



Leibniz Centre for Tropical Marine Research

Jacobs University Bremen

PhD Proposal

Modeling phytoplankton community changes in the Cariaco Basin, Venezuela

Author:
Benjamin POST

Dissertation Committee:
Prof. Dr. Agostino MERICO
Prof Dr. Marc-Thorsten HÜTT
Dr. Esteban ACEVEDO-TREJOS
Prof Dr. Andrew D. BARTON

August 16, 2019

Abstract

Driven by anthropogenic influences, the environmental conditions on our planet are changing at an unprecedented rate and so are the communities and organisms inhabiting it. As oceanic primary producers, phytoplankton plays a central role in the global biogeochemical cycles and are a key component of computational models exploring the biogeochemistry of the ocean.

The field of phytoplankton modeling has seen great advances in the last 20 years, in particular the “trait-based” approach promises ecologically meaningful descriptions of biodiversity by moving away from treating species explicitly, but instead looking at the way organisms interact with the environment (i.e. their traits)

Current modeling approaches are moving towards an explicit treatment of biodiversity to explore the effects on ecosystem function. The CARIACO time-series, located in the Cariaco Basin off the coast of Venezuela, provides long-term observations of the biogeochemistry of a tropical coastal ecosystem including measurements of phytoplankton taxonomic and functional diversity. Over 20 years of data showed clear trends of warming and a marked shift in the phytoplankton community. Yet, there have been no comprehensive ecosystem modeling studies performed using this data. The project aims to further our understanding of phytoplankton communities by using computational ecosystem modeling applied to the CARIACO time-series data. In the first study a phytoplankton functional-type (PFT) model will be used to investigate aggregate biomass changes between contrasting environmental regimes. I have developed a modeling framework in the programming language Python. The development of this modeling framework as an open-source package will form the basis of a technical publication.

This flexible modeling framework will then be used to investigate the relationships between biodiversity and ecosystem function (BDEF) with an explicit treatment of biodiversity and testing the model mechanism against the relationship observed in the CARIACO time-series data.

Contents

1 General Introduction	2
1.1 The ocean, phytoplankton and why it matters	2
1.2 Characterizing phytoplankton communities	5
1.3 Modeling phytoplankton communities	8
1.4 The Cariaco Basin and the CARIACO time-series	11
1.5 Aims of the proposed PhD project	17
2 Understanding phytoplankton community shifts in the eastern Cariaco basin	19
2.1 Regime shift in the Cariaco Basin	19
2.2 Methods	26
2.3 Preliminary Results	32
3 PhytoMFTM - a flexible object-oriented PFT model	34
3.1 Python ecosystem model package development	34
4 Further work	37
4.1 BDEF in Cariaco	37
4.2 Time Table	39
Bibliography	41

1. General Introduction

1.1 The ocean, phytoplankton and why it matters

The complexity of the ocean and its vast ecosystems has fascinated scientists to this day and most likely will continue to do so far into the future. Myriad life forms are embedded in a matrix so far removed from our mostly dry existence on top the earth's crust. In the ocean, life moves in dilution, and the equivalents of forests and grasslands are hard to spot unless the concentration of tiny phytoplankton is so large, that deep blue turns into a milky green.

The term phytoplankton refers to microscopic marine photosynthetic organisms. These microorganisms form the basis of the oceanic food web and are primary producers of planetary scale, contributing roughly half of the oxygen in our atmosphere through photosynthesis (Field et al. 1998). Phytoplankton consists of mostly single-celled organisms, prokaryotes and eukaryotes from a highly diverse evolutionary background (Falkowski et al. 2004). This large phylogenetic diversity is accompanied by a remarkable range of survival strategies, biogeochemical roles, shapes and sizes within the polyphyletic phytoplankton (see Figure 1.1 for a size comparison). The emergence of such a large variety of organisms and the mechanisms sustaining their persistence has been one of the key questions in phytoplankton ecology over the last 50 years.

Most phytoplankton species rely on a few nutrients, co₂ and light for growth Liebig's law, nutrient competition in sea water they are relatively well mixed and in direct competition!

In the liquid medium of sea water it was expected that competitive species would eas-

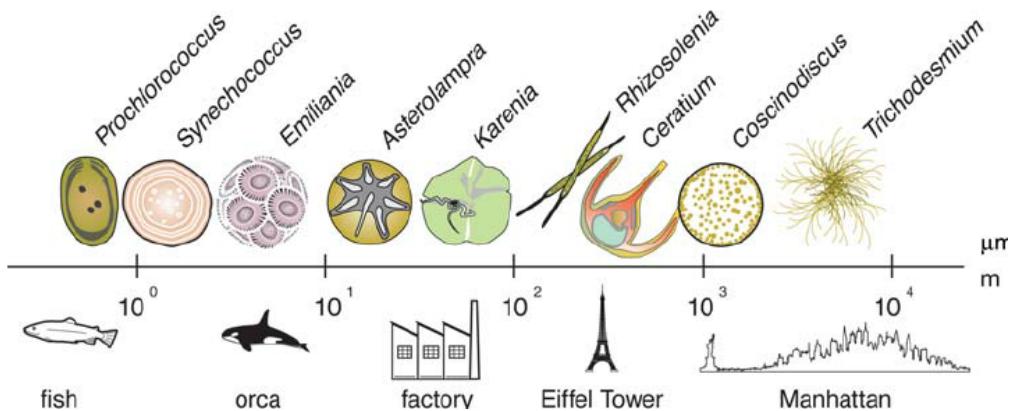


Figure 1.1: "A comparison of the size range (maximum linear dimension) of phytoplankton relative to macroscopic objects." from Finkel et al. (2010)

ily replace others through competitive exclusion, however this clearly was not the case (Hutchinson 1961). Hypotheses explaining the paradox are: - grazing/predation/viruse - nonequilibrium conditions/ time scales - mutualistic interactions - competitive ability (multiple nuts, different affinities) - emergent chaos

In fact it is these forces that make the environment for phytoplankton quite variable!

The distribution of phytoplankton is driven by the complex physical forces that govern ocean currents and the chemistry of the bodies of water they move. The key components are macronutrients (e.g. nitrogen & phosphorus) and micronutrients (e.g. iron & cobalt) welling up from the deeper ocean or flushed in from continental sources. Wherever there are sufficient nutrients available within the euphotic zone, from the surface to the depth where photosynthetically available radiation (PAR) is 1% of the surface value, planktonic life begins to thrive. Planktonic life is further bound by the mixed layer depth (MLD), which describes the depth of the relatively homogenous surface layer of the ocean maintained by wind stress and surface heat fluxes.

If the MLD is above the euphotic depth most phytoplankton species can spend their life actively photosynthesizing, capturing carbon dioxide and nutrients to grow and divide.

The biomass produced by phytoplankton is mostly consumed by higher trophic lev-

els and either assimilated or excreted. -feeding herbivorous zooplankton, carnivorous zooplankton, larvae, fish, mammals, fisheries and humans!

Phytoplankton growth indirectly feeds a considerable part of earth's population through fisheries (Stock et al. 2017)

Another large portion of the phytoplankton experiences natural mortality and viral lysis. Microbial degradation drives remineralization within the euphotic zone, which fuels regenerated production (Eppley & Peterson 1979)

A small fraction of biomass sinks out of the photic layer as fecal or detrital matter to the deeper ocean and an even smaller fraction reaches the sea floor as sediment (roughly 1 %) and remains there over geological times (Honjo et al. 2008). This process has been termed the biological carbon pump.

Carbon sequestered this way is removed from the ocean-atmosphere system for potentially millions of years. Given the projected rise of atmospheric CO₂ levels, it is of grave importance to understand how changes in the phytoplankton community at the surface, driven by anthropogenic stressors and climate change, will affect the carbon burial potential of oceanic ecosystems.

Ecosystems along continental margins provide a particularly productive habitat, with only 10% of total ocean surface area covered by continental margins, but 10-15% of marine primary production and more than 40% of carbon export to the seabed occurring along coastal lines (Yool & Fasham 2001; Muller-Karger et al. 2005).

even shapes the elemental composition of oceanic water itself (Redfield 1958).

Studies have both reported a global declining trend in marine primary production (Boyce et al. 2012) and increasing trends in long-term ocean time series (Chavez et al. 2010). In order to answer questions of how phytoplankton will respond to a changing climate it is necessary to look the diverse phytoplankton community in greater detail.

1.2 Characterizing phytoplankton communities

From the early days of oceanographic research, scientists have been interested in the microscopic organisms that were floating in samples of sea water. These communities contain many species each and in total there are tens of thousands of species of phytoplankton that inhabit the surface ocean (De Vargas et al. 2015). All phytoplankton species use chlorophyll or bacteriochlorophyll to harvest light as the energy source to fix organic carbon, but there is wide variation in virtually all their other properties (Litchman & Klausmeier 2008). In addition to the complex community composition, there are many factors affecting measurements of their bulk properties in the ocean, such as the viral and bacterial community and the influence of diverse grazers, all within the complex three-dimensional physical environment that is the ocean. Where earlier phytoplankton ecologists focused on identifying individual species, decoding their phylogeny or growing them in controlled lab cultures, recent research is trying to integrate the insights gained from these approaches and quantify the diversity on higher levels of organization in relation to other properties of the ecosystem. The focus has shifted towards trait diversity both within and across species and within and across phytoplankton groups. In order to scientifically describe this perplexing diversity the concepts of trait-based ecology and functional types have been developed (Tilman 2001; McGill et al. 2006; Violle et al. 2007).

Functional types, traits and diversity

In the following I will try to clarify the complementary terms of phytoplankton traits, functional traits, and functional types.

The trait-based approach to phytoplankton ecology has been growing in popularity. Part of the fascination evoked by this term stems from its origin in evolutionarily theory. Over the last three decades, it has been adopted by ecologists trying to understand communities and ecosystems. In this new context, the concept of traits has been stretched far beyond its original meaning, which can lead to some confusion surrounding the scope of trait-based methods (Violle et al. 2007). In the simplest definition, a trait is a surrogate of organismal performance. In the ecological context this has been expanded to surrogates for the performance of populations, communities and entire

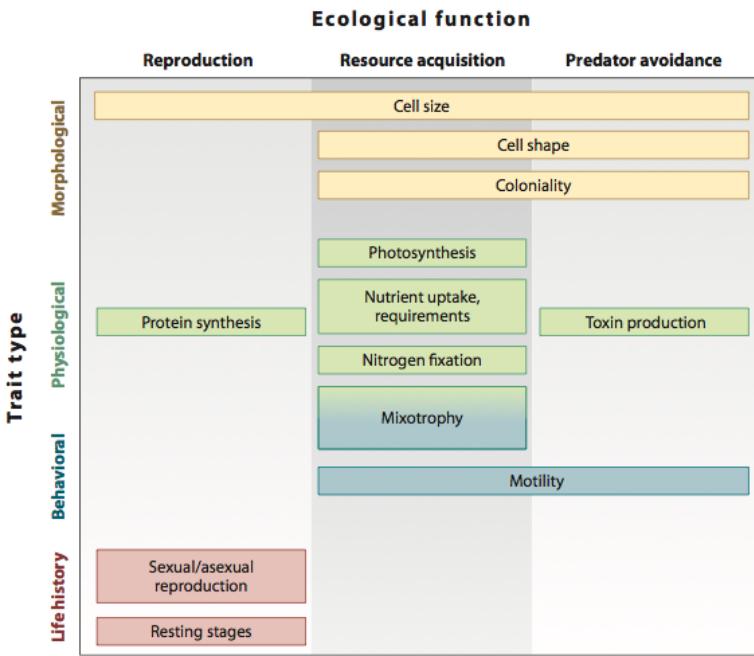


Figure 1.2: "A typology of phytoplankton functional traits" from Litchman & Klausmeier (2008)

ecosystems. This can include ecophysiological traits, life-history traits, demographic traits or response and effect traits of ecosystems (see Figure 1.2 for a selection of phytoplankton traits). Theoretically, any property of an organism or ecosystem could be defined as a trait, but ideally a trait should be functional. Functional traits are defined by Violette et al. (2007) as "morpho-, physio- or phenological traits which impact fitness indirectly via their effects on growth, reproduction and survival". An important facet of the trait-based approach is to describe organismal function via trade-offs between traits. For example when competing for multiple nutrients, phytoplankton species are thought to be constrained by trade-offs in their competitive ability for one over another resource (Tilman 1990).

Phytoplankton are extremely diverse and the trait based approach lends itself to generalizations, as traits and ecological trade-offs can be defined and explored irrespective of species or taxa boundaries (McGill et al. 2006). However, depending on the study and hypotheses to be tested, it can be very helpful to structure the diversity of organisms into distinct groups. Major taxonomic groups of phytoplankton can

be classified based on their ecological or biogeochemical roles within the ecosystem (Iglesias-Rodríguez et al. 2002; Flynn et al. 2015). The concept of functional groups is not in contrast to a trait-based perspective of phytoplankton ecology, but can be complementary to it. By broadly sampling relevant traits across phytoplankton groups and species, functional types can be defined by functional traits and trade-offs and therefore extend the trait-based approach by another level of organization (Litchman et al. 2007). An early example is the work of Ramón Margalef. Margalef used observations of important traits, such as sinking rates and nutrient utilization to build the concept called "Margalef's mandala" to organize phytoplankton functional types (PFTs) on a spectrum of nutrient availability versus turbulence (Margalef 1978).

