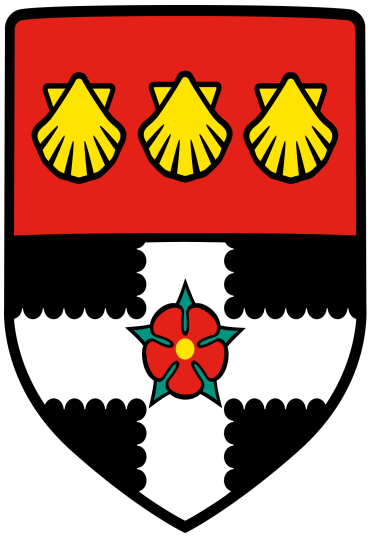
UNIVERSITY OF READING

Department of Geography and Environmental Science



Exploring structural sensitivity in a 1-D marine biogeochemical model

Prima Anugerahanti

A thesis submitted for the degree of Doctor of Philosophy

20 September 2019

Abstract

Equations used to describe the main biological processes determine the dynamics of biogeochemical models. From previous studies, altering the form of these process 'structure functions' has been shown to produce larger differences than changing the values of the parameters used in the models. This study explores the effect of this structural sensitivity in a marine biogeochemical model by generating an ensemble of runs. We use a 1-D Model of Ecosystem Dynamics, nutrient Utilisation, Sequestration, and Acidification (MEDUSA) ensemble, where each member has a different combination of key biogeochemical process equations, each of which tuned to be as similar as possible to the default functions. The model is run at five oceanographic stations spanning three different biogeochemical regimes or provinces: oligotrophic, coastal, and abyssal plain. Marine biogeochemical models are also sensitive to the physical environment, so we also explore the relative impact of altering the physical input and biogeochemical process equations, separately and together.

The impacts of perturbing the biogeochemistry and physics are quantified using statistical metrics, chlorophyll depth distributions, and phytoplankton bloom phenology. We explored the signature characteristics of the different ensembles by examining the anomaly correlations between different ensemble members and also the nitrogen fraction in phytoplankton across different ensemble members.

We found that even small perturbations in model structure can produce a large ensemble spread in many metrics that then mostly easily encompasses the in situ observations. This perturbed biogeochemistry ensemble (PBE) also has an improved RMSE between observations and the ensemble mean, compared to a single deterministic model default run. Perturbing the physics does not generate as large an ensemble range in many of the metrics studied, and cannot always encompass the in situ chlorophyll observations.

From exploring the signature characteristics of the different ensembles, very different characteristics are produced from the two ensembles. Perturbing biogeochemistry alters exchange fluxes between biogeochemical compartments, whereas perturbing the physics only alters the nutrient supply to the biological compartments. Therefore, the perturbed biogeochemistry ensemble provides better representations of uncertainty. We discuss how this might be useful for interpreting discrepancies against observational data.

Declaration of Authorship

Declaration:

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Prima Anugerahanti.

Acknowledgements

Firstly, I would like to thank my supervisors, Shovonlal Roy and Keith Haines, for their infinite patience, support, and guidance throughout my PhD. Thank you for the insights, knowledge, and all the things you have taught me. I would also thank my Panel Chair, Steve Robinson, for all his useful input, and support at our meetings. Additionally, I would like to thank Kevin White for his advice on remote sensing and PhD admins, John Hemmings, who gave me the most recent code of MarMOT-1.1, and Denise Cummings who provided the phytoplankton in situ data at L4.

There are many other people I would like to thank for making the 'battery farm' office less depressing. In particular, thanks to my cohort: Robin Boyd, Ben Jones, and Callum Nolan, who have made coffee breaks and lunchtime full of laughter. I would also like to thank Sarah, Harriet, and Heather for their support during moments of despair in the office and answering questions on administrative stuff. They also fill in on tv show gossips. Additionally, Alice Coles, my desk mate who has allowed me to do the cover illustration for her book, and being a very inspiring person. Thank you, everyone, in the Archway Lodge and Reading Islamic Society for their hospitality and providing delicious meals after salah. Additionally, thanks to my housemate, Demitrius, who has made the house in order as well as providing some deep and funny conversations at dinner time.

A big thank you for my partner and best friend, Thomas Hockham, for all the emotional support, proof-reading feedbacks, delicious dinner, cleaning, bike repairs, useful advice, being a good paddle boarding and swimming partner, and driving me everywhere. Thanks for putting up with the outbursts and temper tantrums, as well as my terrible taste in music. I want to give a huge thank you to my family for their prayers, enthusiasm, and encouraging me to finish my PhD. Thanks to my brother, KB, and my cousins for the silly chats over the course of my PhD; from programming language to politics. Finally, I would like to thank Mama and Papa for their prayers, financial supports, and setting a good example on work ethics. I dedicate this thesis for you.

Contents

List of Figures

List of Tables

Abbreviations

ALOHA A Long-term Oligotrophic Habitat Assessment

AMT Atlantic Meridional Transect

BATS Bermuda Atlantic Time-series Study

CARIACO Carbon Retention in a Coloured Ocean

CEDA Centre for Environmental Data Analysis

CMIP6 Coupled Model Intercomparison Project 6

DCM Deep Chlorophyll Maximum

Diat-HadOCC Diatoms - Hadley Centre Ocean Carbon Cycle model

DIN Dissolved Inorganic Nitrogen

DIC Dissolved Inorganic Carbon

ENSO El Nino-Southern Oscilliation

RSEM European Regional Seas Ecosystem Model

FOAM Forecast Ocean Assimilation Model

HadCM3 Hadley Centre Coupled Model version 3

HadGEM2-ES Hadley Global Environmental Model 2- Earth System

HadOCC Hadley Centre Ocean Carbon Cycle model

HOTS Hawaii Ocean Time Series

GCM General Circulation Model

IPCC Intergovernmental Panel on Climate Change

JGOFS Joint Global Ocean Flux Study

MarMOT Marine Model Optimisation Testbed

MEDUSA Model of Ecosystem Dynamics nutrient Utilissation Sequestration and Acidification

MLD Mixed Layer Depth

NABE North Atlantic Bloom Experiment

NADR North Atlantic Drift

NaN Not a Number

NAO North Atlantic Oscillation

NECS Northeast Atantic Continental Shelf

NEMO Nucleus for European Modelling of the Ocean

NPZ Nutrient Phytoplankton Zooplankton model

NPZD Nutrient Phytoplankton Zooplankton Detritus model

NRR Normalised RMSE Ratio

PAP Porcupine Abyssal Plain

PlankTOM Plankton Type Ocean Model

PISCES Pelagic Interaction Scheme for Carbon and Ecosystem Studies

PBE Perturbed Biogeochemistry Ensemble

PCA Principal Component Analysis

PBPE Perturbed Biogeochenistry Physics Ensemble

PFT Phytoplankton Functional Types

PPE Perturbed Physics Ensemble

RMSE Root Mean Square Error

RLM Replaced Linear Mortality

SeaWiFS Sea-viewing Wide Field-of-view Sensor

SST Sea Surface Temperature

UK-ESM United Kingdom Earth System Model

Chapter 1

Introduction

1.1 Observing the Marine Environments

Human life depends on the ocean ecosystem for food, jobs, medicine, and recreation. The ocean also acts as a carbon sink, and therefore modulates the atmospheric CO_2, which are controlled by the microbial communities (Follows and Dutkiewicz, 2011). In order to understand how the ocean ecosystem works, empirical observations of the physics, chemistry, and biology have been made since the 19th century, pioneered by the Challenger Expedition between 1872-1876 (Hawkins et al., 2013). This was followed recently by other oceanographic expeditions such as the North Atlantic Bloom Experiment (NABE, 1989), Atlantic Meridional Transect (AMT, 1995-present), and Worldwide Marine Radioactivity Studies (WOMARS, 1995-2000). Following the growing realisation of the important role the oceans have in regulating the climate, the Joint Global Ocean Flux Study (JGOFS)), focusing on the ocean's carbon cycle and air-sea fluxes of carbon dioxide, was conducted. The study includes a global survey of oceanic properties, such as temperature, salinity, nitrogen, and chlorophyll, and also long term time-series observations at strategic sites. The Hawaii Ocean Time Series (HOTS) in the Pacific, Bermuda Atlantic Time Series (BATS) near the Sargasso Sea are two of the long term observation sites that have been instrumented and carried on until present. There are other long-term observation stations in the North Atlantic, such as the Western Channel Observatory, station L4 and E4, off the coast of Plymouth, UK. Furthermore, the advancement of technology has allowed satellites to observe sea surface temperature (SST), altimetry, and ocean colour. Compared to *in situ* observation, satellite remote sensing has provide data with higher frequency (1 day), larger coverage, and with high spatial resolution ($<$ 1 km) (Platt and Sathyendranath, 2008).

Long term regional observations have deepened our understanding about the biogeochemisty of the

ocean, for example the occurrence of two phytoplankton blooms, which have different phytoplankton group composition in station L4, (Smyth et al., 2010) (shown on Figure 1.1). Furthermore, multiple regional long term observation stations have made comparison study possible, showing for example how the mixed layer depth in the oligotrophic North Atlantic region is deeper during winter mixing than in the Pacific (Dave and Lozier, 2010). Long term observations can also capture trends of the physics and biogeochemistry of the ocean, for instance the increase in sea surface temperature over the past decades (Laffoley and Baxter, 2016; Doney et al., 2012), and also the increase in CO_2 concentrations over the past 17 years, that caused a decrease in the calcium carbonate saturation state at station ALOHA (Doney et al., 2009). Furthermore, using satellite remote sensing, changes in the global ocean can also be observed; a parallel increase of chlorophyll and SST have occurred mostly in the Indian Ocean between the 1980s to early 2000s, but in the Pacific ocean, an increase in SST and decrease in chlorophyll have been observed (Martinez et al., 2009; Blondeau-Patissier et al., 2014). Chlorophyll and SST observation using remote sensing can also distinguish regions where the change in chlorophyll is due to the change in biomass or due to physiological adjustment in intracellular chlorophyll concentrations, because SST covary with chlorophyll concentrations as it reflect light and nutrient conditions at productive regions. At the subtropical region, the increase in intracellular chlorophyll concentrations are dependent on the nutrient levels (Siegel et al., 2013).

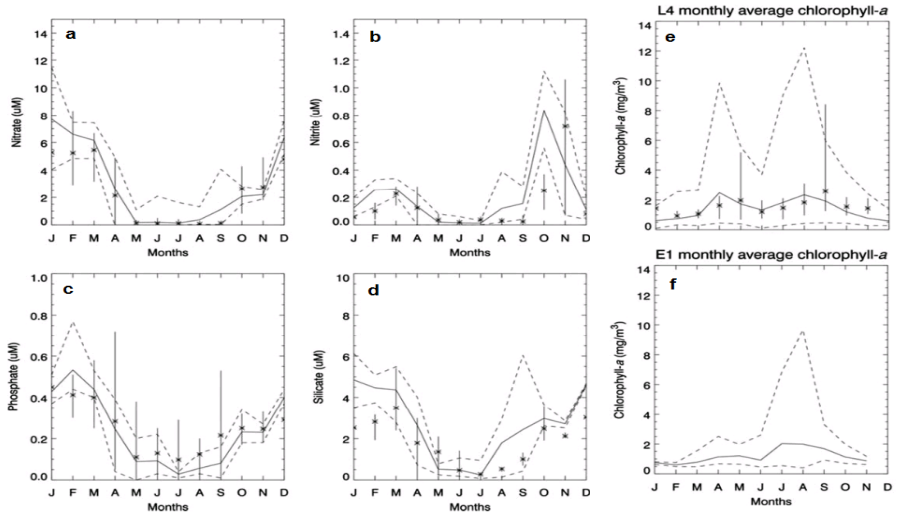


Figure 1.1: Monthly averaged surface nutrients at station L4 and E1 (a, b, c, and d) and phytoplankton concentration (e for L4 and f for E1). The solid line and dashed line on a, b, c, and d are the monthly averaged nutrients and the maximum and minimum concentrations respectively, at station L4. The asterisks and bars denotes the averaged nutrients and the minimum and maximum concentrations, respectively, at station E1. The solid line on e and f shows the mean monthly averaged nutrients at both stations and the minimum and maximum values are enveloped by the dashed line. In e, Satellite derived chlorophyll and its range at L4 are represented by the asterisks and bars, respectively. Figures are originally shown in Smyth et al. (2010)

In situ observations have also shown that the physics of the ocean controls the chemistry and therefore

the biology, for example, based on the 24 year time series, the change in marine fish communities in the eastern English Channel is influenced by the warming of the North Atlantic Ocean due to a switch in the Atlantic Multidecadal Oscillation from a cold to warm phase (Auber et al., 2015). Higher occurrence of coastal algal bloom in recent decade have also been reported to coincide with an increase of nutrient pollution (Parsons and Dortch, 2002; Heisler et al., 2008). The changes in upper ocean stratification caused by climate indices such as El Nino/ Southern Oscillation (ENSO) can influence the rates of primary production and export of particulate matter (Corno et al., 2007). The seasonal cycle of phytoplankton bloom can also influence other marine organism. With the help of satellite ocean colour and long term dataset of haddock recruitment, it is shown that the timing of maximum phytoplankton concentration (bloom peak) can influence the survival of fish larvae (Platt et al., 2003).

Studying the marine environment has not only been done from observations, but it has also been done by mathematical models (Gentleman, 2002). Different models can be developed depending on what scientific questions need to be addressed (Follows and Dutkiewicz, 2011). In order to explore further how the timing of the bloom affects the survival of fish larvae, it is possible to use an individual based model to quantify the survival rates of the fish. Using temperature reanalysis data, a modelling study shows that the increase in survival of fish larvae during warm years can increase to 154% in the North sea (Kristiansen et al., 2011). Using a model which explicitly represent phytoplankton, nutrients, and settling speed, it is possible to explain the occurrence of the high chlorophyll concentration below the mixed layer in the oligotrophic regions (Fennel and Boss, 2003; Cullen, 2014). A combination of plankton model and satellite-derived chlorophyll have been used to study the mechanism of the phytoplankton blooms (Behrenfeld et al., 2013; Behrenfeld and Boss, 2014). Using mathematical models, feedbacks and complex interconnections between the ecology and biogeochemistry can be untangled, and conceptual understanding can be quantified (Follows and Dutkiewicz, 2011). With growing concern over the increasing anthropogenic CO_2, eutrophication, and climate change, it is essential to synthesise observations and model in order to understand how the marine ecosystem may be affected by these factors, and therefore predict the state of the ocean, so that preventive measures can be devised.

Most marine biogeochemical models focus on the lower trophic levels of the ecosystem, where the first level of the marine food-web, such as phytoplankton and zooplankton, is explicitly represented, as the ecosystem is impacted first by changes in the physics or biogeochemistry. The first marine biogeochemical model was developed by Fleming to study the temporal variability of phytoplankton concentration under grazing pressure using differential equations (Fleming, 1939). The model showed that if phytoplankton density and growth rates in the field are known, it is possible to estimate the grazing rate. This simple model is then refined by Riley, who observed plankton production in Georges Bank (Gentleman, 2002). From the observation, Riley further improved Fleming's model by making the phytoplankton

growth rate dependant on the abundance of nutrients, zooplankton, temperature, light, mixing, and euphotic depth, which was the first coupled biogeochemical-physical model of the marine environment (Riley, 1946). Furthermore, instead of using the observed population and nutrient density as an input to run the model Riley separates the nutrients from the biology (phytoplankton and zooplankton) in a food-chain-like manner (Riley et al., 1949). The nutrient concentration is dependent on phytoplankton consumption and plankton respiration. The zooplankton abundance is the balance between phytoplankton grazing, predation from higher predator and losses due to natural death. This three compartment model is the basis of all of the marine biogeochemical models today.

1.2 Early marine biogeochemical model: the NPZ model

The model type nutrient-phytoplankton-zooplankton (NPZ) covers the simplest processes that describe the plankton dynamics. This type of model usually uses nitrogen as the currency as it is the limiting macronutrient to phytoplankton growth. This model consists of phytoplankton growth due to light, or light limited growth (L_I), nutrient uptake, or nutrient limited growth (U_N), zooplankton grazing (G_z), phytoplankton (ρ), and zooplankton (ζ) mortality terms. When constructing the simple NPZ, we need functional forms that describe these processes, and are dependent on the state variables (which are the nutrient, phytoplankton, and zooplankton), parameter values, physical oceanography information such as mixed layer depth or horizontal and vertical velocity which supply nutrients, and the initial conditions. This section will discuss the different equations that have been used in NPZ models, 1.2.1 and the model inputs that are needed in order to run the model, 1.2.2.

1.2.1 NPZ model structure, equations, and parameter

The NPZ model equations can be generalised as:

$$\frac{dP}{dt} = (L_I \times U_N)\,P - G_Z P - \rho P \tag{1.1}$$

$$\frac{dN}{dt} = \rho P + \zeta Z + (1-\lambda) G_Z P - U_N P \tag{1.2}$$

$$\frac{dZ}{dt} = G_z P Z - \zeta Z \tag{1.3}$$

Which describe the rate of change in phytoplankton (P), nutrient (N), and zooplankton (Z) with time. The general equations also shows how the changes in each compartment affect each other. Unlike a physical oceanography model which has the navier-stokes equations, the biogeochemical model does not have a known set of equations that describe different processes, which allows modellers to choose or

derive their own functional forms to describe each biogeochemical process, as long as the model fits with experimental data (Flora et al., 2011). All of the governing process functional forms range from linear equations, that describe constant rate of process (density-independent), to nonlinear forms, where at first the rate increases with concentrations, and then saturates when it reaches the maximum rate (density-dependent) (Franks, 2002), as shown in Figure 1.2. Nonlinear forms have both a maximum rate (μ) and a shape-defining coefficient (usually a half saturation constant, k), and linear functions only have a constant rate. This makes the choice of functional form crucial as it can affect the dynamics of the model, and the choice of the parameter values that dictate the rate of grazing or nutrient uptake rates at a certain concentration. The next subsections will discuss the different functional forms and their parameters for each governing process within an NPZ model.

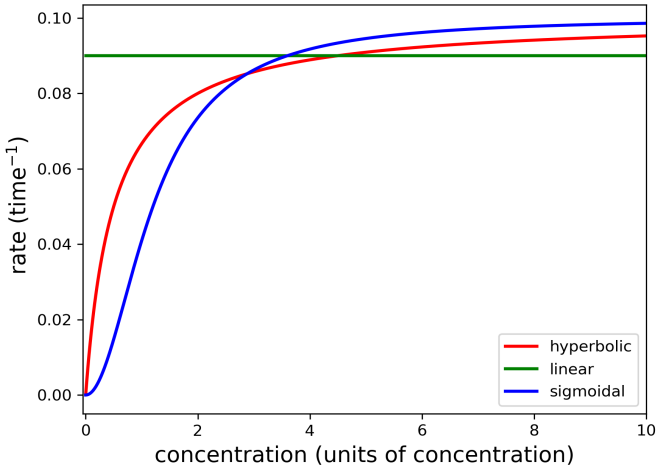


Figure 1.2: Common shapes of functional forms that describe the key processes in biogeochemical models. Saturating response is shown in hyperbolic and sigmoidal.

Phytoplankton Growth (L_I and U_N)

From equation 1.1, the phytoplankton growth is limited by light and nutrients. Phytoplankton growth rate in response to irradiance can be calculated using linear, saturating (using Hyperbolic, trigonometric, and exponential functions), or saturating with photoinhibiting response, summarised in Table 1.1, with a parameter to determine the irradiance at photosynthesis maximum, I_o (Franks, 2002). In a more complex NPZ model, where the mixed layer depth is taken into consideration (which assumes that the biological aspect of the model sits above the deep homogeneous layer with constant nutrient and no plankton (Anderson et al., 2015)), the total growth due to light has to be averaged over the mixed layer, described by:

$$J(t,M) = \int_0^{24h} \int_0^M L_I(I(t,z))dzdt \tag{1.4}$$

where J is the light limited phytoplankton growth, t is the time (days), M is the mixed layer depth, and $24h$ is the day length. The linear and saturating response functions can be chosen to match observed change in phytoplankton light limited growth rate over one day or how well they represent the nonlinear response of photosynthesis to irradiance (Franks, 2002). In order to make the functional forms described in Table 1.1 more similar to the laboratory experiment, instead of only using the term I_o, it is possible to split this term into two different parameters, such as in Smith (1936) function, which is similar to a hyperbolic saturating response. In Smith (1936) function, the term α/V_p^{max} is used to describe I_o, where α is the initial slope of the P-I curve, and maximum photosynthesis rate is described using V_p^{max}. Similarly, the I_o term in exponential function can also be split into α/V_p^{max}, and this have been used in the recent NPZ model to described the light limited growth of phytoplankton (Anderson et al., 2015).

The most common function to represent the nutrient uptake is the logistic saturating rectangular hyperbolic, also known as the Michaelis-Menten or Monod formulation, based on enzyme kinetics. As shown in Figure 1.2, this equation has a shape defining parameter, k_u, and a maximum rate, in V_u^{max}. This kind of function is computationally cheap as it is not explicitly linked to resource availability (Flynn, 2018). However, phytoplankton have been shown to store nutrients in an internal pool before they are used for growth (Droop, 1973, 1983). The quota (Q) describes the amount of a substance in a unicellular organism, and this controls the growth of phytoplankton. Q varies between the subsistence quota (Q_{min}), in which it is impossible to use the nutrient for growing, and a growth maximum rate (Q_{max}) (Franks, 2002). It should be noted that the Q_{max} here is not for the final phytoplankton growth as this is also limited by light, as described in equation 1.1.

Zooplankton Grazing (G_z)

In a simple NPZ model, zooplankton would only graze one type of prey, the phytoplankton. Most of the formulations are based on laboratory experiment where populations of predators are acclimatised to different prey densities (Holling, 1959). From these experiments, there are four 'Types' of zooplankton grazing; (i) linear variations of grazing rate with prey, (ii) curved variation grazing rate with prey (Type II), which used rectangular hyperbolic function, and are often called Holling type II, (iii) another curved variations of grazing, which is similar to the sigmoidal curve, described in Figure 1.2 and is often referred to Holling type III, and (iv) grazing rate that reaches maximum at an intermediate density and decreases as the density gets higher, which mimics when the prey is toxic or the predator gets confused (Holling, 1959; Gentleman et al., 2003). The difference between type II and III is that the grazing rate in the former is higher at lower density prey concentration, compared to the latter, as shown in Figure 1.2. However, Flynn and Mitra (2016) argued that these hyperbolic functional forms do not explicitly represent the

effect of physical oceanography processes (such as turbulence) when encountering prey, or how the motility of zooplankton changes when satiation or saturation occurred.

The first type of grazing has no satiation, and is formulated using a linear function, but there is no theoretical basis to this equation. The second type is described using a Disk Equation, which is based on predator-prey theory, and defined by handling and attack rate. The Michaelis-Menten (rectangular hyperbolic equation), is mathematically equivalent to the Disk equation, however instead of defining the equation using attack, a, and handling time, h, rectangular hyperbolic has maximum rate (g_{max}) and half saturation constant (k_g), where in this case, g_{max} becomes $\frac{1}{h}$, and k_g becomes $\frac{1}{ah}$ (Gentleman et al., 2003). The Michaelis-Menten type function has been used in early NPZ models (Fasham et al., 1990; Steele and Henderson, 1992) and even a more complex model with various types of phytoplankton (Follows et al., 2007; Prowe et al., 2012), rather than the actual Disk equation. Similar response curves can also be formulated using Ivlev functional forms (Ivlev, 1961), however the rate of change in Ivlev is different to the Disk equation, and has a shape defining functions λ instead of k_g or a and h (Gentleman, 2002). The third type can also be described using a Disk equation, however, the attack rate would vary linearly with prey density, according to a constant (a= constant $\times$ prey). Using a Michaelis-Menten type equation, this could be described using a sigmoidal (Holling type III) equation, with $g_{max} = 1/h$ and $k_g = 1/\sqrt{ch}$. The Holling type III function is more commonly used in later biogeochemical models (Fasham et al., 1993; Edwards and Brindley, 1996; Edwards and Yool, 2000; Palmer and Totterdell, 2001; Anderson et al., 2015), and is referred as Holling type III. The fourth type can be described using Holling type II, but with an additional term in the denominator that results in maximum and half saturation constants depending on complicated functions of other parameters (Gentleman et al., 2003). These equations are described in Table 1.1.

Plankton Mortality (ρ and ζ)

The mortality terms in the NPZ model allow the recycling of nutrients from phytoplankton and zooplankton. This is usually modelled using either linear, quadratic, hyperbolic, or sigmoidal functions. The linear mortality has a constant rate independent of plankton concentrations, which may represent higher predators when this is used for zooplankton, whose biomass does not fluctuate (Edwards and Yool, 2000). In an NPZ model, the phytoplankton mortality term is usually modelled using a linear or a quadratic function (Steele and Henderson, 1981, 1992; Fasham et al., 1990; Fasham, 1995; Edwards and Brindley, 1996), although it is possible to represent phytoplankton mortality using a Michaelis-Menten type equation (Fasham et al., 1993). However in a more recent NPZ model, such as that of Anderson et al. (2015), both linear and quadratic phytoplankton mortality is used. This is because the linear mor-

tality account for metabolic losses or natural mortality, and the density dependent loss (non-linear) may represent mortality due to infection or viruses (Palmer and Totterdell, 2001; Yool et al., 2011, 2013; Anderson et al., 2015).

