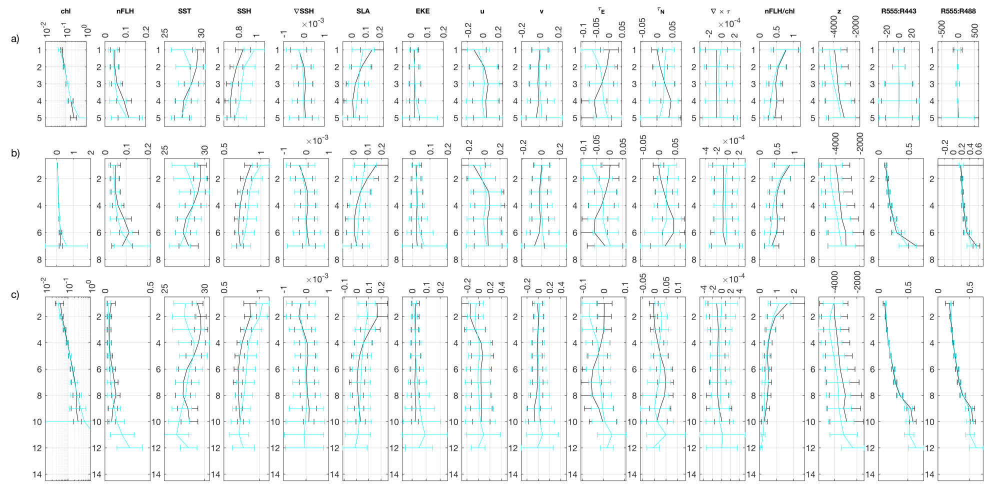

Figure 9: Variability of remotely sensed parameters in the MPA and z1000 domains: Top row = N5; Middle row = N8; Lower row = JSM; black=MPA; cyan=whole domain with depth > 1000 m (z1000).

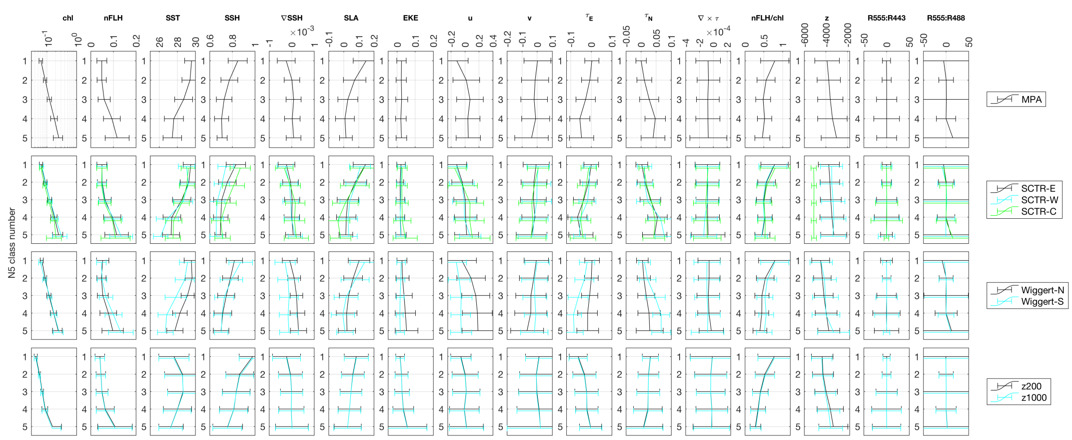

Figure 10: Variability of remotely sensed parameters for each sub-domain using the N5 classification.

