Université Pierre et Marie Curie

Doctoral School of Environmental Sciences

Laboratory LOCEAN / Research group PEPS

Planktonic biodiversity hotspots in the open ocean: detection, drivers and implications at the global scale.

of Alice Soccodato

PhD dissertation in Marine Sciences

Supervised by Francesco d'Ovidio, Silvia De Monte, Marina Lévy

Publicly defended on the 18th December 2014

Against the committee composed by the following members:

Philippe Koubbi, President, Professor UPMC, Paris

Francesco d'Ovidio, Supervisor, Researcher LOCEAN, Paris

Marina Lévy, Supervisor, Researcher LOCEAN, Paris

Silvia De Monte, Supervisor, Researcher ENS, Paris

Chris Bowler, Adjudicator, Researcher ENS, Paris

Cristobal Lopez, Reviewer, Professor IFISC, Spain

Maurizio Ribera D'Alcalà, Reviewer, Professor SZN, Italy

“Sometimes a silent partner, sometimes an irruptive presence,
nature has always been my teacher.

Shared between rational observation and unintelligible enchantment,
my mind has
been absorbed by her beauty and power.

The overwhelming curiosity to undercover the mechanism underneath everything
led me to the most vast and intriguing world of the natural sciences,
βιολογία, the study of life...” Alice S.

Acknowledgements:

They require way more work than the thesis itself!!!

Résumé

Les patterns de biodiversité et les mécanismes qui les maintiennent ont toujours intéressé les biologistes et ont été abordés en considérant des facteurs géologiques, évolutifs et écologiques. Les processus écologiques qui déterminent la co-occurrence des espèces diffèrent en fonction de l'environnement physique de l'écosystème. De nombreuses théories ont proposé des relations entre les tendances observées dans la diversité des espèces et les caractéristiques physiques de l'environnement à grande échelle. Dans les milieux terrestres et aquatiques, l'impact de la température sur la distribution de la biodiversité compte parmi les facteurs les plus influentes et étudiés. Toutefois, de nombreux taxa marins représentent des exceptions à cette influence primaire de la température, alors qu'une fraction dominante des espèces marines est planctonique ou à larves dispersibles. La dispersion par le transport physique a certainement un impact majeur sur les patterns d'abondance des espèces dans l'environnement marin. Certains courants océaniques peuvent en effet contraindre la distribution des stades planctoniques de certaines espèces, même lorsque les paramètres démographiques et physiologiques des espèces sont insensibles aux propriétés de l'eau. Les mécanismes de transport peuvent donc influencer la distribution de la diversité à toutes les échelles, de l'individu aux populations jusqu'aux espèces. Contrairement aux écosystèmes terrestres, les écosystèmes en milieu marin sont sujets à une variabilité dont les échelles spatiales et temporelles sont dictées par les processus du transport physique turbulent. Cet aspect complique l'obtention d'informations synoptiques sur la distribution des espèces marines au niveau global et à haute résolution, alors que cette vision globale est essentielle pour pouvoir comprendre les patterns de biodiversité et les mécanismes impliqués dans leurs variations. En outre, les hotspots de biodiversité sont d'importance primaire pour les efforts de conservation.

Les objectifs de cet étude sont les suivants: identifier les hotspots de biodiversité pélagique des producteurs primaires à l'échelle globale et à haute résolution; déterminer les processus physiques de l'océan qui contrôlent la dynamique spatio-temporelle des hotspots, en se focalisant sur les mécanismes de transport, de dispersion, advection et mélange; étudier l'influence de ces mécanismes de structuration de la biodiversité sur les niveaux trophiques supérieurs.

Pour obtenir ces résultats, les informations sur les parcelles d'eaux aux caractéristiques biophysiques cohérentes ('niches fluido-dynamiques') obtenues par satellite sont utilisées pour

identifier les hotspots de biodiversité microbienne comme région de forte variabilité spatiale de ces niches. Ces hotspots et le rôle du transport dans leur structuration sont étudiés par l'analyse des modèles écologiques et biophysiques de circulation globale (Modèle ECCO2-Darwin) et par l'examen de données moléculaires et morphologiques sur la structure de la communauté in-situ collectées par l'expédition Tara Oceans et Atlantic Meridional Transect. Les possibles effets 'bottom-up' de la diversité des producteurs primaires sur les niveaux supérieurs de la chaine trophique sont évalués par comparaison avec des modèles globaux qui intègrent des bases de données in situ.

Les modèles écologiques couplés avec la circulation océanique identifient comme hotspots de biodiversité des producteurs primaires les domaines parmi les plus dynamiques de l'océan mondial, caractérisés par une turbulence et un mélange développés et la présence de tourbillons. Ces caractéristiques océanographiques peuvent améliorer la productivité locale par le transport des nutriments dans la zone euphotique et augmenter la biodiversité par la juxtaposition des espèces typiques des différentes masses d'eau. De plus, les cartes de biodiversité microbienne suggèrent une propagation bottom-up de la biodiversitè au sein de l'écosystème, les hotspots des producteurs primaires étant corrélés positivement avec la distribution des espèces de prédateurs supérieures.

Les effets du couplage biophysique sur la distribution de la biodiversité n'ont jamais étés observées auparavant à l'échelle globale, car la résolution spatio-temporelle des cartes de biodiversité estimée au milieu pélagique était insuffisante. La disponibilité des cartes satellitaires, l'accessibilité aux modèles de simulation globale à haute résolution et une forte collaboration interdisciplinaire entre océanographes, biologistes moléculaires et modélisateurs des écosystèmes va permettre de les explorer. En plus de l'impact scientifique sur la compréhension des processus structurant la biodiversité, les résultats de cette thèse ouvrent la voie à l'utilisation de données satellite pour la planification des sanctuaires marins internationaux et pour la gestion des ressources marines pélagiques.

Summary

Patterns of biodiversity and the mechanisms that maintain them have always interested biologists and have been addressed considering geological, evolutionary and ecological factors. Ecological processes that determine the co-occurrence of species differ according to the physical environment of the ecosystem. Many theories have proposed relationships between patterns in species diversity and large-scale physical features. In terrestrial and aquatic environments, the impact of temperature on the distribution of biodiversity is among the most influent and studied factors. However, many marine taxa are exceptions in the primary influence of temperature, since a large fraction of marine species is planktonic or with dispersible larvae. In the marine environment, dispersal through physical transport has a major impact on patterns of species abundance. Some ocean currents can indeed determine the distribution of planktonic stages of some species, even when demographic and physiological features of the species are unaffected by water properties. Transport mechanisms may therefore influence the distribution of diversity at all scales, from the individual to populations and species. Contrarily to the terrestrial environment, marine ecosystems are characterized by a variability that has spatial and temporal scales defined by specific biophysical processes of turbulent transport. This aspect makes it challenging to provide synoptic information on the distribution of marine species at the global level and at high resolution, features that are essential to understand patterns of biodiversity and the mechanisms involved in their changes. Moreover, hotspots of biodiversity are of primary concerns for conservation efforts.

The objectives of this study are therefore: to identify biodiversity hotspots of pelagic primary producers on a global scale and at high resolution; to determine the physical ocean processes that control the spatial and temporal dynamics of such hotspots, focusing on transport-driven mechanisms like dispersion, advection and mixing; study the role of these mechanisms in the structuring of biodiversity at higher trophic levels.

To obtain these results, information on water masses with coherent biophysical characteristics ('fluid-dynamical niches') obtained by remote sensing are used to identify hotspots of microbial biodiversity as regions of strong spatial patchiness. These hotspots and the role of transport in shaping their structure are studied by analysing ecological and biophysical global circulation models (Model-ECCO2 Darwin), together with molecular and morphological data on the

structure of the community, obtained using in-situ data collected during the Tara-Oceans expedition and Atlantic Meridional Transect. The possible bottom-up effects of the diversity of primary producers on the upper levels of the food chain are evaluated by comparing them with global models integrated with data collected in situ.

The ecological models coupled with ocean circulation, identified as biodiversity hotspots of primary producers the most dynamic areas of the global ocean characterized by increased turbulence, mixing and the presence of vortices. These oceanographic features can improve local productivity by transporting nutrients in the photic zone and increase biodiversity by the mixing of species typical of different water masses. In addition, maps of microbial biodiversity suggest a bottom up propagation of biodiversity across the ecosystem, hotspots for primary producers being positively correlated with regions where highest number of top predator species are observed.

The effects of the biophysical coupling on the distribution of biodiversity have never been observed before at the global scale, because the spatial and temporal resolution maps of the estimated pelagic biodiversity were insufficient. The availability of satellite maps, access to high-resolution models and a strong interdisciplinary collaboration between oceanographers, molecular biologists and ecosystem modellers allow to explore them. In addition to the impact on scientific understanding of the processes structuring biodiversity, the results of this thesis may pave the way to the use of remote-sensing observations for the planning of international marine sanctuaries and in the management of pelagic marine resources.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

In this Chapter I will briefly define the conceptual framework of my thesis. My work has been an attempt to understand how global patterns of phytoplankton biodiversity can be detected and described for the pelagic open ocean realm by combining remote sensing, models, and in situ observations. In this introductory part I will briefly present a) the concept of biodiversity and its measures, b) the biology and biogeography of plankton, b) the interaction of plankton with physical processes acting at different spatial and temporal scales, and c) the techniques available to study biophysical processes and microbial responses in the ocean. More specifically, in this chapter, I will highlight how turbulent physical structures in the ocean, lead to persistent structures in marine populations. This mechanistic understanding has been applied to coupled biological–physical oceanographic models, which are one set of tools increasingly being used to predict dispersal patterns of marine species, and to satellite data.

1.1 Summary of the context

Knowledge about Earth's ecosystems has always privileged terrestrial environments due to their relative easy accessibility and primary interest for resource exploitation. For this reason, our view at oceanic systems has always been a terrestrial one, i.e., a look based more on the extrapolation of terrestrial ecology than on the development of concepts specific to the oceanic environment.

Far from the coast the vast extension of the ocean surface was nothing more than a limitless desert of water to human eyes, a habitat far too extreme to support high diversity of life. This led to misleading conclusions about the features that characterize the oceans, the environment that constitutes 72% of the surface of the planet. Especially, open oceans were considered as homogeneous and less diverse environments and their importance for the planet and the non-endless nature of their resources was still not clarified and recognized.

In the last decades however, the advance in technology has allowed us to look at the ocean from a different angle: from the sky (satellite), from the surface (research vessel and buoys) and within the ocean (sounder, submarine, etc.). This new prospective has brought us one step closer to understand the true nature of the ocean, a highly dynamic, extremely heterogeneous and invaluable yet poorly known source of biological diversity.

Oceans are the lifeblood of planet Earth and humankind. Oceans are responsible for the balance of biogeochemical cycles and are key contributors to ecosystem services and to the homeostasis (complex set of interacting reactions which maintain the planet in a revolving balance) of the planet. Ocean covers nearly three-quarters of the entire planet earth, and hold about 97% of the planet's water. Moreover they provide food and natural resources exploited worldwide.

Oceans have a strong influence on the world climate as well. In the light of the on-going changes for the maintenance/viability of natural resources and ecosystem services, as well as in the perspective of a changing climate (Hutchinson 1959), the lack of information and still overall poor understanding of the marine ecosystem is now a strong challenge inspiring researchers all around the world.

Most recent marine ecosystem studies have brought to the general attention the existence of a strong and complex interaction of marine biota with their fluid environment. However, one of the most common questions, how biodiversity is distributed across the oceans, still remains

unresolved. This question can re rephrased as: where are biodiversity hotspots geographically located? It is now clear that the ocean turbulent structures shape biodiversity, yet the studies attempting to analyse this process at both physical and biological level are few.

In this study I used remote sensing, ocean circulation models and molecular techniques to explore global patterns of biodiversity in the open ocean. My main result is the development of a novel method for the identification of biodiversity hotspots for plankton (primarily phytoplankton and autotrophic bacteria also known as primary producers) at the global scale. Additionally I also identified which are some of the possible drivers of these hotspots and the implications for the conservation of diversity at the higher levels of the trophic chain.

The existence of phytoplankton is of a fundamental importance for the stability and functioning of the oceanic ecosystem and ocean biogeochemical cycles (Ptacnik *et al.* 2008) as they contribute to 90% of ocean primary production corresponding to the 50% of the global oxygen production (Falkowski *et al.* 1998). Additionally together with the zooplankton form the base of the aquatic food webs, providing an essential ecological function for all aquatic life. The identification of biodiversity hotspots will allow to promote the conservation of areas including a wider range of species laying in the higher trophic level that directly or indirectly rely on plankton for their survival.

The challenge is now how to integrate and interpret remote sensing observation with in-situ biological data to advance our understanding of the whole ecosystem and the processes underneath the newly discovered features.

1.2 Biodiversity in the open ocean

1.2.1 Biodiversity Hotspots: history, definition and challenges

The concept of Biodiversity Hotspot was introduced to provide scientists and policy makers with a practical tool to face the problem of the current species extinction crisis (Koh *et al.* 2004; Briggs 2011). Since its introduction, the concept has received substantially coverage in the scientific literature, resulting in several slightly different definitions. The first definition was coined by Myers as “areas featuring an exceptional concentration of endemic species and experiencing an exceptional loss of habitat” (Myers 1990; Myers *et al.* 2000). In general

Biodiversity Hotspots have been used to identify areas particularly rich in number of species, in rare, in threatened or in endemic species. Despite numerous modifications of the original definition, high species richness appears to be the basic requirement for an area to be classified as a hotspot. The drawback of this approach is that the taxonomic classification at species level is extremely difficult and time consuming, and often requires a multidisciplinary approach (morphological, genetic and molecular) and advanced technologies. Therefore rather than relying on taxonomy at species level to reflect species richness, more recent studies have investigated other measures potentially correlated to species richness, such as habitat heterogeneity, higher taxonomic levels, primary productivity, and functional diversity (Gaston 1996b; Kerr & Packer 1997; Mittelbach *et al.* 2001; Heino & Soininen 2007).

The assessment and application of some of these approaches are far more difficult in the marine environment than on land. The identification of pelagic hotspots in open ocean has been limited due to the difficulty of a) obtaining long-term distribution data on pelagic species; b) the dynamic character of pelagic habitats, and c) the difficulties in establish international relationships (Game *et al.* 2009). The ocean is a highly dynamic environment in constant change, therefore habitat based approaches are hard to apply. Additionally the remoteness nature of the ocean limits in-situ sampling to localised areas, limiting the scale of taxonomy based study.

Today, thanks to the introduction and increased availability of remote sensing information, some of these challenges have been overcome (Zacharias & Roff 2001; Louzao *et al.* 2011). New imagery and data sets are now enabling remote sensing, in conjunction with ecological models, to shed more light on some of the fundamental questions regarding biodiversity hotspot. For example recent studies noted that large assemblages of oceanic predators coincided with areas of frontal systems, that aggregate also plankton species (Olson & Hood 1994; Cotté *et al.* 2007; Kai *et al.* 2009; Cotté *et al.* 2013). In this study I proposed a new biodiversity proxy based on seascape patchiness of dominant phytoplankton types to estimate global marine biodiversity hotspot of primary producers. These hotspots seem to be related to high diversity assemblages of higher trophic level species.

1.2.2 What is biodiversity and why is important

The diversity of Nature has fascinated humankind since ancient times. Several Greek philosophers, among which Aristoteles, Thales, Pythagoras, Heraclitus, Parmenides, and Democritus, reflected on questions such as “Why are there so many kinds?”, “What is the relation of a kind to its individual representatives?”, “Are these kinds arranged in systematic ways?”, and “Why is there order in nature?”. Since the advent of Modern Science, attempting to describe and estimate biodiversity, along with understanding ecosystem characteristics that may determine/influence its patterns, has always been of primary concern for ecologists and conservationists. Darwin *in primis* observed and questioned the origins of geographical differences in the distribution of abundance and type of species (Darwin 1859). Concepts of rare and common species and high and low diversity assemblages emerged. Intuitively, assertion of inequality of species abundances and type, and changes with consequent loss of them have been present in the human thought.

However, defining and quantifying biological diversity, above all over space and time, in a way that allows an explanation for the observable natural patterns, is a challenging task with uncertain outputs. The reason for this difficulty is that biological diversity is a multifaceted concept that cannot be captured by a single parameter, and hence it has been defined and measured in different ways.

The first definition with juridical value appeared in 1992 in the Convention on Biological Diversity (http://www.cbd.int/), presented during the United Nations Conference on Environment and Development in Rio ("Earth Summit"). Here "Biological diversity" was defined as “the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems”.

More widely, “biological diversity” refers to the variety of life at every level of complexity of its organization, from molecules to biomes (defined as contiguous areas with similar climatic conditions on the Earth, such as communities of plants, animals, and soil organisms) (Wilson 1992; Gaston 1996a; Mooney 2002; Hamilton 2005). In general it can be interpreted as the complexity of a system. Therefore evaluating biodiversity gives indications on the state of the ecosystem in a broad sense (Chesson 2000; Hamilton 2005).

Nowadays, the need for attention on this aspect of biological studies is even more urgent, due to the increasing loss, at an alarming rate, of diversity at all levels of the ecosystem (Myers *et al.* 2000; Myers & Worm 2003; Butchart *et al.* 2010; Mora *et al.* 2011; Cardinale *et al.* 2012; Hooper *et al.* 2012). Indeed, evidence is increasing that biodiversity influences productivity, ecosystem processes and services and its stability and resilience, underlining the importance of characterization and preservation of diversity-rich environments (McGrady-Steed *et al.* 1997; Peterson *et al.* 1999; Tilman 2000; Cardinale *et al.* 2006; Worm *et al.* 2006; Cadotte *et al.* 2011; Flynn *et al.* 2011).

1.2.3 What do we know about global patterns of biodiversity in the open ocean

Robert May noted that if aliens visited our planet, one of their first questions would be, "How many distinct life forms (species) does your planet have? He also pointed out that we would be "embarrassed" by the uncertainty in our answer underlying our limited progress with this research topic thus far"(Hamilton 2005; Mora *et al.* 2011).

In summary, we know very little about biodiversity in the open ocean, this is primarily due to the extent and remoteness of the oceanic environment coupled with the difficulty in acquiring detailed information about the whole marine community, particularly over a spatio-temporal scale that can be useful to clarify the contribution of environmental processes in determining biodiversity patterns.

Recent predictions indicate that at least 8.7 million species of eukaryotes are found on Earth, including 2.2 million marine species, of which only about 9% have been taxonomically classified (Mora *et al.* 2011). However, these are luckily to be only underestimations, as the more recent phylogenetic studies have revealed an overall underestimated biodiversity in most groups of oceanic species. In particular, higher diversity has been detected in planktonic species, which questions the traditional morphological approach applied to define species in this group.

Limited information is also available on the distribution of biodiversity and what influences its patterns in the open ocean (Tittensor *et al.* 2010). Previous studies about global patterns of ocean biodiversity found general relationships with latitude, temperature, energetic gradients (areas of sudden variation in flow direction or strength) and with environmental variability (so-called

macro-ecological patterns) (Gaston 2000). The general rule of life on earth is a decreasing diversity with increasing latitude (Longhurst 2010). In several taxonomic groups a so called "latitudinal species diversity gradient" has been observed (Rombouts *et al.* 2009; Tittensor *et al.* 2010; Berke *et al.* 2014). This pattern was also reported for plankton (Yasuhara *et al.* 2012), although several exceptions exist. Indeed, the latitudinal gradient may vary once other variables such as for example longitudinal shifts, depth and topography come into play.

Temperature in particular appeared to be one of the most important variables in describing biodiversity patterns (Rutherford *et al.* 1999). The temperature hypothesis postulates that: "higher temperatures increased metabolic rates may promote higher rates of speciation leading to greater diversity, or that range limits are set by thermal tolerance, with more species tolerant of warm conditions" (Rohde 1992; Currie *et al.* 2004; Allen *et al.* 2007). However, the positive correlation between diversity and temperature has not been observed within all species and regions (Yasuhara *et al.* 2012). Indeed, at local level also the effect of temperature on biodiversity may lose value for more important physical and biological forces. In these areas, for example, high biodiversity values may be better explained by other more important ecological drivers such as turbulent features, nutrient availability and stratification. Upwelling regions are among the most classic examples; here cold waters rich in nutrients are brought to the surface layers by local physical dynamics, creating indirectly areas of high biodiversity (Barton *et al.* 2010).

Finally, a no-less important mechanism which may explain local variation in species diversity is dispersal, commonly defined as the active or passive movement of individuals from its birth site to a new area (Clayton *et al.* 2013; Levy *et al.* 2014). However, dispersal is a very complex mechanism which may result either in a declining diversity, through direct competition for habitat and resources, or in increasing diversity for those communities whose composition is driven by species interactions and competition (Cadotte 2006).

In summary there is no single mechanism that can adequately explain any given biodiversity pattern. Biodiversity patterns may be affected by different forces at different spatio-temporal scales, larger scale processes may affect smaller ones, and finally variability is at the base of biodiversity.

1.2.4 Which kind of biodiversity to measure

In the previous section I described what biodiversity is, and highlight some knowledge gaps. The first step needed in order to proceed further, is to pin the concept down. We cannot even think of superficially carve the mystery of biodiversity and even less understand accurately how much and fast are we loosing biodiversity, if we cannot measure it. Several indices, metrics, algorithms and models have been proposed as tools to study biodiversity. However, any effort to measure biodiversity has to rapidly face a common problem: what kind of biodiversity do we want to measure?

Today, the increasing number of definitions available in the scientific literature is probably the only increasing component of biodiversity. The concept of "species richness" as basic arithmetic count of the number of species found in a specific area, still remains the more simple and pure facet of biodiversity. For many years taxonomy has relied only on the use of morphological features to identify and distinguish different species. However, as molecular and microscopic techniques became widely available, it became evident that morphological features alone cannot always guarantee a correct classification at species level. In some cases for example, such as for cryptic species (species morphologically identic but reproductively isolated) morphological features alone cannot guarantee a correct taxonomic distinction. On the other hand, phylogenetic techniques are not always applicable.

To overcome this problem, biologists around the world have found different strategies to explain observed biodiversity patterns. For example in microbial studies, where the classification at species level is highly questionable, higher-taxon or trait-based classifications are commonly used as proxy of overall diversity (Green *et al.* 2008). The main levels of biodiversity traditionally considered are: morphological, phylogenetic, molecular and functional diversity. These levels however are not totally separated but rather one level may depend on the diversity of the other (Lankau & Strauss 2007).

Morphological diversity (assessment of diversity based on differences in physical features) is the most straightforward, it is often used to identify taxonomic classes, such as species, or to identify some functional traits used to classify functional groups. When differences or limitation exist in the methodologies applied to identify certain physical features, morphological diversity becomes a misleading method to estimate biodiversity of certain communities such as plankton. In these

cases diversity must be assessed at genetic level.

Phylogenetic diversity derives from modern phylogenies that use DNA sequence data and explicit evolutionary relationships among taxa. The diversity is based on the evolutionary distance between the taxa, which is the time that each taxa has evolved independently. This distance represents a proxy for the magnitude of phenotypic differences between any two taxa (Cavender-Bares *et al.* 2009). This kind of diversity has received great attention for conservation and community ecology studies. It also gives insights on the processes determining the structure and assemblage of the community (e.g. communities constrained by competition interactions are more likely to be composed by distant related taxa while communities constrained by tolerance to environmental conditions are more likely to be composed by closely related taxa). However, as it is essential to investigate the greatest possible variety of biological features, the preferred approach would be to support evidence in phylogenetic diversity with morphological distinctness (Vane-Wright *et al.* 1991). In some cases, such as for microbic communities, the combination of genetic and morphological analyses still does not guarantee a correct taxonomic classification. That is why for microbial communities molecular or functional diversity are most used.

Molecular diversity (definition of taxa on the basis of DNA or RNA characteristics or biochemical compounds) is the finest and most critical level of diversity because it is the origin of all diversity (Lankau & Strauss 2007). Molecular methods are an advantage when organisms are rare, cryptic or difficult to identify (Appeltans *et al*. 2012*)*. New high throughput techniques allow to use metagenomics and DNA barcoding to study biodiversity of all environments (Hingamp *et al.* 2013). These techniques detect DNA sequence variations in particular regions of mitochondrial, chloroplast and nuclear DNA depending on the resolution required, as markers. Nuclear genes are more conserved compared to the others and have a slow evolutionary rate, therefore they are used to reveal more ancient evolutionary processes (Simon *et al.* 1994). The most rapidly evolving sequences are non-coding regions of DNA, used for instance for population studies. For protist and bacterial diversity, ribosomal genes are used (Zehr 2011).

Functional diversity (Cadotte *et al.* 2011) is based on the degree to which coexisting species vary in terms of their functional traits. It emphasizes the phenotypic difference among taxa while discounting phylogenetic relatedness, even if trait diversity is often closely associated with phylogenetic diversity. It affects species performances and ecosystem functioning (Chapin III *et*

al. 2000; Garnier *et al.* 2007). Its influence is primarily expected through complementarity or enhancement of ecosystem processes caused by increased efficiency and specialisation of resource use by organisms with a high degree of trait dissimilarity. Experimental, theoretical and observational studies show that the maintenance of ecosystem processes depend primarily on the functional diversity rather than the overall diversity (Hooper et al. 2005).

1.2.5 How do we measure biodiversity: Indexes, Rank Abundance Distributions (RADs), Q-matrixes

As pointed out in the previous paragraph, identifying meaningful measures to evaluate biodiversity is essential for detecting changes that could pinpoint acting ecological processes. In this regard diversity analysis in ecology has always been a highly debated field with many different ideas on how to numerically characterize biological diversity (Magurran & McGill 2011). Multiple methods and biodiversity index have been used and introduced in ecological studies (Magurran & McGill 2011). Generally all these methods have been tested and advantages and limitations clearly described on either empirical or theoretical grounds, however recommendations of experts differ in describing which method to use. The choice of the method to be used depends on the aspect being investigated. Biodiversity can be described in terms of numbers of entities (e.g. how many genotypes, species, or ecosystems), the evenness of their distribution, the differences in some of their characteristics, often related to functional traits, and their interactions (Magurran 2004). The best approach, to have a complete understanding of an ecosystem, is to test different measures of diversity.

The main metrics used in ecology to study biodiversity can be classified depending on the amount of information we are able to detect and on the type of investigation that can suit. Following an ascending order of the amount of information on the community captured by the method we can use indices, rank abundance distributions (RADs) and ecological distances (Q-matrices).

Indices differ primarily in the importance they give to the number of categories and their abundance (Magurran 2004). The categorization can include not only species, but it can reflect guild composition, trophic structure, functional diversity, phylogenetic diversity, molecular diversity and morphological diversity (Magurran 2004). As meaningful aspects of complexity,

involving interactions among individuals and between individuals and their environment, depend on population abundance, the frequencies of the different classes are considered when processes determining the structure and functioning of an ecosystem are investigated. Diversity indices establish equivalence relations among communities, depending on the aspect of compositional complexity measured. Among the diversity indices that consider both species richness and abundance the most used one are the Shannon Entropy Index (Shannon 1948), where species are weighted by the logarithm of their abundances, and the Simpson Index (Simpson 1949), used primarily to quantify the biodiversity of an habitat by taking into account the number of species present, as well as the abundance of each species (Table 1).

Table 1 Summary of the most common diversity indexes and characteristics

Index	Equation	Description
Richness Metrics		
Richness (S):	$S = \sum_{i=1}^{R} n_i$	Total number of species (or other categories) identified in the samples.
Chao Estimated Diversity	$S_{chao} = S+S_1^2/ 2S_2$	Allows to compare species richness (S) between sites with different sample sizes. S_1 and S_2 are singletons and doubletons respectively.
Diversity metrics		
Shannon-Wiener diversity	$H_{Shannon}= -\sum_{i=1}^{R} p_i \ln p_i$	Shannon's information theory can be used to calculate the information of a community as an estimate of diversity.
Simpson diversity	$H_{Simpson} = \sum_{i=1}^{R} p_i^2$	Gives the probability that two individuals drawn at random from an infinite community would belong to the same species.
Linearized Shannon Index	$H_{expSh} = \exp (H_{Shannon})$	The rate of change is a linear function; allows to directly compare changes among communities.
Taxonomic distinctness	$\Delta= \left[\sum\sum_{i<j} \omega_{ij}\right]/[n(n-1)/2]$	Describes the average taxonomic distance (ω) between two randomly chosen organisms through the phylogeny of all the species (n) in a data-set.
Q-diversity Statistic	"inter-quartile slope" of the cumulative species abundance curve	A bridge between the abundance models and diversity indices. Provides an indication of community diversity. No weighting towards very abundant or rare species.

Evenness metrics		
Shannon evenness	$E_{Shannon} = H_{Shannon}/ln(S)$	Diversity (H) is a mixture of richness and evenness, therefore removing richness (S) should produce evenness.
RADs (β)	β = slope of log abundance vs. ascending rank	The slope of the Rank Abundance Distribution of the community.
p_i is the real population frequency of the i-th species, with *i* ranging from 1 to S, which is the total number of species present.		

RADs are a representation of the community composition indicating its species diversity and abundance. Assuming there are S species at one site n, with n = (n1, n2, . . . , nS), n_k $(1 \leq k \leq S)$ is the relative abundance of the k^{th} species at this site, ordered from the most to the least abundant one (Wilson 1992). The advantage of this approach is that it is applicable to all environments (Gaston 1996a). RADs allow to compare samples taken from geographically separated locations that have few or no species in common. RADs are important because they take into account the community structure and therefore the type of abundance relationships between species.

A Q-matrix is a squared symmetric distance matrix indicating the distance between samples, calculated from a n x p matrix indicating the abundance of each species (n) per sample (p). They are most used to estimate changes and rate of changes of both type of species and abundance among communities under certain spatial and temporal intervals. The most used in ecology is the Bray-Curtis Dissimilarity Index (Bray & Curtis 1957) which allows to describe modal relationships and estimate dissimilarity in community composition based only on taxa that occur at least in one sample.

In contrast with indexes, RADs and Q-matrices allow to apply quantitative analyses to study the shape of the community. It can be related to processes acting in the environment, species interactions, etc. The distribution of abundant species (common and intermediate) is the first mark of characterization and the one less subject to sampling issues, therefore the most studied in classical ecology. However, advances in technology allow to define RADs as characterized by very long tails, indicating that the rare biosphere is a very important part of the community, especially for microbial assemblages, likely influencing the distribution of the abundant classes in spatio/temporal successions (Purvis & Hector 2000).

Abundance distribution models have been developed to classify community structures based on

hypothesis on environmental resource partitioning. Geometric and logseries distributions (Motomura 1932; Fisher *et al.* 1943) are typical of species-poor communities characterized by minimal cooperativity where the community is structured by one or few factors and there are only one or few dominant species. Log-normal distributions (Preston 1948) are instead typical of large, mature communities where sequential breaking of empty niche space through ecological or evolutionary processes have created rare, intermediate and common classes (Sugihara 1980). Brocken stick (or Dirichlet with a=1) (MacArthur 1957) distribution hypothesizes a random niche boundary with no real relationship between original species diversity and the size of the habitat after subsequent arrivals. This is typical of narrowly defined communities of taxonomically related organisms. The abundance distribution model of the community under study can influence the value of diversity indices making them useless in certain conditions. It is therefore important to understand which kind of community we are investigating and avoid the most biased indexes.

For more information about the index applied in this study, the Shannon-Wiener index, RADs, and Q-matrices (Bray-Curtis dissimilarity index) refer to Chapter 2 (Materials and Methods).

1.2.6 Introducing the concept of functional types to study plankton community ecology

Several studies have now demonstrated that simplifying biodiversity, to a level manageable by available mathematical models, by using the concept of "functional types", is an acceptable and robust approach in global oceanographic studies. Furthermore there is an increasing evidence that functional diversity is more important in the maintenance of ecosystem functioning than species diversity (Cadotte *et al.* 2011).

Plankton can then be subdivided based on common morphological and physiological traits into functional groups, i.e. group of species that, irrespective of taxonomic relatedness, share similar functional traits (Mora *et al.* 2011). A functional trait is "a defined, measurable property of organisms, usually measured at the individual level, and used comparatively across species" (McGill *et al.* 2006). Plankton functional type (PFT) based models are the most recent in a series of coupled ocean-ecosystem models developed to achieve a deeper understanding of ocean biogeochemistry.

Firstly, plankton can be divided into three general groups: bacterioplankton, autotrophic phytoplankton and heterotrophic zooplankton. However, the most common traits used to define functional groups in plankton is the size (Fig. 1). Based on size, plankton can be distinguished in picoplankton (less than 2 µm in diameter), nanoplankton (2-20 µm), microplankton (20-200 µm), mesoplankton (200-500 µm) and macroplankton (more than 500µm) (Karsenti *et al.* 2011).

Plankton can be further subdivided based on the length of the planktonic life stage in: holoplankton, meroplankton and tychoplantkon. Holoplankton comprises organisms whom entire life cycle is planktonic, such as marine protists. Meroplankton includes organisms that spend only a portion of their life or life cycles in the plankton, such as planktonic larvae of benthic invertebrates, chordates and crustaceans. While Tychoplankton is composed by demersal zooplankton or even benthic diatoms that can be periodically inoculated into the plankton by bottom currents, waves and bioturbation.

Finally, phytoplankton can also be distinguished by their biogeochemical roles, not considered under the size-only approached in: Nitrogen-fixers, Silicifiers and Calcifiers (Nair *et al.* 2008). Nitrogen-fixers are characterized by the ability to fix atmospheric nitrogen and have therefore an important impact on the nitrogen cycle and climate change. In the ocean they are represented by a variety of organisms among which the most abundant group is that of cyanobacteria (Tyrrell 1999; Monteiro *et al.* 2010). Silicifiers are united by the use of silica to form their cell theca and include four taxonomic groups: Chrysophyta, Silicoflagellates, Xanthophyta and Bacillariophyta (Brownlee & Taylor 2002). Among the Silicifiers, diatoms (Bacillariophyta) are the dominant. While Calcifiers or Coccolithofores are characterised by the presence of external plates, called coccoliths, made of calcium carbonate (Nair *et al.* 2008).

The major phytoplankton taxa can also be grouped using their pigment contents (chlorophylls a, b, c and carotenoids) as functional traits (Roy *et al.* 2011) (Table 2). Pigment content can be determined using the High Performance Liquid Chromatography for in vitro measurements (Uitz *et al.* 2006).

Table 2 Major marker pigment used for the classification of phytoplankton groups. The most common used pigment alga-class association are in bold.

Pigment	Taxonomic significant
Fucoxanthin	**Diatoms**, Prymnesiophytes, Chrysophytes, Dinoflagellates
Peridinin	**Dinoflagellates**
19'-Hexanoyloxyfucoxanthin	**Prymnesiophytes**, Chrysophytes, Dinoflagellates
19'-Butanoyloxyfucoxanthin	**Pelagophytes**, Prymnesiophytes
Alloxanthin	**Crysophytes**
Chlorophyll-b	**Chlorophytes**, Prasinophytes
Divinyl-Chlorophyll-b	Prochlorophytes
Zeaxanthin	**Cyanobacteria**, Chlorophytes, Prasinophytes, Crysophytes, Eunglenophytes

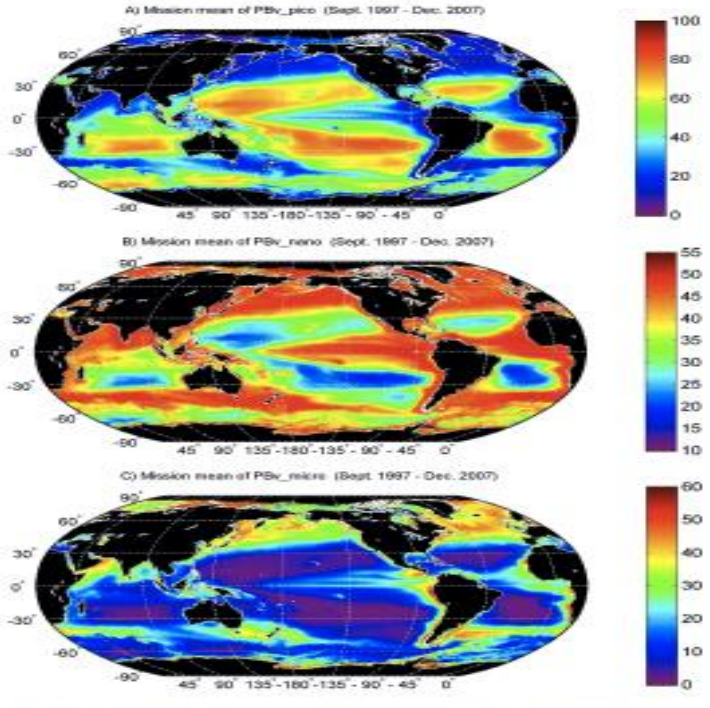


Figure 1. SeaWiFs mission mean (Sep 1997-Dec 2007) maps of the PFT's: (A) the percent volume concentration contribution of picoplankton-sized particles (0.5.2 μm in diameter), (B) the percent volume concentration contribution of nanoplankton-sized particles (2-20 μm) and (C) percent volume concentration contribution of microplankton-sized particles (20-50 μm). The percent contributions were calculated out of the total volume (0.5-50 um diameter range) as a function of the PSD slope ε. The figure and caption is taken from (Kostadinov *et al.* 2010).

1.3 Plankton as keystone component in the functioning of the marine ecosystem

1.3.1 Overview of plankton ecology and biogeography

Marine biological communities can be broadly distinguished into two functional groups: nekton and plankton. The term nekton includes all marine organisms that can actively move against the current. Whereas the term "Plankton" was coined to identify a highly diverse group of marine organisms such as viruses, bacterioplankton, phytoplankton and zooplankton (Fig.2) united by the small size and inability to swim against the current. As a result the movement of plankton in the water is primarily driven by transport processes such as advection, turbulence, upwelling and mixing. The large majority (45.5%) of the plankton community is composed by nanoplankton, followed by picoplankton (43.6%) and microplankton (10.9%) which is mainly found in the mid and high latitudes (Samuelsson *et al.* 2002).

Plankton biogeography is controlled by the physical, chemical, and meteorological characteristics that shape ocean ecosystem dynamics. As a result plankton spatial distribution is patchy, with patches ranging from few meters to several kilometres. Patchiness depends on advective effects of water movements, physical-chemical boundary conditions, grazing and reproductive rates of the communities. Among these factors, ocean currents have been identified as one of the major mechanisms defining plankton distribution and linking biogeographical regions by influencing carbon, nutrient and primary production, eggs and larvae and fish populations by determining the genetic flow. Mesoscale physical structures for example are able to generate environmental heterogeneity that may persist long enough to drive planktonic species distribution and biodiversity (d'Ovidio *et al.* 2010; Levy *et al.* 2014), often resulting in large pelagic assemblages (Alvain *et al.* 2005; Longhurst 2010). This mechanism is one of the proposed processes that concur to explain the 'paradox of the plankton' (Hutchinson 1961), where under the competitive exclusion principle (Hardin 1960), the coexistence of so many species using the same few resources would be impossible, but it is observed (Huisman *et al.* 2001).

In the ocean, physical and chemical characteristics of the water masses promote the dominance of certain plankton groups that delineate therefore different communities. Patchiness of the biophysical environment (abiotic plus microbes) varies in the ocean within spatial and temporal scales determined by the response of the demography of the plankton to the physical and

chemical characteristics. Under this view, regions characterized by these features have the attribute of ecotonal zones (regions of transitions between two biomes where two distinct communities may overlap) and transition zones (Sournia 1994), where alloctone populations (population displaced from its original natural habitat), are advected abundantly enough and with sufficient regularity to influence community structure and persist through time. Ecotones occur at multiple spatial scales, ranging from transitions between biomes to local small-scale transitions (Kark & Van Rensburg 2006). Ecotonal diversity is enhanced by additive blending of the different sources and by the emergence of ecotonal specialists. Zoogeographic studies of the North Atlantic revealed a peak in biodiversity resulted from a mixing of different faunal elements due to the major front between the South Atlantic Central Water and the North Atlantic Central Water (Angel 1993). The Gulf Stream region is an example of an ecological transition zone, where local physical structures can act as an interspecific barrier or as a mean of dispersal and distribution extension or again as a disruptor of abundance, distribution and genotypes frequency. The same happens for the southward extension of the California Current, for Kuroshio, Aghulas, Somalia and Easter Australian Currents. Patchiness by eddy formation is also favoured by the fast-moving equatorial recirculation systems (Longhurst 2010).

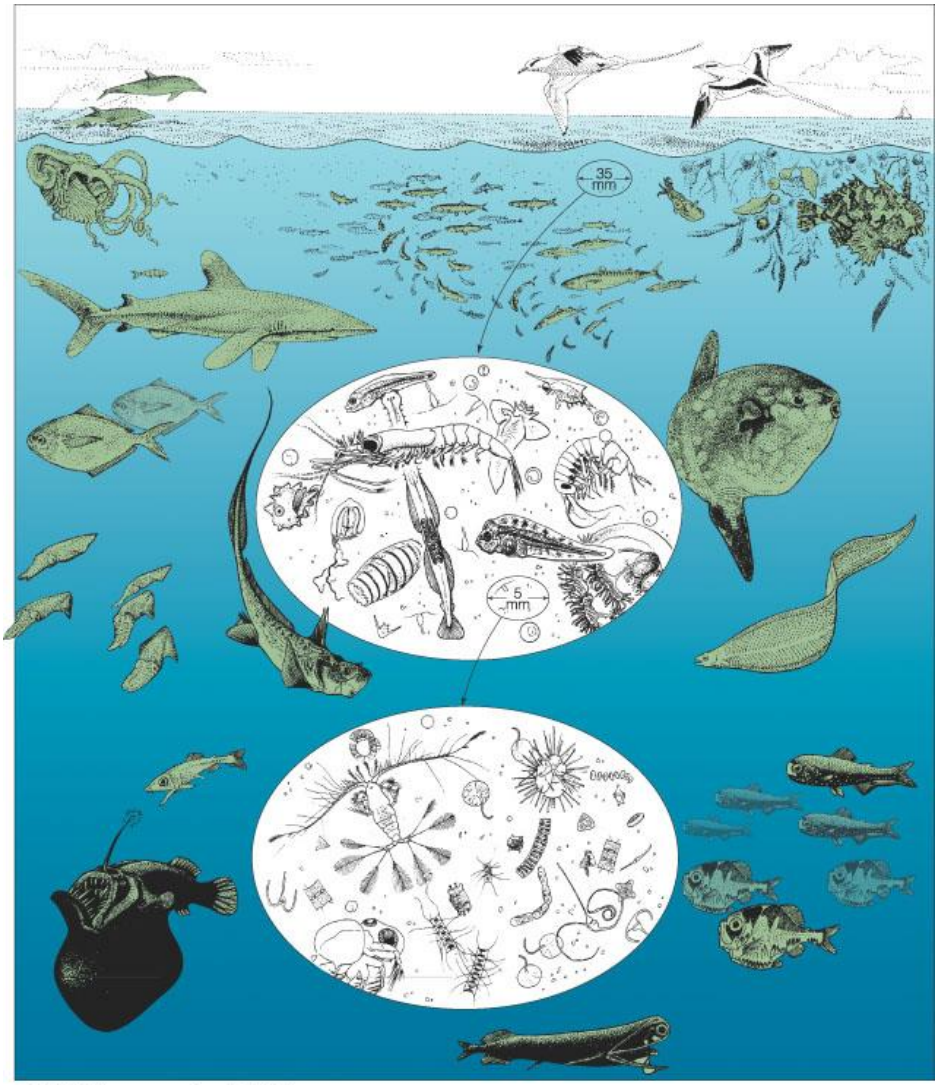


Figure 2. Graphic representation of the interaction between plankton and the biotic part of the marine ecosystem. This thesis focuses mostly on the phytoplanktonic component.

1.3.2 Phytoplankton biology and distribution

This study is primarily focused on phytoplankton and to a smaller extent on autotrophic bacterioplankton, also referred to as primary producers. Primary producers are responsible for almost all (97%) organic matter production in the sea, and indirectly to provide the food resources for the higher trophic levels. Primary productivity is defined as: "the rate at which radiant energy is stored by photosynthetic and chemosynthetic activity of producer organisms in the form of organic substances which can be used as food materials" (Leith & Whittaker 1975).

Primary producers play also active roles in global biogeochemical cycles and climate variability through the export of carbon between the surface ocean and the atmosphere to ocean depths, the transfer of energy and organic matter to higher trophic levels and the drawdown of nutrients essential for primary production (Falkowski *et al.* 1998).

Phytoplankton are the major primary producers in the marine ecosystem and include a number (on the order of tens of thousands) of species (Jeffrey & Vesk 1997) which can be broadly divided in: diatoms, dinoflagellates, coccolitophores, silicoflagellates, cyanobacteria and green algae (Table 3).

Phytoplankton take inorganic materials, such as nitrogen and carbon, and convert them into biomass via photosynthesis. Since photosynthesis relies on light, phytoplankton distribution is limited to the superficial euphotic layer, where sun light can penetrate. In the ocean, this layer roughly corresponds to the depth of the mixed layer (varying seasonally with latitude), where the motion of particles is driven by an intense turbulent mixing.

Table 3. Summary of the major phytoplankton groups found in the ocean

General Groups	Description
Diatoms	Diatoms belong to the Phylum Ochrophyta, Class Bacillariophyceae and are distinguished in Centrales and Pennales. They vary in size between from few microns up to 1 mm in length. Dominate phytoplankton communities in high latitudes, in the neritic zone of boreal and temperate waters and in areas of upwelling. They consist of single cells or cells chains with an external rigid silicate skeleton (frustule) which encases the vegetative protoplast.
Dinoflagellates	Dinoflagellates belong to the Phylum Dinophyta, Class Dinophycaea, and depending on the presence or absence of cellulose plates, they are distinguished in Peridiniales and Gymnodiniales. They vary in size between 5 μm to more than 200 μm. Widely distributed, dominate phytoplankton communities in tropical and subtropical environments and in temperate and boreal autumn assemblages. Unicellular biflagellates mainly autotrophs but are presents also heterotrophic and symbiotic species. Some species produce toxins and are able to create massive mortality events (red tides).
Coccolitophores	Coccolitophores belong to the Phylum Haptophyta, Class Coccolithophyceae. Range in size from 5 to 100 μm. Mostly autotrophs, few taxa are heterotrophs in the aphotic zone.

	They generally show maximum abundance in tropical and subtropical open ocean waters. They are unicellular biflagellates algae covered by calcareaous plates, embedded in a gelatinous sheath surrounding the cell. Coccolithophores reproduce asexually through binary fission.
Silicoflagellates	Silicoflagellates belong to the Phylum Heterokontophyta, Class Dictyochophyceae. Unicellular, uni or biflagellates less than 30 μm in diameter. Secrete an internal skeleton of siliceous spicules. The majority are photosynthetic but some heterotrophic taxa are present. Can predominate in temperate waters but are most abundant in cold nutrient rich environments. Asexual reproduction.
Cyanobacteria (blue-green algae)	Cyanobacteria belong to the Phylum Cyanophyta. At lower level the taxonomy is still highly debated. Cyanobacteria include prokaryotic organisms, less than 1 μm, but can form bigger chains. They are photosynthetic organism characterised by chitinous walls. Frequently found in shallow nearshore tropical seas. Asexual reproduction and spores.
Green algae (Chlorophyta)	Green algae is the most diverse group of algae. Chlorophyceae are one of the classes in which green algae are divided. They come in a wide variety of shapes and forms including unicellular, filamentous or colonial algae. They are found in both flagellated and non-flagellated forms, generally less than 70 μm in diameter. They are commonly found in estuaries and enclosed seas.

Although with geographic and temporal variations, primary productivity displays an overall pattern of higher values in upwelling and shallow temperate waters, due to greater nutrient supply, and lower values in tropical regions. On a local scale the centres of ocean gyres with characteristically depleted nutrients have generally low productivity. In these regions a stably stratified thermocline precludes the vertical exchange of water in the ocean gyres, thereby preventing the replenishment of nutrients for phytoplankton in the photic zone. On the contrary, upwelling areas, where nutrient-rich deep cold waters replace on the surface the warmer and generally poor surface waters, are areas with high levels of primary productivity and thus rich in higher trophic level species (Chavez *et al.* 2011). Further, analysis of monthly climatologies show that intensive blooms are not stable but rather may occur only in specific periods and regions throughout the year (D'ortenzio *et al.* 2012). For example comparisons between January and June data showed that diatoms blooms occur in June (Fig. 3, 4, 5) only along coastal areas within mid and high latitudes. Although, a number of small patches dominated by diatoms can also be found

in open oceans associated with turbulent flows such as eddies.

Light, temperature, nutrients and grazing are recognised as the most important environmental variables controlling photosynthesis and therefore primary production and phytoplankton growth. The photosynthesis is more efficient over certain range of intermediate to low light intensities, while can be inhibited at high light intensities. Despite the direct correlation between light and temperature, temperature dependent only reaction also exists. These reactions involve enzymes, such as RUBISCO (MacIntyre *et al.* 1996) which are responsible for catalysing the conversion of inorganic carbon to organic matter. The photosynthesis rate increases until the enzymes reached optimal temperature and stop when enzymes denaturate. The increasing temperature during spring is also responsible for the formation of a mixing layer that maintains both nutrient and phytoplankton in the euphotic zone therefore increasing productivity (Sverdrup 1953).

Inorganic nutrients like nitrogen and phosphorus, silicon, macroelements and microelements and organic nutrients in both particulate and dissolved forms also control phytoplankton growth (Elmgren & Larsson 2001; Smith 2003), biomass (Cloern 2001; Bledsoe *et al.* 2004) and species composition (Duarte *et al.* 2000; Smayda & Reynolds 2001). Arguably Nitrogen and Phosphorus are the most important elements regulating phytoplankton growth (Tyrrell 1999). Nitrogen is the chief limiting element to primary production in surface waters (Conley *et al.* 2009). Whereas Phosphorus in the form of Orthophosphate ($HPO4^{-2}$ and $PO4^{-3}$) is the principal limiting nutrient of phytoplankton growth in coastal waters under the influence of freshwater discharges (Labry *et al.* 2002).