The terms functional group and functional type are used interchangeably, with functional groups more often referring to the grouping of species and the functional type describing the group as a whole, often as implemented in computational models. In fact, the simplification of the phytoplankton community into functional types has been widely used for the design and interpretation of computational models that try to recreate or make predictions about the biogeochemical cycling, biogeographic distribution, productivity and other ecosystem functions of phytoplankton (Gregg et al. 2003; Le Quéré et al. 2005). Biogeochemically defined functional types are most often used, as these functional traits can usually be well defined within an ecosystem model. Typical examples of such functional groups are silicifiers, which broadly correspond the phylogenetic group of diatoms, and calcifiers, which are usually represented by coccolithophores. Such functional types are always simplifications of the natural phytoplankton diversity. Silicoflagellates create silicified skeletons like diatoms, but are often not explicitly included because they rarely dominate modern phytoplankton assemblages. The choice of which functional groups to include in a model can also be driven by biogeography or analytical considerations concerning the measurement instrumentation used for a particular study (Irwin & Finkel 2017).

In biodiversity research, the trait-based approach has been readily adopted. It used to be that species diversity (i.e. the number of species) was the most important metric, but now it is functional diversity, which can be described by the variance in the value of a functional trait of the community or ecosystem. It is important to keep in mind that functional types are often composed of many species with a possibly large

variance in trait values. Recent research is trying to understand the effects of diversity within functional types and within species (Violle et al. 2012, 2017; Des Roches et al. 2018).

1.3 Modeling phytoplankton communities

Given the complexity of the ocean ecosystem, it is necessary to aggregate our knowledge of the many smaller parts into comprehensive ecological models in order to test mechanistic hypotheses and investigate their full-scale implications. Computational models of phytoplankton growth have been developed since the 1940s and have greatly increased in sophistication and complexity since then, co-evolving with the rise in computational resources (Gentleman 2002). Phytoplankton modeling started with formulations based on the Lotka-Volterra equations of predator-prey dynamics (Fleming 1939). From these relatively simple descriptions models evolved to describe the oceanic physical environment and the ecosystem it contains including multiple trophic levels. The nutrient-phytoplankton-zooplankton (NPZ) and nutrient-phytoplankton-zooplankton-detritus (NPZD) models, originally developed by John Steele with a model ocean split in two layers, succeeded in reproducing the typical annual bloom dynamics observed in the temperate ocean (Steele 1958; Evans 1988; Fasham et al. 1990). Further developments have been in more exact physiological descriptions of phytoplankton based in cellular metabolism and energy allocation (Geider et al. 1997) and both simple and more complicated ecosystem formulations driven by local and global 3D circulation models (Lacroix et al. 2007; Hirata et al. 2013).

However, in their simplified approach, these models unavoidably limit the characterization of a diverse phytoplankton community (Bruggeman 2009). These plankton ecosystem models are typically highly aggregated, such that a single variable determines the response of a diverse assemblage of phytoplankton species (Franks 2009). Implementing a meaningful treatment of biodiversity in ecological models is a key challenge in the field of phytoplankton modeling (Queirós et al. 2015). The most apparent way of implementing this within the framework of established NPZD models is to include multiple equations and state variables for different phytoplankton functional

types (Le Quéré et al. 2005). For every group that fulfills a distinct ecosystem function, a new set of parameters has to be added, which complicates the model structure and increases computational costs. This somewhat intuitive approach, however, does lead to problems. First and foremost, there is a lack and inherent uncertainty of data from field and culture experiments to constrain functional types, which leads to the difficulty of validating the model output in light of insufficient information, leading multiple authors to criticize the PFT modeling approach as attempting to "run before we can walk" particularly when used for extrapolating into the future (Anderson 2005; Shimoda & Arhonditsis 2016).

The current scientific discussion can seem intimidating to an early career scientist, as both the most obvious future directions of ecosystem model design as exemplified by PFT models, as well as very basic assumptions, such as using Monod nutrient uptake kinetics in phytoplankton models, have come under scrutiny (Flynn 2010; Smith et al. 2014; Hellweger 2017).

Nevertheless, there are also examples of modeling approaches that show a way forward. To name an alternative to the modeling paradigm discussed so far, there is individual based modeling (IBM). In IBM the phytoplankton are explicitly represented as individual agents, allowing for a diverse and spatially interactive phytoplankton community (Hellweger & Bucci 2009). However, the computational cost and structural complexity of this approach makes both model construction and interpretation difficult.

Another approach is to extend traditional NPZD models via moment-based estimation of aggregate properties (Mericó et al. 2009). A specific implementation of this is the PhytoSFDM model developed by Acevedo-Trejos et al. (2016). Instead of modeling multiple size-classes of phytoplankton explicitly, the community is described not only by the biomass, but also by the mean size and size variance. Size is used as a master trait, with size variance being used as a proxy for functional diversity. Trade-offs related to nutrient uptake, grazing and sinking structure the phytoplankton community along the size spectrum as driven by the physical forcing. The model was used to investigate latitudinal diversity gradients in the Atlantic Ocean (Acevedo-Trejos et al. 2018). This approach allows a relatively simple model structure and is computationally efficient, but

the PhytoSFDM model structure currently only allows for a single phytoplankton community with a log-normal distribution in cell size. This lends the PhytoSFDM model more to studies of large scale processes and biogeographic patterns rather than to local ecosystem modeling studies, where a more fine-scaled resolution of the phytoplankton community might be required.

A major advancement in the field of phytoplankton modeling was the DARWIN model developed at MIT (Follows et al. 2007). In the general framework of the DARWIN model, large numbers of phytoplankton types are initialized with equal biomass but with different parameters for the most important traits, namely those related to light harvesting, temperature dependence, and nutrient acquisition. These parameters are chosen stochastically from broad ranges of values, based on laboratory and field data, and constrained by simplified allometric functions describing ecological trade-offs. Different functional types are prescribed in the model via varying nutrient utilization traits (e.g. small phytoplankton that cannot assimilate nitrate as Prochlorococcus analogs). Over multi-annual runs this community self-assembles through ecological competition and physical changes produced by the simulated environment of a global circulation model (GCM). In the random initialization of phytoplankton types, the DARWIN model allows for the emergence and development of diverse phytoplankton communities. This approach of modeling biodiversity has been termed “selection-based” (Follows & Dutkiewicz 2011). The model framework is continually modified and expanded, for example for exploring the effect of grazing formulations (Prowe et al. 2012), the biogeography of phytoplankton traits (Barton et al. 2013) and the influence of ocean acidification at a global scale (Dutkiewicz et al. 2015). A study of particular interest is the size-structured food-web model component developed by Ward et al. (2012), because the modeling approach combines the trait-based approach of scaling parameters allometrically along cell size with a PFT modeling approach. Each functional type is assigned a different allometry based on the size ranges and relationships taken from data. Together with the selection-based biodiversity representation that the DARWIN framework provides, this seems to be a promising direction for future models. The model has however only been applied and compared to data at a global scale and the code of this specific implementation is not publicly available.

The field of phytoplankton modeling provides diverse approaches and no consensus

has been reached as to which computational tools and model structures are the most promising. In describing an ecosystem in computational models, a balance needs to be found between complexity and simplicity. The natural ecosystems are invariably very complex, however models that capture too much of the detail are difficult to interpret, and overtly simplistic models risk not capturing the essential dynamics present. Another important aspect of model development, no matter where it lies on the scale of complexity, is model validation using ecological data (Rykiel Jr 1996). In fact model complexity might generally have to be adapted to the scientific question at hand, which often emerges from patterns observed in ecological data.

1.4 The Cariaco Basin and the CARIACO time-series

At the beginning of my PhD I set out to find publicly available ocean time-series data that could be used as the basis for modeling work about phytoplankton diversity. Initially the plan was to choose multiple locations, to compare results from model applications in contrasting environments. Not surprisingly the search did quickly yield results, among the most prominent: the Bermuda Atlantic Time-series Study (BATS) and the Hawaii Ocean Time-Series (HOTS). Quickly the issues of public ocean time-series data became apparent. Links often lead to defunct sites and servers were sporadically maintained. In particular for my application the problem was that the basic physical parameters and bulk properties such as total chlorophyll were readily available, but more specific phytoplankton functional type or taxonomy data was harder to find, if not missing. In particular the type of phytoplankton data that was available differed widely between the stations and would have not allowed for a straightforward comparison. It was only later in my search that I came across the CARIACO time series, an acronym for "CCarbon Retention In A Colored Ocean", located in the Cariaco Basin off the coast of Venezuela. The data is available through the University of South Florida (USF) at <http://imars.marine.usf.edu/cariaco> and includes a wealth of data that was collected since 1995. Most importantly to my purposes, the data included both detailed phytoplankton pigment measurements and taxonomic data of the phytoplankton community at monthly intervals. It was the first ocean time-series with such detailed public phytoplankton data and soon I decided to focus my work

on the Cariaco Basin. Previous work specifically using the phytoplankton dataset has to my knowledge been limited to statistical modeling and analysis (Taylor et al. 2012; Mutshinda et al. 2013a,b; Pinckney et al. 2015; Irwin et al. 2015). Apart from an early study looking at carbon fluxes, where the only resolved functional type in the ecosystem model were diatoms (Walsh et al. 2002), there has been no study using a mechanistic ecosystem model applied to the CARIACO time-series, allowing for novel applications and insights from my PhD thesis.

The CARIACO time-series station has historically been a site of scientific interest and served as a natural laboratory in biogeochemical research for more than 60 years (Muller-Karger et al. 2019). It has played important roles in understanding the nutrient stoichiometry of the ocean (Redfield 1963). The CARIACO Ocean Time-Series program was established in 1995 off the coast of Venezuela ($10^{\circ} 30' N$, $64^{\circ} 40' W$, see Figure 1.3). Located in the south-eastern Caribbean Sea, the Cariaco Basin is a 160 km long and 70 km wide tectonic depression, reaching up to 1400 m in depth. The two deeper parts of the basin are separated by a saddle of 900 m depth, with the time-series mooring located in the eastern part. The entire basin is bound to the west and north by a shallow ridge at 100 m depth, restricting the exchange of deep water with the Caribbean Sea. The restricted circulation and high productivity at the surface resulted in anoxic conditions below 250 m depth within the basin (Richards & Vaccaro 1956). The Cariaco Basin is the worlds largest truly marine anoxic basin (Wakeham et al. 2012). The hydrography at the surface is influenced by Guyana and North Equatorial currents that flow into the Caribbean Sea from a south-eastern direction, but this exchange is restricted to the two channels above the 100 m ridge. Observed and modeled horizontal surface water velocities within the basin are relatively weak, indicating a minimal influence of horizontal transport at the mooring site (Alvera-Azcárate et al. 2009).

The time-series data

The program was established as a joint-project of the Venezuelan Fondo Nacional de Ciencia, Tecnología e Investigación (FONACIT) and the US National Science Foundation (NSF), with the particular interest in creating a time-series of surface ocean biogeochemistry that could be linked to satellite observations and the sedimentation

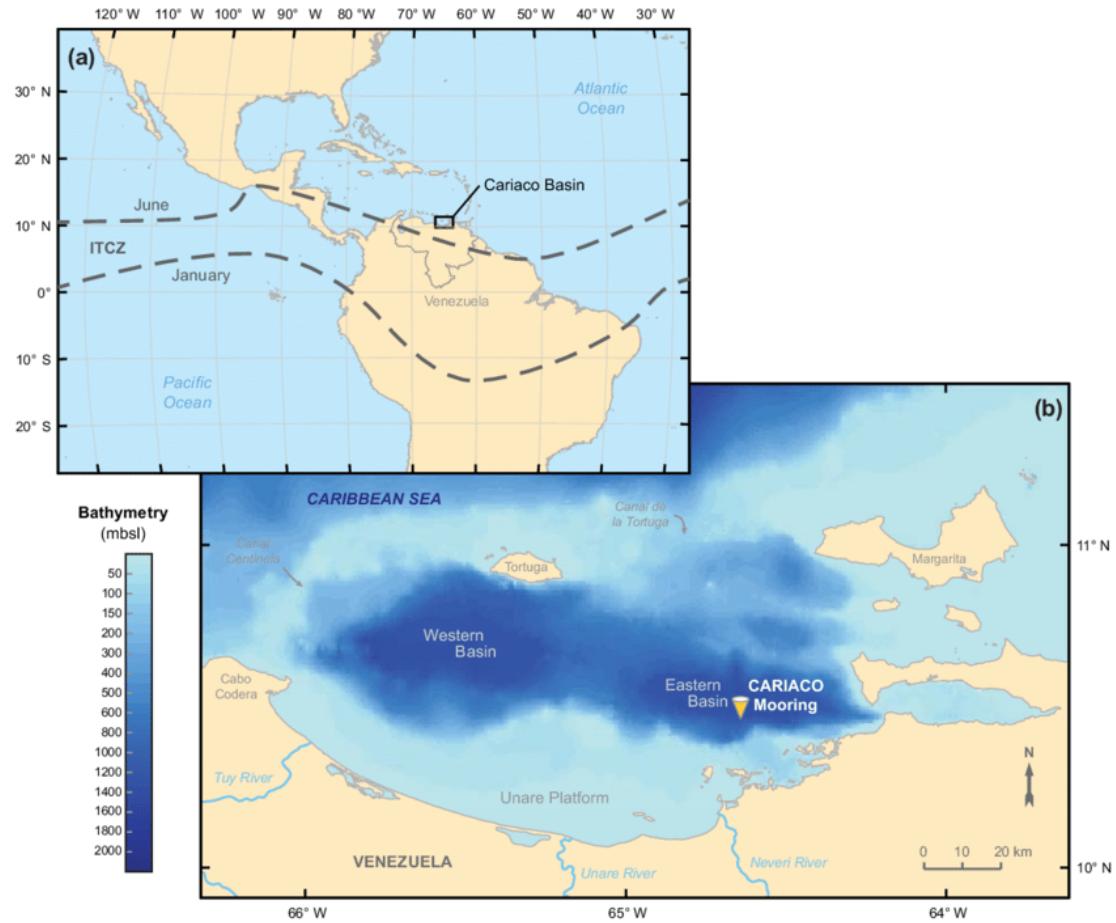


Figure 1.3: "Study area. A. Location of the Cariaco Basin off the Venezuelan coast in the southern Caribbean Sea, with January and June positions of the Intertropical Convergence Zone (ITCZ). B. Location of the CARIACO station in the eastern sub-basin, general bathymetry and local rivers emptying in the basin (bathymetric data from GEBCO_08 Grid)" from Bringué et al. (2019)