Zooplankton mortality terms have been shown to alter model dynamic more than altering the parameter values when the equations are altered. Similar to phytoplankton mortality, in an NPZ model, this process is usually represented by linear mortality, however instead of producing a steady state, it produces oscillations over a wider parameter range (Edwards and Yool, 2000; Franks, 2002). The range of different functional forms for zooplankton mortality may be due to the various feeding strategies of zooplankton. When cannibalism occurs within the zooplankton compartment, or if the population of higher predators changes proportionally with zooplankton, the mortality term may be represented using quadratic mortality. This functional form will have a mortality rate that depends on the zooplankton biomass. The hyperbolic and sigmoidal forms mimic the higher predator grazing strategy (Edwards and Yool, 2000), similar to the zooplankton grazing. Non-linear mortality functions do not produce limit cycles (Steele and Henderson, 1992), and therefore are more preferred. When compared to in situ data, using quadratic mortality produces the largest deviations compared to sigmoidal and hyperbolic (Mitra, 2009). Although it is suggested that using only one mortality term is inappropriate and necessary in a planktonic food web model as the description of zooplankton loss processes is biologically inaccurate (Mitra, 2009), more recent simple NPZD model (e.g. Anderson et al. (2015)), and even more complex biogeochemical model such as Palmer and Totterdell (2001); Dutkiewicz et al. (2009); Halloran et al. (2010) still use these density dependent mortality functions, whose equations are summarised on Table 1.1.

1.2.2 Coupling NPZ model to a physical model

We also need to consider the physical oceanography aspect before running an NPZ model, as both light and nutrients limit phytoplankton growth and are dependent on the physical oceanography. In the ocean, nutrient is generally stored in the deeper layers of the water column, therefore a physical transport mechanism is needed. One of the most important physical inputs is the mixed layer depth, which marks the depth at which temperature, salinity, and density is generally constant throughout the top layer. This depth affects the changes in nutrient and light availability for phytoplankton in the upper layer. For example, in high latitudes, the depth of the mixed layer is relatively deep in the winter due to cooling and wind mixing, resulting in high nutrient concentrations in the surface waters. Due to the deep mixing, the phytoplankton cells are moving vertically from the shallow to the deeper part of the water column. However since the day length is short and heating is absent, despite being near the surface, the growth of phytoplankton is low. In spring, when day length is a longer, and the mixed layer becomes shallower,

Table 1.1: Functional forms that describe the main processes of an NPZ model.

Process	Functional form	Description
Light limited growth	$\frac{I}{I_0+I}$	Hyperbolic (Saturating response)
(L_I)	$\tanh\left(-\frac{I}{I_0}\right)$	Trigonometric (Saturating response)
	$1-\exp\left(-\frac{I}{I_0}\right)$	Exponential (Saturating)
	$\frac{I}{I_o}\exp\left(-\frac{I}{I_0}\right)$	Exponential (saturating and photo inhibiting response). Parameter I_0 irradiance at photosynthesis maximum
	$V_p^{max}\left(1-\exp\left(-\alpha I/V_p^{max}\right)\right)$	Exponential, but I_o is replaced with α/V_p^{max}
	$\frac{V_p^{max}\alpha I}{\sqrt{(V_p^{max})^2+\alpha^2 I^2}}$	Smith (Saturating response)
Nutrient limited growth	$\frac{V_n^{max}N}{k_s+N}$	Michaelis-Menten (Saturating response)
(U_N)	$V_n^{max}\frac{(1+KQ)(Q-Q_{min})}{(Q-Q_{min})+KQ(Q_{max}-Q_{min})}$	Droop cell quota, KQ is a shape defining parameter
Zooplankton grazing	$g_{max}P$	Type I Linear, g_{max} is maximum grazing rate
(G_z)	$g_{max}\frac{P}{P+k_g}$	Holling type II (Saturating response), k_g is half saturation constant
	$\frac{aP}{1+ahP}$	Disk Equation Type II (Saturating response), where a and h are the attack and handling, respectively
	$g_{max}[1-exp(-\lambda P)]$	Ivlev (Saturating response), λ is the shape defining parameter
	$g_{max}\frac{P^2}{P^2+k_g^2}$	Holling type III (Saturating response)
	$\frac{\hat{a}P}{1+\hat{a}hP}$	Disk Type III (Saturating response) $\hat{a}=cP$, where c is constant
	$\frac{P}{\kappa+P+\beta P^2}$	Type IV (Threshold), where κ and β, are parameters that controls the decline of population
Plankton Mortality	μ	Linear
(ρ, ζ)	μX	Quadratic
	$\mu\frac{X}{X+k_X}$	Hyperbolic (Saturating response)
	$\mu\frac{X^2}{b^2+X^2}$	Sigmoidal (Saturating response). X is the concentration of zooplankton or phytoplankton.

the phytoplankton concentration increases, as well as the zooplankton, (Behrenfeld and Boss, 2014). In coastal and other open ocean systems, the nutrients are injected to the upper layer by upwelling, or vertical currents that move upwards (Fennel and Neumann, 2014).

If the NPZ or a more complex biogeochemical model is coupled with a physical model using the advection-diffusion equation, the change in concentrations (C) for a state variable in a 3D system is described by:

$$\frac{dC}{dt} = -(w+w_s)\frac{dC}{dz} - u\frac{dC}{dx} - v\frac{dC}{dy} + \kappa_h\left(\frac{d^2C}{dx^2}\frac{d^2C}{dy^2}\right) + \kappa_v\frac{d^2C}{dz^2} + \text{biological equations} \tag{1.5}$$

The physical components from this equation are determined by the physical input or the ocean circulation model. u, and v, are the horizontal, and w the vertical, velocities and w_s is the vertical swimming or sinking speed of the state variable, κ_v and κ_h are the vertical and horizontal eddy diffusivities, which control the mixed layer. If the model is run in 1D, only the first $\left(-(w+w_s)\frac{dC}{dz}\right)$ and last term $\left(\kappa_v\frac{d^2C}{dz^2}\right)$ of the equations are going to be used. Apart from using the advection diffusion equations, the physical processes can be parameterised even more simply by using a slab ocean mixed layer model (Parekh et al., 2005; Anderson et al., 2015). This physical model has a simple structure which consist of two vertical layers. The depth of the upper layer, which represents the mixed layer, can vary seasonally and can be determined from observations. The bottom layer is an unchanging bottom layer that contains fixed nutrient concentrations (Anderson et al., 2015). This slab model is zero-dimensional and more computationally efficient than using advection-diffusion schemes. It is also possible to permanently set the mixed layer depth, similar to that of Steele and Henderson (1981); Edwards and Brindley (1996), where mixing is parameterised as $P_{max}/(b+cP)M$, where P_{max} is the maximum phytoplankton growth under optimum light condition, M is the mixed layer depth, b is the light attenuation by water, and c is the phytoplankton shelf-shading coefficient. This approach to parameterising the physical forcing is usually used in simple NPZ models to conduct sensitivity studies.

Apart from coupling the biogeochemical equations to the advection-diffusion equation, we need another input to initialise the model, this is called the initial conditions. This input can be obtained from in situ observations or from spinning up the model, where the model is run for a long time, using either an ocean general circulation model output, or physical climatology input until it has reached its steady state. This approach is used because the nature of the system is often excitable and results in high chlorophyll and zooplankton concentrations at the first few days or time of the model run (Edwards and Brindley, 1996). After embedding the biological equations (described from equations1.1 to 1.3) to the advection-diffusion equation and deciding on the initial conditions, the NPZ model can now be run (Steele and Henderson, 1981; Murray and Parslow, 1999; Edwards and Brindley, 1996; Edwards and Yool, 2000).

As computers become more powerful, apart from seeking to represent observations at particular sites (Fasham et al., 1990; Robinson et al., 1993), NPZ models can also used to conduct sensitivity studies. Conducting sensitivity analysis in a simple model is important in order to investigate the range of model behaviour, before moving to a more complicated scenario (Franks, 2002), such as when adding different types of nutrients, such as ammonia (NH_4) and nitrite (NO_2), or before coupling with a physical model. These studies include: varying the biogeochemical parameters and how they affect the model dynamics (Ruan, 2001; Edwards and Brindley, 1996; Edwards, 2001), the range of possible ecosystem behaviour before coupling to a physical model (Franks, 2002), and what will happen if different functions are used to describe a biogeochemical process (Steele and Henderson, 1981; Murray and Parslow, 1999; Franks, 2002; Edwards and Yool, 2000). However, the lower trophic ecosystem is more complicated than the NPZ model represents, therefore adopting better model formulations to represent key processes and system feedbacks is necessary (Raick et al., 2006).

1.2.3 More complicated biogeochemical models

Riley improved Fleming's model by adding the nutrient compartments and separating the nutrient field from phytoplankton and zooplankton. As computing power and knowledge of the ocean biogeochemistry increases, the addition of more model compartments to a marine biogeochemical model becomes more common, but the core structure of lower trophic levels remains. For example the addition of detritus, which facilitates the export of organic material back to the nutrient pool, and bacteria to quantify the nutrient fluxes through the microbial loop (Fasham et al., 1990, 1993; Palmer and Totterdell, 2001) have been included into marine biogeochemical models, making the NPZ model into an NPZD. Additionally, different phytoplankton have different biogeochemical roles in the marine environment. For example, calcifiers influence atmospheric CO_2 through ocean alkalinity and carbonate chemistry, which makes these phytoplankton types essential when addressing ocean acidification. Other types, such as diatoms, the silicifiers, are important in terms of carbon cycle, as this phytoplankton type contributes more to the carbon export than smaller phytoplankton (Le Quere et al., 2005). Representation of other macronutrients apart from DIN is therefore necessary to include in a model such as silicate (Yool et al., 2011, 2013) and phosphate (Butenschön et al., 2016; Aumont and Bopp, 2006; Le Quere et al., 2005) as different phytoplankton functional types (PFTs) require different nutrients to grow. Furthermore, there are regions where macronutrients are abundant, for example in the North Pacific, where nitrate concentration is high, but the surface water is dominated by picoplankton, which is usually found in nutrient poor water, then according to laboratory experiment, the phytoplankton growth is limited by iron (Martin and Fitzwarer, 1988). Due to this early study, iron is often included in the model in order to permit spatial variability in phytoplankton concentrations (Yool et al., 2011, 2013; Butenschön et al., 2016; Le Quere et al., 2005).

As the marine environment is rapidly changing, marine biogeochemical models have been devised to be used as a tool to understand how the ocean ecosystem responds to climate change (Bopp et al., 2005) and to project trends in ocean ecosystems under the changing global environment (Bopp et al., 2013; Gehlen et al., 2015). This is because the microbial communities are regulating the environment and export organic matter to the deep ocean, which is critical for the global atmospheric CO_2 sequestration and modulation (Follows and Dutkiewicz, 2011). Since most of these models explicitly represent lower trophic levels, these can also be used to monitor the health of ocean ecosystems, and assess the potential impact of climate change on higher trophic levels (Stock et al., 2011; Gehlen et al., 2015). Sudden changes in the ocean environment would affect various biogeochemical pathways, therefore it is necessary to include more biogeochemical processes or model compartments. Models that include dissolved inorganic carbon (DIC), alkalinity, and different PFT, such as coccolithophores may be able to address some questions regarding how ocean acidification may affect marine calcifiers, the effect of increasing CO_2 emission and stratification on the primary production, and export of particulate matter (Steinacher et al., 2009; Yool et al., 2013). Therefore, the information from future projections of these complex models can be provided to policy makers or fisheries managers (Hyder et al., 2015). Biogeochemical models, which results can be provided to end users are called operational models (Ford et al., 2012). The operational model may be embedded in an earth system model, so that biogeochemical feedbacks can be simulated during climate change scenarios. These include some questions about the contribution of the ocean to the atmospheric trace gas composition in changing climate and the release of methane hydrate that could be triggered by climate change (Kwiatkowski et al., 2014).

Operational biogeochemical models vary in complexity, but the model currency is mostly nitrogen, similar to the simpler NPZ models. The Diatom Hadley Centre Ocean Carbon Cycle model (Diat-HadOCC) (Halloran et al., 2010), which is the ocean biogoechemistry component of HadGEM2-ES, the Met Office's Earth system model, has nitrogen as the model currency, and has two different phytoplankton types, diatoms and 'other phytoplankton'. Apart from the two plankton types, zooplankton, and dissolved inorganic nitrogen (DIN), this model explicitly represents DIC, total alkalinity, silicate, and dissolved iron, and for cloud feedbacks, a dimethyl sulfide submodel is added (Cox and Kwiatkowski, 2013). This model is developed from a simpler model HadOCC, which is essentially an NPZD model with DIC and alkalinity explicitly represented. Although simpler than Diat-HadOCC, this model was used as the first ever coupled carbon-climate study in the UK, by embedding in it to the physical coupled general circulation model (GCM) HadCM3 (Cox et al., 2006).

Another global marine biogeochemical model with similar structure is the Model of Ecosystem Dynamics nutrient Utilisation Sequestration and Acidification (MEDUSA-2.0), preceded by MEDUSA-1.0, which is described as 'beyond the standard NPZD model formulations' (Yool et al., 2011). Al-

though similar to the Diat-HadOCC with DIC, alkalinity, silicate, and diatoms explicitly represented, in MEDUSA-2.0 phytoplankton and zooplankton are separated into two large (diatoms and mesozooplankton) and small (non-diatoms and microzooplankton). MEDUSA also has dynamic Silicon:Nitrogen and Carbon:Chlorophyll ratios, unlike Diat-HadOCC. As explained in the previous paragraph, besides expanding plankton into different types or sizes, the nutrients can also be expanded. For example, a similarly structured model with two phytoplankton and zooplankton types, the Pelagic Interaction Scheme for Carbon and Ecosystem Studies (PISCES), have expanded its DIN compartments into ammonium and nitrate, as well as adding phosphate to the model compartments (Aumont and Bopp, 2006). These models have good levels of complexity that is close enough to the observations and can address some important questions about climate change. For example the PISCES model has been used to address the iron fertilisation questions and how it may affect CO_2 levels if the experiment done for a long term (100 years), and found that iron fertilisation does not reduce as much CO_2 as previously thought (Aumont and Bopp, 2006). Using MEDUSA, it is possible to predict the shift of bloom timing in response to the projected 21st century climate change (Henson et al., 2017). With the representation of a dimethyl sulfide submodel, Diat-HadOCC embedded in the HadGEM2 model is able to simulate how increasing albedo can affect the mean radiative forcing (Crook et al., 2016).

In order to address further questions and have better representation of real ocean ecosystems, some models include even more processes, ecosystem representations, and PFTs. For example, the European Regional Seas Ecosystem Model (ERSEM) was initially designed for shelf seas, but has expanded to a global ocean. The model is essentially still an NPZD model, but in its most recent version, it has four phytoplankton and three zooplankton functional types, which are explicitly represented. Additionally, there are five nutrient compartments, along with carbon, oxygen, and alkalinity. The biogeochemical rates are computed within the pelagic and benthic systems, for oxygen and carbon the gas transfer is computed on the sea surface, and the fluxes across the seabed (Butenschön et al., 2016). Overall, ERSEM has 57 model tracers, shared between the different ecosystem representations. For global models, the addition of different PFTs is also important to represent different phytoplankton types in different oceanic regions. The PlankTOM10 model has five phytoplankton types including diatoms, coccolitophores, mixed phytoplankton, nitrogen fixers, and picophytoplankton, and three zooplankton types, protozooplankton, mesozooplankton, and macrozooplankton (Le Quere et al., 2005). Similar to ERSEM, PlankTOM10 explicitly represents carbon, nitrogen, oxygen, phosphate, silicate, and iron cycles, along with air-sea fluxes of CO_2, oxygen, dimethyl sulfide and nitrous oxide. With more representations of phytoplankton and zooplankton, it represents global satellite-derived chlorophyll concentration better than simpler models (such as Diat-HadOCC and MEDUSA-2.0) (Kwiatkowski et al., 2014). More complex models are also able to reproduce emergent relationships, for example, ERSEM can reproduce higher diatom fractions

in the northern part of the North Atlantic better than simpler models (Holt et al., 2014), as well as the distribution of the phytoplankton community structure (De Mora et al., 2016).

An ocean biogeochemical model is also needed for the next generation of UK Earth System Model (UKESM1). In order to find a suitable model, an inter-comparison project was conducted. In this project, six UK marine biogeochemical models were assessed based on their ability to reproduced bulk properties of the ocean, such as nutrient and carbon distributions (Cox and Kwiatkowski, 2013; Kwiatkowski et al., 2014), using the same ocean GCM and their computational cost. This study was done in a global 3-D setting, and explored biogeochemical dynamics from an NPZD model (Palmer and Totterdell, 2001) to multiple PFTs (Le Quere et al., 2005). From this project, simpler models with fewer state variables showed better agreement with observational biogeochemical data (such as CO2, dissolved inorganic nitrate, and alkalinity). From these results, the inter-comparison project recommended MEDUSA-2.0 to be adopted as the ocean biogeochemical component for UKESM1, given the model shows relatively good performance and offer a good compromise between complexity and computer time (Kwiatkowski et al., 2014; Cox and Kwiatkowski, 2013).

The problem with complex models is that some of the ecology is still poorly understood, for example the optimum conditions for coccolitophore blooms, and the justification to divide zooplankton into different sizes (Anderson, 2005), which makes some of the PFT representations lack robustness. Furthermore, as discussed in section 1.2.1, some of the functional forms that represent biogeochemical processes are biologically inaccurate (Flynn and Mitra, 2016; Mitra, 2009). There is also no consensus on which processes should be represented in a biogeochemical model. Different physical models can also give alternative simulations (Doney et al., 2009; Sinha et al., 2010). Additionally, parameter values have to be estimated by fitting the model to data (Ward et al., 2013; Spence et al., 2016), and therefore this adds another source of uncertainty.

1.3 Sources of uncertainties

Marine biogeochemical models are uncertain for many reasons; there is limited knowledge of the biogeochemical and physical processes and how they work, limited ability to parameterised the known processes into the models, and limitations in the in situ observations to provide for model initialisation, parameter optimisation (making the model simulation more similar or closer to the observations by changing the parameter values), and validation. Therefore, when constructing a marine biogeochemical model, we need a few things to consider:

- Structure: which includes how many model compartments are represented in the model (complex-

ity) and what model equations or functional forms are chosen to represent the biogeochemical processes.

- Parameters: what parameter values are going to be assigned to the model formulations.
- The physical input: which physical model or input are we going to use to force the biogeochemical model.

These area also the components that contribute to the model uncertainty. These uncertainties are interlinked, as they often emerge because of one or another, for example, the higher the model compartments (complexity), the more uncertain the parameters are, and the harder it is to optimise without sometimes resulting unrealistic parameter values (Friedrichs et al., 2006). In a multi PFTs biogeochemical model, some processes are poorly understood, and the lack of in situ data, means parameter optimisation for a particular model compartment would not be feasible. Earlier studies have also shown that complex biogeochemical models are sensitive to GCMs because of the difference of mixing strength between different GCMs (Sinha et al., 2010).

1.3.1 Parametric and structural uncertainty due to model complexity

Each of the functional forms in the model equations contain parameters (or constants), which are adjustable, as these values are sometimes obtained from tuning the model with observational data or from laboratory experiments (Anderson et al., 2015). However, the corresponding parameter values are variable in different regions, times, and across taxa (Hemmings and Challenor, 2012). Indeed, model parameter values are mostly constants, however adding dependence on time and space will require more parameters to be adjusted. These parameters can be varied in order to analyse how they affect the model output dynamics, as some parameters can affect model results more than others. This practice is called sensitivity analysis, and is usually carried out in modelling studies before deciding the correct parameters to be assigned. For example, in an unforced NPZ system (the model is run without time-varying physical input) varying maximum zooplankton mortality rate can produce oscillations in model dynamics, instead of reaching a steady state, due to bifurcation (Edwards and Brindley, 1996; Edwards, 2001; Denman, 2003). Increasing the zooplankton mortality rate from 1 to 1.5 day^{-1}, in a similar model, would produce oscillations between the N, P, and Z concentrations, with a period of 35 days, instead of a steady state (the NPZ compartments becoming constant), as shown in Figure 1.3(a) and (b) (Edwards and Brindley, 1996). Adding model compartments also alters the steady state conditions, summarised in Figure 1.3(c) and (d). When detritus is added to the model compartments and is grazed by zooplankton, the oscillations still occurred, when the rate is changed to 1.25 day^{-1}, with the period of oscillation 59

days (Edwards, 2001). This shows that oscillations still occur when the zooplankton mortality parameter is varied, even though a model compartment is added, although the period of oscillations changed.

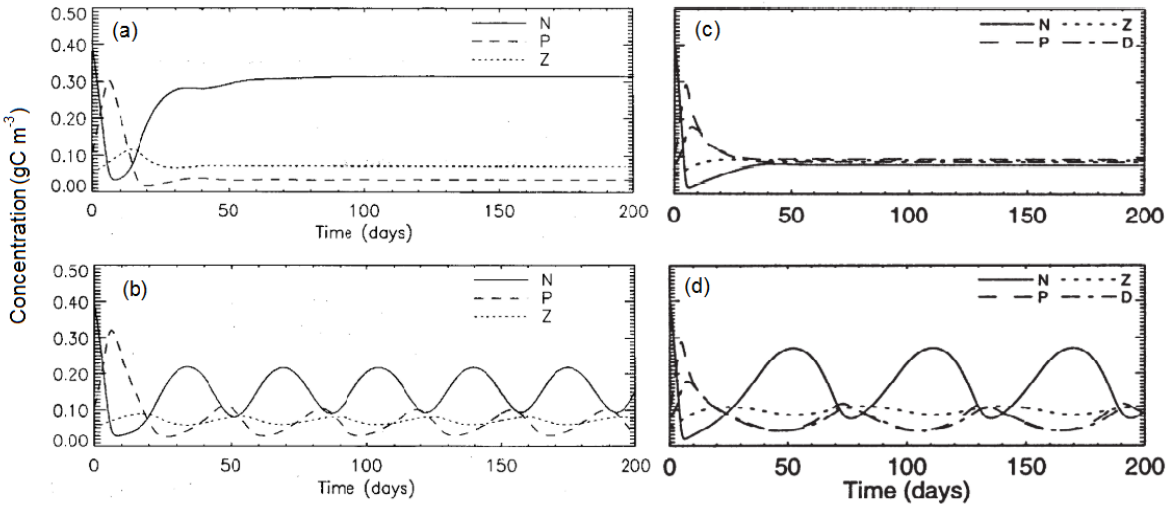


Figure 1.3: Time series of nutrient, phytoplankton, and zooplankton (a and b), and detritus (c and d) output from and NPZ and NPZD model respectively. Subfigure a, shows the steady states in an NPZ model when maximum zooplankton mortality d=1.0 day^{-1}, b shows the oscillations between model compartments d=1.5 day^{-1} (these subfigures are obtained from Edwards and Brindley (1996)). Subfigure c shows the steady states in the NPZD model when d=1.5 day^{-1}, and d shows the unforced oscillations when d=1.25 day^{-1} (These subfigures are obtained from Edwards (2001)).

Since parameter values are adjustable, it is possible to use optimisation techniques to adjust model parameters to get closer to observations. However, parameter fitting can hide underlying model problems, such as biologically inaccurate model equations (Anderson and Mitra, 2010). Additionally, during the optimisation process parameter values that have been obtained could be unrealistic, with parameters outside the observation or experimental range (Anderson et al., 2015). Calibrated parameters that are different from the initially hypothesised values make it difficult to assess whether it is reasonable to to use them or whether they compensate for an error in another parameter (Robson et al., 2018). Even after optimisation, it is important to ensure that the model is giving a good fit to the observation for the right reasons (Anderson, 2005; Friedrichs et al., 2006). Parameter sensitivity is closely linked to model complexity. Complex models are less sensitive to varying parameters (Friedrichs et al., 2007; Kriest et al., 2010, 2012), as the higher the number of model compartments the more parameters are constraining the model dynamics. Yet not all parameter values can be optimised using observations, as this can result in unrealistic values and some of the parameter values cannot be obtain from observation, such as mortality rate (Ward et al., 2010). There are also parameters, such as sinking rate, plankton density independent mortality, and remineralisation of small detritus, that can be removed without changing the model's ability reproduce biogeochemical tracers and primary productivity (Ward et al., 2013). However, the optimum parameters are specific to the assimilated observation and physical environment, showing that the optimum parameters are generally not portable (Ward et al., 2013), which is problematic for

a global 3-D model. This may mean that for diagnostic purposes simpler model performed better, but not necessarily for operational purposes as they do not represent all the necessary biological processes (Ward et al., 2013). Although complex models can reproduce better surface nutrients, in the global ocean interior simpler models are also able to reproduce oxygen and nutrient distributions (Kriest et al., 2012). Furthermore, in a model inter-comparison project, more complex models do not always reproduce bulk in situ observations, such as total DIN and DIC, better than the simpler models (Kwiatkowski et al., 2014). This leads to a conclusion that the discrepancy between more complex model and the observation is due to poorly represented physical modelling (Friedrichs et al., 2006; Hemmings and Challenor, 2012; Ward et al., 2013).