Primary biomass is also indirectly affected by grazing (Lundry *et al.* 1997). A substantial decline in chlorophyll-a concentration has been observed in some areas, part of this decline was associated with an increase in the grazing pressure of microzooplankton associated to global warming (Chen *et al.* 2012). Additionally, selective feeding by zooplankton can indirectly affect primary production by modifying the composition of phytoplankton community. Biomass growth rates increase from picophytoplankton to medium-sized phytoplankton cells (Chen & Liu 2010). The extent to which each group of zooplankton contribute to grazing vary spatially within the bloom because the response rate of each group is quite different.

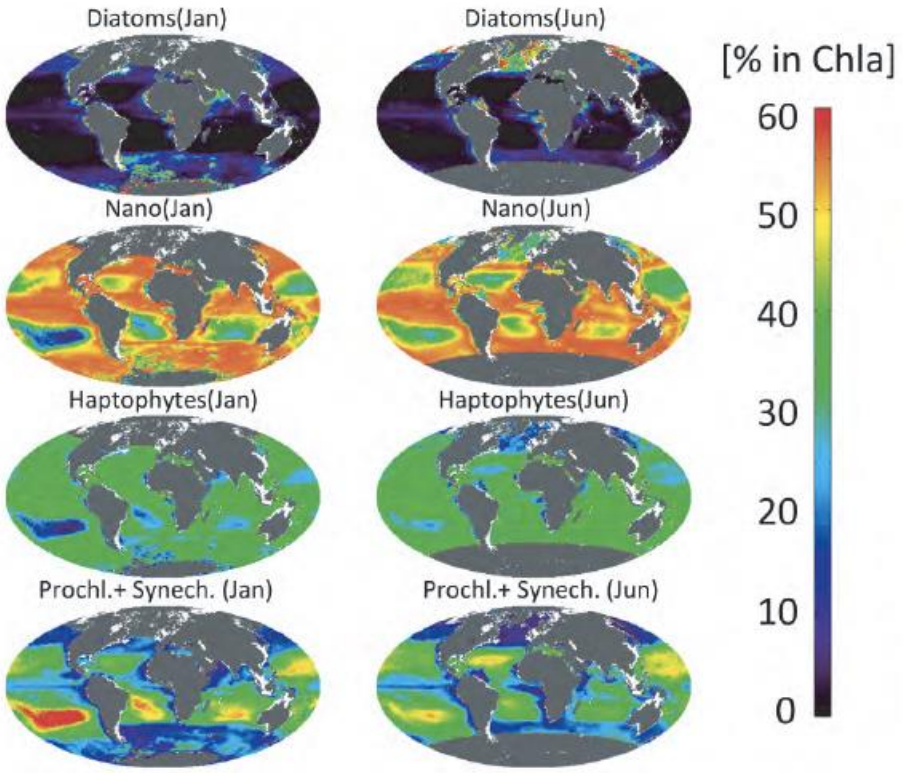


Figure 3. Partial contributions of diatoms, nanoplankton, haptophytes and *Synechococcus* + *Prochlorococcus* derived using the Hirata *et al.* (2011) method with SeaWiFS data from 1998-2010 to obtain the January and June climatologies (IOCCG 2014).

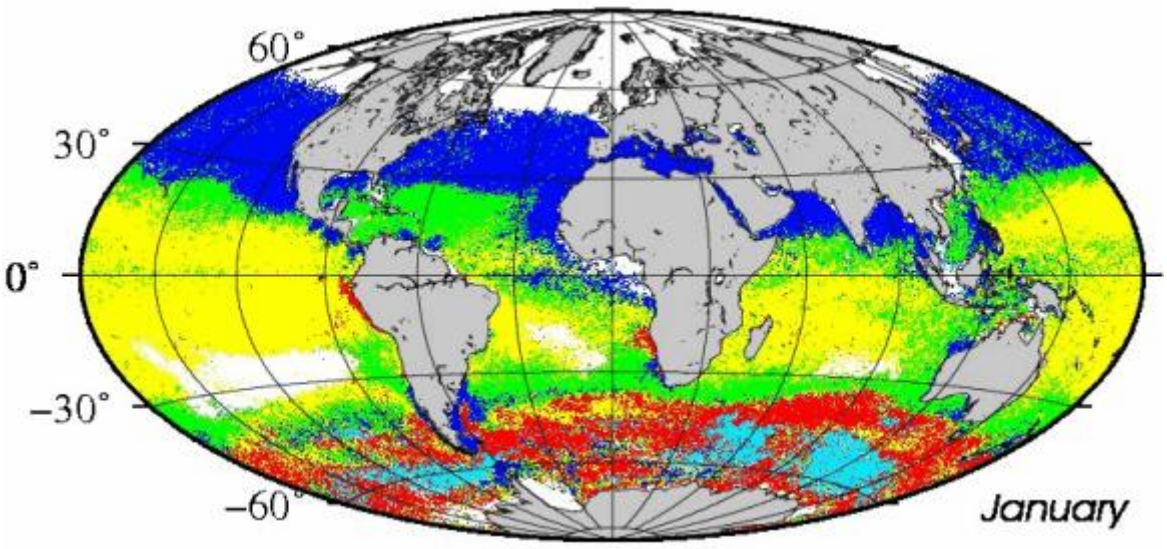


Figure 4. Maps of the dominant phytoplankton group taken from Alvain *et al*. (2005) (haptophytes in cyan, *Prochlorococcus* in green, SLC in yellow, and diatoms in red and nanoplankton in blue) for January 2002 obtained from the standard PHYSAT method of Alvain *et al*. (2005).

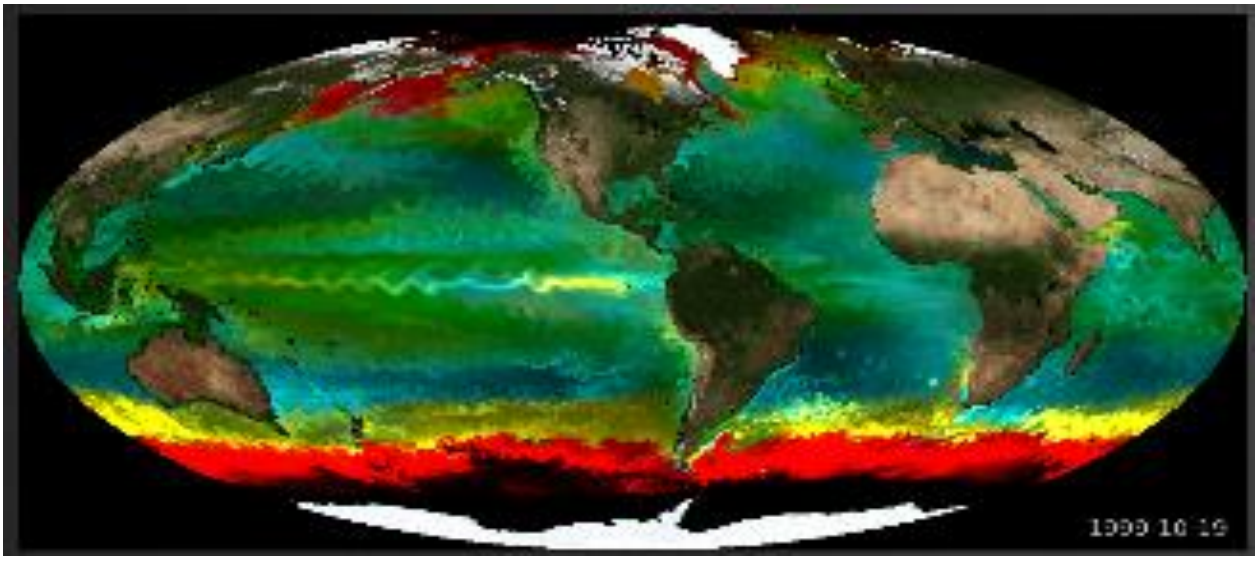


Figure 5. Global map of phytoplankton biogeography as predicted by the ECCO2-Darwin model. In the map are shown 4 functional phytoplankton groups: diatoms (red), *Prochlorococcus* (cyan), large eukaryotes (yellow) and small autotrophs (green). Courtesy of Mick Follows.

1.3.3 Oceanographic structures shaping plankton distribution

The ocean is a complex dynamic system regulated by a number of different processes (physical, chemical, biological, and atmosphere-ocean interactions) which operate and interact at different spatial and temporal scales (Cullen et. al 2002). On the spatial scale, ocean process can be distinguished in large scale (>1000 km), mesoscale (10-100 km) submesoscale (1-10 km) and microscale (cm) (Fig. 6).

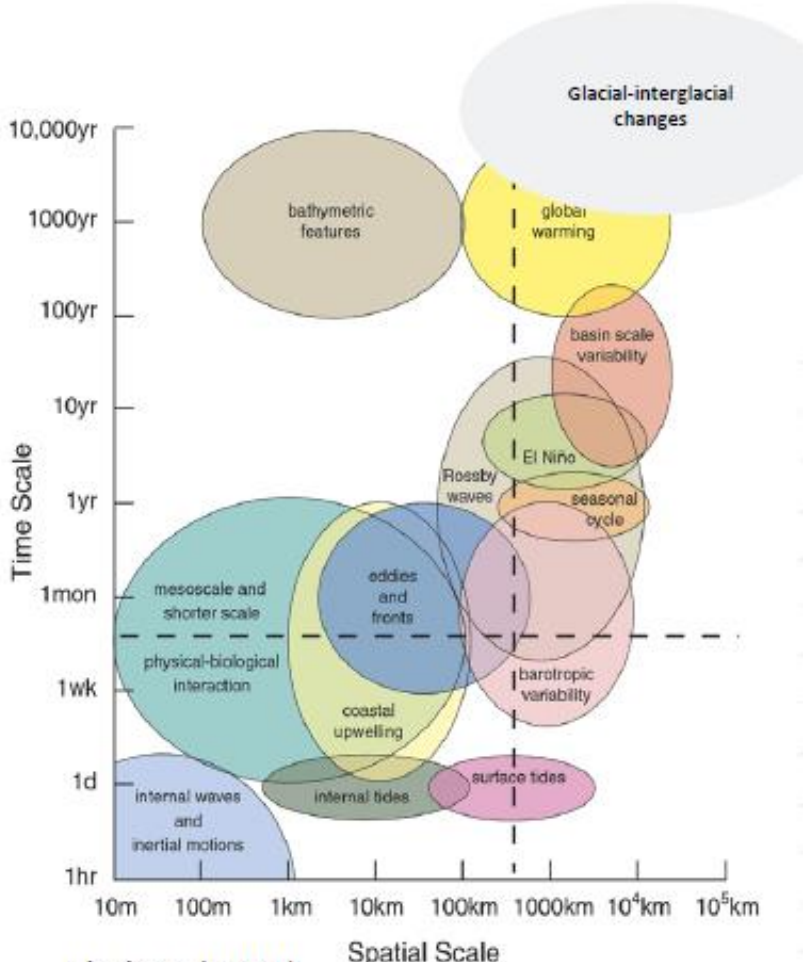


Figure 6. Spatial and temporal scales in the ocean and associated processes from Chelton et al. 2001 (Chelton *et al.* 2001).

The dynamics of the ocean at large scales are highly variable and are driven by multiple forcing factors (Earth motion, solar insolation, winds, tides and freshwater input). The instability of this large-scale circulation leads to the development of meso and submeso scale dynamics such as eddies (Garçon *et al.* 2001), fronts (Belkin 2009) and filaments (Lévy *et al.* 2001; Nieto *et al.* 2012).

At large scale plankton distribution is determined primarily by the distribution of macronutrients and large scale current systems (Fager & McGowan 1963; Pollard *et al.* 2007). At mesoscale eddies, fronts and filaments help to promote the concentration of nutrients (Horne & Platt 1984; Le Fèvre & Frontier 1988). These areas of high energy gradients provide optimal growth conditions for plankton and several other organisms at the base of the food chain (Hernández-

Garcıa & López 2004; Levy *et al.* 2014). Eddies, fronts and filaments play a significant role in ocean-scale biogeochemistry (Peterson et al. 1999; Lapeyre & Klein 2006) by affecting a range of biological processes, mostly related to the transport of organisms in the waters such as distribution and dispersal trough the physical processes of advection, diffusion, stirring and mixing.

Eddies are physical hydrographic features widely distributed across the oceans, with radius ranging from tens to hundreds of kms, characterised by strong vorticity and with lifetime of few weeks up to more than one year. These structures are generated through barotropic and baroclinic instabilities of the mean currents at mesoscale and account for a large portion of ocean turbulent kinetic energy (McWilliams 1985; Stammer 1997).

Eddies are made of a core and a periphery, these regions are characterised by different dynamics and therefore have substantially different effects on the dispersion of fluid parcels (Swearer *et al.* 1999). The core of eddies is associated with areas of regular motion that act as a sink for the particles for time comparable with the eddy lifetime. Whereas peripheries are associated to a more turbulent motion with advection being the primary process controlling and promoting the transport and mixing of plankton (Babiano *et al.* 1994). Eddy can be divided in cold-core and warm-core eddies (Hyrenbach *et al.* 2000). Cold-core eddies are characterised by the upwelling of cold nutrient-rich waters in the centre, and downwelling of warm waters at the periphery. As a result upwelling within cold-core rings supports enhanced primary production. On contrary in Warm-core eddies the upwelling of cold nutrient-rich water occurs in the periphery.

In contrast to the eddies observed at the mesoscale, high resolution observations show elongated filaments as the predominant features at the submesoscale (Pascual *et al.* 2010). The size of these structures is about 1 to 50 km wide and up to 100 km long and is extended vertically below the mixed layer. The time scale associated to filaments is of the order of days/weeks. Filaments affect different processes of the ocean dynamics. Horizontally the elongation of water masses in this structure intensifies local gradient enhancing dispersion and mixing (Lapeyre & Klein 2006). But they can also act as a barrier to water movement (Joseph & Legras 2002). Filaments can affect plankton pattern formation through lateral stirring (Lehahn *et al.* 2007).

Fronts are highly dynamic narrows regions which can be identified by the presence of sharp gradients in hydrographic properties such as temperature, salinity, and nutrient concentrations

(Belkin 2009). For example a front could be an interface between two water masses of different temperature or between two current systems. Such interaction produces areas of highly energetic mesoscale and submesoscale activity such as for example downwelling and upwelling flows. As a result fronts, often associated with intense vertical motions, play an important role in the vertical transport of nutrients or plankton. The extent of fronts is dependent on the structure involved in its formation, and can vary from few to hundreds of kilometres laterally and to 10^3 m in depth and persist from few hours to virtual permanence (Owen 1981).

1.3.4 The interaction between phytoplankton and the turbulent dynamics of the ocean

One of the reasons for the inhomogeneity found in the marine environment lies in the presence of strong horizontal advection, associated with mesoscale turbulence, coupled with the spatial inhomogeneity of the nutrient input and the nonlinear dynamics of plankton populations (López *et al.* 2001; Seuront *et al.* 2001). In the ocean the spatial distribution of all marine organisms is driven by the interaction between organism behaviour and physical structure and processes. To what extent behaviour or physical process affect organisms distribution depends on the swimming ability of the species under investigation (McManus & Woodson 2012). For microscopic organisms, such as plankton, behavioural movements have little effect on their larger scale distribution patterns and consequently the motion of the individual organisms and, on a larger scale, the distribution and concentration of entire populations are determined by fluid dynamics (Fig.7).

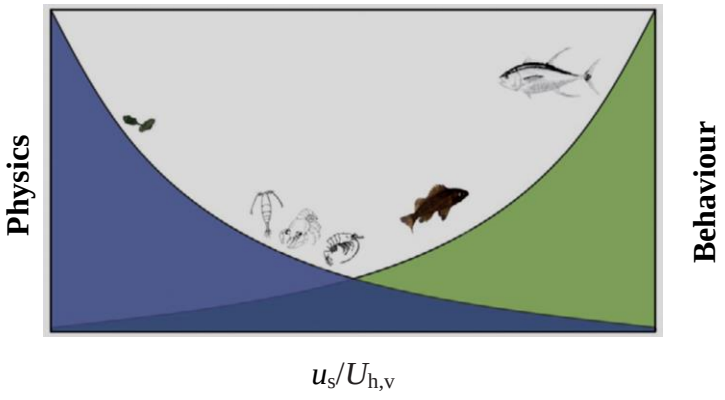


Figure 7. In this figure the influence of physics and behaviour in driving the distribution of organisms is plotted against the ratio of swimming velocity (u_s) to characteristic velocity ($u_s/U_{v,h}$) where $U_{v,h}$ refers to vertical and horizontal flow, respectively. Organism groups are shown in

approximate locations across this continuum. Pictures taken from McManus & Woodson 2012 (McManus & Woodson 2012).

The extent to which behavioural and physical forces affect the distribution of marine organisms can be investigated analysing the Reynolds numbers (Re) and the velocity ratio (the ability of an organism to swim against a flow) (Guasto *et al.* 2012):

Re = $u_s L/v$; *Velocity Ratio* = $u_s/U_{v,h}$,

where u_s is the organism's swimming velocity, L is the organism's length and v is kinematic viscosity and $U_{v,h}$ refers to vertical and horizontal flow. For plankton both the Reynolds numbers and the velocity ratio are typically very small (<<1), therefore in these conditions physics is the dominating force over behaviour (McManus & Woodson 2012) (Fig.7).

In the marine environment the velocity component can be expressed as the sum of the main component of the velocity (U) and its fluctuation term (u^1) or turbulence (Reynolds decomposition). If a) the turbulence is considered homogeneous (i.e. the variance of the velocities does not change in space), b) flow speed is constant ($du/dt = dv/dt = dw/dt = 0$), c) horizontal velocities are larger than vertical ($w << u, v$), d) gravity is the only active force, and e) the resistance of the fluid is small, the motion of fluid in the ocean can be simply described using the geostrophic equation

$$\frac{\partial p}{\partial x} = \rho f v; \frac{\partial p}{\partial y} = \rho f u; \frac{\partial p}{\partial z} = \rho g;$$

where $f = 2\Omega\sin\varphi$ is the Coriolis parameter and ρ is the concentration field.

In reality fluids have a turbulent nature dominated by advection, diffusion stirring and mixing. Advection, diffusion, mixing and stirring are among the primary physical processes involved in the patchiness distribution of plankton in the euphotic layer (Platt 1972) in both the horizontal and vertical dimensions (Abraham 1998). Advection has been identified by many authors as a key process for shaping patterns observed in plankton bloom distributions at meso and submeso scales (Olascoaga *et al.* 2008; Neufeld *et al.* 2010; Pérez-Muñuzuri & Huhn 2010). In simple terms, advection indicates the transport of a particle from one place to another under the fluid motion. Such transport can occur both on vertical, which is involved in the transport of nutrient in

the surface, and horizontal, responsible for the transport of plankton. Diffusion is the transport associated to irregular motions. Diffusion has two primary properties: it is random in nature, and transport is from regions of high to low concentration. Mixing can be described as the combination of advection and diffusion and is responsible for reducing inhomogeneities in the properties of a fluid medium. While stirring produces the stretching of the fluid resulting in strong gradients in the concentration field enhancing the effect of the diffusion and the mixing. The interplay of stirring and mixing is important in the ocean for temperature, salinity, chlorophyll and other natural tracers.

In this turbulent environment the movement of organisms in the fluid (in two-dimensional turbulence) can be described by the reaction-diffusion-advection equations with the reaction terms representing the biological interactions:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{u}\nabla\rho = \boldsymbol{f}(\rho) + D\nabla^2\rho,$$

where, $\boldsymbol{u}$ the velocity vector field, D is the dispersion coefficient (assumed constant for all the components of the concentration field) and ρ is the concentration field. The vector quantity ρ is the concentration of different scalar fields undergoing biological interactions, and the vector quantity f(ρ) represents the reactions between the various components of ρ (Pasquero *et al.* 2001).

1.3.5 Remote sensing a promising tool to investigate phytoplankton distribution

In this paragraph I briefly described remote sensing, an invaluable tool to investigate phytoplankton distribution at global level.

In situ observation and sampling (from oceanic expeditions, processed by microscopy, flow cytometry, chromatographic and genetic analysis) could never provide information at sufficient spatial frequency and short enough time interval to capture a synoptic view of oceanic features. The recognition of this issue stimulated scientists to find ways to identify oceanic physical structures and biological functional groups using remote sensing.

Remote sensing is generally defined as the use of electromagnetic radiation to acquire information about a specific surface such as the ocean (Gordon & Morel 1983). In remote sensing data, in form of reflected radiation, images and datasets are collected indirectly with the aid of remote sensors mounted on the satellite. As satellites are not in direct contact with the

environment under study, data must be extrapolated by the energy signal of the emitted and reflected radiations using appropriate algorithms and models. Such data are therefore particularly useful in ocean global studies.

The radiations used in remote sensing are those directly emitted from the ocean and derived from reflected solar radiation which can be recorded only during daylight by passive sensors, and reflected radiation derived from the pulse energy directly emitted by the satellite using active sensors.

Satellite sensors are now able to receive and record a broad range of the electromagnetic spectrum from the visible to infrared and microwave wavelength regions. The visible spectrum only allows to obtain information on colour variation associated to phytoplankton or inorganic suspension up to 100 m in depth but it is limited to daylight hour and cloud free conditions. In the infrared it is possible to obtain information throughout the entire day but relative only to the superficial layer of the ocean and in cloud free conditions. Whereas, in the microwave spectrum the ocean surface can be viewed in all atmospheric conditions.

In the past 40 years the introduction of remote sensing has revolutionized our vision and understanding of the ocean circulation. Today information available from remote sensing such as Sea Surface Height (SSH), Sea Surface Temperature (SST) and "Ocean Colour" provides a global, high resolution description of the sea level, ocean circulation variations and phytoplankton distribution at a resolution capable of resolving most of the mesoscale processes. More detailed information on SSH, SST and Ocean Colour are given in paragraph 2.2.1

1.4 Objectives

Biodiversity patterns are challenging to elucidate for the pelagic ocean realm, and causal effects of environmental processes on biodiversity are even more difficult to disentangle. With this work, I suggest an alternative and new approach to integrate biodiversity studies for the pelagic oceanic environment at the global level. I rely on recently developed knowledge about oceanic ecosystem processes to reduce remarkably the coverage, resolution, detectability and spatiotemporal scale issues common to most of the biodiversity studies.

The primary objectives of my investigation are:

1. Can we describe plankton biodiversity hotspots at the global level?
2. Can we estimate these hotspots based on an ecosystem functioning approach?
3. Which are the drivers of these hotspots?
4. Which are the implications for diversity at higher levels of the trophic chain?

The answers to these questions require an interdisciplinary approach that utilizes the most advanced techniques and available information. That is why I chose to use global ecological and circulation models, remote sensing, and large in situ datasets of species occurrence based on morphological and genomic approaches. This broad application allows us to overcome issues of fragmented and heterogeneous information, combining biological and environmental context at the global level. I follow a 'Study it all' principle with the aim of integrating complementary sources towards a synoptic view of the ecosystem.

I first explore the potential of a new method of estimating biodiversity of primary producers in the open ocean based on fluid dynamical processes, using a global ecological and circulation model. Then I apply this method to remote sensed data to highlight real spatial and temporal patterns of biodiversity and analyse match and mismatch with in situ information of ecological plankton communities. Finally I investigate the potential of bottom up effects of biodiversity trough the trophic chain in terms of biogeographical regions as stable hotpots.

The organization of the manuscript is the following: in Chapter 2 I am going to illustrate the different methods. Chapter 3 is dedicated to the validation of the approach using modelling. In Chapter 4 I present the transfer of the approach on real remote sensed data. In Chapter 5 I extrapolate information about biodiversity hotspots of the whole ecological pelagic community. Chapter 6 presents case studies as perspective works that can be deepen to investigate the genomic biodiversity of plankton and its relation to other methods of estimating biodiversity and its meaning. Finally in Chapter 7 I present a detailed conclusion of the study with suggestions for its applications and insights.

Chapter 2 Materials and Methods

In contrast with the terrestrial realm, marine ecosystems are embedded into a fluid dynamical environment whose characteristic timescales often overlap with the ecological ones. Therefore, a concerted effort in the fields of biology and physics is essential. Novel methodological and technological advancements have increased biological resolution: global circulation models (GCMs), ocean observing systems (OOS) and remote sensing, from the submeso- to the global-scale. At the same time, high throughput -omics techniques permit to explore the depth of the microbial community structure in unprecedented details. This study has exploited both kind of advancements in technology, integrating global models, remote sensing and next generation molecular information about plankton communities to better understand biodiversity patterns in the open ocean. In this section I present a description of each method and dataset used in my thesis.

2.1 Modelling approaches to measure global biodiversity

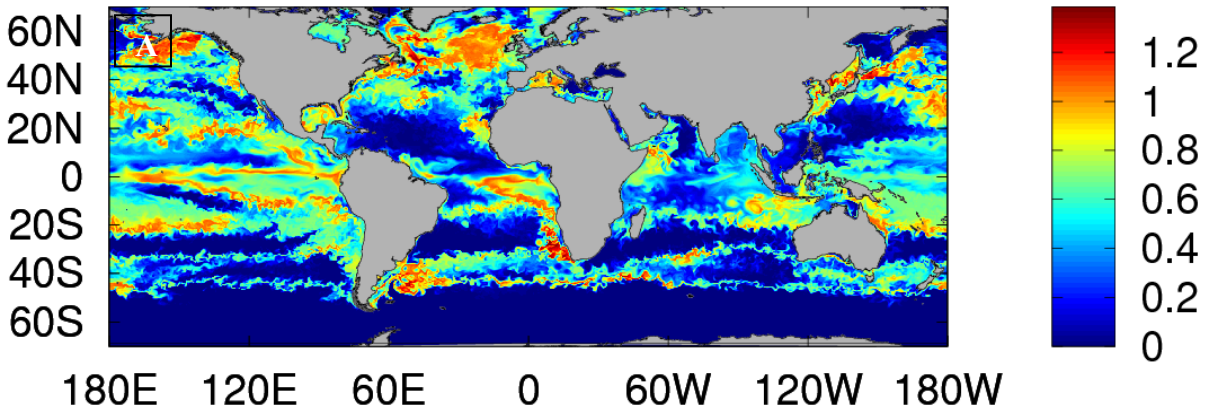
2.1.1 The Darwin model: coupled physical and ecological model of the global ocean

In order to develop a novel measure of biodiversity, I have used a global ecological and physical circulation model, the MIT ECCO2-Darwin. In the context of this thesis, this model can be seen as a virtual environment where the physical and biological state of each point of the ocean is fully known. This model has allowed to validate several proxies of biodiversity against the modelled ground truth in the global ocean, and to study the reliability of these proxies against biogeographical or physical parameters. The peculiarity of the ECCO2-Darwin model is to create a global, high-resolution self-organizing planktonic ecosystem following a function-based approach where the community structure is an emergent property of the ecosystem itself. The Darwin model is based on the Massachusetts Institute of Technology general circulation model (MITgcm) (Marshall *et al.* 1997a; Marshall *et al.* 1997b), coupled with a biogeochemical and ecological component as described in (Dutkiewicz *et al.* 2009). The physics of the ocean circulation is simulated with a grid horizontal resolution of ca 20km, that resolves mesoscale structures in the tropical regions along with subpolar submesoscale features, and 50 vertical layers ranging in depth from 10 m near the surface to approximately 450 m at a maximum model depth of 6150 m. The ecological component includes 78 different phytoplankton types and 2 classes of grazers. The planktonic community structure emerges from a range of possibilities determined by processes of dispersal, competition for resources, predation and physiological characteristics. Physiological parameters are determined stochastically from a broad range of realistic parameters that describe biomass growth as regulated by light and temperature sensitivity and nutrients affinity with allometric constrains. Biomass loss includes a linear mortality, sinking and dispersal and zooplankton predation, which is determined by a Holling type II functional response with prey preferences according to size and palatability. Depending on the species types physiology, 5 main phytoplankton functional groups are recognizable: diatom analogs as large cells that require silica, two *Prochlorococcus* analogs as cells that can or cannot assimilate nitrate, small photo-autotrophs and large eukaryotes.

The model is nominally integrated from 1992 to 1999 and constrained by observations of hydrography and altimetry. All the biological and chemical components are modelled as tracers: organic and inorganic forms of nitrogen, phosphorous iron and silica distributions are initialised

from observed climatologies and plankton types are initialised with the same biomass in each grid. The diversity of the successful population is self-selected during the initial adjustment (spin off) of the ecosystem model. The system that emerges converges to acceptable patterns of surface nutrients, biomass, primary and export production and phytoplankton biogeography (Follows *et al.* 2007), showing that this virtual ecosystem is sufficiently complex to reflect processes analogous to those structuring the real ocean (Fig. 8).

In this study, I consider only the surface layer, which corresponds to the ocean layer that is available to remote detection for comparison with other analysis (see Chapter 4 and 5). In any location of the surface layer, I retrieved from the model the biological and physical parameters on which all the analysis are based: abundance of every virtual species and of every functional group, total chlorophyll (Chl is diagnosed using the approach of Cloern et al. (Cloern & Jassby 1995): the Chl:C ratio is computed as a function of light, temperature and nutrient-limited growth rate for each type at every time step) (µg/C), Sea Surface Temperature (degrees Celsius), nutrients concentration (nitrites, nitrates, ammonium, silicates, phosphorous, iron) and Eddy Kinetic Energy (m/s). EKE is a diagnostic of the mesoscale variability in the oceans and provides information of the turbulent component of the flow. The EKE per unit of mass is given by $EKE = \frac{1}{2}\langle u'^2 + v'^2 \rangle$, where u' and v' are the instant deviations in zonal and meridional velocities from the average over the selected period.



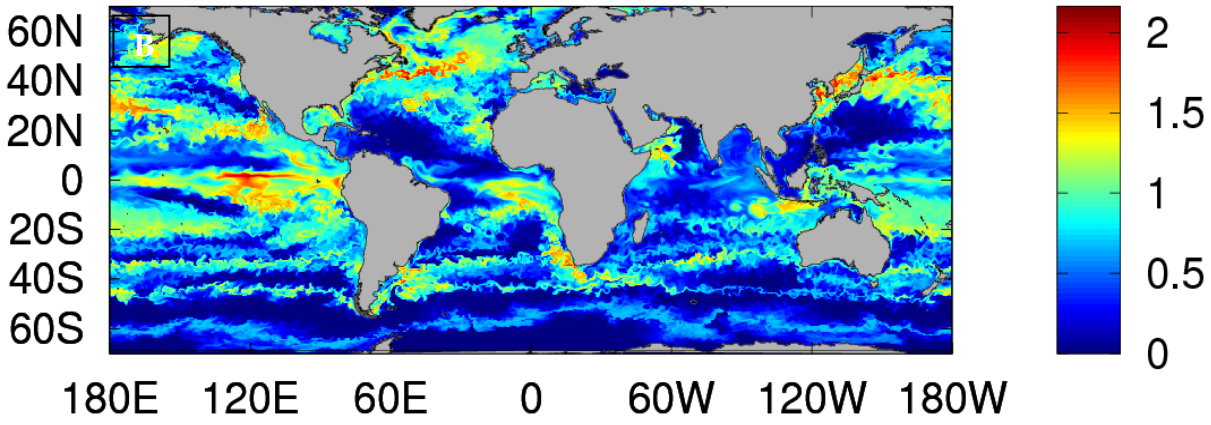


Figure 8. Daily snapshot of local (alpha) biodiversity estimated by the classical Shannon index using A) the functional groups and B) all the 78 phytoplankton types of the ECCO2-Darwin model.

2.1.2 Computation and relationship of the 'local' and 'seascape' diversity in the model

I chose to estimate the biodiversity of the virtual ecosystem using the Shannon-Wiener entropy index $\mathbf{H= -\sum p_i \log (p_i)}$ (Shannon & Weaver 1948), where both the virtual species and the functional groups are weighted by the logarithm of the frequency of their abundances ($\boldsymbol{p_i}$). I used this definition to calculate the classical local (alpha) diversity for each pixel. In my work I have extended this index to a 'seascape' diversity, spatial-based variant of the Shannon index. This novel index is calculated over a constant area of a 1 degree disk radius (ca 100Km) in which the frequencies of the species are substituted by the frequencies of the pixels occupied by each different most dominant phytoplankton group inside the given area (Fig. 9) (see Chapter 2.2). The switch from local abundances (traditional definition) to nearby abundances of only dominant types is an important conceptual and practical change. Conceptual, because it assumes that local biodiversity reflects the patchiness of the nearby communities; and practical, because it requires only information on the spatial distribution of dominant types. Note that the frequency of occurrence of a dominant type reflects spatial patchiness rather than local community diversity.

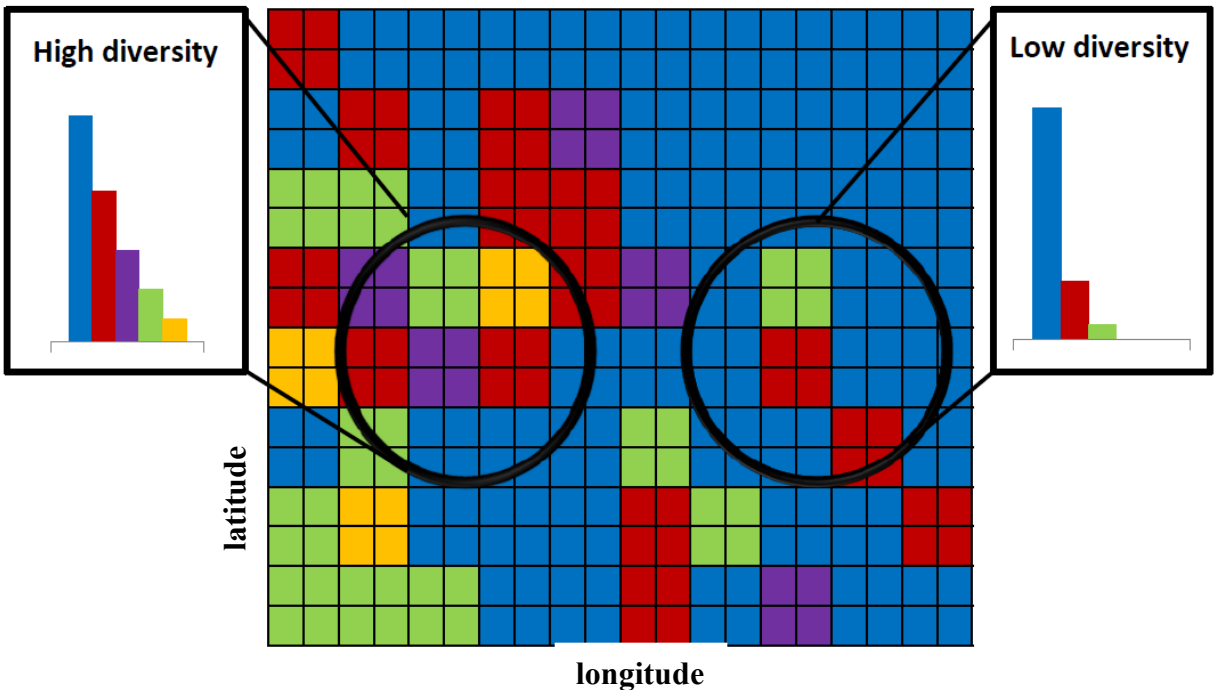


Figure 9. The definition of a novel index of biodiversity has been a central step in my thesis work. This figure provides a schematic representation of how this spatial-based diversity index, proxy for planktonic biodiversity, is calculated. Squares of different colours represent the different most dominant species in that location. Species-abundance distribution histograms show the abundance of the dominant species only, in terms of occupied space in the given area. The more the most dominant species in the communities are heterogeneous over the considered disk area, therefore the species-abundance distributions are broad, the more the index will be high and vice versa, identifying regions of high and low diversity. The black circles which define the region over which the histogram is computed have been optimized to 1 degree (see Chapter 3 for details).

To analyse how seascape diversity relates to measures of local biodiversity I calculated daily local biodiversity patterns based both on phytoplankton types and on functional groups of the model for 1997-1999. I then compared their climatology with annual climatologies of seascape diversity over the same temporal and spatial scale. Quantitative comparisons have been approached using zero-order correlation analysis and regression analysis to determine direction and strength of the relationship. The fitted regression model was chosen to be a non-linear function intercepting the y axis at a very small positive value (as local biodiversity cannot realistically be zero) and reaching the value of potential maximum biodiversity for that system. The maximum potential value of the Shannon index is the logarithm of the total number of species or functional groups in the ecosystem. The regression model combines an initial

exponential increase and a linear tendency towards the maximum possible value of biodiversity after normalization for the maximum theoretical Shannon index reachable by our system for both indexes of local and seascape diversity. The function is : $Y = b1 + b2\ (1-e^{-45x}) + b3\ (1-x)\ (1-e^{-45x})$, w here b1 = 0.06, b2 = 0.39, b3 = -0.31 vs types, $p < 10^{-4}$ and b1 = 0.04, b2 = 1, b3 = -0.80 vs groups, $p < 10^{-4}$.

The choice of the best spatial scale (disk radius) to use on the global analysis have been done using the highest regression fits obtained using radia from 1 to 5 degrees and analysing the global maps of the residuals.

Residuals obtained by regression analysis were studied in relation to the physico-chemical characteristics of the environment: Eddy Kinetic Energy, Chlorophyll, Sea Surface Temperature and its gradient and nutrients. To underline possible patterns, the values of the annual climatologies of the environmental variables where binned following equal intervals. The residuals of the regression analysis were averaged depending on the pixel location having environmental values corresponding to the bins. The bins of the environmental variables were plotted against the averaged residuals. The percentage corresponding to the surface cover of the ocean of the extreme values has been computed dividing by the total number of values representing the whole ocean.

Biodiversity hotspots have been spatially defined, and mismatches between local and seascape diversity compared, using the seascape threshold value that maximizes the difference between hotspots and non-hotspots in local biodiversity.

2.2 Remote sensing approaches to measure global biodiversity

2.2.1 Remote sensing information to describe the marine environment

To conduct ecological studies at local level, in situ sampling is fundamental to assess for instance possible consequences of human activity and climate change, but only in rare occasions can be used for large scale studies. Most importantly, in situ data cannot be collected to provide a synoptic view of the ocean. Today the use of remote sensing allows to have a real time, broad view of the entire ocean circulation and large and mesoscale biological processes. In my study I principally used three kinds of remote sensed biophysical data: Sea Surface Height (SSH), Sea

Surface Temperature (SST) and Chlorophyll concentration (Chl) from ocean color.

Whereas altimetry data are more used to extrapolate information about currents and transport processes, SST and Chl can be used as tracers to detect boundaries of existing physical features in the ocean. Indeed, water masses of "colour" and temperature different to the average background values can be used to track dynamic physical features in the ocean over very large scale using sequential satellite imagery.

2.2.1.1 The Sea Surface Height

The Sea Surface Height (SSH = height of the ocean surface measured relative to the geoid) is provided by highly accurate satellite altimeters (TOPEX/Poseidon, Envisat, Jason-1, and OSTM/Jason-2). The SSH is derived by the combination of the time needed by the signal emitted from the satellite radar to travel to the surface and back to the antenna and the precise satellite location data.

Today highly accurate real time SSH data are available thanks to the combination of measurements from two multiple satellite altimeters, the Topex/Poseidon (T/P) and the European Remote Sensing Satellite (ERS-1) (Ducet *et al.* 2000). This dataset includes information for about 17 years which can be downloaded from the French Archiving Validation and Interpretation of Satellite Oceanographic (AVISO) data centre.

SSH is quantified from the reference geoid assuming that fluids are in a geostrophic balance (the pressure gradient is in balance with the Coriolis force). If sea depth is not taken into consideration and the sea surface is assumed to be homogeneous, SSH can be calculated as the difference between the satellite height (S) and the altimetric range (R): SSH= S - R.

Altimetry has been reliably used to study mesoscale geostrophic circulation structures (fluid motion in balance between the pressure gradient force and the Coriolis force) with lifetimes of at least a week. In fact the AVISO dataset revealed the prevalence of mesoscale structures (in particular thousands of eddies) throughout most of the world's ocean (Chelton *et al.* 2007). On the contrary, structures not in geostrophic balance (where the active forces are not in balance, such as upwelling and tidal mixing) are ephemeral and may disappear or evolve in bigger structures (mesoscale).

2.2.1.2 Principles to investigate transport processes: Eulerian and Lagrangian

Lagrangian and Eulerian analysis constitute the conceptual framework of the analysis I applied to the topology and transport of the ocean flows using altimetry data.

Fluid dynamics in the ocean can be observed and described using two principal methods: Lagrangian, which requires the observation of each individual molecule trajectory in the flow, and Eulerian, which is based on the observation of the fluid velocity at established locations. The Lagrangian methods are the most appropriate and efficient to describe the fluid dynamics, additionally the physical conservation laws are inherently Lagrangian since have been developed to describe a moving fluid. However, the Eulerian approach was chosen for the development of all the theory in fluid mechanics for its simplicity.

Given the Eulerian velocity vector, v(x,t), which describes the fluid velocity at any space (x) and time (t), the motion of a fluid particle with initial localization x(0) is given by the following equation:

$$\frac{dx}{dt} = \vec{u}(\mathrm{t}) = \vec{v}(\vec{x}, \mathrm{t}).$$

When the location and time of the particle are known, the above equation is also valid to estimate the velocity of particles in a Lagrangian system. Several diagnostics are available to identify Eulerian and Lagrangian coherent structures (Boffetta *et al.* 2001b). In my study I used the Lyapunov exponent (Lagrangian) and the Eddy Kinetic Energy (Eulerian). Differences exist between the various diagnostics. The Lyapunov exponent performs better than other diagnostics in extracting, from geophysical flows, regions where the stretching rates are maximal, which are in turn candidate locations of tracer frontal structures and transport barriers (Haller & Yuan 2000). While EKE is the preferred diagnostic to identify mesoscale variability in the oceans (Ferrari & Wunsch 2010).

2.2.1.3 Lagrangian coherent structures to study transport processes

I used altimetry data to be able to detect transport processes like fronts and mixing regions, by calculating from them Lagrangian diagnostics like the Lyapunov exponents. Here is a description of the principles and calculation.

The Lagrangian theory is funded on the property that Lagrangian particles display a chaotic

motion, also known as ‘‘chaotic advection’’, over a two-dimensional time dependent flow (Lapeyre 2002). The computation of the exponential rate of particle separation (also known as Lyapunov exponent) allows to quantify this chaotic non diffusive mixing. The Lyapunov theory states that “in the asymptotic limit in time, fluid particles separate with the same exponential growth rate”.

Considering two particles trajectories, $x(t_0) = \mathbf{x_0}$ and $x(t) = x_0 + \Delta x_0$, each of which will generate an orbit separated by a distance $\Delta x(t_0)$, the global Lyapunov exponent is defined by

$$\lambda = \lim_{t\to\infty} \lim_{\delta x(t_0)\to 0} \frac{1}{t} ln \frac{[\delta x(t)]}{[\delta x(t_0)]}$$

The Lyapunov exponent "λ" is useful for distinguishing among the various types of orbits. It works for discrete as well as continuous systems. The λ values is an indication of the stability of the system. Negative exponents are characteristics of stable systems with decreasing value indicating greater stability. On the contrary positive values of the Lyapunov exponent indicate an unstable and chaotic system. Finally, a Lyapunov exponent of zero indicates steady state mode systems.

The Lyapunov exponent can be calculated using two different methods: the finite-time Lyapunov exponent (FTLE), and the finite-size Lyapunov exponent (FSLE). The FTLE and FSLE are based on similar principles but provide slightly different information. The FTLE provides a measure of particles separated over a specific timeframe period, and for this reason has been proposed by Haller as an indicator of Lagrangian coherent structures (LCSs) (Haller & Yuan 2000). Whereas the FSLE is inversely proportional to the time at which two particles reach a prescribed separation. FSLE provides an alternative to FTLE as indicator for LCSs, and is commonly applied in oceanographic studies. Direct comparisons between FTLE and FSLE have been attempted and if appropriately calibrated, give similar results (Sadlo & Peikert 2007). However Boffetta et al. (Boffetta *et al.* 2001a) argued that: “the FTLE is not capable of recognizing the relevant structures, namely the boundaries between chaos and large-scale mixing”. Therefore FSLE technique appears to be the preferred option for oceanographic applications.

More precisely, $\lambda(x, t, \delta_0, \delta_f)$, the FSLE, at position x and time t, computed from the time τ, it takes for a trajectory starting at time t at a distance δ_0 from x to reach a separation δ_f from the

reference trajectory that started at x:

$$\lambda(x, t, \delta_0, \delta_f) = \frac{1}{\tau} \log \frac{\delta_f}{\delta_0}$$

The FSLE depends on the choice of two length scales: the initial separation δ_0 and the final one δ_f. Since I am interested in mesoscale structures, the other length, δ_f, will be chosen as $\delta_f < 1$ degree, i.e., separations of less than ca. 110 Km. In this way the FSLE represents the inverse time scale for mixing up fluid parcels between length scales δ_0 and δ_f (Bettencourt *et al.* 2012).

2.2.1.4 Sea Surface Temperature

SSH measurements are now widely used to recover surface velocities especially in the open ocean in any meteorological conditions. However, the low accuracy of the results, due to the high signal noise can induce important errors in the location of mesoscale features. On the contrary, in clear sky conditions, Sea Surface Temperature (SST = temperature of the water on the superficial oceanic layer) measurements are able to accurately locate ocean structures either at mesoscale (microwave SST) or at higher spatial resolution (infrared SST).

SST satellite measurements allow to get data over large regions of the oceans in near-real time (Wunsch & Schreiber 1992; Robinson 2004; Dzwonkowski *et al.* 2010). At the beginning of the satellite adventure, SST was indirectly recorded during clear days by a series of thermal infrared sensors, such as Advanced Very High Resolution Radiometer (AVHRR), deployed on operational meteorological satellites to record images of cloud top temperatures (Conway 1997). This problem was overcome with the launch on the AQUA, a multi-national NASA scientific research satellite, of AMSR-E, a new microwave thermal sensor. AMSR-E functions on microwave instead of infrared data, limiting weather influence (emission and attenuation by water vapor and clouds), particularly for the low-frequency channels (6.9 GHz and 10.7 GHz). However, AMSR-E cannot give information about SST close to the coast.

Such accurate large-scale, long-time dataset is important to conduct a wide range of studies on climate change as well as studies on fish ecology, habitat modification and water acidification (Gentemann *et al.* 2003; Purkis & Klemas 2011).

2.2.1.5 Ocean Colour

The term ocean colour has been introduced to describe a particular effect derived from the

interaction between light and the particles within the upper ocean layer. In simple terms when the energy is emitted by the sun and hits the water surface, part of it is absorbed by the ocean while part of the energy is reflected at different intensities depending on the particles in suspension in the upper layer of the ocean. The amount of reflected light of different wavelength (within a range of 400-700 nm) can be measured on satellite through water color sensors (CZCS, SeaWiFs, MODIS among the most cited). Thus the "color of the ocean" is related to the spectrum of visible light emitted from the object focus of the study. This particular feature has been intensively used to study phytoplankton distribution.

Phytoplankton uses pigment antennae (i.e. the sum of chlorophyll-a, divinyl-chlorophyll-a, and chlorophyllide a) to capture the energy of photons. Chlorophyll-a absorbs the red and blue wavelengths and reflects the green ones which are received by satellite. For this reason, in presence of high Chl-a concentration, in the satellite imagery the colour of the ocean changes from blue (true colour of the ocean when light is not absorbed by particles on the surface) to green. Additionally, the intensity of the absorption also provides a proxy of phytoplankton biomass.

In the recent past, the analysis of ocean colour satellite data has moved beyond the estimation of chlorophyll-a concentration to include new parameters, such as the ability to determine the dominant phytoplankton groups in the surface waters (Aiken *et al.* 2009). This is possible because different phytoplankton groups with similar biogeochemical functions absorb light at different wavelengths accordingly to their pigment composition resulting therefore in blooms of different colour. For example, a characteristic, and easily visible from the satellite, milky-turquoise, is produced by coccolithophores, whereas some cyanobacteria produced a characteristic golden yellow colour bloom.

The first order signal retrieved from the ocean sensors colour, also known as the normalized water-leaving radiance (nLw), is particularly useful to quantify the ocean productivity as it depends only from Chl-a. However, as Chl-a is the only pigment present in all phytoplankton species, the first order signal does not provide any information on the composition of phytoplankton blooms. To bypass this limitation, several methods to retrieve PFT information from satellite ocean color data have been proposed (Uitz *et al.* 2006; Nair *et al.* 2008). For example, Alvain et al. (Alvain *et al.* 2005; Alvain *et al.* 2008) developed the PHYSAT method,

also applied in this study, which is based on the second-order spectral effects of nLw and in-situ diagnostic pigment data to classify normalized spectra into six dominant phytoplankton groups (diatoms, nanoeukaryotes, *Synechococcus*, *Prochlorococcus*, *Phaeocystis*-like and Coccolitophorides) (See paragraph 2.2.4 for more information).

Phytoplankton maps developed with this method can be used as proxy indicators of algal biomass distribution which ultimately supply food for fish. Combination of chl, SSH and SST is the most useful in understanding fish behaviour, allowing to collect important information on fish stock recruitment that is dependent on relative time of spawning and seasonal phytoplankton blooms (Cushing 1990). As well as to study the habitat utilization of higher predators in relation to the environment scales and also the impact of climate change on production.

2.2.2 Remote sensing information about biodiversity

To estimate a way of measuring global biodiversity in the real ocean by considering the spatial patchy distribution of diverse planktonic communities over a given scale, I used remote sensed information (Fig. 10). Several algorithms aim at extracting biologically relevant information from spectral anomalies in remotely sensed optical signals (Ciotti & Bricaud 2006; Raitsos *et al.* 2008; Kostadinov *et al.* 2009; Brewin *et al.* 2010; Hirata *et al.* 2011) or their heterogeneity (Rocchini *et al.* 2010).