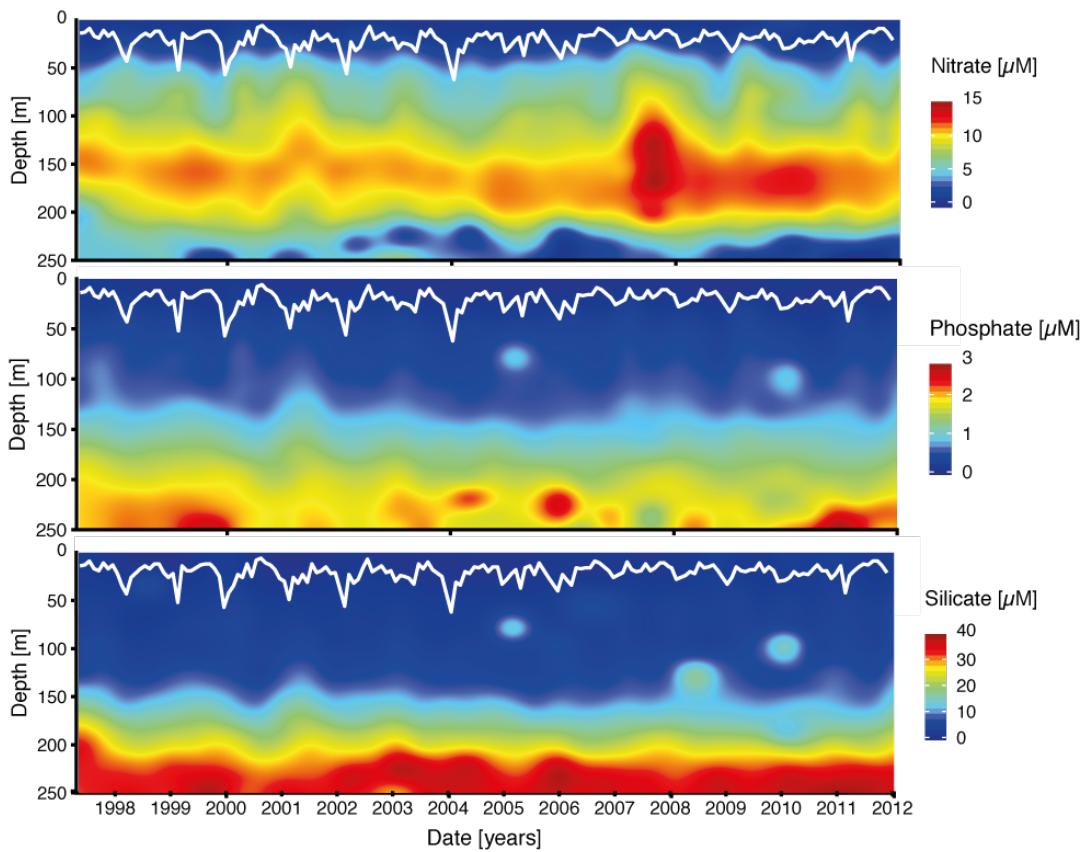


Figure 1.4: Contour plots of Nitrate, Phosphate and Silicate of the CARIACO time series down to 250 m. White line shows depth of MLD.

accumulating in the anoxic basin. Since November 1995 there were 232 core cruises at mostly monthly intervals, in addition to sediment trap and microbial-biogeochemistry process cruises. The full set of measurements and determination methods taken can be found in the manual of methods that was published by the coordinators of the time-series (Astor et al. 2013). Of particular interest to my applications are the detailed nutrient measurements taken using a Niskin bottle sampler and continuous flow analysis (see Figure 1.4 for contour plots of Nitrate, Phosphate and Silicate), as well as the phytoplankton taxonomy (see Figure 2.3 for PFT abundances) and high-pressure liquid chromatography (HPLC) measurements (see Figure 2.2 for HPLC derived total chlorophyll a *chl a*) also taken at discrete depth intervals over the duration of the time series.

The general dynamics observed over the course of the time-series are comprehensively discussed in a recent review paper by Muller-Karger et al. (2019). In the same publication it was also announced that the time-series was officially ended, due to a lack of funding. However, in the two decades that it existed the time-series data created a detailed description of the biogeochemistry of the Cariaco Basin. Despite the tropical location of the time series, the data reveals a seasonal cycle driven by upwelling along the southern coastline of the Caribbean Sea occurring usually between November and August, as trade winds intensify with the southward migration of the Intertropical Convergence Zone (ICTZ) (see Figure 1.3, A). The influx of nutrients provides the basis for high productivity (320 to $628 \text{ g C m}^{-2} \text{ y}^{-1}$) in the surface waters of the basin. This drives vertical export, with 9 to $10 \text{ g C m}^{-2} \text{ y}^{-1}$ reaching the bottom sediments, which amounts to 1-3 % of primary production (PP) (Muller-Karger et al. 2019). The waters are inhabited by a diverse community of microorganism, in particular at the oxic-anoxic interface at roughly 250 m depth where novel eukaryotes have been found (Stoeck et al. 2003).

Long-term trends show a reduction in upwelling intensity in the time from 2003 to 2013, which has been linked to a weakening trend in the trade winds and the north-easterly movement of the Atlantic centroid of the ICTZ (Taylor et al. 2012). This reduction in upwelling is visible in a reduction nitrate concentration and coincides with an increase in temperature across the upper mixed layer (see Figure 1.5).

The change in physical conditions was accompanied by a shift in the biotic community. Phytoplankton bloom intensities were reduced, while phytoplankton diversity and zooplankton densities showed an increasing trend. Interestingly, vertical export remained at a similar level, despite the fact that phytoplankton taxonomy data shows a marked reduction in larger phytoplankton species. (Taylor et al. 2012; Pinckney et al. 2015). The increase in zooplankton densities was linked to a collapse in the local sardine fisheries (see Figure 2.1). The biogeochemistry of the deeper waters reflects a shift in the biotic community at the surface (Scranton et al. 2014).

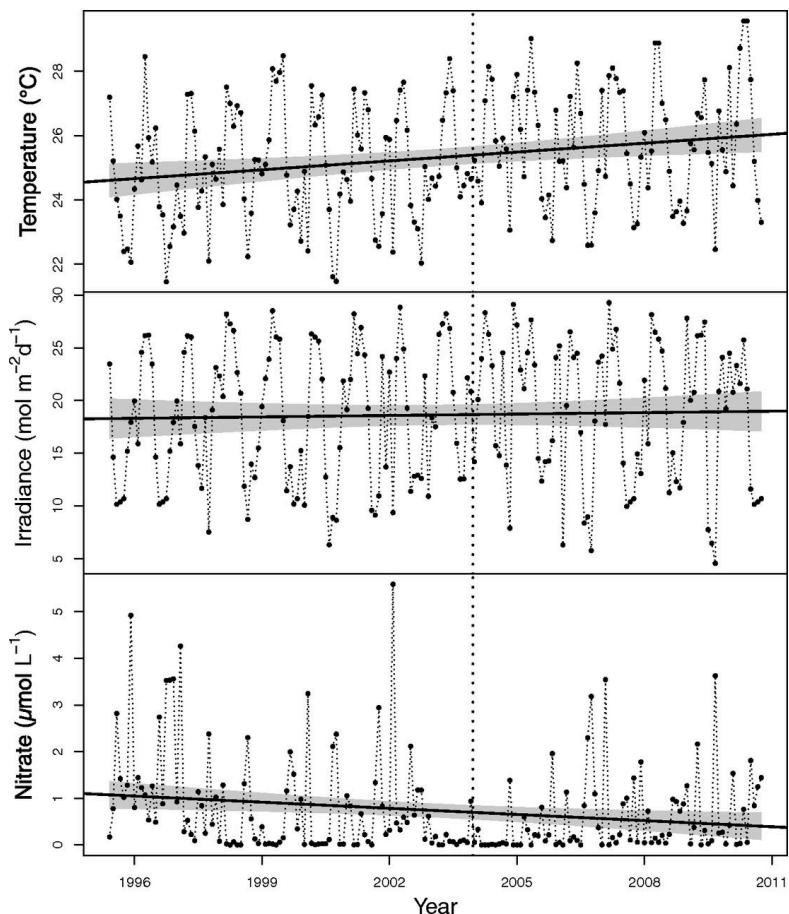


Figure 1.5: "Monthly environmental conditions averaged over the upper mixed layer (1, 7, 15, and 25 m depth) from the CARIACO Ocean Time-Series Program: temperature ($^{\circ}\text{C}$), irradiance ($\text{mol m}^{-2} \text{d}^{-1}$), and nitrate concentration ($\mu\text{mol L}^{-1}$). The vertical dotted line is drawn at the boundary (January 1, 2004) between the cool and warm periods. The straight lines are linear regressions: temperature = $(24.6 \pm 0.3) + (0.09 \pm 0.03)t$, $R^2 = 0.05$, $P < 0.005$; irradiance = $(18.1 \pm 0.9) + (0.05 \pm 0.11)t$, $R^2 = 0.001$, $P = 0.65$; nitrate = $(1.06 \pm 0.14) - (0.045 \pm 0.017)t$, $R^2 = 0.04$, $P = 0.03$, where t is time in years since January 1, 1996, errors are one SE, and the shaded region is the 95% confidence interval on the line. The R^2 is very low because of the tremendous interannual variation relative to the trend." from Irwin et al. (2015)

1.5 Aims of the proposed PhD project

The marked dynamics captured in the CARIACO time-series provide a wealth of data to explore during my PhD. With the upcoming collaboration with Andrew Barton at the Scripps Institution of Oceanography I will be working under the supervision of a phytoplankton modeling expert who has direct experience of the CARIACO dataset, as he himself worked on a proposal for an ecosystem modeling study in the region that was unfortunately not funded. I am also in contact with James Pinckney and Claudia Benitez-Nelson from the University of South Carolina, who have been heavily involved in the scientific project of CARIACO and have been kind enough to share specific datasets that were not publicly available. The general aim of my PhD Project, as exemplified by the three planned manuscripts, is to build a framework to explore the shifts observed in the phytoplankton community in the Cariaco Basin.

- The first study will focus on the biomass dynamics observed in the HPLC-derived phytoplankton pigment data. I have built an ecosystem model that resolves multiple nutrients and multiple functional types of both phytoplankton and zooplankton. The goal is to explore the drivers of community composition. Hypothesis of bottom-up and top-down processes (a decrease in upwelling and an increase in zooplankton grazing respectively) driving these changes have been put forward in the literature. The ecosystem model would allow for a mechanistic exploration of the effects of these processes on the phytoplankton community, with a direct comparison to the dynamics observed in the CARIACO data.
- During the model construction of the first study, I developed a technical concept to implement a flexible PFT model framework in the open-source programming language Python. The code structure allows for any combination and amount of functional types to be added to the model structure, which lends itself very well to the study of biodiversity effects, as well as simply implementing different types of ecosystem models. This structure is already used within the first study, however I aim to fully develop the code as a Python package, to release to the scientific community. The python package will form the basis of a technical manuscript to be submitted to the journal Geoscientific Model Development.

- As the first study uses a PFT modeling approach, but otherwise does not explore the effects of biodiversity on ecosystem function, the third study will aim to address this issue with the help of the developed modeling tool. The CARIACO time-series provides a rich dataset to test hypotheses, in particular of the effects of intra- and inter-PFT diversity. The collaboration with Andrew Barton is central to the formulation the specific scientific questions and model implementations. I aim to complete this project in the third year of my PhD.

2. Understanding phytoplankton community shifts in the eastern Cariaco basin

This is the current state of progress towards the first manuscript

2.1 Regime shift in the Cariaco Basin

The CARIACO time-series has been collecting detailed data on the phytoplankton community in the Cariaco Basin from 1995 to 2017 (see Section 1.4 for a full description). What has been a particular focus of the research based on this data set is the apparent changes in environmental conditions documented in both the physical boundary conditions as well as the biological data. Taylor et al. (2012) described a state shift in the biotic community. Using the fluorometric *chl a* data collected their analysis proposes a reduction in phytoplankton biomass, coinciding with a reduction in nutrient upwelling. An increase in zooplankton abundances (although the measurements of this were only commenced in 2001) was linked to a collapse in sardine fisheries in the region (see Figure 2.1). Pinckney et al. (2015) further expanded on this work by looking in detail at both the phytoplankton taxonomy data, as well as HPLC-derived *chl a* data. The HPLC-derived data showed no clear reduction in total *chl a* and spoke against a reduction in phytoplankton biomass. In comparison to the fluorometric method, HPLC can measure the concentrations of multiple pigment types and the total chlorophyll signal is less influenced by the specific pigment composition of the phytoplankton community. What the data showed specifically was that instead of a reduction in total chlorophyll within the water column, overall bloom intensity (i.e. the variability in biomass) decreased and phytoplankton moved to greater depth (see Figure 2.2). This

is also supported by irradiance measurements showing a reduction in the depth of the euphotic zone (Pinckney et al. 2015). Using the software CHEMTAX, the individual pigment composition can be used trace the proportion of total chlorophyll biomass that is substituted by the individual functional groups. This data was kindly supplied by James Pinckney and forms the basis of the planned study.

The two regimes described by Pinckney et al. (2015) do not describe a formally analyzed time point, but rather are the two time periods where HPLC data coverage was given. The previous analysis by Taylor et al. (2012) places the changing point into the year 2005, which coincides with the time point right before the second time period where HPLC-data was continuously measured. The term regime shift is actually not very well defined and has been used uncritically in the scientific literature (DeYoung et al. 2004). There are global trends and indications of a regime shift, but methods to identify regime shifts are not well established and have been critically discussed in the literature (Steele 2004; Mantua 2004; Litzow et al. 2016). To my knowledge no formal exploration of a potential regime shift has been performed with the CARIACO data, therefore the term regime shift is used here to describe the observed changes in the phytoplankton community and physical environment without presupposing a formally defined state shift in the entire ecosystem.