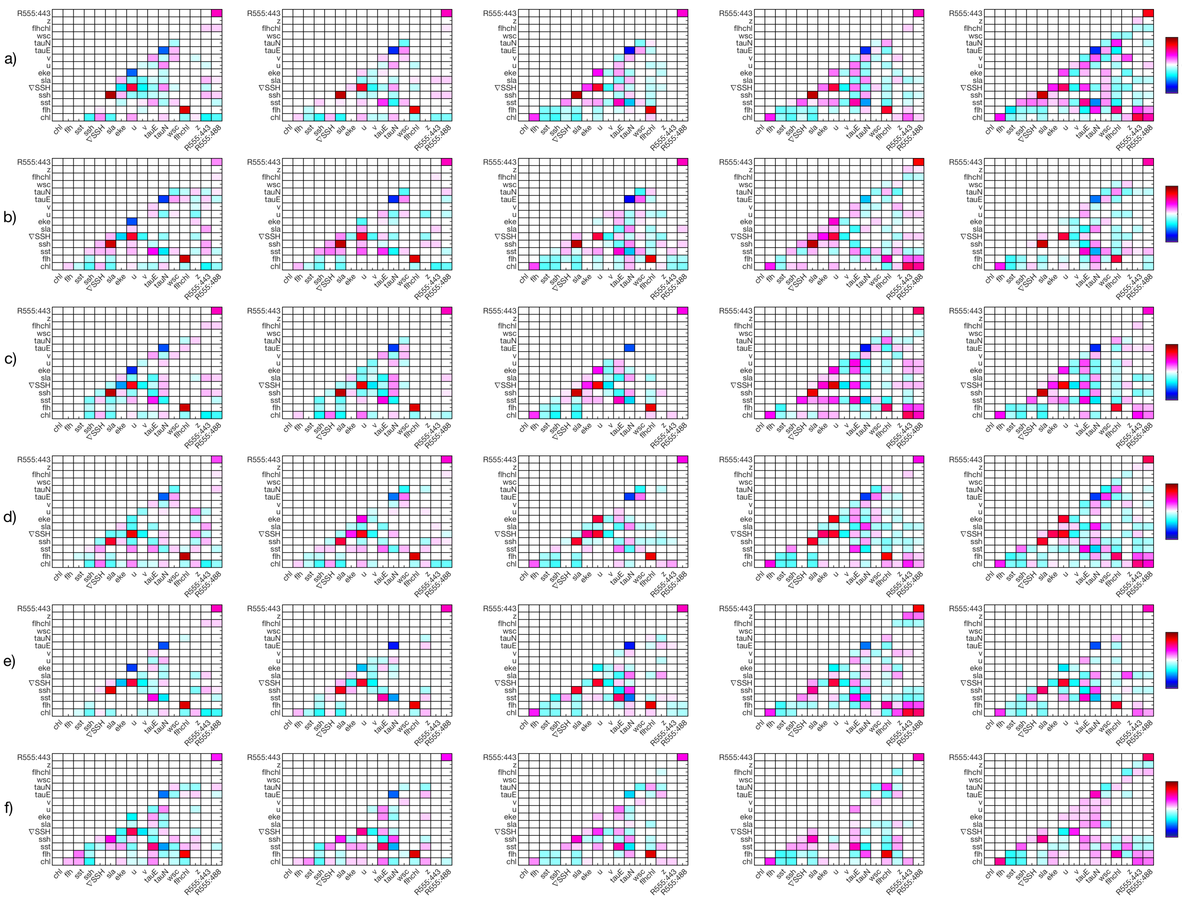

Figure 11: Averaged Mann-Kendall-Sen correlation coefficients between variables within each class for each domain: a) MPA, b) SCTR-W, c) SCTR-C, d) Wiggert-N, e) Wiggert-S, f) z1000. Only results with a p-value < 0.01, confirmed using bootstrapping to remove autocorrelation and subsampling effects, are shown.

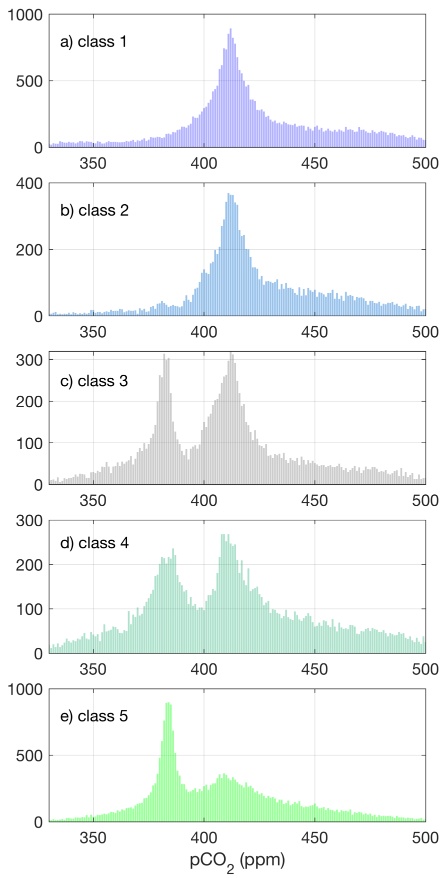

Figure 12: OCO-2 lowest altitude pCO_2 distributions within the z1000 domain for the N5 classes applied to MODIS Level 2 data, January 2015.