1.3.2 Uncertainty due to physical forcing input

The physical environment controls the growth of phytoplankton by providing mixed layer depth, temperature, irradiance, which affect the growth rate of phytoplankton, and vertical velocity which brings the nutrients to the euphotic zone (Doney, 1999). Physical forcings for marine biogeochemical model are often obtained from ocean GCMs. Similar to the biogeochemical models, ocean GCMs also have deficiencies, such as some physical processes are not explicitly embedded in the model equations, uncertain initial or boundary conditions, structural errors in the formulations, and uncertain model parameters, such as mixing and diffusion rates (Griffies et al., 2000).

Uncertainties in these fields can affect various biogeochemical processes such as nutrient supply and vertical transport of dissolved organic matter (Doney, 1999; Najjar et al., 2007). Using assimilated NEMO vertical velocity and mixing data, instead of from NEMO output has also been shown to enhance primary production, but may overestimate nutrient concentration, due to the overall upwelling resulted from the assimilated data (Raghukumar et al., 2015). Furthermore the resolution of a model can also affect the biological compartments, as it has been shown that phytoplankton production is affected by sub-mesoscale processes (Lévy et al., 2001). For example in the oligotrophic gyres, in order to enhance productivity in the biogeochemical model, nutrient supply can be increased by introducing eddy-induced nutrient pumping, which requires super high resolution physical model (Popova et al., 2006; McGillicuddy et al., 2003). Furthermore when a complex marine biogeochemical model with multiple PFTs is run using two different ocean circulation models (NEMO and OCCAM), although the bulk properties, such as the total phytoplankton biomass, do not show much difference, the distribution of PFTs between the two runs can be markedly different, due to the different strength of the mixing processes (Sinha et al., 2010). This discrepancy between simulations may be due to complex models exhibiting a wide range of behaviours, which are not easily predictable before actually running the model (Sinha et al., 2010).

There have been some uncertainty studies done especially for marine biogeochemical models, which examine the effect of parameteric, complexity, and physical model uncertainties to the model results. The next subsection will briefly summarise some of these studies and what has been found.

1.3.3 Examining uncertainties in marine biogeochemical models

As discussed in section 1.3.1, one way of examining uncertainties in a biogeochemical model is to vary the parameter values. A different approach was tried by Steele and Henderson (1992), instead of varying the parameter values, the equation to describe zooplankton mortality was varied between linear and quadratic. It was shown that when the quadratic equation is used, limit cycle oscillations do not occur. A further study by Edwards and Brindley (1996) and Edwards (2001) varied the parameter values and found that the quadratic equation is able to produce oscillations as shown in Figure 1.3, in both NPZ and NPZD models. Similar sensitivity studies have been applied to the seven compartment biogeochemical model of Fasham et al. (1993), and limit cycle oscillations still occurred with all mortality functions, even with sigmoidal and hyperbolic forms, when the half saturation coefficients are low (Edwards and Yool, 2000). This shows that oscillations are produced regardless of the functional forms that are used to describe zooplankton mortality, but it is more common with linear terms (Edwards and Yool, 2000). Therefore the occurrence of limit cycle oscillations are dependent on the choice of parameter values. These studies show how the choice of functional forms and parameter values can change the model dynamics, however it only explores one particular process, the zooplankton mortality.

A more comprehensive study by (Kriest et al., 2012) assess both sensitivity in parameter values and model complexity, by changing only some essential parameters (for example: the maximum growth rate, half saturation constant for light, nutrients, and sinking speed). The model complexity varies from nutrients only, to NPZD-DOP and is run globally. The parameters are varied by halving and doubling the typical values for each parameter, but with a smaller range for the sinking speed, so that the experiment is computationally feasible. Compared to the in situ nutrients, in this case phosphate, simpler models show higher sensitivity to variations in parameters. The regional misfits also differ across different complexity models, where more complex models underestimate nutrients in the North Pacific and the Southern Ocean, but the simpler models overestimate nutrients in the equator, low latitudes, and northern North Atlantic. Overall the more complex model produces a better fit with observations at the surface, but all models are sensitive to sinking parameters (Kriest et al., 2012). The results from this study agrees with a previous study by Friedrichs et al. (2007), where the performance of 1-D models with varying complexity, from NPZD to a 24 component model, is assessed. The most complex model shows the best fit with the observations, with or without data assimilation to optimise the parameters, as long

as only a few key biogeochemical parameters, such as growth and remineralisation rates, are tuned to the observation. However if the number of parameters that are optimised are not constrained, the success of reproducing an independent dataset may decrease (Friedrichs et al., 2006, 2007). Furthermore underdetermined parameters are harder to fit than those with *a priori* values from observation (Ward et al., 2010), making complex global biogeochemical models highly uncertain, and difficult to fit, as they are computationally expensive and must cover various biogeographical regions. This also begs the question whether it is still necessary to run a single deterministic model on a global level, where parameters values are often dependent on regions, and across taxonomic groups.

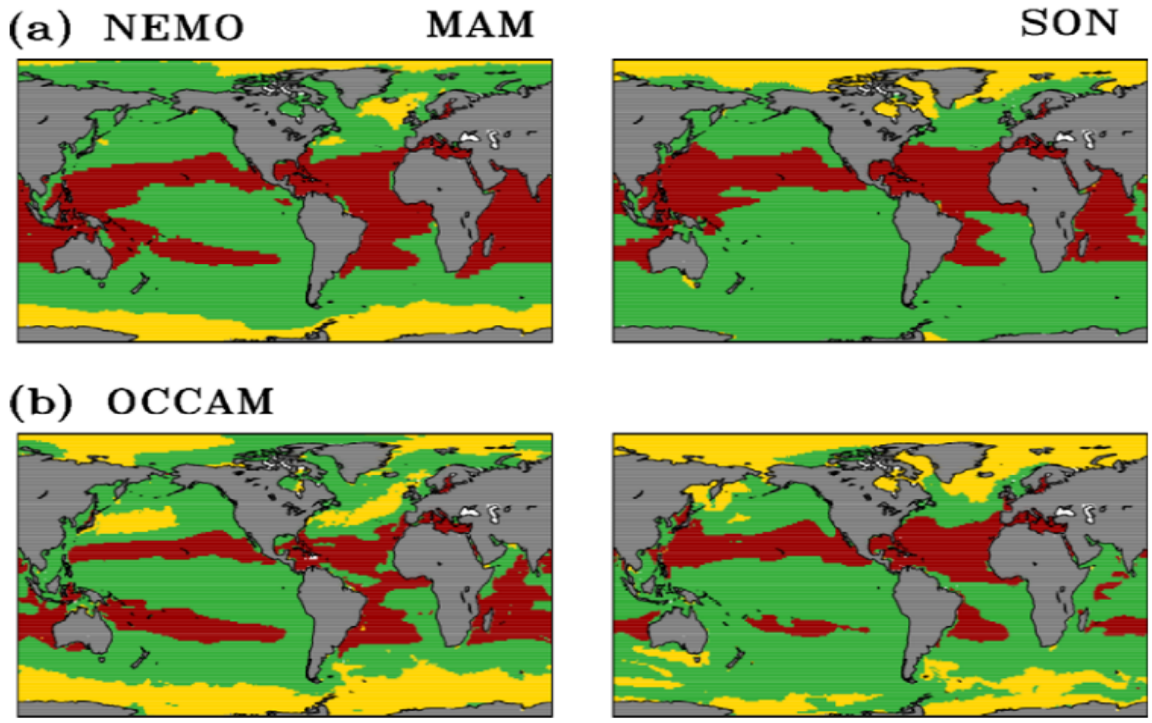


Figure 1.4: Distribution of dominant PFT for March April May (MAM) and September, October, November (SON). Mixed phytoplankton, diatoms, and coccolithophores are shown in green, yellow, and brown respectively in 1994 for NEMO (a) and OCCAM (b) simulation (These subfigures are obtained from Sinha et al. (2010)).

In a 1-D model, chlorophyll and zooplankton biomass distribution is more sensitive to the model complexity than the choice of physical forcing input, if the biogeochemical model is not assimilated. Yet, after assimilation, the distribution of zooplankton and chlorophyll concentrations are more sensitive to the physical input (Friedrichs et al., 2006). This study may be relevant in 1-D regional models, however in a global operational 3-D biogeochemical model, optimisation may be computationally expensive (Kriest et al., 2012). Furthermore, as previously discussed, the emergent properties, such as the distribution of PFTs in PlankTOM5.2, a complex biogeochemical model with five PFTs, are sensitive to the choice of ocean GCM (Sinha et al., 2010), as shown in Figure 1.4. When NEMO is used, mixed-phytoplankton and mesozooplankton dominated the North Atlantic, North Pacific, and Southern Ocean

between September to November, but in OCCAM diatom and microzooplankton are the dominant types in these regions. This discrepancy may be due to enhanced mixing in OCCAM, but it further emphasises that correct formulation of PFTs is important as this determines the emergent properties of a biogeochemical model (Sinha et al., 2010). Anderson et al. (2010) conducted a study using similar model (PlankTOM5.2) where the zooplankton grazing equation is varied using four different functional forms, namely rectilinear, Holling type II, Holling type III, and Ivlev. The study uses similar parameter values for each functional response to show how different behavioural assumptions is translated into functional dependencies on diet and food density. The model results shows that the surface chlorophyll concentration during autumn varies greatly, especially between Holling type III and Holling type II (2.73 mmol C m^{-3} and 1.58 mmol C m^{-3}, respectively), but not so much during spring time (2.96 mmol C m^{-3} and 2.12 mmol C m^{-3}, respectively). These results show that subtle differences among functional responses can affect significantly the predicted phytoplankton densities.

It is clear from these studies, both the model functional forms that describe the biogeochemical processes and the parameters play a big role in contributing to the model uncertainty. In the next subsection, we will discuss how different model equations can alter the model dynamics.

1.4 Why structural sensitivity needs to be addressed

From the studies mentioned previously, changes in model dynamics can occur when the parameter values are perturbed during conventional parameter sensitivity analyses. However, much larger changes in system dynamics can result from changes in the structural formulation of process functions, despite these formulae producing similarly shaped curves (Wood and Thomas, 1999; Fussmann and Blasius, 2005; Levin and Lubchenco, 2008; Flora et al., 2011; Adamson and Morozov, 2013; Aldebert et al., 2016). This is because a change in formulation may affect the function shape, and this also affect the stable state of the system (Aldebert et al., 2016). This is called structural sensitivity (Wood and Thomas, 1999; Flora et al., 2011; Adamson and Morozov, 2013). A study by Fussmann and Blasius (2005), demonstrated that in a simple Rosenzwig-McArthur predator and prey model, using similarly shaped prey uptake functions could produce different predator and prey dynamics. In the study, the parameters are chosen so that the functions are phenomenologically similar, as shown in Figure 1.5.

Although similarly shaped, these functions are formulated differently:

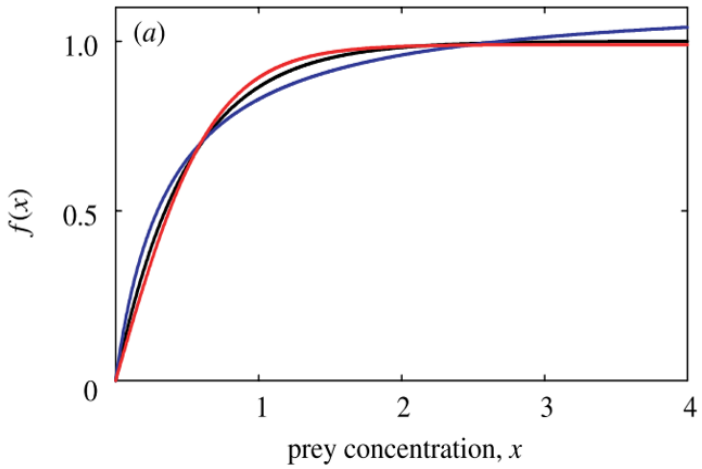


Figure 1.5: Three similarly shaped prey uptake functions from (Fussmann and Blasius, 2005). The rectangular hyperbolic (1.6) is shown in blue, hyperbolic tangent (1.7) is shown in red, and the ivlev (1.8) function is shown in black. Figure originally from Fussmann and Blasius (2005).

$$f_H(x) = \frac{a_{rh}x}{1 + b_{rh}x} \tag{1.6}$$

$$f_T(x) = a_{ht}\tanh(b_{ht}x) \tag{1.7}$$

$$f_I(x) = a_i(1 - \exp(-b_i x)) \tag{1.8}$$

Using equation 1.6, resulted in oscillations with high amplitude of (~ 0.8), and predator and prey concentrations dropping close to zero, conversely equations 1.8 produced much lower amplitude (~ 0.2) of oscillation. Equation 1.7 produced the most striking difference as it produces a steady state, summarised in Figure 1.6. These differences in model dynamics show structural sensitivity due to model functional forms (hereafter structural sensitivity).

Structural sensitivity may be less significant in models built on well-tested mechanisms such as those in the physical sciences. However, in a biogeochemical model, any mathematical functions that describe a process are likely to be an oversimplified representation of that process. This is because biogeochemical processes are mostly complex, and involve many interactions between diverse individuals, across different regions and temporal resolutions in an environment that is changing rapidly (Adamson, 2015). This is even more problematic if the process itself is not well understood so that theoretical justification for the specific representation is weak (Adamson and Morozov, 2013). Often it is difficult to implement the functional relations that are observed in the laboratory into a large scale ecosystem with heterogeneous populations (Englund and Leonardsson, 2008). From simple predator-prey models, applying similarly shaped equations can also give completely different stability and oscillatory model dynamics (Fussmann

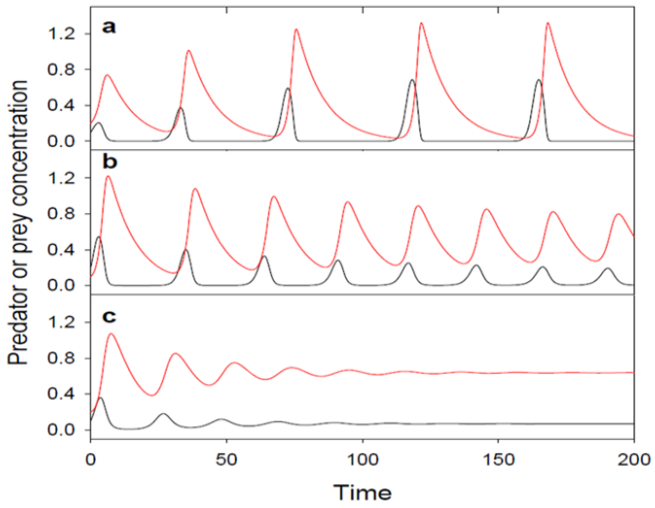


Figure 1.6: Different dynamics produced from similarly shaped prey uptake function 1.5. Rectangular hyperbolic (1.6) is shown in a, ivlev (1.8) is shown in b, and hyperbolic tangent (1.7) is shown in c. Figure is originally from Fussmann and Blasius (2005).

and Blasius, 2005; Roy and Chattopadhyay, 2007). Moreover, a specific functional form may not capture all details of the biological processes, for example, the rectangular hyperbolic function for grazing, commonly known as the 'Holling Type II', fails to correctly describe what happens to grazer's movements when satiation has been reached (Flynn and Mitra, 2016). The formulation that is adequate from a theoretical point of view, does not necessarily describe the data quantitatively (Aldebert et al., 2018). In a recent study, although the complex model promotes the survival of a species, such as phytoplankton, the variability of dynamics in the foodweb is more strongly affected by the choice of functional forms (Aldebert et al., 2016).

Some studies have investigated the effects of different process formulations on biogeochemical models, e.g. Edwards and Yool (2000); Yool et al. (2011) have demonstrated that in a simple NPZ model and an intermediately complex model, linear density dependent mortality produces the most significant differences when applied to diatoms, compared with sigmoidal, quadratic, or hyperbolic forms, as stated in section 1.2. The choice of zooplankton grazing equations can also affect phytoplankton concentration dramatically in a model with five plankton types, PlankTOM5.2 (Anderson et al., 2010),which has been discussed previously, and also in self-assembling ecosystem models (Prowe et al., 2012). The Holling type II grazing function produces 30% less total surface phytoplankton concentration compared to the Holling type III functions in the North Atlantic and North Pacific (Anderson et al., 2010). Nevertheless, not all processes give significantly different model output. Anderson et al. (2015) also shows that when two similarly shaped photosynthesis-irradiance curves, namely, Smith and the exponential function, were used in an NPZD model, the concentration of chlorophyll during the spring bloom was only slightly higher (0.2 mg m^{-3}) for the exponential function (Anderson et al., 2015). For these studies, only one

process function is altered in the model, despite multiple processes that might be structurally sensitive. Therefore, a thorough assessment on how structural sensitivity affects the results of an intermediately-complex or an operational model, with more biological processes that are varied, is needed (Aldebert and Stouffer, 2018).

There are models that avoid the issue of structural sensitivity by constructing a generalised bifurcation diagram in a generalised parameter space, to explore the possible model dynamics (Gross and Feudel, 2006). However, this kind of model does not address alternative stable states that can be affected by structural sensitivity (Aldebert et al., 2016). Further, the discrepancies reported from simple interaction models suggest that the dynamics of complex biogeochemical models need to be tested by altering their default functional forms (Anderson and Mitra, 2010; Anderson et al., 2010). To properly address the structural sensitivity in an operational biogeochemical model, it is possible to generate a model ensemble.

1.4.1 Generating ensembles to address uncertainty

In order to address the alternative simulations representing structural uncertainty, it is possible to generate an ensemble. Here, an ensemble means multiple model variations are run to acquire a range of future predictions, or simulations, with its uncertainties represented by the spread of outcomes. The ensemble approach has also been used to inform policy makers and planners to estimate uncertainty associated with physical model projection, so that appropriate strategies for adaptation could be identified (Murphy et al., 2007). This also means that the ensemble should represent the uncertainties that may also occurred in the true value from the observation.

In climate modelling, perturbed physics ensembles have been developed to investigate multiple parameter uncertainty (Murphy et al., 2007; Tinker et al., 2016). The ensemble can be generated by varying uncertain parameters according to prior studies of relevant physical processes, or using an ensemble Kalman filter (Murphy et al., 2007). Another approach to ensemble modelling is by exploring multiple parameterization (functional) uncertainties (Subramanian and Palmer, 2017). Additionally, it is possible to generate an ensemble which consists of multiple climate models, or multimodel ensembles, although since some models share similar process representations and parameterisations, it is often that nominally different models might still have similar biases, because the models are not independent (Abramowitz et al., 2019). Although uncertainty is always present in every model, it is possible to utilise this so to improve the skill of climate models. The next section will introduce data assimilation, where uncertainties are used to make model more consistent with observation.

1.5 Utilising uncertainties: Data Assimilation

Data assimilation aims to provide the best possible descriptions of the time varying ocean circulation, and therefore helping to improve the skill of climate models (Stammer et al., 2015). Some quantities in the ocean are difficult to observe, such as vertical velocity, mixing, and viscosity coefficients, and data assimilation can help to estimate these quantities (Stammer et al., 2015). In order to apply this method, an error estimate of the model and available observations are required. Data assimilation is also used in numerical weather prediction to create initial conditions for atmospheric models designed to forecast over short time scales (within hours or days) (Bouttier & Courtier (1999)). This approach is adopted by oceanographers to initialise ocean and seasonal forecasts (e.g. Anderson et al. (1996)).

Ocean data assimilation converts the available information from the observations, which include making a model more consistent with the observations, within the error estimates of both the model and the observations. In order to carry out this approach, a model that is being constrained by quality controlled data, in which error information about both the model and observations are available is required. The mechanism for filtering and extracting the useful information from inaccurate and noisy data and combining them with the model is provided from the assimilation method. The mechanism in carrying the information forward from earlier observations to the analysis time and beyond is provided by the model (Anderson et al., 1996). There are two main methods, one is sequential where model and data are combine at given (analysis) times, then integrated forward to the next assimilation time by using the analysis as initial conditions for the model, and the other is a trajectory method, where the best model trajectory that fits the data that are gathered during a particular time interval is selected. In this method the model needs to be varied so that the trajectory analysis can be carried out.

As explained in previous sections some model parameters, for example the half saturation coefficient, vertical mixing coefficient, or the external forcings, such as vertical velocities, are uncertain. In the trajectory method, these uncertainties are utilised to form model trajectories, to then be used in minimising the 'cost function' which measures the difference between the model trajectory and the observations. As seen in the pilot experiment in section 2.9, the variability arising by altering different functional forms is large, in terms of time period and also the mean overall concentration. This high range may be able to encompass the time varying observational data and errors, without having to vary further parameter values.

In marine biogeochemical models, data assimilation has been used with ocean colour data, such as SeaWIFs (Ford et al., 2012; Ford and Barciela, 2017; Ciavatta et al., 2014). The results show that the model bias is reduced, root mean square error, and also the correlation with GlobColor products is improved. However if the physical forcing for the model is assimilated independently, although resulting in better

correlations with satellite-derived chlorophyll, the bias is increased, especially for nutrients. This may be because the fluctuation of vertical nutrient transport due to spurious vertical velocity, or gravity wave generation in the assimilation cycle initialisation (Ourmières et al., 2009; Raghukumar et al., 2015). This evidence will be presented in section 2.5.3.

From section 1.3, studies have shown that the effect of model uncertainties can vary across different oceanic regions. The next section will explore the characteristics of different oceanic regions based on oceanographic stations, where physical, chemical, and biological oceanographic data have been observed and analysed regularly over a long period of time. These stations are used to monitor the changes in the ocean environment and how it impact the ecosystem, and also provide observational data to construct, run, optimise, or validate biogeochemical models.

1.6 Variability of the biogeochemistry at different oceanic regions

Different regions in the ocean are characterised by different physical and biogeochemical signatures. The North Pacific Tropical Gyre is characterised by low nutrients and strong stratification (Barber, 1988). The sudden increase in phytoplankton biomass, also called a bloom, at the surface only occurs when the upper ocean stratification is changed (Corno et al., 2007), however high phytoplankton concentrations can also occur below the mixed layer, also known as deep chlorophyll maxima (DCM). The concentrations of the DCM is often higher than at the surface. The coastal region, is often characterised by high nutrient availability at shallower depths and its transport to the euphotic zone is controlled by upwelling (Walsh et al., 1999). In the subpolar region, the phenomenon of the spring bloom has fascinated oceanographers since the 1950s (Sverdrup, 1953). It has been shown that there is a link between winter mixing depth and the occurrence of this phytoplankton bloom, which implies that weaker winter mixing can reduce the bloom cycle (Behrenfeld et al., 2013).

These differences in physical and biogeochemical processes can be partitioned into four biomes, and further into various biogeographical provinces, as shown in Figure 1.7. According to Longhurst et al. (1995), seasonal phytoplankton growth observed from satellite-derived chlorophyll and physical processes in the upper ocean are used to partition the provinces. There are four primary domains of the global pelagic ecosystem: Polar, Westerlies, Trade Winds, and Coastal. The first domain is characterised by the seasonal sea ice that results in brackish water that occur during spring and summer. The active bloom occurs as soon as ice break-up occurs. The defining characteristic of the westerlies is the deepening of the mixed layer in the winter that is caused by the westerlies wind stress. This region is further split into 16 secondary biogeochemical provinces, which include the North Atlantic Subtropical

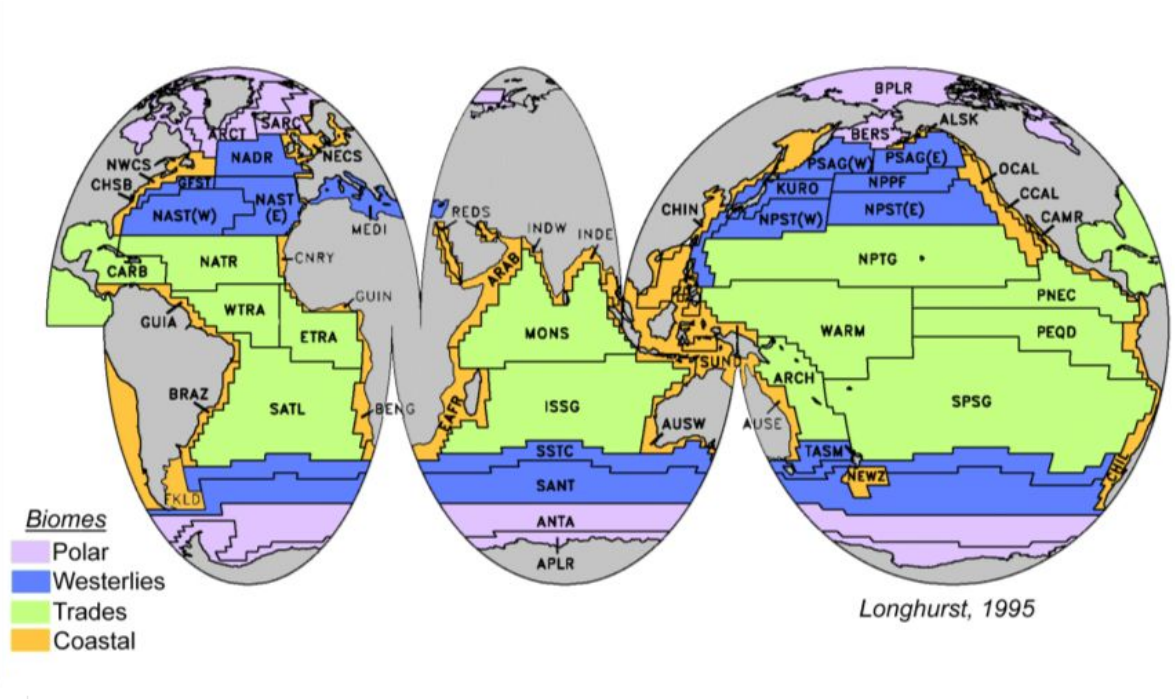


Figure 1.7: Biogeographical provinces, acronyms are defined by Longhurst et al. (1995). The provinces are essentially partitioned into four biomes: polar (pink), westerlies (blue), trade winds (lime green), and coastal (amber). The figure is taken from Ecological Geography of the Sea (Longhurst, 2007).