The variability in environmental conditions in Cariaco between the two regimes can be seen in the aggregated model forcing (see Figure 2.5). There is a clear reduction in MLD variability, which allows me to use this as one of the drivers of a possible regime shift, however I plan on performing a more formal analysis of the regime forcing in the future. The other main hypothesis that has been put forward as a possible explanation of the shifts in the phytoplankton community, in particular towards smaller sizes of phytoplankton, is the collapse in the sardine fisheries and the subsequent increase in zooplankton biomass.

Although the zooplankton data does not cover the the first time-period of HPLC-data coverage, there measurements of the *Phaeo* pigment over the entire time period. *Phaeo* pigment is a results of chlorophyll degradation and can be used as a proxy for zooplankton grazing pressure. The overall dynamics point towards an increase in grazing pressure, coinciding with the increase in zooplankton abundance that was

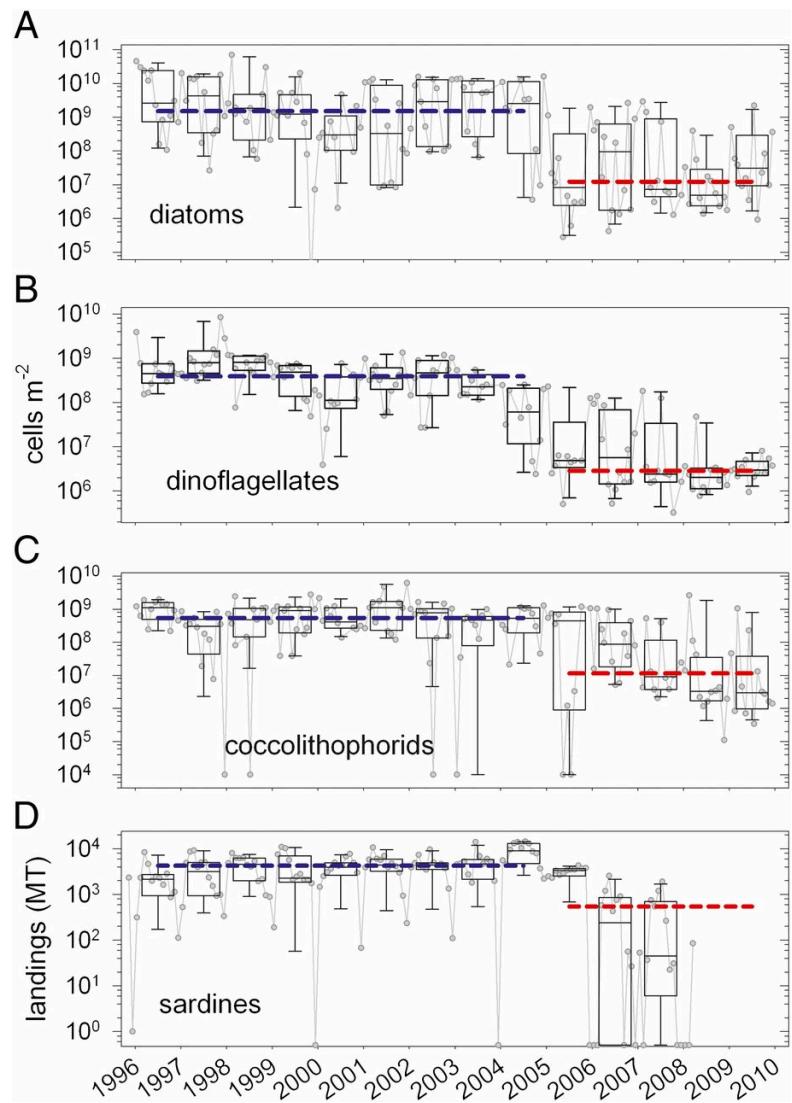


Figure 2.1: "Shifts in phytoplankton community composition and sardine landings from the southeastern Margarita Island fishery. Monthly observations presented as gray symbols. Box and whisker plots depict binned annual variations in diatom (A), dinoflagellate (B), coccolithophorid (C) inventories integrated over the upper 55 m and sardine fishery landings (D) in metric tons. Boxes represent the interquartile range of all observations (25th to 75th percentiles). Internal horizontal lines and whiskers are medians and 10th to 90th percentiles, respectively. Blue and red horizontal lines represent the grand medians of all observations between 1996 and 2004 and between 2005 and 2009, respectively. Data in early and late bins are significantly different in all cases (ANOVA; $P < 0.001$). [Fishery data are courtesy of L. W. González (Universidad de Oriente, Boca de Río, Isla de Margarita, Venezuela); zero values artificially set at 0.5 for plotting purposes.]" from Taylor et al. (2012)

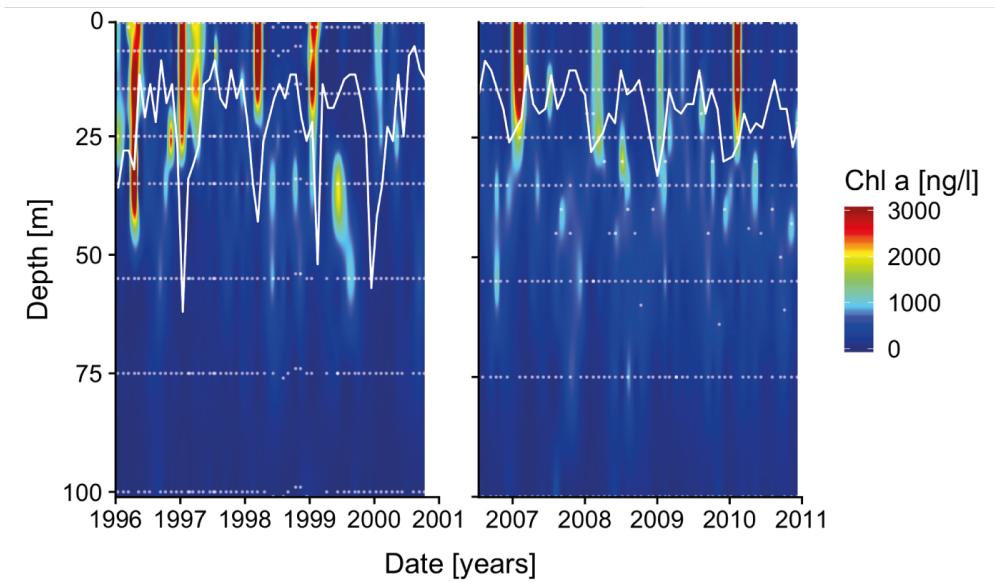


Figure 2.2: Contour plot of HPLC-measured *chl a* for the two time periods with full data coverage (January 1996 to October 2000 and July 2006 to December 2010). Light white dots indicate data points. White line shows the depth of the mixed layer. HPLC Data and MLD depth was received from James Pinckney and Claudia Benitez-Nelson.

documented (Pinckney et al. 2015). In the model the collapse of the sardine fisheries, an important grazer of zooplankton, can be directly implemented through the higher-order mortality term acting on zooplankton (see Figure 2.4).

As there has to date been no ecosystem modeling study performed in the region, I am using a relatively simple model formulation without an explicitly trait-based treatment of diversity. Instead I focus on resolving the phytoplankton and zooplankton functional types that were directly measured in the time-series. For phytoplankton, the data used is the functional-type resolved HPLC-derived *chl a* data presented by Pinckney et al. (2015). The CHEMTAX analysis resolved 7 functional types in total, based on the pigments that were measured. To simplify the implementation and parametrization of the model, these were aggregated to 5 groups as presented in Figure 2.3. As the groups of Cryptophytes, Prasinophytes and Chlorophytes do not present a unique biogeochemical role and showed similar dynamics, they were aggregated to the "Others" group. This is a preliminary grouping for initial analysis. Zooplankton were measured using bongo net tows with two different mesh sizes of 200 and 500 microns. The

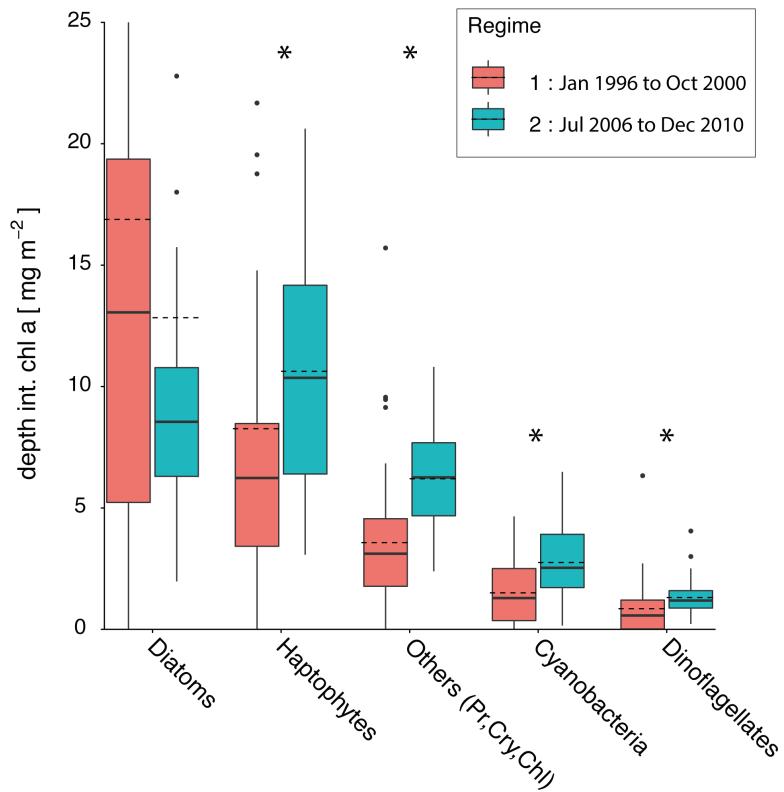


Figure 2.3: Boxplots of HPLC-measured *chl a* depth integrated to 100 m for the two time periods with full data coverage. Boxplots illustrate the 25th and 75th quartiles (the box), the whiskers show the 5th and 95th percentiles, full line shows the median, dotted line shows the mean value per regime. Asterisk * indicates a significant difference between Regimes 1 and 2 in a one-sided t-test ($p < 0.05$). Functional type *chl a* was separated using the software Chemtax as described in Pinckney et al. (2015)

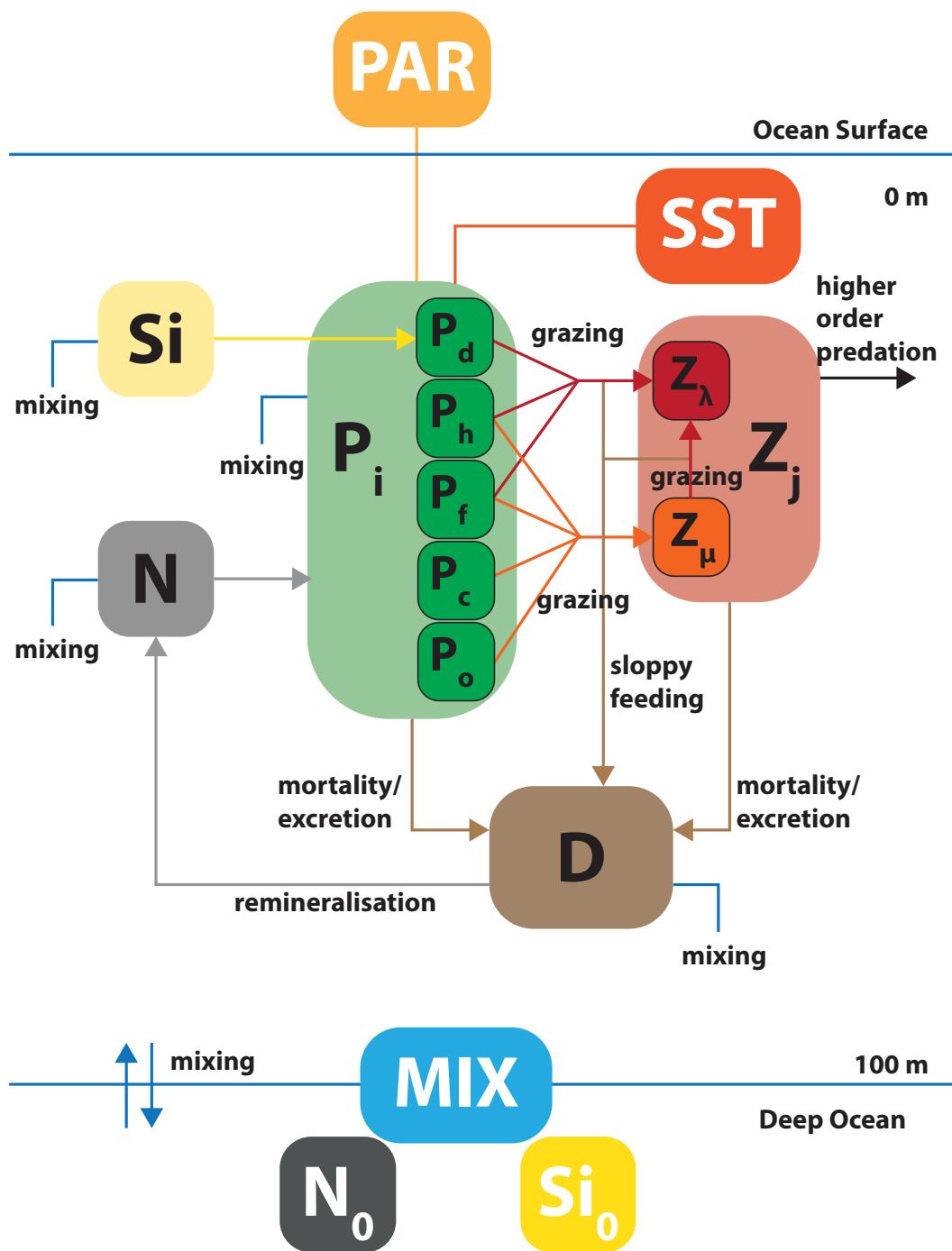


Figure 2.4: Model schematics of current iteration of the ecosystem model. The ecosystem is embedded in a box of fixed depth. Nutrients below are assumed to be constant, mixing is a function MLD to simulate seasonal upwelling. See the methods description for a full explanation of model components.