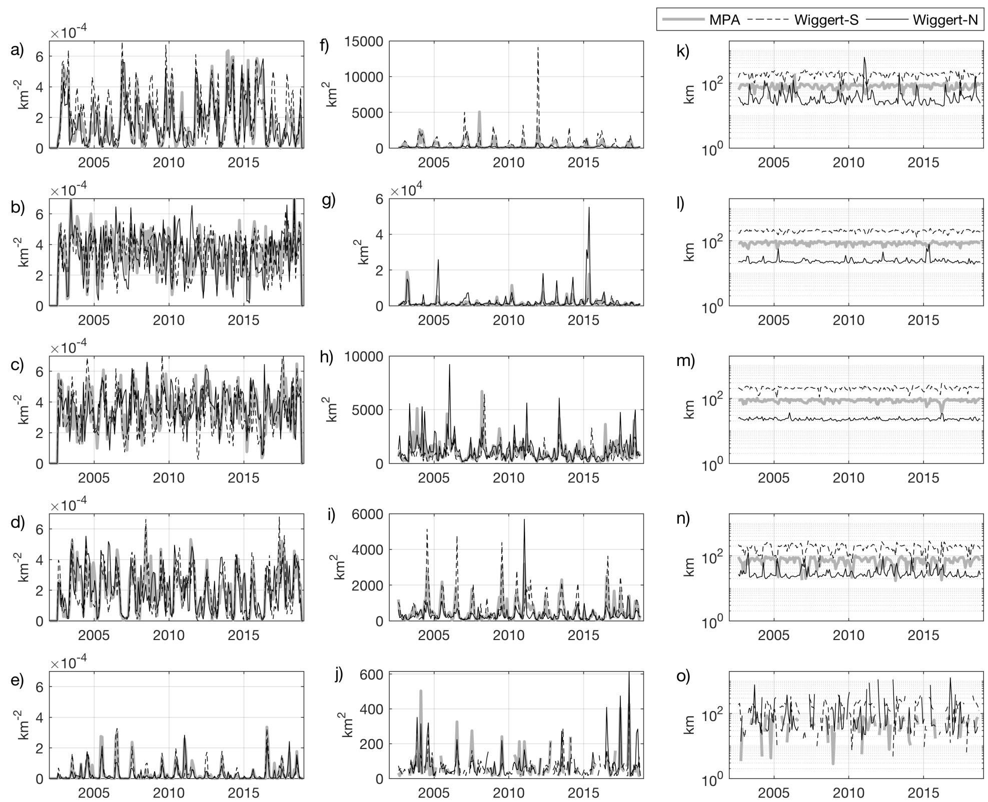

Figure 13: a) to e) Density of patches of each N5 class (a=class 1, e=class 5); f) to j) Average patch area (f=class 1, j=class 5); k) to o) Average distance between patches (k=class 1, o=class 5) for the MPA, Wiggert-N and Wiggert-S domains. Note different y-axis limits are used to show detail.

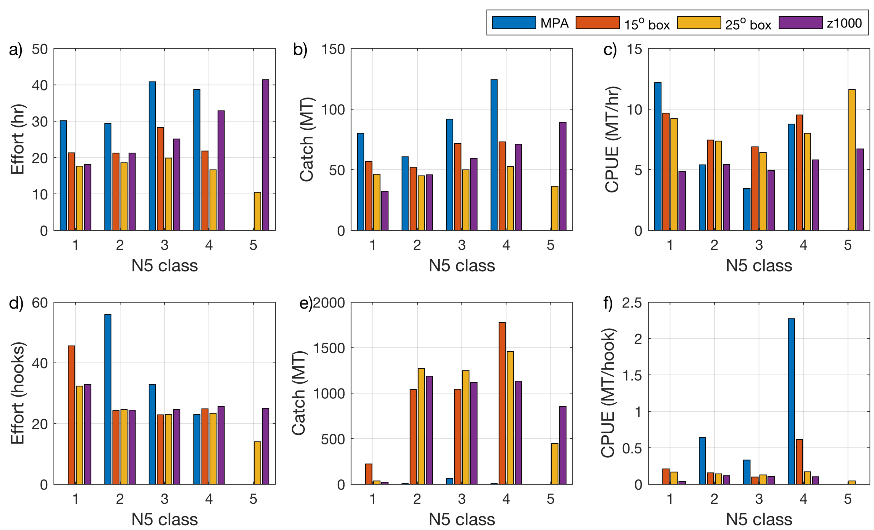

Figure 14: IOTC fishing catch and catch per unit effort for the Indian Ocean surface (top row; a, b,c) and longline (lower row; d, e, f) fisheries.