Gyre (NAST(W)) and North Atlantic Drift (NADR). In the trade wind domain, it is expected that the mixed-layer algal blooms are not light limited, the seasonal change in mixed-layer depth changed due to a geostrophic response to the wind field, and the nutrient renewal in the mixed-layer only occurs at strong divergences (Longhurst et al., 1995). This biome is further partitioned into 12 provinces, which includes the Caribbean (CARB), North Atlantic Tropical Gyre (NATR), and North Pacific Tropical Gyre (NPTG). Finally the coastal domain, has significantly modified circulation produced because of the coastal topography and coastal wind regimes (Longhurst et al., 1995). Often in this domain, coastal upwelling and anticyclonic eddy fields dominate. However, these biogeographical regions can be fluid. For example the unification of tropical and subtropical Pacific regions during ENSO events (Barber 1998). Barber (1998) also distinguish further the biogeographical region into six ocean basin ecosystems, which includes Coastal upwelling, Low latitude gyre, Equatorial upwelling, subarctic gyre, southern ocean, and eastern boundary currents. These regions are distinguished by stratification strength and duration, nutrient level and source, the amount of primary productivity, and processes that regulate the productivity. For example the low latitude gyre is characterised by strong and permanent stratification, with oligotrophic (nutrient limited concentrations) conditions. This results in low to medium primary productivity, which is regulated by the grazing and physical processes (Ducklow, 2003).

As discussed earlier, observing the marine environment has been done in order to understand a variety of interacting processes via positive or negative feedbacks (Dickey and Bidigare, 2005). Due to challenges

and problems faced by oceanographers, such as climate change, pollution, and eutrophication, ocean observational programmes have become inter-disciplinary, in order to understand the interaction between physical, biological, chemical, and geological processes. Therefore in some of oceanographic cruises, inter-disciplinary observations are common, for example the North Atlantic Bloom Experiment (NABE) spanning from April to May 1989 (Harrison et al., 1993), AMT cruise (Aiken et al., 2000) that takes place every year since 1995 between summer and fall, and the Ocean Surface Mixing, Ocean Submesoscale Interaction Study (OSMOSIS) that took place between April to September 2015 (Hemsley et al., 2015). However, these observations only take place for a short period. Although AMT cruises, take place every year, it only covers a certain period (e.g. between September to October), and therefore only the long term trend for that season can be drawn from the cruise. In one of the first ocean biogeogchemical studies, in the Northeast Pacific Ocean, significant variability in rates of primary production and particle fluxes have been observed (Karl and Lukas, 1996). Yet, there are no clear relationships between new production and primary production. Although the intensity and scope of the research project is comperhensive, high sampling frequency is needed in order to resolve the natural variability in the oligotrophic ocean (Karl and Lukas, 1996). This further emphasises the needs to have long term frequent holistic integrative measurements.

1.6.1 Long term observation stations

Long term regional ocean monitoring has been done in order to gain a deeper understanding of the variability of the natural habitat, recording multi year trends, and how global change of the environment that might arise from human activities affect the variability (Karl and Lukas, 1996). It has documented the increase of CO_2 over the years (Doney et al., 2012). In fisheries management, long term data are also essential for informing a precautionary approach (Hawkins et al., 2013). As one of the objectives of JGOFS, these regional monitoring stations are now spread across different biogeographical provinces: station ALOHA, which is located at the North Pacific Tropical Gyre within the trade wind region; station BATS, located in the North Atlantic Subtropical Gyre within the westerlies, but also the trade wind region of North Atlantic Tropical Gyre; L4 and Cariaco which are located in the Northeast Atlantic Continental Shelf and Guiana Current respectively, within the coastal region; and station PAP, situated in the North Atlantic Drift, within the westerlies. These stations have available chlorophyll and DIN data at least at the surface or depth profiles. Apart from monitoring the changes in the ocean environment, the also provide observational data to build, run, or validate biogeochemical models. This section will explain the characteristics of each station, such as the average DIN and chlorophyll concentrations, along with the mixed layer depth, and the measuring strategies.

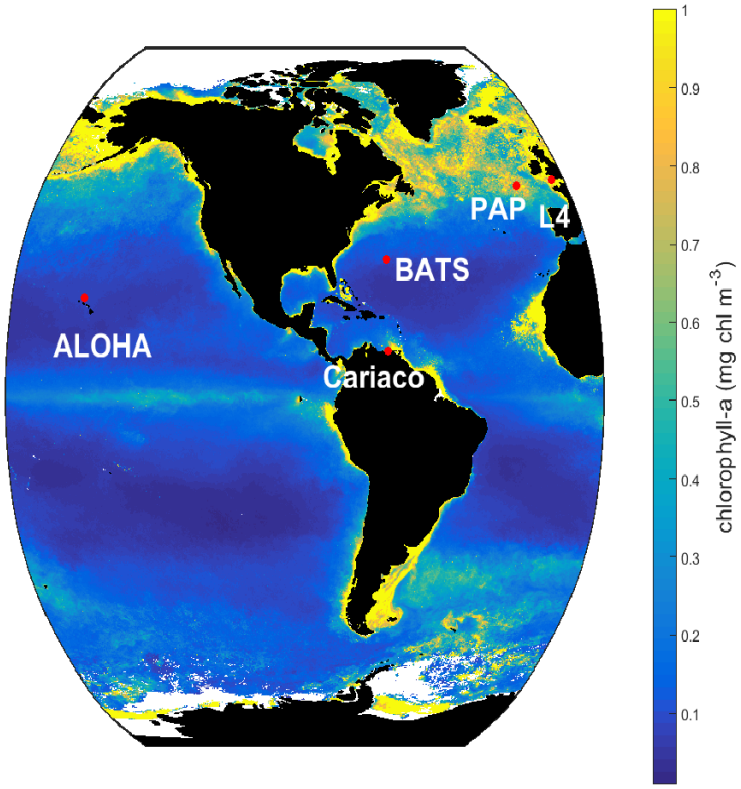


Figure 1.8: SeaWIFs-derived mean 1998 chlorophyll-*a* (mg m^{-3}) overlain with the 5 oceanographic stations time series site (Red dots). These stations are located in different oceanic regions: oligotrophic (ALOHA and BATS), coastal (L4 and Cariaco), and abyssal plain (PAP).

Station ALOHA

Station ALOHA (A long term Oligotrophic Habitat Assessment) is a deep-water hydrostation in the NPTG (22°45'N 158°00'W). The first water sample was taken on October 1988, and the following samples are taken monthly. These are collected from the surface to 50m above the seafloor (4800m). The measurements that are conducted at this station include dissolved nitrate and nitrite, silica, and chlorophyll-a (Karl and Lukas, 1996). However recently, other measurements have been added (such as alkalinity, pH, dissolved organic nitrogen, see Hawaii Ocean Time Series (2017)). The main objective of the measurements is to provide a database for modeller to validate and develop biogeochemical models (Karl and Lukas, 1996).

At station ALOHA, surface warm water and deep permanent pycnocline and nutricline are common features. In the upper 100m of the water column, nutrient concentration is very low ($<$ 15nM), and the nutricline (2 μmol kg^{-1} horizon of nitrate plus nitrite concentrations) depth varies seasonally with a range of 170 – 191 m (Dave and Lozier, 2010). The annual mean mixed layer depth (MLD) at this station has increased from 1997 to 2004 from 56m to 66m, which may be caused by the El-nino Southern Oscillation (ENSO) (Corno et al., 2007). The seasonal range of MLD is between 36-96 m, where deeper mixed layers occur in winter (Dave and Lozier, 2010). A continuous DCM, that usually occurs below the mixed layer is observed at 74 – 140 m, because light-attenuating compounds at the surface are rare, and therefore photosynthetically available radiation (PAR) is sufficient to support it, with DCM concentration could go up to 0.4 mg m^{-3} (Letelier et al., 1993), as shown in Figure 1.9.

Although chlorophyll levels are low in the top 25m (chlorophyll mean $\sim$ 0.08 mg m^{-3}), phytoplankton blooms have been observed to occur during late March, where concentrations can go up to three times of the mean (Letelier, Karl, Abbott, Flament, Freilich, Lukas and Strub, 2004), with another small peak during summertime, which comprises of nitrogen-fixing phytoplankton. Due to the low nitrogen concentration in the euphotic zone, nitrogen fixation (diazotrophy) is an important source of nitrogen, and supports the summer chlorophyll booms, although the bloom is not as high as the one in spring (Villareal et al., 2012). The increase of chlorophyll concentration at ALOHA is not always due to the entrainment of deep nutrients, but it may be due to photo-adaptation and changing photosynthetic efficiency with deeper mixing that occurred during winter months(Dave and Lozier, 2010). In the absence of large-scale climate events, such as ENSO, the North Pacific has stronger stratification than the North Atlantic (Dave and Lozier, 2010). This reduces vertical mixing and restricts deep nutrient inputs to the surface (Letelier et al., 1993), as station ALOHA is located at the relatively calm southern and eastern recirculation zone of the North Pacific Subtropical Gyre. In the event of low mixed layer depth variability and detectable pulses of nutrients, the proliferation of diazotrophs occurred and the pelagic ecosystem shifted from

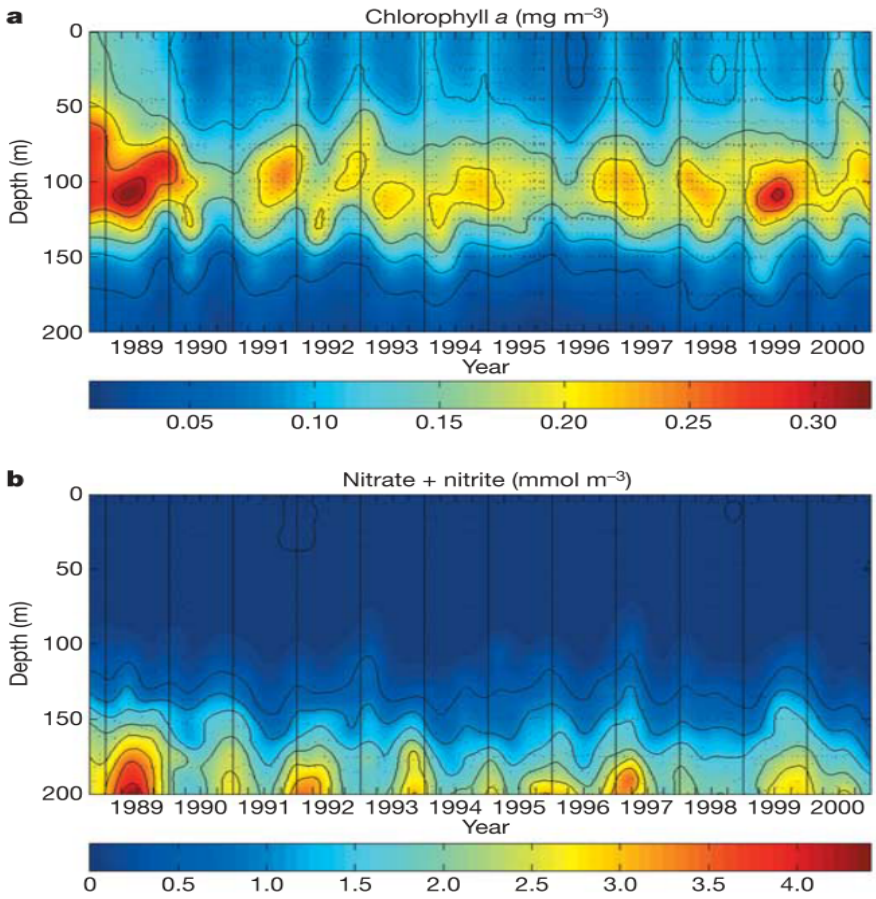


Figure 1.9: Subsurface chlorophyll maxima at station ALOHA (a) and dissolved inorganic nitrate and nitrite (b) at station ALOHA between 1989 to 2000. The figure is obtained from Huisman et al. (2006).

nitrogen limited to phosphorus (Letelier, Karl, Abbott, Flament, Freilich, Lukas and Strub, 2004).

Phytoplankton community structure can also be influenced by physical changes. A study Corno et al. (2007), reported an increase in primary production and biomass at station ALOHA, that coincide with changes in upper ocean stratification, which includes the change in MLD, which may link to larger-scale climate forcing, such as ENSO and Pacific Decadal Oscillation. These changes are potentially influencing the nutrient dynamics in the upper ocean, and light availability (increasing mixing resulted in decreasing daily integrated light in the mixed layer). These changes are accompanied by a shift in plankton assemblages from larger eukaryotes to smaller prokaryotes in the upper ocean.

Station BATS

The Bermuda Atlantic Time-series is located in the northwest Sargasso Sea, in the subtropical gyre of the western North Atlantic (32° 50'N, 64° 10'W, NASW(W)), bounded on the west and northwest by the Gulf Stream and to the south by the North Atlantic equatorial current (Steinberg et al., 2001). At this station hydrographic properties, such as temperature, salinity, and conductivity, and biological parameters, such as nitrate, silicate, and chlorophyll-a have been sampled monthly since 1988. Surface temperature and mixed-layer depth varies seasonally by 8°–10° C and from >200m in the winter to <10 m during summer respectively (Bates, 2001).

During April to October, high heat flux and low wind stress causes a strong thermal stratification. During this period, nutrients are absent within the euphotic zone (Lipschultz, 2001), making primary production rates low, but a subsurface chlorophyll maximum is present between 60 and 120m (Steinberg et al., 2001). The DCM at BATS which consists of chlorophyll-a and chlorophyll-b, peaks at 100m, which can go up to 0.7 mg m^{-3} (Steinberg et al., 2001), as shown in Figure 1.10. The underlying Subtropical Mode Water (STMW), that is found in a thick layer of nearly uniform temperature around 18°C, controls the nutrient availability of nutrients in the North Atlantic Subtropical Gyre. The thickening of STMW layers depletes the nutrient in the Gyre, as it push the nutrients deeper (Palter et al., 2005). However in the winter deep vertical mixing, brings nutrients into the euphotic zone, making the nitrate concentration increase up to 0.5 mmol m^{-3} (Lipschultz, 2001), resulting in a short period of spring bloom between January to March, making the initiation occur in late December, when the MLDs were deeper than the euphotic zone, and termination is between April–May, before the MLDs starts to deepen. The phytoplankton bloom at BATS occur due to the increase in phytoplankton group, instead of the increase in biomass of a particular group that is present in the water (Steinberg et al., 2001). From these description, although the trade winds region is not light limited which is similar to BATS, the mixed layer deepening in winter is more similar to the westerlies domain (Longhurst et al., 1995).

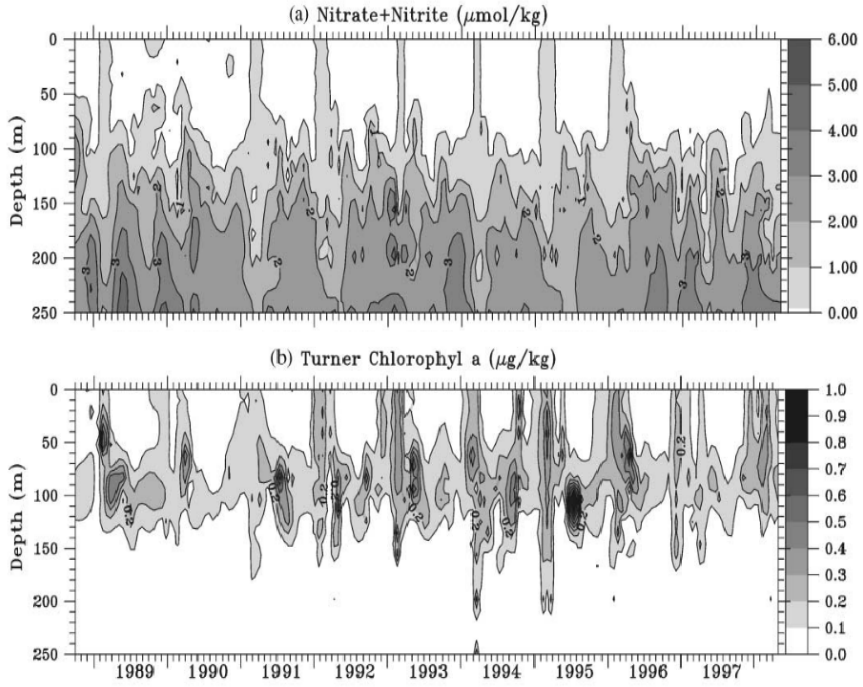


Figure 1.10: Dissolved nitrate and nitrite (a) and chlorophyll a concentration (b) at station BATS between 1989 to 1997. The figure is obtained from (Steinberg et al., 2001).

Biogeochemistry at Station BATS has also been shown to be driven by climate variability, such as ENSO (Bates, 2001). Similar to the North Pacific, in the Sargasso Sea, there has been an increase in integrated chlorophyll, primary production, and prokaryotic phytoplankton between 1996-2007, which coincide with large-scale climate forcing. However, unlike the North Pacific, there are no significant changes in MLD and upper ocean stratification (Lipschultz, 2001). Nevertheless there was an increase in vertical mixing due to the negative anomalies in the North Atlantic Oscillation (NAO). This NAO index is the dominant climate mode in the region, and negative values would result in intensified westerlies. This changes the mixing therefore increasing the nitrate gradients, which also raise the concentration of cyanobacteria, *Synechococcus* (Lomas et al., 2010). Additionally in a modelling study, it is suggested that the thinning layer of STMW, which in turn increases the nutrient supply in the upper ocean, also contributes to the increase in chlorophyll and therefore primary production at BATS (Saba et al., 2010).

Station Cariaco

This station in located on the Venezuelan continental shelf, in the eastern part of the Cariaco basin (10°30'N, 64°40'W). It has been the site of sampling for marine biogeochemical and ecological observation since 1995, as part of the Carbon Retention in a Coloured Ocean (CARIACO) Time-Series study (Muller-Karger et al., 2001). It is bound to the north by a sill, connecting Margarita Island to Cabo

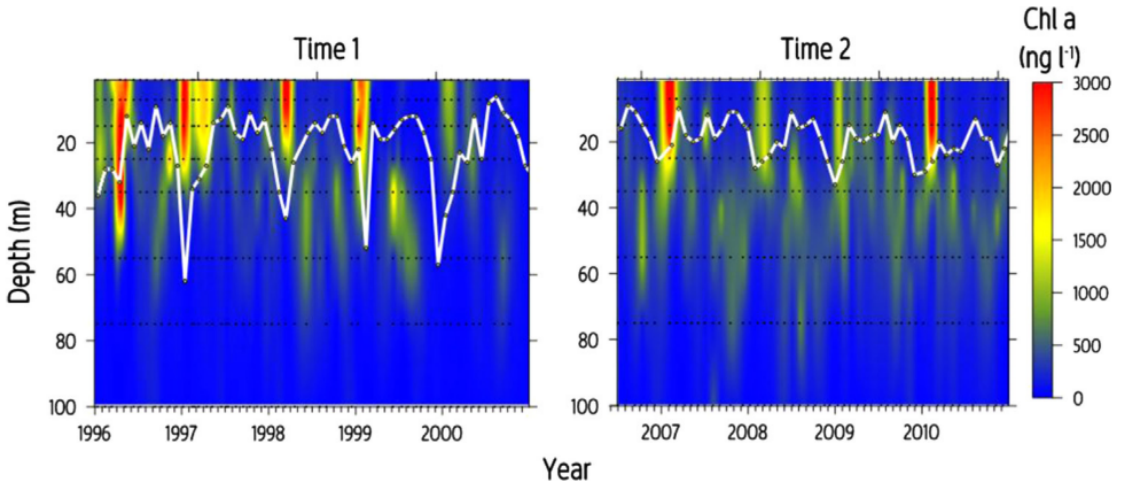


Figure 1.11: HPLC-measured chlorophyll from January 1996 to October 2000 (Time 1) and July 2006 and December 2010 (Time 2). Dots represent the data points from the in situ observation at station Cariaco. This figure is obtained from Pinckney et al. (2015)

Codera, at a mean depth of about 100m with two channels breaching this sill. It restricts any water exchange below 140m, therefore below 275m, permanent anoxia occurs because of restricted circulation and high primary productivity (Muller-Karger et al., 2001).

Coastal, wind-driven upwelling in the southern Caribbean sea has been linked to the high productivity in the Cariaco Basin (Walsh et al., 1999). The upwelling season occurs between December and April due to trade winds, and accounts for $\sim$ 70% of annual depth integrated primary production that is dominated by microphytoplankton consisting of diatoms, dinoflagellates, and coccolithophores (Muller-Karger et al., 2001). During these months, the total chlorophyll-a increases (up to 3 mg m^{-3}) (Pinckney et al., 2015), and peaks between April and May exceeding 6 mg m^{-3} (Muller-Karger et al., 2001). The chlorophyll concentration then declines to $<$0.2 mg m^{-3}, in June and December (Muller-Karger et al., 2001). Subtropical Underwater (SUW), that lies north of 14.5° N between 100 and 200m, depth and shallower than 150m south of 14.5° N, is characterised by high salinity ($\sim$ 36.85 PSU). Intrusion of deeper water from the Caribbean sea into the Basin occurrs when the salinity gradient tilts upward towards the southern Caribbean around the same time as the upwelling events (December and the following June) (Morrison and Smith, 1990). Within the layer, Nitrate of 5-10 μM is stored (Morrison and Nowlin, 1982). When there is an upward tilt, nitrate is therefore elevated to the surface, which leads to elevated nutrients, stimulating phytoplankton growth (Muller-Karger et al., 2001). However in May-November, the rainy season caused an increase in stratification of surface waters, reducing primary productivity (Lorenzoni et al., 2013).

There were weaker trade winds between 2006-2010 which changed the physical and biogeochemical conditions of the basin. Furthermore by 2010, the average SST has increased by $\sim$1.0°C, compared to 1995, and the mixed layer depth increased (Taylor et al., 2012; Pinckney et al., 2015). The surface

chlorophyll decreased in concentrations between 2006 and 2010. However, the 100m integrated chlorophyll between 1996 and 2000 is significantly lower than between 2006 and 2010, especially when only integrated from 55-100m. These are summarised in Figure 1.11. This increase in subsurface chlorophyll is associated with less intense upwelling and greater light penetration due to the decrease in primary production. A decrease in diatoms has been reported, but other microphytoplankton have increased in abundance within the top 100m (Pinckney et al., 2015). This is consistent with earlier studies where a decrease in large cells can be triggered by a small increase in temperature (Mousing et al., 2014). These observations further shows that phytoplankton community structure can be shaped by the physics of the ocean.

Station L4

Station L4 is located 13 km off the coast of Plymouth (50° 15'N, 4° 12.3'W; 50m depth) in the Western English Channel where its dynamics is influenced by tides. It stratifies constantly in summer where nutrients are depleted and a distinct thermocline is observed (Smyth et al., 2010). During periods of strong south westerly wind the hydrography is influenced by both oceanic and also from the riverine flood water from the nearby rivers (Smyth et al., 2015). Additionally, multi-decadal oscillations of several environmental parameters at L4 are driven by the NAO (Reygondeau et al., 2013), such as the seasonal variability, timing, and amplitude of wind speed and direction and sea surface temperature (SST) and stratification (Molinero et al., 2013).

Surface measurements of Chlorophyll-a and inorganic nutrients are taken weekly. Chlorophyll a is available from 1999, and surface nutrients have been collected since 2000 at station L4. From the satellite image, the background chlorophyll a concentration in the Western English Channel is typically 1 mg m^{-3} throughout the year. In the middle of the channel, the bloom occurred for a long time, from June to August, which mainly consist of dinoflagellates (Smyth et al., 2010). However, two bloom events occur every year, with sharp peaks for short durations; in May and September where it could go up to 8 mg m^{-3} in the autumn, but the typical value is observed to be around 2 mg m^{-3}, as shown on Figure 1.1. Diatoms mainly dominates the spring bloom whereas are found dinoflagellates in autumn, as they can access the nutrients below the nutricline and swim back up to access light at the surface (Ryan et al., 2009; Smyth et al., 2010).

Another aspect that controls the length of the phytoplankton bloom is the availability of nutrients. In January, surface nitrate can get up to 8μM, however between May and July, it falls below the detectable limit. Before the spring bloom in March, a maximum of 0.25μM has been measured, although it then falls below the detectable limit until July, because it has been taken up by the phytoplankton. In Septem-

ber and October, when thermocline vanishes and stratification has been broken, nitrate is replenished in the surface layer (Pingree et al., 1977), which causes the early autumn bloom.