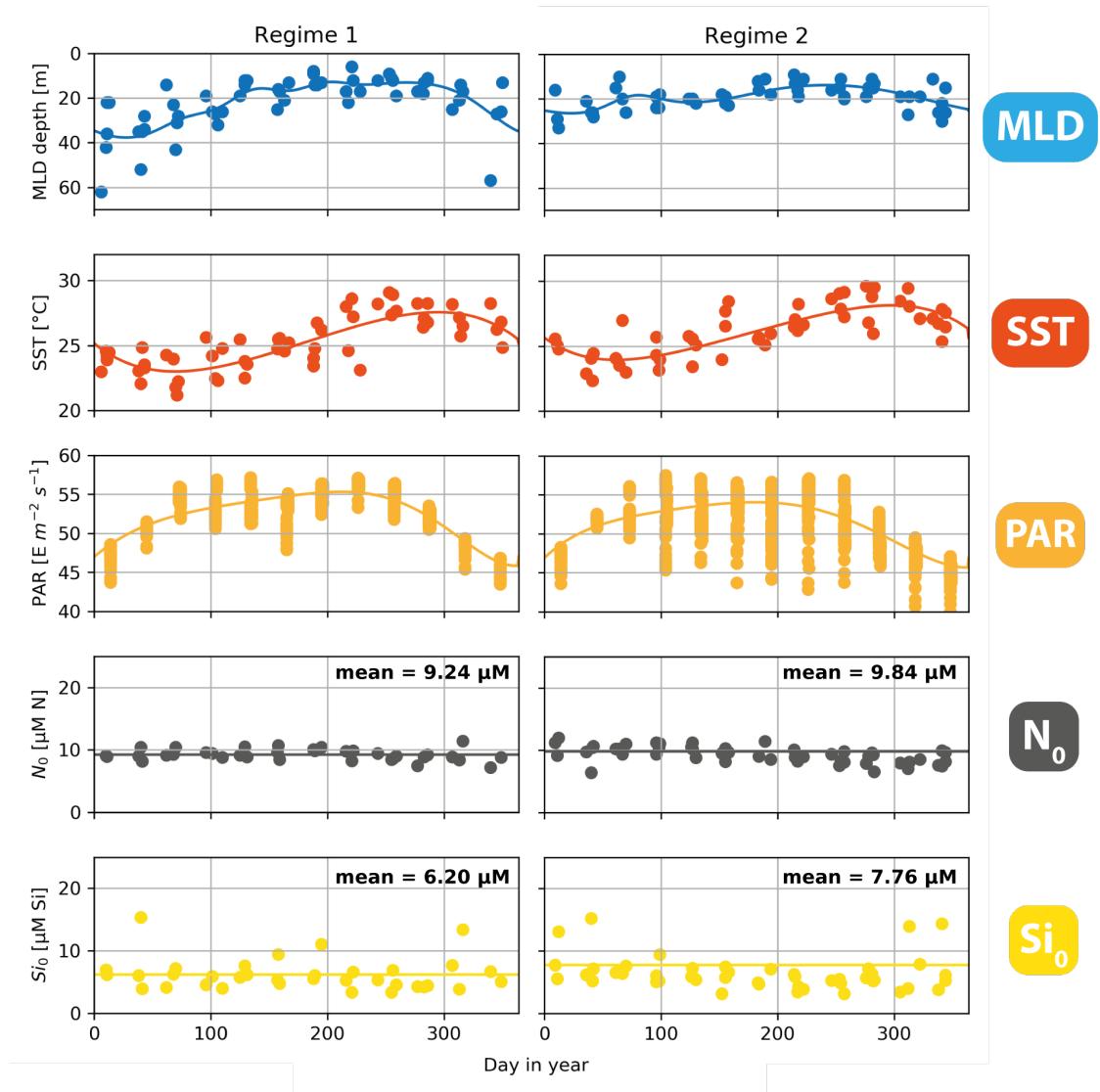


Figure 2.5: Aggregated forcing per the two regimes adapted from Pinckney et al. (2015). Regime 1 from January 1996 to October 2000, Regime 2 from July 2006 to December 2010, values are aggregated to a single year. Depth profiles of Nitrate and Silicate were interpolated to depth and averaged between 100 and 150 m. Continuous line shows interpolated values of MLD, PAR and SST used for model forcing and mean values for N and Si due to the assumption of a constant value at depth. SST and PAR data is taken from SeaWiFS satellite monthly climatological data.

model resolves these as two separate zooplankton types.

2.2 Methods

Model description

NPZD models are simplified marine ecosystem models that can be adapted to different physical settings and food web structures. For this model, the basic ecosystem structure is inspired by the models of Fasham (1990). The physical setting of the model uses a zero-dimensional structure where a fixed volume of water resides above a deep homogenous layer. The code structure is the PhytoMFTM model (as explained in Section 3) written in the open source programming language Python, which provides a flexible framework for NPZD-type models with multiple functional types of phytoplankton and zooplankton. The model code and all statistical scripts are available publicly on Github (<https://github.com/ben1post/BennyPhD>).

The model framework was adapted to the setting of the CARIACO time-series in the Cariaco basin of the coast of Venezuela. The data includes phytoplankton functional-type resolved HPLC data and two size-classes of zooplankton, which were included in the model as the 5 phytoplankton types, and 2 zooplankton types. The phytoplankton types include Diatoms P_d , Haptophytes P_h , Dinoflagellates P_f , Cyanobacteria P_c and Others (representing a generalized small green algae) P_o . There are two Zooplankton types split by size class, named Mikrozooplankton Z_μ and Mesozooplankton Z_λ .

Nitrogen N (and Silicate Si for Diatoms) is assimilated by the phytoplankton types P_i , which are grazed by the zooplankton types Z_j . Mortality of and excretion from phytoplankton and zooplankton, and sloppy feeding by zooplankton contribute to Detritus D . In addition to the linear mortality of P_i and Z_j , there is an additional quadratic mortality term acting on Z_j , which represents higher-order predation on zooplankton. Other nutrients are not implemented yet, but might be as the work progresses. The functional types differ only in their nutrient uptake dynamics and growth rates. Nitrogen-fixation of Cyanobacteria, which has been documented in the Cariaco Basin (Montes et al. 2013), could also be explicitly included, but has not been to date.

Model physics in a tropical coastal setting

The ecosystem component of the model is set within a zero-dimensional physical environment. The water column is divided into a box of 100 meters depth, above a homogenous deep layer. There is no lateral advection, but vertical mixing is modeled as a function of mixed layer depth (MLD) over time $M(t)$. Temperature depth profiles have been used to reconstruct the MLD at the investigated location. The derivative of MLD over time is given as $h(t) = dM(t)$. Exchange between the two layers is described by the two processes of turbulent diffusion and upwelling, with the depth of the MLD as a proxy for the strength of exchange between the two layers. The effects of entrainment and detrainment on nutrients, are given by the term $h^+(t) = \max[h(t), 0]$. Phytoplankton and detritus are assigned a constant sinking rate as of now, but this could be also be modified as a function of MLD depth. Zooplankton is assumed to be able to maintain themselves within the model box, therefore Z_j is not modified by the upwelling term. Diffusive mixing between the layers has been parameterized with a constant factor k . The entire upwelling term is thus

$$\kappa = \frac{k + h^+(t)}{M(t)} \quad (2.1)$$

In addition to the MLD interpolated from time series data, the model is externally forced with sea surface temperature (SST) and interpolated from monthly to daily values and photosynthetically active radiation (PAR) from monthly averaged SeaWiFS satellite data. This model construct is a preliminary modification of the typical slab model, as described in Anderson et al. (2015). A slab model describes the ecosystem component above a variable MLD, however this is not suitable in the tropical coastal location of the Cariaco Basin, as exemplified in Figure 2.6. Since the MLD is much shallower than in temperate settings, there is no constant nutrient concentration below the MLD, instead it is highly variable. In addition to this, in a typical slab model phytoplankton below the MLD is not described. In temperate settings the MLD can reach far below 100m and thereby far below the euphotic zone. In Cariaco the MLD is generally above the euphotic zone and there is a considerable concentration of phytoplankton growing below the MLD. Therefore the box model structure, resolving the top 100 m was adopted.

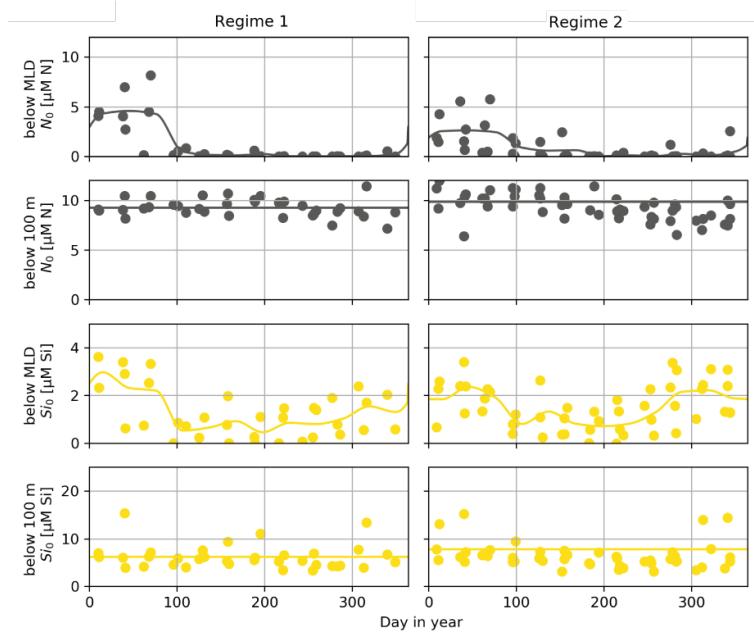


Figure 2.6: Nutrient concentrations below MLD and averaged between 100 and 150 m depth for both regimes aggregated to one year. For values below MLD the continuous line represents interpolated forcing, for values below 100 m the line shows the mean value.

Phytoplankton

Phytoplankton growth is a function of light (PAR), temperature (SST) and nutrients. These factors are assumed to independently limit growth, so that (exemplary for P_d , i.e. diatoms) the growth term is

$$\mu_d = \mu_d^{max} \cdot U_d(N, Si) \cdot L_d(PAR) \cdot T_d(SST) \quad (2.2)$$

where μ_d^{max} is the maximum growth rate per day and $T(PAR)$ is Eppleys formulation for temperature dependent growth, given as $T(SST) = e^{0.063*SST}$ with temperature in $^{\circ}C$. The light-limiting term $L(PAR)$ represents the integrated photosynthesis within the mixed layer as a function of incident irradiance at the surface I_0 . Light attenuation is calculated using the Lambert-Beer law with irradiance at depth z equal to

$$I(z) = I_0 \cdot e^{-k_{PAR} \cdot z} \quad (2.3)$$

Here, k_{PAR} is calculated as the sum of the constant attenuation coefficient of water

k_w and the self-shading of phytoplankton k_c with the unit μM^{-1} multiplied by total phytoplankton biomass P , i.e. $k_{PAR} = k_w + k_c P$. This model uses the Smith PI curve as a basis for the calculation, with V_p representing the photosynthetic rate, α , the initial slope of the PI curve and V_p^{Max} , the maximum photosynthetic rate

$$V_p = \frac{\alpha \cdot I \cdot V_p^{Max}}{\sqrt{(V_p^{Max})^2 + \alpha^2 \cdot I^2}} \quad (2.4)$$

Combining equation (2) and (3) as presented in Anderson et al. (2015), the integrated photosynthesis \bar{V}_p over depth z is calculated as

$$\bar{V}_p(z) = \frac{V_p^{Max}}{k_{PAR} \cdot z} \cdot \ln \left(\frac{\alpha \cdot I_0 + \sqrt{(V_p^{Max})^2 + (\alpha \cdot I_0)^2}}{\alpha \cdot I(z) + \sqrt{(V_p^{Max})^2 + (\alpha \cdot I(z))^2}} \right) \quad (2.5)$$

where \bar{V}_p equals the light-limiting term L in the growth equation (2).

Nutrient limited growth of the phytoplankton community is described via a Monod equation.

$$U(N) = \frac{N}{k_N + N} \quad (2.6)$$

For diatoms P_d the nutrient limiting term depends on both nitrogen and silicate concentration within the upper layer. According to Liebig's law of the minimum, always the lower nutrient availability limits Diatom growth:

$$U_d(N, Si) = \min \left(\frac{N}{k_d^N + N}, \frac{Si}{k_d^{Si} + Si} \right) \quad (2.7)$$

All other phytoplankton types are nutrient-limited only by available Nitrogen as in equation (6). Phytoplankton mortality and excretion are parameterized as a linear constant rate mo . Phytoplankton sinking is defined by the parameter v . With G_μ as grazing by Microzooplankton and G_λ as grazing by Mesozooplankton (defined below), the equations for all phytoplankton types P_i can be written as

$$\frac{dP_i}{dt} = \mu_i \cdot P_i - mo_i \cdot P - G_\mu(P_i) - G_\lambda(P_i) - v \cdot P_i \quad (2.8)$$

Zooplankton

Two zooplankton types are Microzooplankton Z_μ and Mesozooplankton Z_λ . Following Anderson et al. (2015) the grazing of, for example, Z_λ on diatoms P_d is formulated as follows

$$G_\lambda(P_d) = \left(\frac{\mu_\lambda^Z \cdot \phi_d^\lambda \cdot P_d}{(k_\lambda^Z)^2 + \phi_d^\lambda \cdot P_d + \phi_c^\lambda \cdot P_c + \phi_{df}^\lambda \cdot P_{df} + \phi_n^\lambda \cdot P_n + \phi_\mu^\lambda \cdot Z_\mu} \right) \cdot Z_\lambda \quad (2.9)$$

$$\phi_d^\lambda = \rho_d^\lambda P_d, \phi_c^\lambda = \rho_c^\lambda P_c, \phi_{df}^\lambda = \rho_{df}^\lambda P_{df}, \phi_n^\lambda = \rho_n^\lambda P_n, \phi_\mu^\lambda = \rho_\mu^\lambda Z_\mu$$

with μ_λ^Z as the maximum grazing rate, k_λ^Z as the half saturation constant of grazing, ϕ_d^λ as the density dependent feeding preference of Z_λ feeding on P_d , defined as $\rho_d \cdot P_d$, with ρ_d^λ as the feeding preference coefficient.