The PHYSAT algorithm (Alvain et al., 2005) provides a classification of bio-optical anomalies of Ocean color (http://oceancolor.gsfc.nasa.gov/) constructed on the detection of specific signatures of normalized water leaving radiances (nLw) measured by ocean color satellites (http://log.univ-littoral.fr/Physat). This empirical method utilizes simultaneous in-situ pigments inventories (i.e. GeP&Co campaigns: Dandonneau et al., 2004) and remote sensing information collected by the SeaWiFS ocean color sensor. The magnitude and -to a lesser extent- the shape of specific anomalies (nLw*) are empirically associated with mainly 6 dominant groups: nanoeucaryotes, *Prochlorococcus*, *Synechoccocus*, diatoms, coccolitophorids and Phaeocystis-like (Alvain et al, 2005; Alvain et al., 2008). The second order variability of the satellite signal, nLw*, is obtained by dividing the actual nLw by a mean nLw reference model which depends only on the standard Chlorophyll-a concentration. To characterize each group sampled in-situ in function of nLw* spectrum, a set of criteria has been defined. When applied to global daily remote sensing

measurements, these criteria allow to retrieve a global map of the most frequent group of dominant phytoplankton. When no dominant groups are detected, pixels are classified as 'unidentified'. The geographical distribution and seasonal succession of PHYSAT groups have been studied using the global Sea-viewing Wide Field-of-view Sensor (SeaWiFS) ten year archive. The information provided by PHYSAT is in accordance with previous studies based on in-situ observations (Alvain *et al.* 2008) and has been validated on independent in-situ measurements to unveil the underlying theoretical explanation of the selected criteria using a realistic model of radiative transfer for surface open ocean waters (Alvain *et al.* 2012).

PHYSAT, along with all the other approaches that try to estimate phytoplanktonic community composition from remote sensing, presents some limitations. Indeed, only a limited number of distinct bio-optical classes can be identified and related to properties of the ecosystem. Moreover, the categorization is subject to errors. Such uncertainty does not compromise the analysis, since the approach of deriving a biodiversity index from spatial heterogeneity does not rely on the identification of a specific group, but only requires regions where different communities are stirred to appear as a mosaic in terms of the different PHYSAT bio-optical anomalies. The misclassification or undetectability of some patches would affect the quantitative value of a biodiversity index at a given point, but not necessarily in a systematic way. Therefore, considering the temporal and spatial scale of the analysis, the maxima and minima of the estimate are generally unaffected.

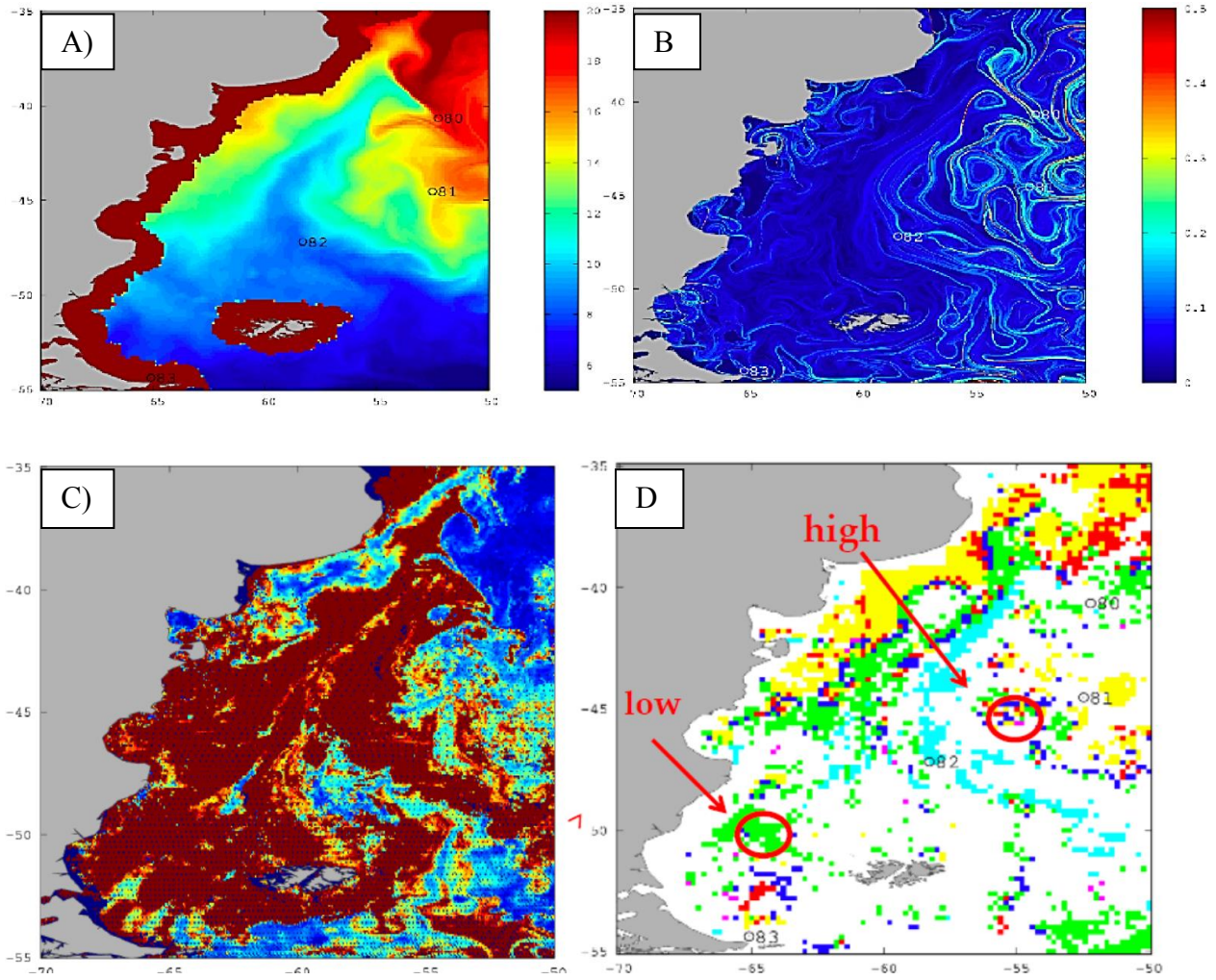


Figure 10. Remote sensed information of biophysical processes characterizing the Brazilian and Malvinas currents confluence zone offshore Patagonia on a weekly snapshot. A) AMSR-E derived SST (C°); B) Lyapunov exponents describing transport fronts and coherent water masses rationally representing fluid-dynamical niches (day $^{-1}$); C) chlorophyll bloom; D) Physat detected dominant phytoplankton groups (different colours) showing zones of high and low spatial diversity (in the disk).

2.2.3 Computation of a spatial-based diversity index and its relationship to local diversity

The main advantage of the seascape biodiversity index defined in 2.1.2 - which is the main methodological result of my thesis - is that it only requires spatial information of the dominant phytoplanktonic types, and therefore can be formally applied to remote sensed biophysical patches like the ones detected by PHYSAT. I used patchiness in maps of PHYSAT bio-optical anomalies as an indicator of niche spatial heterogeneity at a given time, and assume that patches that are about 100 kms apart are typically mixed within a few weeks by oceanic turbulence

(d'Ovidio *et al.* 2010). With these assumptions a new index based on remote sensing as a proxy for diversity induced by horizontal stirring is proposed under the form of an area-based Shannon entropy of the distribution of bio-optical classes:

$\tau = -\sum \boldsymbol{p_i} \log (\boldsymbol{p_i})$

The frequency p_i is the fraction of pixels, within a circular area, falling in the i-th bio-optical class. τ is maximal when the bio-optical classes are all present and evenly distributed inside the area, and minimal when only one class is observed. The radius of 100 kms appears to be the best for a global analysis by numerical simulations (see Chapter 3), and for remote sensing data it is sufficiently large to provide a statistically meaningful frequency distribution, and sufficiently small for stirring to mix the area content on the time scale of a planktonic bloom. The index τ is a local measure of the heterogeneity in the ocean surface bio-optical properties at a certain scale, that is the radius of a disk over which the occurrence frequencies of remote-sensed bio-optical frequencies are estimated.

The basis for considering τ a possible proxy of biodiversity derives from the expectation of finding higher diversity at one point in the ocean, the more heterogeneous the distribution of planktonic communities - in this case, emerging as bio-optical anomalies - around that point. On the contrary, if the conditions at a given location are such to support high diversity, this high diversity will reflect on the ability of that local community to prosper in nearby niches shaped by transport, thus creating high spatial heterogeneity in the nearby environment. Water masses that are 100 kms apart are mixed on a time scale of the order of few weeks, comparable to that of a planktonic bloom (d'Ovidio et al., 2010), so that each of the communities present in the area are expected to be locally represented. On the other hand, communities that are locally present will be given the chance to colonize niches in that area before the bloom season is over. The concept of niches that are fixed in time, typical of terrestrial landscapes, does not apply to the open ocean environment. This is especially true at the scale ~10s of km (meso- and submesoscale), where microbial communities are disturbed (Connell 1978) by stirring on the same time scale of their demography.

2.2.4 Data analysis

Spatial patterns of remote sensed diversity and regional validation.

To obtain a global map of spatial patterns of biodiversity, I used 1-week global PHYSAT composites at 9 km resolution and for a focal point of the ocean surface, I computed the spatial-based index τ. The obtained patterns were validated for the Atlantic Ocean using Atlantic Meridional Transect in situ data (for description of AMT data see Chapter 2.3.1). The AMT measures of biodiversity have been compared to the τ index by zero-order correlation analysis at their resolution, which is lower than that obtainable through the satellite measures. I used the τ index at the same geographical position of the AMT-2 stations and computed both the annual climatology (2003-2010) and the climatology restricted to the months in which the transect was run. The stations were also divided according to Northern and Southern Hemisphere locations and a linear relationship with τ evaluated.

Both τ and AMT diversity were compared to average SST (2003-2010) by quadratic regression to evaluate consistency in already known macroecological patterns related to temperature.

Temporal patterns of remote sensed diversity

Seasonal climatologies were computed to evaluate seasonal changes in the τ index. I quantified seasonal changes using a Signal-to-Noise ratio. This ratio is the ratio of mean (average climatology) and standard deviation of a measurement or signal:

$$\mathrm{SNR} = \frac{\mu}{\sigma}.$$

Commonly used in image processing with a spatial meaning, I transferred it to a temporal meaning to highlight zones of instability of the τ diversity through time. Following the *Rose criterion,* an SNR close to 5 allows to distinguish the signal with very low uncertainty.

Annual successional patterns of plankton dominant community were estimated using Shannon index for each pixel counting how many functional groups appeared at each location through time.

Annual successional diversity and τ diversity patterns were compared to average Chlorophyll climatology (2003-2010) by regression analysis.

2.3 In situ global biodiversity information

The main challenge in global biodiversity patterns is that exploitable information about marine organisms shows a sparse resolution, especially for plankton, due to logistical reasons. Sampling design often lacks sufficient spatial and temporal consistency, replicates, consistent protocol, databases organization and diffusion. To gather large scale information about morphological related taxa diversity, here I exploited the most comprehensive one, the Atlantic Meridional Transect (AMT) that runs since 1997. On the other hand, to exploit the possibilities of new technological approaches in resolving biodiversity at finer levels, I referred to the Tara Oceans project.

Recently, a relatively large amount of data has been collected by tracking the movement of marine predators. Regions with increased probability of localization of different species of marine predators have been identified (Block *et al.* 2011; Kaschner *et al.* 2011), and considered as hotspots of biodiversity for higher levels of the trophic chain. I used these literature derived information to globally compare remote sensed diversity of primary producers with diversity of top predators species. Moreover, to compare with diversity at higher levels of the trophic chain, I used integrated compilations of global biodiversity in situ observations with statistical models based on habitat predictions in order to have smooth global complete information. Specifically, Aquamaps integrates the most comprehensive databases and the best habitat modelling approach.

2.3.1 Atlantic Meridional Transect

2.3.1.1 The project, the sampling design and collection

The Atlantic Meridional Transect programme (http://amt-uk.org/) samples the Atlantic Ocean twice a year, measuring biological, chemical and physical parameters. I consider here the AMT-2 transect, that was run from 22 April to 28 May 1996, when the RRS *James Clark Ross* navigated from the Falkland Islands to the UK. The main aim of the second AMT cruise was to study biological processes in the open Atlantic Ocean over very extensive spatial scales. The transect crosses different ecosystems, from sub-polar to tropical and from euphotic shelf seas and upwelling systems to oligotrophic mid-ocean gyres (Fig 11).

Stations are about 200 Km apart. Water samples were collected using Niskin bottles for 5-6 depths for each station. 100 ml were taken at each sampling depth for counting and identification of nanoplankton and microplankton species (cyanobacteria, phytoplankton, heterotrophic plankton, ciliates). Samples were preserved in lugol iodine solution and examined following Uthermol sedimentation technique under inverted microscope. Bacteria were counted using epifluorescence technique.

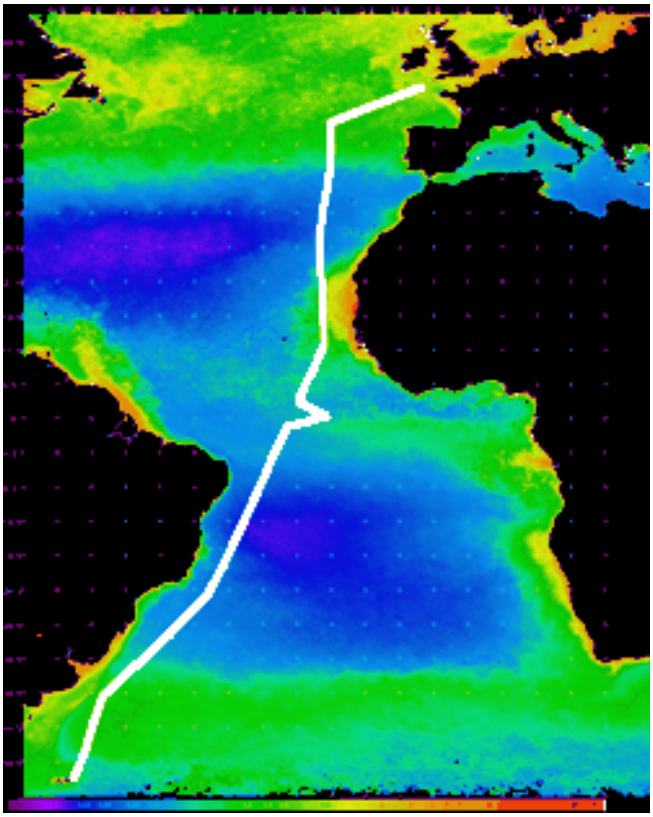


Figure 11. Map of the Atlantic Meridional Transect overlayed to ocean color data about chlorophyll concentration showing eutrophic (orange, yellow and green) and oligotrophic blue and violet) regions. From http://amt-uk.org.

2.3.1.2 Morphological diversity from inverted microscopy analysis

Community composition was measured in both cellular (n cells m^{-3}) and carbon biomass (mg C m^{-3}) by inverted microscopy. These data have already been analysed to investigate planktonic community structure (Marañón *et al.* 2000; Irigoien *et al.* 2004; Cermeño *et al.* 2008). I

computed the Shannon index by considering the abundance (integrated over the mixed layer) of the organisms classified in functional groups, and weighting organisms belonging to different size fractions by their carbon biomass, following Cermeño et al., 2008. The index was normalized with respect to the maximal value that would be attained if all genera observed in the cruise were equally represented at a single location. While considering depths separately, or using higher levels of taxonomic classification, the overall pattern is the same irrespective of the particular method of computation of the Shannon index.

2.3.2 Aquamaps

2.3.2.1 The model approach and data integration from global databases

AquaMaps (www.aquamaps.org) (Kaschner *et al.* 2013) is a model that provides large-scale predictions of natural occurrences of marine species. It estimates environmental preferences of marine species with respect to depth, water temperature, salinity, primary productivity, and association with sea ice or coastal areas. These estimates of species preferences, called environmental envelopes, are derived from large sets of occurrence data available from online databases (i.e. GBIF www.gbif.org and OBIS www.iobis.org), and from independent knowledge from the literature about a given species distribution and its available habitat usage, (i.e.in FishBase, SeaLifeBase and AlgaeBase for non-fish). The environmental envelopes are matched against local environmental conditions to determine the suitability of a given area in the ocean for a particular species. The AquaMaps approach of incorporating species occurrences and expert knowledge into an environmental envelope is modified from an ecological niche model originally developed by Kaschner et al. 2006 (Kaschner *et al.* 2006) for predicting global distributions of marine mammals. Its advantage is to allow corrections for biases in occurrence data such as non-representative coverage of a species' distribution, biases in sampling effort and data provision, and species misidentifications. Moreover, this approach is applicable to a wide range of marine organisms, both fish and non-fish species. AquaMaps predictions have been validated using independent and effort-corrected survey data (Ready *et al.* 2010). The performance of the model was comparable with other presence-only species distribution models such as GARP-Genetic Algorithm for Rule Set Production (Anderson *et al.* 2003), Maxent -Maximum Entropy Modeling (Phillips et al., 2006), GLMs-generalised linear models and GAMs-generalized additive models

(McCullagh & Nelder 1989).

To compare the information about diversity of primary producers and the diversity of species of higher trophic levels, I compared the τ index calculated from remote sensed information of dominant plankton groups (see Chapter 2.2) with global species richness calculated from Aquamaps using just the species with probability higher than 0.5 of being present for each pixel.

2.3.2.2 Relationship of remote sensed diversity of primary producers with diversity of consumers

To analyse how remote sensed diversity of primary producers relates to measures of local biodiversity of the global ecosystem including multispecies distributions of higher levels of the trophic chain, I used 1) regional maps of top marine predators diversity estimated from fishery data from Worm et al. 2003 and from telemetry tracking from Block et al. 2011 for a qualitative comparison and 2) global predicted species richness of consumers for a quantitative comparison with annual climatologies (2003-2010) of remote sensed diversity. Quantitative comparisons have been approached using zero-order correlation analysis and regression analysis to determine direction and strength of the relationship. The best fitted regression model is a linear function.

Residuals obtained by regression analysis were studied in relation to the physical and biological characteristics of the environment: Total Kinetic Energy (m/s), Chlorophyll (µg/l), Sea Surface Temperature (C°) and temperature gradient (C°). All the environmental variables are derived by computed annual climatologies (2003-2010) of remote sensed data L3: 9 km sea surface currents from altimetry data from AVISO, 9 km weekly composites of Chl from SeaWiFS and 9 km daily images of SST from AVHRR.

To underline possible patterns, the values of the annual climatologies of the environmental variables where binned following equal intervals. The residuals of the regression analysis where averaged depending on the pixel location having environmental values corresponding to the bins. The bins of the environmental variables were plotted against the averaged residuals. The percentage corresponding to the surface cover of the ocean of the extreme values has been computed dividing by the total number of values representing the whole ocean.

To identify hotspots of biodiversity detected by both indices, I used a congruence analysis. In

each map, I defined biodiversity hotspots as the top-valued pixels covering 20% of the total ocean surface. This percentage is set in agreement with the target chosen by the Convention on Biological Diversity, which aims at protecting at least 10% of the global ocean extension by 2020 (Aichi Target 11). Spatial congruence is evaluated calculating the percentage of overlapping computed.

2.3.3 Tara ocean expedition and global high throughput information

The Tara Oceans project (oceans.taraexpeditions.org) is a 3 year (2009-2011) global-scale investigation of morphological, genetic and functional biodiversity of plankton organisms in relation to the dynamical physico-chemical parameters of the oceans. One of the primary objectives is to map biodiversity across scales spanning five orders of magnitude, from viruses to bacteria, archaea, protists and metazoans (Karsenti *et al.* 2011).

The advantage of Tara Oceans, compared to other global scale genomics studies already launched (i.e. Global Ocean Sampling (GOS) expedition) is the integration of the genetic, morphological and functional diversity, and its environmental context at global scale and at multiple depths from viruses to zooplankton. Sampling was conducted over 153 stations (Fig. 12) and multiscaled according to allometric criteria. Plankton were collected from up to three depths: near the surface (~5 m), at the depth of maximum chlorophyll *a* fluorescence (20–200 m) and in the mesopelagic layer (200–1000 m) to capture deep oceanographic features.

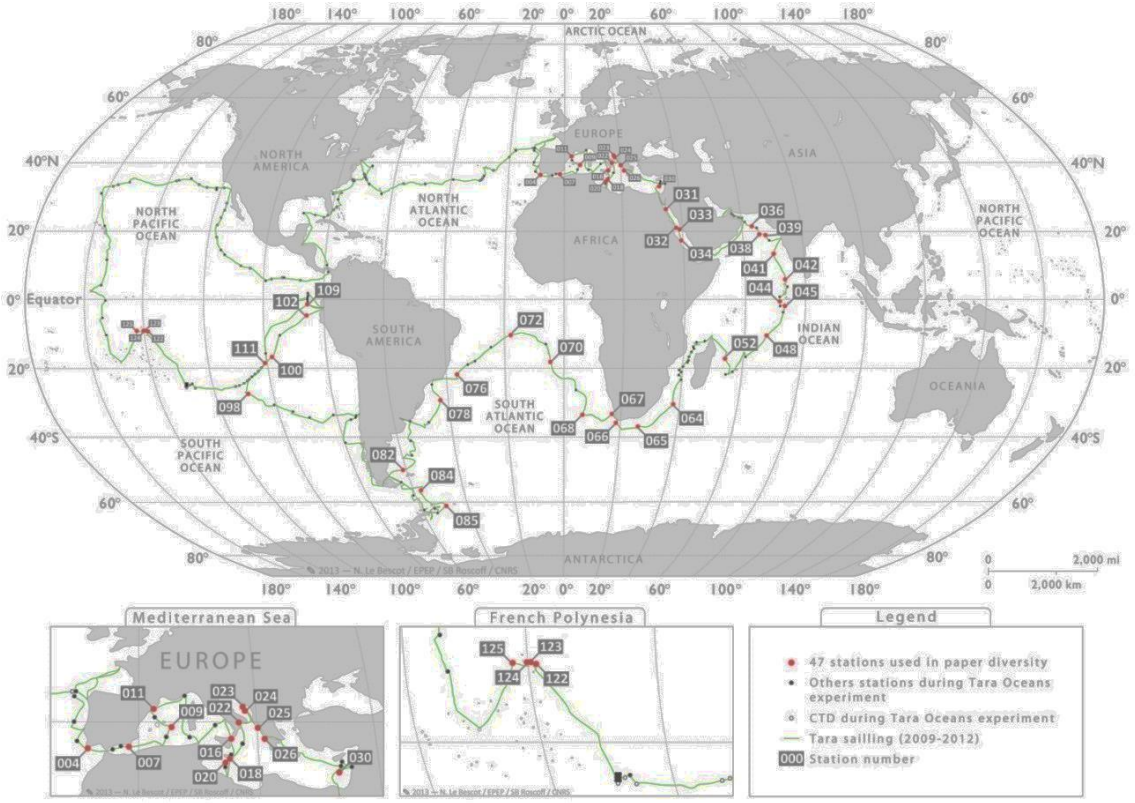


Figure 12. Worldwide layout of the *Tara* Oceans project transect and sampling stations. Courtesy of Tara Oceans project.

Sampled plankton was size fractioned by 7 serial filtrations and volume proportional to organism size. When possible, samples were collected using a large peristaltic pump, otherwise using Niskin bottles mounted on a rosette equipped with physico-chemical sensors. Bongo nets were used for larger plankton (Fig. 13). 100 liters of seawater from each depth were first passed through 200- and 20-µm mesh filters to collect larger plankton, then passed in series through 5, 2, and 0.5 µm mesh filters.

Samples were either collected on glass/microfibers polycarbonate filters and resuspended in vials for genetic analysis, morphological analysis with live or fixed organisms. Samples were then frozen accordingly.

DNA was extracted from the filters using a modified CTAB (hexadecyltrimethylammonium bromide) protocol (Winnepenninckx *et al.* 1993). Between 100 and 250 µl of purified total metagenomic DNA was recovered per sample.

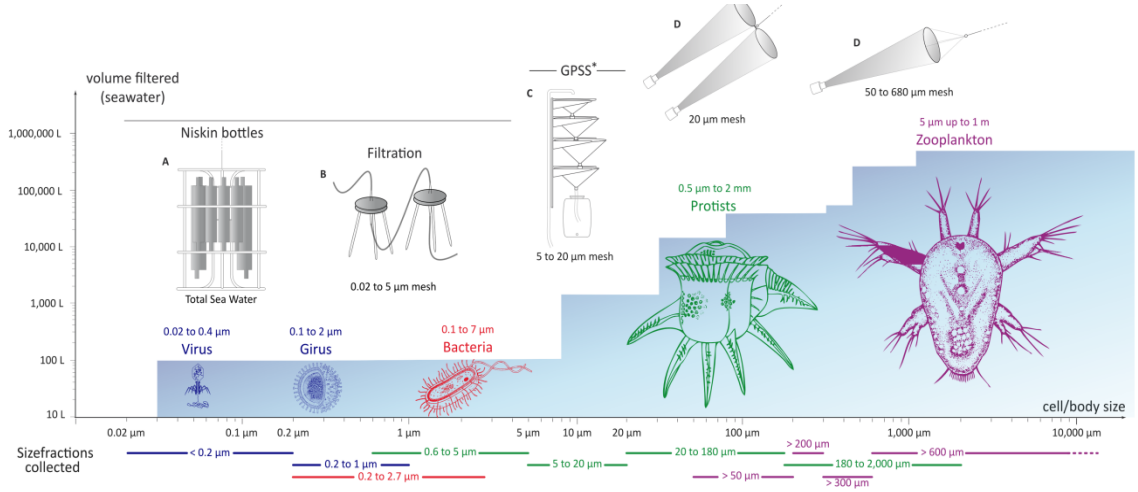


Figure 13. Graphical display of the plankton sampling process highlighting the different methods used for each size fractions. From Karsenti et al. 2011 (Karsenti *et al.* 2011).

2.3.3.1 Morphological diversity from high throughput imaging

Morphological data were gathered using four different high throughput imaging techniques targeting diverse plankton communities according to allometric constrains (Fig. 14). I focused on data elaborated by two of them: flow cytometry and FlowCam.

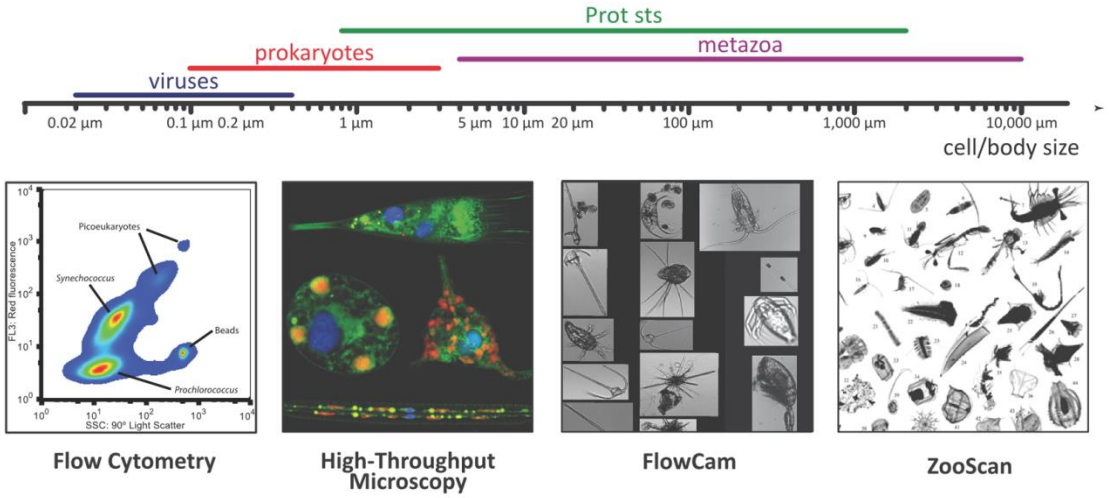


Figure 14. Schematic example of high throughput imaging applied depending on the size classes. From Karsenti et al. 2011.

2.3.3.2 Seaflow

2.3.3.2.1 Description

New high frequency flow cytometers -SeaFlow (Ribalet *et al.* 2010) allow long continuous real-time shipboard observations of microbial cells of 0.5-20 microns in size that can be automatically reconducted to predefined phytoplankton populations based on size and pigments and reflectance characteristics (Swalwell *et al.* 2011). Fluorescence and light-scatter properties of cells in liquid suspension passing one by one through a light field are measured. Scattering depends on the size, shape and refractive index of the cells. The scattering and autofluorescence properties are exploited to identify different phytoplankton. The pico (0.2–2 μm) and nano (2−20 μm) eukaryotes produce a greater lightscatter signal and brighter red fluorescence than the prokaryotic picoplankton and can be distinguished from them (Dubelaar & Jonker 2000). Prokaryotic picoplankton of similar sizes, such as *Synechococcus* and *Prochlorococcus*, can be distinguished based on the orange fluorescence signal produced by the phycoerythrin pigments present in large concentrations in *Synechococcus*.

2.3.3.2.2 Data analysis in the Mediterranean transect case study

I explored the Seaflow data from the Tara transect during the Mediterranean lag run from 23 September to 1 October 2009 provided from F. Ribalet. Data were recorded each 3min interval along with GPS position. Data were checked for optical alignment and fluid stability. The FlowPhyto (Ribalet *et al.* 2011) software was set up to cluster six phytoplankton populations, of which the following were found during the cruise: Nanoplankton, *Synechococcus*, Ultraplankton, Picoplankton. Only validated data were used to cluster a pre-defined number of phytoplankton populations. Here is an example that shows the four populations on different cytograms (Fig. 15, provided by F. Ribalet).

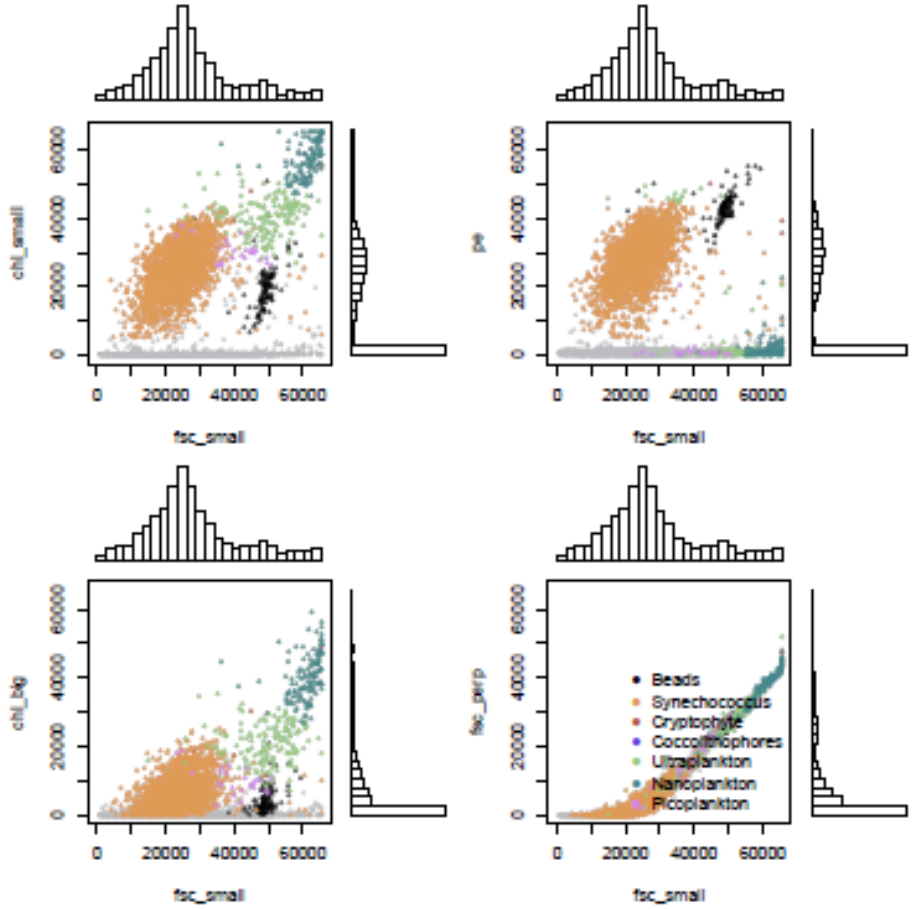


Figure 15. Flow cytometric signatures of four phytoplankton populations: Nanoplankton, *Synechococcus*, Ultraplankton, Picoplankton. 'fsc_small' represents the forward angle light scatter, which is roughly proportional to cell size; 'ch_small' represents the red fluorescence from chlorophyll; 'pe' represents the orange fluorescence from phycoerythrin and is used to identify *Synechococcus* and Cryptophytes; 'fsc_perp' represents the polarized light scatter and is used to identify Coccolithophores. Courtesy of F. Ribalet.

Indexes that use cytometric diversity capture the structure of a microbial community and indirectly its richness in physiological and genetic variations. They are based on the bio-optical properties of the community measured at the single cell level by flow cytometry. CytoDiv R package was used to compute the cytometric diversity indices as described in Li 1997 (Li 1997). I used the exponential of the Shannon index, also called linearized Shannon index, to qualitatively compare changes in microbial diversity along the transect, according to physical characteristics of the water masses. These characteristics were highlighted with multisatellite measurements in the context of the mesoscale fluid dynamical landscape. Satellite data used, corresponding to the

same time period of the cruise, were ocean color (SeaWiFS, ENVISAT and MODIS and MERIS) and surface currents from altimetry (Jason-2,Envisat). Multisatellite ocean color daily images at 4 km resolution were used for surface chlorophyll-a distribution. Surface currents were analysed with Lagrangian diagnostics, in order to extract transport properties like transport fronts, mixing regions, and origin of water masses. This was achieved by constructing particle trajectories and computing Lyapunov exponents (FSLE). Physical data (Lagrangian diagnostics), biophysical data (chlorophyll) and biological data (cell type, density and assemblages) were combined together to investigate the contribution of horizontal transport in shaping planktonic community structure along a continuous surface transect.

2.3.3.3 FlowCam

2.3.3.3.1 Description

The FlowCam is composed by a black and white digital camera (1024 X 768 pixels) equipped with a 4x objective, a 300 um depth flow cell, a green laser (532 nm), two channels of fluorescence detection (575 nm and >650 nm), a digital signal processor and an imaging processing software such as VisualSpreadsheet. The FlowCam can detect the image of particles in the range of 12 μm to 300 μm. Total Particles are detected and imaged at a regular user defined interval, typically at a rate of 7 camera images per second. The processing capability is 1-3 ml/min and a density of 50,000 particles/ml. In real time cells are counted, subimages are saved (segmentation), background is subtracted from the image and cell sizes are measured directly in images. More than 25 different image parameters are collected. Basic shape measurements include: Equivalent Spherical Diameter (ESD), Area Based Diameter (ABD), Length, Width, Aspect Ratio, Area, Volume Advanced; Morphology Measurements Include: Circularity, Elongation, Compactness, Circle Fit, Perimeter, Convex Perimeter, Edge Gradient, Fiber Curl; Gray-Scale and Color Measurements Include: Intensity, Average Intensity, Sigma Intensity, Transparency, Average Red, Green, Blue, R/G Ratio, R/B Ratio, G/B Ratio. Based on the acquired image parameters, taxonomic classification of the images utilizes a training set, machine learning algorithm and expert system (Fig. 16).

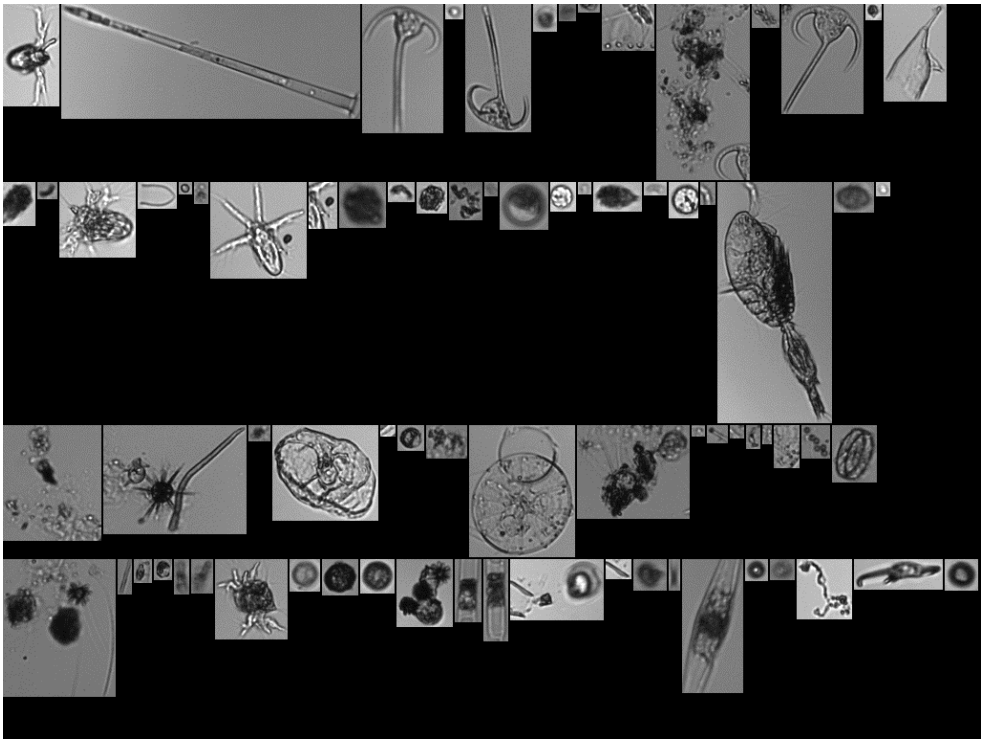


Figure 16. Example of plankton detected from FlowCam. Courtesy of Tara Oceans project.

2.3.3.3.2 Data analysis in the Mediterranean Sea case study

For this preliminary study on coherence of biodiversity information and community structure from different techniques, I focused on diatoms collected from three Mediterranean stations (Tara st 7, 23, 30, see Fig.12). I used surface samples of the 20-180 μm size fraction --as the size fraction containing most of the diatoms species— taxonomically identified with FlowCam in Auto Image mode at 10,000 particles for the morphological data (classification of FlowCam images improved by Noan Le Bescot with the help of expert taxonomists). For molecular data I used the V9 18S rDNA barcoding tags (see Chapter 2.3.3.4) obtained from the same samples. Only tags corresponding to diatoms species according to SILVA reference database were retained. As the morphological method does not have the same resolution as the genetic one, to make the two methods comparable, classes were collapsed at the genus level. Note however that most of the genera presented just one predominant species and another one or two very rare species. Therefore this re-categorization was not expected to have an influence on the comparison. Relative abundance distribution curves were created and compared to reference models for the distribution (e.g. exponetial, log-normal, etc). The slope of the interpolation model

curve was calculated to determine if the different groups could have been represented by the same distribution, characterized by its type and slope. K-dominance curves (cumulative relative abundances per genus rank) were compared for each group: usually the community having the lowest saturating curve is considered to be the most diverse. To capture various aspects of biodiversity and to be able to compare biodiversity between stations in a consistent way, genus richness, Q index, Shannon index, linearized Shannon index, Chao's estimation of Shannon index and taxonomic distinctness were considered (Magurran & McGill 2011). These indices are chosen because Q doesn't suffer the bias of the most rare and the most abundant species, Shannon weights the classes by their true frequency, the linearized Shannon index allows for comparisons of effect size, Chao's estimation (Chao 1984) includes also the probability of rare species not being detected and taxonomic distinctness gives different importance to the classes based on their taxonomic uniqueness in the community. Ecological distances such as Bray-Curtis Dissimilarity index was also calculated to compare the stations based on the change in composition of the community structure.

2.3.3.4 Molecular diversity from barcoding

2.3.3.4.1 Description

Divers molecular techniques allow to explore different levels of biological organization such as the type of organism, its genome, and its expressed genes. Organismal composition of Tara Oceans samples was defined through sequencing of phylogenetic markers and environmental gene content/expression was derived through direct metagenomics/metatranscriptomics sequencing.

Bacterial 16S and eukaryotic 18S ribosomal DNA (rDNA) genes contain nine "hypervariable regions" (V1-V9) that are proved to show substantial sequence diversity among different taxa. Species-specific sequences within a given hypervariable region represent valuable targets for biodiversity studies and other scientific investigations. No single region can differentiate among all bacteria and all eukaryotes; therefore, systematic studies that compare the relative advantage of each region are needed (Chakravorty *et al.* 2007).

Multiplexing high-throughput sequencing was used to analyse the metabarcode, allowing to massively sequence key genetic markers for fast and semi-quantitative assessment of community composition. Diversity barcodes (300-400bp) have been developed (Roscoff and SILVA

database) for the total eukaryotic community (V4 or V9 regions of 18S rDNA), the total prokaryotic community (V3 and V6 region of 16S rDNA), the photosynthetic part of protistan community (chloroplastic 16S rDNA fragment), and the fine taxonomic structure of particularly important groups (fragments of mitochondrial COX1 gene or nuclear 28S rDNA) (Fig. 17). These sequences exhibit sufficient base conservation for the design of PCR primers of broad taxonomic range, but at the same time, enough variability to precisely identify the taxon they belong to. Barcodes were PCR-amplified from total-DNA or total-RNA (cDNA) extracts, and multiplexed and sequenced using *ILLUMINA* technology. An informatics pipeline for the primary analyses of the generated data (extraction of high-quality sequences/chimera detection and removal/taxonomic assignation using a home-made reference database), produced usable data for quantitative biodiversity analysis.

Metagenomic tags were mapped to operational taxonomic units (OTUs) on 97% similarity clustering of reference sequences (SILVA and Roscoff databases). This cut off is used to group taxa to species level, but in microbes it can refer also to genus levels (Mende et al. 2013). Moreover, issues exist such as being related to the heuristic nature of the tags, intra-genomic gene copy number variation, etc. Constrains are discussed in Chapter 6.3.

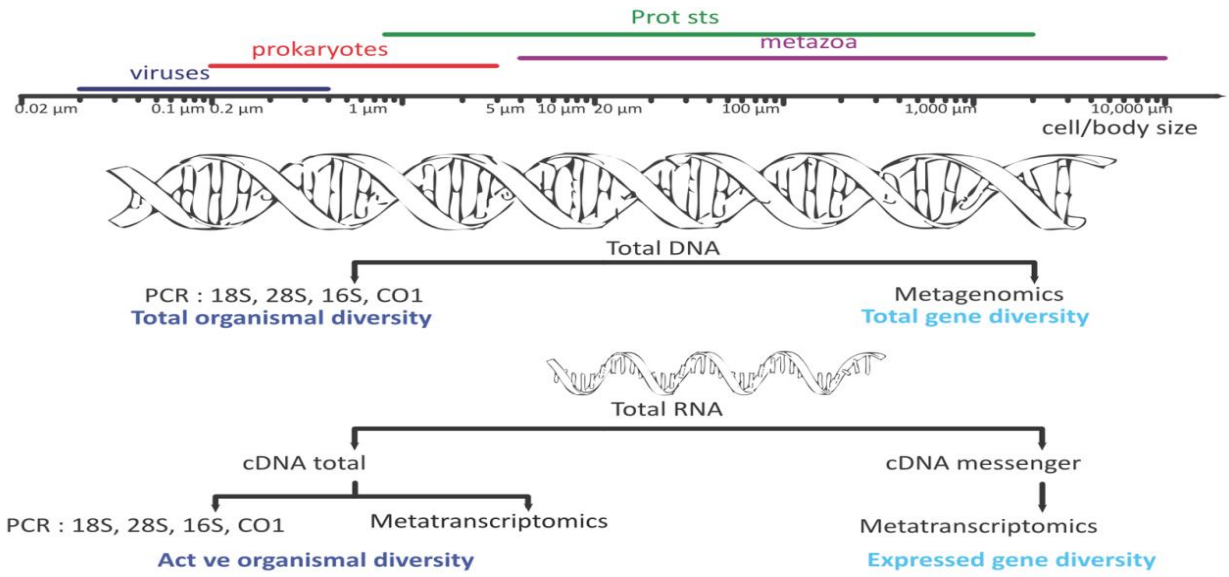


Figure 17. Shotgun Paired-end ILLUMINA sequenced genomic markers. From Karsenti et al. 2011.

2.3.3.4.2 Analysis of phytoplankton community structure from metabarcoding

I used metabarcode data to derive information about the planktonic community structure and diversity. As the rare biosphere plays a challenging role in biodiversity estimation, moreover if molecular diversity is considered, I investigated the community as composed by rare and common species and look at the difference of shape in the distribution. The slope of the fitted curve of the distribution was used as a parameter to compare communities between stations and to compare with the rare community counterpart. I constructed Rank Abundance Distributions for each of the 45 surface stations using just V9 barcodes of photosynthetic eukaryotic plankton. The variable axes were log-transformed. The noise and discretization problems of the curves were eliminated by discarding the classes with abundance less than log(10) reads. Different cutoff values were tested to define the tail of the distribution, therefore to define the rare species and common species abundance curves of the community. The chosen (and best) cutoff is at the classes with log(5000) reads and below. This cutoff was used to fit separately the part of the community composed by common and rare species. The distributions were fitted by a power-law function and the slope calculated. A Z-test was applied on the slope values to determine if all the curves had similar distributions across the different stations among common and rare species. A t-test was applied to determine if the average value of the slope was different between common and rare part of the community. To investigate if the extreme values corresponded to stations with particular environmental conditions at that time, I used multivariate analysis. First I investigated how similar stations were in terms of physical and biological environmental conditions. I used five remote sensed variables (Sea Surface Temperature, Chlorophyll, FSLE, latitudinal component of the Total Kinetic Energy and longitudinal component of the Total Kinetic Energy) and I used a Principal Component Analysis to eliminate redundancy of information and detect which variable was contributing the most to the dissimilarity of the stations in terms of the biophysical environment. These variables were used to graphically represent the dissimilarity among stations by use of a Multidimensional Scaling plot.

Chapter 3 Definition and robustness of a new biodiversity proxy. A study based on the ECCO2-Darwin circulation model.

Biodiversity of an ecosystem is quantified by indicators of species richness or by statistics based on the abundance of different types, such as the Shannon index, that require sampling the local community to sufficient depth. If dispersal is not a primary force in shaping species repartition, these indicators provide information about the processes that have shaped a community at a given location. In the open ocean, dispersal due to mixing acts on spatiotemporal scales that are shorter than in most land ecosystems, thus making local ecosystem composition heavily dependent on the surrounding ecological landscape. The analysis of a realistic numerical model for phytoplankton biogeography shows that local biodiversity can be estimated by spatial information on only dominant functional types. Such 'seascape' diversity index best represents local diversity in regions of medium to high energy currents, where mixing efficiently converts patchiness into local diversity. This result supports the use of spatially extended data with poor community resolution to evaluate and monitor biodiversity hotspots at the global scale.

3.1 Introduction

Biodiversity plays a key role for the homeostasis of the planet and for the resilience of many ecosystem services of key societal relevance (McGrady-Steed *et al.* 1997; Naeem & Li 1997). Biodiversity however is not evenly distributed and understanding what are the global patterns of its variability is a central issue in ecology, as well as an important step for setting effective conservation priorities (Hutchinson 1959; Gaston 2000; Myers & Worm 2003). Being at the base of the trophic chain, primary producers have received special attention in biodiversity studies (Jetz *et al.* 2009). The correlation between their diversity and that of consumers makes the understanding of phytoplankton biodiversity distribution particularly relevant for conservation of marine ecosystems (Duffy 2003; Jenkins *et al.* 2013).

Primary production in the ocean is largely provided by planktonic microbes, and this feature poses several challenges to the establishment of biodiversity maps. Observation-wise, traditional

measures of biodiversity – typically used for macroscopic organisms with well-defined taxonomic classification – cannot be straightforwardly applied to microbial communities. These are indeed increasingly characterized by genomic methods, where no consensus has yet been reached on how to detect, or even define, microbial species (Konstantinidis *et al.* 2006; Achtman & Wagner 2008; Giovannoni *et al.* 2013). More fundamentally, the geographical arrangement of marine ecosystems, and planktonic ones in particular, changes on an extremely rapid time scale, that of the so-called 'ocean weather' (Williams *et al.* 2007; Levy *et al.* 2008). At the mesoscale (10-100 kms and days-weeks scale), the seascape's dynamics stands as a source of perturbation for the planktonic ecosystems, and turbulent mixing enhances dispersal, both factors potentially increasing biodiversity (Connell 1978; Cadotte & Fukami 2005; Clayton *et al.* 2013). In the open ocean, such transport-related processes are likely to prevail over mechanisms that root biodiversity in the adaptation to local environmental features. Finally, the dynamic nature of niches arrangement (d'Ovidio *et al.* 2010) blurs the species ranges and makes it impossible to neglect transport when assessing the global biogeography of species.

3.2 What models and observations tell us about biodiversity and its drivers in the ocean?

Model and observational studies have recently started to tackle these problems and to integrate ocean physics in the description of the biogeography of planktonic species. In the past decade, circulation models have yielded accurate and quantitative physical simulations of the global ocean dynamics down to the mesoscale and below; more recently, the ocean physics has been complemented with modules describing phytoplankton ecology in increasing detail (Follows *et al.* 2007; Barton *et al.* 2010; Levy *et al.* 2014). Such coupled models provide reliable virtual testbeds for studying the processes underlying global biodiversity patterns of oceanic primary producers and their relation to physical determinants like transport and mixing induced by mesoscale turbulence.

Observations have also been providing new opportunities to the identification of biodiversity patterns in the ocean. Collections of decade-long observational records, together with advancements in statistical methods for habitat modelling, have produced global maps of species occurrence and richness for consumers and higher trophic levels, with resolutions in the range of 1000-10000 kms (i.e., averaging out mesoscale variability) (Irigoien *et al.* 2004; Worm *et al.*

2005; Tittensor *et al.* 2010; Kaschner *et al.* 2011; Acevedo-Trejos *et al.* 2013; Soininen & Luoto 2014). Sampling of microbial organisms has however started much later, and still remains too sparse to yield global biogeographical maps for phytoplankton. In situ sampling of marine microbial communities provided either coarse-grained macroecological biodiversity patterns (like latitudinal or temperature dependence (Fuhrman *et al.* 2006; Chust *et al.* 2013; Martiny *et al.* 2013) or very detailed local snapshots.