Changes in phytoplankton community structure and bloom phenology have been observed in station L4. A prolonged duration and amplitude of summer SST have occurred since the early 1990s, which has also been observed in the east North Atlantic deep water. This significant environmental change results from the decline of temperature gradient that occurrs between 1995-2007 and therefore determines the dynamics of the food web (Falkowski and Oliver, 2007). Due to the shift in NAO index, small summer and autumnal peaks occurs, but the spring bloom peak decreases between 1995-2000 (Reygondeau et al., 2015). However, the two bloom events returned in 2001-2007, but with lower average chlorophyll and lower spring bloom amplitude (Molinero et al., 2013). Earlier observations have also shown that in the mid-late 1990s there was a decrease in large diatoms and an increase in phyto-flagellates (Widdicombe et al., 2010).

PAP-SO

The Porcupine Abyssal Plain sustained Observatory (PAP-SO) is located in the Northeast Atlantic, (49° N, 16.5° W), which is in the westerlies biomes in the NADR province. At this site, temperature and salinity have been measured using Seabird SBE 37-IM MicroCAT recorders at 30m. Biogeochemical parameters such as nitrate, chlorophyll-a fluorescence and $p(CO_2)$, were also measured using sensors on the frame, between 2002 and 2007, at a single depth varying from 20 to 225, which fluctuates with the current (Hartman et al., 2015). The sample for nitrate is taken twice daily, between 2003 to 2005. The mooring was deployed three times between 2003 to 2005; between 12 July 2003 to 16 November 2003, 17 November 2004 to 16 June 2004, and 22 June 2004 to 18 July 2005 (see Körtzinger et al. (2008)).

The characteristics of station PAP are similar to the North Atlantic. Observed nitrate from the mooring shows an increase in concentration during the winter and depletion between spring and summer due to phytoplankton consumption. From shipboard measurements, winter nitrate concentrations has experienced a decline (from 8.3 to 4.9 mmol m^{-3} between 2003-2005) over the three year period (Hartman et al., 2010). The North Atlantic Current (NAC) influences station PAP. The changes in nutrient concentration in 2003 to 2005 coincide with the shift in NAC from a southern mode to a more northerly crossing (Bower and von Appen, 2008). The northern mode made the PAP site more exposed to subtropical condition, with lower nitrate concentrations (Hartman et al., 2010, 2015). This declines results in a further decrease in annual new production over the productive period (from March to August), from 85.4 $gCm^{-2}a^{-1}$ in 2003 to 40.3 $gCm^{-2}a^{-1}$ in 2005, which also agrees with the satellite observation from SeaWiFS.

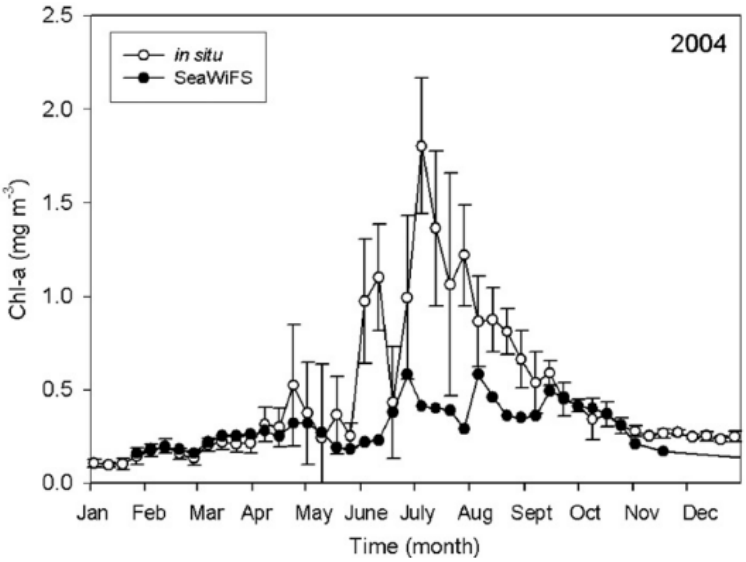


Figure 1.12: Weekly mean ± 1 standard deviation of chlorophyll-a in 2004 from in situ (at an average ± 1 standard deviation depth of 37.56 ± 14.82 m) on the mooring at the PAP observatory sites (open circles), compared with the SeaWIFS-derived chlorophyll concentration (black circles) within 200km radius of the PAP observatory site. The figure is obtained from Hartman et al. (2010)

As at L4, at PAP, two blooms occur through the year, one in spring (April) and another in summer. This has been observed by satellite and also the mooring sensors. However the satellite-derived product underestimates the *in situ* chlorophyll, which may be due to the development of a seasonal DCM which cannot be detected by the satellite. The spring bloom of 2004 peaked in late April, as shown in Figure 1.12, although it shifted into mid April in 2005 (Hartman et al., 2015). This may be caused by the shoaling of the MLD that occurred earlier in the year in 2005 than in 2004 (Hartman et al., 2010). Although the data at PAP is limited to short period, these biogeochemcal changes further show that at this station, a prolonged warming period that affected the bloom timing also occurred, consistent with that reported at station L4.

1.6.2 Similarities and differences

At most of the oceanographic stations, phytoplankton blooms occur when light and nutrients are sufficient, which usually occurs in spring. The availability of nutrients in the euphotic zone is controlled by the upwelling and mixing. Not long after an episode of deep winter mixing, an increase in phytoplankton concentration in the surface usually occurs. Therefore when there has been a change in the physics, such as ENSO, the weakening of trade winds, or an increase in temperature, the timing and the taxonomic composition of the phytoplankton bloom would change. At most stations these physical changes have caused an increase in primary productivity in the last two decades (Corno et al., 2007; Saba et al., 2010).

The oligotrophic ocean is characterised by very low nutrients in the euphotic zone, and therefore low chlorophyll concentration (surface mean $\sim$ 0.08 mg m^{-3}). Small phytoplankton then dominates the water, and during times of low nitrogen, the diazotrophs start to increase. Higher concentrations of chlorophyll are often observed below the mixed layer, due to the absence of light attenuating compounds. The DCM at ALOHA occurs deeper than at BATS and the MLD at ALOHA, although it has increased in depth over the past decade, is shallower than the MLD at BATS during the winter. High nutrient concentrations in the euphotic zone at BATS occurred in early spring after winter mixing, which is then followed by an increase in chlorophyll concentration in the euphotic zone. At ALOHA, although the increase in subsurface chlorophyll is around a similar time, it is not due to the increase in nutrients, as the nutricline occurred deeper than the deepest MLD, but due to photo-adaptation and therefore more efficient photosynthesis.

In the coastal and abyssal plain regions, the mean chlorophyll and nutrient concentration in the euphotic zone is higher than the oligotrophic (surface DIN mean can go up to 8 mmol m^{-3} at the coastal station and in the oligotrophic region, the highest DIN concentration is only $\sim$ 0.5 mmol m^{-3}). The phytoplankton bloom in the light limited regions, such as L4 and PAP occurs twice. First one during spring and second one in autumn or summer at L4 and PAP respectively, when light and nutrients are sufficient. At Cariaco, the bloom occurs in late spring (between April and May) after upwelling in December to April. Usually in the coastal regions, the bloom is dominated by large phytoplankton such as diatoms and coccolitophores. However due to changes in the physical conditions, a community shift to smaller phytoplankton has occurred at both coastal stations, showing how temperature can change phytoplankton community structure. At these stations, the peak of the phytoplankton bloom concentration can go up to 8 mg m^{-3}, although at Cariaco, there has been a decline in the concentration at the bloom peak between 2007 to 2010, and at L4 a decline in chlorophyll concentrations has been observed since 1995.

The oceanographic stations discussed above show how variable the oceanic conditions are at different sites. Even though some stations are located in the oligotrophic regions, the distribution of nutrients and therefore phytoplankton are different. This further shows the challenges in deterministic global biogeochemical modelling to reproduce the observations and therefore provide future predictions under different climate scenarios. From these observations, a change in the timing of a phytoplankton bloom can reflect a change in both physical and chemical environment. Therefore in order to assess the impact of changes due to uncertainties in marine biogeochemical models, the timing of phytoplankton bloom can be used to assess how structural uncertainty affect the model prediction. Furthermore, to provide uncertainties in the phytoplankton bloom in regions where harmful algal blooms occur, an ensemble can be used to inform policy makers and to devise adaptation strategies (Hyder et al., 2015) or early warning system. Therefore a method of generating an ensemble, as well as assessing biogeochemical model

uncertainty, is needed. The ensemble model would also need to be able to capture the observations better than a single deterministic model.

1.7 Aims and Objective

Marine biogeochemical models are uncertain, and there are different sources of uncertainty: parametric, structural - which consist of model complexity and the choice of model equations - and physical input. From previous studies, altering the model process equations have been shown to produce larger differences in model dynamics, compared to altering a large range of parameters. Assessing the model sensitivity thoroughly can be done by using the ensemble method, where multiple model variations are run to acquire a range of model simulations. The sensitivity of the climate model have been assessed using ensemble method, by perturbing the physical model parameter (Murphy et al., 2007; Tinker et al., 2016) or by running multiple climate models, also known as multi-model ensemble (Tebaldi and Knutti, 2007; Abramowitz et al., 2019), but using an ensemble method is not yet common in marine biogeochemical modelling.

This thesis aims to explore the structural sensitivity in a marine biogeochemical model arising from the model process equations (structures) by generating a perturbed biogeochemistry ensemble. Structural sensitivity has been assessed in marine biogeochemical models in previous studies (Ward et al., 2013; Yool et al., 2011; Anderson and Mitra, 2010), however, these studies have not assessed all the key biogeochemical processes either separately or together. In this study, we aim to assess the uncertainty arising from a moderately complex biogeochemical model ensemble of outputs generated using all possible functional forms within the NPZ compartments. This is because these compartments are the heart of most of the marine biogochemical model regardless of the complexity. This is a new approach to generate biological model ensemble, which is different to varying the parameter approach. Generating biogeochemical model ensemble, as done in this thesis, may also contribute to data assimilation (e.g., for Trophic Diffusive Model (TDM) (Cossarini et al., 2009) and NPZ model (Roy et al., 2012)), where model ensemble are generated routinely. Marine biogeochemical models are also highly sensitive to the physical forcings that drive the biogeochemical tracers (Doney, 1999; Sinha et al., 2010; Friedrichs et al., 2006). For the completeness of this approach, we will explore the impact of perturbing physics and biology separately and together. Signature characteristics from the ensembles are also explored, so that it may be possible to infer which processes (physical, biological, or both) may be responsible for the temporal changes seen from in situ observations.

The main objectives of this project are:

1. To develop a method to generate a biogeochemistry ensemble by perturbing the key biogeochemical process formulations in an intermediately complex biogeochemical model (Chapter 2).

2. To evaluate the performance of the new ensemble model and default single structure model with the available in situ observation and satellite-derived products. Additionally, we will explore how:
   a) altering the biogeochemical process formulations affect the model predictions of plankton bloom and
   b) how the effect of structural sensitivity vary across different regions.

3. Extend the perturbed biogeochemistry methodology to include perturbing the physics in the model and generating perturbed physics-biology ensemble. Further, we will also develop methods to minimise the computational cost of such combine perturbations systematically. (Chapter 4)

4. To compare the impact of perturbing the biogoechemistry and physics on model outputs in different regions of the ocean. More specifically we will look into the signatures and characteristics of the ensemble generated from perturbing the biogeochemistry and physics. (Chapter 5)

1.8 Thesis Outline

In this chapter, the construction of a marine biogeochemical model, its evolution and sensitivities to model inputs and equations have been examined. The concept of structural sensitivity and how to examine the effect of this uncertainty in an intermediately-complex biogeochemical model in different regions using a 1-D model have also been explored.

Chapter 2 will describe the methods of exploring the structural sensitivity in more detail, including the model output, various functional forms that made up the ensemble, which will address objective 1. The model metrics, including how to determine phytoplankton bloom phenology, will also be described.

Chapter 3 will examine the effect of structural sensitivity of the ensemble, by using the metrics described in Chapter 2 at different oceanographic stations identified in Chapter 1. This chapter will address the research questions in objective 2. Along with some discussion of the biases against the observation data and possible further work towards 3D modelling.

Since a 3D model is computationally expensive, it is necessary to reduce the number of the perturbed biogeochemistry ensemble, while still retaining the ensemble range and other statistical properties. Furthermore, as described in the introduction chapter, biogeochemical models are also sensitive to the physical input. Therefore the method for generating perturbed physics ensemble, by adding different strength noise to the vertical velocity will be explained in 4. This chapter will discuss the method for objective

4 as well as compare the effect of perturbing the physics, and both biogeochemistry and physics to the model dynamics using the statistical metrics described in Chapter 2.

Chapter 5, will compare and examine the characteristic signatures from the two ensembles described in chapter 4. The ensemble spread, how this affects the chlorophyll distributions, anomaly correlations, and nitrogen proportions between phytoplankton and DIN pool will be explored. This chapter mainly addressed objective 5.

Chapter 6 will generally discuss the results shown in Chapters 3 and 5, and what can be improve on if there is a further study.

Conclusion of the study and possible future directions for structural sensitivity studies, especially in a 3D operational biogeochemical model and how we may utilise the ensemble model will be discussed in Chapter 7.

Chapter 2

Methods: Generating a Perturbed Biogeochemistry Ensemble [1]

The previous chapter has introduced the marine biogeochemical models from the most basic to the more complex models. We discussed how the marine biogeochemical models are uncertain and there are different sources of uncertainty: parametric, structural; which consist of model complexity and the choice of model equations, and physical input. From previous studies, altering model equations have been shown to produce different model dynamics, compared to altering parameters. Assessing structural sensitivity thoroughly can be done by using the ensemble method, which has also been used in climate modelling. Since this method can be computationally expensive, it is possible to use a 1-D model to minimize the computational cost. Calibration based on specific locations using a 1-D model has been shown to improve the predictive skill of 3-D models (Oschlies and Schartau, 2005; Kane et al., 2011; McDonald et al., 2012). Structural uncertainty can provide a form of ensemble to inform policy makers to devise adaptation strategies, and it is possible to use this approach in order to explore different versions of a marine biogeochemical model where each version contains a different combination of functional form. It is possible to assess the impact of the uncertainty using interannual means of biogeochemical properties, such as chlorophyll and DIN, as well as the timing of phytoplankton blooms. At oceanographic stations where biogeochemical and physical observations are available for a long period, a disturbance in the environment, such as a rise in temperature or deepening of the mixed layer, affects the timing as well as concentration of the phytoplankton bloom (Taylor et al., 2012; Pinckney et al., 2015; Reygondeau et al., 2015; Molinero et al., 2013).

As noted, a model intercomparison project, showed that MEDUSA-2.0 has a good balance of complex-

[1] This chapter is based on parts of the paper Anugerahanti, P., Roy, S. and Haines, K. (2018), 'A perturbed biogeochemistry model ensemble evaluated against in situ and satellite observations, Biogeosciences **15**(21), 66856711.

ity, the ability to reproduce bulk marine biogeochemical properties, and also the computation cost, and therefore becomes the ocean biogeochemical model for the first UKESM (Kwiatkowski et al., 2014). However, in this study, we used MEDUSA-1.0, because the model is less complex, and compared to MEDUSA-2.0, there is not much difference in the NPZ simulations between the two models (such as both models show an underestimation of chlorophyll in the oligotrophic gyre and a poor match between the modelled and SeaWiFS-derived chlorophyll (Yool et al., 2013)). The MEDUSA model will be discussed further, along with its default functional forms in section 2.1. Furthermore, the 1-D version of MEDUSA-1.0 is available within the Marine Model Optimisation Toolbox (MarMOT) (Hemmings and Challenor, 2012; Hemmings et al., 2015) where simulation of large ensembles can be done with sufficient computational cost, which will be explained in section 2.2. The method of generating the ensemble and optimising the functional forms so that they are equivalent to the MEDUSA default functional forms will be described in section 2.3. The model parameters are described in section 2.4. Running the 1-D MEDUSA using MarMOT is shown in section 2.5, along with the physical input, the initial conditions, and the validation data. The model statistical metrics are described in sections 2.7 and 2.8.

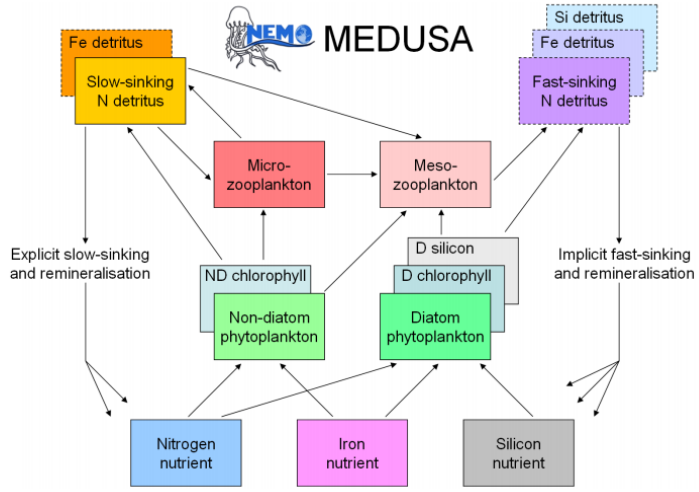
2.1 MEDUSA

MEDUSA is a biogeochemical model which spans beyond the basic NPZD model. The first version of this model (MEDUSA-1.0) has 11 state variables, including four plankton groups distinguished by size, silicic acid, chlorophyll, and iron. The state variables are:

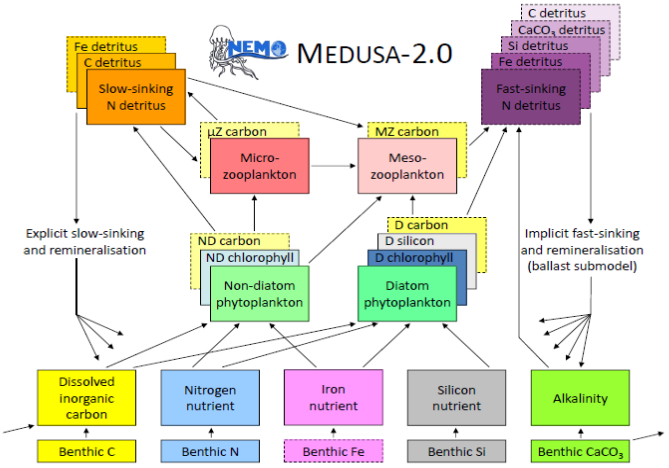
Non-diatom phytoplankton	mmol N m^{-3}
Diatom phytoplankton	mmol N m^{-3}
Chlorophyll in non-diatoms	mg m^{-3}
Chlorophyll in diatoms	mg m^{-3}
Silicon in diatoms	mmol Si m^{-3}
Microzooplankton	mmol N m^{-3}
Mesozooplankton	mmol N m^{-3}
Slow sinking detritus	mmol N m^{-3}
Dissolved Inorganic Nitrogen (DIN)	mmol N m^{-3}
Silicic acid	mmol Si m^{-3}
Iron	mmol Fe m^{-3}

Although nitrogen is the model currency, a diatom silicon state variable is explicitly represented which allows for dynamic Si:N ratio in diatoms, based on the model of Mongin et al. (2006). In order to permit regional growth limitation by the iron, an additional iron cycle is present. The second version, MEDUSA-2.0 incorporates the carbon and oxygen cycles by including total dissolved inorganic carbon, alkalinity, dissolved oxygen, and detrital carbon. The fluxes of carbon in MEDUSA-2.0, are calculated

from those of nitrogen using the Redfield ratio (Redfield, 1934). State variables for both versions of MEDUSA and their interactions are shown in figure 2.1.



(a) MEDUSA-1.0



(b) MEDUSA-2.0

Figure 2.1: Schematic diagrams showing state variables their interactions in MEDUSA 1.0 (2.1a) and the more complex MEDUSA-2.0 (2.1b). Explicitly modelled components are shown in boxes with solid borders, whereas dashed borders are implicit. Component that are modelled using other currencies (e.g chlorophyll, silicon) shown overlapped with other box. (Yool et al., 2011, 2013)

From Figure 2.1, although intermediately complex, the interaction between nutrients, phytoplankton, and zooplanktons are similar to that in an NPZ model, but with multiple nutrients and plankton groups. The four groups of plankton consist of two phytoplankton and two zooplankton groups which are distinguished by size. The small component includes prokaryotic nanophytoplankton (represented as non-diatoms) and microzooplankton which includes protists and metazoan larvae. The larger component consists of diatom phytoplankton and mesozooplankton. Diatoms are chosen to represent the larger phytoplankton since diatom is a key component and predominates in highly productive areas (Mann, 1999;

Yool et al., 2011). This means that the term non-diatom in this model merely ignores other larger species such as coccolithophorids or flagellates (Cox and Kwiatkowski, 2013). In zooplankton grazing, micro-zooplankton strongly control the population of smaller phytoplankton, whereas the mesozooplankton slowly graze the larger phytoplankton. Another assumption is that the small phytoplankton have better nutrient uptake kinetics and therefore faster growth (Furnas, 1990), which also applies to zooplankton (Baird and Suthers, 2007). Maximum phytoplankton growth, V_{p^T} is also dependant on temperature, by applying the Eppley (Eppley, 1972) function:

$$V_{p^T} = V_p \times 1.066^T \tag{2.1}$$

where T is temperature and V_p is the maximum phytoplankton growth rate at 0°C. However, this is not applied to zooplankton grazing nor growth. Chlorophyll quotas in phytoplankton are explicitly represented so that its response to irradiance will depend of the chlorophyll content .

MEDUSA-1.0 has 11 differential equations (15 in MEDUSA-2.0) that describe the biogeochemical processes in the model. Regardless of the horizontal or vertical positions, these equations are applied within every ocean grid cell in the physical model (Yool et al., 2011), unlike earlier model studies which distinguish between photic and aphotic zones (e.g. Popova et al. (2006)). Both versions of MEDUSA use the classic rectangular hyperbolic Michaelis-menten for nutrient limited phytoplankton growth, and given the initial slope of P-I curve $\hat{\alpha}_p$ and maximum phytoplankton growth V_{p^T}, Smith function (Smith, 1936) as described in Table 1.1, is used to calculate light limited phytoplankton growth rate, J_p, given local irradiance, I, the equation becomes:

$$J_p = \frac{V_{p^T}\ \hat{\alpha}_p\ I}{(V_{p^T}^2 + \hat{\alpha}_p^2\ I^2)^{1/2}} \tag{2.2}$$

In MEDUSA-2.0, I is the sum of two components of photosynthetically available radiation (PAR), red, and green-blue, which are derived from the simpler LOBSTER model (Lévy et al., 2001). The two different wavebands are attenuated separately by seawater and chlorophyll. Nutrient uptake and primary production via Redfield coupling is determined using a multiplicative term that brings together light and nutrient limitation factors.

Diatom growth is limited by the availability of silicic acid since they use this macronutrient to construct their cell walls or frustules such as spines and girdle bands (Martin-Jezequel et al., 2000). If the ratio is less than the critical value $R^0_{Si:N}$, diatoms are unable to grow (Martin-Jezequel et al., 2000). This makes diatoms have flexibility in their requirement of silicon. Above this minimum value, the uptake of iron and nitrogen is scaled by a factor of Si:N ratio, and if the ratio is three times the minimum, growth is unrestricted. In silicon uptake, the maximum rate is achieved when the Si:N ratio is below a critical threshold $(3 \times R^0_{Si:N})^{-1}$. Silicon uptake is linearly decreased to another threshold value, $(R^0_{Si:N})^{-1}$, and

above this threshold, no silicon is taken up by diatom cells, but diatom nitrogen can still increase and alter the Si:N ratio. The loss of silicic acid occur everywhere in the water column since the modern ocean is undersaturated with it (Yool and Tyrrell, 2003). This is represented by a linear loss rate following (Mongin et al., 2006).

Zooplankton grazing is separated by size and as explained previously, larger zooplankton have a broader prey range. This results in grazing preference. MEDUSA adopts the (Fasham et al., 1990) method of grazing preference, where an ingestion function balances the availability of prey items. It is assumed that both zooplankton groups prefer higher quality food items; non-diatoms in microzooplankton and microzooplankton and diatoms in mesozooplankton. Nevertheless this prey selectivity is biologically questionable (Cox and Kwiatkowski, 2013), because the predator-prey interactions are complex; including prey capture and satiation feedback (Flynn and Mitra, 2016). The nitrogen and carbon component ingested by zooplankton are separated. Inefficient grazing, due to zooplankton messy feeding, is returned back to dissolved nutrient. Then, the C:N ratio for the ingested material can be derived, and compared with the ideal ratio preferred by zooplankton. Using the assimilation and metabolism submodel of (Anderson and Pondaven, 2003), growth, excretion, and respiration are balanced according to which nutrient is limiting. Apart from grazing, all four plankton types have secondary losses to other processes, which are the density-independent loss terms, modelled using a linear function, and a density-dependent loss term which are modelled using rectangular hyperbolic terms as it provides a source of stabilising feedback (Steele and Henderson, 1992).