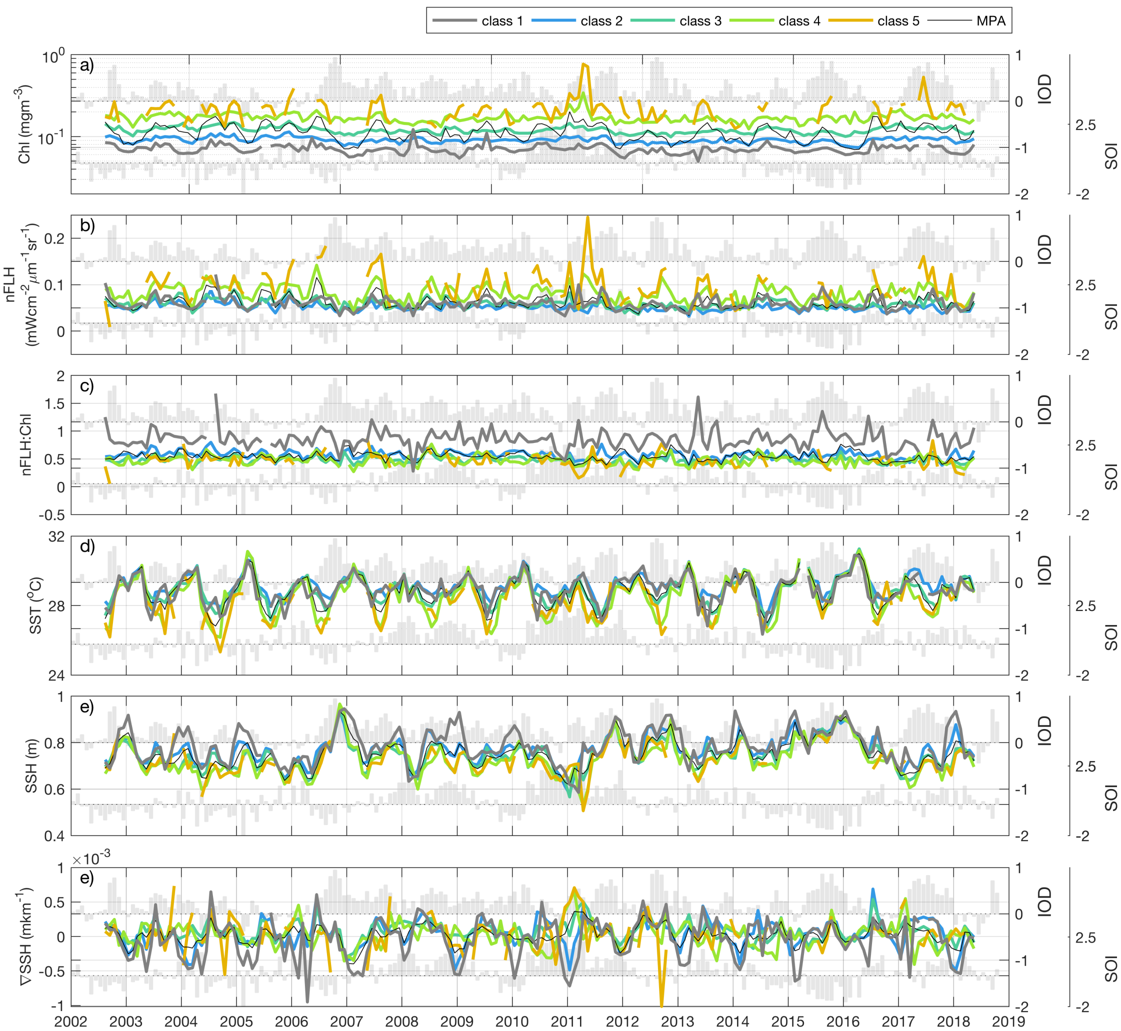

Figure 15: Time-series for each N5 class within the MPA for a) Chl, b) FLH, c) FLH:Chl, d) SST, e) SSH and f) ∇SSH.