Satellite data, and in particular bio-optical anomalies derived from ocean color, have been recently exploited as a possible source of information on global phytoplankton diversity patterns (De Monte *et al.* 2013). Compared to in situ observations, data acquired by remote sensing contain very approximate and indirect biological information. On the other hand, though, they offer a high resolution (km/day or smaller) and have global coverage.

3.3 Objectives

In this study we use a numerical model to study the extent to which an index defined on minimal -but spatially resolved- information about the planktonic community -the 'seascape' diversity index- reliably assess local biodiversity. Both the physical and ecological dynamics underpinning global planktonic diversity patterns are accessible in the model, allowing us to uncover the physical basis of the correspondence between indexes, and to identify the regions of the global ocean where such a correspondence is most accurate. We show that the 'seascape' diversity index agrees with different, locally computed biodiversity indicators in great part of the world ocean, and it is best in moderately to highly energetic regions and out of blooms.

3.4 How a local alpha diversity and an area-based diversity relate and why

3.4.1 Quantitative relationship between local and seascape diversity of virtual species

A 'seascape' biodiversity index can be defined as the Shannon index of the distribution of the relative spatial extension, within a disc of given radius, of areas dominated by a given community (see Chapter 2.1.2). The choice of basing an index on such minimal ecological information stems from the features of the data that are and will be available by remote sensing. Algorithms processing ocean color indeed attribute to every pixel a tag that reflects some properties of the local ecosystem (different dominant planktonic types, e.g. diatoms and coccolitophores, are

distinguished in the classification), but have no power in describing finer features of the community (Alvain *et al.* 2008; Alvain *et al.* 2012).

We use numerical simulations issued by a physical global circulation model coupled to a planktonic virtual ecosystem of 78 phytoplanktonic ‘types’ belonging to 5 distinguished 'functional groups', and two zooplanktonic predators (see Chapter 2.1.1) (Follows *et al.* 2007; Dutkiewicz *et al.* 2009; Barton *et al.* 2010). This model has proved to reproduce realistic biogeographical patterns of dominant species, and grants a simultaneous access to the local community composition and to the seascape index. We can thus compare directly different biodiversity indexes over the whole ocean extent, thus including regions with different physical and ecological dynamic regimes.

Local biodiversity is evaluated by computing the Shannon index based on the abundances of virtual species or of functional groups (Hooper *et al.* 2005; Follows & Dutkiewicz 2011; Barton *et al.* 2013). We consider different levels of 'taxonomic' resolution because the former best reflects the measures obtained by in situ sampling and sequencing, whereas the latter is expected to correlate better with the seascape index, being both derived from group-level abundances. The computation of diversity indexes are performed on daily maps at each grid point (1 degree resolution for most of the analysis presented here) and then averaged over 3 years of model runs in order to obtain the climatological global maps (Fig. 18 and 19).

The comparison of the local and seascape biodiversity reveals strong correlations (Pearson’s correlation test $r = 0.67$ vs types and $r = 0.76$ vs functional groups, $p < 10^{-9}$). It is surprising that even though the seascape diversity is better related to local diversity computed at the level of functional groups, a drastic increase in taxonomic resolution of the community (from 5 classes to 78 virtual types) is not associated to a drop in such a correlation.

The bivariate histogram plots of the seascape and local biodiversity indexes (Figs. 18 and 19) displays a nonlinear relationship: the seascape index decreases at a faster rate than its local counterparts, while still providing a correct estimate when the Shannon index is strictly zero (only one species is present in the area). This is an expected consequence of the fact that the finite number of pixels in the considered area for the seascape diversity imposes a lower limit to the resolution of the seascape diversity indicator. Moreover close to the limit of resolution finite-size

effects render the estimation less effective. This limit does not exist for the local indexes, computed on continuous abundances. In order to correct for this bias and to improve the agreement in regions of high diversity, where the relationship between the indexes is expected to be linear, we fit the data with a nonlinear function that goes close to the origin and saturates to a straight line (see Chapter 2.1.2 and Fig. 18 and 19). This relationship allows to explain 54% and 69% of the local biodiversity variability in terms of the seascape index ($R2 = 0.54$ vs types and $R2 = 0.69$ vs functional groups, $p < 10^{-9}$).

We repeated the previous analysis changing the radius of the disk upon which seascape diversity is evaluated. Being seascape diversity generated by horizontal transport, we expect the goodness of the match to local diversity to drop off when such a distance exceeds the reach of mesoscale mixing. On the other hand, such a radius cannot be too close to the spatial resolution, otherwise the signal would be blurred by noise. The goodness of fit indeed steadily degrades when the radius is increased until 5 degrees, and to a larger extent for species than for groups (radius 100Km $R2 = 0.69$; 200Km $R2 = 0.68$; 300Km $R2 = 0.67$; 400Km $R2 = 0.65$; 500Km $R2 = 0.63$ vs functional groups; $p < 10^{-4;}$ radius 100Km $R2 = 0.54$; 200Km $R2 = 0.52$; 300Km $R2 = 0.50$; 400Km $R2 = 0.49$; 500Km $R2 = 0.47$ vs phytoplankton types; $p < 10^{-4}$). This is however accompanied by a latitudinal shift in the distribution of errors, reflecting the fact that the radius associated to mesoscale turbulence, the Rossby radius, is smaller at higher latitudes, whereas larger radii are more appropriate to describe mixing in tropical regions of the ocean (Fig. 20 and 21). This dependence can be taken into account when applying this method of biodiversity estimation to a particular region of the ocean, whereas for global studies a radius of one degree appears appropriate.

To be able to define biodiversity hotspots using seascape, we distinguished hotspots regions of local biodiversity based on the global threshold of the spatial-based index = 0.16 (Fig. 22), that determines just hotspots locations in local biodiversity (Fig. 23). A majority of consistency was evident between hotspots identified by seascape and local diversity (Fig. 24), but few regions exist where high local and seascape diversity do not perfectly overlap. A fixed threshold for a global analysis can underestimate hotspots in regions characterized by high biodiversity with respect to regional maxima but not with respect of global maxima. Indeed, the main trend of the local biodiversity global pattern is detected for both the functional groups and phytoplankton

types with minor misclassifications, even if the degree of diversity inside the hotspots may not be the same.

Although based on minimal knowledge of the biological community composition, the seascape index compensates the lack of ecological detail with spatial resolution. The trade-off between these two different sources of information is possible due to transport and mixing that transform spatial patchiness into local diversity on a time scale that is shorter than that of competitive exclusion. This implies that the estimate will not be equally accurate in all regions of the ocean, since the intensity of turbulent dynamics, as well as the distribution and demography of dominant types vary considerably.

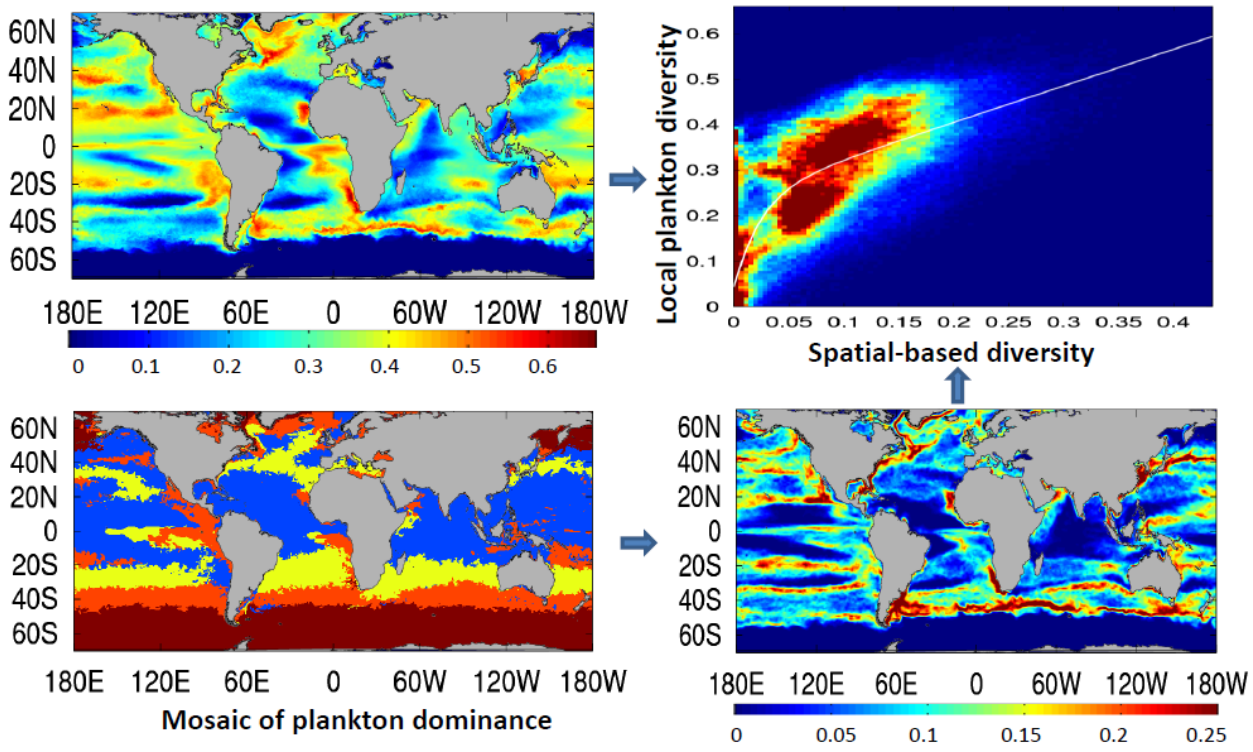


Figure 18. Bivariate histogram showing values frequency (color scale) and relationship between seascape diversity information (estimated at 1 degree disk radius) (x axis) with its relative map and local biodiversity estimated by Shannon index based on ECCO2-Darwin model functional groups (y axis) with its relative map. The distribution of the dominant groups (mosaic of plankton dominance) resulted from the model is also showed.

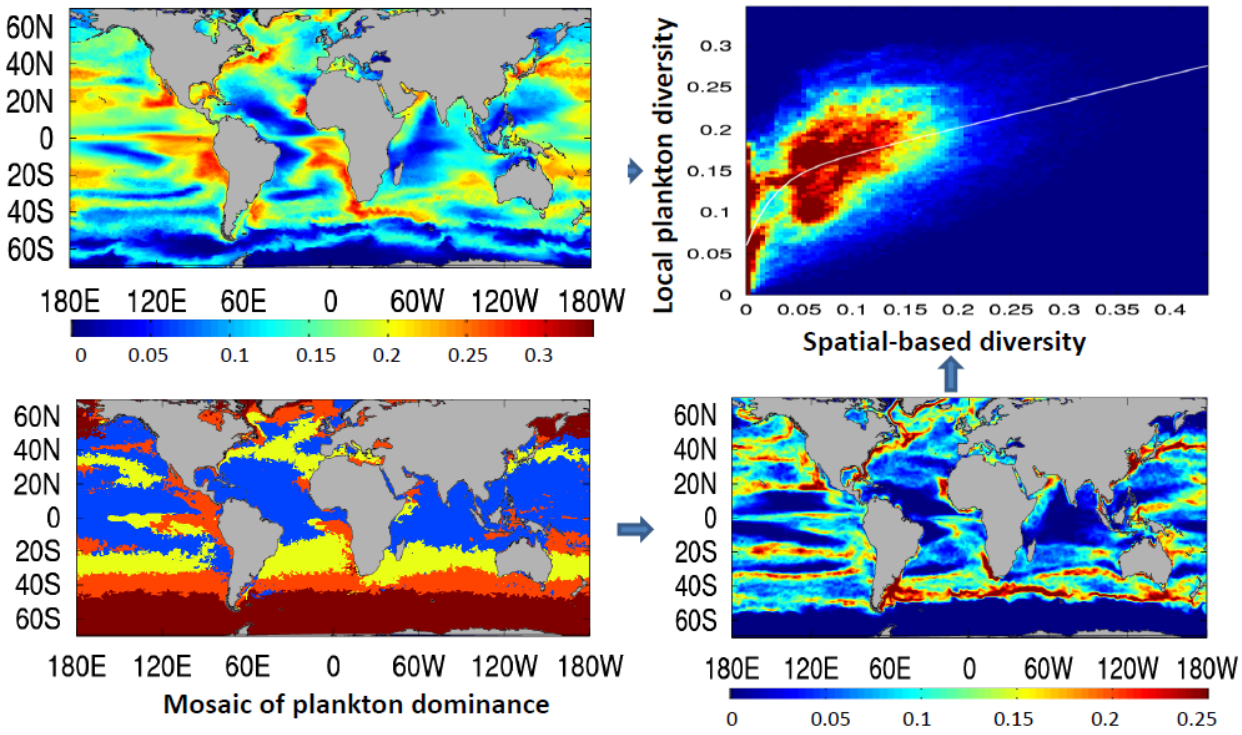


Figure 19. Bivariate histogram showing values frequency (color scale) and relationship between seascape diversity information (estimated at 1 degree disk radius) (x axis) with its relative map and local biodiversity estimated by Shannon index based on ECCO2-Darwin model phytoplankton types (y axis) with its relative map. The distribution of the dominant groups (mosaic of plankton dominance) resulted from the model is also showed.

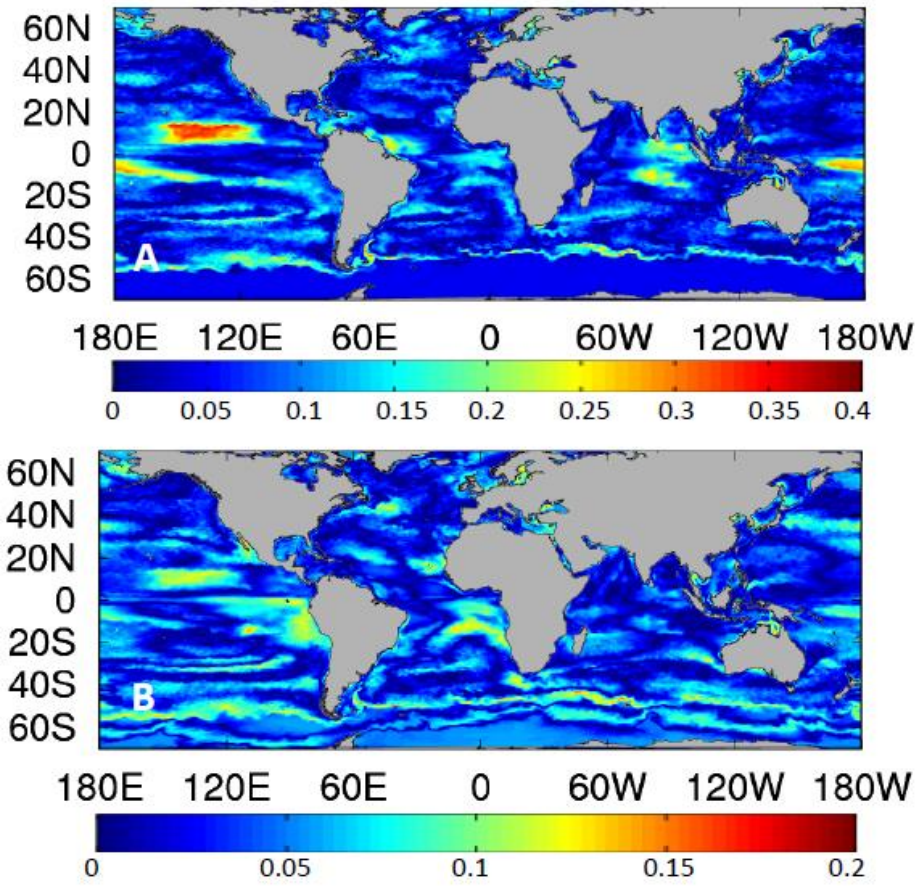


Figure 20. Global map of residuals after regression analysis between seascape diversity information (estimated at 1 degree disk radius) and local biodiversity estimated by Shannon index based on ECCO2-Darwin model functional groups (A) and phytoplankton types (B).

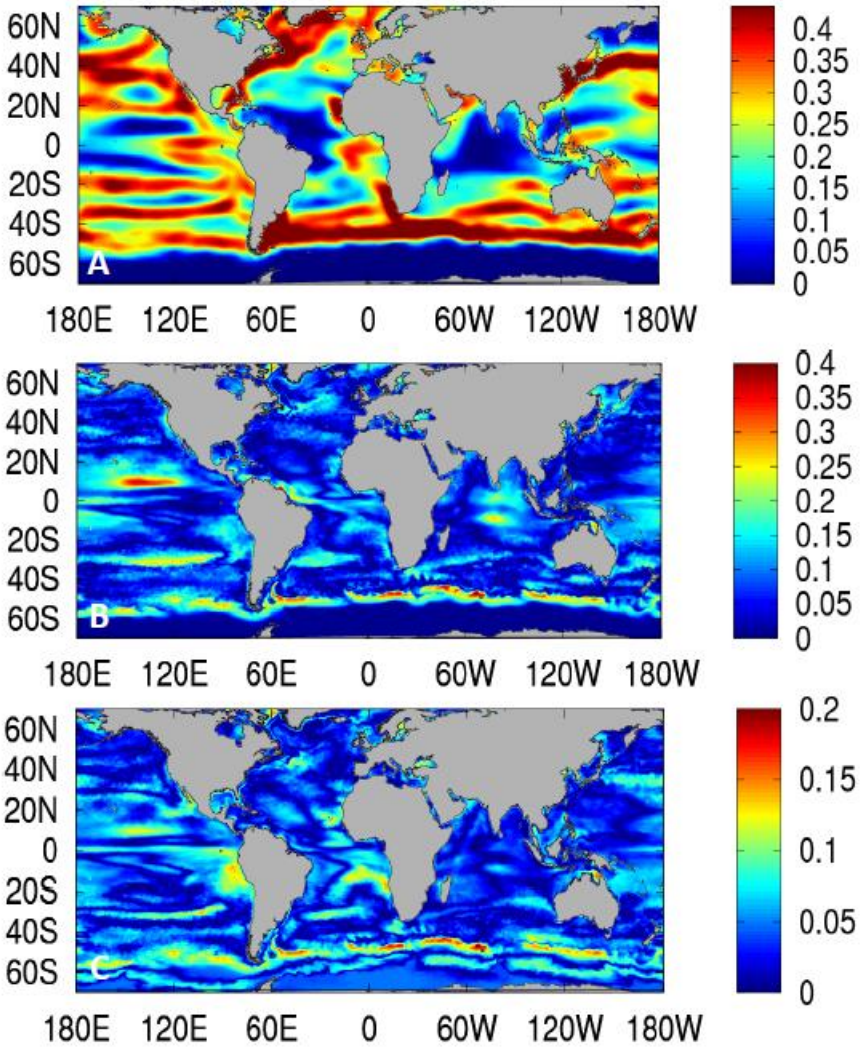


Figure 21. A) Global annual climatology map of biodiversity hotspots estimated by seascape diversity information at 5 degree disk radius. Global map of residuals after regression analysis between seascape diversity and local biodiversity estimated by Shannon index based on ECCO2-Darwin model functional groups (B) and phytoplankton types (C).

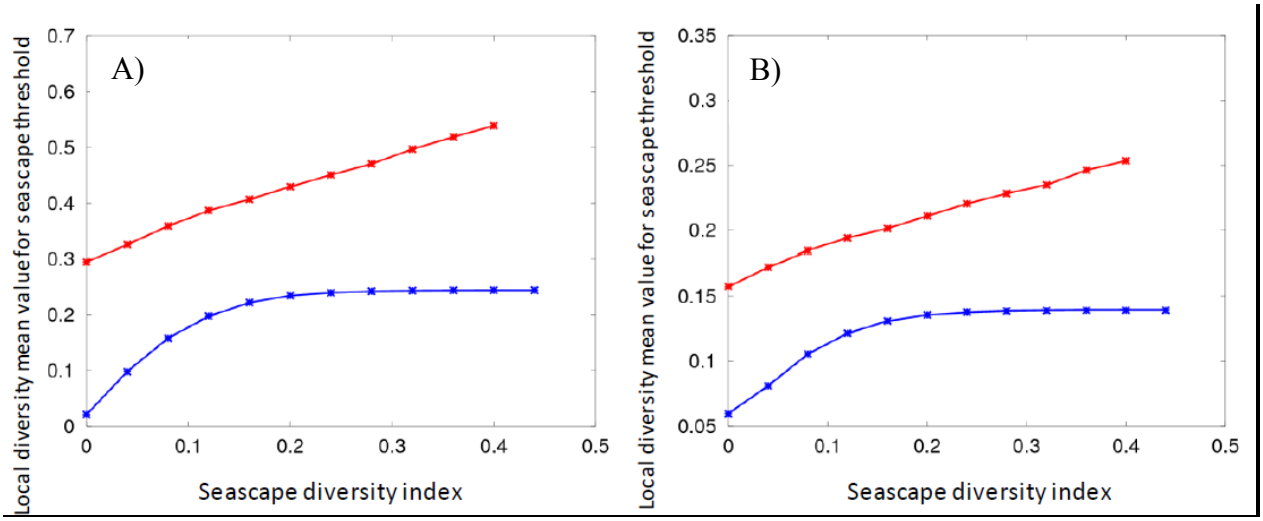


Figure 22. Determination of the threshold used to define biodiversity hotspots based on seascape diversity. X axis are values of seascape diversity index, y axis are mean values of local biodiversity A) estimated by Shannon index based on ECCO2-Darwin model functional groups and B) phytoplankton types, for hotspots locations (red) and non-hotspots locations (blue) defined by the seascape diversity threshold value in x. The optimal threshold value coincides with the non-hotspot mean starting to be asymptotic while the hotspots mean continues to increase, indicating that after that threshold, only hotspots locations for both indexes are considered and they coincide.

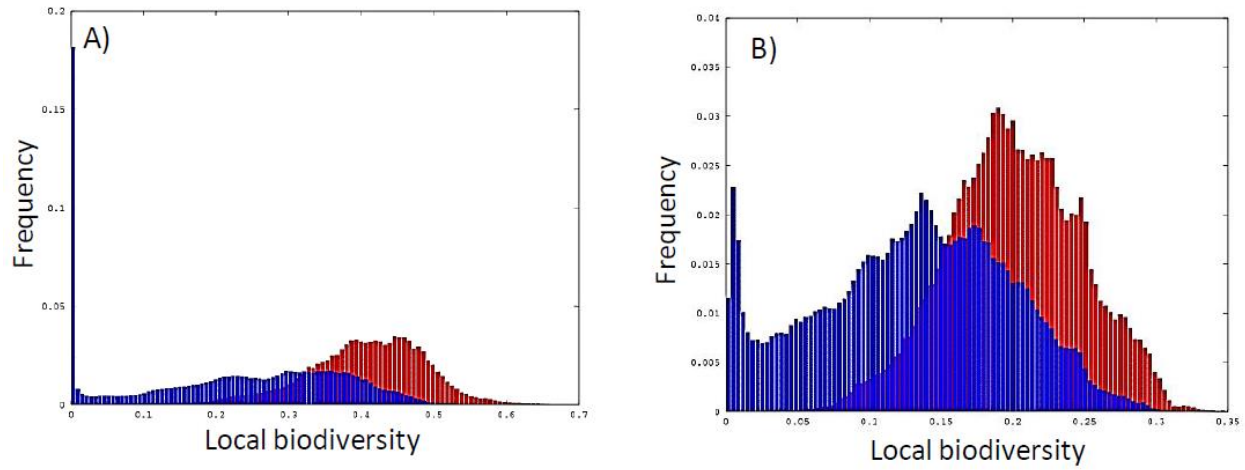


Figure 23. Frequency histogram for values of local biodiversity classified as hotspots (red) and non-hotspots (blue) based on the optimal seascape diversity threshold 0.16 estimated in Fig 21. A) local biodiversity from functional groups, B) local biodiversity from phytoplankton types.

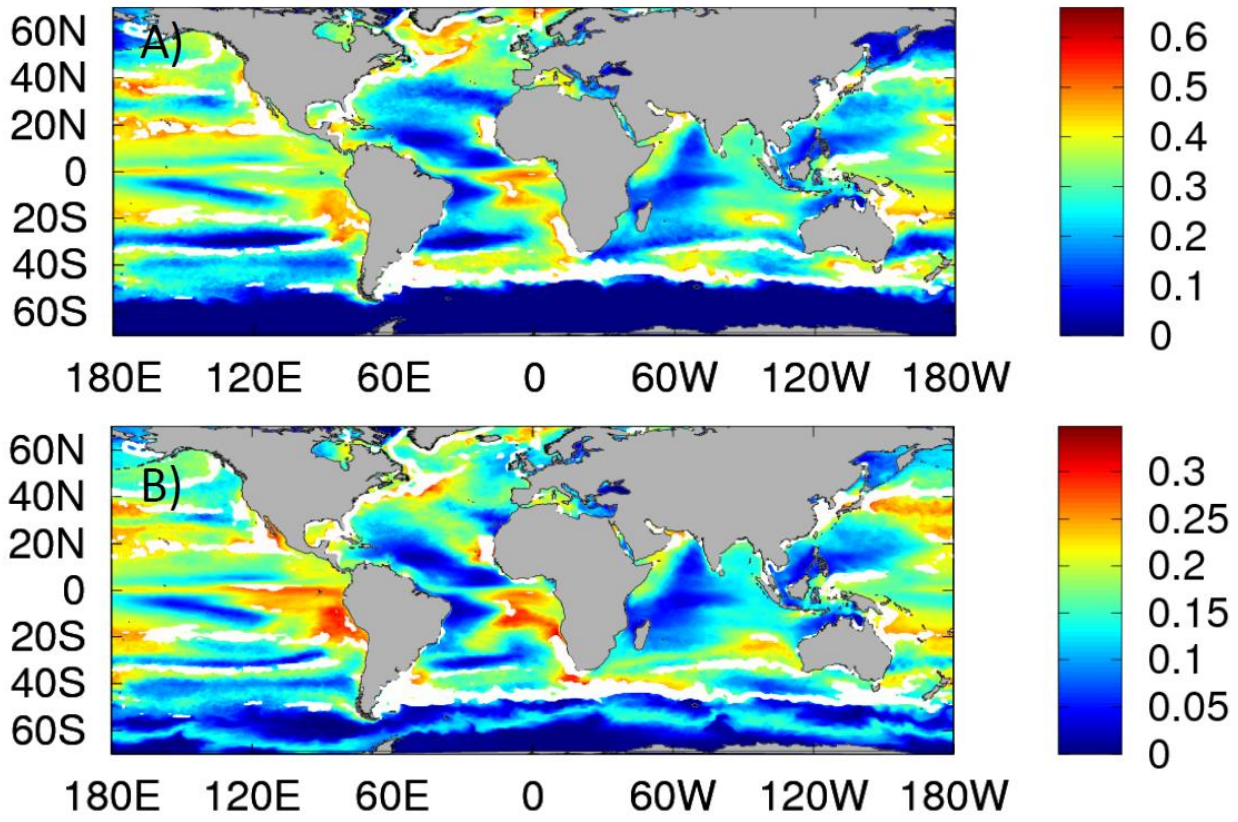


Figure 24. Hotspots (in white) estimated by seascape diversity information threshold 0.16 (estimated at 1 degree disk radius) are overlaid to local biodiversity estimated by Shannon index based on ECCO2-Darwin model functional groups (A) and phytoplankton types (B).

3.4.2 Environmental factors and accuracy of the estimation of the proxy

For every location in the ocean, the model provides, besides biodiversity, a number of physical and ecological parameters. These parameters correspond to features that are globally available by remote sensing, and can be used to predict the expected error associated to using the seascape diversity as a proxy for local diversity. Since environmental variables are inhomogeneously distributed in the global ocean, the statistical analysis of their associations to errors in the estimate can uncover trends, useful to point out the mechanistic base of the correspondence between indices, rather than provide a method for quantitative error estimation.

We analyzed the residuals of the fit in Fig.18 and 19, whose global distribution is displayed in Fig. 20, with respect to three physical variables (Sea Surface Temperature, SST gradient and Eddy Kinetic Energy), nutrient concentrations (nitrites, nitrates, ammonium, phosphates,

silicates, iron) and a biological variable (surface total chlorophyll concentration) (Fig. 25 and 26). The average error introduced by the seascape proxy is then evaluated for points corresponding to the same range of the environmental variables. As shown in Figure 25, the general trends in the dependence of the error on environmental parameters is the same for the two local proxies at different taxonomic resolution, with the exception of the EKE that we discuss below.

The error made in using the seascape diversity index as a proxy for local diversity is largely independent of temperature, indicating no latitudinal bias in the estimation (Fig 25a). However, it depends on SST gradient, and declines as the temperature difference increases (Fig 25b). This result is consistent with the fact that the seascape index evaluates the biodiversity component that is generated by mixing, as SST gradients characterize transport fronts corresponding to the boundaries between water masses of different origin. Regions with a low SST gradient typically have low spatial diversity, and thus follow in the range where the seascape index hits the lower boundary of its resolution, thus clumping in the same class cases having different, and low, local diversity (Levy *et al.* 2014). The entity of residues decreases with EKE (Fig. 25c), that quantifies the energy that is dissipated in small scale turbulence and fuels mixing below the mesoscale. Apart from a few points of extremely high EKE, that cover 0.01% of the ocean surface, the error in estimating species-based diversity declines as the energy increases. Contrary to the dependence on the SST gradient, there is a sudden drop in the error at about 0.2 m/s. This implies that the seascape diversity index represents best the local species diversity for moderate to high EKE values, that is if mixing occurs on a sufficiently rapid time scale. If the local biodiversity is instead evaluated based on types, then the error remains low for all EKE values, apart from the regions of very high energy that cover 0.05% of the ocean surface. This result hints to the fact that in regions of low mixing there is a possible coexistence of different virtual species within the same functional group. By mapping such regions in the global ocean, these regions correspond to major ocean gyres, where coexistence of different Phrochlorococcus-analogue ecotypes is known to occur in the model and in the real ocean (Follows *et al.* 2007). More puzzling is the increase of the error with chlorophyll concentration (Fig 25d), which may be a consequence of the fact that concentrations above 0.5 mg/ml are typical of regions characterized by strong planktonic blooms. In those cases, that correspond to 1.8% of the ocean surface, competitive exclusion is probably the dominant ecological process. According to what happens also in in situ measures of diversity

during a bloom, local diversity as estimated by the Shannon index drops because of the overwhelming abundance of the dominant type. High levels of chlorophyll concentrations are typically discarded by classification algorithms based on ocean color because of lack of resolution in the spectral anomalies used to detect the functional groups. It is hence likely that regions where the seascape proxy incurs an increase in error are actually not represented in the estimations by remote-sensing.

The dependence of the estimation error on nutrients does not yield any pattern. Indeed it is not expected that single chemical species are good indicators of the amount of turbulence or disturbance of the ecosystem.

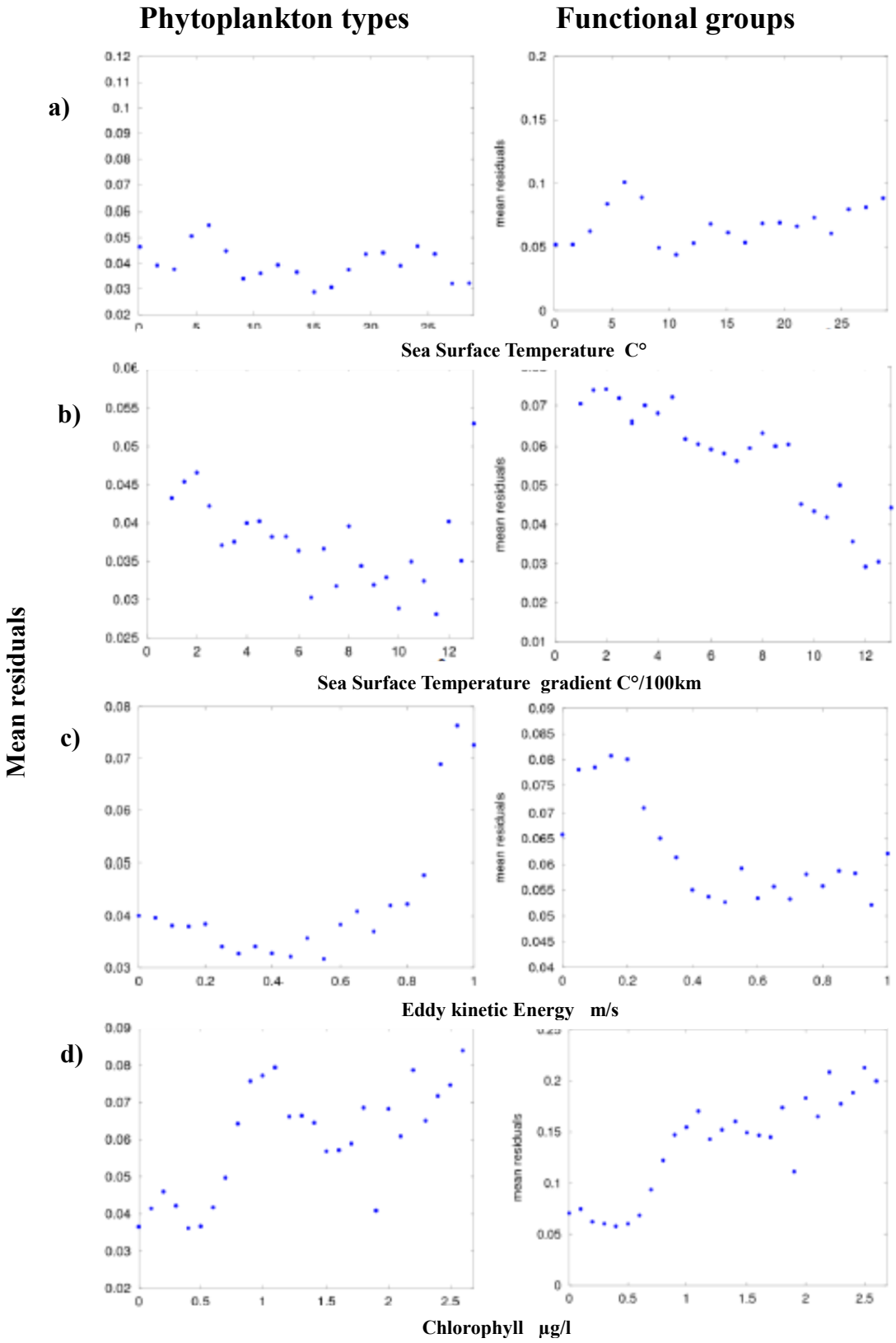


Figure 25. Study of the residuals of the regression between seascape diversity information

(estimated at 1 degree disk radius) and local diversity estimated by Shannon index based on ECCO2-Darwin model. The residual values have been averaged inside classes depending on the environmental characteristics present at each location. Extreme bin values that include less than 0.1‰ of the ocean extent are not showed. A) annual climatology of Sea Surface Temperature (C°); B) annual climatology of Sea Surface Temperature gradients (C°/100km); C) annual climatology of Eddy Kinetic Energy (m/s); D) annual climatology of Chlorophyll (µg/l).

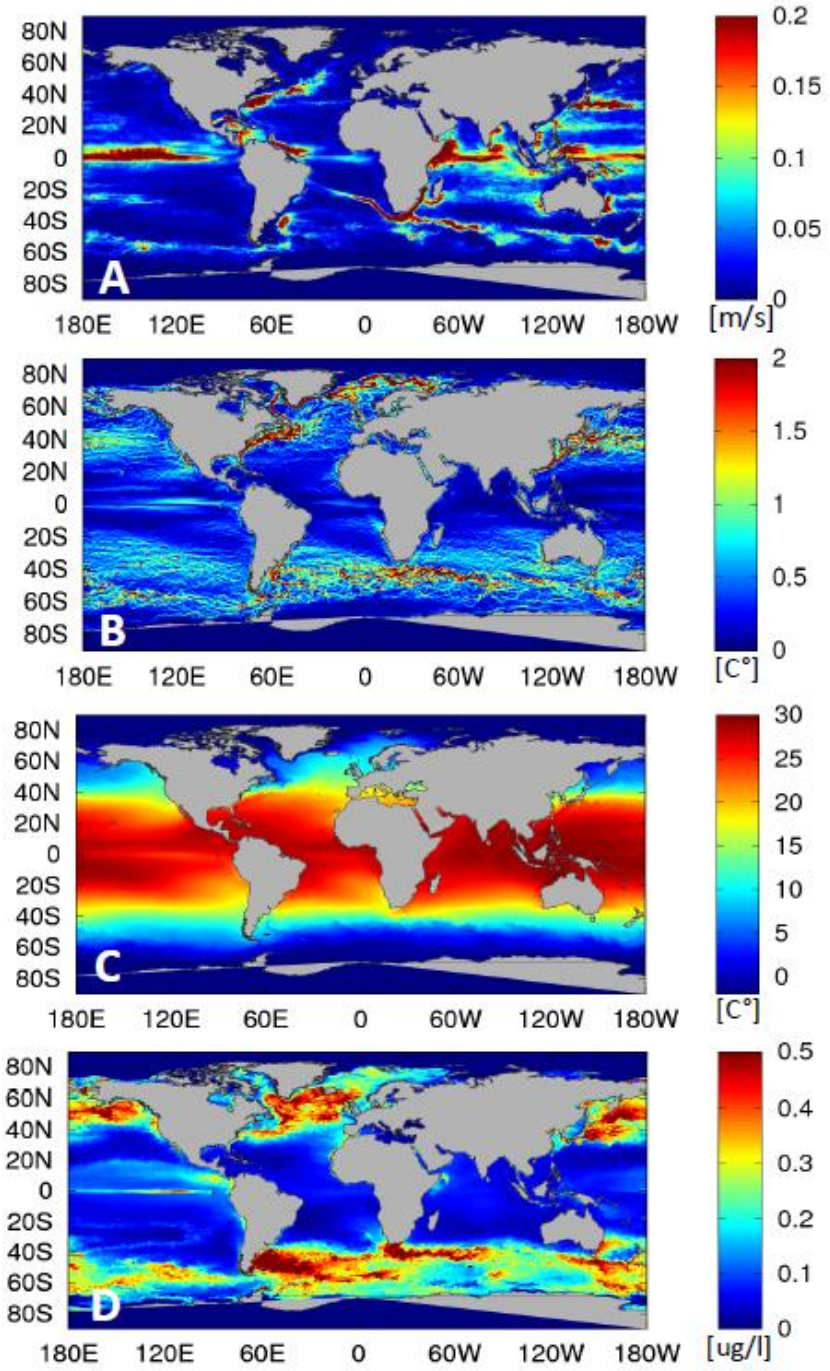


Figure 26. A) Map of average Eddy Kinetic Energy of the model for the 3 years of integration. B) Map of the average SST gradients of the model for the 3 years of integration; C) Map of the average SST of the model for the 3 years of integration; D) Map of the average Chlorophyll concentration of the model for the 3 years of integration.

3.5 Discussion and conclusion

The patchy nature of the environment has been recognized long ago as a possible determinant of higher biodiversity for terrestrial organisms (Hutchinson & Macartur 1959). At that time, the open ocean upper waters were considered much less rich in diversity and "uniform and quasi horizontal" (Hutchinson 1958). In the following fifty years, remote sensing of the open ocean has revealed that inhomogeneities as large as those observed across basins also occur on much smaller spatial scales (~10-100km) and on temporal scales comparable to the demography of phytoplankton (days to weeks). These seascape inhomogeneities form a mosaic analogous to what one sees on land - but dynamically changing much faster than terrestrial environmental patches. Concomitantly, oceanic microbial diversity has been found as being higher than evaluated previous to the high-throughput analysis of microbial communities (Venter *et al.* 2004), and at the same time the species range to be so vast to question the importance of local adaptation in structuring the community. Together, these observations point to the fact that physical forcing and plankton demography concur in building up diversity through environmental fluctuations and dispersal. The interactions between transport and ecology have been shown to take place during blooms, that are events when one would expect competitive exclusion to take place faster (d'Ovidio *et al.* 2010). However, when the physical and demographic time scales entwine, the ecosystem is kept away from its local optimum by changes in the features of the water masses, as well as by immigration from the surrounding regions. If it is impossible to use remote sensing to follow the rearrangement of the local planktonic community, one can use information about the spatial variability of the ecosystem to infer the part of biodiversity that is a consequence of horizontal mixing alone. Although an index based on such an idea has been proposed (De Monte *et al.* 2013) and shown to follow a macroecological pattern in its temperature dependence, the demonstration that the underlying ansatz is valid at the global scale seems out of reach.

We have therefore turned to numerical models to demonstrate that a seascape diversity index -- measurable by satellite -- indeed reflects the local diversity accessible by in situ sampling. The Darwin model, that simulates simultaneously ocean circulation and the ecology of planktonic

ecosystems, was used as testbeds to determine the extent to which local biodiversity can be inferred starting from crude, but spatially extended, measures of the planktonic community.

Our results show that, in spite of the minimal ecological detail, a proxy for local biodiversity can be derived by information on the spatial distribution of dominant phytoplanktonic types. Our analysis evidenced that the correspondence between local and seascape diversity is not limited to considering the same depth of community description: similar results hold if local diversity is assessed at the level of virtual species or of functional groups, whereas only the latter are used to estimate the seascape index. Analogously, one can expect that the same mechanism holds for every species that is passively drifting, and not only for photosynthetic ones, so that hotspots identified by using bio-optical signatures may extend from primary producers to other component of the ecosystem, such as zooplankton and heterotrophic microbes.

Considered in relation to environmental variables, the discrepancy between local and seascape diversity indexes appears to be minimal in regions where the turbulent energy of horizontal currents enhances mixing at the submesoscale. These regions can be identified by remote-sensing measurements of the eddy kinetic energy, and cover 30% of the ocean surface. This study hence paves the way to the systematic use of remote sensing in detecting and monitoring regions where high biodiversity is enhanced by turbulent mixing.

The mechanism of transport-driven biodiversity enhancement is expected to be effective for passively transported organisms whose time scale of competitive exclusion is longer than the typical lifetime of submesoscale filaments (of the order of weeks). Whether this scenario is only possible under conditions of neutral selection (Hellweger *et al.* 2014), or it is compatible with selection acting on ephemeral niches (Sauterey *et al.* 2014), is an interesting evolutionary problem that is still open.

Hotspots of seascape diversity are fed by communities that prosper in neighboring locations, and therefore are expected to be maintained even if the ocean circulation is modified as a consequence of climatic changes. The opportunity of identifying such hotpots as priorities for conservation thus depends crucially on the extent to which transport-induced diversity propagates to higher levels of the trophic chain. In order to assess such a possibility, further investigations are needed on the link between biodiversity hotspots of species that respond differently to the

physical forcing of oceanic currents.

Chapter 4 Ecological relevance of remote sensing. Biodiversity hotspots estimated from space.

Understanding the variability of marine microbial biodiversity is a fundamental challenge for biologist and oceanographers. Current observational programs are based on in situ studies, but their implementation at the global scale is difficult, due to the ocean extent, its temporal variability and the heterogeneity of the data sources on which collections are built. Here, I present the possibility of identifying phytoplankton biodiversity hotspots using remote sensing. I define a Shannon entropy index based on patchiness in ocean color bio-optical anomalies. This index provides an increased resolution compared to most available approaches, global coverage and temporal variability. It shows a relation to temperature and mid-latitude maxima in accordance with those previously evidenced in microbiological biodiversity model and observational studies. Regional maxima are in remarkable agreement with several known biodiversity hotspots for plankton organisms, as well as with some in situ planktonic biodiversity estimates (from Atlantic Meridional Transect cruise). These hotspots do not show a relationship with chlorophyll concentration, although temporal variability estimated as successions of dominant types are instead more related to chlorophyll. These results encourage to explore marine biodiversity by integrating approaches coming from different research disciplines.

4.1 Introduction

Defining global patterns of marine microbial diversity is a fundamental ecological issue, with important implications in understanding biogeochemical and climatic processes. However, the nature of the open ocean realm does not allow to strictly define boundaries for marine biodiversity hotspots. This difficulty is due to the multiplicity of ecological and physical scales involved in shaping oceanic ecosystems (Levin *et al.* 1992).

Plankton biogeography and biodiversity is commonly assessed using in situ sampling. The locally observed species can be characterized at the morphological, functional, and molecular level of biological organization, while covariates related to their presence are found in the local physico-chemical context (Jones 2007). Nonetheless, in situ sampling remains still too sparse to produce

global maps of phytoplankton species distribution. Until now, plankton sampling has unveiled either macroecological patterns (like latitudinal or temperature dependence of biodiversity) or very detailed but local snapshots of microbial diversity.

The relationship between biodiversity and latitude and temperature gradients has been highlighted in many studies on both land and marine ecosystems (Hillebrandt 2004). The kinetic energy or temperature hypothesis predicts positive correlations with temperature, especially for ectoterms. It postulates that at higher temperatures increased metabolic rates may promote higher rates of speciation, leading to greater diversity. Further, range limits are set by thermal tolerance, with more species tolerant of warm conditions. In the ocean, latitudinal gradients have been shown to exist for species richness of planktonic marine bacteria (Fuhrman *et al.* 2008), and for functional diversity (Raes *et al.* 2011). However, quantitative differences in the latitudinal profiles, related to the relatively small number of observations and the heterogeneity of the conditions in which these have been performed, do not allow to go much beyond the qualitative conclusion that diversity decreases at high latitudes and that a mid-latitude maximum may exist.

Another well studied aspect of biodiversity is the relationship with the productivity of an ecosystem. Large meta-analyses of terrestrial and aquatic ecosystems suggest that the shape of the productivity–diversity relationship is generally either positive or unimodal (Dodson *et al.* 2000; Irigoien *et al.* 2004; Gillman & Wright 2006). For marine phytoplankton, there are less data available but a few large-scale studies also suggest a unimodal response with maximum diversity peaking at intermediate levels of productivity and minimum diversity during massive blooms that escape grazing predation (Vallina *et al.* 2014). Nonetheless, agreement exists on the fact that the trend of this relationship depends on the considered spatial and temporal scale (Chase & Leibold 2002; Whittaker & Heegaard 2003).

The global distribution of phytoplankton has mostly been addressed by remote sensing of ocean at various spatial and temporal scales. Beyond the estimation of chlorophyll concentration, different spectral analyses allow to characterize phytoplankton biogeography at a resolution of few tens of kilometers and few days. Among those, the PHYSAT algorithm (Alvain *et al.* 2008) clusters spectral anomalies into classes of bio-optical properties. These classes correlate with the dominant phytoplankton types and provide a simple tool for identifying niches that sustain

distinct planktonic communities (d'Ovidio *et al.* 2010) and pave the way to studies that may go beyond the description of macroecological and biogeographical patterns.

4.2 Objectives

Here I propose an index based on remote sensing as a proxy for diversity induced by horizontal stirring, and show that such an index is consistent with many features of in situ studies of global biodiversity and macroecological patterns in the open ocean. To meet these objectives, I combine information on biological heterogeneity and fluid dynamics in order to estimate presumed biodiversity hotspots. I use patchiness in maps of bio-optical anomalies as an indicator of niche spatial heterogeneity, and assume that patches that are about 100 kms apart are typically mixed within a few weeks by oceanic turbulence (see Chapter 2.2).

4.3 Plankton community dominance and diversity

4.3.1 Reanalysis of chl spectra: towards an information on biogeography and diversity of planktonic communities

Global distributions of functional dominant phytoplankton groups derived from the Physat reanalysis of chlorophyll spectra (Fig. 27) have been used to derive information about diversity of the planktonic community. The diversity index created, τ index, is a local measure of the heterogeneity in the ocean surface bio-optical properties at a certain scale, that is the radius of a disk over which the occurrence frequencies of remote-sensed bio-optical frequencies are estimated.

The possibility to observe a higher diversity at one point in the ocean, the more heterogeneous the distribution of planktonic communities - in this case, emerging as bio-optical anomalies - around that point, is at the base of the conception of the τ index as a possible proxy of biodiversity. Consequently, if the conditions at a given location are such to support high diversity, these conditions will reflect on the ability of that local community to prosper in nearby niches shaped by transport, thus creating high spatial heterogeneity in the nearby environment.

The global map of τ (Fig. 28) shows strong global patterns that extend across oceanic basins. High τ diversity occurs both in productive systems (upwellings; western boundary currents; islands'wakes; North Sea) and in moderately productive and oligotrophic regions (SW and NE

Pacific; North Pacific transition zone), highlighting two kinds of hotspots: confluence regions - where nearby communities are mixed by horizontal stirring - and geographical features, such as islands or reefs.

The diversity index τ has been computed within disks of 100 km radius. Quantitatively, water masses that are 100 kms apart are mixed on a time scale of the order of few weeks, comparable to that of a planktonic bloom, so that each of the communities present in the area may also be locally represented. On the other hand, communities that are locally present will be given the chance to colonize niches in that radius before the bloom season is over. This measure exploits therefore a specific feature of oceanic ecosystems at the submesoscale, where horizontal transport and mixing occur on the same time scale of planktonic blooms.