Solving method

The system of differential equations was solved numerically using the fourth-order Runge-Kutta method in the `odeint` function of the `scipy` package in python 3.7.

Full System of equations

$$\begin{aligned}
\frac{\partial N}{\partial t} &= \kappa \cdot (N_0 - N) + \delta_D^N \cdot D - \sum_{i=1}^{n_P} [\mu_i \cdot U_i(N_0, Si_0) \cdot L_i(PAR) \cdot T_i(SST) \cdot P_i] \\
\frac{\partial Si}{\partial t} &= \kappa \cdot (Si_0 - Si) - \mu_{dt} \cdot U_{dt}(N_0, Si_0) \cdot L_{dt}(PAR) \cdot T_{dt}(SST) \cdot P_{dt} \\
\frac{\partial P_i}{\partial t} &= \mu_i \cdot U_i(N_0, Si_0) \cdot L_i(PAR) \cdot T_i(SST) \cdot P_i - m_i \cdot P_i \\
&\quad - \sum_{j=1}^{n_Z} [I_j^{tot} \frac{p_j^i \cdot P_i}{R_j} Z_j] - v \cdot P_i \\
\frac{\partial Z_\mu}{\partial t} &= \delta_Z \cdot I_\mu^{tot} \cdot Z_\mu - \mu_\lambda \frac{Z_\mu}{Z_\mu + k_\lambda} Z_\lambda - m_\mu \cdot Z_\mu - g_\mu \cdot Z_\mu^2 \\
\frac{\partial Z_\lambda}{\partial t} &= \delta_Z \cdot I_\lambda^{tot} \cdot Z_\lambda + \delta_\lambda \cdot \mu_\lambda \frac{Z_\mu}{Z_\mu + k_\lambda} Z_\lambda - m_\lambda \cdot Z_\lambda - g_\lambda \cdot Z_\lambda^2 \\
\frac{\partial D}{\partial t} &= \sum_{j=1}^{n_Z} [(1 - \delta_Z) I_j^{tot} \cdot Z_j] + (1 - \delta_\lambda) \cdot \mu_\lambda \frac{Z_\mu}{Z_\mu + k_\lambda} Z_\lambda \\
&\quad - \sum_{j=1}^{n_Z} [m_j \cdot Z_j] + \sum_{i=1}^{n_P} [m_i \cdot P_i] - v_D \cdot D - \delta_D^N \cdot D
\end{aligned}$$

where:

N_0 = Nitrogen concentration right below mixed layer [μM],

N = Nitrogen concentration above mixed layer [μM],

v = sinking rate of P_i [$m day^{-1}$],

$M(t)$ = mixed layer depth at time point t [m],

$\kappa = \frac{1}{M(t)} \cdot (h^+(t) + \kappa)$ Constant that parameterizes diffusive mixing,

$h^+(t) = \max(0, \frac{d}{dt} M(t))$ Function that describes upwelling as function of MLD,

δ_D^N = Remineralization rate of nitrogen component of detritus D [$\mu M d^{-1}$],

v_D = sinking rate of D [$m day^{-1}$],

μ_i = Growth rate of phytoplankton type i [d^{-1}],

$U_i = \begin{cases} \min\left(\frac{N}{N+U_i^N}, \frac{Si}{Si+U_i^{Si}}\right), & \text{if P-type is Diatom} \\ \frac{N}{N+U_i^N}, & \text{otherwise} \end{cases}$ Nutrient uptake of P_i ,

$L_i = \frac{1}{M(t) \cdot k_w} \cdot \left(e^{1 - \frac{PAR(t)}{Opt_i^L}} + e^{1 - \frac{PAR(t) \cdot e^{-M(t) \cdot k_w}}{Opt_i^L}} \right)$ Light dependence of P_i ,

$$T_i = e^{0.063 \cdot SST} \text{ Temperature dependence of } P_i,$$

P_i = Biomass of phytoplankton type i [μMN],

m_i = Mortality/excretion rate for phytoplankton type i ,

$I_j^{tot} = \mu_j^Z \cdot \frac{R_j}{R_j + k_j^Z}$ Total intake of zooplankton type j ,

k_j^Z = Half saturation constant of zooplankton type j ,

$R_j = \sum_i (p_{ij} \cdot P_i)$ Total ressource density of zooplankton type j ,

p_j^i = Feeding preference of zooplankton type j feeding on phytoplankton type i ,

$R_\mu = p_\mu^n \cdot P_n + p_\mu^{dn} \cdot P_{dn} + p_\mu^c \cdot P_c$ Total ressource density of Mikrozooplankton Z_μ ,

$R_\lambda = p_\lambda^{dt} \cdot P_{dt} + p_\lambda^{dn} \cdot P_{dn} + p_\lambda^c \cdot P_c$ Total ressource density of Mesozooplankton Z_λ ,

Z_j = Biomass of zooplankton type j [μMN],

δ_Z = Grazing efficiency of zooplankton on phytoplankton (represents sloppy feeding),

g_i = Higher order predation on zooplankton (quadratic),

m_j = Mortality/excretion rate for zooplankton type j ,

2.3 Preliminary Results

The model as described is up and running, preliminary results are shown in Figure 2.7. As these are the results of using a parameter fitting routine implemented in the PhytoMFTM framework, the results are yet not entirely reasonable. One zooplankton type dominates, whilst the other is outcompeted. However the model is able to roughly recreate the nutrient dynamics observed in the first regime. The regimes are as defined by the HPLC data coverage, and the first regime will be used to create a baseline run to test the hypotheses on. Upwelling of nutrients can be modeled using the MLD forcing of the second regime and the collapse in sardine fisheries can be approximated with a reduction in the higher-order mortality term of zooplankton. As the model physics are not yet settled, this will have an effect on the observed dynamics and comprehensive parameter fitting will be performed later on.

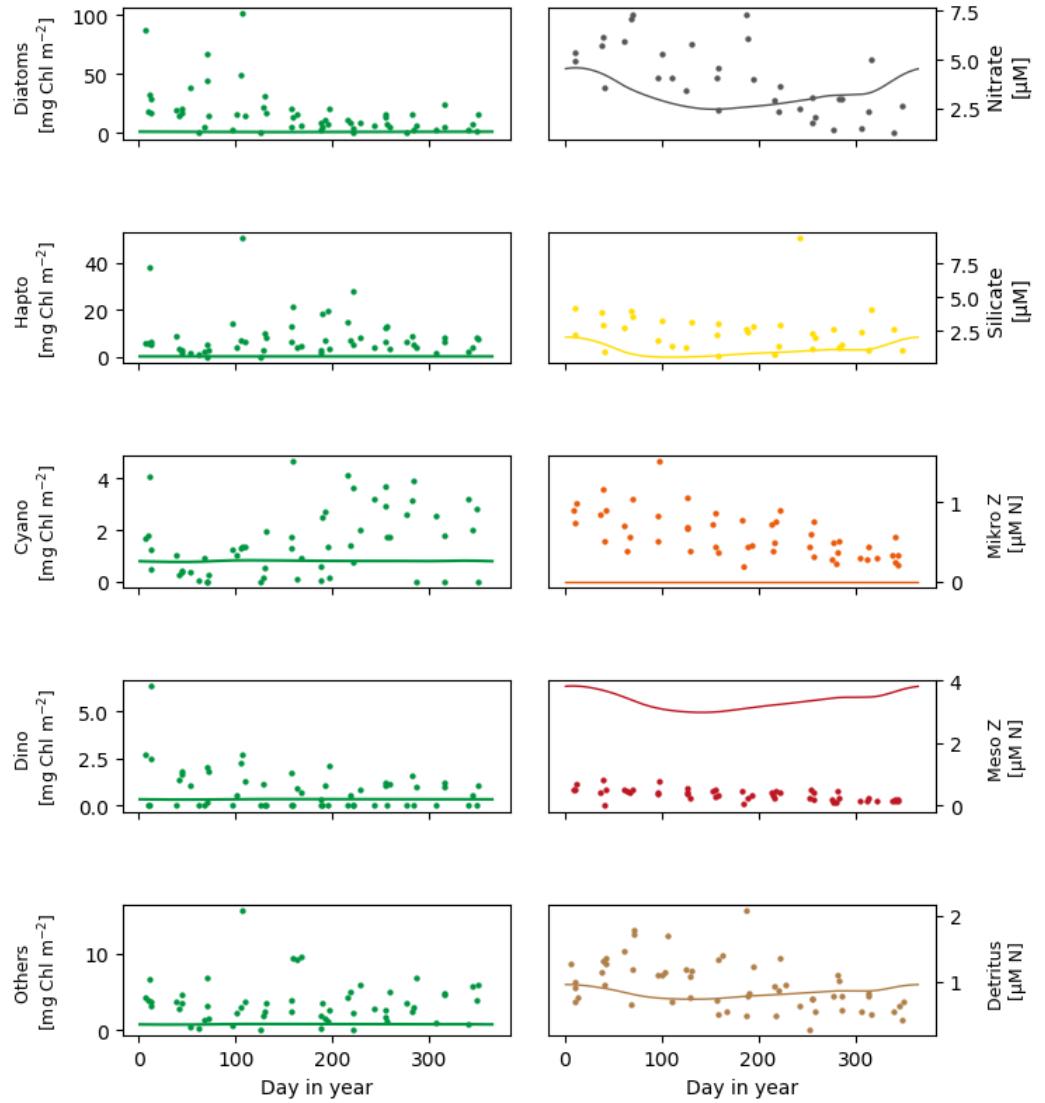


Figure 2.7: Current best fit of preliminary model. Data is shown as dots, model results by continuous line. Model output is in units of μM converted to integrated chlorophyll to compare to depth-integrated PFT *chl a* pigment data. Nutrient and detritus data was measured in μM . Zooplankton dry weight is converted by a carbon content of 32% and a constant C:N ratio of 5.625. Clearly the model parameters are not ideal yet, however a general agreement with the PFT data and nutrient dynamics can be observed. Zooplankton parameters need to be adjusted to allow for a more realistic response.

3. PhytoMFTM - a flexible object-oriented PFT model

This is the work towards the second manuscript, to be submitted to Geoscientific Model Development (GMD)

3.1 Python ecosystem model package development

The developments in computational capacities and tools have revolutionized the field of ecology. Where models were first only mathematically solved, soon they could be run on highly specialized computers until today where anybody can code or run a model on their personal computer, smartphone or even a cloud platform. This allows much greater reproducibility and makes sharing model code, results and statistical analysis much easier. Science can greatly benefit from this, even though there are many barriers to completely "Open Science". Open Science refers to transparency at all stages of the scientific process (Hampton et al. 2015).

Given that code is easily shared, this can be well implemented in ecological modeling research. There are different attempts at unified modeling frameworks, but most researchers still use different tools depending on their experience and computational literacy (Michener & Jones 2012). The development of open-source modeling packages is a step in the direction of greater reproducibility and collaboration in ecological research. The modeling community is moving towards open-source tools and in particular the programming language Python is growing in popularity (Lin 2012). The PhytoMFTM model structure itself will allow for flexible implementations of different

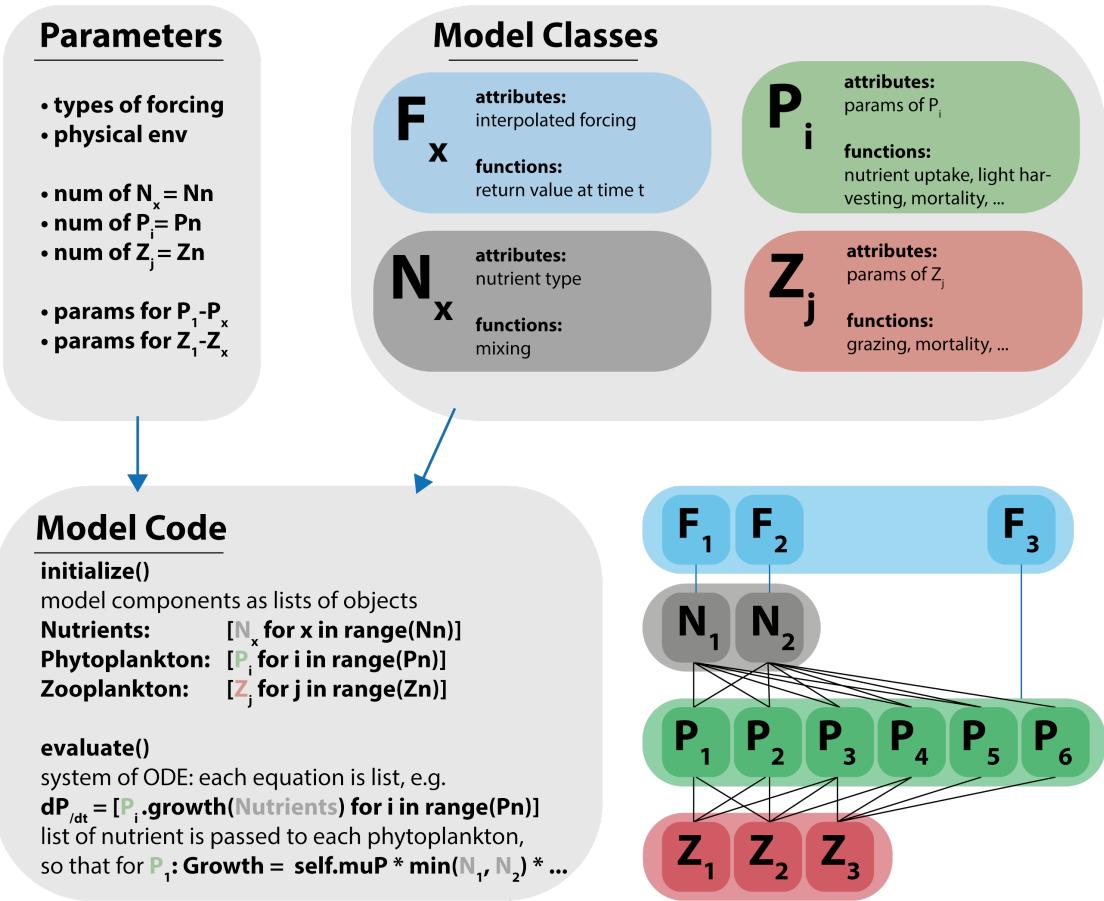


Figure 3.1: Schematic of code structure for PhytoMFTM model. Square brackets ("[]") denote a list object in Python, "for i in range(x)" within brackets creates list of length x of whatever comes before. Model classes are used to store both the parameters associated with each instance, e.g. different nutrient uptake parameters for PFTs and the functions that are used within the ODE. Each equation is also a list structure that calls every instance of the required objects using list comprehensions, so that all interactions are calculated according to the number of model components that were initialized in the parameters.

kinds of ecosystem models.