a)

		Chl	FLH	FLH:Chl	R555:R443	R555:R488	SST	SSH	∇SSH	u	v	EKE	∇×τ	z
MPA		-234	-234	-25	-4			+2345		+1				+5
z1000		-12345	-1234	-23	+2			+12345		+2	-1			
SCTR-W		-1234	-1234					-1234						
SCTR-E		-234	-1234	-2345			+5	-2345						+5
SCTR-C		-2	-12	-12			+12		+2	+12				
Wiggert-N		-1234	-12345	-345	+2		+3	+12345	+12	+12				
Wiggert-S		-1234	-1234	-123				+34	+5	+45	+3	-45	+2	+5
Regression coefficient, S:	Min.	-0.096	-0.099	-0.099	-0.105	-0.096	0.105	0.102	0.099	0.097	-0.145	-0.127	-0.107	0.110
	Max.	-0.261	-0.313	-0.221	0.201	-0.097	0.210	0.550	0.127	0.152	0.108	-0.134	0.119	0.157
Trend (per annum):	Min.	-7.55×10^{-6}	-1.25×10^{-5}	-7.68×10^{-5}	-5.68×10^{-4}	-7.55×10^{-4}	1.27×10^{-4}	4.76×10^{-6}	3.30×10^{-9}	7.97×10^{-6}	-1.28×10^{-5}	-9.49×10^{-6}	5.32×10^{-8}	1.36
	Max.	-1.26×10^{-4}	-5.21×10^{-5}	-2.65×10^{-4}	2.22×10^{-4}	-8.18×10^{-4}	1.87×10^{-3}	1.51×10^{-4}	3.52×10^{-7}	2.34×10^{-4}	3.28×10^{-5}	-1.26×10^{-5}	-2.07×10^{-5}	2.14

- negative trend, whole domain, raw values only
- negative trend, whole domain, anomalies only
- negative trend, whole domain, raw & anomalies
- posative trend, whole domain, raw values only
- posative trend, whole domain, anomalies only
- posative trend, whole domain, raw & anomalies

b)

		Chl	FLH	FLH:Chl	R555:R443	R555:R488	SST	ΔSST	SSH	∇SSH	u	v	EKE	τ_E	τ_N	∇×τ	z
MPA						-3		+12345	+12345	+12		-1	-2	-12345	+345	-1235	-3
z1000			-2					+12345	+234	$+1_{-35}$	+15	+3	$+5_{-23}$	-234	+4		
SCTR-W						+3	-1	+12345	+1234			-2		-12345	+12345	-1234	-2 / +3
SCTR-E		-34						+12345	+12345	+12	+12	-2		-12345	+12345		
SCTR-C		-234	-12		+2 / -3	-3	-3	-123	+12345	-34	-234		-135	-12345	+34	+4	-4
Wiggert-N		-123				-3		+12345	+12345	-123			-234	-12345	+34	-34	+34
Wiggert-S					-2	5	-45	+123	+12345	+123	+123	-12		-12345	+345		
Regression coefficient, S:	Min.	-0.097	-0.098	+0.100	-0.172	-0.231	-0.134	-0.181	+0.105	-0.215	-0.209	-0.265	-0.206	-0.098	-0.099	-0.158	-0.116
	Max.	-0.293	-0.190	+0.143	+0.173	+0.132	+0.254	+0.266	+0.279	+0.245	+0.181	+0.127	+0.117	-0.283	+0.165	+0.163	+0.116

- negative trend, whole domain, raw values only
- negative trend, whole domain, anomalies only
- negative trend, whole domain, raw & anomalies
- posative trend, whole domain, raw values only
- posative trend, whole domain, anomalies only
- posative trend, whole domain, raw & anomalies

Figure 16: a) Summary of the temporal trends found within each domain ($p < 0.05$), including ranges of the Sen regression coefficient and rates of change. Trends for specific classes are denoted by class number; trends for the entire domain are indicated by block colour (red=positive trend; blue=negative trend). b) Summary of significant ($p < 0.05$) correlations between remotely sensed variables and the Indian Dipole Mode Index, including ranges of the Sen regression coefficient. Correlations for specific classes are denoted by class number; correlations for the entire domain are indicated by block colour (red=positive trend; blue=negative trend). Full correlation results, including 95% confidence intervals on the regression slopes, are given in Table S1.