MEDUSA-2.0 is chosen as the ocean biogeochemical part of UKESM1 because the model shows a better overall match with the observational data, and it is computationally inexpensive compared to other more complex models. This study also shows no evidence that the more complex models produce a better match with the bulk properties. Complex marine biogeochemical models may represent the ocean ecosystem better, however, some of the ecology is still poorly understood (Anderson, 2005). For example, from the experiment, the degradation of dissolved organic matter in the ocean is not dependent on only one group of bacteria, and therefore only having one compartment of bacteria is not very accurate (Anderson, 2005). Adding more complexity also means increasing the number of equations and therefore the model parameters. The magnitude of g_{max} or μ may influence how much phytoplankton is lost due to grazing and mortality. As summarised in Table 1.1, each process in an NPZ model can be modelled using a variety of functional forms, so adding more process may also means more functional forms to choose. Additionally, different ocean circulation models may result in different simulations of the state variables, as the physical processes controls the biogeochemical processes. These problems show that the marine biogeochemical model has considerable uncertainties.

2.1.1 The default MEDUSA functional forms

The base of both MEDUSA-1.0 and 2.0 include nutrient uptake, zooplankton grazing, and plankton mortalities, which are represented using functional forms. The way the MEDUSA differential equations and default functional forms of MEDUSA is most similar to the NPZDB model by Fasham et al. (1990) and NPZD model by Popova et al. (2006). However, unlike that in the NPZDB in MEDUSA, instead of having two types of DIN (ammonium and nitrate), and bacteria, the NPZD compartments of MEDUSA has DIN, silicon, and iron, along with four types of plankton.

For nutrient limited growth, the model uses the classic Michaelis-Menten (rectangular hyperbolic) equation, where phytoplankton take up the ambient nutrient. Since the maximum growth is decided using the eppley function, the parameter that is adjustable for this term is the half saturation constant. The phytoplankton growth is brought together in a multiplicative term between light limited, J_p, as described in equation 2.2 and nutrient limited growth. Primary production of diatom PP_d can then be calculated via Redfield coupling:

$$\mathrm{PP}_d = J_p \times \frac{N}{k_N + N} \times \frac{Fe}{k_{Fe} + Fe} \times \frac{Si}{k_{Si} + Si}$$

The first term of the primary production equation is the light limited growth. The second, third, and fourth terms are the nutrient limited growth modelled using Michaelis-Menten term, with N, Fe, and Si representing the nutrients DIN, Iron, and silicon respectively.

In terms of grazing, instead of using multiple resource grazing using Holling type II, as in (Fasham et al., 1990), the functional form that describes zooplankton grazing is Holling type III, as (Ryabchenko et al., 1997). Multiple prey grazing $G_{p,q}$ can be expressed as a weighted fraction of the total ingestion $G_{T,p}$, therefore:

$$G_{p,q} = G_{T,p} \frac{w_{p,q} D_p}{D_{T,q}}$$

where:

$$D_{T,q} = \sum_q w_{p,q}, D_p$$

Where D_p is the density of prey type p and $D_{T,q}$ is a weighted measure of the total available food. The term $w_{p,q}$ is the preference for prey type p grazed by zooplankton q, which corresponds to the relative proportion that prey p occurs in the diet of zooplankton q compared to the relative proportion that prey p occurs in the environment (Gentleman, 2002). This makes Holling type III function for

microzooplankton, Z_{mi} that into:

$$G_{Pr} = \frac{g_{max} \cdot w_{Pr} \cdot Pr^2 \cdot Z_{mi}}{k^2 + w_{Pn}^2 \cdot Pn^2 + w_D^2 \cdot D^2}$$

where w_{Pr} is the microzooplankton grazing preference for a prey Pr, which could be non-diatom, Pn, or detritus, D. Since non-diatoms are assumed to be more nutritious compared to detritus, the weighting on non-diatoms is higher (0.75) than on detritus (0.25). As explained previously, for plankton mortality terms, MEDUSA has both density independent and density dependent mortality rates for all the phytoplankton and zooplankton types. Density-independent loss is modelled by a linear function representing plankton metabolic loss. Density-dependent loss, which includes processes such as higher-trophic grazing and disease, is modelled using the hyperbolic function of plankton concentration (Fasham et al., 1993). All of the default functional forms are summarised in Table 2.1.

Both MEDUSA-1.0 and MEDUSA-2.0 have similar default functional forms which have been described above. As discussed above, MEDUSA-1.0 is the simpler version of MEDUSA-2.0 and only focuses on the NPZD compartments, and the results from MEDUSA-2.0 and MEDUSA-1.0 are quite similar in terms of geographical distribution of primary productivity and limiting nutrients. However slightly higher surface phytoplankton and microzooplankton concentrations are observed globally from MEDUSA-1.0, but for nutrients and mesozooplankton, MEDUSA-2.0 globally produces slightly higher concentrations (+9.1% in DIN, +5.2% in silicon, and +5.8% in Iron) especially in the Atlantic and the Indian Oceans (Yool et al., 2013). As a consequence of the reduced phytoplankton concentrations, MEDUSA-2.0 shows less pronounced seasonality at higher latitudes in the Northern Hemisphere compared to MEDUSA-1.0 (Yool et al., 2013). These differences may be caused by the different simulation years between MEDUSA-2.0 (145 years) and MEDUSA-1.0 (40 years), and also the differences in sinking rate, where MEDUSA-2.0 uses 2.5 m d^{-1} instead of 3 m d^{-1} and the updated iron deposition, by using a 'present day' field produced by Mahowald et al. (2005). Due to these minor differences, we have decided to use the simpler MEDUSA-1.0 model.

Before moving to other functional forms that have similar properties to these default functions and generating the ensemble, it is possible to address uncertainty in MEDUSA using MarMOT, which has the 1-D version of MEDUSA-1.0 embedded in it. This will reduce computational cost and alternative versions of the MEDUSA model e.g. an ensemble, can be run using a standard PC. The next section will introduce MarMOT and how to run MEDUSA using MarMOT.

2.2 Addressing model uncertainty using Marine Model Optimisation Testbed

Marine Model Optimisation Testbed (MarMOT) is a site-based emulator which relies on an array of 1-D simulators of the target model dynamics. The full description of MarMOT can be found in Hemmings and Challenor (2012) and Hemmings et al. (2015). This model optimisation testbed is used to examine the impact of error in a biogeochemical model's environmental input data, including the physical environment and adjustable parameters. MEDUSA is integrated into a 1-D scheme where the environmental input from the site-based simulation is acquired from the physical and biogeochemical information from the 3-D model. By using MarMOT, various physical frameworks can be used as an input as the biogeochemical parameters are independent of the physics. However the biogeochemical environment is dependent on the parameters. The Marine Model Optimization testbed facility is used to configure the 1-D simulator for MEDUSA (Hemmings et al., 2015).

Using MarMOT, the biogeochemical state variable (i) concentration (c_{ik}) is calculated at each depth level (k) in the 1-D simulator, using the first and fifth term of equation 1.2.2:

$$\frac{d_{c_{ik}}}{dt} = -(w_p + w_i)\frac{\partial c_i}{\partial z} + \frac{\partial}{\partial z}\left(K_\rho \frac{\partial c_i}{\partial z}\right) + \mathrm{SMS}_{ik}(\boldsymbol{C}, \boldsymbol{F}) \tag{2.3}$$

The rate of change due to vertical flux divergence is described in the first term, where w_p and w_i are the vertical velocity of the water, provided by the physical framework, and the biological material relative to the water respectively, with K_ρ the turbulent mixing diffusion coefficient, which is obtained from the physical framework. The source-minus-sink term, SMS_{ik} from the MEDUSA plankton model is embedded in the third term as a function of the state vector $\boldsymbol{C}$ and forcing vector $\boldsymbol{F}$, including temperature, downwelling solar radiation at sea surface, which are provided from the physical framework, and soluble iron from atmospheric dust deposition input. Since light available for phytoplankton photosynthesis and the nutrient sources from the remineralisation of fast-sinking detritus depend on tracer concentrations at one level shallower than the current level, SMS_{ik} is depth-dependent. w_i is a constant sinking rate for detritus, which corresponds to the slow sinking parameter from MEDUSA. Initially, MarMOT is developed to explore robustly the behaviour of global biogeochemical models by using perturbed parameter ensemble with multiple members on a basis of analysis performed on representative arrays of 1-D simulators. This will also allow large parameter space investigation (Hemmings et al., 2015).

Since MarMOT allows running the 1-D version of a marine biogeochemical operational model, it is possible to run a large number of ensemble members containing different functional form combinations, which describe the essential biogeochemical processes. Furthermore, it is also possible to vary the physical input and compare these uncertainty within the same model using MarMOT. The full method for

simulating the 1-D MEDUSA in MarMOT is described in appendix A.

2.3 Generating the ensemble: Optimising the functional forms

Different marine biogeochemical models share common processes: the interaction between nutrients, phytoplankton, and zooplankton shown on figure 2.2. These processes could be modelled using different equations. For example multiple predator uptake has two functional forms that could be used to represent the process; using Michaelis-Menten expression or Holling type II (Fasham et al., 1990) and Holling type III (Ryabchenko et al., 1997). They have similar saturation curves despite being described by different mathematical equations. Structurally different analytical forms could be used interchangeably, given the uncertainty with which resource uptake by real organisms is measured (Fussmann and Blasius, 2005). Functions that show similar shapes to MEDUSA's default functional forms are chosen to ensure that they describe the biogeochemical processes in a similar way.

The non-linear least square method is used to enhance the phenomenological similarity between the default functional forms. The parameters associated with the functions, apart from the maximum rate, are obtained by using the command `'scipy.optimize.curve_fit'`, which uses non-linear least squares, in python. Here we made the maximum rate values for other functional forms similar to MEDUSA's default parameter, so that these functions would become saturated at the same concentration. Different functional forms and their parameters are then embedded into the model code. The ensemble model has 128 members, which resulted from combining four phytoplankton nutrient uptake, two zooplankton grazing, four phytoplankton, and four zooplankton mortalities functions.

All possible combinations of functional forms that describe the main biogeochemical processes (such as nutrient uptake, grazing, and mortality, explained in section 2.3.1 to 2.3.3) can be generated as an ensemble, whereby each member contains a combination of functional forms similar to the default MEDUSA functions. These combinations are then embedded into the 1-D MEDUSA code. The same process function is always used for both diatoms and non-diatoms, or mesozooplankton and microzooplankton. Each ensemble member has at least one functional form changed from the default functions. This provides a total number of 128 combinations, arising from 4 types of nutrient uptake, 4 phytoplankton mortality formulations, 2 types of zooplankton grazing, and 4 zooplankton mortalities.

2.3.1 Nutrient uptake ensemble

Alongside light, nutrient concentration limits the growth of phytoplankton. In MEDUSA the standard hyperbolic monod, hereafter U_h, function is used as the default function. The growth of cells monotoni-

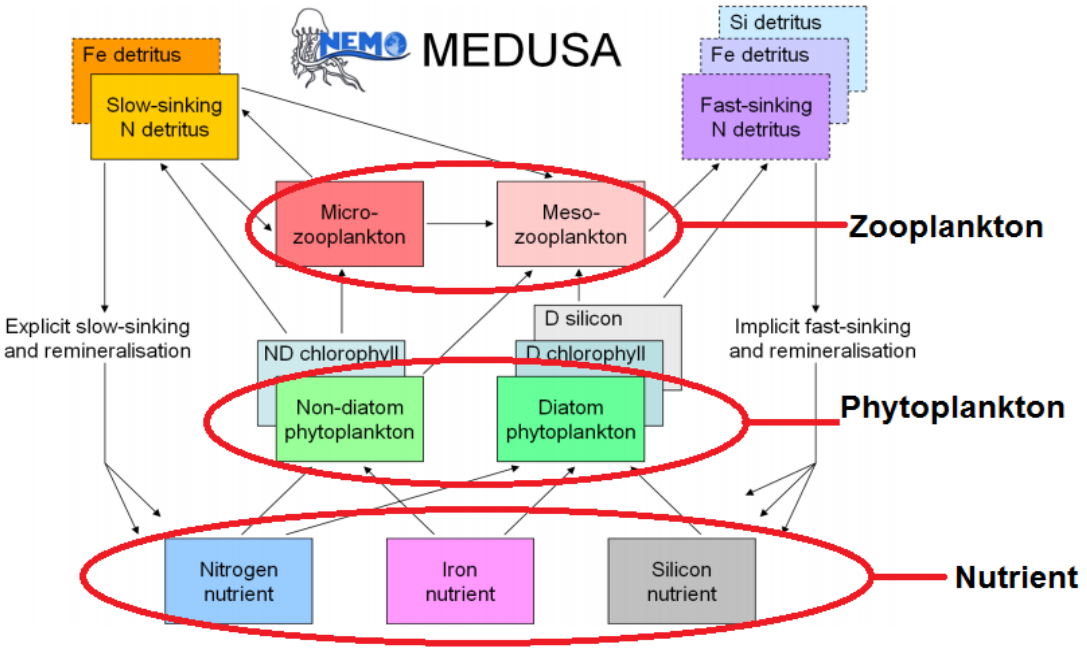


Figure 2.2: Schematic diagrams of MEDUSA-1.0 adopted from Yool et al. (2011), red circles shows core biogeochemical process. These are represented by nutrient uptake by phytoplankton, zooplankton grazing, and plankton mortality, which could be parameterised by different equations

cally increases with ambient nutrient concentration, and halts when nutrients become scarce. If nutrient concentrations are high, the rate of uptake saturates. Other mathematical functions show similar properties including (i) Sigmoidal (Fennel and Neumann, 2014) U_s, (ii) the exponential (Ivlev, 1961), U_e, and (iii) trigonometric functions (Jassby and Platt, 1976), U_t. All these functions include a shape-defining parameter, k, which for monod and sigmoidal can be interpreted as a half saturation constant, and a maximum uptake rate, described in equation 2.1. As explained in section 2.1, MEDUSA has silicon and iron nutrients as well as diatoms and non diatoms. The uptake of different phytoplankton types and nutrients use similar functions but different parameter values for k, summarised in Table 2.1, obtained by minimising the sum squared difference with U_h. The nutrient uptake functions after optimization are shown in Figure 2.3(a). The fit is done over the nutrient concentration ranging from 0.001 to 20 mmol m^{-3} for DIN and silicon. These are discretised into 1000 intervals. The difference in shape of the optimised functional forms are more obvious between 0.1 to 1 mmol N m^{-3}.

2.3.2 Zooplankton grazing ensemble

In MEDUSA, both phytoplankton and zooplankton are grouped into ”small” and ”large” categories. The small zooplankton, represented by the microzooplankton, graze on non-diatoms and detritus, with the more nutrient rich, higher quality, non-diatoms preferred over detritus. Larger zooplankton, represented by mesozooplankton have a broader range of prey, including both microzooplankton and diatoms, which

are higher quality food sources compared to non-diatoms and detritus.

Describing multiple grazing using mathematical forms can be done by defining the zooplankton grazing rate using the hyperbolic Michaelis-Menten or Holling type II, hereafter G_2 expression and weighted preference on the different food sources p_n (Fasham et al., 1990). Suppose the specific grazing rate, G, is described by:

$$G = g\frac{F}{k_g + F} \tag{2.4}$$

where F is the total food with their preferences by grazers ($p_a Pa + p_a Pb$), g is the maximum grazing rate, and k_g is the half-saturation constant for grazing. When Pa is grazed constantly, equation (2.4) is then scaled into the zooplankton's food preference by substituting the preference scaled prey concentration:

$$G_{Pa} = g\frac{p_a Pa}{k_g + (p_a Pa + p_b Pb)} \tag{2.5}$$

Since zooplankton preferences will change throughout the year, the assigned preference should change as a function of the food ratio. This could be achieved by defining the weighted preference, p_a^* and p_b^*:

$$p_a^* = \frac{p_a Pa}{p_a Pa + p_b Pb} \text{ and } p_b^* = \frac{p_b Pb}{p_a Pa + p_b Pb} \tag{2.6}$$

Substituting p_a^* and p_b^* for p_a and p_b, in equation (2.5), grazing on Pa is described by:

$$G_{Pa} = g\frac{p_a Pa^2}{k_g(p_a Pa + p_b Pb) + p_a Pa^2 + p_b Pb^2} \tag{2.7}$$

Another method of multiple grazing parameterisation is based on sigmoid Holling type III function, which is defined as:

$$G = \frac{gF^2}{k_g^2 + F^2} \tag{2.8}$$

In the case of two resources, equation (2.8) for Pa uptake becomes:

$$G_{Pa} = g\frac{p_a Pa^2}{k_g^2 + p_a Pa^2 + p_b Pb^2} \tag{2.9}$$

Equation (2.9) is the default functional form used in MEDUSA for zooplankton grazing. As shown in figure 2.3(b) equation(2.7) and (2.9) have similar trends where grazing rate becomes constant as it reaches a certain phytoplankton concentration and a half saturation constant k_g. During the fitting process, the range of phytoplankton and microzooplankton concentrations used was 0.001 to 10 mmol m^{-3}, discretised in 1000 intervals equally. At low zooplankton concentrations (between $\sim$0.01 to $\sim$0.5 mmol m^{-3}) the Holling type III response has lower grazing rates than the hyperbolic, however as the phytoplankton concentration increases, the Holling type III curve has a more rapid increase in predation rate before

becoming saturated (Edwards and Yool, 2000) compared to the Holling type II, shown on Figure 2.3(c). Preferences for food types are kept the same as MEDUSA's default parameters, with terms summarized in Table 2.1.

2.3.3 Plankton mortality ensemble

Alternative functions can describe the density-dependent mortality, and in this study the combinations of hyperbolic (ρ_h, ζ_h), linear (ρ_l, ζ_l), quadratic (ρ_q, ζ_q), and sigmoidal (ρ_s, ζ_s) functions to describe the phytoplankton (ρ) and zooplankton (ζ) mortalities are used (equations and abbreviations are shown on Table 2.1). Similar to grazing and nutrient uptake, the functional forms have different maximum rates for each plankton type. These maximum rates are made the same for all the different functions.

Of the four different mortality functions, linear and quadratic functions are most different in shape, as shown on Figure 2.3(c). Using the linear term is similar to constant removal of plankton at the same rate as the maximum mortality (μ). To make the linear function similar to the sigmoidal and hyperbolic functions, the maximum mortality rate is set so that the total loss integrated over the range of phytoplankton concentrations (calculated as the area below the function representing the total loss in linear terms, between 0.001 to 10 mmol m^{-3}) is similar to that for the hyperbolic curve. The quadratic term, instead of asymptoting, continues to grow with plankton abundance. In order to keep this similar to other forms, after reaching a certain concentration the function is switched to linear, so that the rate plateaus at high abundance. For sigmoidal mortality, the default μ are not changed but the half-saturation constant, k_M is optimised. The optimised mortality functions are shown in Figure 2.3(c). The range of phytoplankton and zooplankton concentrations used during the fitting process was between 0.001 - 10 mmol m^{-3}, and discretised within 1000 intervals equally. A distinctive feature of these functional forms after optimisation is that the quadratic mortality rate remains low until phytoplankton concentration reaches 1.0 mmol m^{-3}, and the linear function shows consistently high plankton mortality (Figure 2.3(c)).

2.4 MEDUSA parameters that have been changed

Apart from sinking rate, maximum growth, and grazing rates, parameters not listed in Table 2.1 are kept at their default values (Yool et al. (2011), described in Appendix A.2). From a previous 3-D MEDUSA run, the oligotrophic regions show a low 'background' chlorophyll concentration (Yool et al., 2011) so to raise this concentration a higher maximum growth rate and lower grazing rate have been used. The maximum uptake rate, V_p , is 0.8 day^{-1}, similar to that in the HadOCC model (Palmer and Totterdell, 2001). For zooplankton grazing, similar to NPZ models (Fasham et al., 1990; Fasham, 1995; Anderson

Table 2.1: Parameter values for resource uptake (U), zooplankton grazing (G), and plankton mortalities (ρ and ζ for phytoplankton and zooplankton respectively), described using similar functional forms (shown in Figure 2.3). In grazing equation, g_m represents maximum grazing rate, P_a is the prey, and p_a denotes the grazing preference. Starred equations are the default functional responses in MEDUSA.

Process/ Plankton type	Symbol	Meaning	Parameter value (mmol m^{-3})			
Nutrient Uptake (U)			Monod* (U_h) $\frac{n}{n+k}$	Sigmoidal (U_s) $\frac{n^2}{n^2+k^2}$	Exponential (U_e) $1-\exp(-\frac{n}{k})$	Trigonometric (U_t) $\frac{2}{\pi}\arctan\left(\frac{n}{k}\right)$
Non-diatom	kN_{nd}	shape defining constant for DIN	0.5	0.74	1.12	0.60
	kFe_{nd}	shape defining constant for iron	0.33 $\times 10^{-3}$	0.49 $\times 10^{-3}$	0.74 $\times 10^{-3}$	0.40 $\times 10^{-3}$
Diatom	kN_d	shape defining constant for DIN	0.75	1.12	1.68	0.91
	kSi_d	shape defining constant for silicon	0.75	1.12	1.68	0.91
	kFe_d	shape defining constant for iron	0.67 $\times 10^{-3}$	0.99 $\times 10^{-3}$	1.50 $\times 10^{-3}$	0.81 $\times 10^{-3}$
Grazing (G)			Holling type III* (G_1) $g_m \frac{p_a Pa^2}{k_g^2+p_a Pa^2+p_b Pb^2}$		Holling type II (G_2) $g_m \frac{p_a Pa^2}{k_g(p_a Pa+p_b Pb)+p_a Pa^2+p_b Pb^2}$	
Microzooplankton	k_{mi}	half saturation constant	0.80		0.46	
	pmi_{nd}	grazing preference for non-diatom	0.75		0.75	
	pmi_{det}	grazing preference for detritus	0.25		0.25	
Mesozooplankton	k_{me}	half saturation constant	0.30		0.17	
	pme_{nd}	grazing preference for non-diatom	0.15		0.15	
	pme_{det}	grazing preference for detritus	0.15		0.15	
	pme_d	grazing preference for diatoms	0.35		0.35	
	pme_{mi}	grazing preference for microzooplankton	0.35		0.35	
Mortality (ρ, ζ)			Hyperbolic* (ρ_h, ζ_h) $\mu \frac{P}{P+k_M} P$	Linear (ρ_l, ζ_l) μP	Quadratic (ρ_q, ζ_q) μP^2	Sigmoidal (ρ_s, ζ_s) $\mu \frac{P^2}{P^2+k_M^2} P$
Non-diatom	μ_{nd}	maximum rate (day^{-1})	0.10	0.09	0.05	0.10
	k_{Mnd}	half saturation constant	0.50	-	-	0.74
Diatom	μ_d	maximum rate (day^{-1})	0.10	0.09	0.05	0.10
	k_{Md}	half saturation constant	0.50	-	-	0.74
Microzooplankton	μ_{mi}	maximum rate (day^{-1})	0.10	0.09	0.05	0.10
	k_{Mmi}	half saturation constant	0.50	-	-	0.74
Mesozooplankton	μ_{me}	maximum rate (day^{-1})	0.20	0.19	0.07	0.20
	k_{Mme}	half saturation constant	0.75	-	-	1.12

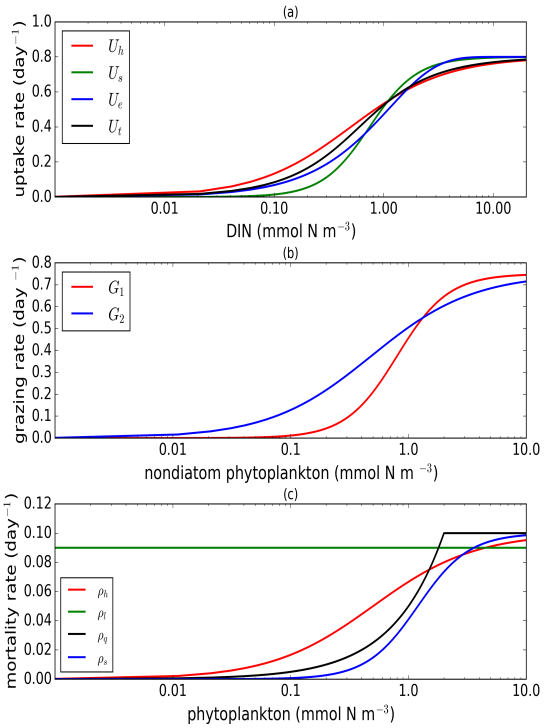


Figure 2.3: Nearly identical curves which describes resource uptake (a), zooplankton grazing (b), and phytoplankton mortality (c). Figure (a) shows four uptake functions, which have been optimised to the default uptake function, monod (U_h). Figure (b) shows two grazing functional forms, the holling type III (G_1) and type II (G_2) functions. Four phytoplankton mortality functions are shown on figure (c), whereby hyperbolic is the default function. The optimisation method is describe in section 2.3.1, 2.3.2, and 2.3.3. The range for DIN in (a) is between 0.001 and 20 mmol m^{-3}, and phytoplankton in (b) and (c) are 0.001 and 10 mmol m^{-3}. Table 2.1 describes the function's equations and parameters.

et al., 2015) we use 1 day^{-1} as the maximum grazing rate, g_m. MEDUSA also parameterises both slow and fast detritus sinking factors. It is assumed that the latter sinks rapidly relative to the model time-step, and remineralisation of the detrital nitrogen and silicon is done implicitly. In the default model 3 m day^{-1} is used for the slow sinking of detritus, however over long runs we found this leads to downward loss of nutrients from the euphotic zone, to the sea floor. Earlier studies have used lower detrital sinking rates (Steele and Henderson, 1981; Fasham et al., 1990; Lacroix and Gregoire, 2002; Raick et al., 2006), between 0 to 1.25 m day^{-1} and a study has even suggested to use 0 m day^{-1}, (Ward et al., 2013). In this study a sinking rate of 0.1 m day^{-1}, towards the lower end of the range of literature values, is chosen to prevent depletion of state variables particularly at the shallower stations.