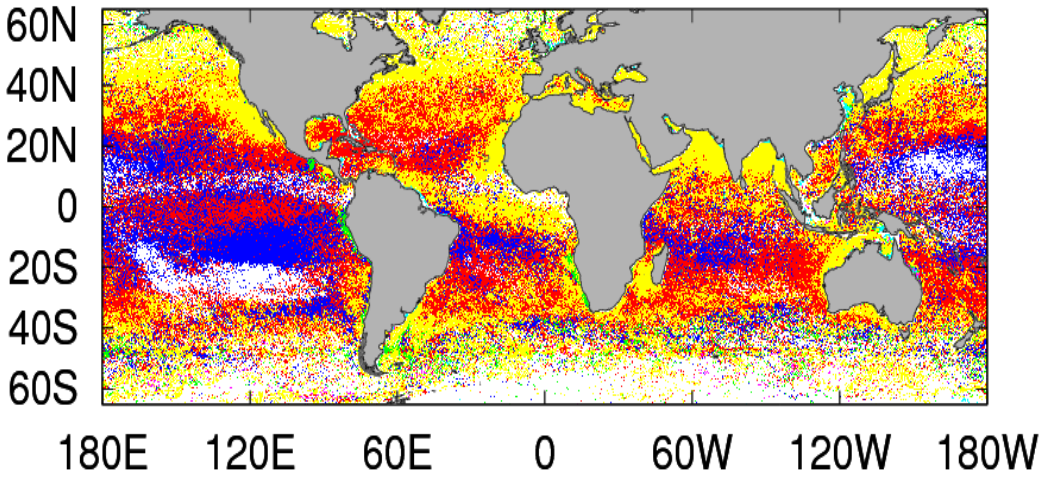


Figure 27. Global winter climatology of phytoplankton dominant functional groups as detected by Physat algorithm. Red = *Synecococcus*-like; Blue = *Prochlorococcus*-like; Green = Diatoms-like; Pink = *Phaeocystis*-like; Cyan= Coccolitophores-like; Yellow = nanoplankton-like.

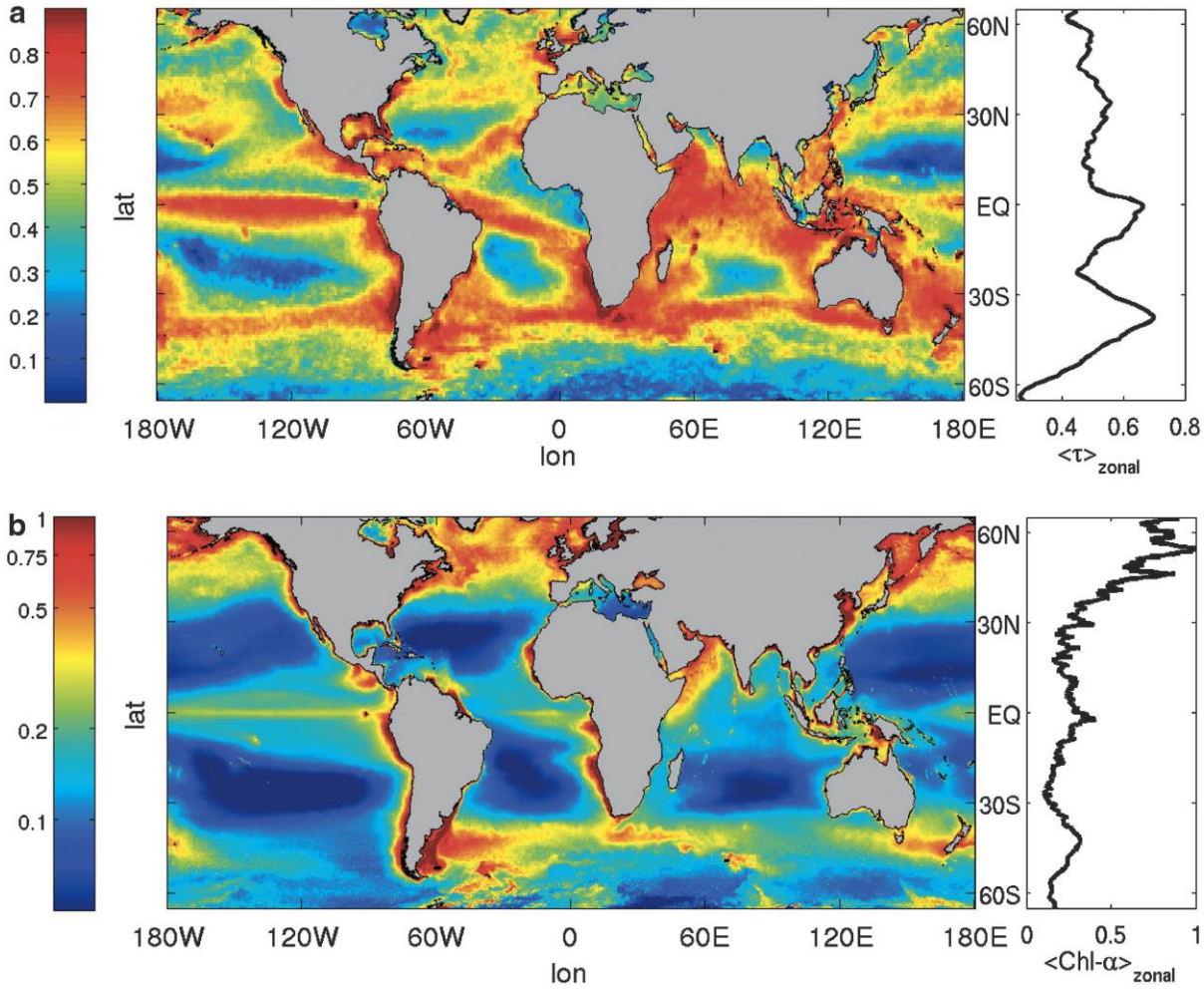


Figure 28. (a) Index τ for the global ocean: average of daily maps of τ computed over 7-day composites of PHYSAT data for the period 2003–2010. The hotspots emerging from this analysis characterize regions with a standing representation of several nearby communities, as identified by their dominant phytoplankton types (PHYSAT algorithm applied to SeaWiFS radiances). (b) Average of daily chlorophyll-a maps (SeaWiFS, 2003–2010). From De Monte, Soccodato *et al.* 2013.

4.3.2 Remote sensed τ diversity and global plankton biodiversity hotspots.

Regions of high values of τ largely overlap with known biodiversity hotspots of primary producers (Barton *et al.* 2010) and even metazoan (Tittensor *et al.* 2010). The zonal average (Fig. 28) of τ shows intermediate latitude maxima that are consistently reported in observational and model biodiversity studies (Raes *et al.* 2011). The comparison with in situ data from the AMT cruise (Fig. 29 and 30) (Irigoien *et al.* 2004; Cermeño *et al.* 2008) shows a latitudinal pattern of

nano- and microplankton biodiversity in agreement with the value of the τ index computed at the sampling locations. Qualitatively, the in situ and satellite-estimated measures of biodiversity show a comparable latitudinal alternation of maxima and minima: from north to south, a shallow minimum centered in 15:30N, then a peak in the subtropics (15:20N); an equatorial peak around 10S followed by a deep minimum; and a recovering trend for both curves when continuing from 20S further poleward (Fig. 30). However, whereas the correlation is significant in the Southern hemisphere (Fig. 31a, Pearson's $r = 0.7$, $p = 4 * 10^{-3}$), in the Northern hemisphere the quantitative agreement is lost. The localization of Northern hemisphere AMT stations includes upwelling systems where vertical processes are more relevant than in the open ocean, and where possibly stirring may not be the dominant structuring process shaping biodiversity hotspots. The importance of the Mauritanian upwelling, that is regularly crossed by the AMT transect, in affecting picophytoplankton community composition has already been reported (Zubkov *et al.* 1998). Additionally, the in situ biodiversity measure in stations 1 and 3 appears to be extreme with respect to the general trend of the temperature dependence of biodiversity (Fig.31b).

The τ index displays a highly significant positive correlation with SST (Fig. 31b, $R2 = 0.26$, $P < 10^{-4}$), with a unimodal relation, corresponding to the latitudinal trend of increase in biodiversity when moving from the poles towards the tropics, consistent with what is previously reported (Rutherford *et al.* 1999; Worm *et al.* 2005; Tittensor *et al.* 2010). Both this trend and the zonal mean of the τ index show that diversity decreases at high latitudes and that a mid-latitude maximum exists. Altogether, the AMT data provide a picture of the variation of diversity with temperature that is consistent to what displayed for the τ index. Both regression lines show a very similar trend with a maximum at intermediate temperatures (Fig 31b).

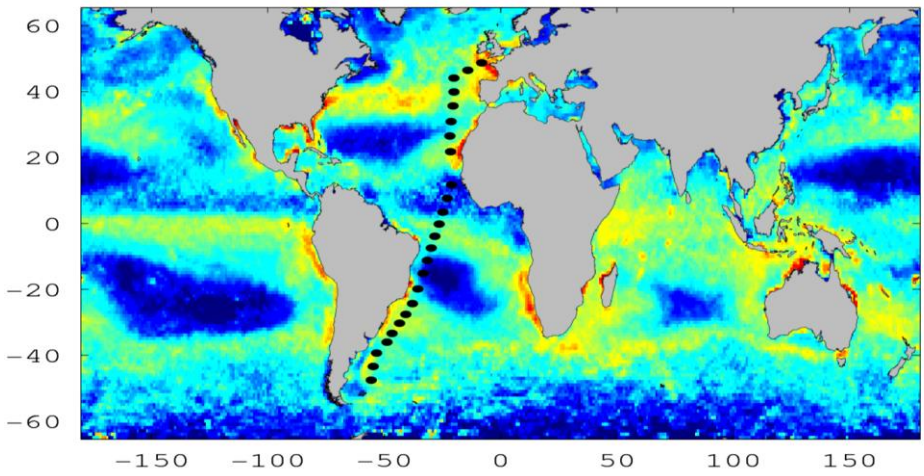


Figure 29. Map of τ diversity for the global ocean showing the Atlantic Meridional Transect run on which the relationship between remote sensed and in situ plankton diversity has been investigated. τ diversity is the average of daily maps of τ computed over 7-day composites of PHYSAT data over the cruise monthly time lag (April-May, from 2003 to 2010). The hotspots emerging from this analysis characterize regions with a standing representation of several nearby communities, as identified by their dominant phytoplankton types (PHYSAT algorithm applied to SeaWiFS radiances).

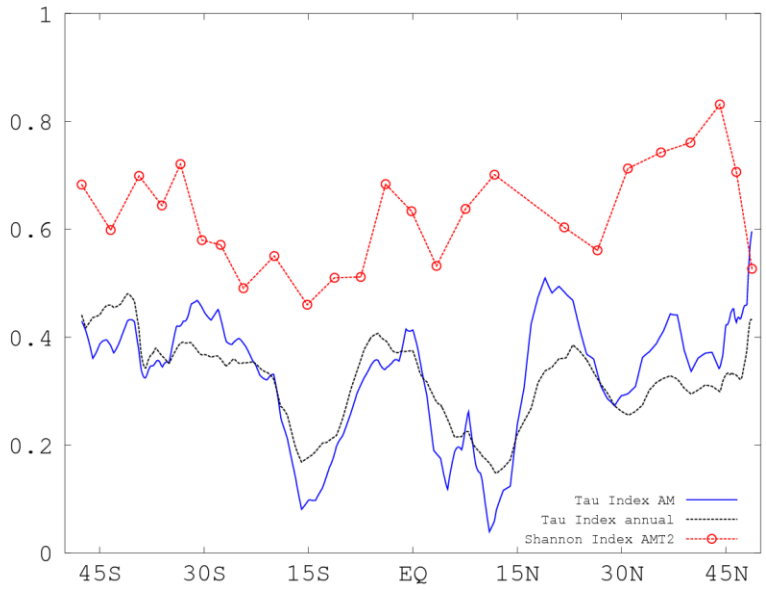


Figure 30. Diversity as a function of latitude: blue, normalized τ index calculated at the AMT transect positions and averaged over the months of April and May from 2003 to 2010; black, normalized τ index calculated at the AMT transect positions over the years 2003 to 2010; red, nano- and micro-plankton normalized Shannon index from the AMT-2 in-situ measures.

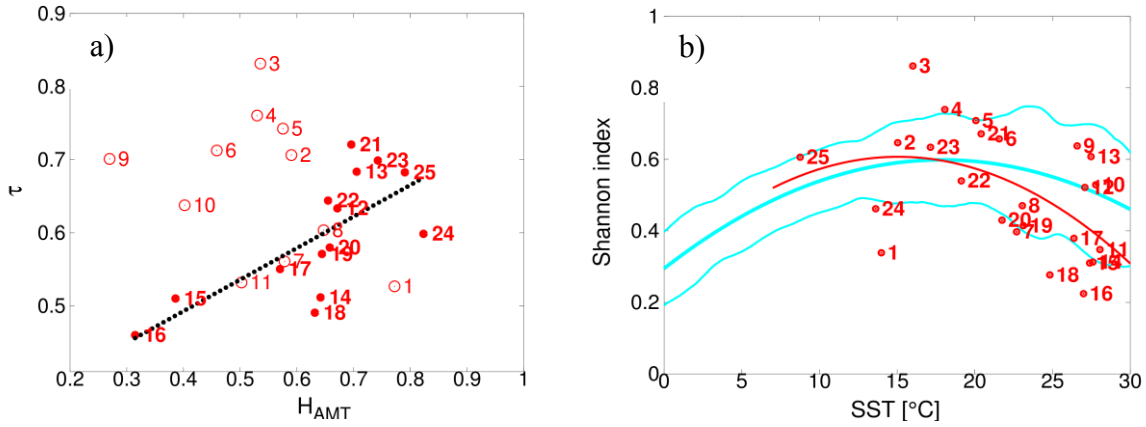


Figure 31. In figure (a) τ index versus AMT-2 cruise Shannon index (filled circles, Southern hemisphere; empty circles, Northern hemisphere). The black line is the best linear fit for the Southern hemisphere data, that shows a statistically significant correlation (r=0.7, p=4 10^{-3}). (b) Shannon index (the number of the station is indicated next to the point) versus average Sea Surface Temperature. The best quadratic fit of the points relative to the in situ measures (red line, R^2=0.26, p=0.038) is displayed together with the fit of the τ index and the confidence interval to one standard deviation (light blue lines).

4.3.3 Remote sensed temporal patterns, stability and ecological successions

Remote sensed biodiversity showed agreements with in situ data and macroecological patterns. In particular, τ index calculated on the same time lag than the AMT transect performed better than its annual average. Indeed, τ index can give also information of temporal variability of plankton diversity in the open ocean, as shown in Fig 32. The temporal change is particularly evident in areas characterized by seasonal regimes such as the temperate and subpolar regions, that host strong diatoms blooms.

The analysis of the temporal stable signal, that indicates permanents hotspots (signal-to-noise ratio), confirms the fact that regions characterized by strong, seasonal, intermittent blooms are transient hotspots of diversity and appear mostly seasonally. On the other hand, stable hotspots of biodiversity are located at the equatorial front, in the Caribbean and Central America, in the Hawaii area, in the Indian Ocean, East of Australia and in general at the transition between the oligotrophic waters of the Pacific and Atlantic gyres and the more nutrient rich waters outside the gyres (Fig. 33).

Transient hotspots regions characterized by strong seasonal blooms typically present diverse

ecological successions of dominant groups. Phytoplankton ecological successions and taxa turnover can also be described by remote sensing as shown in Fig.34, where the diversity of the dominant groups appearing at one location trough time is calculated and temperate and upwelling regions stand out. These regions are also considered among the most productive regions of the globe.

To be able to investigate the relationship between global remote sensed diversity and productivity, we compared both the τ index and the temporal diversity in successions with average climatologies of chlorophyll concentration as proxy for ocean productivity. No significant correlation was evident between diversity and productivity if all the extreme high chlorophyll values were retained. On the other hand, for low to intermediate/high regimes of productivity (average chl < 0.5 μg/l), overall quantitative statistical analyses showed a significant unimodal function such as diversity of temporal successions increasing with chlorophyll concentration and peaking at intermediate levels of productivity (Fig. 35 and 36; regression R2 = 0.25; $p < 10^{-9}$; $y = 0.562 + 4.297x - 7.005x^2$). The unimodal distribution can also significantly represent the relationship between τ index and productivity (Fig. 37 and 38; regression R2 = 0.14; $p < 10^{-9}$; $y = 0.323 + 1.970x - 3.437x^2$), although the fit is less accurate compared to the diversity of temporal successions. The relationship is stronger in the transition zones between oligotrophic and more nutrient rich waters, as showed in the map of the residuals (Fig. 38).

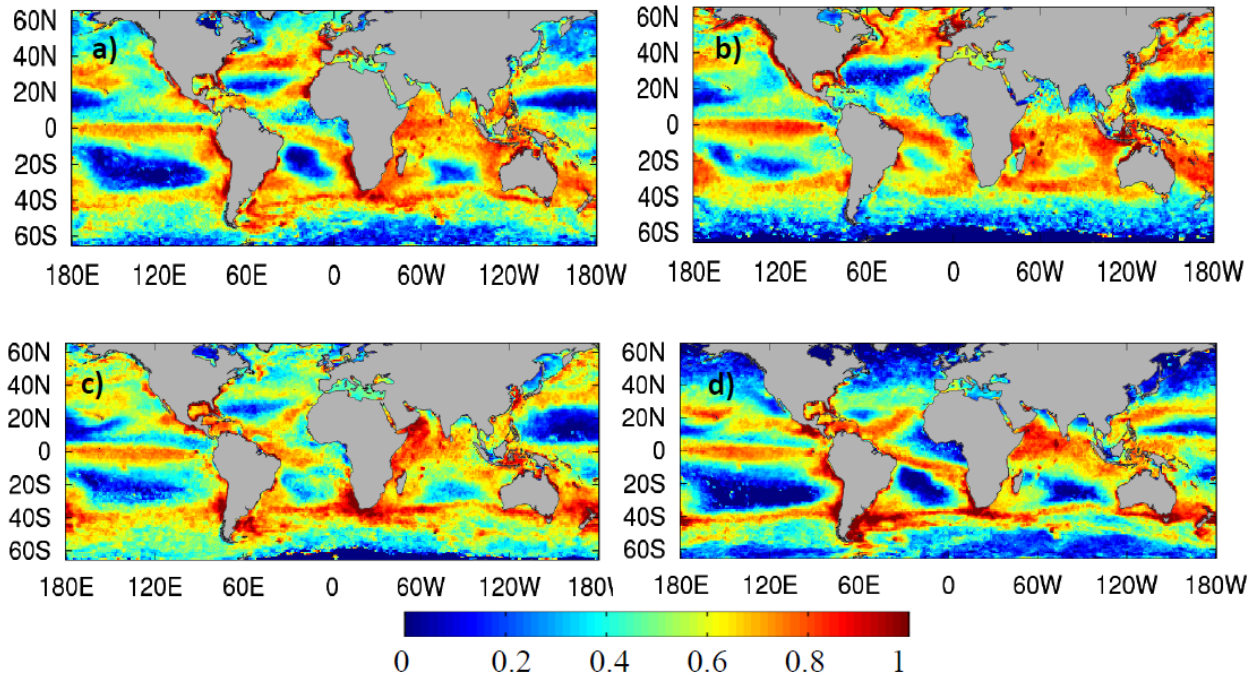


Figure 32. Index τ for the global ocean: average of daily maps of τ computed over 7-day composites of PHYSAT data for the period 2003–2010. a) Boreal Spring; b) Summer; c) Fall; d) Winter.

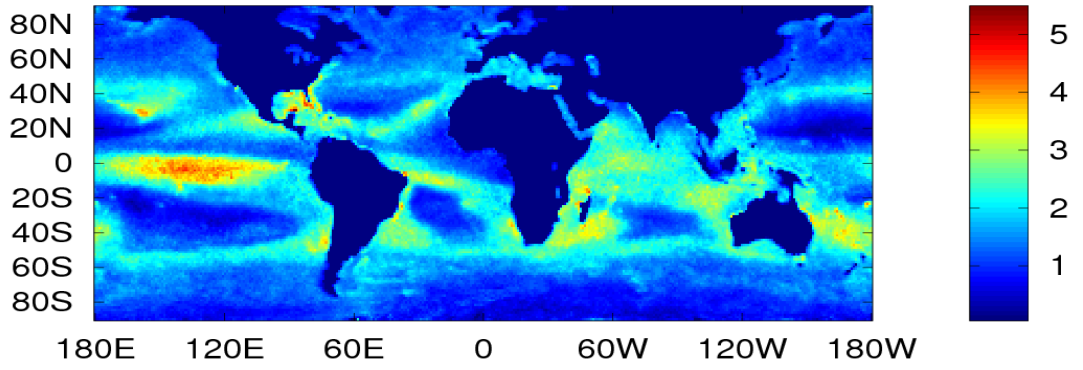


Figure 33. Map of signal-to-noise ratio indicating the stability (strength of the signal) of the τ diversity index through time (2003-2010).

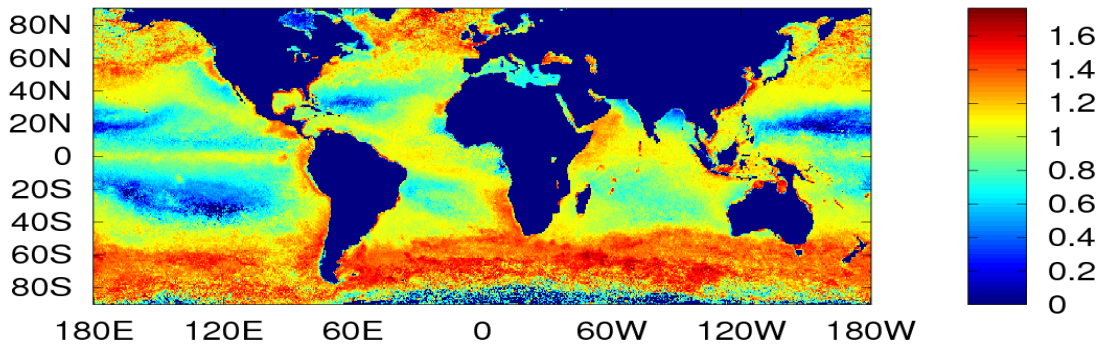


Figure 34. Average diversity (Shannon index) of temporal successions of dominant groups over the period 2003-2010.

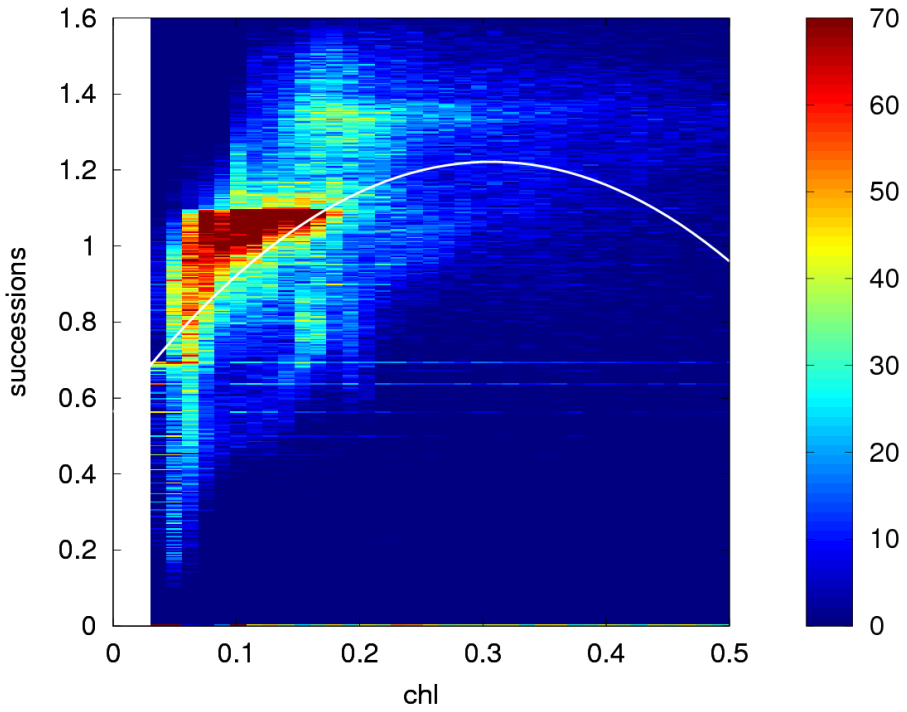


Figure 35. Bivariate histogram showing values frequency (color scale) and relationship between average climatology (2003-2010) of chlorophyll concentration (µg/l) at low and intermediate/high regimes of productivity and average diversity of remote sensed temporal successions (regression R2 = 0.25; $p < 10^{-9}$).

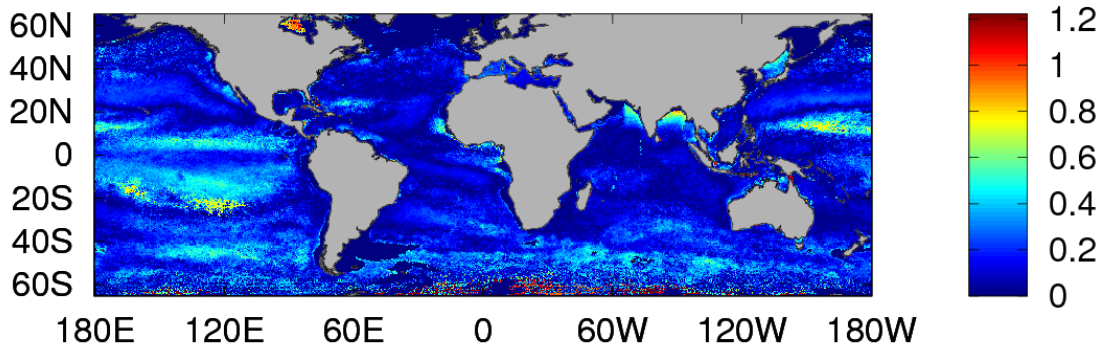


Figure 36. Global map of residuals after regression analysis between average diversity of remote sensed temporal successions and average climatology (2003-2010) of chlorophyll concentration for low and intermediate/high productivity regimes.

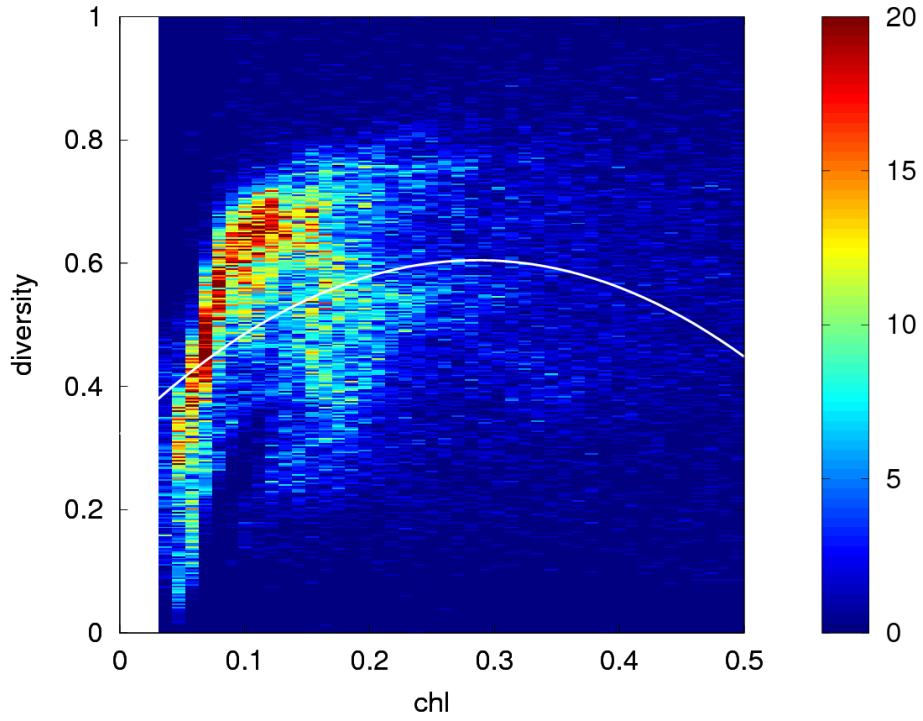


Figure 37. Bivariate histogram showing values frequency (color scale) and relationship between average climatology (2003-2010) of chlorophyll concentration (µg/l) at low and intermediate/high regimes of productivity and average climatology of remote sensed τ diversity. (regression R2 = 0.14; $p < 10^{-9}$).

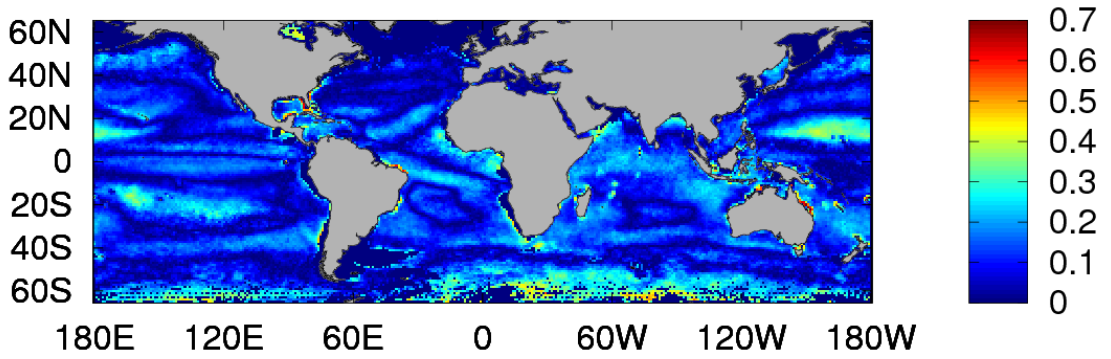


Figure 38. Global map of residuals after regression analysis between average climatology of remote sensed τ index and average climatology (2003-2010) of chlorophyll concentration for low and intermediate/high productivity regimes.

4.4 Discussion and conclusion

A new index of biodiversity based on remote sensed information about the spatial distribution of dominant phytoplankton groups has been proposed. This measure exploits a specific feature of oceanic ecosystems at the submesoscale, where horizontal transport and mixing occur on the same time scale of planktonic blooms. Indeed, the concept of niches that are fixed in time, typical of terrestrial landscapes, does not apply to the open ocean environment.

This stirring-mediated diversity, which this approach tries to capture, is only one component of oceanic biodiversity at a given location, since other processes may occur, that do not appear to the level of the dominant type or cannot be captured by this area-based index. One example of physical mechanisms other than horizontal stirring which is likely to provide a strong structuring effect to marine communities is the vertical dynamics occurring in upwelling regions (Abbott & Zion 1985; García-Reyes *et al.* 2014).

Upwelling regions, along with temperate and high latitude zones appear as transient biodiversity hotspots through time. Here the phytoplankton community structure is dominated by a strong ecological succession of different dominant types (Lindenschmidt & Chorus 1998; Cury *et al.* 2000; Rossi *et al.* 2008). In the open ocean these regions correspond to regions of intermediate/high productivity where diversity of temporal ecological successions is related to chlorophyll concentration by a unimodal function already seen in numerous studies about

productivity-diversity relationship (Dodson *et al.* 2000; Irigoien *et al.* 2004; Gillman & Wright 2006; Vallina *et al.* 2014). Also the τ index shows, to a less degree and mainly for low and intermediate level of productivity, a global relationship with chlorophyll. Both cases are in agreement with i) the functional diversity hypothesis, that states that productivity of a system can be maximized and maintained by the complementarity and redundancy of ecological roles (Tilman *et al.* 1997), and ii) the potential energy or 'productivity-richness' hypothesis, that predicts a positive effect of primary productivity on richness at coarse grain sizes, such as facilitating larger population sizes that avert extinctions or support niche specialists (Currie *et al.* 2004; Tittensor *et al.* 2010). The non-relationship between diversity and productivity for high/extreme values of chlorophyll may be due to the fact that they are mostly typical of coastal waters and well localized strong blooms, where remote sensed derived information is expected to underperform and productivity to be influenced by additional factors (Longhurst 2010).

Global remote sensed diversity patterns are also in agreement with latitudinal and temperature patterns already showed in macroecological studies, both in models and in in situ observations (Hillebrandt 2004; Fuhrman *et al.* 2008; Tittensor *et al.* 2010; Raes *et al.* 2011).

Remote sensing provides a synoptic global-scale information with relatively high spatiotemporal resolution. Due to this advantage, its derived products can be useful to define other remotely sensed proxies of biodiversity. Together with the τ index I proposed here, such proxies would allow to describe the ecological context of in situ observations and to optimize sampling strategies. Moreover, the availability of long term ocean color data offers the opportunity of studying trends in biodiversity patterns from a new, global perspective. This work shows that remote sensing products may contain unexploited potential for biodiversity studies. However, only interdisciplinary efforts gathering together remote sensing specialists and marine biologists will allow to assess the power of satellite-based indexes in resolving biodiversity patterns.

Chapter 5 Can we derive information on higher levels of the trophic chain?

Describing biodiversity patterns starting from primary producers have proven to be a powerful tool for understanding what determines the variety of life on Earth as well as for setting priorities for conservation efforts. Despite the importance of acquiring this information, the ecological characterization of important environments such as the open ocean is challenging, because of its extent and remoteness, its fluid dynamical environment and the mostly planktonic form of its primary producers. Identification of pelagic biodiversity hotspots still relies on sparse in-situ observations, or on global models. Here I use a remote sensed approach to show that most global pelagic biodiversity hotspots can be identified by considering the spatiotemporal features of oceanic niches sustaining different types of plankton. I attempt to assess the correspondence between the remote sensed diversity of primary producers estimated by τ index with the biodiversity patterns of upper trophic levels. These global patterns of biodiversity are shown to be highly congruent with consumers' hotspots of species richness, suggesting that remote sensing, when opportunely treated, has the potential of providing reliable candidates of areas of special biological interest for the entire pelagic marine ecosystem.

5.1 Introduction

Biodiversity plays a key role for the functioning and the resilience of many ecosystems, and provides several services of societal relevance. However, biodiversity is not evenly distributed and understanding the global patterns of its variability is a central issue in ecology and an important step for setting effective conservation priorities (Hutchinson & Macartur 1959; Gaston 2000; Myers *et al.* 2000). It is generally recognized that top predators are receiving inadequate protection. This is more evident for species living in the open oceans (Game *et al.* 2009) where the effectiveness of the few existing conservation plans is hindered by the weakness of the international agreements. Indeed, large marine vertebrates are highly mobile, ranging great distances across different countries over the course of their lives.

Migrations occur across entire ocean basins and habitat utilization varies according to regional oceanography and taxon-specific life-history characteristics. Regardless of the species under study, one of the primary drivers for their aggregation and movements is foraging. The formation and propagation of pelagic foraging habitats depends on complex oceanographic dynamics (Hazen *et al.* 2013), so habitat in the marine environment does not always refer to fixed geographical space, but to preferentially used areas that may shift and are difficult to identify (Scales et al. 2014). Finding a tool to identify these dynamical areas of aggregation that influence marine vertebrate distributions and diversity is therefore crucial for the effective conservation of the 'High Seas' and top predators (Hooker *et al.* 2011).

As highlighted in the Introduction, meso- (10 - 100 kms) and sub-mesoscale (<10 kms) oceanographic dynamics are responsible of the development of ecologically significant features such as fronts and eddies (Godø *et al.* 2012). Within these turbulent features, convergence zones can enhance, by various mechanisms, plankton biomass, which is advected from surrounding water masses, driving bottom-up processes across multiple trophic levels including apex predators (Graham *et al.* 2001; Bakun 2006). The mechanisms linking physical processes, prey dynamics and top predator foraging are complex and scale dependent (Fauchald 2009).

Being the base of the trophic chain, primary producers have received special attention in biodiversity studies (Jetz *et al.* 2009). It has been shown that their diversity strongly correlates with the diversity of consumers (Jenkins *et al.* 2013), and in some cases it has been suggested

that biogeographical diversity patterns estimated for primary producers may be used as a surrogate of patterns for higher levels of the trophic chain (Duffy *et al.* 2007).

Global biodiversity patterns for primary producers have been extensively studied for terrestrial or coastal systems (Kerswell 2006). Although the amount of primary production in the ocean and on land is comparable, off-shore marine biodiversity patterns are comparatively much less known. Phytoplankton provide a similar contribution to the planetary primary production as terrestrial plants do, playing a key role in the climate system, and sustain critical services for humankind. Although the amount of primary production in the ocean is roughly comparable to that on land, there are two main differences that have challenged the application of terrestrial methods, such as Eulerian sampling and belt transects, to the open ocean: i) primary production is largely provided by microscopic organisms, whose definition of diversity in terms of number of taxonomic units is still under debate (Komárek 2006), and ii) the ocean landscape is not static on ecological timescales, but has a strong transport processes and mixing dynamics which overlaps with the demographic time scales of phytoplankton.

In the past years, circulation models and observational studies along with remote sensed information have circumvented some aspects of the two problems, although these solutions have not been integrated together yet. In the past decade circulation models have yielded accurate and quantitative physical simulations of the oceanic global dynamics up to the mesoscale. More recently the accuracy of these models have increased with the introduction of complex ecological components for phytoplankton (Follows *et al.* 2007; Stock *et al.* 2011).

Observations have been providing new opportunities to the identification of biodiversity patterns in the ocean. Collections of observational records together with advancements in habitat modelling and satellite tracking of large vertebrates has led to global maps of cross-taxa species occurrence and richness at least at the 1000-10000 km scale (i.e., averaging out the mesoscale variability) (Selig *et al.* 2014). Complementary to in situ observations, observations by remote sensing contain very approximate and indirect biological information but have higher resolution (km/day) and global coverage.

It is for this reason that satellite data, and especially ocean color, have been recently explored as a possible source of primary producers diversity information (De Monte *et al.* 2013). It has been

shown that a Shannon biodiversity index can be formally defined on remote sensed ocean color by using an area-based algorithm which estimates the patchiness of putative different communities (identified for their different bio-optical anomaly) (De Monte *et al.* 2013). Mesoscale regions containing a mosaic of contrasted communities are marked by this algorithm as biodiversity hotspot candidates, assuming that they become mixed by the mesoscale turbulence. This index provides global maps which satisfy basic macroecological relations and responded positively to a global validation based on a physical and ecological circulation model (see Chapter 3 and 4). These maps can therefore be compared to information about the distribution of top predator species and biodiversity of consumers at global level.

5.2 Objectives

In this study, I used remote sensing, ecological models and observations to investigate the extent to which a biodiversity index defined on remote sensed information about the planktonic community (the τ diversity index) reliably assesses biodiversity patterns at multiple levels of the trophic chain. Both the physical and ecological dynamics underpinning global diversity patterns of primary producers are indirectly embedded in the index (see Chapter 3 and 4), allowing a kind of end-to-end approach.

More specifically, to achieve such objective, I compared global biodiversity patterns of primary producers and consumers, taking into account that there is a transition from direct effects of oceanography (depending on the characteristics of the life stages of the species) on the distribution of organisms, to the indirect effects via predator and prey distribution and behavior. Hotspots of biodiversity of primary producers correspond to regions of aggregations of top predator species and in general to regions of higher cross-taxa biodiversity. These results suggest that the remote detection of marine biodiversity hotspots can be a tool in monitoring the health of marine ecosystems at the global scale.

5.3 Bottom up effect of plankton seascape diversity on higher levels of the trophic chain.

5.3.1 Remote sensing phytoplankton biodiversity and top predators aggregation

A remote sensing biodiversity index τ has been defined as the Shannon index of the distribution of the relative spatial extension, within a disc of given radius, of areas dominated by a different

dominant phytoplankton community (see Chapter 2.2.3). Algorithms processing ocean color attribute to every pixel a tag that reflects some properties of the local ecosystem (different dominant planktonic types, e.g. diatoms and coccolitophores, are distinguished in the classification), but it is not possible to resolve finer features of the community. This approach provides global maps of phytoplankton biodiversity that have been regionally validated.

In contrast to the observations of phytoplankton biodiversity, which provide a sparse coverage for many remote regions of the global oceans, a relatively larger amount of data has been collected in recent years by tracking the movement of marine predators. Regions with increased probability of localization of different species of marine predators have been identified (Block *et al.* 2011; Kaschner *et al.* 2011), and are considered as biodiversity hotspots of high trophic level species.

Diversity hotspots estimated by the τ index appear to match several independent observations relative to marine top predators. The τ index is particularly high in the North Pacific Transition Zone (NPTZ) (Fig. 39). The NPTZ is a highly dynamic region that delineates the boundary between warm, oligotrophic subtropical gyres and cold, productive subarctic gyres, characterized by contrasted phytoplankton communities, and is a marine biodiversity hotspot of global significance (Sydeman *et al.* 2006). Numerous marine vertebrates with contrasting life histories preferentially use areas of the NPTZ, including northern elephant seals (*Mirounga angustirostris)*, salmon shark (*Lamna ditropis*) and blue shark (*Prionace glauca)*, Bluefin (*Thunnus thynnus*) and albacore tuna (*Thunnus alalunga*), Laysan *(Phoebastria immutabilis)* and black-footed albatrosses (*Phoebastria nigripes)*, and loggerhead turtles (*Caretta caretta*) (Scales *et al.* 2014). The NPTZ is identified by Block et al., 2011 as a region where a high number of predator species concentrate, and attains high τ values more to the West, in proximity with the California Current large marine ecosystem. This ecosystem is one of the major Eastern boundary upwellings (e.g. Canary Current, Benguela Current, Humboldt Current) and is characterized by intense surface frontal activity between cool, nutrient-rich upwelled water and warmer oligotrophic waters further offshore and is considered also a hotspot of marine biodiversity (Chavez & Messié 2009). Bioaggregation in upwelling-driven frontal structures attracts foragers from diverse foraging guilds (Nur *et al.* 2011; Sabarros *et al.* 2013). On the contrary the area between the California Coast and the Hawaii is a region with minimum predator density (Block *et al.* 2011), and accordingly attains low τ values (Fig. 39).

Worm et al., (2003) (Worm *et al.* 2003) estimated the biodiversity of top predator species such as tunas and billfishes by fishing catches in few regions of the globe. Although data are not uniformly distributed, a large scale pattern correspondence between high τ values and the diversity rich zones of Hawaii and in the Caribbean Sea and low τ values with less biodiversity rich areas north of Hawaii and in the offshore Atlantic Ocean is evident (Fig. 40). In Central America, a strong latitudinal change in biodiversity (τ) is also evident in proximity of the Galapagos. It is in correspondence of the Pacific Equatorial Front, manifesting between the equatorial upwelling to the South and warmer tropical waters to the North. It is characterized by steep gradients in temperature, salinity and nutrients (Ballance *et al.* 2006). It has been shown that planktivorous seabirds strongly associate with this feature, which entrains zooplankton in surface layers (Spear *et al.* 2001).

On a more global scale, the τ index map shows similarities with the maps of marine mammal occurrence and species richness derived by predictions of global distributional ranges and empirical observations (Fig. 4 of Kaschner et al., 2011). Similarities are also evident considering coarse global patterns of species richness encompassing several phyla, from zooplankton to marine mammals (see Tittensor et al., 2010, Fig. 2d, reproduced in Fig 41). Global hotspots of biodiversity are described around the Philippines, Japan, China, Indonesia, Australia, India and Sri Lanka, South Africa, and the Caribbean and southeast USA. Cross-taxon oceanic diversity shows consistent bands of high average diversity at 30° latitude in both hemispheres. Within these latitudes, oceanic diversity peaks closer to the continents and along boundary currents such as the Gulf Stream and Kuroshio Current, probably owing to the availability of favorable habitat for foraging. The same is found in biodiversity patterns described by τ index and explanations were given in Chapter 3 and 4.

The matching of areas considered hotspots for predator diversity with transport-induced hotspots is a feature that supports the bottom-up propagation of biodiversity across the trophic chain, and thus further supporting the possible use of the τ index for detecting biodiversity hotspot as well as habitat shift and degradation for higher tropic levels.

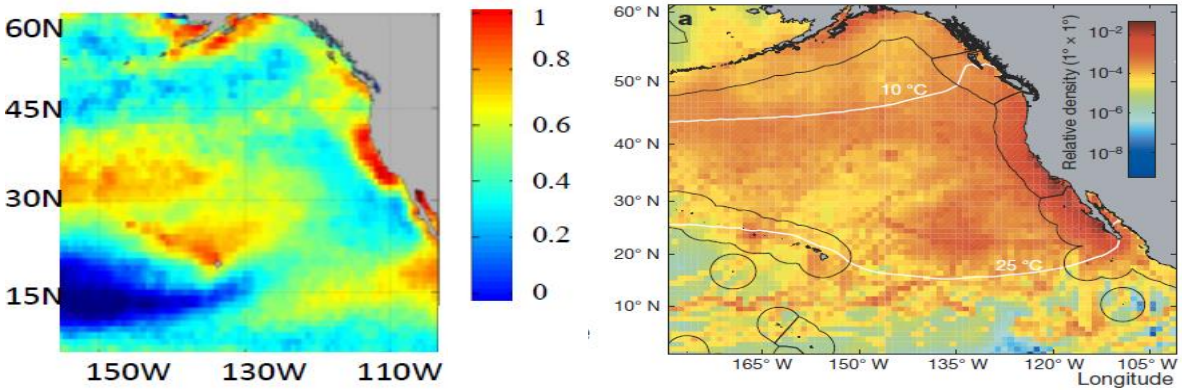


Figure 39. τ index map (on the left), centered in the North Pacific, where the North Pacific Transition Zone, the region East of Hawaii and the California current stand out as hotspots in agreement with biodiversity hotspots found by Block et al., 2011 in the Pacific (on the right).

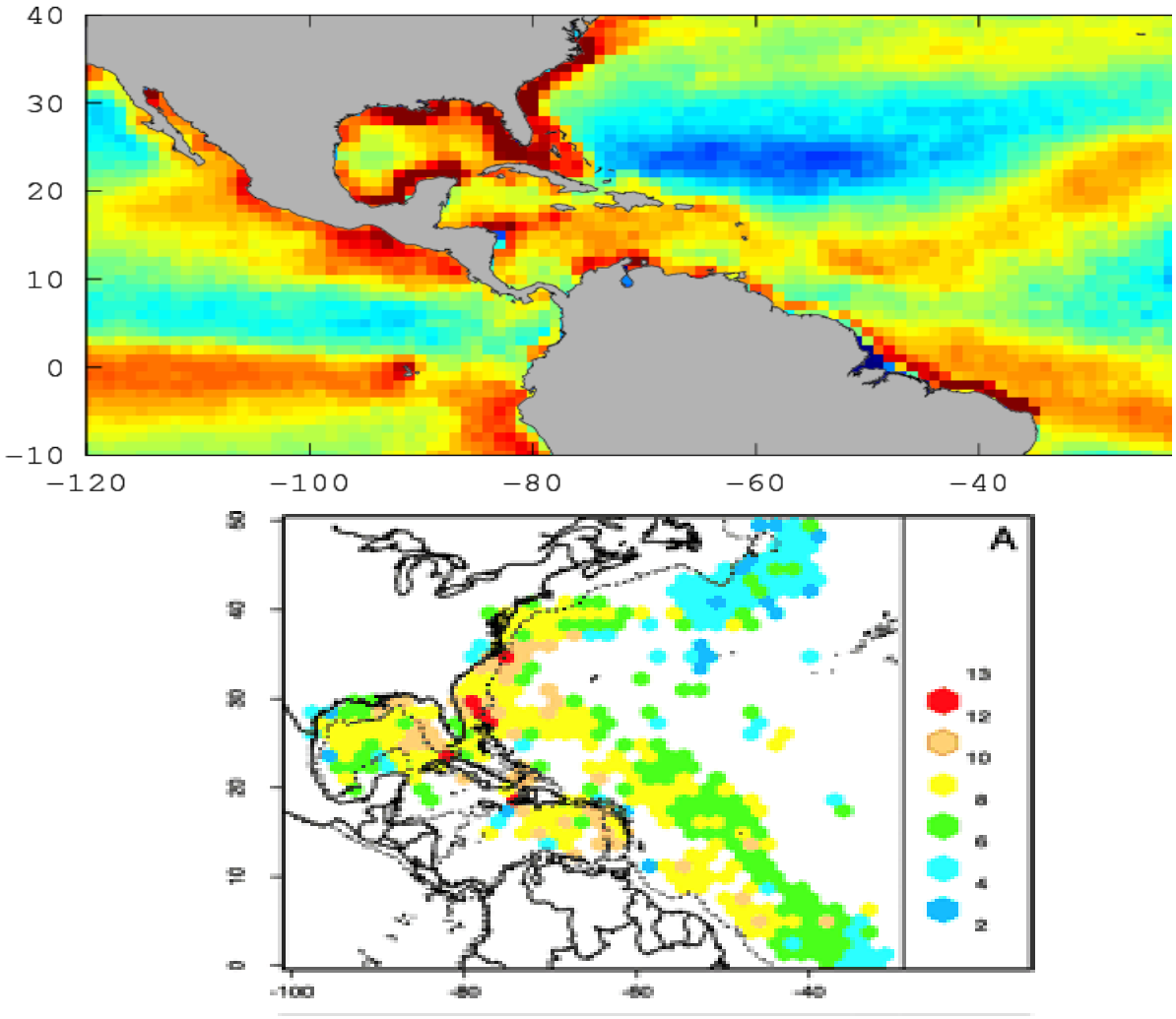


Figure 40.Top: τ index map, centered in Central America, where hotspots of biodiversity stand out in agreement with top predators biodiversity hotspots found by Worm et al., 2003 (A, species richness).

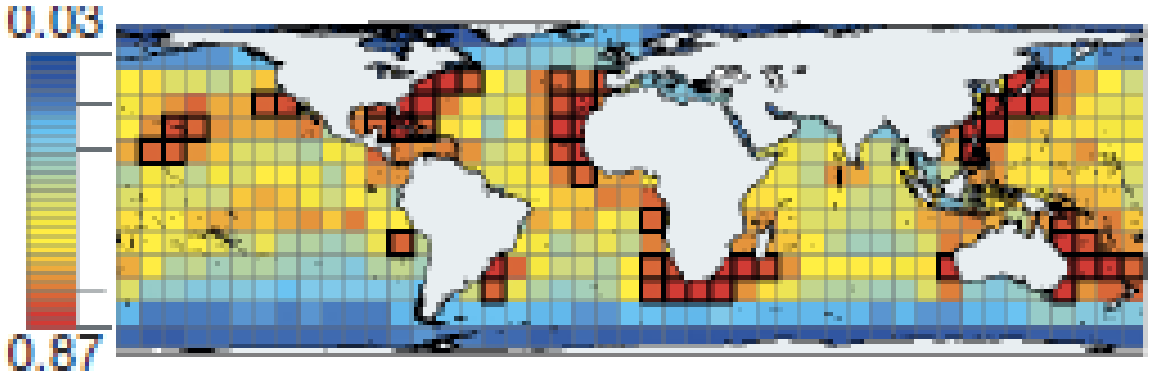


Figure 41. Global marine species richness for oceanic taxa. Cells with a bold outline are hotspots (defined as the 10% of cells with highest mean richness). Horizontal tick marks on color-bars indicate quartiles. From Tittensor *et al*. 2010.