Object-oriented structure

The model code used in the previous section was developed with flexibility in mind. The main reason for building this code was that at the start of the project, I did not know yet how many functional types will have to be implemented. Instead of manually adding a new equation, each equation is implemented as a list of equations (see Figure 3.1). These lists contain the functions to be solved and iterate over a list of all model components. Each model class, like a phytoplankton type for example, contains the parameters assigned to it so that each individual interaction can be called upon as the equation iterates over the contained list. In theory this allows for an arbitrary number of model components, only constrained by computational costs that increase with each instantiated object. The object-oriented code allows for efficient calculation, particularly in relatively simple applications. For a model structure as used in Section 2 a five year run can be calculated in 2.5 seconds.

4. Further work

4.1 BDEF in Cariaco

The scientific question of the third planned study will be the relationship between diversity and ecosystem function in the Cariaco Basin.

This study will benefit from the CARIACO data for validating the diversity patterns simulated by the model. However, this will require a more detailed analysis of the diversity data found in the time series, as the available species counts are not directly comparable to the size-based functional type diversity simulated by the planned model. In addition to the conversion of species counts to functional types and size classes, a more rigorous investigation of BDEF relationships within the data is necessary to guide the modeling study and provide testable hypotheses. Upon arriving at Scripps, I will spend the first few months discussing the model structure with Andrew Barton and implementing the capabilities do use such a model in the PhytoMFTM package. Then I will focus on the data analysis and hypotheses testing parts of this project, which will then lead to the third manuscript. The model implementation could potentially be inspired by Loreau (1998) to investigate relationships between biodiversity and resource use efficiency. Of particular interest in the Cariaco time series are two proxies for diversity. Firstly the taxonomy data that was collected (see Figure 4.2), and secondly the proxy for functional WHAT? that is pigment diversity as measured via HPLC (see Figure 4.1).

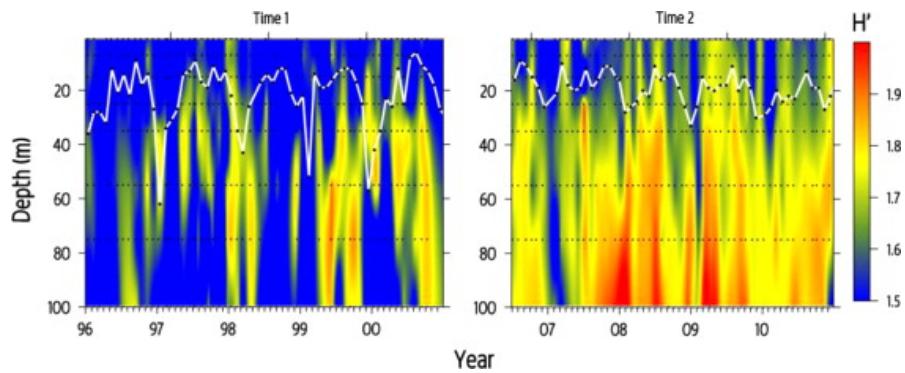


Figure 4.1: "Time series contour plot of photopigment diversity index (H). Data points are indicated by dots on the plot and the white line shows the mixed layer depth." from Pinckney et al. (2015)

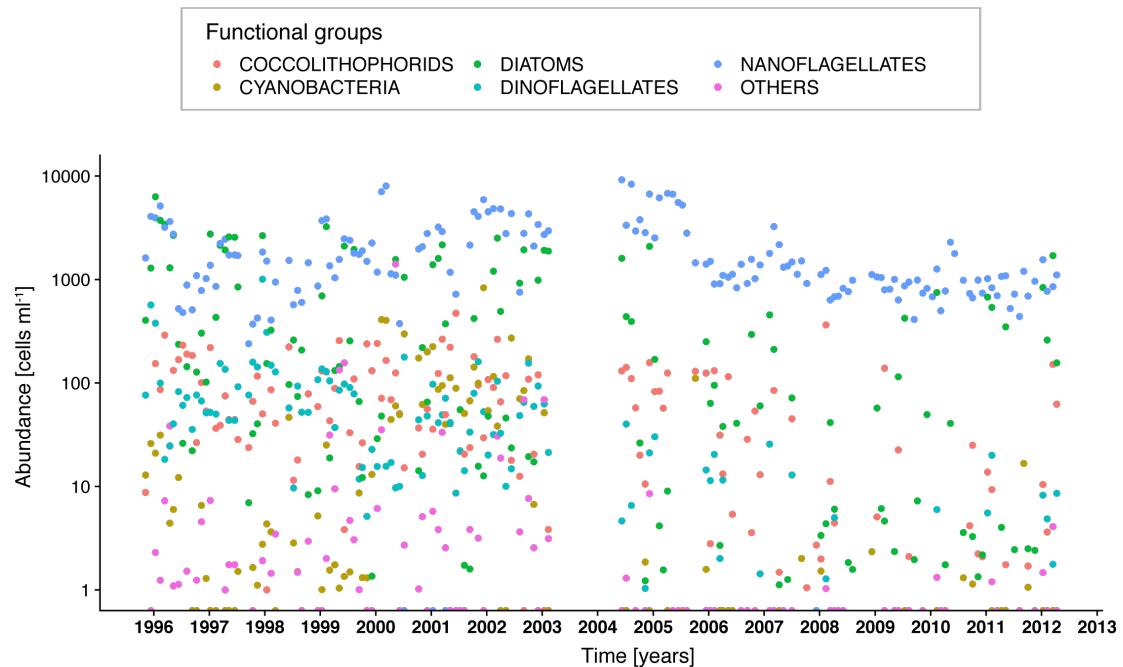
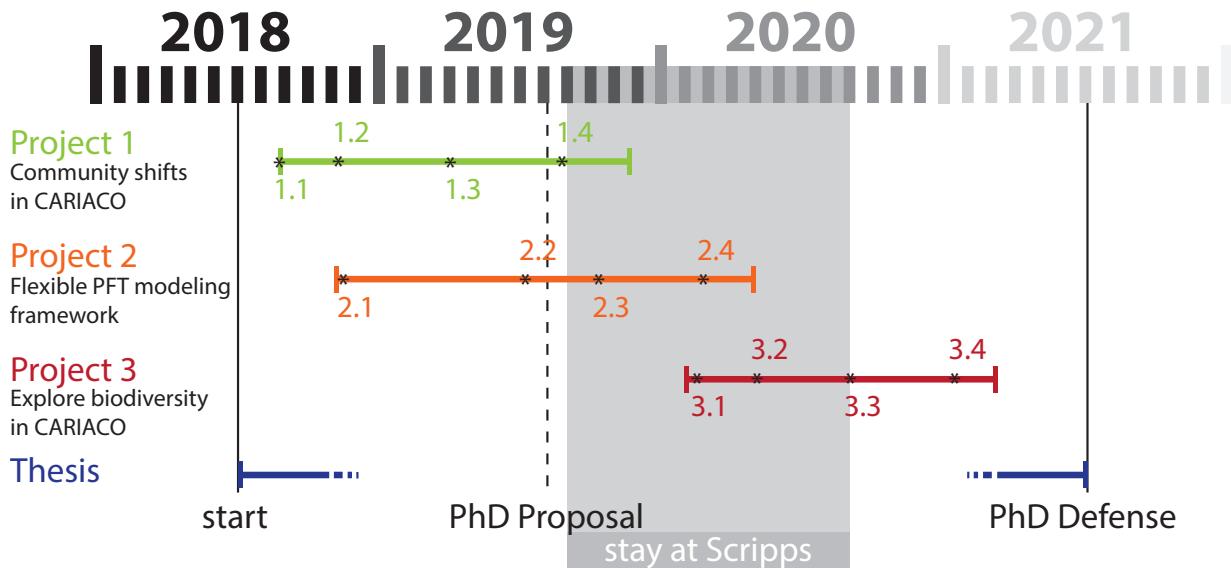


Figure 4.2: Abundances per functional group in phytoplankton taxonomy data of CARIACO time-series. Phytoplankton species were identified using light microscopy and identified down to genus or species level.

4.2 Time Table



The first 2 months of the PhD were spent working on a review of the current scientific literature .

Project 1 - Understanding community shifts

- 1.1: Finding public ocean time-series to use for phytoplankton ecosystem modeling.
- 1.2: Analyzing CARIACO time-series data and creating a model construct to test hypotheses.
- 1.3: After receiving HPLC data from James Pinckney and Claudia Benitez-Nelson, the model ecosystem structure was finalized.
- 1.4: Model physics need to be finalized, to create baseline run with the regime 1 forcing. Then start hypothesis testing, generating final results and writing up the manuscript in the following months.

Project 2 - PhytoMFTM

- 2.1: Idea for flexible object-oriented model was developed, first code testing.
- 2.2: Code structure was optimized for usage for Project 1.
- 2.3: Upon arriving at Scripps, I will discuss possible model structures for Project 3

with Andrew Barton and modify the model package accordingly.

2.4: Prepare model code for package deployment and start writing technical manuscript describing code function and structure.

Project 3 - BDEF in Cariaco

3.1: Settle on hypothesis and model structure together with Andrew Barton, start specific data analysis of CARIACO time-series data relevant to the question.

3.2: Finish data analysis of diversity related data and start implementing the model with the finished PhytoMFTM framework.

3.3: Start hypothesis testing and parameter fitting.

3.4: Finish up results and start writing manuscript.

The last 3 months of the PhD are reserved for writing up the thesis and preparing PhD Defense.

Bibliography

- Acevedo-Trejos E, Brandt G, Lan Smith S, Merico A (2016) PhytoSFDM version 1.0.0: Phytoplankton Size and Functional Diversity Model. *Geoscientific Model Development* 9(11):4071–4085
- Acevedo-Trejos E, Marañón E, Merico A (2018) Phytoplankton size diversity and ecosystem function relationships across oceanic regions. *Proceedings of the Royal Society B: Biological Sciences* 285(1879):20180621
- Alvera-Azcárate A, Barth A, Weisberg RH (2009) A nested model of the Cariaco Basin (Venezuela): Description of the basin's interior hydrography and interactions with the open ocean. *Ocean Dynamics* 59(1):97–120
- Anderson TR (2005) Plankton functional type modelling: Running before we can walk?
- Anderson TR, Gentleman WC, Yool A (2015) EMPOWER-1.0: An Efficient Model of Planktonic ecOsystems WrittEn in R. *Geoscientific Model Development* 8(7):2231–2262
- Astor YM, Fanning K, Guzmán L, Li X, Lorenzoni L, Masserini R, Muller-Karger F, Tappa E, Varela R (2013) Handbook of Methods for the Analysis of Oceanographic Parameters at the CARIACO Time-series Station: Cariaco Time Series Study.
- Barton AD, Pershing AJ, Litchman E, Record NR, Edwards KF, Finkel ZV, Kiørboe T, Ward BA (2013) The biogeography of marine plankton traits. *Ecology Letters* 16(4):522–534
- Boyce DG, Lewis M, Worm B (2012) Integrating global chlorophyll data from 1890 to 2010. *Limnology and Oceanography: Methods* 10(11):840–852
- Bringué M, Pospelova V, Tappa EJ, Thunell RC (2019) Dinoflagellate cyst production in the Cariaco Basin: A 12.5year-long sediment trap study. *Progress in Oceanography* 171:175–211
- Bruggeman J (2009) Succession in plankton communities: a trait-based perspective. *Department of Theoretical Biology* :158