2.5 Running the Model and Generating the Ensemble

MEDUSA is run in the Marine Model Optimization Testbed (MarMOT-1.1 alpha) (Hemmings and Challenor, 2012; Hemmings et al., 2015). As explained in section 2.2, MarMOT was developed to investigate

the effect of sensitivity in plankton model simulations, especially in regard to parameter and environmental inputs (Hemmings and Challenor, 2012). Despite some uncertainties associated with the differences in physical forcing, fluxes, and initial values of biogeochemical properties, using 1-D simulations to approximate 3-D model behaviour for calibrating models based on specific sites has improved the predictive skill of 3-D models (Oschlies and Garçon, 1999; Oschlies and Schartau, 2005; Kane et al., 2011; McDonald et al., 2012). As discussed in section 1.3, parameter and physical forcing uncertainties vary across different regions, as the biogeochemistry in the regions have different physical controls. In order to robustly examine the structural uncertainty in a marine biogeochemical model, it is essential to examine this uncertainty in different regions or sites. Since MarMOT permits running multiple ensembles at different sites, the 1-D MEDUSA is run at five oceanographic stations, described in section 1.6. The ensemble at each station is initialised using in situ measurements of chlorophyll, DIN, silica, and iron, and is run over 10 years starting from January 1998, following the test-data and availability of NEMO output for physical input, and in situ data. In the next subsections, the model inputs for MarMOT will be described in the next subsections.

2.5.1 Physical input

Physical input files consist of gridded values of vertical velocity (w, m day^{-1}), vertical diffusion coefficient (vdc, m^2 day^{-1}), and temperature (T, °C), which are applied at each depth level. Additionally, time series of downwelling solar radiation (W m^{-2}) and mixed layer depth (m) are also used as input. These are obtained from the 5-day mean output of the Nucleus for European Modelling of the Ocean (NEMO) model, using the Met Office Forecast Ocean Assimilation Model (FOAM), which controls the physical parameters and therefore the biogeochemical tracers every 5 days. The FOAM-NEMO system assimilates satellite-derived sea surface temperature, sea-level anomaly, sea-ice concentration, temperature, and salinity profile data, in order to make the physical system more realistic (Storkey et al., 2010).

However, assimilating physical data into a coupled physical biogeochemical model does not improve the simulation of the ecosystem. A study by Ourmières et al. (2009) using the LOBSTER model, shows that although assimilating physical data has improved the primary production at the Labrador Sea (due to increasing eddy activity), it does not follow the pattern from SeaWiFS derived chlorophyll-a. When FOAM is used with the 3-D HadOCC model, it simulated an overestimation of nutrient concentration, compared to NEMO without data assimilation, due to the change in isopycnal levels caused by spurious vertical velocities (Ford et al., 2012; Ourmières et al., 2009). During the initial run, when NEMO-FOAM output is used to run MEDUSA, a large discrepancy was observed between the in situ and modelled nutrient distributions, especially at ALOHA, BATS, and Cariaco. The effect of using NEMO-FOAM on

the nutrients in 1-D MEDUSA is going to be explored in subsection 2.5.3.

Since input data on the vertical diffusivity was not stored in FOAM, we used values from NEMO ORCA025-N102 output from January 1998-December 2001 and from ORCA0083-N01 from January 2002-December 2007, both obtained from the CEDA Group workspace web (`http://gws-access.ceda.ac.uk/public/nemo/#_top`). These physical inputs are 5-day averaged and are available at 63 depth levels (from 6 to 5800m) for NEMO ORCA025-N102, with the resolution of 1/4° and ORCA0083-N01, with finer resolution at 1/12° and more depth levels (from 0.5 to 6000m). The level thickness increases exponentially with depth.

2.5.2 Biogeochemical input

The 1-D MEDUSA ensemble is run at five oceanographic stations: PAP, ALOHA, BATS, Cariaco, and L4, as these stations provide relatively long term biogeochemical observation data that are sampled regularly, such as chlorophyll and DIN. The input for the biogeochemical environment are the initial conditions for the 11 primary tracers (state variables) including; DIN, non-diatom, diatom, silicon in diatom, silica, detritus, microzooplankton, mesozooplankton, non-diatom chlorophyll, diatom chlorophyll, and iron (mmol m^{-3}), along with the model parameter values. Initial conditions for chlorophyll, silicate, iron, and DIN concentrations and are taken from the in situ data at the five oceanographic stations. We did not use spin up runs when initialising, which will be discussed in section 2.6. Location coordinate, data source, and maximum depth are summarised in Table 2.2 and the stations locations are shown in Figure 1.8. After initialization, in situ data from these stations are used to compare the in situ observation with the model results. For station PAP, because surface chlorophyll data is not available, we also use SeaWIFS-derived chlorophyll-a data with 9 km spatial resolution and 8-day averages provided by GlobColor (`http://hermes.acri.fr/`) for comparing the surface chlorophyll.

Table 2.2: Location, data source, and available depth range for the five oceanographic stations

Station	Location	Source	depth range
ALOHA	22°45'N, 158°00'W	`http://hahana.soest.hawaii.edu/hot/hot-dogs/interface.html`	5-5000 m
BATS	32°50'N, 64°10'W	`http://bats.bios.edu/`	4-4000 m
Cariaco	10°30'N, 64°40'W	`http://imars.marine.usf.edu/cariaco`	1-1310 m
L4	50°15'N, 4°12.3'W	`http://www.westernchannelobservatory.org.uk/data.php` (available upon request)	surface
PAP	49°N, 16.5°W	`http://projects.noc.ac.uk/pap/data`	7-400m

At these stations, the DIN consists of ammonia, nitrate, and nitrite, however at oligotrophic stations like ALOHA the ammonium is below the detection limit (Hawaii Ocean Time Series, 2017), and therefore DIN only consists of nitrate and nitrite. At PAP we use the initial condition from one of MarMOT's test stations, located at 50°N, 20°W in the Atlantic (Hemmings et al., 2015), since the nitrate data were only

collected between 30-400 m. At station L4 chlorophyll and DIN data were collected from the surface from 1999-2008. Since the maximum depth in this station is only 50 m deep, the initial concentrations for chlorophyll and DIN are the same at every depth (total chlorophyll = 0.27 mg m^{-3}, DIN = 6 mmol m^{-3}). Other inputs that are not available at the websites mentioned above, such as microzooplankton, mesozooplankton, and detritus were taken from the nearest test stations. In the oligotrophic stations, 75 % of total chlorophyll was allocated initially to the non-diatom phytoplankton since these dominated the water column (Villareal et al., 2012). At the other stations half of the total chlorophyll goes into the diatoms.

For validation of the model, we consider the total chlorophyll-a concentration, instead of separating diatoms and non-diatoms. The model is simulated at 37 depth levels, spanning from 6-1200 m, instead of the full 63 depth level (6-5800 m) to minimise computational cost, apart from station L4, with maximum depth of 50 m, and Cariaco, where the maximum depth for the physical input is available down to 500 m, although the depth at which nutrients are sampled are down to 1310 m. The boundaries for the depth levels are as follows: 6, 12, 19, 25, 32, 39, 46, 54, 62, 71, 80, 90, 100, 112, 124, 137, 152, 168, 187, 207, 229, 254, 281, 312, 347, 386, 429, 477, 531, 591, 656, 729, 809, 896, 991, 1093, and 1200 m. At the lowest level, vertical velocity and diffusion are set to zero and this level is a sink for detritus. Stations that have shallower maximum depths are run with fewer depth levels. Additionally, apart from the physical input files a time series for soluble iron flux from dust deposition is applied, but this is constant using the average value from (Mahowald et al., 2009), but the iron profiles data are available from the websites mentioned in Table 2.2.

2.5.3 Uncertainty in physical input: NEMO FOAM

As explained in section 2.5.1, directly assimilating the physical oceanography model cause an overestimation of nutrients concentration. Therefore, to avoid overestimation the vertical velocities from the FOAM system were capped at the 90^{th} and 10^{th} quantiles, and the 10-year mean of the vertical velocity is also removed. This means that the time mean of vertical velocity is zero, summarised in Figure 2.4. These adjustments gave a better long-term vertical structure to the nutrient and other distributions.

In this subsection, we will examine the effect of using NEMO FOAM to the DIN distribution and concentration at the five oceanographic stations that has been described in Table 2.2. We will consider the correlations, p values, RMSE, and bias from the default run, ensemble median, and mean, which are summarised in table 2.3. Primarily in MEDUSA, nitrogen is the model currency (Yool et al., 2011). Inorganic nitrate, NO_3, and nitrite, NO_2, were used as an initial condition for DIN, however if the station has ammonia, NH_4, this nutrient is also added to the DIN pool.

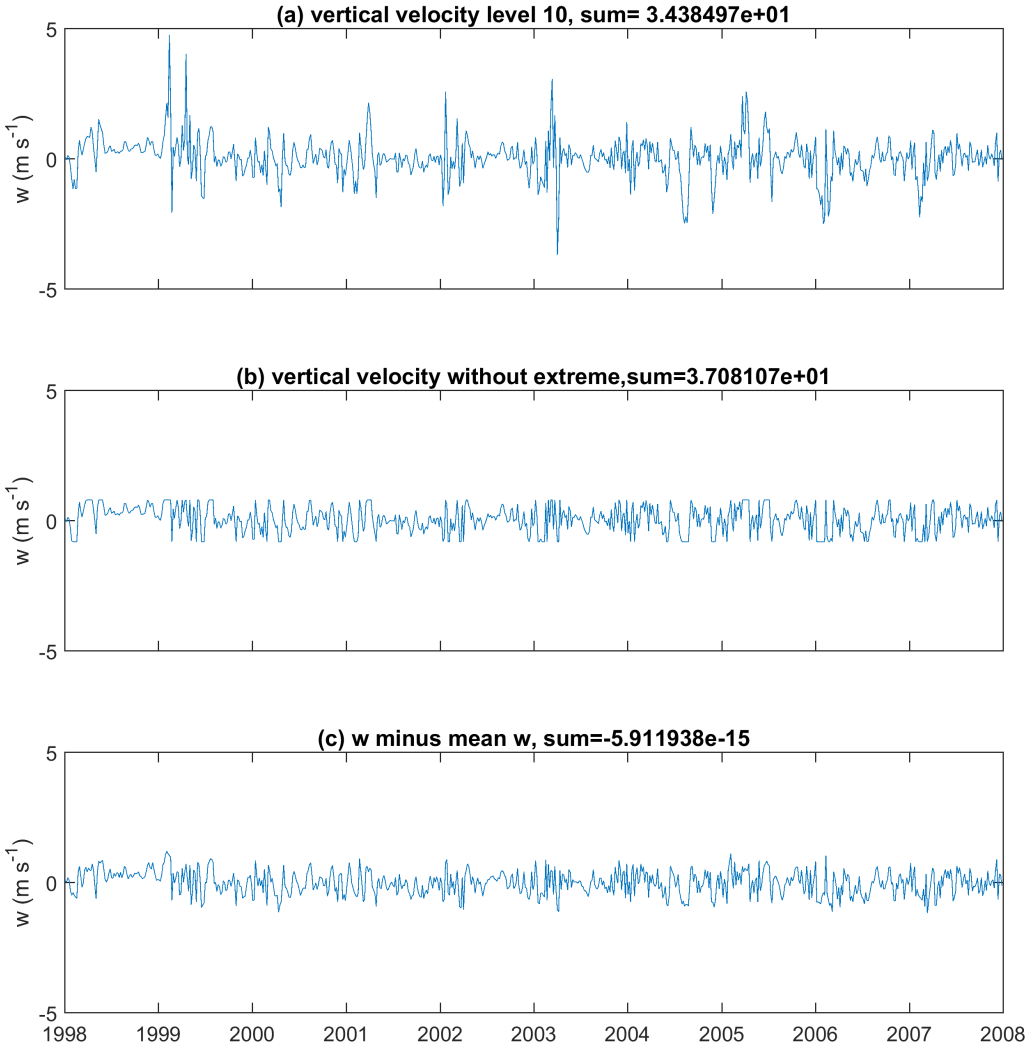


Figure 2.4: Vertical velocity, w, at station ALOHA at Level 10 (71m). Subfigure (a) shows the original vertical velocity, (b) shows the capped vertical velocity, and (c) shows the capped vertical velocity with its annual mean subtracted, which is the chosen physical input

At station ALOHA, the distribution of DIN changes over time in the simulation, especially after 1999, nitrocline has become shallower, going from 400m to 200m, whereas in situ DIN distribution show a consistent nitrocline at 400m. Nonetheless, low nutrients of > 4 mg m^{-3} in the top 150m is represented well in the model simulation. In contrast, simulation at station BATS shows a deepening in nitrocline by the end of 2002, to the point where it gets deeper than 1000m. However from mid 1999 to the end of 2001, the nitrocline got shallower from 500m to 200m 2.5, whereas in situ DIN shows a stable nitrocline over the years, apart from occasional shallowing in 2001 and 2007. Exhaustion of nutrient at the first 50m is also not well represented in the model simulation at BATS, especially after the deepening of the nitrocline. These are shown on figure 2.5

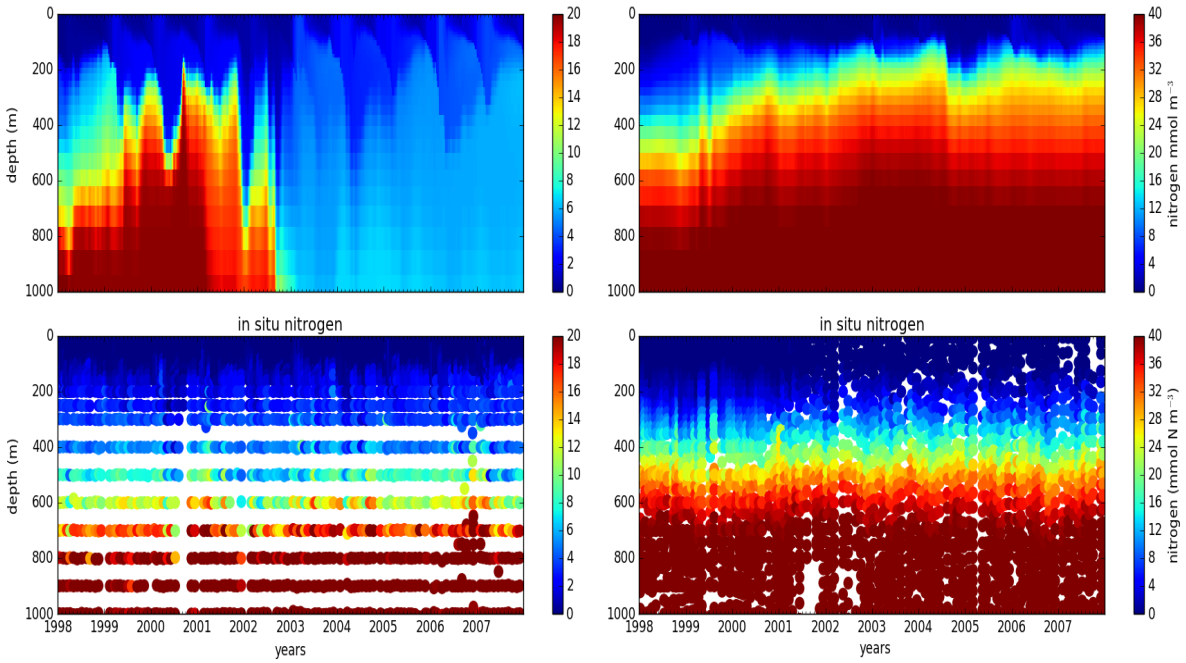


Figure 2.5: (top) ensemble mean and (bottom) *in situ* DIN profiles in station BATS (left) and ALOHA (right).

In simulating DIN, there are no noticeable differences between the ensemble mean, median, and default run. Simulated DIN at Station ALOHA shows RMS differences of 8.06, 8.05 and 8.16 mmol m^{-3} for ensemble mean and median and default run, respectively, and is overestimated compared to in situ nitrate. DIN at station BATS has a smaller bias than at station ALOHA, and is slightly underestimated. Despite the discrepancy in the DIN distribution, the RMS difference and bias at station BATS is lower that those in station ALOHA. Despite the discrepancy in distributions, both stations show relatively high correlations between in situ DIN and all the ensemble members ($r > 0.7$, $p < 0.05$ and $r > 0.9$, $p < 0.05$ at station BATS and ALOHA respectively).

DIN distribution at station Cariaco is rather unique. In this region, as discussed in Chapter 1, section 1.6.1, a nutrient rich layer with high salinity occurs between 100 and 200m, called the Subtropical

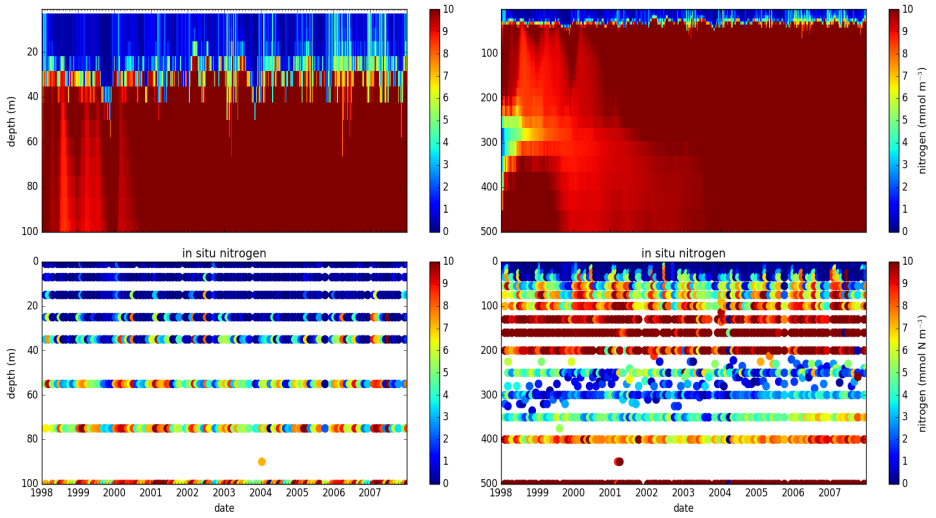


Figure 2.6: (top) ensemble mean DIN at the first 100m (right) and 500m (left) and (bottom) *in situ* profiles in station Cariaco

Under Water (SUW). Strong upwelling happened during December to April, and intrusion of deeper water to shallower depths brings this layer to the surface to feeds nutrient pools, therefore enhancing the productivity (Walsh et al., 1999). The Cariaco nutrient profile has two layers of nutrients from the SUW in the top 200m, and one below 400m the anoxic region, as shown on figure 2.6.

From the model output, DIN is visibly overestimated, in the first nutrient rich layer, shown in Figure 2.6. High DIN concentration (>10 mmol m $^{-3}$) is only observed below 100 - 200m, and below 400m, whereas in the model it occurs at all depths below 30m. The two nutrient layers that occurs at 100 to 250m, and below 400m are also not captured in the simulation. In the first four months, the DIN layers remain separated, but they eventually merge in 1999. The simulated first layer is also deeper, with a nitrocline of 30m and 275m, and the first layer starts to get deeper within six months. This discrepancy is also reflected in the weak correlations of DIN between in situ and the ensemble mean and the default run, where r =0.48 and 0.5, respectively, summarised on Table 2.3). This discrepancy between ensemble simulation and in situ DIN is expected because the simulation of the two layers is only possible if we use the 3-D version of MEDUSA where horizontal advection can be simulated. An increase in DIN in the first half of the year is also simulated, as a 'burst', only lasting a few weeks for the first five years, but more defined in later years. However, these bursts happen at very shallow depths (in the top5m) compared to in situ profiles, where they happened in the top 10m.

As explained from section 1.6.1 in situ DIN from station PAP was only collected from mid 2002 to mid 2004, to a maximum depth of 400m. Strong seasonality of nitrate is captured well. High nitrate at the surface occurs in the winter (December - April) and declines in the summer. However, below 400m, a mostly continuous high ($>$ 10 mmol m^{-3}) nitrate layer is found. This seasonal pattern also coincide

with the mixed layer; when the mixed layer gets deeper, nutrient concentrations get higher in the surface and decline as it shallows. This may also coincide with light availability during the summer. These are shown on figure 2.7. Compared to in situ DIN, the model overestimates the concentration and does not show similar seasonality. This is shown by a very weak correlation shown in table 2.3, however it shows the lowest RMSE compared to other stations.

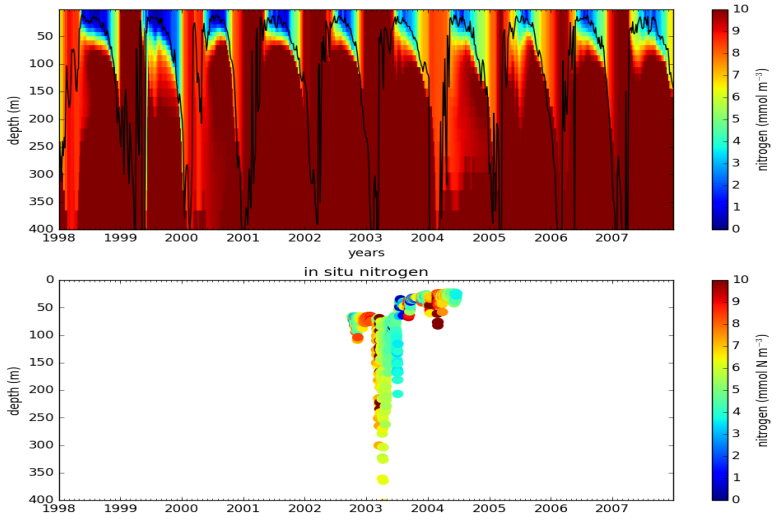


Figure 2.7: (top) ensemble mean DIN at the first 400m depth. Black line shows mixed layer depth (bottom) *in situ* DIN at PAP

The overestimation of DIN concentration also happens at the coastal station L4, after we decrease the sinking speed to 0.1 m day^{-1}. The overestimation occurs during summer when in situ DIN is close to 0, but the ensemble still simulate high DIN concentration (>4 mmol m^{-3}) shown on figure 2.8. The pattern of DIN is dictated by the mixed layer depth, similar to station PAP. However, despite the overestimation, the correlations between in situ and the ensemble mean and median at this station is still relatively high.

From these results, it is evident that vertical velocity from NEMO-FOAM has changed the distribution of nutrient in the water column, especially at stations ALOHA, BATS, and Cariaco. It also shows an overestimation of DIN, from the surface to 200m, except for station BATS where large downwelling current push the DIN to even deeper depth.

2.6 Model spin up

In this experiment, spin up runs are not used to initialise the model because when the model is run for a very long time, a steady increase of DIN concentration occurs over the time series. In this section, we run the 1-D MEDUSA at station ALOHA using two different physical inputs and spinup periods: (i) using the first year's physical inputs, described in 2.5.1, which have been capped as described in the

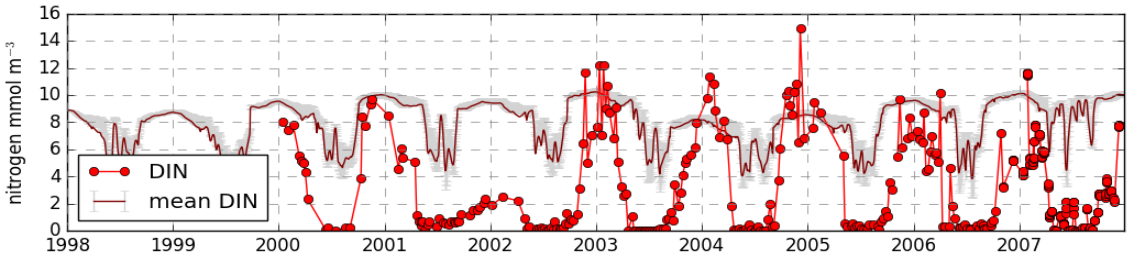


Figure 2.8: surface nitrate at station L4, the solid and dotted red lines are the ensemble mean and the in situ DIN concentration. The shaded grey is the 75th and 25th quartiles of the ensemble spread.

Table 2.3: Ensemble mean, median, and default run error statistic for DIN profiles at different stations for the top 1000m, from 1st January 1998 to 31st December 2007. For L4 the statistical metrics are taken from the surface. Bias is calculated by subtracting the model output with in situ data, therefore negative values denotes overestimation.