5.3.2 Covariance and congruence of primary production and consumers' diversity

This possibility of mapping phytoplankton biodiversity from satellite-detected patchiness of their bio-optical signature allows to test whether the relation between global diversity patterns of primary producers and consumers, already observed on land, is present in the ocean. For consumers, I used AquaMaps species richness (see Chapter 2.3.2). In contrast with maps based on observations only, AquaMaps integrates records of species occurrence and knowledge of habitat modelling together, providing smoother patterns with a higher resolution that matches better the satellite-derived phytoplankton diversity maps. Covariance of the two global maps is presented in Fig. 42. A linear regression model indicates a strong correlation (Pearson's r = 0.61; linear regression fit R2 = 0.38; $p < 10^{-9}$, intercept -0.04; b1 = 1.697). In order to test whether this is a confounding effect due to a common environmental determinant, I studied the residuals of the regression analysis (Fig. 43) versus SST, gradient of SST, Chl, and TKE and also computed regressions of the AquaMaps diversity versus the same environmental variables (Fig 44). None of the environmental variables scored better than the satellite-derived diversity, indicating more a direct relationship than a common environmental determinant (TKE R2 = 0.06; SST R2 = 0.29; SSTgrad R2 = 0; Chl R2 = 0.004; $p < 10^{-9}$). The goodness of fit for high-diversity regions is confirmed by the substantial overlap of the biodiversity hotspots defined according to the different indexes. In each map, I defined biodiversity hotspots as the top-valued pixels covering 20% of the total ocean surface. This percentage is set in agreement with the target chosen by the Convention on Biological Diversity, which aims at protecting at least 10% of the global ocean

extension by 2020 (Aichi Target 11). Indeed, the seascape biodiversity hotspots overlap substantially with hotspots of diversity at multiple levels of the ecosystem covering 12% of the globe for 20% coverage of the highest values for each index (Fig. 45).

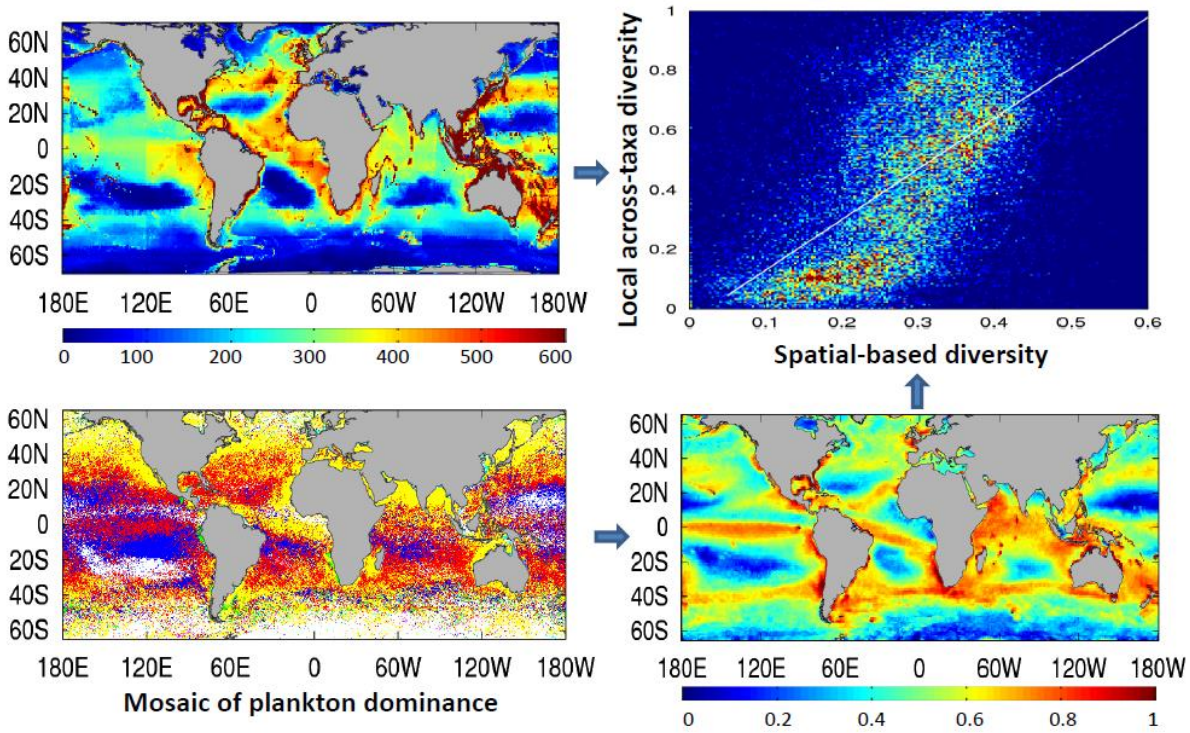


Figure 42. Bivariate histogram showing values frequency (color scale) and relationship between τ diversity information (estimated at 1 degree disk radius) (x axis) with its relative map and cross-taxa global biodiversity estimated by species richness based on AquaMaps (y axis) with its relative map. The distribution of the dominant phytoplankton groups resulted from PHYSAT reanalysis is also showed.

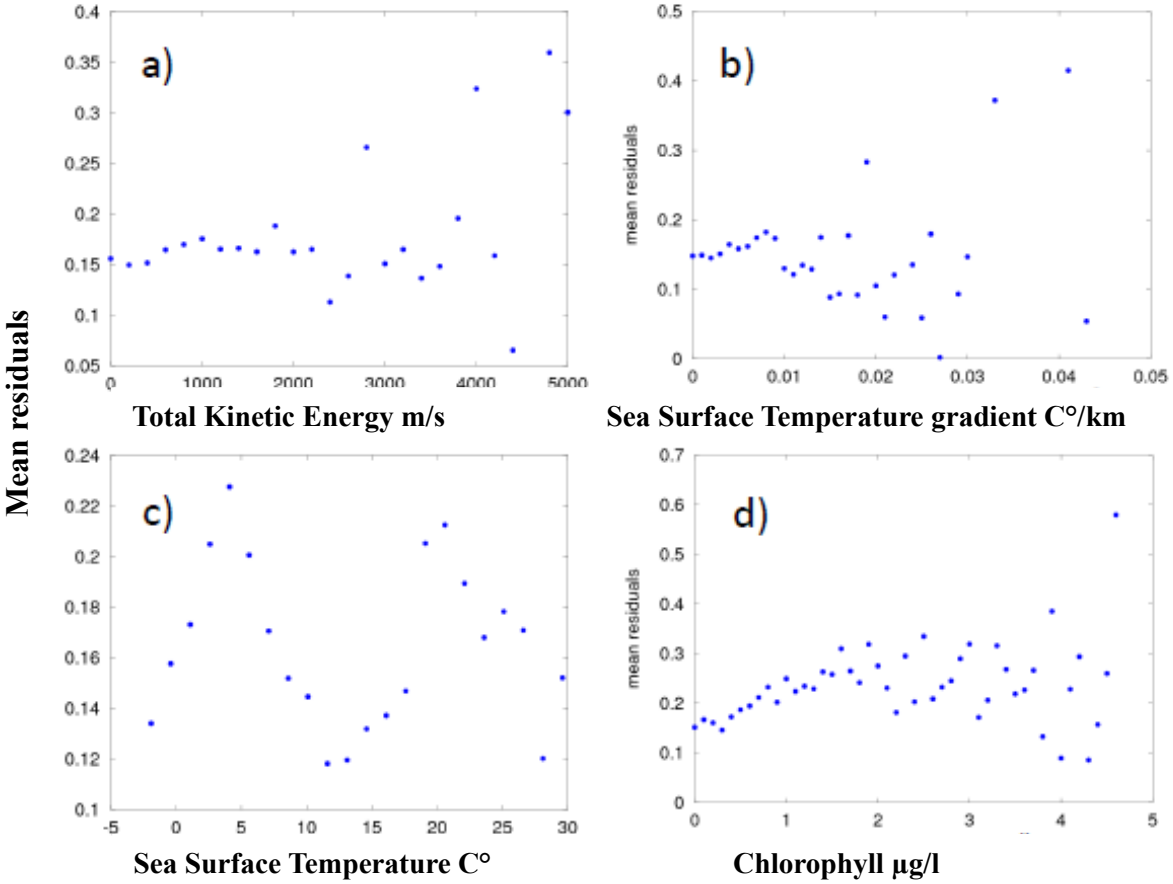


Figure 43. Study of the residuals of the regression between τ diversity information (estimated at 1 degree disk radius) and global cross-taxa biodiversity estimated by species richness based on Aquamaps. The residual values have been averaged inside classes depending on the environmental characteristics present at each location. A) annual climatology of Total Kinetic Energy (m/s); B) annual climatology of Sea Surface Temperature gradient (C°/km); C) annual climatology of Sea Surface Temperature (C°); D) annual climatology of Chlorophyll (µg/l).

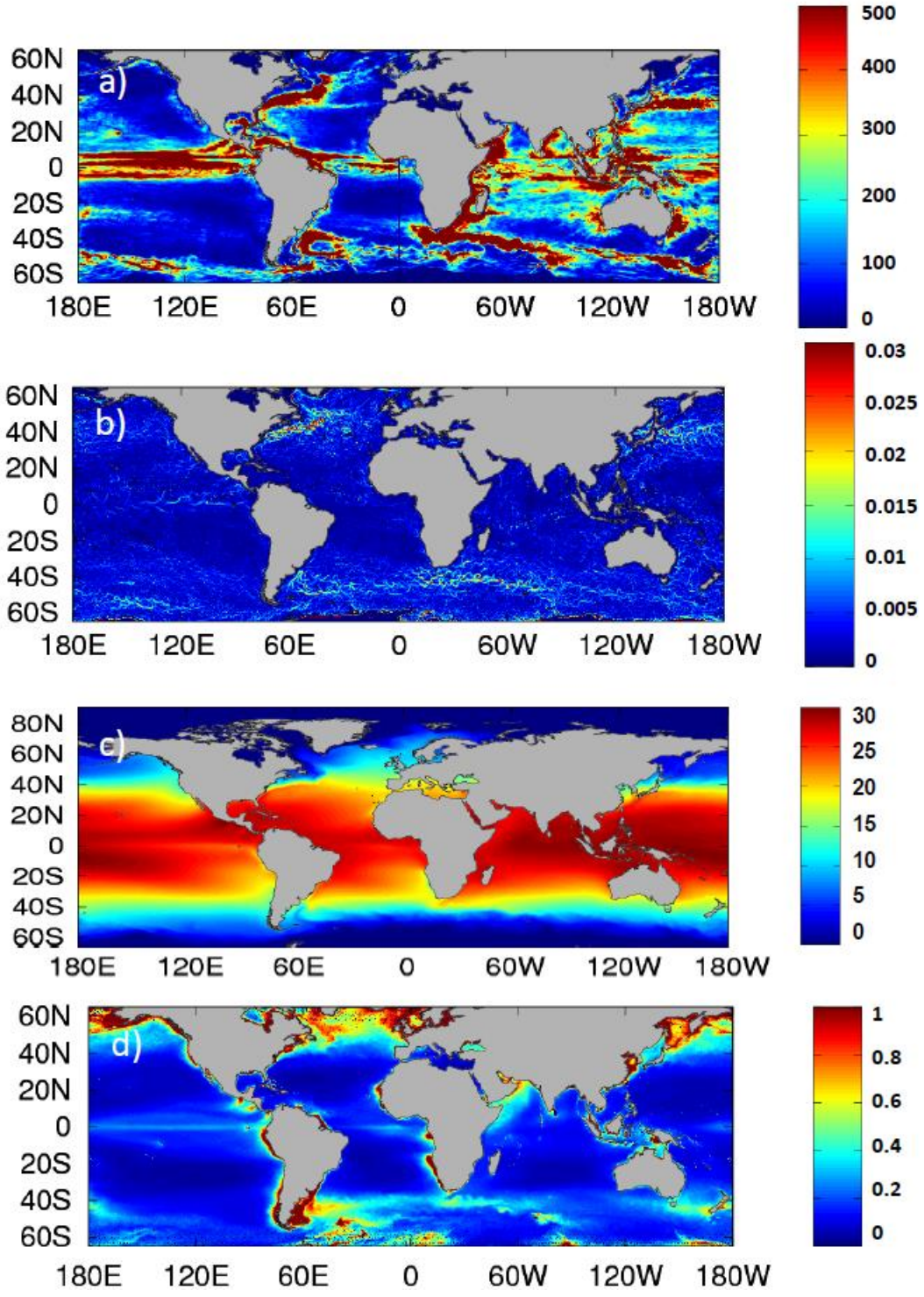


Figure 44. Remote sensed annual climatologies (2003-2010) of a) Total Kinetic Energy (m/s), b) Sea Surface Temperature gradients (C°/km), c) Sea Surface Temperature (C°), d) Chlorophyll concentration (μg/l).

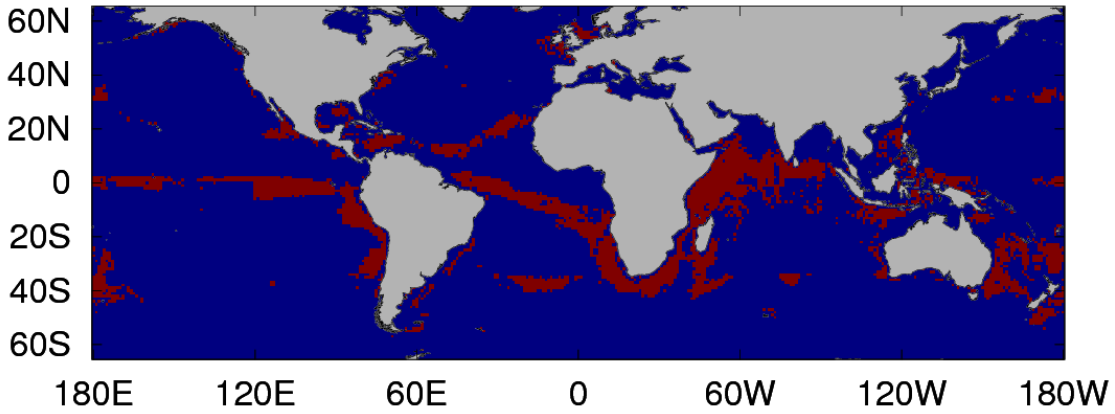


Figure 45. Biodiversity hotspots covering 10% of the global ocean as estimated by overlaying the highest values of τ remote sensed diversity of primary producers and biodiversity of consumers estimated by AquaMaps.

5.4 Discussion and conclusion

Remote sensing have been showing that inhomogeneities as strong as basin scale also exists in the ocean on much shorter spatial scales (~10-100km) and on temporal scale comparable to the demography of phytoplankton (days to weeks). These inhomogeneities form a mosaic patchiness, similar to that on land, but more dynamic. Recently, the mesoscale environmental mosaic patchiness has been shown to reflect into a mosaic of contrasted optical properties that can be detected by satellite sensors and that are considered signature of phytoplankton communities dominated by different types. Spatial analysis of these remote sensed, fine scale mosaics formally yields a Shannon index that has been proposed to represent a proxy of marine biodiversity at the scale of 100kms (De Monte *et al.* 2013). Here I studied this index in relation with top predators and global consumers' diversity, finding that the satellite-derived biodiversity recognizes most areas of aggregation for top predator species and has a good correspondence with consumer diversity in terms of both covariance and pattern congruence. However, not all the areas of high species richness overlap with diversity of primary producers, among them coastal areas of the Indonesian archipelago and waters surrounding Azores archipelago in the North Atlantic Ocean. These habitats are likely to be characterized by other ecological processes with strong influence on the species community structure. Moreover, the approach developed here seems to be more suitable for detecting areas of aggregation of large vertebrate pelagic species, because they base

their movements on the tracking of ephemeral and persistent oceanic hydrographical features that lead to the fast formation and dissipation of areas of high prey aggregation and diversity.

Despite few mismatches, overall these results suggest that spatially extended data even with poor community resolution is a promising tool to evaluate and monitor biodiversity hotspots at the global scale. They also suggest that the planktonic biodiversity enhancement by physical forcing extends beyond primary producers to other levels of the trophic chain, even if the exact mechanisms are not fully elucidated (Leibold *et al.* 1997; Bakun 2006). They may support aggregation and higher biomass and diversity through the trophic chain and relax density dependent interactions by offering a wider set of competitive opportunities (Huston 1979; Strong 1983; Griffin *et al.* 2008). Indeed, density and diversity of predators and preys affect the efficiency of trophic energy transfer. Under constant resource supply, the density of a species determines the strength of intraspecific and interspecific competition, which can also influence the effect of biodiversity on ecosystem functioning (Cardinale *et al.* 2006; Duffy *et al.* 2007). In case of environmental and resource heterogeneity, this strength can be weaken promoting coexistence through spatial and temporal niche partitioning and diversification of patterns of resource utilization.

This result is somehow surprising because previous studies showed, for instance, little association between phytoplankton and zooplankton diversity at macroecological scales (Irigoien *et al.* 2004). Possible causes of this discrepancy are the scale and the number of replicates of the sampling (regional in this case, i.e. local in Irigoien et al. 2004), the end to end approach that considers the ensemble of the trophic chain and not only zooplankton, and the two-dimensionality of the maps. For instance, when computing the correlation between our satellite-derived proxy and Aquamaps biodiversity georeferencing only in terms of the latitude, I still find a good association (Pearson's $r = 0.77$, $p < 10^{-9}$), showing that scale and number of replicates and an end-to-end approach are likely to be essential in global biodiversity studies.

The strong association between phytoplankton biodiversity and top predators distributions and AquaMaps biodiversity suggests that phytoplankton biodiversity can be used as a surrogate first guess for cross-trophic biodiversity hotspots and indirectly to monitor the health of marine ecosystems at the global scale. The identification of priority areas, such as 'hotspots', has been an essential tool for conservation planning. However, identifying spatially explicit areas of high

biodiversity in the open ocean has been limited by the lack of data and in particular the small geographic coverage. Considering that the phytoplankton biodiversity was retrieved by remote sensing and had global coverage, this approach is therefore particularly promising to improve global conservation strategy efforts.

Several studies can be envisaged in the future. Field studies assisted by real time processing of satellite data should help to better understand the strength and limitations of remote sensing of biological diversity, and suggest in which direction to develop next generation sensors for satellites. In particular, they may explore whether the reliability of the remote sensed proxy and the association between phytoplankton and consumer biodiversity hold at higher spatiotemporal resolutions. Moreover, specific trophic levels depending on the regional system under study can be isolated and the indirect influence of planktonic biodiversity investigated. This tool would be of invaluable importance to nominate areas of potential interest for conservation in the open ocean, as their establishment and enforcement is still highly debated (Halpern *et al.* 2008; Game *et al.* 2009; Grantham *et al.* 2011). The main issues are how to cense marine life in such a remote, vast and dynamical environment, how to recognize and define areas of aggregation of species characterized by different life stages and habits, how to enforce protection in zones beyond national jurisdiction. Because of these reasons, being able to describe and monitor the ecosystem at such temporal and spatial scales would represent a great step into the establishment of marine protected areas in the high seas.

Chapter 6 Perspectives-linking remote sensing to in-situ high-throughput information of plankton community.

6.1 Introduction

Comparing biodiversity in different places and times requires methods to quantify its extent and to track its changes. Many indicators of biodiversity are currently in use, most of which have been developed primarily for terrestrial macroscopic organisms (Magurran 2004). Whereas measuring the diversity of microscopic organisms especially those living in the marine environment such as phytoplankton is notoriously more problematic (Carstensen *et al.* 2005; Haegeman *et al.* 2013). Planktonic communities indeed include large numbers of species, mostly at very low abundance, whose quantification involves the application of field intensive sampling protocols. Sampling design must also take into account that phytoplankton exibit high species turnover as a response to changes in the environment, making their structure spatially and temporally variable (Dybern & Hansen 1989; Pannard *et al.* 2008). Furthermore the taxonomic classification of many species requires specialized and time-consuming methods (e.g. electron microscopy and high level of taxonomical expertise). Finally, the analytical methodologies available for analysing biodiversity pose several limitations depending on the used method with the result that, for instance, some organisms could be overrepresented while others never detected (Gotelli & Colwell 2001; Magurran 2004). Together, natural variability and human error may substantially affect the consistency of the quantification, thus invalidating expensive sampling efforts and analyses. The increasing availability of throughput imaging and sequencing technologies at acceptable costs (e.g. FlowCam and Illumina) (Logares *et al.* 2012) is now enabling the exploration of microbial diversity at an unprecedented scale. Indeed, they allow to analyse rapidly and cost effectively "en masse" large number of samples and to have a detailed view of the community composition. However, issues also exist related to these techniques. For instance, imaging better detects organisms with evident and distinctive morphological features and in a certain size range, whereas molecular tags can give different classifications (depending on the considered loci of the genome), are subjected to sequencing errors, require reference sequences for correct assignation of the taxon and usually define species boundaries based on a priori percent of sequence similarity with available databases (Moritz & Cicero 2004; Piganeau *et al.* 2011).

The objective of this chapter is to open perspectives onto the use of new generation high throughput techniques for studying how data type and analysis methods may bias biodiversity estimation. Indeed, in order to compare in-situ with satellite-based diversity estimations, we need to bridge the gap between two different techniques; the former which yields very detailed, but geographically and temporally sparse data and the latter which provides data at high spatiotemporal resolution, but that is associated to an extremely shallow view of the community.

In this Chapter I will investigate a) variation in morphological diversity in presence of physical structures like hydrographic fronts, b) the difference in biodiversity information obtained using a morphological or a genetic approach, and c) the shape of the community structure using Rank Abundance Distributions. To do so I analysed different sources of data collected during the Tara ocean expedition (see Chapter 2.3.3). There is not one single way of looking at microbial communities in the ocean, but combining different approaches requires to first understand which features of the community different data allow to tackle (Magurran 2004).

To analyse the **morphological diversity structured by hydrographic context** (fronts), I used data from a 100 km transect in the Balearic Sea (see Chapter 2.3.3). This region is characterized by complex circulation and hydrography with a high temporal and spatial variability (Mooney 2002). Well-developed mesoscale filaments and eddies (Alborean gyres, Catalan fronts) have been described in the shelf/slope region of the continental side (Fig. 46). These are instabilities associated with a thermohaline front lying along the continental slope. A second thermohaline front is found along the islands slope. Exhaustive analysis of the spatial variability associated with these structures showed intensified density gradients associated with vertical motions up to 50 m/day, about two orders of magnitude higher than the classical vertical motions typical of coastal upwelling conditions (Send *et al.* 1999). Within the Balearic sea, all four trophic regimes which characterize the Mediterranean Sea have been identified (coastal, blooming, intermittent and non-blooming) (D'ortenzio & Ribera d'Alcalà 2009). Together with the strong seasonality ruling the basin, such a diversity of regimes creates optimal conditions for the alternation of phytoplankton populations dominated by different functional groups and species (Chesson 2000). The West Mediterranean Sea is therefore a good case study to investigate how morphological diversity may be influenced by hydrographic context even at small scales.

The second aspect that I investigated in this chapter is the **robustness of taxonomic**

classification performed through a morphological or a genetic approach. To conduct this analysis I used only diatom data collected from three stations in the Mediterranean Sea (the first three stations of the Tara Oceans expedition where the sequencing data were produced in a pilot study for protist metabarcoding). Although there have been some studies comparing marine community structures obtained using different molecular techniques (Piganeau *et al.* 2011), few studies compared diatoms communities identified morphologically and genetically (Amato *et al.* 2007). Conventionally, phytoplankton taxonomy has been based on morphological features, with the underlying assumption that these defining characteristics were distinctive to a single species and related to specific physiological function (Sournia 1994; Tett & Barton 1995). Recently molecular techniques have been successfully applied to classify bacterial groups, plankton communities and eukaryotic diversity. These molecular techniques have revealed new lineages (Not *et al.* 2007) and suggested that the bacterial and eukaryotic diversity may be orders of magnitude greater than previously expected (Vaulot & Marie 1999; Massana & Pedrós-Alió 2008). Morphological identification has a lower taxonomic resolution with respect to phylogenetic and molecular techniques, thus the direct comparison of observations based on these different techniques are intrinsically difficult. Here, I computed the most commonly applied biodiversity index to data issued by metabarcoding and by morphological classification.

The **structure of a community** is also an important component of biodiversity to investigate, especially for highly complex community structures such as that of plankton. Quantitative analysis allows to study the abundance distribution of species composing the community. The Rank Abundance Distributions (RADs) can be related to processes acting in the environment as well as ecological processes determined by species interaction (MacArthur 1957). The count of abundant species (common) is the first mark of characterization of a community and the one less subject to sampling issues, therefore the most studied in classical ecology. However, advances in technology have allowed to observe microbial RADs possessing very long tails, indicating that the rare biosphere is a very important part of the community. Indeed, rare species are likely to influence the distribution of the abundant classes in spatio-temporal successions (Purvis & Hector 2000; Lennon & Jones 2011). Rare species are often represented by organisms, in a quiescent state or with slow growing processes, that will burst when certain environmental variables will provide the optimum conditions to outcompete the dominant types. The bulk of the empirical

work has recognised two facts: RADs follow a “hollow curve” (on arithmetic scale), with many rare species and few common species in every system studied and within this broad constraint, there is great variation in the details (McGill *et al.* 2007). Several authors have also highlighted that many theoretical models can produce the same diversity distributions curve (McGill 2003; Mouquet & Loreau 2003), however to suggest hypothesis about the implicit model responsible of the shape of the RADs is not the purpose of this study. It is widely proposed that the abundance distribution of the species in a community, as numerous other biological phenomena (Solé *et al.* 1999; Solé *et al.* 2000; Allen *et al.* 2001) can be represented by a power-law function (McGill 2003; Pueyo 2006). Power-law distributions are characterized by very long heavy tails and are strictly related to logseries and lognormal distributions as they are all successive terms in a Taylor series expansion. I used this function to study the slope of the distribution of photosynthetic protists as an informative metric for the diversity of community structure to be eventually related to remote sensed diversity indexes.

6.2 Tara cases study caveats and pitfalls

6.2.1 Results: Morphological diversity structured by hydrographic context

I analyzed horizontal stirring effects on the distribution of microbial marine communities due to hydrodynamical processes along a ~100Km transect in the Western Mediterranean Sea. The phytoplankton community has been characterized by cytometric diversity using Seaflow (see Chapter 2.3.3.2). Satellite chlorophyll images do not show any variation or trend along the transect, therefore the region seems homogeneous from the viewpoint of production (Fig. 46a). On the other hand, the calculation of the longitudinal advection shows a remarkable heterogeneity of the origin of water masses (Fig. 46b) and the calculation of the Lyapunov exponents shows the presence of eddies in the region and two strong fronts on the trajectory of the transect (Fig. 46c). Changes in biodiversity appear in correspondence with these ocean fronts. In particular, four different zones are identified: 1) a region of low biodiversity in the inner side of the eddy, 2) a region of higher biodiversity in the outer side, 3) a transition zone from low to high diversity close to the strongest front and 4) a sustained high biodiversity close to the coast

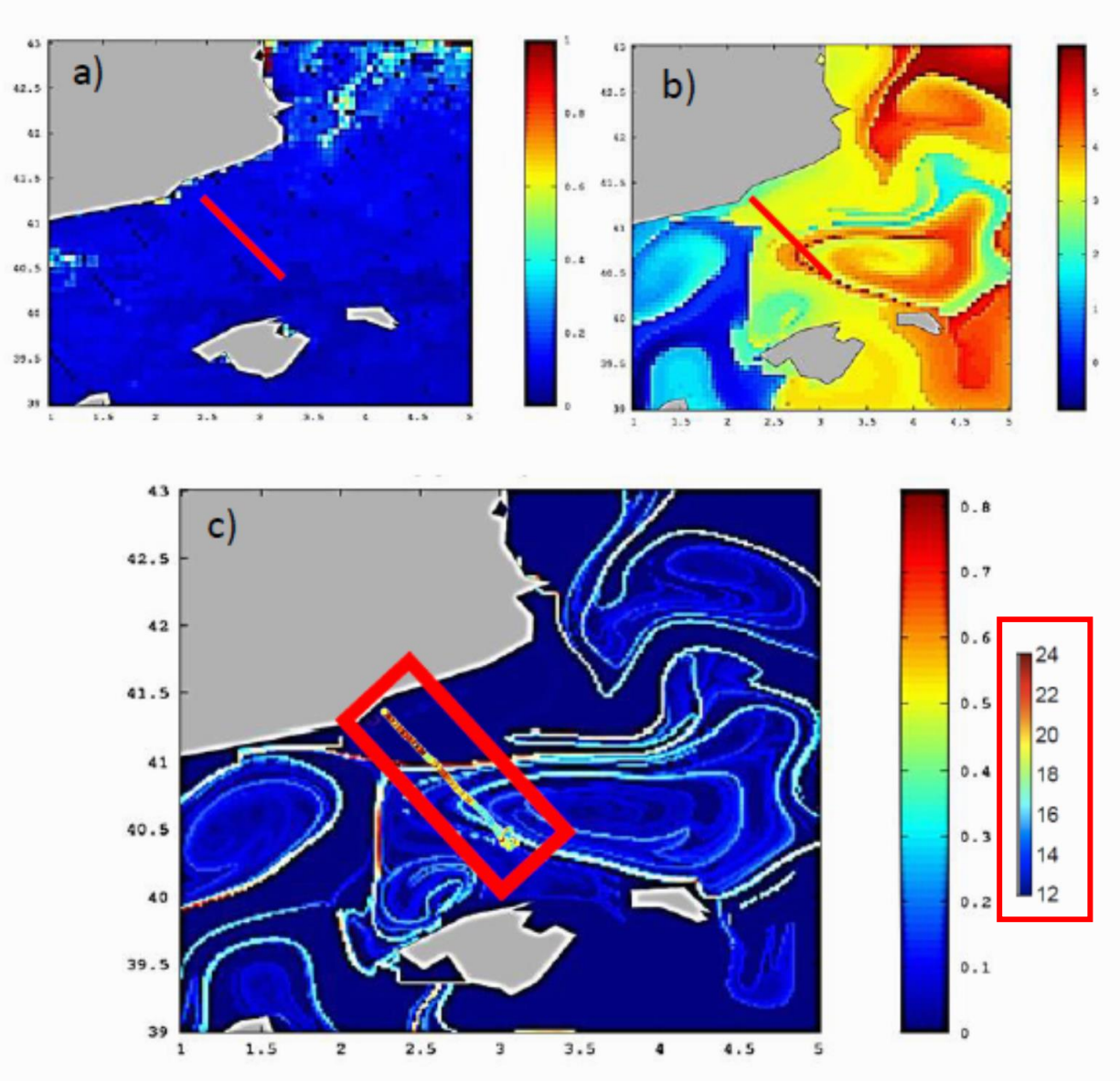


Figure 46. a) Chrorophyll concentration (µg/l); b) 15 days backwards longitudinal advection of water masses showing their origin; c) Lyapunov exponent showing transport barriers and overlayed differences in planktonic biodiversity along the transect, estimated by Seaflow (exp Shannon index).

6.2.2 Results: Morphological and genetic diversities: crossvalidation

I compared samples from 3 different stations in terms of methods used to estimate biodiversity

robustness of taxonomic classification performed through a morphological or a genetic approach

I compared biodiversity estimates obtained from in situ observations of samples from three

different stations of the Tara Oceans Expedition. These three stations in the Mediterranean Sea were the first to be released and constituted a pilot study for the development of different observation techniques. In particular, I used preliminary counts of diatoms obtained via morphological classification and via metabarcoding (see Chapter 2.3.3.3 and 2.3.3.4) and looked at consistency in the estimation of biodiversity using classification at genus level from the two methods.

Station 7 stands out as the less biodiverse among the three Mediterranean stations, consistently through the different samplings and statistical methods. Stations 23 and 30 usually display very similar statistics. rDNA tags identify 46 diatom genera for station 23, 38 genera for station 30 and 26 genera for station 7. K-dominance curves calculated on genetic tags show more similarity between station 23 and 30 compared to station 7, but with 23 being the most diverse (Fig. 47). Morphological analysis showed lower resolution (Fig. 48, 49, 50). Less than 10 diatom genera were detected by the FlowCam for station 23 and 30 (7 not available). Diatom surface diversity from K-dominance curves calculated on morphological data is very similar between stations 23 and 30, only the first dominant species has a different relative abundance (Fig. 48).

Biodiversity indices, ecological distance (β diversity metrics) and the decay exponent of the RADs are calculated for the surface samples (Magurran 2004; see paragraph 1.2.5). All the distributions fit a geometric model ($p < 0.001$), presenting almost only one dominant genus characterizing the community, even if the dominant genus changes using different methods for station 30 (*Chaetoceros vs Hemiaulus*) and, within the rDNA tags method, changes for station 7 from *Chaetoceros* to *Dytilum*. This condition is in agreement with abundance distribution model theories, where geometric and logseries shapes (Motomura 1932; Fisher et al. 1943) are typical of species-poor communities characterized by minimal cooperativity where the community is structured by one or few factors and there are only one or few dominant species. This is in accordance also with a comprehensive study of RADs of Mediterranean plankton where the entire community is generally well represented by power law functions but only diatoms are best represented by logseries (Pueyo 2006). All the considered indices rank the stations consistently in terms of diversity, with some variation concerning station 23 and 30 (Table 4 and 5). Station 7 results to be the less diverse for all the indices and its community structure is less similar to station 23 and 30, probably due to the influence of Atlantic water masses that come inside the

Mediterranean Sea through the Gibraltar Strait. Considering the rDNA Tags, station 23 appear to be the most diverse, but changing the method to FlowCam identification, station 30 results to be the most diverse. Taxonomic distinctness instead describes station 30 as the most taxonomically diverse, thus genus identified are less closely related compared to genus present in the other stations.

In spite of the general agreement on the differences in biodiversity between stations, important inconsistencies appear when the relative abundance of diatom genera, identified by morphological and genetic methods, are compared within each station. Some genera identified by the FlowCam are not identified by the rDNA tags, even if the sequence for that genus was known (SILVA); the relative abundance of the dominant genus *Chaetoceros* is very different between the two methods for station 23 and moreover the dominant genus change completely in station 30, where *Hemiaulus* is enormously present in the FlowCam data and absent in the rDNA tags, even if the sequence is known (Fig. 49 and 50). These inconsistencies may reflect the preliminary nature of the assignation, that has proven particularly problematic for marine eukaryotes, due to the fact that a large proportion of species, and sometimes of genera, are still unknown. Some assignation problems of the Tara Dataset have been solved later, when the dataset has expanded to include more stations.

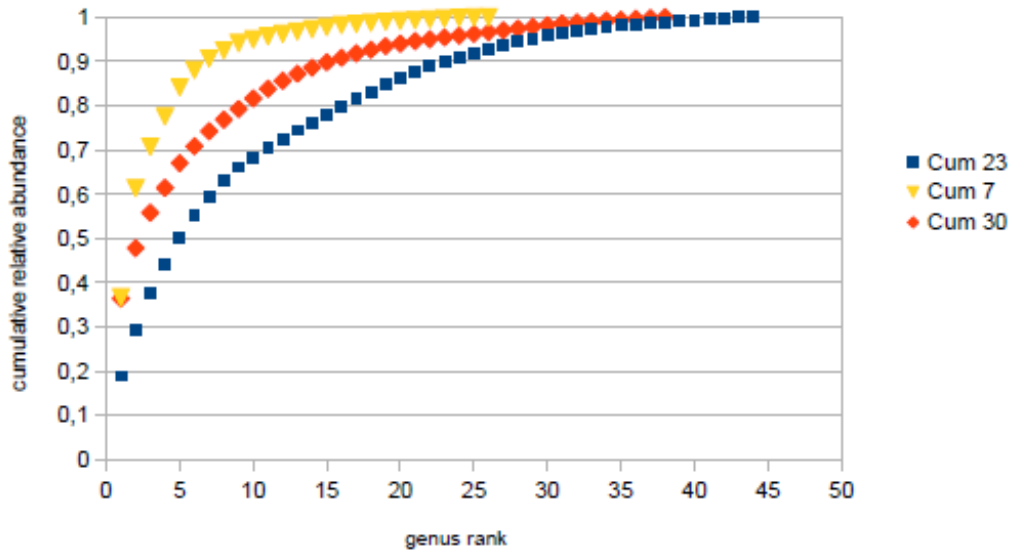


Figure 47. Comparison between K-dominance curves of station 7, 23 and 30 estimated by genetic diversity.

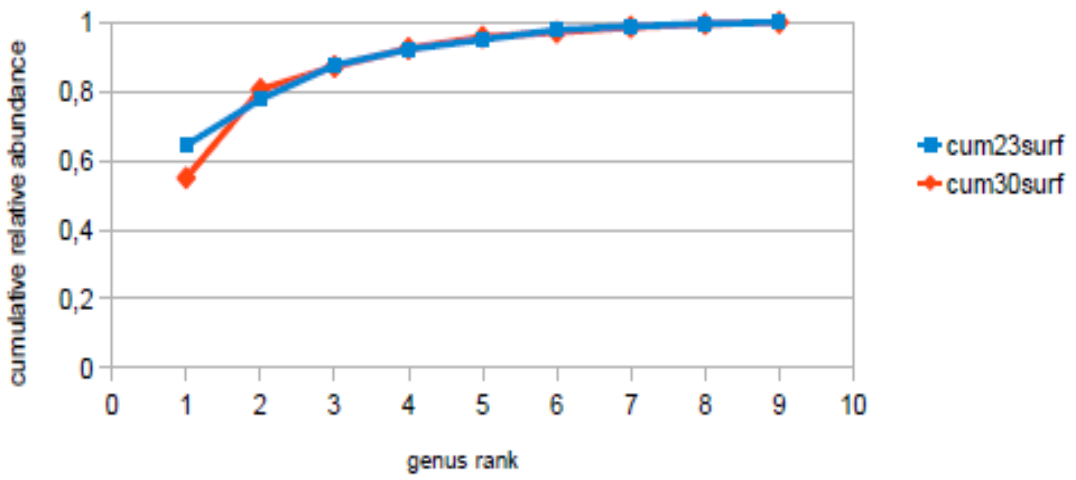


Figure 48. Comparison between K-dominance curves of station 23 and 30 estimated by morphological diversity.

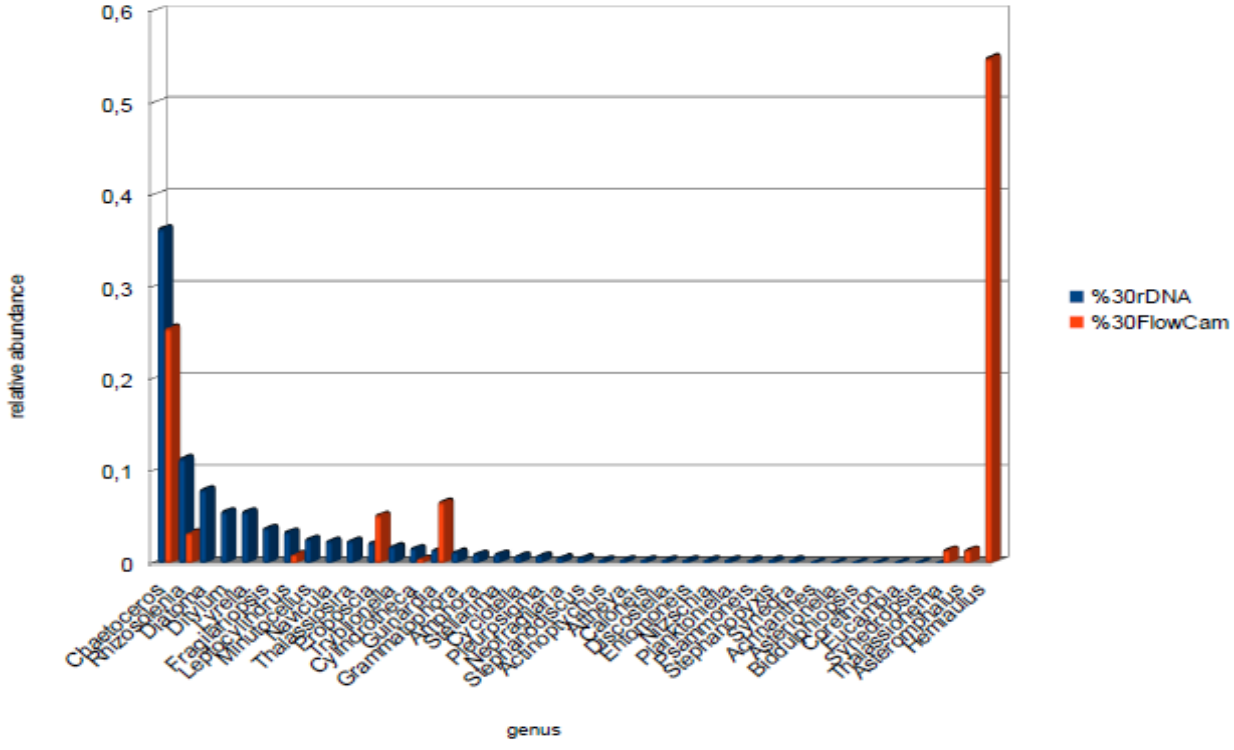


Figure 49. Comparison between diatoms community structure detected by morphological and genetic diversity for station 30.

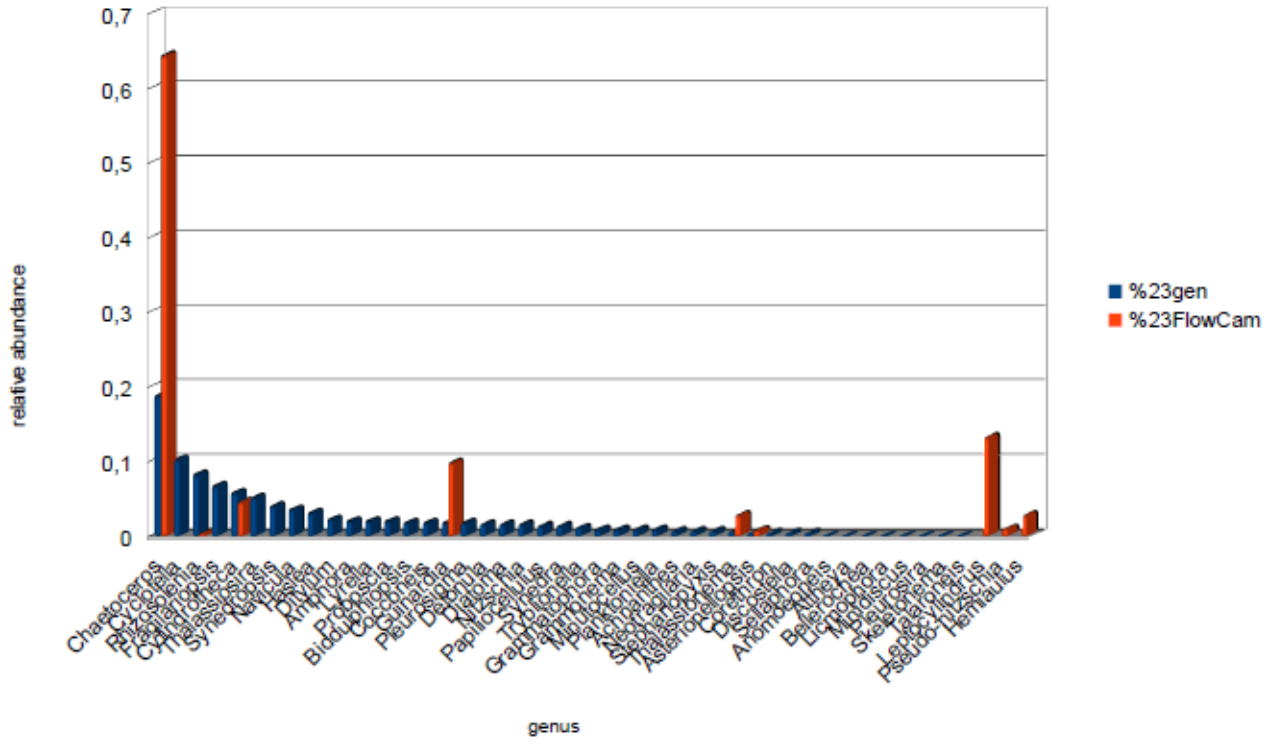


Figure 50. Comparison between diatoms community structure detected by morphological and genetic diversity for station 23.

Table 4. Biodiversity indexes calculated for the sampling stations 7, 23, 30 for morphological (red) and genetic (blue) diversity.

	S (genus richness)	Q (Kempton & Taylor 1976)	Shannon-Wiener (Shannon 1948)	Linearized Shannon (Jost 2006) (95% CI)	Estimated Shannon (Chao & Shen 2003) (95% CI)	Taxonomic distinctness (Warwick & Clarke 1995)	Interpolation slope
St 7 diatoms rDNA genus	26	21,52	1,97	7,17 (6,85-7,9)	1,99 (1,93-2,07)	4,96	0,81
St 23 diatoms rDNA genus	46	30,87	3,05	21,1 (20,89-25,94)	3,15 (3,05-3,26)	5,17	0,92
St 30 diatoms rDNA genus	38	31	2,49	12,06 (11,69-15,18)	2,6 (2,47-2,73)	5,35	0,9
St 23 diatoms FlowCam genus	9	6,73	1,22	3,39 (3,09-3,78)	1,23 (1,13-1,33)	3,9	0,61
St 30 diatoms FlowCam genus	9	6,73	1,3	3,67 (3,26-4,39)	1,34 (1,19-1,49)	3,86	0,6

Table 5. Ecological distances (Bray-Curtis Dissimilarity Index) between 7, 23, 30 sampling stations for morphological and genetic diversity (blue) and between stations for genetic diversity only (red).

Bray-Curtis Dissimilarity Index				
St 23 diatoms FlowCam genus	0			
St 30 diatoms FlowCam genus	0,62	0		
St 23 diatoms rDNA genus	0,63	0,67	0	
St 30 diatoms rDNA genus	0,54	0,67	0,24	0
St 7 diatoms rDNA genus	0			
St 23 diatoms rDNA genus	0,68	0		
St 30 diatoms rDNA genus	0,64	0,24	0	

6.2.3 Results: the shape of the community by abundance distributions

I analysed the plankton community structure of all available Tara stations considering both the abundance and the diversity of photosynthetic eukaryotes at the taxonomic resolution of 97% OTUs (see Chapter 2.3.3.4). The shape of the RADs of the sampled communities varies from station to station. Differences in the RADs shapes are mostly related to the most abundant OTUs (Fig. 51), while the heavy tail is always present. All the distributions seem to follow a power-law function ($p < 0.001$) (Fig. 52), accordingly to results from comprehensive study of the distribution function of total phytoplankton community (Pueyo 2006). When the RAD is split in common and rare components, the means of the common versus rare slope are significantly different ($p < 0.001$). The interpolation curve presents similar slopes for the rare part of the community (Z-test $p > 0.05$) and different slopes for the common part (Z-test, $p < 0.001$) (Fig. 53).

I compared the slope of the interpolation curves of the common OTUs to the environmental context of the stations to see if the extreme values of the slope corresponded to extreme values of the environmental conditions. The non-redundant variables identified by the Principal Component Analysis (Fig. 54) describing the dissimilarities among stations are represented in the

Multi-Dimensional Scaling plot in Fig. 55. Stations outside the cloud are stations where the slope of the common OTUs curve are further away from the mean (i.e. station 85, 82, 84, 67, 6, 49). This is consistent with the expectation that the distribution of the common part of the photosynthetic community strongly responds to the contextual environment, thus affecting the shape of the RADs. This aspect can have evident impact on the interpretation of biodiversity indexes if the distribution of the community is not known. The most straightforward example is, if biodiversity indexes that consider both richness and abundance are applied on samples of communities in a blooming phase, they will be disproportionally influenced by the dominant species and will not correctly represent the true diversity of the community.

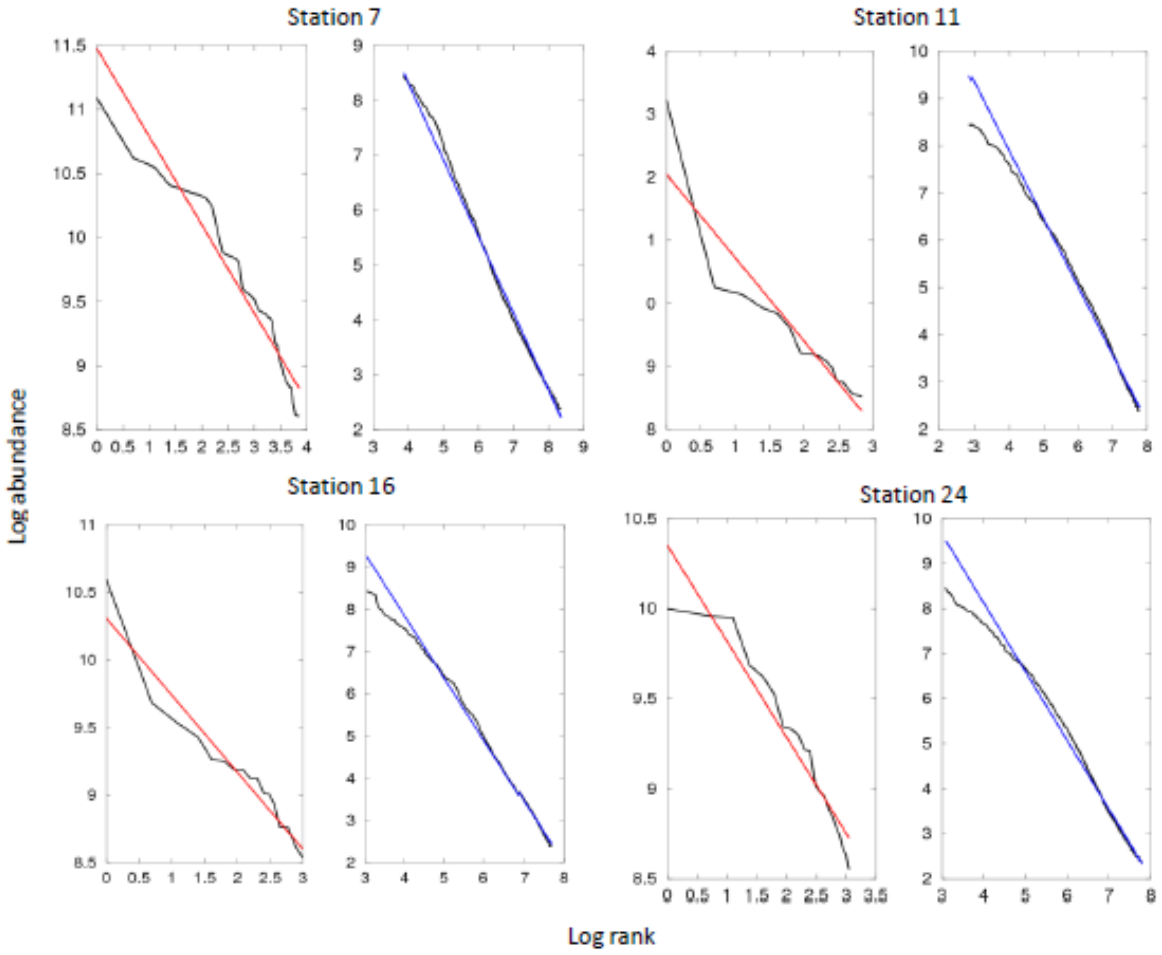


Figure 51. Rank Abundance Distributions for each surface station using V9 barcodes of photosynthetic eukaryotic plankton. Classes with abundance less than log(10) reads have been

eliminated. Classes with less than log(5000) reads are considered rare. Common (red) and rare (blue) curves of the community are showed separately with their respective interpolation.