- Chavez FP, Messié M, Pennington JT (2010) Marine Primary Production in Relation to Climate Variability and Change. *Annual Review of Marine Science* 3(1):227–260
- De Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, Lara E, Berney C, Le Bescot N, Probert I (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science* 348(6237):1261605
- Des Roches S, Post DM, Turley NE, Bailey JK, Hendry AP, Kinnison MT, Schweitzer JA, Palkovacs EP (2018) The ecological importance of intraspecific variation. *Nature Ecology and Evolution* 2(1):57–64
- DeYoung B, Harris R, Alheit J, Beaugrand G, Mantua N, Shannon L (2004) Detecting regime shifts in the ocean: Data considerations. *Progress in Oceanography* 60(2-4):143–164
- Dutkiewicz S, Morris JJ, Follows MJ, Scott J, Levitan O, Dyhrman ST, Berman-Frank I (2015) Impact of ocean acidification on the structure of future phytoplankton communities. *Nature Climate Change* 5(11):1002–1006
- Eppley RW, Peterson BJ (1979) Particulate organic matter flux and planktonic new production in the deep ocean. *Nature* 282(5740):677
- Evans GT (1988) A framework for discussing seasonal succession and coexistence of phytoplankton species. *Limnology and Oceanography* 33(5):1027–1036
- Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJ (2004) The evolution of modern eukaryotic phytoplankton. *Science* 305(5682):354–360
- Fasham MJR, Ducklow HW, McKelvie SM (1990) A nitrogen-based model of plankton dynamics in the oceanic mixed layer. *Journal of Marine Research* 48(3):591–639
- Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* 281(5374):237–240
- Finkel ZV, Beardall J, Flynn KJ, Quigg A, Rees TAV, Raven JA (2010) Phytoplankton in a changing world: Cell size and elemental stoichiometry. *Journal of Plankton Research* 32(1):119–137
- Fleming RH (1939) The control of diatom populations by grazing. *ICES Journal of Marine Science* 14(2):210–227
- Flynn KJ (2010) Ecological modelling in a sea of variable stoichiometry: Dysfunctionality and the legacy of Redfield and Monod. *Progress in Oceanography* 84(1-2):52–65
- Flynn KJ, St John M, Raven JA, Skibinski DOF, Allen JI, Mitra A, Hofmann EE (2015) Acclimation, adaptation, traits and trade-offs in plankton functional type models: reconciling terminology for biology and modelling. *Journal of Plankton Research* 37(4):683–691

- Follows MJ, Dutkiewicz S (2011) Modeling Diverse Communities of Marine Microbes. *Annual Review of Marine Science* 3(1):427–451
- Follows MJ, Dutkiewicz S, Grant S, Chisholm SW (2007) Emergent biogeography of microbial communities in a model ocean. *Science* 315(5820):1843–1846
- Franks PJ (2009) Planktonic ecosystem models: Perplexing parameterizations and a failure to fail. *Journal of Plankton Research* 31(11):1299–1306
- Geider RJ, MacIntyre HL, Kana TM (1997) Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a: carbon ratio to light, nutrient-limitation and temperature. *Marine Ecology Progress Series* 148:187–200
- Gentleman W (2002) A chronology of plankton dynamics in silico: How computer models have been used to study marine ecosystems. *Hydrobiologia* 480:69–85
- Gregg WW, Ginoux P, Schopf PS, Casey NW (2003) Phytoplankton and iron: Validation of a global three-dimensional ocean biogeochemical model. *Deep-Sea Research Part II: Topical Studies in Oceanography* 50(22–26):3143–3169
- Hampton SE, Anderson SS, Bagby SC, Gries C, Han X, Hart EM, Jones MB, Lenhardt WC, Mac- Donald A, Michener WK, Mudge J, Pourmokhtarian A, Schildhauer MP, Woo KH, Zimmerman N (2015) The Tao of open science for ecology. *Ecosphere* 6(7):art120
- Hellweger FL (2017) 75 years since Monod: It is time to increase the complexity of our predictive ecosystem models (opinion). *Ecological Modelling* 346:77–87
- Hellweger FL, Bucci V (2009) A bunch of tiny individuals-Individual-based modeling for microbes. *Ecological Modelling* 220(1):8–22
- Hirata T, Hashioka T, Shigemitsu M, Yamanaka Y, Saux-Picart S, Allen JI, Aita-Noguchi M, Sumata H (2013) A comparison between phytoplankton community structures derived from a global 3D ecosystem model and satellite observation. *Journal of Marine Systems* 109–110:129–137
- Honjo S, Manganini SJ, Krishfield RA, Francois R (2008) Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: A synthesis of global sediment trap programs since 1983. *Progress in Oceanography* 76(3):217–285
- Hutchinson GE (1961) The Paradox of the Plankton. *The American Naturalist* 95(882):137–145
- Iglesias-Rodríguez MD, Brown CW, Doney SC, Kleypas J, Kolber D, Kolber Z, Hayes PK, Falkowski PG (2002) Representing key phytoplankton functional groups in ocean carbon cycle models: Coccolithophorids. *Global Biogeochemical Cycles* 16(4):47–1–47–20

- Irwin AJ, Finkel ZV (2017) Phytoplankton functional types: A trait perspective. *bioRxiv* :148312
- Irwin AJ, Finkel ZV, Müller-Karger FE, Troccoli Ghinaglia L (2015) Phytoplankton adapt to changing ocean environments. *Proceedings of the National Academy of Sciences* 112(18):5762–5766
- Lacroix G, Ruddick K, Park Y, Gypens N, Lancelot C (2007) Validation of the 3D biogeochemical model MIRO&CO with field nutrient and phytoplankton data and MERIS-derived surface chlorophyll a images. *Journal of Marine Systems* 64(1-4):66–88
- Le Quéré C, Harrison SP, Prentice IC, Buitenhuis ET, Aumont O, Bopp L, Claustre H, Cotrim Da Cunha L, Geider R, Giraud X, Klaas C, Kohfeld KE, Legendre L, Manizza M, Platt T, Rivkin RB, Sathyendranath S, Uitz J, Watson AJ, Wolf-Gladrow D (2005) Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology* 11(11):2016–2040
- Lin JWB (2012) Why python is the next wave in earth sciences computing. *Bulletin of the American Meteorological Society* 93(12):1823–1824
- Litchman E, Klausmeier CA (2008) Trait-Based Community Ecology of Phytoplankton. *Annual Review of Ecology, Evolution, and Systematics* 39(1):615–639
- Litchman E, Klausmeier CA, Schofield OM, Falkowski PG (2007) The role of functional traits and trade-offs in structuring phytoplankton communities: Scaling from cellular to ecosystem level. *Ecology Letters* 10(12):1170–1181
- Litzow MA, Hobday AJ, Frusher SD, Dann P, Tuck GN (2016) Detecting regime shifts in marine systems with limited biological data: An example from southeast Australia. *Progress in Oceanography* 141:96–108
- Loreau M (1998) Biodiversity and ecosystem functioning: a mechanistic model. *Proceedings of the National Academy of Sciences* 95(10):5632–5636
- Mantua N (2004) Methods for detecting regime shifts in large marine ecosystems: A review with approaches applied to North Pacific data. *Progress in Oceanography* 60(2-4):165–182
- Margalef R (1978) Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanologica acta* 1(4):493–509
- McGill BJ, Enquist BJ, Weiher E, Westoby M (2006) Rebuilding community ecology from functional traits. *Trends in Ecology and Evolution* 21(4):178–185
- Merico A, Bruggeman J, Wirtz K (2009) A trait-based approach for downscaling complexity in plankton ecosystem models. *Ecological Modelling* 220(21):3001–3010

- Michener WK, Jones MB (2012) Ecoinformatics: supporting ecology as a data-intensive science Trends in Ecology and Evolution,. *Trends in Ecology and Evolution*, 27(2)(2):85–93
- Montes E, Thunell R, Muller-Karger FE, Lorenzoni L, Tappa E, Troccoli L, Astor Y, Varela R (2013) Sources of $\delta^{15}\text{N}$ variability in sinking particulate nitrogen in the Cariaco Basin, Venezuela. Deep-Sea Research Part II: Topical Studies in Oceanography 93(September):96–107
- Muller-Karger FE, Varela R, Thunell R, Luerssen R, Hu C, Walsh JJ (2005) The importance of continental margins in the global carbon cycle. *Geophysical Research Letters* 32(1):1–4
- Muller-Karger FE, Astor YM, Benitez-Nelson CR, Buck KN, Fanning KA, Lorenzoni L, Montes E, Rueda-Roa DT, Scranton MI, Tappa E, Taylor GT, Thunell RC, Troccoli L, Varela R (2019) The Scientific Legacy of the CARIACO Ocean Time-Series Program. *Annual Review of Marine Science* 11(1):annurev-marine-010318-095150
- Mutshinda CM, Finkel ZV, Irwin AJ (2013a) Which environmental factors control phytoplankton populations? A Bayesian variable selection approach. *Ecological Modelling* 269:1–8
- Mutshinda CM, Troccoli-Ghinaglia L, Finkel ZV, Müller-Karger FE, Irwin AJ (2013b) Environmental control of the dominant phytoplankton in the Cariaco basin: A hierarchical Bayesian approach. *Marine Biology Research* 9(3):247–261
- Pinckney JL, Benitez-Nelson CR, Thunell RC, Muller-Karger F, Lorenzoni L, Troccoli L, Varela R (2015) Phytoplankton community structure and depth distribution changes in the Cariaco Basin between 1996 and 2010. *Deep-Sea Research Part I: Oceanographic Research Papers* 101:27–37
- Prowe AE, Pahlow M, Dutkiewicz S, Follows M, Oschlies A (2012) Top-down control of marine phytoplankton diversity in a global ecosystem model. *Progress in Oceanography* 101(1):1–13
- Queirós AM, Bruggeman J, Stephens N, Artioli Y, Butenschön M, Blackford JC, Widdicombe S, Allen JI, Somerfield PJ (2015) Placing biodiversity in ecosystem models without getting lost in translation. *Journal of Sea Research* 98:83–90
- Redfield AC (1958) The biological control of chemical factors in the environment. *American scientist* 46(3):230A–221
- Redfield AC (1963) The influence of organisms on the composition of seawater. *The sea* 2:26–77
- Richards FA, Vaccaro RF (1956) The Cariaco Trench, an anaerobic basin in the Caribbean Sea. *Deep Sea Research* (1953) 3(3):214–228
- Rykiel Jr EJ (1996) Testing ecological models: the meaning of validation. *Ecological modelling* 90(3):229–244

- Scranton MI, Taylor GT, Thunell R, Benitez-Nelson CR, Muller-Karger F, Fanning K, Lorenzoni L, Montes E, Varela R, Astor Y (2014) Interannual and subdecadal variability in the nutrient geochemistry of the Cariaco Basin. *Oceanography* 27(1):148–159
- Shimoda Y, Arhonditsis GB (2016) Phytoplankton functional type modelling: Running before we can walk? A critical evaluation of the current state of knowledge. *Ecological Modelling* 320:29–43
- Smith SL, Merico A, Wirtz KW, Pahlow M (2014) Leaving misleading legacies behind in plankton ecosystem modelling. *Journal of Plankton Research* 36(3):613–620
- Steele JH (1958) Plant production in the northern North Sea. HM Stationery Office
- Steele JH (2004) Regime shifts in the ocean: Reconciling observations and theory. *Progress in Oceanography* 60(2-4):135–141
- Stock CA, John JG, Rykaczewski RR, Asch RG, Cheung WWL, Dunne JP, Friedland KD, Lam VWY, Sarmiento JL, Watson RA (2017) Reconciling fisheries catch and ocean productivity. *Proceedings of the National Academy of Sciences* 114(8):E1441–E1449
- Stoeck T, Taylor GT, Epstein SS (2003) Novel eukaryotes from the permanently anoxic Cariaco Basin (Caribbean Sea). *Applied and Environmental Microbiology* 69(9):5656–5663
- Taylor GT, Muller-Karger FE, Thunell RC, Scranton MI, Astor Y, Varela R, Ghinaglia LT, Lorenzoni L, Fanning KA, Hameed S, Doherty O (2012) Ecosystem responses in the southern Caribbean Sea to global climate change. *Proceedings of the National Academy of Sciences* 109(47):19315–19320
- Tilman D (1990) Constraints and Tradeoffs: Toward a Predictive Theory of Competition and Succession. *Oikos* 58(1):3
- Tilman D (2001) Functional diversity. *Encyclopedia of Biodiversity* 3(1):13
- Violle C, Navas ML, Vile D, Kazakou E, Fortunel C, Hummel I, Garnier E (2007) Let the concept of trait be functional! *Oikos* 116(5):882–892
- Violle C, Enquist BJ, McGill BJ, Jiang L, Albert CH, Hulshof C, Jung V, Messier J (2012) The return of the variance: Intraspecific variability in community ecology. *Trends in Ecology and Evolution* 27(4):244–252
- Violle C, Thuiller W, Mouquet N, Munoz F, Kraft NJ, Cadotte MW, Livingstone SW, Mouillot D (2017) Functional Rarity: The Ecology of Outliers. *Trends in Ecology and Evolution* 32(5):356–367
- Wakeham SG, Turich C, Schubotz F, Podlaska A, Li XN, Varela R, Astor Y, Sáenz JP, Rush D, Sinninghe Damsté JS, Summons RE, Scranton MI, Taylor GT, Hinrichs KU (2012) Biomarkers, chemistry and microbiology show chemoautotrophy in a multilayer chemocline in the Cariaco Basin. *Deep-Sea Research Part I: Oceanographic Research Papers* 63:133–156

Walsh JJ, Dieterle DA, Müller-Karger FE, Bohrer R, Bissett WP, Varela RJ, Aparicio R, Díaz R, Thunell R, Taylor GT, Scranton MI, Fanning KA, Peltzer ET (2002) Simulation of carbon-nitrogen cycling during spring upwelling in the Cariaco Basin. *Journal of Geophysical Research: Oceans* 104(C4):7807–7825

Ward BA, Dutkiewicz S, Jahn O, Follows MJ (2012) A size-structured food-web model for the global ocean. *Limnology and Oceanography* 57(6):1877–1891

Yool A, Fasham MJR (2001) An examination of the “continental shelf pump” in an open ocean general circulation model. *Global Biogeochemical Cycles* 15(4):831–844