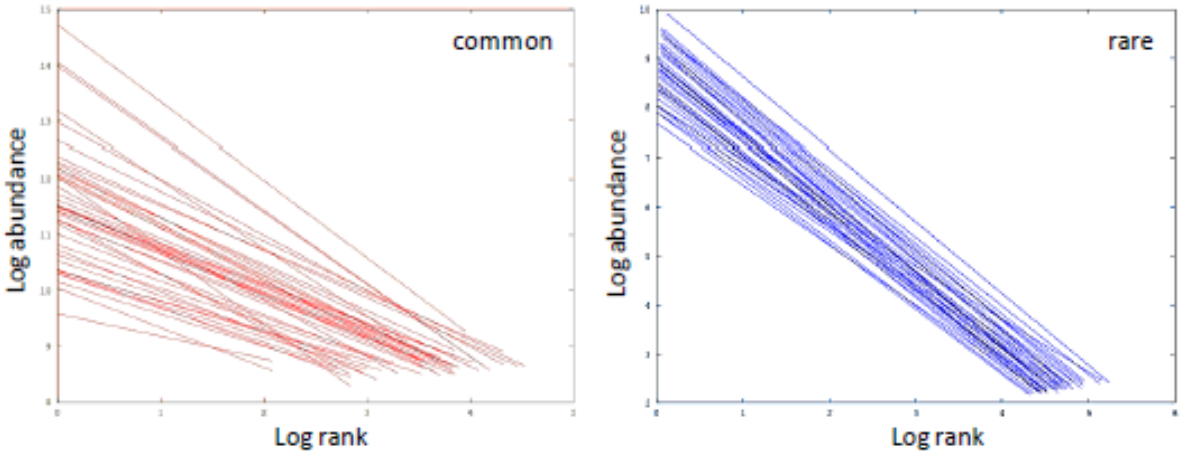


Figure 52. Overlay of the interpolation curves for all the stations. Classes with abundance less than log(10) reads have been eliminated. Classes with less than log(5000) reads are considered rare (blue).

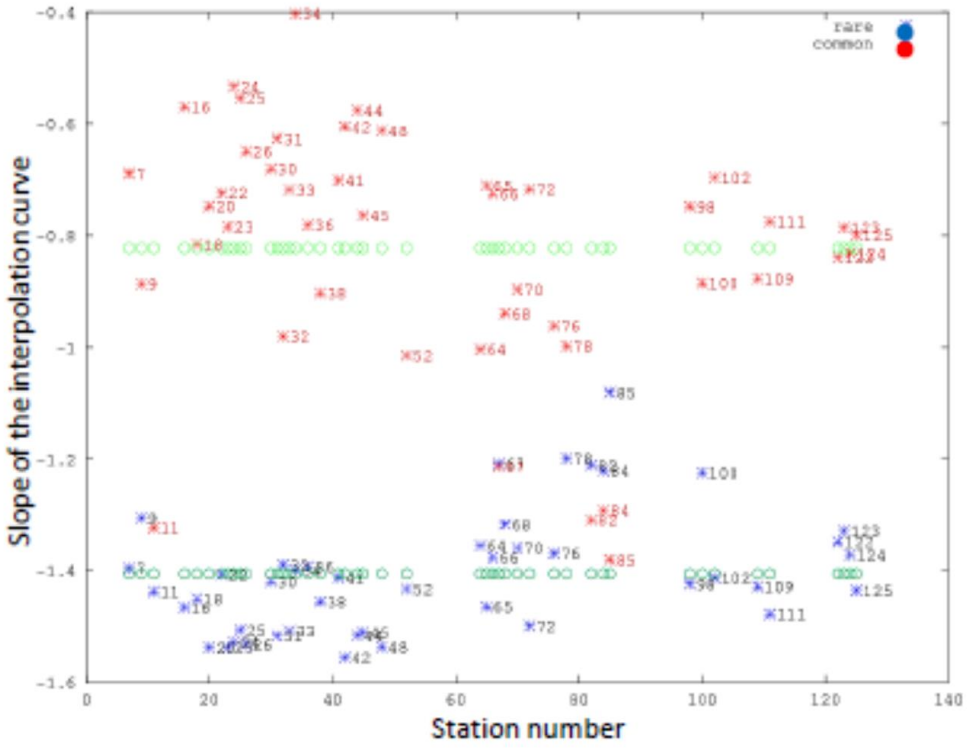


Figure 53. Scatterplot of the slopes for the common (red) and rare (blue) community interpolation curves for each station. Green circles represent the mean of the slope for the common and rare community.

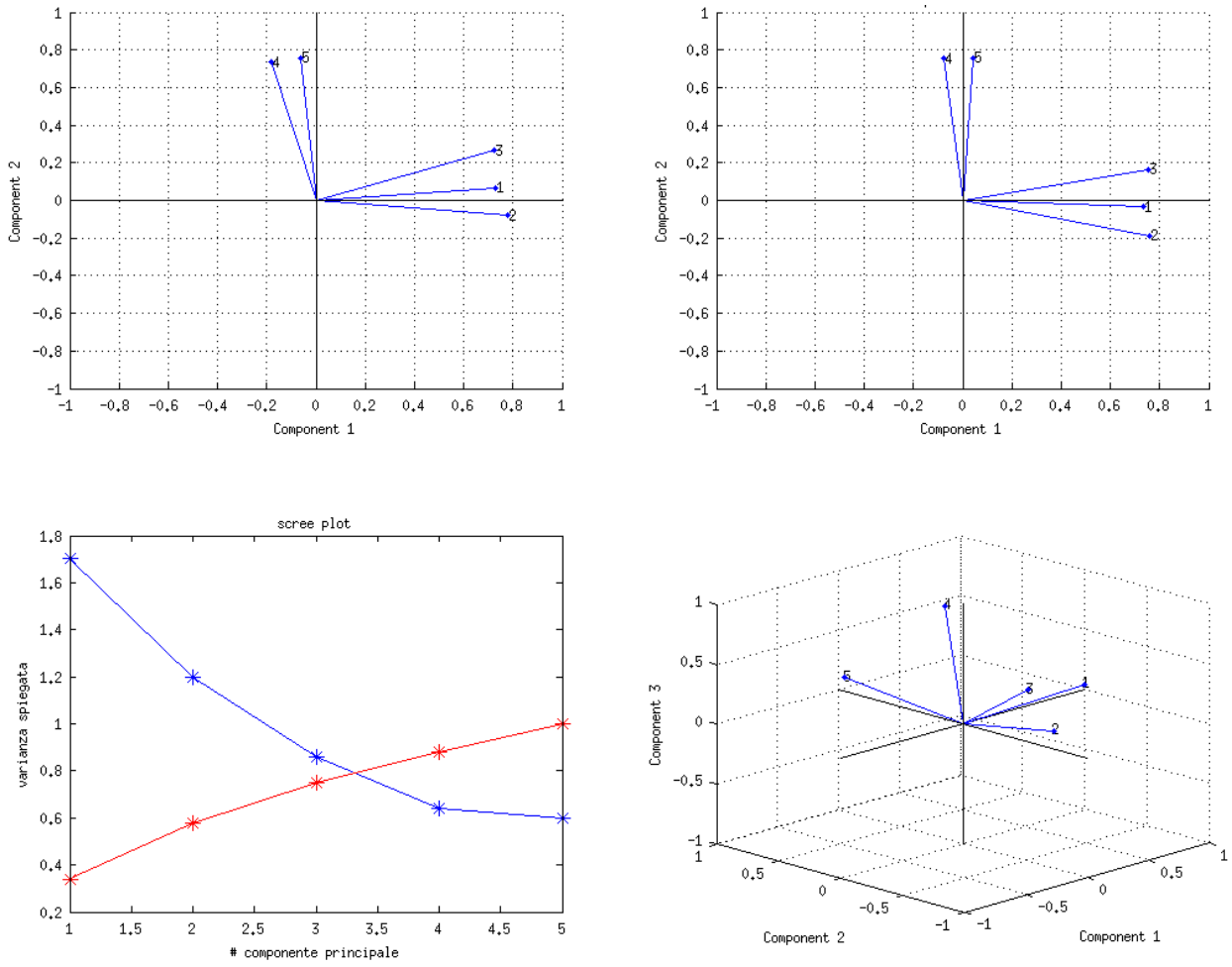


Figure 54. Principal Component Analysis using satellite variables (1=Sea Surface Temperature, 2=Chlorophyll, 3=Finite Size Lyapunov Exponent, 4=latitudinal component of Total Kinetic Energy and 5= longitudinal component of Total Kinetic Energy) indicates which variables are contributing the most in describing the dissimilarity among the stations in terms of their biophysical environmental context.

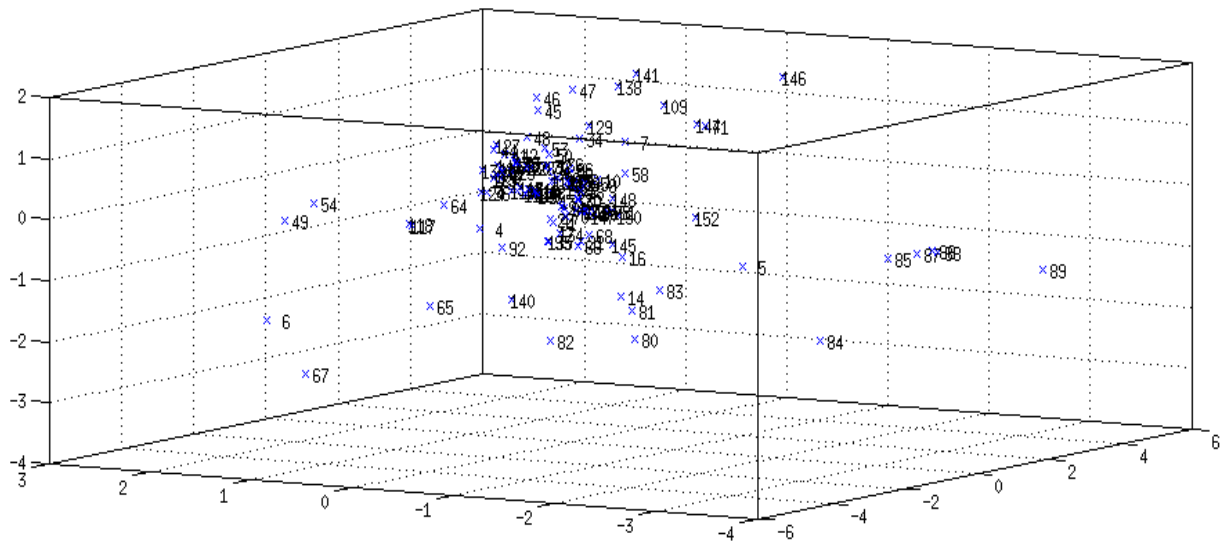


Figure 55. MultiDimensional Scaling showing a good representation (SStress = 0.19) of the dissimilarity among stations in terms of environmental context.

6.3 Discussion and conclusion

With this prospective chapter my overall intention was to highlight how biodiversity studies, weather based on remote sensing data, or morphological and molecular analyses, may be seriously biased if environmental variability and population characteristics are not taken into account. The complexity of this problem is further increased by the multiple ways proposed to measure biodiversity. Despite the advantage and disadvantages of all methods have been properly highlighted, an overall guide of how and when to use a specific method is not available. Different methods highlighted different aspects and all have limitations.

In particular I focused on the high variability observed in data collected in-situ. In-situ data provide only a snapshot in time and space and if not properly contextualized with the local environmental and physical variability, it is difficult to extrapolate processes linked to ecosystem functions. The major problems associated to in-situ data are: a) high environmental heterogeneity even at small scale, b) high data noise, c) uncertainty in the comparisons across studies. In this study I showed how high variability may occur in biodiversity estimated even at small scale, suggesting that sampling replicates, if not conducted exactly within the same homogeneous location, may yield substantially different results. Consistency across replicates may also be affected by the increasingly variable weather patterns associated to climate change. I also

compared taxonomic classification based on morphological analyses and genetic techniques to show how data heterogeneity may also arise by differences in methodologies. As a result in-situ dataset may include an high signal noise which must be properly minimized or modeled in order to obtain reliable and robust results. In this respect remote sensing information could provide an important tool to improve in-situ sampling as it provides the global view helpful to develop appropriate sampling design by taking into account local heterogeneity. But most importantly, data from remote sensing can be integrated with information provided in situ to have a broader view of the context and correctly interpret the locally available information from in-situ data.

The Mediterranean Sea is characterized by complex circulation and hydrography with characteristic water types present at different geographic locations (d'Ovidio *et al.* 2004; D'ortenzio & Ribera d'Alcalà 2009). These variable hydrographic patterns, in conjunction with plankton's individual life histories, are important to shape the distribution of plankton communities across the basin (Volpe *et al.* 2007). Therefore in regions where diversity may be largely influenced by the hydrographic context even at small scale (i.e. Mediterranean Sea), the choice of the scale and the sampling method are fundamental aspects to consider in biodiversity studies along with incorporating information about the biological and environmental context.

In the first case study, the data obtained by a continuous sampling technique (SeaFlow) evidenced that, on relatively small spatial scales (- in this case, ~10 kms), physical structures can constrain plankton distribution and affect biodiversity patterns. The change in biodiversity was particularly evident passing from the inner to the outer side of the eddy, and at the border of the filament constituting a two-fold barrier between different communities.

In the second case study I analysed the variation in the structure of diatoms communities among three stations considering morphological and molecular methods. I found substantially different structures across stations which can be related to the history of the water masses they belong to. Station 7 is likely to be strongly influenced by water masses coming from the Atlantic Ocean, which present peculiar physical and chemical characteristics and also transport species that are typical of another basin (Coll *et al.* 2010), therefore its dissimilarity compared to the other stations may not be surprising. The difference in the sampling time is also likely to have an effect. Indeed September, the month during which station 7 was sampled, is usually characterized by the late summer bloom (Garcés *et al.* 1999; Barale *et al.* 2008). According to the available

total Chl concentration data, station 23 is the most productive probably because of its proximity to the coast and anthropogenic activities. On contrary to station 7 and 23 which were chosen randomly, station 30, which also showed high biodiversity, was selected based on a satellite-guided sampling strategy aimed at targeting the core of an hydrodynamically very stable and recurrent eddy (Menna *et al.* 2012). As a result, with the available data it is not possible to infer if these hydrographical conditions influenced or not the community structure of station 30. Other stations close to Cyprus would be worthwhile to be investigated. Another way to infer if physical controls are capable of structuring the communities and promote differentiation among stations is to consider more phytoplankton groups beyond diatoms because different groups have different responses to local environmental conditions. Furthermore, this would allow comparison with remote sensed data of phytoplankton dominant types and diversity.

Some issues have been identified regarding the methods used in the pipeline leading from the real ecosystem to its representation in terms of biodiversity indicators. Errors occurring at any step between sampling and bioinformatic analysis may lead to inconsistent results in comparing various aspects of biodiversity at the different stations. First of all, it is important that the sampling effort is equal for all the stations and that the area is well represented (i.e. is environmental heterogeneity present?) (Cao *et al.* 2002). The statistical representativeness of the samples could be strengthened in different ways. One could use more replicates (i.e. Tara protocol provides 2 tubes for each fraction for each analysis, but replicates are not available yet) and include probabilistic analysis (Gotelli & Colwell 2001). Contextualizing with satellite and in-situ physical and chemical environmental data would moreover allow to include spatial and temporal variability in the analysis (i.e. sampling was made in different times, spanning also different seasons, thus not directly comparable).

Secondly, the need of a reference sequence for the identification of V9 rDNA tags and the statistical method used for the assignment leave out a great number of sequences that cannot be identified yet, but that should be quantified and included in the analysis in order to have a correct account of the abundances (Wilson et al. 2011). Moreover, multiple copies of the same sequence per individual could lead to an overestimation of abundances and a risk of change in the class frequencies (Kembel *et al.* 2012). Therefore, it is important to have an idea of the proportion of the number of copies for each taxa, or at least size fraction, to be able to do correct quantitative

analysis to study the ecology of the environment.

Finally, some technical problems have been identified for FlowCam data coming from the Mediterranean stations. The first problem was that the lens focus is optimized for zooplankton and big diatoms and dinoflagellates. Therefore most of the images on which the FlowCam data of this study is based are blurry and not identifiable. For this reason species have unequal probabilities of being detected (i.e. risk of unreliable abundances) and a subjective misclassification (see *Hemiaulus*) may be present. Moreover, the protocol at 10,000 particles is not sufficient for having a good profile of the community composition at high taxonomic levels (see comparison between diatom tags and FlowCam) because 90% of the particles, and in some case more, are rejected as not cells or not identifiable, images of the living cells are difficult to identify, therefore the result is an underestimation of biomass, abundance and richness. Increasing the number of particles analysed, increasing the size and the resolution of the images captured to include all the features of the cells (avoiding blur) and refine the classification with the help of expert taxonomists will improve the power of this method. It would allow to use it not only for monitoring purposes (because only the dominant and big components of the community can be detected with very approximate quantitative indications), but also to study the community composition and the ecological relationships structuring it with a robust approach that can give results more comparable with the genetic method. Indeed, it is not possible to give the FlowCam method the same resolution of the genetic ones, but considering its data as a subsample of the more resolved DNA tag data, both should measure the same differences in the community structure and show a certain degree of agreement. Moreover, morphological data are needed for identifying new species or new ecotypes that cannot be resolved by the genetics because of recent divergence or of lack of reference sequence (Rynearson & Armbrust 2000).

In the third case study I showed that the structure of the ecological community in terms of shape given by richness and abundance of its components is also an important aspect to be studied. Indexes are strongly influenced by the patterns of both common and rare species and they can imply important processes acting on the community assemblage. The shape reflects the changes in the community structure due to environmental context or acting ecological processes (Magurran 2004). If just a number (i.e. the value of an index) is considered, then all these changes could be masked or misinterpreted. Indexes are already challenging to interpret as the

change in their value is not proportionally related to a particular quantitative change in the community. Preliminary analyses run for all the stations suggested a power-law distribution of the community structure, with slope influenced by environmental context for the commons and similar for the heavy tails. Similarity of the slope of the rare part of the community can be due to: i) intrinsic property of the microbial ribosomial diversity (involving evolutionary processes) ; ii) intrinsic properties of the ocean (involving mixing and fluid-dynamical niches) ; iii) fractal organization of nature (involving biological and ecological processes related to body sizes such as metabolism and resource competition). To better investigate the law that best describes the rank abundance distributions and answer the above questions, OTUs should be classified at different percentages of similarity. Moreover, the distributions should be analysed for different size fractions and guilds separately. Finally, the relationship between OTUs present in the common's part and OTUs present in the heavy tails should be investigated to elucidate aggregation or connectivity patterns.

Clearly more dedicated studies are needed to better understand how biodiversity index behave in different situation and under different constraints. Arguably the most compelling issues are to determine if the remote sensed biodiversity resemble the "real" diversity and to develop new methods to interpolate data derived from all this sources.

Chapter 7 General Conclusions and perspectives

7.1 General conclusions and perspectives

The ocean has always fascinated humankind for its inaccessibility and mysterious content. Only relatively small areas of the ocean were long well known, above all for the interest in the exploitation of their resources, whereas open ocean environments were first ignored and considered homogeneous and poor of life (Hutchinson 1958). With the advent of technological development these ideas started to be reversed. Global explorations, observing systems and the development of remote sensing allowed to uncover a complex, dynamical and diversity rich environment (Longhurst 2010). Ecological and environmental studies showed the influence of the physical processes of the ocean on the growth and distribution of the species at the base of marine life: the plankton (Hernández-Garcıa & López 2004). Moreover, even species of higher levels of the trophic chain were described to aggregate in correspondence of particular hydrographic structures such as fronts and eddies (Hyrenbach *et al.* 2000; Rossi *et al.* 2008; Kai *et al.* 2009; Cotté *et al.* 2013). Nonetheless, determining which physical mechanisms and ecological interactions act to shape life and biodiversity in the pelagic ocean raises several technological and logistic problems that only recently, with the advent of high throughput techniques, high resolution global circulation models and multiple satellite missions for remote sensing, have started to be addressed. This improvement of the quality of data has allowed to resolve the interplay between physics and biology at the submesoscale. Indeed, at this scale, ocean variability is in accordance with the variability of the demography of the plankton. Not just one method is sufficient to describe and understand the physical and related biological processes acting in the marine environment and concurring to shape biodiversity. This understanding is however fundamental to manage natural resources and predict ecosystem responses.

In the marine like in the terrestrial environments, remote sensing has emerged as a powerful tool for measuring and monitoring biological resources (Hardman-Mountford *et al.* 2008; Hansen *et al.* 2010). On land its use was pushed much further, allowing to estimate biodiversity of primary producers based on spectral heterogeneity of optical signals (Rocchini *et al.* 2010), paving the way to broad scale ecological studies. In the ocean, remote sensing has provided incomparable views of global spatio-temporal patterns of phytoplankton distribution and biomass (Field *et al.*

1998). However, even remote sensed data cannot alone elucidate the complexity of the oceanic realm. For example while primary productivity can be derived from satellite data, extrapolating information about biodiversity of primary producers and of consumers or inferring spatial and temporal variability in biological aggregations from these data, is an open challenge.

In this work I tried to integrate the complementary advantages of models, in situ observations and remote sensing for shedding some light on the mechanisms that influence the formation of marine biodiversity rich areas and on the geographical identification of global biodiversity hotspots in the open ocean. To do this, I had to face several issues such as how to appropriately define biodiversity in the marine plankton ecosystem; which ecologically relevant spatial and temporal scales to measure; which synoptic environmental measure to produce and what technological limitations could have been overcome. Due to the dynamic nature of the vast ocean environment and to the microbic nature and passive dispersal of plankton, it is not possible to define local (alpha) diversity of the ecological community as on the terrestrial environment and for the same reasons it is difficult also to estimate the total regional (gamma) diversity and its turnover (beta). Therefore, I used an approach that considers the spatial distribution of water masses characterized by different plankton communities distinguished by their dominant phytoplankton group. This proxy can be seen as a mixture of alpha and beta diversities as the community changes with its dispersal mean. As the patchy distribution of plankton and its demography are influenced by the environmental physical heterogeneity of mesoscale and submesoscale structures, I considered spatial scales of few 100 kms and temporal scales of few weeks to estimate biodiversity. I used a remote sensing proxy to have a synoptic information that allows higher resolution and temporal variability. I also considered various levels of complexity of life when estimating biodiversity, comparing morphological, functional and molecular measures, integrating the different sources to overcome technological limitations.

In Chapter 3 I used a physical and ecological global circulation model of a virtual plankton ecosystem to design and implement a mesoscale marine ecosystem metric and I applied it to identify areas of enhanced plankton biodiversity. The main advantage of this metric is that it can be defined without the need of having information about the whole community structure but just spatial information of dominant functional groups. I conceptualized a spatial-based index of biodiversity, analogous to other diversity indexes, such as the well-known Shannon entropy index

(Shannon & Weaver 1948; Shannon & Weaver 1949), that is predictably associated with local biodiversity of primary producers. I verified that the best global spatial scale to use to predict such a diversity is in the range of 1 degree. This scale is typical of the formation of hydrographical structures capable of creating fluid-dynamical niches where diverse dominant phytoplankton groups prevail and where stirring and mixing can act to merge different communities and/or favor the appearance of specialists, thus increasing biodiversity. The patchiness of the biophysical environment (abiotic plus microbes) varies in the ocean within the spatial and temporal scale determined by the response of the demography of the plankton to the physical and chemical characteristics (d'Ovidio *et al.* 2010). Turbulence allows mixing of communities at spatial and temporal scales notably inferior to scales of evolutionary and competitive exclusion, and the emergence of biodiversity hotspots. It is not surprising that the correspondence between spatial-based and local diversity is maximal for regions characterized by high kinetic energy, where biophysical patchiness and mixing are stronger. Meso and submesoscale structures depend on the scale of the radius of deformation of the flow due to rotational effects (Rossby radius). This radius varies according to latitude (decrease towards higher latitudes), therefore, in case of regional studies, the scale of the radius used to calculate the diversity proxy can be adjusted. In this modeled environment predation is very simple and plankton species are approximated by functional types and groups, therefore quantitative mismatches in the representation of global biodiversity compared to real data are envisageable. This is the reason why then I directly used remote sensing data to derive spatial-based diversity and compared them with corresponding in situ data for local diversity.

In Chapter 4 I applied this approach to satellite-derived information and inferred a synoptic view of global biodiversity distribution of primary producers. Remote sensing data are able to show the patchy distribution of the different dominant phytoplankton groups, that is shaped by the physical structures on which the passive dispersal of plankton depends (Martin 2003; Alvain *et al.* 2008; d'Ovidio *et al.* 2010). I validated the global patterns of satellite-based diversity by comparison with independent observations at the regional scale, and in particular using in situ morphological data. Local biodiversity is enhanced by the contribution of physical dynamical processes such as advection, dispersal and mixing that typically create patches of water masses characterized by different communities but that can mix over short spatiotemporal scales. Indeed,

this enhancement is maximized in regions characterized by fronts, upwelling, boundary currents, biogeographical transition zones and island reefs. I found that the spatial-based diversity index displays relations with temperature and productivity in agreement with already known macroecological patterns where unimodal distributions show maximum biodiversity at intermediate levels of temperature and productivity (Hillebrandt 2004; Tittensor *et al.* 2010; Raes *et al.* 2011; Fuhrman *et al.* 2008; Dodson *et al*. 2000; Irigoien *et al.* 2004; Vallina *et al.* 2014). I showed seasonal variation in biodiversity patterns and I established the temporal stability of hotspots. Stable hotspots are mainly areas of transition among different ocean biogeographical provinces where spatial environmental heterogeneity stands out and exclude regions of strong ecological successions, typically strong upwelling and high temperate zones, where spatial heterogeneity depends strictly on seasonal temporal changes (Longhurst 2010). The fundamental role of heterogeneity for biological diversity at all levels of ecological organization and scale is long known (Levin & Paine 1974; Ives & May 1985; Tilman 1994). In landscape, patchiness highlights the spatial matrix of ecological processes and the fluxes of its biotic and abiotic components. Patchiness influences resource availability, niche partitioning, dispersal and connectivity of populations and succession of communities. Studies on terrestrial and marine ecosystems have pointed out the ecological perspective of habitat spatial and temporal heterogeneity being related to higher biodiversity due to the potential of more ecological niches to be exploited by different species. Spatiotemporal changes disrupt the continuity of the biological communities and are able to reduce competitive exclusion on wider spatiotemporal scales and promote species persistence and coexistence (e.g. communities on intertidal rocky platforms (Paine 1966; Levin & Paine 1974; Paine 1974), coral reefs and tropical forests (Porter 1972; Connell 1978; Sousa 1984), deep seas (Dayton & Hessler 1972), general (Levin & Hesshaimer 2000; Tews *et al.* 2004). In the pelagic realm, transport processes are mainly responsible of these spatiotemporal changes. Global ocean biogeography is firstly defined by large scale circulations. Instabilities of large scale circulations give rise to meso and submesoscale structures that are able to create environmental heterogeneity. Plankton demography and distribution is in resonance with these structures. This peculiarity is at the base of the conception of this study aimed at estimating biodiversity using spatial-based information detectable from remote sensing.

In Chapter 5 I interpreted the association between remote sensed diversity and local biodiversity to reflect biophysical processes acting on the plankton community and beyond. Indeed, these processes act also indirectly on other levels of the ecosystem, promoting aggregation and trophic interactions among diverse species. Diversity at one trophic level can influence the effects of diversity at adjacent and nonadjacent trophic levels, and is also influenced by environmental context and intraguild predation (Bruno & O'Connor 2005). Therefore, relations between physical dynamics, diversity of primary producers and consumers both of the entire trophic chain and for separate guilds are to be elucidated as few studies exists that consider the effects of diversity at multiple levels. Biodiversity hotspots are areas of aggregation where density dependent effects may be reduced, negative interactions such as competition may be modified due to the possibility of differentiation in resource utilization and of development of positive interactions among diverse species. The new satellite based diversity index defined in Chapter 4 may imply the fundamental food-web mechanisms claimed by Bakun 1996 (Bakun 1996), which include nutrient enrichment, concentration of prey, and aggregation of predators. The congruence showed between most of the spatial-based diversity hotspots of primary producers and local diversity of predators and in general of cross-taxa diversity, may also have implications for energy transfer from lower- to other upper-trophic-level species. For example, Ainley et al. (2009), Olson (1989), Sournia (1994), Young et al. (2014), and Cotté (2013 under revision) (Olson & Olson 1989; Sournia 1994; Ainley & Siniff 2009; Cotté *et al.* 2013; Young *et al.* 2014) showed spatial co-occurrence of divers predators, pelagic nekton and plankton at frontal systems, which are related to remote sensing diversity as they are present in areas of high biophysical heterogeneity. Submesoscale and mesoscale structures, such as eddies and fronts, can result in shear zones between different water masses or current flow speeds that allow bioaccumulation of preys (Menkes *et al.* 2002). This hydrographically induced bioaccumulation can give rise to the creation and dissipation of diversity hotspots faster than classical bottom-up processes (Prairie *et al.* 2012).

The purpose of Chapter 6 was to investigate which kind of biological information should be used to compare in-situ and satellite-based biodiversity proxies. Indeed these two measures provide very different ways of looking at the planktonic community: one that is very detailed, but geographically and temporally sparse, the other that has high spatiotemporal resolution, but that

is associated to an extremely shallow view of the community. I showed that a continuous in situ sampling technique evidenced plankton biodiversity patterns associated to physical frontal structures. I highlighted caveats and pitfalls of the case-study cross-validations, such as the not univocal correspondence between morphological and genetic diversity measured on the same sample, and the problem of applying biodiversity indexes to communities that are characterized by very different abundance distributions, depending on the time and place of the sampling.

Certainly, the satellite-based diversity has limitations too as it can only represent transport-related features of the phytoplankton community assembly. Indeed, it mainly detects zones of increased diversity due to merging of different communities and it is not able to give information on the degree of diversity of the original communities. Therefore it is more related to proximate causes (physico-chemical and ecological) than ultimate causes (evolutionary based) of diversity. The satellite-based diversity proxy discussed here relies on the detection of bio-optical patches characterized by different biophysical features, therefore the classification is limited to such detectable features. Other algorithms exist that propose classifications of bio-optical signals (Volpe *et al.* 2007; Raitsos *et al.* 2008; Hirata *et al.* 2011; Palaczet *et al.* 2013). However, each method was developed for different sensors and has intrinsic proper limitations, beyond problems related to atmospheric corrections and instrument calibrations. Moreover, the area over which the diversity of patches is calculated cannot be too small, to allow statistical power, so a lower limit exist for the spatial resolution provided by this method so far. This limit is intrinsically related to the available resolution of ocean color satellite data. Limitations are also evident for high latitudes that are data-deficient for a considerable part of the year, when illumination levels are low and cloud coverage is high, limiting the power of observation closer to the poles. Furthermore, in case of very high chlorophyll concentration, the distinction of dominant groups can be biased (due to change in size abundance distribution of the plankton community during blooms and overwhelming abundance of chl-a), altering the validity of the remote sensed diversity index in highly productive zones, thus not allowing a correct estimation of local plankton diversity. Finally, satellite data can provide information only on the first 10s meters depth of the ocean (depth that vary depending on the region of the globe according to turbidity (Robinson 2004), but it is generally assumed that the plankton community that satellites detect can be representative of the mixed layer. Moreover, horizontal front that define (sub)mesoscale

structures responsible of the fluid-dynamical niches detected by satellite have associated vertical dynamics. To explore the role of vertical dynamics in the structuring of biodiversity patterns, congruencies between vertical velocities and biodiversity hotspots could be investigated.

Future studies to be addressed include the improvement and ad hoc tuning of reanalysis of ocean color to be applied to specific oceanic regimes and biogeographical regions. This development would allow remote sensing diversity to be estimated also for ecosystems characterized by particular environmental conditions and that are of strong interest for management purposes. Among them, global coastal waters are subjected to high levels of suspended organic and inorganic matter thus turbidity, which hinders optical signal capacity. They are also the most exploited ecosystems of the world due to anthropic activities and the ones most in need of insight on the effects of biodiversity on ecosystem functioning and services. Future research should also target global ocean expeditions that provide so far high throughput molecular information to quantify the relationship between seascape patchiness and the very fine level of biological organization. Indeed, high throughput molecular information provides, when opportunely treated, a high resolution image of the community with the potential to unveil changes in acting ecological processes. Finally, to better evaluate the bottom up effect of trophic transfer of biodiversity from plankton to mid- and higher levels, firstly, single guilds at a time should be examined. Then, regional studies that consider time series data of both key oceanographic features and successions of communities and aggregations of biomass and diversity at subsequent levels of the trophic chain could be used for comparisons with changes in remote sensed diversity. Upwelling systems typical of eastern boundary currents would represent a plausible candidate for such a kind of study, due to their well-known seasonal patterns and well-known supporting biological communities. Moreover, time specific data may be more valuable than time averaged data in studying such a highly dynamical environment in deeper detail.

7.2 Implications for management and conservation

The protection of biodiversity has been addressed by management policies around the world as the primary conservation concern. Like terrestrial and coastal ecosystems, the open ocean is not homogeneous, rather is a mosaic of high biodiversity areas, where many species aggregate, and areas with low biodiversity (Worm *et al.* 2003; Worm *et al.* 2005; Sydeman *et al.* 2006). Despite

the increasing conservation awareness, today open ocean waters remain poorly protected with only between 0.08% and 0.65% of the open ocean currently falling within Marine Protected Areas (MPAs) (Wood *et al.* 2008). Identifying offshore and pelagic areas of primary interest for management and protection is challenging but possible, especially if the mechanisms of persistence can be identified (Hyrenbach *et al.* 2000). Pelagic areas of ecological importance are found to be dependent on static bathymetric features (i.e. reefs, shelf break, submarine canyons, seamounts, islands), persistent hydrographic features (fronts, boundary currents and gyres) and more ephemeral transport features (eddies, filaments, upwelling plumes) (Hyrenbach et al. 2000). The range of distribution of nektonic species mostly mirrors large scale oceanographic regimes and currents, known to influence productivity and plankton biogeography (Sund *et al.* 1981; Brodeur *et al.* 1999). At meso and submesoscale, predators' distribution is influenced by dispersion and availability of preys (Hunt Jr & Schneider 1987; Rose & Leggett 1990). Dispersal and availability of preys depends on the bioaggregation and enhanced productivity effects generated by specific hydrological structures such as eddies, fronts and filaments (Franks 1992; Larson *et al.* 1994; Olson & Hood 1994). Novel methodologies to predict areas characterized by these more ephemeral structures are needed. Moreover, new approaches to describe connectivity patterns and source and sink areas of biodiversity in the open ocean should be developed.

Despite the challenges typical of the conservation of natural resources in the open ocean, the improvement of integrative methods including remote sensing seems promising. For instance, a similar approach is already applied by the Pacific Hawaiian Fisheries to daily predict loggerhead turtle habitat based on their regular occurrence along offshore frontal systems (Etnoyer *et al.* 2006). Developing applicable metrics that combine remote sensing data on diversity of primary producers with consumer distributions could be a valuable approach for understanding the spatial organization of large marine ecosystems and predict areas of species aggregations at particular point in time. These predicted areas require an adaptive management as their boundaries may vary in space and time in response to system dynamics (Hobday & Hartmann 2006). These kind of pelagic marine protected areas are under consideration for ecosystem protection (Halpern *et al.* 2008), but synoptic global level information, especially at the scale of the whole ecosystem, is needed to efficiently and wisely choose the locations of potential protected sites.

Because of its relatively high temporal and spatial resolution provided with minimum effort at the

global level and because of the variety of information detectable (both abiotic and biotic) about the ecosystem, including the diversity proxies proposed in this study, remote sensing applications can be strongly envisaged for the Ecosystem Based Management and Marine Spatial Planning frameworks of the high seas (Ardron *et al.* 2008; Game *et al.* 2009). On the other hand, remote sensing may not be appropriate to identify geographically small areas (<< 10000 km^2) of particular high biodiversity values, such as hotspots of endemic species or of specific conservation priorities such as nursery and breeding grounds, which are not strictly related to trophic interaction processes. With the aim of providing a global biodiversity proxy, I applied the satellite-derived τ index to identify relatively large areas (~10000 km^2) of predicted elevated biodiversity, which seems to concentrate mid- to upper-trophic-level consumers, particularly top predators, despite the inherent variability in the system. Areas identified as hotspots are mainly frontal zones, eastern boundary currents, biogeographical transition zones and coral reefs, most of them already known as areas of importance for aggregation and migration of endangered large marine vertebrates and for the presence of very rich biota (Sournia 1994; Hyrenbach *et al.* 2000; Cotté *et al.* 2007; Kai *et al.* 2009). Indeed, hotspots are predicted along boundary currents such as the California current, the Gulf Stream, the Kuroshio Current, upwelling regions off the coast of North America, South America and Africa, Agulhas Retroflection and Malvinas-Brasil Confluence zones and subpolar frontal systems, Indian and Pacific Ocean atolls and small islands and Australian and Indonesian coral reefs. Because these areas of high diversity are mainly located in biogeographic transition zones and in zones where transport regimes are of primary importance, the ability to monitor their spatial shifts and temporal persistence is of great concern in the perspective of climate change. Indeed, the same areas are predicted to be highly affected mainly due to temperature driven species range shifts and shift or strengthening or weakening of oceanographic processes such as upwelling, convergent fronts and current systems (Bellard *et al.* 2012; Doney *et al.* 2012). In this light, the detection of pelagic biodiversity hotspots using remote sensing information (τ index) as a surrogate, can provide a feasible process based tool for the biological characterization of large areas at the global level without the need of in-situ census (Myers et al. 2000).

Regardless of the advantages and limitations of the novel method here described, for conservation practices to be successful, international agreements to define protected areas and

ensure their enforcement must be developed. Indeed, the open ocean is part of the Areas Beyond National Jurisdiction, therefore no legal power is recognized and can be applied in these zones. Political forces often blame the potential fallibility of scientific knowledge and use this as an excuse not to act. At this point, making use of the best available science is the best available choice. Nonetheless, the true solution to the problem seems not to be found in the natural sciences, but rather in humanity. The ocean continues to suffer from the tragedy of the commons (Hardin 1968) and the world is not realizing that we are all actors of the same tragedy. Which will be the end depends on all of us. Let's prove the most clever man in history was wrong: 'Only two things are infinite, the universe and human stupidity, and I am not sure about the former' (A. Einstein).

Chapter 8 References

Abbott M.R. & Zion P.M. (1985). Satellite observations of phytoplankton variability during an upwelling event. *Continental Shelf Research*, 4, 661-680.

Abraham E.R. (1998). The generation of plankton patchiness by turbulent stirring. *Nature*, 391, 577-580.

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

Achtman M. & Wagner M. (2008). Microbial diversity and the genetic nature of microbial species. *Nature Reviews Microbiology*, 6, 431-440.

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

Ainley D.G. & Siniff D.B. (2009). The importance of Antarctic toothfish as prey of Weddell seals in the Ross Sea. *Antarctic Science*, 21, 317-327.

Allen A.P., Gillooly J.F. & Brown J.H. (2007). Recasting the species-energy hypothesis: the different roles of kinetic and potential energy in regulating biodiversity. *Scaling biodiversity*, 1.

Allen A.P., Li B.L. & Charnov E.L. (2001). Population fluctuations, power laws and mixtures of lognormal distributions. *Ecology letters*, 4, 1-3.

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

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

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

Amato A., Kooistra W.H., Levialdi Ghiron J.H., Mann D.G., Pröschold T. & Montresor M. (2007). Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist*, 158, 193-207.

Anderson R.P., Lew D. & Peterson A.T. (2003). Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological modelling*, 162, 211-232.

Angel M.V. (1993). Biodiversity of the pelagic ocean. *Conservation Biology*, 7, 760-772.

Appeltans W., Ahyong S.T., Anderson G., Angel M.V., Artois T., Bailly N., Bamber R., Barber A., Bartsch I. & Berta A. (2012). The magnitude of global marine species diversity. *Current Biology*, 22, 2189-2202.

Ardron J., Gjerde K., Pullen S. & Tilot V. (2008). Marine spatial planning in the high seas. *Marine Policy*, 32, 832-839.

Babiano A., Boffetta G., Provenzale A. & Vulpiani A. (1994). Chaotic advection in point vortex models and two-dimensional turbulence. *Physics of Fluids (1994-present)*, 6, 2465-2474.

Bakun A. (1996). *Patterns in the ocean: ocean processes and marine population dynamics. California Sea Grant College System, National Oceanic and Atmospheric Administration, in cooperation with Centro de Investigaciones Biológicas del Noroeste*.
Bakun A. (2006). Fronts and eddies as key structures in the habitat of marine fish larvae: opportunity, adaptive response and competitive advantage. *Scientia Marina*, 70, 105-122.
Ballance L.T., Pitman R.L. & Fiedler P.C. (2006). Oceanographic influences on seabirds and cetaceans of the eastern tropical Pacific: a review. *Progress in Oceanography*, 69, 360-390.
Barale V., Jaquet J.-M. & Ndiaye M. (2008). Algal blooming patterns and anomalies in the Mediterranean Sea as derived from the SeaWiFS data set (1998–2003). *Remote Sensing of Environment*, 112, 3300-3313.
Barton A.D., Dutkiewicz S., Flierl G., Bragg J. & Follows M.J. (2010). Patterns of Diversity in Marine Phytoplankton. *Science*, 327, 1509-1511.
Barton A.D., Pershing A.J., Litchman E., Record N.R., Edwards K.F., Finkel Z.V., Kiørboe T. & Ward B.A. (2013). The biogeography of marine plankton traits. *Ecology letters*, 16, 522-534.
Belkin I.M. (2009). Rapid warming of large marine ecosystems. *Progress in Oceanography*, 81, 207-213.
Bellard C., Bertelsmeier C., Leadley P., Thuiller W. & Courchamp F. (2012). Impacts of climate change on the future of biodiversity. *Ecology letters*, 15, 365-377.
Berke S.K., Jablonski D., Krug A.Z. & Valentine J.W. (2014). Origination and Immigration Drive Latitudinal Gradients in Marine Functional Diversity. *Plos One*, 9, e101494.
Bettencourt J.H., López C. & Hernández-García E. (2012). Characterization of coherent structures in three-dimensional flows using the finite-size Lyapunov exponent. *arXiv preprint arXiv:1207.1975*.
Bledsoe E.L., Phlips E.J., Jett C.E. & Donnelly K.A. (2004). The relationships among phytoplankton biomass, nutrient loading and hydrodynamics in an inner-shelf estuary. *Ophelia*, 58, 29-47.
Block B.A., Jonsen I.D., Jorgensen S.J., Winship A.J., Shaffer S.A., Bograd S.J., Hazen E.L., Foley D.G., Breed G.A., Harrison A.L., Ganong J.E., Swithenbank A., Castleton M., Dewar H., Mate B.R., Shillinger G.L., Schaefer K.M., Benson S.R., Weise M.J., Henry R.W. & Costa D.P. (2011). Tracking apex marine predator movements in a dynamic ocean. *Nature*, 475, 86-90.
Boffetta G., Lacorata G., Redaelli G. & Vulpiani A. (2001a). Detecting barriers to transport: a review of different techniques. *Physica D: Nonlinear Phenomena*, 159, 58-70.
Boffetta G., Lacorata G., Redaelli G. & Vulpiani A. (2001b). Detecting barriers to transport: a review of different techniques. *Physica D: Nonlinear Phenomena*, 159, 58-70.
Bray J.R. & Curtis J.T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecological monographs*, 27, 325-349.
Brewin R.J.W., Sathyendranath S., Hirata T., Lavender S.J., Barciela R.M. & Hardman-Mountford N.J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecological Modelling*, 221, 1472-1483.
Briggs J.C. (2011). Marine extinctions and conservation. *Marine biology*, 158, 485-488.
Brodeur R.D., Mills C.E., Overland J.E., Walters G.E. & Schumacher J.D. (1999). Evidence for a substantial increase in gelatinous zooplankton in the Bering Sea, with possible links to climate change. *Fisheries Oceanography*, 8, 296-306.

Brownlee C. & Taylor A.R. (2002). Algal calcification and silification. *eLS*.
Bruno J.F. & O'Connor M.I. (2005). Cascading effects of predator diversity and omnivory in a marine food web. *Ecology Letters*, 8, 1048-1056.
Butchart S.H., Walpole M., Collen B., van Strien A., Scharlemann J.P., Almond R.E., Baillie J.E., Bomhard B., Brown C. & Bruno J. (2010). Global biodiversity: indicators of recent declines. *science*, 328, 1164-1168.
Cadotte M.W. (2006). Dispersal and Species Diversity: A Meta-Analysis. *The American Naturalist*, 167, 913-924.
Cadotte M.W., Carscadden K. & Mirotchnick N. (2011). Beyond species: functional diversity and the maintenance of ecological processes and services. *Journal of applied ecology*, 48, 1079-1087.
Cadotte M.W. & Fukami T. (2005). Dispersal, spatial scale, and species diversity in a hierarchically structured experimental landscape. *Ecology letters*, 8, 548-557.
Cao Y., Larsen D., Hughes R., Angermeier P. & Patton T. (2002). Sampling effort affects multivariate comparisons of stream assemblages. *Journal of the North American Benthological Society*, 21, 701-714.
Cardinale B.J., Duffy J.E., Gonzalez A., Hooper D.U., Perrings C., Venail P., Narwani A., Mace G.M., Tilman D. & Wardle D.A. (2012). Biodiversity loss and its impact on humanity. *Nature*, 486, 59-67.
Cardinale B.J., Srivastava D.S., Duffy J.E., Wright J.P., Downing A.L., Sankaran M. & Jouseau C. (2006). Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature*, 443, 989-992.
Carstensen J., Frohn L.M., Hasager C.B. & Gustafsson B.G. (2005). Summer algal blooms in a coastal ecosystem: the role of atmospheric deposition versus entrainment fluxes. *Estuarine, Coastal and Shelf Science*, 62, 595-608.
Cavender-Bares J., Kozak K.H., Fine P.V. & Kembel S.W. (2009). The merging of community ecology and phylogenetic biology. *Ecology letters*, 12, 693-715.
Cermeño P., Marañón E., Harbour D., Figueiras G.F., Crespo G.B., Huete-Ortega M., Varela M. & Harris P.R. (2008). Resource levels, allometric scaling of population abundance, and marine phytoplankton diversity. *Limnology and Oceanography*, 53, 312-318.
Chakravorty S., Helb D., Burday M., Connell N. & Alland D. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of microbiological methods*, 69, 330-339.
Chao A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of statistics*, 265-270.
Chapin III F.S., Zavaleta E.S., Eviner V.T., Naylor R.L., Vitousek P.M., Reynolds H.L., Hooper D.U., Lavorel S., Sala O.E. & Hobbie S.E. (2000). Consequences of changing biodiversity. *Nature*, 405, 234-242.
Chase J.M. & Leibold M.A. (2002). Spatial scale dictates the productivity–biodiversity relationship. *Nature*, 416, 427-430.
Chavez F.P. & Messié M. (2009). A comparison of eastern boundary upwelling ecosystems. *Progress in Oceanography*, 83, 80-96.
Chavez F.P., Messié M. & Pennington J.T. (2011). Marine primary production in relation to climate variability and change. *Annual review of marine science*, 3, 227-260.
Chelton D.B., Ries J.C., Haines B.J., Fu L.-L. & Callahan P.S. (2001). Satellite altimetry. *International Geophysics*, 69, 1-ii.

Chelton D.B., Schlax M.G., Samelson R.M. & de Szoeke R.A. (2007). Global observations of large oceanic eddies. *Geophysical Research Letters*, 34.

Chen B., Landry M.R., Huang B. & Liu H. (2012). Does warming enhance the effect of microzooplankton grazing on marine phytoplankton in the ocean. *Limnol. Oceanogr*, 57, 519-526.

Chen B. & Liu H. (2010). Relationships between phytoplankton growth and cell size in surface oceans: Interactive effects of temperature, nutrients, and grazing. *Limnology and oceanography*, 55, 965.

Chesson P. (2000). Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, 343-366.

Chust G., Irigoien X., Chave J. & Harris R.P. (2013). Latitudinal phytoplankton distribution and the neutral theory of biodiversity. *Global Ecology and Biogeography*, 22, 531-543.

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

Clayton S., Dutkiewicz S., Jahn O. & Follows M.J. (2013). Dispersal, eddies, and the diversity of marine phytoplankton. *Limnology & Oceanography: Fluids & Environments*, 3, 182-197.

Cloern J.E. (2001). Our evolving conceptual model of the coastal eutrophication problem. *Marine Ecology Progress Series*, 210, 223-253.

Cloern J.E. & Jassby A.D. (1995). Year-to-year fluctuation of the spring phytoplankton bloom in south San Francisco Bay: an example of ecological variability at the land-sea interface. In: *Ecological time series*. Springer, pp. 139-149.

Coll M., Piroddi C., Steenbeek J., Kaschner K., Ben Rais Lasram F., Aguzzi J., Ballesteros E., Bianchi C.N., Corbera J., Dailianis T., Danovaro R., Estrada M., Froglia C., Galil B.S., Gasol J.M., Gertwagen R., Gil J., Guilhaumon F., Kesner-Reyes K., Kitsos M.-S., Koukouras A., Lampadariou N., Laxamana E., López-Fé de la Cuadra C.M., Lotze H.K., Martin D., Mouillot D., Oro D., Raicevich S., Rius-Barile J., Saiz-Salinas J.I., San Vicente C., Somot S., Templado J., Turon X., Vafidis D., Villanueva R. & Voultsiadou E. (2010). The Biodiversity of the Mediterranean Sea: Estimates, Patterns, and Threats. *Plos One*, 5, e11842.

Conley D.J., Paerl H.W., Howarth R.W., Boesch D.F., Seitzinger S.P., Havens K.E., Lancelot C. & Likens G.E. (2009). Controlling eutrophication: nitrogen and phosphorus. *science*, 323, 1014-1015.

Connell J.H. (1978). Diversity in tropical rain forests and coral reefs. *Science*, 199, 1302-1310.

Conway J. (1997). Evaluating ERS-1 SAR data for the discrimination of tropical forest from other tropical vegetation types in Papua New Guinea. *International Journal of Remote Sensing*, 18, 2967-2984.

Cotté C., d'Ovidio F., Dragon A., Guinet C. & Levy M. (2013). Top predator preference in a complex environment: Elephant seals target distinct mesoscale features of the Antarctic Circumpolar Current, Progress Ocean., in revision, Dec 2013.

Cotté C., Park Y.-H., Guinet C. & Bost C.-A. (2007). Movements of foraging king penguins through marine mesoscale eddies. *Proceedings of the Royal Society B: Biological Sciences*, 274, 2385-2391.

Currie D.J., Mittelbach G.G., Cornell H.V., Field R., Guégan J.F., Hawkins B.A., Kaufman D.M., Kerr J.T., Oberdorff T. & O'Brien E. (2004). Predictions and tests of climate-based

hypotheses of broad-scale variation in taxonomic richness. *Ecology letters*, 7, 1121-1134.
Cury P., Bakun A., Crawford R.J., Jarre A., Quiñones R.A., Shannon L.J. & Verheye H.M. (2000). Small pelagics in upwelling systems: patterns of interaction and structural changes in "wasp-waist" ecosystems. *ICES Journal of Marine Science: Journal du Conseil*, 57, 603-618.
Cushing D. (1990). Plankton production and year-class strength in fish populations: an update of the match/mismatch hypothesis. *Advances in marine biology*, 26, 249-293.
D'ortenzio F., Antoine D., Martinez E. & Ribera d'Alcalà M. (2012). Phenological changes of oceanic phytoplankton in the 1980s and 2000s as revealed by remotely sensed ocean-color observations. *Global Biogeochemical Cycles*, 26.
D'ortenzio F. & Ribera d'Alcalà M. (2009). On the trophic regimes of the Mediterranean Sea: a satellite analysis. *Biogeosciences*, 6, 139-148.
d'Ovidio F., Fernández V., Hernández-García E. & López C. (2004). Mixing structures in the Mediterranean Sea from finite-size Lyapunov exponents. *Geophysical Research Letters*, 31.
d'Ovidio F., De Monte S., Alvain S., Dandonneau Y. & Lévy M. (2010). Fluid dynamical niches of phytoplankton types. *Proceedings of the National Academy of Sciences*, 107, 18366-18370.
Darwin C. (1859). On the origins of species by means of natural selection. *London: Murray*.
Dayton P. & Hessler R. (1972). Role of biological disturbance in maintaining diversity in the deep sea. In: *Deep Sea Research and Oceanographic Abstracts*. Elsevier, pp. 199-208.
De Monte S., Soccodato A., Alvain S. & d'Ovidio F. (2013). Can we detect oceanic biodiversity hotspots from space&quest. *The ISME journal*, 7, 2054-2056.
Dodson S.I., Arnott S.E. & Cottingham K.L. (2000). The relationship in lake communities between primary productivity and species richness. *Ecology*, 81, 2662-2679.
Doney S.C., Ruckelshaus M., Duffy J.E., Barry J.P., Chan F., English C.A., Galindo H.M., Grebmeier J.M., Hollowed A.B. & Knowlton N. (2012). Climate change impacts on marine ecosystems. *Marine Science*, 4.
Duarte C.M., Agustí S., Gasol J.M., Vaqué D. & Vazquez-Dominguez E. (2000). Effect of nutrient supply on the biomass structure of planktonic communities: an experimental test on a Mediterranean coastal community. *Marine Ecology-Progress Series*, 206, 87-95.
Dubelaar G. & Jonker R. (2000). Flow cytometry as a tool for the study of phytoplankton. *Scientia Marina*, 64, 135-156.
Ducet N., Le Traon P.-Y. & Reverdin G. (2000). Global high-resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and-2. *Journal of Geophysical Research: Oceans (1978–2012)*, 105, 19477-19498.
Duffy J.E. (2003). Biodiversity loss, trophic skew and ecosystem functioning. *Ecology letters*, 6, 680-687.
Duffy J.E., Cardinale B.J., France K.E., McIntyre P.B., Thébault E. & Loreau M. (2007). The functional role of biodiversity in ecosystems: incorporating trophic complexity. *Ecology letters*, 10, 522-538.
Dutkiewicz S., Follows M. & Bragg J.G. (2009). Modeling the coupling of ocean ecology and biogeochemistry. *Global Biogeochemical Cycles*, 23.
Dybern B. & Hansen H. (1989). Baltic Sea patchiness experiment PEX'86. *ICES Coop. Res. Rep*, 163, 1-100.
Dzwonkowski B., Lipphardt Jr B., Kohut J.T., Yan X.-H. & Garvine R.W. (2010). Synoptic

measurements of episodic offshore flow events in the central mid-Atlantic Bight. *Continental Shelf Research*, 30, 1373-1386.
Elmgren R. & Larsson U. (2001). Nitrogen and the Baltic Sea: managing nitrogen in relation to phosphorus. *The Scientific World Journal*, 1, 371-377.
Etnoyer P., Canny D., Mate B.R., Morgan L.E., Ortega-Ortiz J.G. & Nichols W.J. (2006). Sea-surface temperature gradients across blue whale and sea turtle foraging trajectories off the Baja California Peninsula, Mexico. *Deep Sea Research Part II: Topical Studies in Oceanography*, 53, 340-358.
Fager E.W. & McGowan J.A. (1963). Zooplankton Species Groups in the North Pacific Co-occurrences of species can be used to derive groups whose members react similarly to water-mass types. *science*, 140, 453-460.
Falkowski P.G., Barber R.T. & Smetacek V. (1998). Biogeochemical controls and feedbacks on ocean primary production. *science*, 281, 200-206.
Fauchald P. (2009). Spatial interaction between seabirds and prey: review and synthesis. *Marine Ecology Progress Series*, 391, 139-151.
Ferrari R. & Wunsch C. (2010). The distribution of eddy kinetic and potential energies in the global ocean. *Tellus A*, 62, 92-108.
Field C.B., Behrenfeld M.J., Randerson J.T. & Falkowski P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *science*, 281, 237-240.
Fisher R.A., Corbet A.S. & Williams C.B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *The Journal of Animal Ecology*, 42-58.
Flynn D.F., Mirotchnick N., Jain M., Palmer M.I. & Naeem S. (2011). Functional and phylogenetic diversity as predictors of biodiversity-ecosystem-function relationships. *Ecology*, 92, 1573-1581.
Follows M.J. & Dutkiewicz S. (2011). Modeling diverse communities of marine microbes. *Annual review of marine science*, 3, 427-451.
Follows M.J., Dutkiewicz S., Grant S. & Chisholm S.W. (2007). Emergent biogeography of microbial communities in a model ocean. *science*, 315, 1843-1846.
Franks P.J. (1992). Sink or swim: Accumulation of biomass at fronts. *Marine ecology progress series. Oldendorf*, 82, 1-12.
Fuhrman J.A., Hewson I., Schwalbach M.S., Steele J.A., Brown M.V. & Naeem S. (2006). Annually reoccurring bacterial communities are predictable from ocean conditions. *Proceedings of the National Academy of Sciences*, 103, 13104-13109.
Fuhrman J.A., Steele J.A., Hewson I., Schwalbach M.S., Brown M.V., Green J.L. & Brown J.H. (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences*, 105, 7774-7778.
Game E.T., Grantham H.S., Hobday A.J., Pressey R.L., Lombard A.T., Beckley L.E., Gjerde K., Bustamante R., Possingham H.P. & Richardson A.J. (2009). Pelagic protected areas: the missing dimension in ocean conservation. *Trends in Ecology & Evolution*, 24, 360-369.
Garcés E., Masó M. & Camp J. (1999). A recurrent and localized dinoflagellate bloom in a Mediterranean beach. *Journal of Plankton Research*, 21, 2373-2391.
García-Reyes M., Largier J.L. & Sydeman W.J. (2014). Synoptic-scale upwelling indices and predictions of phyto-and zooplankton populations. *Progress in Oceanography*, 120, 177-188.
Garçon V.C., Oschlies A., Doney S.C., McGillicuddy D. & Waniek J. (2001). The role of

mesoscale variability on plankton dynamics in the North Atlantic. *Deep Sea Research Part II: Topical Studies in Oceanography*, 48, 2199-2226.
Garnier E., Lavorel S., Ansquer P., Castro H., Cruz P., Dolezal J., Eriksson O., Fortunel C., Freitas H. & Golodets C. (2007). Assessing the effects of land-use change on plant traits, communities and ecosystem functioning in grasslands: a standardized methodology and lessons from an application to 11 European sites. *Annals of botany*, 99, 967-985.
Gaston K.J. (1996a). Biodiversity-congruence. *Progress in physical geography*, 20, 105-112.
Gaston K.J. (1996b). Species richness: measure and measurement. *Biodiversity: a biology of numbers and difference*, 77-113.
Gaston K.J. (2000). Global patterns in biodiversity. *Nature*, 405, 220-227.
Gentemann C.L., Donlon C.J., Stuart-Menteth A. & Wentz F.J. (2003). Diurnal signals in satellite sea surface temperature measurements. *Geophysical Research Letters*, 30.
Gillman L.N. & Wright S.D. (2006). The influence of productivity on the species richness of plants: a critical assessment. *Ecology*, 87, 1234-1243.
Giovannoni S., Temperton B. & Zhao Y. (2013). Giovannoni et al. reply. *Nature*, 499, E4-E5.
Godø O.R., Samuelsen A., Macaulay G.J., Patel R., Hjøllo S.S., Horne J., Kaartvedt S. & Johannessen J.A. (2012). Mesoscale eddies are oases for higher trophic marine life. *Plos One*, 7, e30161.
Gordon H.R. & Morel A.Y. (1983). *Remote assessment of ocean color for interpretation of satellite visible imagery: A review*. American Geophysical Union.
Gotelli N.J. & Colwell R.K. (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology letters*, 4, 379-391.
Graham W.M., Pagès F. & Hamner W.M. (2001). A physical context for gelatinous zooplankton aggregations: a review. *Hydrobiologia*, 451, 199-212.
Grantham H.S., Game E.T., Lombard A.T., Hobday A.J., Richardson A.J., Beckley L.E., Pressey R.L., Huggett J.A., Coetzee J.C. & van der Lingen C.D. (2011). Accommodating dynamic oceanographic processes and pelagic biodiversity in marine conservation planning. *Plos One*, 6, e16552.
Green J.L., Bohannan B.J.M. & Whitaker R.J. (2008). Microbial Biogeography: From Taxonomy to Traits. *Science*, 320, 1039-1043.
Griffin J.N., De La Haye K.L., Hawkins S.J., Thompson R.C. & Jenkins S.R. (2008). Predator diversity and ecosystem functioning: density modifies the effect of resource partitioning. *Ecology*, 89, 298-305.
Guasto J.S., Rusconi R. & Stocker R. (2012). Fluid mechanics of planktonic microorganisms. *Annual Review of Fluid Mechanics*, 44, 373-400.
Haegeman B., Hamelin J., Moriarty J., Neal P., Dushoff J. & Weitz J.S. (2013). Robust estimation of microbial diversity in theory and in practice. *The ISME journal*, 7, 1092-1101.
Haller G. & Yuan G. (2000). Lagrangian coherent structures and mixing in two-dimensional turbulence. *Physica D: Nonlinear Phenomena*, 147, 352-370.
Halpern B.S., Walbridge S., Selkoe K.A., Kappel C.V., Micheli F., D'Agrosa C., Bruno J.F., Casey K.S., Ebert C. & Fox H.E. (2008). A global map of human impact on marine ecosystems. *science*, 319, 948-952.
Hamilton A.J. (2005). Species diversity or biodiversity? *Journal of Environmental Management*, 75, 89-92.
Hansen M.C., Stehman S.V. & Potapov P.V. (2010). Quantification of global gross forest cover loss. *Proceedings of the National Academy of Sciences*, 107, 8650-8655.

Hardin G. (1960). The competitive exclusion principle. *science*, 131, 1292-1297.
Hardin G. (1968). The tragedy of the commons. *science*, 162, 1243-1248.
Hardman-Mountford N.J., Hirata T., Richardson K.A. & Aiken J. (2008). An objective methodology for the classification of ecological pattern into biomes and provinces for the pelagic ocean. *Remote Sensing of Environment*, 112, 3341-3352.
Hazen E.L., Jorgensen S., Rykaczewski R.R., Bograd S.J., Foley D.G., Jonsen I.D., Shaffer S.A., Dunne J.P., Costa D.P. & Crowder L.B. (2013). Predicted habitat shifts of Pacific top predators in a changing climate. *Nature Climate Change*, 3, 234-238.
Heino J. & Soininen J. (2007). Are higher taxa adequate surrogates for species-level assemblage patterns and species richness in stream organisms? *Biological conservation*, 137, 78-89.
Hellweger F.L., van Sebille E. & Fredrick N.D. (2014). Biogeographic patterns in ocean microbes emerge in a neutral agent-based model. *science*, 345, 1346-1349.
Hernández-Garcıa E. & López C. (2004). Sustained plankton blooms under open chaotic flows. *Ecological Complexity*, 1, 253-259.
Hillebrandt H. (2004). On the generality of the latitudinal diversity gradient. *American Naturalist*, 163, 192-211.
Hingamp P., Grimsley N., Acinas S.G., Clerissi C., Subirana L., Poulain J., Ferrera I., Sarmento H., Villar E. & Lima-Mendez G. (2013). Exploring nucleo-cytoplasmic large DNA viruses in Tara Oceans microbial metagenomes. *The ISME journal*, 7, 1678-1695.
Hirata T., Hardman-Mountford N., Brewin R., Aiken J., Barlow R., Suzuki K., Isada T., Howell E., Hashioka T. & Noguchi-Aita M. (2011). Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences*, 8, 311-327.
Hobday A. & Hartmann K. (2006). Near real-time spatial management based on habitat predictions for a longline bycatch species. *Fisheries Management and Ecology*, 13, 365-380.
Hooker S.K., Cañadas A., Hyrenbach K.D., Corrigan C., Polovina J.J. & Reeves R.R. (2011). Making protected area networks effective for marine top predators. *Endangered Species Research*, 13, 203-218.
Hooper D., Chapin Iii F., Ewel J., Hector A., Inchausti P., Lavorel S., Lawton J., Lodge D., Loreau M. & Naeem S. (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological monographs*, 75, 3-35.
Hooper D.U., Adair E.C., Cardinale B.J., Byrnes J.E., Hungate B.A., Matulich K.L., Gonzalez A., Duffy J.E., Gamfeldt L. & O'Connor M.I. (2012). A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature*, 486, 105-108.
Horne E. & Platt T. (1984). The dominant space and time scales of variability in the physical and biological fields on continental shelves. *Rapp. PV. Reun. Cons. Int. Explor. Mer*, 183, 8-19.
Huisman J., Johansson A.M., Folmer E.O. & Weissing F.J. (2001). Towards a solution of the plankton paradox: the importance of physiology and life history. *Ecology Letters*, 4, 408-411.
Hunt Jr G. & Schneider D. (1987). Scale-dependent processes in the physical and biological environment of marine birds. *Seabirds: feeding ecology and role in marine ecosystems*, 7-41.
Huston M. (1979). A general hypothesis of species diversity. *American naturalist*, 81-101.
Hutchinson G. (1958). Concluding remarks. Populations studies: Animal ecology and

demography. Cold Spring Harbourd. In: *Symposia on Quantitative Biology*, pp. 415-427.
Hutchinson G. & Macartur R. (1959). A theoretical ecological model of size distributions among species of animals. *American naturalist*, 117-125.
Hutchinson G.E. (1959). Homage to Santa Rosalia or why are there so many kinds of animals? *American naturalist*, 145-159.
Hutchinson G.E. (1961). The paradox of the plankton. *American naturalist*, 137-145.
Hyrenbach K.D., Forney K.A. & Dayton P.K. (2000). Marine protected areas and ocean basin management. *Aquatic conservation: marine and freshwater ecosystems*, 10, 437-458.
IOCCG (2014). Phytoplankton Functional Types from Space. Sathyendranath, S. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 15, IOCCG, Dartmouth, Canada. In.
Irigoien X., Huisman J. & Harris R.P. (2004). Global biodiversity patterns of marine phytoplankton and zooplankton. *Nature*, 429, 863-867.
Ives A.R. & May R.M. (1985). Competition within and between species in a patchy environment: relations between microscopic and macroscopic models. *Journal of Theoretical Biology*, 115, 65-92.
Jeffrey S. & Vesk M. (1997). *Introduction to marine phytoplankton and their pigment signatures*. na.
Jenkins C.N., Pimm S.L. & Joppa L.N. (2013). Global patterns of terrestrial vertebrate diversity and conservation. *Proceedings of the National Academy of Sciences*, 110, E2602-E2610.
Jetz W., Kreft H., Ceballos G. & Mutke J. (2009). Global associations between terrestrial producer and vertebrate consumer diversity. *Proceedings of the Royal Society B: Biological Sciences*, 276, 269-278.
Jones S.E. (2007). The marine microbial world: a voyage of discovery. *ISME J*, 1, 469-470.
Joseph B. & Legras B. (2002). Relation between kinematic boundaries, stirring, and barriers for the Antarctic polar vortex. *Journal of the atmospheric sciences*, 59, 1198-1212.
Kai E.T., Rossi V., Sudre J., Weimerskirch H., Lopez C., Hernandez-Garcia E., Marsac F. & Garçon V. (2009). Top marine predators track Lagrangian coherent structures. *Proceedings of the National Academy of Sciences*, 106, 8245-8250.
Kark S. & Van Rensburg B.J. (2006). Ecotones: marginal or central areas of transition? *Israel Journal of Ecology & Evolution*, 52, 29-53.
Karsenti E., Acinas S.G., Bork P., Bowler C., De Vargas C., Raes J., Sullivan M., Arendt D., Benzoni F., Claverie J.-M., Follows M., Gorsky G., Hingamp P., Iudicone D., Jaillon O., Kandels-Lewis S., Krzic U., Not F., Ogata H., Pesant S., Reynaud E.G., Sardet C., Sieracki M.E., Speich S., Velayoudon D., Weissenbach J., Wincker P. & the Tara Oceans C. (2011). A Holistic Approach to Marine Eco-Systems Biology. *PLoS Biol*, 9, e1001177.
Kaschner K., Tittensor D.P., Ready J., Gerrodette T. & Worm B. (2011). Current and Future Patterns of Global Marine Mammal Biodiversity. *PLoS ONE*, 6, e19653.
Kaschner K., Watson R., Trites A. & Pauly D. (2006). Mapping world-wide distributions of marine mammal species using a relative environmental suitability (RES) model. *Marine Ecology Progress Series*, 316, 285-310.
Kaschner K.J., Rius-Barile K., Kesner-Reyes C., Garilao S.O.K., Rees T. & R. F. (2013). AquaMaps: Predicted range maps for aquatic species. World wide web electronic publication, www.aquamaps.org, Version 08/2013.
Kembel S.W., Wu M., Eisen J.A. & Green J.L. (2012). Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. *PLoS*

computational biology, 8, e1002743.
Kerr J.T. & Packer L. (1997). Habitat heterogeneity as a determinant of mammal species richness in high-energy regions. *Nature*, 385, 252-254.
Kerswell A.P. (2006). Global biodiversity patterns of benthic marine algae. *Ecology*, 87, 2479-2488.
Koh L.P., Dunn R.R., Sodhi N.S., Colwell R.K., Proctor H.C. & Smith V.S. (2004). Species coextinctions and the biodiversity crisis. *science*, 305, 1632-1634.
Komárek J. (2006). Cyanobacterial taxonomy: current problems and prospects for the integration of traditional and molecular approaches. *Algae*, 21, 349-375.
Konstantinidis K.T., Ramette A. & Tiedje J.M. (2006). The bacterial species definition in the genomic era. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361, 1929-1940.
Kostadinov T., Siegel D. & Maritorena S. (2010). Global variability of phytoplankton functional types from space: assessment via the particle size distribution. *Biogeosciences Discussions*, 7, 4295-4340.
Kostadinov T.S., Siegel D.A. & Maritorena S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *J. Geophys. Res.*, 114, C09015.
Labry C., Herbland A. & Delmas D. (2002). The role of phosphorus on planktonic production of the Gironde plume waters in the Bay of Biscay. *Journal of Plankton Research*, 24, 97-117.
Lankau R.A. & Strauss S.Y. (2007). Mutual feedbacks maintain both genetic and species diversity in a plant community. *science*, 317, 1561-1563.
Lapeyre G. (2002). Characterization of finite-time Lyapunov exponents and vectors in two-dimensional turbulence. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 12, 688-698.
Lapeyre G. & Klein P. (2006). Dynamics of the upper oceanic layers in terms of surface quasigeostrophy theory. *Journal of physical oceanography*, 36, 165-176.
Larson R.J., Lenarz W.H. & Ralston S. (1994). The distribution of pelagic juvenile rockfish of the genus Sebastes in the upwelling region off central California. *Calif. Coop. Ocean. Fish. Investig. Rep*, 35, 175-221.
Le Fèvre J. & Frontier S. (1988). Influence of temporal characteristics of physical phenomena on plankton dynamics, as shown by North-West European marine ecosystems. In: *Toward a Theory on Biological-Physical Interactions in the World Ocean*. Springer, pp. 245-272.
Lehahn Y., d'Ovidio F., Lévy M. & Heifetz E. (2007). Stirring of the northeast Atlantic spring bloom: A Lagrangian analysis based on multisatellite data. *Journal of Geophysical Research: Oceans (1978–2012)*, 112.
Leibold M.A., Chase J.M., Shurin J.B. & Downing A.L. (1997). Species turnover and the regulation of trophic structure. *Annual review of Ecology and Systematics*, 467-494.
Leith H. & Whittaker R. (1975). Primary production of the biosphere. *Ecological Studies, Springer-Verlag, Berlin*, 14.
Lennon J.T. & Jones S.E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology*, 9, 119-130.
Levin I., Bösinger R., Bonani G., Francey R.J., Kromer B., Münnich K., Suter M., Trivett N.B. & Wölfli W. (1992). Radiocarbon in atmospheric carbon dioxide and methane: global distribution and trends. In: *Radiocarbon after four decades*. Springer, pp. 503-518.
Levin I. & Hesshaimer V. (2000). Radiocarbon-a unique tracer of global carbon cycle dynamics.

Levin S.A. & Paine R.T. (1974). Disturbance, patch formation, and community structure. *Proceedings of the National Academy of Sciences*, 71, 2744-2747.
Levy H., Schwarzkopf M.D., Horowitz L., Ramaswamy V. & Findell K. (2008). Strong sensitivity of late 21st century climate to projected changes in short-lived air pollutants. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 113.
Lévy M., Klein P. & Treguier A.-M. (2001). Impact of sub-mesoscale physics on production and subduction of phytoplankton in an oligotrophic regime. *Journal of Marine Research*, 59, 535-565.
Levy O., Ball B.A., Bond-Lamberty B., Cheruvelil K.S., Finley A.O., Lottig N.R., Punyasena S.W., Xiao J., Zhou J. & Buckley L.B. (2014). Approaches to advance scientific understanding of macrosystems ecology. *Frontiers in Ecology and the Environment*, 12, 15-23.
Li W. (1997). Cytometric diversity in marine ultraphytoplankton. *Limnology and oceanography*, 42, 874-880.
Lindenschmidt K.-E. & Chorus I. (1998). The effect of water column mixing on phytoplankton succession, diversity and similarity. *Journal of Plankton Research*, 20, 1927-1951.
Logares R., Audic S., Santini S., Pernice M.C., de Vargas C. & Massana R. (2012). Diversity patterns and activity of uncultured marine heterotrophic flagellates unveiled with pyrosequencing. *The ISME journal*, 6, 1823-1833.
Longhurst A.R. (2010). *Ecological geography of the sea*. Academic Press.
López C., Neufeld Z., Hernández-García E. & Haynes P.H. (2001). Chaotic advection of reacting substances: Plankton dynamics on a meandering jet. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 26, 313-317.
Louzao M., Pinaud D., Péron C., Delord K., Wiegand T. & Weimerskirch H. (2011). Conserving pelagic habitats: seascape modelling of an oceanic top predator. *Journal of applied ecology*, 48, 121-132.
Lundry M.R., Barber R.T., Bid, are R.R., Chai F., Coale K.H., Dam H.G., Lewis M.R., Lindley S.T., McCarthy J.J. & Roman M.R. (1997). Iron and grazing constraints on primary production in the central equatorial Pacific: an EqPac synthesis. *OCEANOGRAPHY*, 42.
MacArthur R.H. (1957). On the relative abundance of bird species. *Proceedings of the National Academy of Sciences of the United States of America*, 43, 293.
MacIntyre H.L., Geider R.J. & McKay R.M. (1996). Photosyntesis and regulation of RUBISCO activity in net phytoplankton from delaware bay. *Journal of phycology*, 32, 718-731.
Magurran A.E. (2004). Measuring biological diversity.
Magurran A.E. & McGill B.J. (2011). Challenges and opportunities in the measurement and assessment of biological diversity. *Biological Diversity: Frontiers in measurement and assessment. Oxford University Press Oxford*, 1-7.
Marañón E., Holligan P.M., Varela M., Mouriño B. & Bale A.J. (2000). Basin-scale variability of phytoplankton biomass, production and growth in the Atlantic Ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 47, 825-857.
Marshall J., Adcroft A., Hill C., Perelman L. & Heisey C. (1997a). A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research: Oceans (1978–2012)*, 102, 5753-5766.
Marshall J., Hill C., Perelman L. & Adcroft A. (1997b). Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *Journal of Geophysical Research: Oceans (1978–2012)*, 102, 5733-5752.

Martiny A.C., Pham C.T., Primeau F.W., Vrugt J.A., Moore J.K., Levin S.A. & Lomas M.W. (2013). Strong latitudinal patterns in the elemental ratios of marine plankton and organic matter. *Nature Geoscience*, 6, 279-283.
Massana R. & Pedrós-Alió C. (2008). Unveiling new microbial eukaryotes in the surface ocean. *Current opinion in microbiology*, 11, 213-218.
McCullagh P.N. & Nelder F. (1989). JA (1989) Generalized linear models. *Monographs on Statistics and Applied Probability*, 37.
McGill B.J. (2003). A test of the unified neutral theory of biodiversity. *Nature*, 422, 881-885.
McGill B.J., Enquist B.J., Weiher E. & Westoby M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, 21, 178-185.
McGill B.J., Etienne R.S., Gray J.S., Alonso D., Anderson M.J., Benecha H.K., Dornelas M., Enquist B.J., Green J.L. & He F. (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology letters*, 10, 995-1015.
McGrady-Steed J., Harris P.M. & Morin P.J. (1997). Biodiversity regulates ecosystem predictability. *Nature*, 390, 162-165.
McManus M.A. & Woodson C.B. (2012). Plankton distribution and ocean dispersal. *The Journal of experimental biology*, 215, 1008-1016.
McWilliams J.C. (1985). A uniformly valid model spanning the regimes of geostrophic and isotropic, stratified turbulence: Balanced turbulence. *Journal of the atmospheric sciences*, 42, 1773-1774.
Menkes C.E., Kennan S.C., Flament P., Dandonneau Y., Masson S., Biessy B., Marchal E., Eldin G., Grelet J. & Montel Y. (2002). A whirling ecosystem in the equatorial Atlantic. *Geophysical Research Letters*, 29, 48-1-48-4.
Menna M., Poulain P.-M., Zodiatis G. & Gertman I. (2012). On the surface circulation of the Levantine sub-basin derived from Lagrangian drifters and satellite altimetry data. *Deep Sea Research Part I: Oceanographic Research Papers*, 65, 46-58.
Mittelbach G.G., Steiner C.F., Scheiner S.M., Gross K.L., Reynolds H.L., Waide R.B., Willig M.R., Dodson S.I. & Gough L. (2001). What is the observed relationship between species richness and productivity? *Ecology*, 82, 2381-2396.
Monteiro F., Follows M. & Dutkiewicz S. (2010). Distribution of diverse nitrogen fixers in the global ocean. *Global Biogeochemical Cycles*, 24.
Mooney H. (2002). The debate on the role of biodiversity in ecosystem functioning. *Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford*, 12-17.
Mora C., Tittensor D.P., Adl S., Simpson A.G. & Worm B. (2011). How many species are there on Earth and in the ocean? *PLoS biology*, 9, e1001127.
Moritz C. & Cicero C. (2004). DNA barcoding: promise and pitfalls. *PLoS biology*, 2, e354.
Motomura I. (1932). On the statistical treatment of communities. *Zool. Mag*, 44, 379-383.
Mouquet N. & Loreau M. (2003). Community patterns in source-sink metacommunities. *The American Naturalist*, 162, 544-557.
Myers N. (1990). The biodiversity challenge: expanded hot-spots analysis. *Environmentalist*, 10, 243-256.
Myers N., Mittermeier R.A., Mittermeier C.G., Da Fonseca G.A. & Kent J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403, 853-858.
Myers R.A. & Worm B. (2003). Rapid worldwide depletion of predatory fish communities.

Nature, 423, 280-283.
Naeem S. & Li S. (1997). Biodiversity enhances ecosystem reliability. *Nature*, 390, 507-509.
Nair A., Sathyendranath S., Platt T., Morales J., Stuart V., Forget M.-H., Devred E. & Bouman H. (2008). Remote sensing of phytoplankton functional types. *Remote Sensing of Environment*, 112, 3366-3375.
Neufeld Z., Hernández-García E. & Hernández E. (2010). A Dynamical Systems Approach.
Nieto K., Demarcq H. & McClatchie S. (2012). Mesoscale frontal structures in the Canary Upwelling System: New front and filament detection algorithms applied to spatial and temporal patterns. *Remote Sensing of Environment*, 123, 339-346.
Not F., Valentin K., Romari K., Lovejoy C., Massana R., Töbe K., Vaulot D. & Medlin L.K. (2007). Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. *science*, 315, 253-255.
Nur N., Jahncke J., Herzog M.P., Howar J., Hyrenbach K.D., Zamon J.E., Ainley D.G., Wiens J.A., Morgan K. & Ballance L.T. (2011). Where the wild things are: predicting hotspots of seabird aggregations in the California Current System. *Ecological Applications*, 21, 2241-2257.
Olascoaga M., Beron-Vera F., Brand L. & Koçak H. (2008). Tracing the early development of harmful algal blooms on the West Florida Shelf with the aid of Lagrangian coherent structures. *Journal of Geophysical Research: Oceans (1978–2012)*, 113.
Olson D.B. & Hood R.R. (1994). Modelling pelagic biogeography. *Progress in Oceanography*, 34, 161-205.
Olson R.R. & Olson M.H. (1989). Food limitation of planktotrophic marine invertebrate larvae: does it control recruitment success? *Annual review of Ecology and Systematics*, 225-247.
Owen R.W. (1981). Fronts and eddies in the sea: mechanisms, interactions and biological effects. *Analysis of marine ecosystems*, 197-233.
Paine R. (1974). Intertidal community structure. *Oecologia*, 15, 93-120.
Paine R.T. (1966). Food web complexity and species diversity. *American naturalist*, 65-75.
Palaczet A., St John M., Brewin R., Hirata T. & Gregg W. (2013). Distribution of phytoplankton functional types in high-nitrate low-chlorophyll waters in a new diagnostic ecological indicator model. *Biogeosciences Discussions*, 10, 8103-8157.
Pannard A., Bormans M. & Lagadeuc Y. (2008). Phytoplankton species turnover controlled by physical forcing at different time scales. *Canadian Journal of Fisheries and Aquatic Sciences*, 65, 47-60.
Pascual A., Ruiz S., Jerome B., Faugere Y., Buongiorno Nardelli B., Larnicol G., Le Traon P.-Y. & Tintore J. (2010). On the combination of satellite and in-situ observations to better understand (sub) mesoscale dynamics. In: *38th COSPAR Scientific Assembly*, p. 237.
Pasquero C., Provenzale A. & Babiano A. (2001). Parameterization of dispersion in two-dimensional turbulence. *Journal of fluid mechanics*, 439, 279-303.
Pérez-Muñuzuri V. & Huhn F. (2010). The role of mesoscale eddies time and length scales on phytoplankton production. *Nonlinear Processes in Geophysics*, 17, 177-186.
Peterson A., Soberón J. & Sánchez-Cordero V. (1999). Conservatism of ecological niches in evolutionary time. *science*, 285, 1265-1267.
Piganeau G., Eyre-Walker A., Grimsley N. & Moreau H. (2011). How and why DNA barcodes underestimate the diversity of microbial eukaryotes. *PLoS One*, 6, e16342.
Platt T. (1972). Local phytoplankton abundance and turbulence. In: *Deep Sea Research and Oceanographic Abstracts*. Elsevier, pp. 183-187.

Pollard R.T., Venables H.J., Read J.F. & Allen J.T. (2007). Large-scale circulation around the Crozet Plateau controls an annual phytoplankton bloom in the Crozet Basin. *Deep Sea Research Part II: Topical Studies in Oceanography*, 54, 1915-1929.
Porter S.C. (1972). Distribution, morphology, and size frequency of cinder cones on Mauna Kea volcano, Hawaii. *Geological Society of America Bulletin*, 83, 3607-3612.
Prairie J.C., Sutherland K.R., Nickols K.J. & Kaltenberg A.M. (2012). Biophysical interactions in the plankton: A cross-scale review. *Limnology & Oceanography: Fluids & Environments*, 2, 121-145.
Preston F.W. (1948). The commonness, and rarity, of species. *Ecology*, 29, 254-283.
Ptacnik R., Solimini A.G., Andersen T., Tamminen T., Brettum P., Lepistö L., Willén E. & Rekolainen S. (2008). Diversity predicts stability and resource use efficiency in natural phytoplankton communities. *Proceedings of the National Academy of Sciences*, 105, 5134-5138.
Pueyo S. (2006). Self-similarity in species–area relationship and in species abundance distribution. *Oikos*, 112, 156-162.
Purkis S.J. & Klemas V.V. (2011). *Remote sensing and global environmental change*. John Wiley & Sons.
Purvis A. & Hector A. (2000). Getting the measure of biodiversity. *Nature*, 405, 212-219.
Raes J., Letunic I., Yamada T., Jensen L.J. & Bork P. (2011). Toward molecular trait-based ecology through integration of biogeochemical, geographical and metagenomic data. *Mol Syst Biol*, 7, 473.
Raitsos E.D., Lavender J.S., Maravelias D.C., Haralabous J., Richardson J.A. & Reid C.P. (2008). Identifying Four Puhytoplankton Functional Types from Space: An Ecological Approach. *Limnology and Oceanography*, 53, 605-613.
Ready J., Kaschner K., South A.B., Eastwood P.D., Rees T., Rius J., Agbayani E., Kullander S. & Froese R. (2010). Predicting the distributions of marine organisms at the global scale. *Ecological modelling*, 221, 467-478.
Ribalet F., Marchetti A., Hubbard K.A., Brown K., Durkin C.A., Morales R., Robert M., Swalwell J.E., Tortell P.D. & Armbrust E.V. (2010). Unveiling a phytoplankton hotspot at a narrow boundary between coastal and offshore waters. *Proceedings of the National Academy of Sciences*, 107, 16571-16576.
Ribalet F., Schruth D.M. & Armbrust E.V. (2011). flowPhyto: enabling automated analysis of microscopic algae from continuous flow cytometric data. *Bioinformatics*, 27, 732-733.
Robinson I.S. (2004). *Measuring the oceans from space: the principles and methods of satellite oceanography*. Springer.
Rocchini D., Balkenhol N., Carter G.A., Foody G.M., Gillespie T.W., He K.S., Kark S., Levin N., Lucas K., Luoto M., Nagendra H., Oldeland J., Ricotta C., Southworth J. & Neteler M. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecological Informatics*, 5, 318-329.
Rohde K. (1992). Latitudinal gradients in species-diversity: the search for the primary cause. *Oikos*, 65, 514–527.
Rombouts I., Beaugrand G., Ibaňez F., Gasparini S., Chiba S. & Legendre L. (2009). Global latitudinal variations in marine copepod diversity and environmental factors. *Proceedings of the Royal Society B: Biological Sciences*, rspb20090742.
Rose G.A. & Leggett W.C. (1990). The importance of scale to predator-prey spatial correlations: an example of Atlantic fishes. *Ecology*, 33-43.

Rossi V., López C., Sudre J., Hernández-García E. & Garçon V. (2008). Comparative study of mixing and biological activity of the Benguela and Canary upwelling systems. *Geophysical Research Letters*, 35.
Roy S., Llewellyn C.A., Egeland E.S. & Johnsen G. (2011). *Phytoplankton pigments: characterization, chemotaxonomy and applications in oceanography*. Cambridge University Press.
Rutherford S., D'Hondt S. & Prell W. (1999). Environmental controls on the geographic distribution of zooplankton diversity. *Nature*, 400, 749-753.
Rynearson T.A. & Armbrust E.V. (2000). DNA fingerprinting reveals extensive genetic diversity in a field population of the centric diatom Ditylum brightwellii. *Limnology and oceanography*, 45, 1329-1340.
Sabarros P.S., Grémillet D., Demarcq H., Moseley C., Pichegru L., Mullers R.H., Stenseth N.C. & Machu E. (2013). Fine-scale recognition and use of mesoscale fronts by foraging Cape gannets in the Benguela upwelling region. *Deep Sea Research Part II: Topical Studies in Oceanography*.
Sadlo F. & Peikert R. (2007). Efficient visualization of Lagrangian coherent structures by filtered AMR ridge extraction. *Visualization and Computer Graphics, IEEE Transactions on*, 13, 1456-1463.
Samuelsson K., Berglund J., Haecky P. & Andersson A. (2002). Structural changes in an aquatic microbial food web caused by inorganic nutrient addition. *Aquatic Microbial Ecology*, 29, 29-38.
Sauterey B., Ward B., Follows M., Bowler C. & Claessen D. (2014). When everything is not everywhere but species evolve: an alternative method to model adaptive properties of marine ecosystems. *Journal of Plankton Research*, Accepted for publication.
Scales K.L., Miller P.I., Hawkes L.A., Ingram S.N., Sims D.W. & Votier S.C. (2014). On the Front Line: frontal zones as priority at-sea conservation areas for mobile marine vertebrates. *Journal of applied ecology*.
Selig E.R., Turner W.R., Troëng S., Wallace B.P., Halpern B.S., Kaschner K., Lascelles B.G., Carpenter K.E. & Mittermeier R.A. (2014). Global Priorities for Marine Biodiversity Conservation. *Plos One*, 9, e82898.
Send U., Font J., Krahmann G., Millot C., Rhein M. & Tintoré J. (1999). Recent advances in observing the physical oceanography of the western Mediterranean Sea. *Progress in Oceanography*, 44, 37-64.
Seuront L., Schmitt F. & Lagadeuc Y. (2001). Turbulence intermittency, small-scale phytoplankton patchiness and encounter rates in plankton: where do we go from here? *Deep Sea Research Part I: Oceanographic Research Papers*, 48, 1199-1215.
Shannon C.E. & Weaver W. (1949). The mathematical theory of communication (Urbana, IL. In. University of Illinois Press IL.
Shannon K. & Weaver W. (1948). A mathematical theory of communication. *Bell System Tehn. J*, 3, 623-637.
Simon C., Frati F., Beckenbach A., Crespi B., Liu H. & Flook P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the entomological Society of America*, 87, 651-701.
Simpson E.H. (1949). Measurement of diversity. *Nature*.
Smayda T.J. & Reynolds C.S. (2001). Community assembly in marine phytoplankton: application

of recent models to harmful dinoflagellate blooms. *Journal of Plankton Research*, 23, 447-461.
Smith V.H. (2003). Eutrophication of freshwater and coastal marine ecosystems a global problem. *Environmental Science and Pollution Research*, 10, 126-139.
Soininen J. & Luoto M. (2014). Predictability in species distributions: a global analysis across organisms and ecosystems. *Global Ecology and Biogeography*, 23, 1264-1274.
Solé R.V., Alonso D. & McKane A. (2000). Scaling in a network model of a multispecies ecosystem. *Physica A: Statistical Mechanics and its Applications*, 286, 337-344.
Solé R.V., Manrubia S.C., Benton M., Kauffman S. & Bak P. (1999). Criticality and scaling in evolutionary ecology. *Trends in Ecology & Evolution*, 14, 156-160.
Sournia A. (1994). Pelagic biogeography and fronts. *Progress in Oceanography*, 34, 109-120.
Sousa W.P. (1984). The role of disturbance in natural communities. *Annual review of Ecology and Systematics*, 353-391.
Spear L.B., Ballance L.T. & Ainley D.G. (2001). Response of seabirds to thermal boundaries in the tropical Pacific: the thermocline versus the Equatorial Front. *Marine Ecology Progress Series*, 219, 275-289.
Stammer D. (1997). Global characteristics of ocean variability estimated from regional TOPEX/POSEIDON altimeter measurements. *Journal of physical oceanography*, 27, 1743-1769.
Stock C.A., Alexander M.A., Bond N.A., Brander K.M., Cheung W.W., Curchitser E.N., Delworth T.L., Dunne J.P., Griffies S.M. & Haltuch M.A. (2011). On the use of IPCC-class models to assess the impact of climate on living marine resources. *Progress in Oceanography*, 88, 1-27.
Strong D.R. (1983). Natural variability and the manifold mechanisms of ecological communities. *American naturalist*, 636-660.
Sugihara G. (1980). Minimal community structure: an explanation of species abundance patterns. *American naturalist*, 770-787.
Sund P.N., Blackburn M. & Williams F. (1981). Tunas and their environment in the Pacific Ocean: a review. *Oceanogr. Mar. Biol. Ann. Rev*, 19, 443-512.
Sverdrup H. (1953). On conditions for the vernal blooming of phytoplankton. *Journal du Conseil*, 18, 287-295.
Swalwell J.E., Ribalet F. & Armbrust E.V. (2011). SeaFlow: A novel underway flow-cytometer for continuous observations of phytoplankton in the ocean. *Limnol. Oceanogr. Methods*, 9, 466-477.
Swearer S.E., Caselle J.E., Lea D.W. & Warner R.R. (1999). Larval retention and recruitment in an island population of a coral-reef fish. *Nature*, 402, 799-802.
Sydeman W.J., Bradley R.W., Warzybok P., Abraham C.L., Jahncke J., Hyrenbach K.D., Kousky V., Hipfner J.M. & Ohman M.D. (2006). Planktivorous auklet Ptychoramphus aleuticus responses to ocean climate, 2005: Unusual atmospheric blocking? *Geophysical Research Letters*, 33.
Tett P. & Barton E. (1995). Why are there about 5000 species of phytoplankton in the sea? *Journal of Plankton Research*, 17, 1693-1704.
Tews J., Brose U., Grimm V., Tielbörger K., Wichmann M., Schwager M. & Jeltsch F. (2004). Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. *Journal of Biogeography*, 31, 79-92.
Tilman D. (1994). Competition and biodiversity in spatially structured habitats. *Ecology*, 75, 2-

16.
Tilman D. (2000). Causes, consequences and ethics of biodiversity. *Nature*, 405, 208-211.
Tilman D., Knops J., Wedin D., Reich P., Ritchie M. & Siemann E. (1997). The influence of functional diversity and composition on ecosystem processes. *science*, 277, 1300-1302.
Tittensor D.P., Mora C., Jetz W., Lotze H.K., Ricard D., Berghe E.V. & Worm B. (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, 466, 1098-1101.
Tyrrell T. (1999). The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature*, 400, 525-531.
Uitz J., Claustre H., Morel A. & Hooker S.B. (2006). Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *J. Geophys. Res.*, 111, C08005.
Vallina S.M., Follows M., Dutkiewicz S., Montoya J.M., Cermeno P. & Loreau M. (2014). Global relationship between phytoplankton diversity and productivity in the ocean. *Nature communications*, 5.
Vane-Wright R.I., Humphries C.J. & Williams P.H. (1991). What to protect?—Systematics and the agony of choice. *Biological conservation*, 55, 235-254.
Vaulot D. & Marie D. (1999). Diel variability of photosynthetic picoplankton in the equatorial Pacific. *Journal of Geophysical Research: Oceans (1978–2012)*, 104, 3297-3310.
Venter J.C., Remington K., Heidelberg J.F., Halpern A.L., Rusch D., Eisen J.A., Wu D., Paulsen I., Nelson K.E. & Nelson W. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *science*, 304, 66-74.
Volpe G., Santoleri R., Vellucci V., Ribera d'Alcala M., Marullo S. & d'Ortenzio F. (2007). The colour of the Mediterranean Sea: Global versus regional bio-optical algorithms evaluation and implication for satellite chlorophyll estimates. *Remote Sensing of Environment*, 107, 625-638.
Whittaker R.J. & Heegaard E. (2003). What is the observed relationship between species richness and productivity? Comment. *Ecology*, 84, 3384-3390.
Williams R.G., Wilson C. & Hughes C.W. (2007). Ocean and atmosphere storm tracks: The role of eddy vorticity forcing. *Journal of physical oceanography*, 37, 2267-2289.
Wilson E.O. (1992). The diversity oflife. *Belknap, Cambridge, Massachusetts*.
Winnepenninckx B., Backeljau T. & De Wachter R. (1993). Complete small ribosomal subunit RNA sequence of the chiton Acanthopleura japonica (Lischke, 1873)(Mollusca, Polyplacophora). *Nucleic acids research*, 21, 1670.
Wood L.J., Fish L., Laughren J. & Pauly D. (2008). Assessing progress towards global marine protection targets: shortfalls in information and action. *Oryx*, 42, 340-351.
Worm B., Barbier E.B., Beaumont N., Duffy J.E., Folke C., Halpern B.S., Jackson J.B., Lotze H.K., Micheli F. & Palumbi S.R. (2006). Impacts of biodiversity loss on ocean ecosystem services. *science*, 314, 787-790.
Worm B., Lotze H.K. & Myers R.A. (2003). Predator diversity hotspots in the blue ocean. *Proceedings of the National Academy of Sciences*, 100, 9884-9888.
Worm B., Sandow M., Oschlies A., Lotze H.K. & Myers R.A. (2005). Global Patterns of Predator Diversity in the Open Oceans. *Science*, 309, 1365-1369.
Wunsch G. & Schreiber H. (1992). *Stochastische Systeme*. Springer-Verlag.
Yasuhara M., Hunt G., Dowsett H.J., Robinson M.M. & Stoll D.K. (2012). Latitudinal species diversity gradient of marine zooplankton for the last three million years. *Ecology letters*, 15, 1174-1179.

Young J.N., Goldman J.A., Kranz S.A., Tortell P.D. & Morel F.M. (2014). Slow carboxylation of Rubisco constrains the rate of carbon fixation during Antarctic phytoplankton blooms. *New Phytologist*.

Zacharias M.A. & Roff J.C. (2001). Use of focal species in marine conservation and management: a review and critique. *Aquatic conservation: marine and freshwater ecosystems*, 11, 59-76.

Zehr J.P. (2011). Nitrogen fixation by marine cyanobacteria. *Trends in microbiology*, 19, 162-173.

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

ANNEX 1

The material presented in this work is subject of peer reviewed publications.

The content of the entire Chapter 3 is under submission for the following research article in Global Ecology and Biogeography: **Soccodato A**., De Monte S., Follows M., Levy M., d'Ovidio F. Identification of global biodiversity patterns of planktonic communities by spatially-resolved observation of dominant functional types.

The content of Chapter 4 is partially included in the following communication: De Monte S., **Soccodato A**., Alvain S., d'Ovidio F., (2013). Can we detect oceanic biodiversity hotspots from space? *ISME J* 7:2054-6 . My contribution to this paper is in the comparison of the satellite diversity index with extended in situ data and with known macroecological patterns related to diversity, such as temperature gradients. The content of the same Chapter relative to temporal variability of the satellite-based diversity and its relation with productivity in the ocean will be part of a future publication.

The content of Chapter 5 is part of a following research article under preparation: **Soccodato A**., De Monte S., Levy M., d'Ovidio F. Seascape mosaic of plankton dominance reveals global biodiversity hotspots across the trophic chain.