Stations		median	mean	default
ALOHA	Correlation	0.926	0.926	0.922
	p	0	0	0
	RMSE (mg m^{-3})	8.051	8.061	8.155
	Bias (mg m^{-3})	-4.860	-4.878	-4.841
	n		2931	
BATS	Correlation	0.734	0.732	0.748
	p	0	0	0
	RMSE (mg m^{-3})	5.060	5.074	4.969
	Bias (mg m^{-3})	0.244	0.199	0.641
	n		3279	
Cariaco	Correlation	0.479	0.482	0.504
	p	0	0	0
	RMSE (mg m^{-3})	6.984	6.980	6.304
	Bias (mg m^{-3})	-5.269	-5.284	-4.627
	n		1963	
PAP	Correlation	0.244	0.246	0.227
	p	0	0	0
	RMSE (mg m^{-3})	5.450	5.308	5.253
	Bias (mg m^{-3})	-3.802	-3.538	-2.722
	n		814	
L4	Correlation	0.616	0.625	0.627
	p	0	0	0
	RMSE (mg m^{-3})	4.604	4.526	4.182
	Bias (mg m^{-3})	-3.699	-3.627	-3.134
	n		267	

previous subsection (2.5.3), repeated over 50 years and (ii) using the climatology (mean from 10 years of the physical inputs), repeated over 250 years. The initial condition for this run is taken from the in situ data, similar to that in the full model.

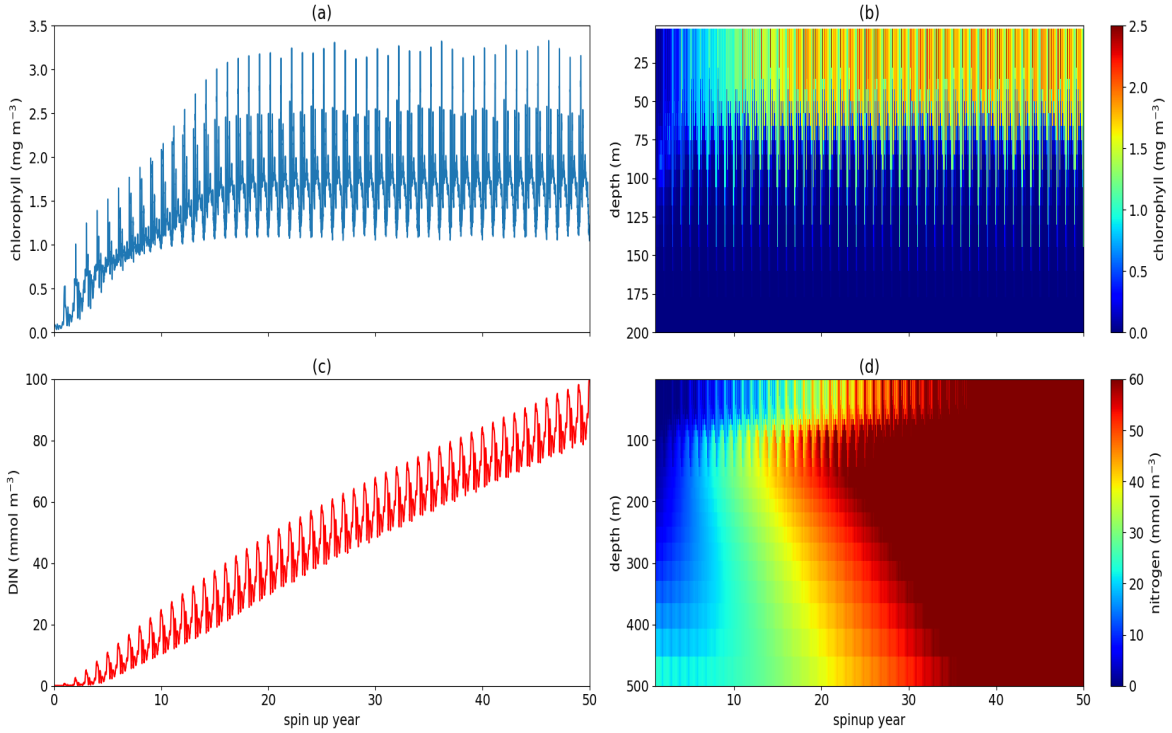


Figure 2.9: Spinning up 1-D MEDUSA using first year condition at ALOHA for 50 years. Chlorophyll surface and profile are shown on (a) and (b) respectively. DIN surface and profile are shown on (c) and (d) respectively.

When repeating the first year's physical input for 50 years, the chlorophyll concentrations reach a repeating cycle after ∼16 years, except for DIN. During the spin up period, surface chlorophyll concentration increases from 0.13 to 1.64 mg m^{-3} in the span of ∼15 years, then reaches a constant cycle for the rest of the spinup period, summarised in Fig. 2.9(a). Additionally, the deep chlorophyll maxima vanish as the concentration increases (Fig. 2.9(b)). For surface DIN, on the other hand, the concentration increases throughout the spinup period, as shown in Fig. 2.9(c). This could be due to the mean upward vertical velocity, bringing the DIN from the depth to the surface, and therefore increasing the concentration.

Another experiment has been carried out using the climatology of the physical inputs of the 10-year period (January 1998 - December 2007) and has been run for 250 years. In this spin up run, the surface chlorophyll decreases from 0.06 to 0.01 mg m^{-3} after four years, and starts to increase over ∼ 70 years to 0.2 mg m^{-3}. Then it starts to decline again for the rest of the time series, summarised in Fig. 2.10(a). When chlorophyll starts to decline, DIN increases drastically over the time series, (Fig. 2.10(c) and (d)). This might be caused by the decline of phytoplankton. The deep chlorophyll maxima also vanishes after ∼ 54 years.

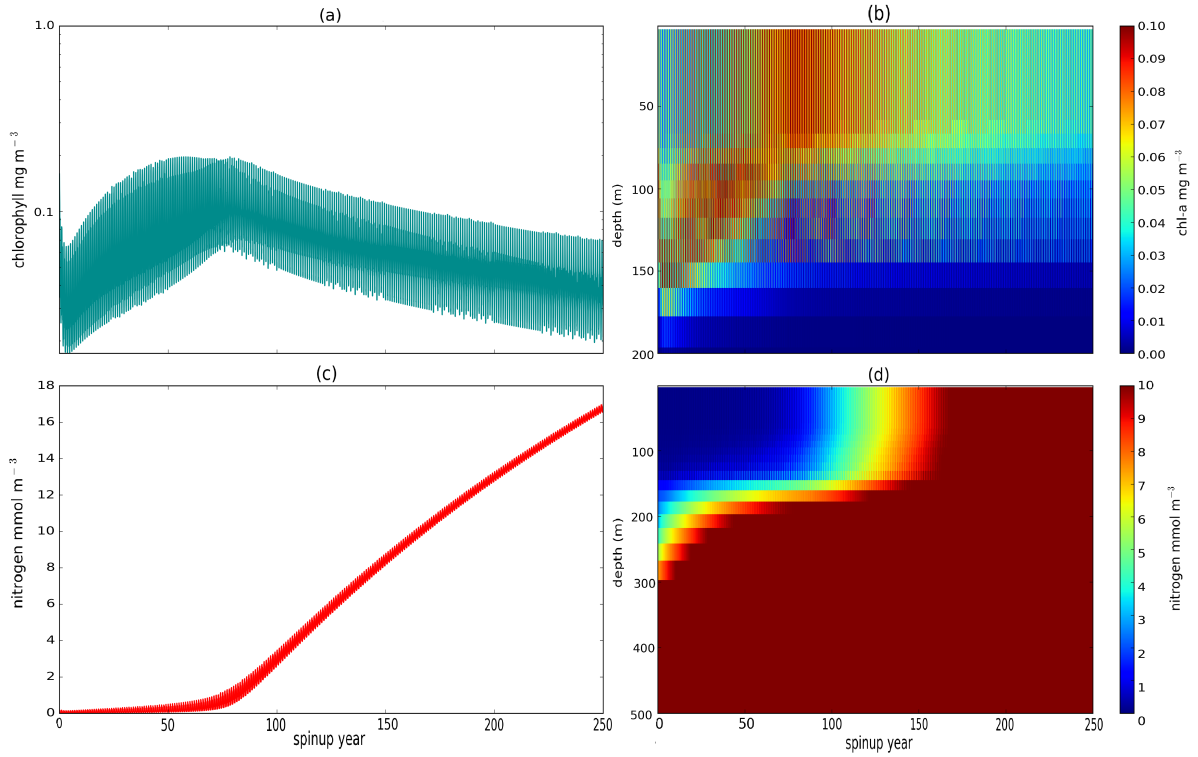


Figure 2.10: Spinning up 1-D MEDUSA using climatology of the physical input for 250 years at ALOHA. Chlorophyll surface and profile are shown on (a) and (b) respectively. DIN surface and profile are shown on (c) and (d) respectively.

From these two spinup runs, we decided not to initialise the model using the spinup run, as both runs show that DIN concentrations keep increasing. This is mainly due to the sum of the first 200m vertical velocity from the climatology at ALOHA is going upwards, making the DIN increase over the years, especially in the first spinup run. This also suggests that the increasing trend may still present in all runs.

2.7 Model metrics

2.7.1 Statistical metrics

Exploring the effect of structural uncertainty can be done by quantifying its bulk properties, such as the total concentration of chlorophyll in the surface, at depth, and DIN concentrations at depth, both seasonal and annual average (Kwiatkowski et al., 2014), and the regional differences of the effect of structural sensitivity. We use statistical metrics including correlation coefficient, root-mean squared error (RMSE), bias, ensemble range, and 10-year mean depth profiles of DIN and chlorophyll and integrated chlorophyll in order to compare the ensemble model with the default model and how well they represent the observations. For surface chlorophyll, apart from the metrics above we use the mean chlorophyll abundance each year in order to see inter-annual variability, and monthly abundance for the seasonal

variations. A similar approach is applied to DIN, apart from station L4 where chlorophyll and DIN are only observed at the surface, we use the averaged DIN over 200 m (integrated DIN / depth) to calculate the inter-annual mean and monthly abundance, because at some stations, especially the oligotrophic, it is often that the top 50m DIN concentration is 0 mmol m^{-3}. These statistical metrics are compared with in situ data.

2.7.2 Other metrics

In an ensemble forecast system, an ensemble with good reliability is one that is statistically consistent with the observations, such that the observations are statistically indistinguishable from the ensemble members. This is something that cannot be done with default run alone. In order to assess the value of the ensemble probability distribution we must assess the consistency of the ensemble spread as well as the ensemble mean error (Moradkhani and Meskele, 2010). A simple method is discussed by Anderson (2001) which takes the ratio R_a of RMSE of the ensemble mean and the mean RMSE of all the ensemble members, which has the expectation value $E[R_a] = \sqrt{\frac{(n+1)}{2n}}$, where n is the number of ensemble members. This is called the Normalised RMSE Ratio (NRR= $R_a/E[R_a]$) where the desirable ensemble spread should have NRR=1. If the NRR >1 then the spread is too small, and NRR <1 indicates that the ensemble spread is too large. We may expect different NRR values for different metrics and also for variability on different timescales, such as monthly or inter-annual data. This method has previously been used to set the number of ensemble members in data assimilation (Moradkhani et al., 2006; Roy et al., 2012).

As explained in section 1.7, one of the research objectives is to determine whether structural sensitivity affects the model prediction. In order to address this question, it is possible to use phytoplankton bloom phenology metrics. These metrics are useful ecological indicators for detecting natural and anthropogenic impacts on the pelagic ecosystem (Platt and Sathyendranath, 2008). In this study, seven phenology indicators are considered; bloom initiation time, where the chlorophyll concentration exceeds a certain threshold, which is 1.05% of the chlorophyll median of a certain year. We also derived a 'termination time', where bloom concentration falls below the same threshold. The number of days when chlorophyll concentration is higher than the threshold is the 'bloom duration'. The concentration at the bloom peak and the date it takes place, are also included as indicators. We also note the amplitude of the bloom, which is half the peak height minus the minimum chlorophyll concentration. The method of deriving the phytoplankton bloom phenology is explained in the next section.

2.8 Phytoplankton Phenology as model metrics

When the environmental conditions are right, a dramatic increase in phytoplankton may occur over large areas of the ocean, called a phytoplankton bloom. This phenomenon has been observed to occur annually all over the world's oceans (Heinrich, 1962; Cole, 2013). In the North Atlantic, this is seen in the springtime and sometimes in the autumn (Raymont, 1980). Two of the most well known hypotheses which explains the occurrence of the North Atlantic bloom are the critical depth hypothesis Sverdrup (1953) and the disturbance recovery hypothesis Behrenfeld et al. (2013).

According to the critical depth hypothesis, the initiation of the bloom occurred when the critical mixing depth is reached, which usually occurred in spring in the North Atlantic. According to this hypothesis, during winter, the loss rate of phytoplankton exceeds the division rates. As spring proceeds, mixed-layer integrated phytoplankton division rates increase with decreasing MLD and increasing sunlight and temperature. This eventually made the division rates equal to the loss rates (which are assumed to be constant over time). The depth at which loss rates is similar to the division rates is called the critical mixing depth (Sverdrup, 1953).

Disturbance recovery hypothesis (Behrenfeld et al., 2013) stated that blooms are initiated by physical processes (disturbance) that disrupt the balance between grazing and phytoplankton division, such as deep winter mixing, upwelling, or polar night. According to this hypothesis, the changes of physical processes in autumn have a more significant impact on grazers, so that by early winter, the division rate of phytoplankton can exceed loss rate by grazing. This makes the depth-integrated phytoplankton biomass increase. Higher phytoplankton division rate is sustained during spring because increasing grazing pressure is in line with increasing phytoplankton division rate (Behrenfeld and Boss, 2014).

The phytoplankton bloom occurs annually, but the size, time which it occurs, and the length of a particular bloom varies yearly. The study of these biological events is called phenology.

2.8.1 Why phenology matters

Interest has grown into phytoplankton phenology because it may affect the survival of fish larvae through a 'mismatch' hypothesis (Cushing, 1990), where the difference between the timing of the spring bloom of phytoplankton and fish spawning time determine the variations in the abundance of adults. In the North Atlantic a link between anomaly in phytoplankton bloom and haddock larval survival has been observed (Platt et al., 2003). In the event of global climate change, bloom timing may be affected (Kahru et al., 2010), and therefore a mismatch between life cycles at higher trophic levels and phytoplankton bloom (Edwards et al., 2004). Since variation in seasonal temperature in the ocean is smaller than the terrestrial,

the change in phenological events that are driven by climate change are expected to occur earlier in the ocean, despite warming in the ocean being slower than terrestrial trends (Burrows et al., 2011). Earlier and longer stratification is expected to occur (Sarmiento et al., 2004) due to the rise in sea surface temperatures; in response to this, phytoplankton in high-latitude regions may initiate earlier blooms and potentially boost primary production (Bopp et al., 2001; Steinacher et al., 2009). In contrast reduced plankton production may be observed in lower and mid-latitudes as increases stratification prolongs the period of nutrient limitation because entrainment of nutrients from below the thermocline is inhibited (Behrenfeld et al., 2006).

Satellite-derived ocean colour data provide a valuable tool to examine phytoplankton blooms and their initiation at the basin scale (Platt et al., 2009). A bloom is characterised by a strong chlorophyll signal in the seasonal variation, and is the most important event in the dynamics of phytoplankton (Platt and Sathyendranath, 2008). As shown in figure 2.11, a bloom can be identified using its peak (maximum), width (initiation, termination, and length of the bloom), and timing of peak. These are the indicators that are associated with phytoplankton phenology. In determining this metrics, there are several methods that fall into three broad categories (Ji et al., 2010; Brody et al., 2013): (i) threshold method (Siegel et al., 2002; Racault et al., 2012; Cole et al., 2012) or when a fitted model to chlorophyll-a data crosses a set of thresholds (Platt et al., 2009; Zhai et al., 2011; Sapiano et al., 2012), (ii) cumulative sum (Greve et al., 2005; Mackas et al., 2012), and (iii) rate of change methods (Sharples et al., 2006; White et al., 2009). These methods will be explained further in the next subsection.

2.8.2 Methods of determining phytoplankton phenology

In the threshold method, initiation occurs when chlorophyll-a concentration exceeds a certain threshold. This is the most widely used to identify bloom initiation and termination (Brody et al., 2013). The threshold itself could be set in absolute terms (chlorophyll concentration) or in relative terms (some percentage of the maximum amplitude). The cumulative sum method identifies a bloom when the cumulative sum of chlorophyll crosses a certain threshold percentile of the total biomass. For the rate of change methods, initiation of the bloom is detected when the increase of chlorophyll is most rapid in a time series or a fitted function. All bloom detection methods have similar characteristics, initiation, peak timing, duration (the difference between start and end of the bloom), and the maximum concentration (bloom height), shown on Figure 2.11.

When determining bloom initiation, some studies use an annual median based threshold (Siegel et al., 2002; Racault et al., 2012; Cole et al., 2012), although in the critical depth hypothesis, the initiation is not determined using this method (Behrenfeld and Boss, 2014). The range is between 1-30% above the

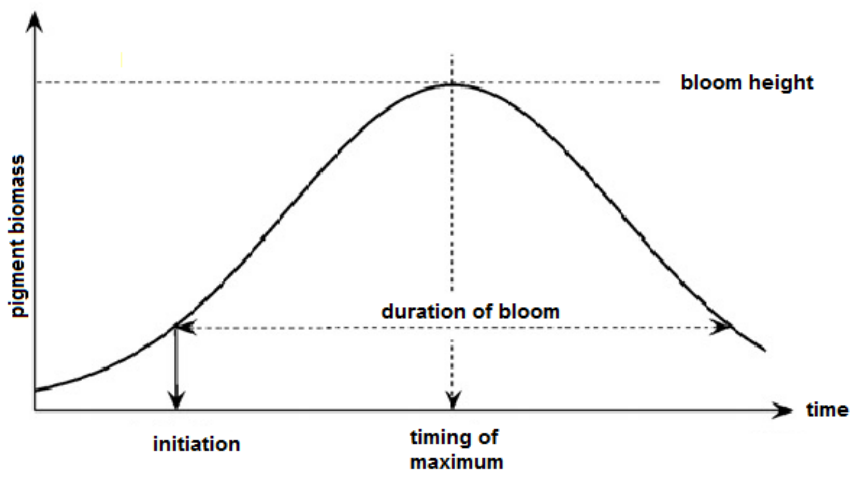


Figure 2.11: Properties of a phytoplankton bloom on a time series of chlorophyll concentration. In this figure, initiation is determined using the threshold method. Figure is obtained from (Platt and Sathyendranath, 2008)

annual median, but a study by (Siegel et al., 2002) found that bloom initiation dates were not sensitive to the choice of percentage and chose 5%. In order to only identify the primary bloom, time series of chlorophyll are often shifted, by starting in the late summer for regions where high primary production occurs in fall (Siegel et al., 2002; Henson et al., 2006; Platt et al., 2009; Thomalla et al., 2011; Cole, 2013). Another option is to find the bloom peak and work backwards to find when chlorophyll concentration goes below the thresholds (Siegel et al., 2002; Racault et al., 2012; Cole et al., 2012). To eliminate noisy data, some studies fit a function or model to the time series before identifying the initiation (Platt et al., 2009; Vargas et al., 2009; Zhai et al., 2011; Sapiano et al., 2012). Others temporally and spatially averaged chlorophyll data and only consider a bloom when the concentration above threshold occurs for more than two weeks long (Brody et al., 2013). Another method is to use the percentage of maximum bloom peak as the threshold. Compared to the median threshold, this method is influenced by the size of bloom peak (Cole et al., 2012). Anomalous spikes of overestimated chlorophyll values can bring the threshold value to more than the annual median.

Similar to threshold methods, the cumulative sum method uses some threshold value to estimate the start and end of the bloom. The method is initially used in zooplankton phenology (Greve et al., 2005). In contrast to the threshold value where the time series is centred around the bloom peak, the cumulative sum method has to set the start date during low chlorophyll so that thresholds are associated with a percentage of the data (i.e initiation when concentration is 15% and termination when the sum is 85% (Greve et al., 2005)). In the rate of change method, there are no threshold values, since it is based on the greatest positive and negative change in chlorophyll concentration for initiation and termination respectively. When this method is used in satellite-derived chlorophyll data where gaps are present, a

model or function may be fitted to the time series to smooth the data.

Aspects of phytoplankton phenology could be used as ecological indicators. These indicators are used as a metric to show how structural sensitivity affects the model prediction. The indicators and their descriptions are summarised in table 2.4. These are derived using the method described above, and applied to all ensemble outputs for each year. In this study, annual median based threshold and curve fitting methods have been used to derive the phenology of phytoplankton, to avoid misidentification for secondary blooms as primary bloom (Brody et al., 2013), which will be explained below.

Table 2.4: Ecological indicators that serve as the ensemble metrics

Indicators	Description
Annual mean	The average concentration of chlorophyll at a particular year
Initiation	The day when chlorophyll concentration exceeds a certain threshold
Bloom timing	The month at which the chlorophyll is at its maximum
Peak height	Concentration of highest bloom
Amplitude	Half of the highest peak
Duration	The length of days when chlorophyll concentration is higher than 50% of the peak height
Termination	The end the first prominent bloom of the year

Threshold method applied to data

The initiation is determined by the day when chlorophyll concentration exceeds a given threshold. Following the work of (Siegel et al., 2002). Using this method alone, determination of phenology is shown on figure 2.12, where we use the MarMOT-MEDUSA test run at 60°N, 20°W, which is run for three years from January 1998 to December 2000. First, the model outputs are separated into three individual years. After finding the maximum peak, initiation and termination values were calculated within six months of the peak, and if these values are outside the range, then it is set as not a number (NaN) and excluded from further analysis.

As seen on figure 2.12, in the second and third year, there is a second peak that occurs near the spring maximum. By using the method explained above, the actual peak height is uncertain. Additionally, the phytoplankton bloom not only occurs during spring time, but also during summer, thus the secondary bloom is undetected. The length of spring bloom becomes as long as 200 days. To resolve this problem the output is then fitted to a 4th order polynomial function if a second bloom is detected, and to a 3rd order if only one prominent bloom occurs during the year.

To fit the curve, spikes in the data that are above the threshold are identified as the points to be fitted. The fitting of a polynomial curve to the data is done by using `polyfit` and `polyval` functions in

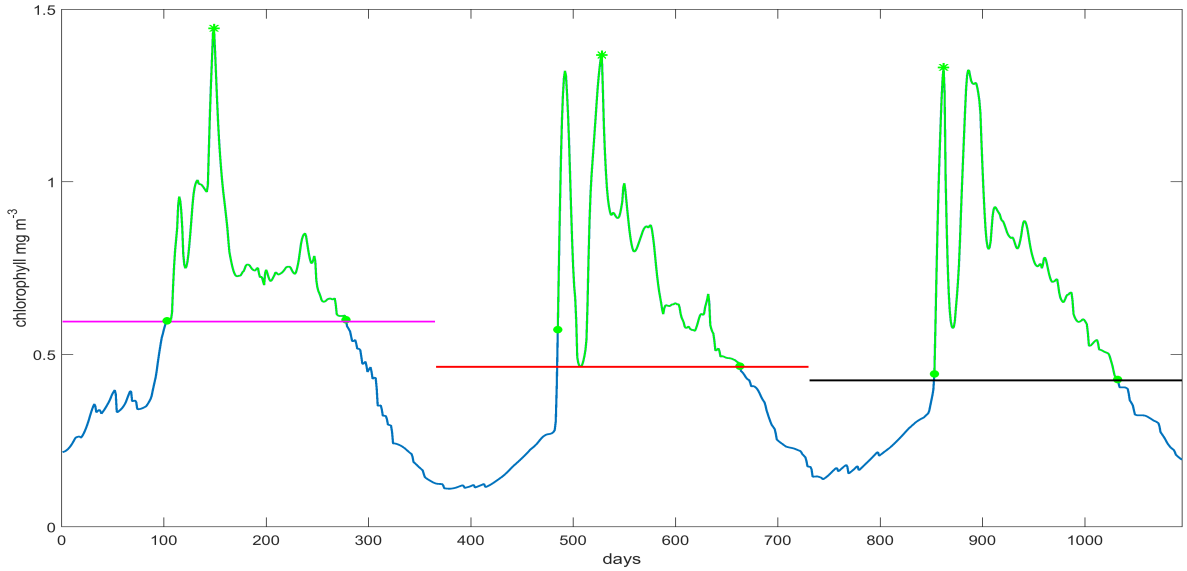


Figure 2.12: Time series of chlorophyll concentration at 60°N 20°W. The black, red, and pink lines denote the threshold value for each year. The green line shows when chlorophyll is above the threshold concentration, and the green star is the maximum chlorophyll. The closed green circles represent initiation and termination time

MATLAB. By using the acquired value (`polyval`), peak height is identified by finding the maximum data point from the fitted curve. Amplitude is then calculated as half of the highest peak (Kostadinov et al., 2017). Then the spring bloom is defined as when the fitted values are 50% of this peak height and therefore the length of this set of data is the duration of the spring bloom. Initiation of spring bloom is similar to the annual threshold method. However, termination is determined when the `polyval` value is below 50% of the peak height, if only one prominent peak is detected. When two peaks are present, the termination of spring bloom is determined when the first bloom reduces to its minimum, just before the second bloom starts (at the first valley). This method is useful to see how the bloom develops and terminates, whether the concentration increases rapidly and decreases slowly or vice versa. The phenology is summarised in figure 2.13. The curve fitting method is only applied if the data show potential outliers especially in higher concentrations. If there is only one prominent bloom each year, such as at stations ALOHA and BATS, and the data are smooth, the regular threshold method (when the concentration is above 50% of the maximum bloom, and the associated initiation and termination times) without fitting the data with a curve is applied. To avoid results being affected by how bloom phenology is determined, the same method is used for determining the metrics from both in situ and model output.

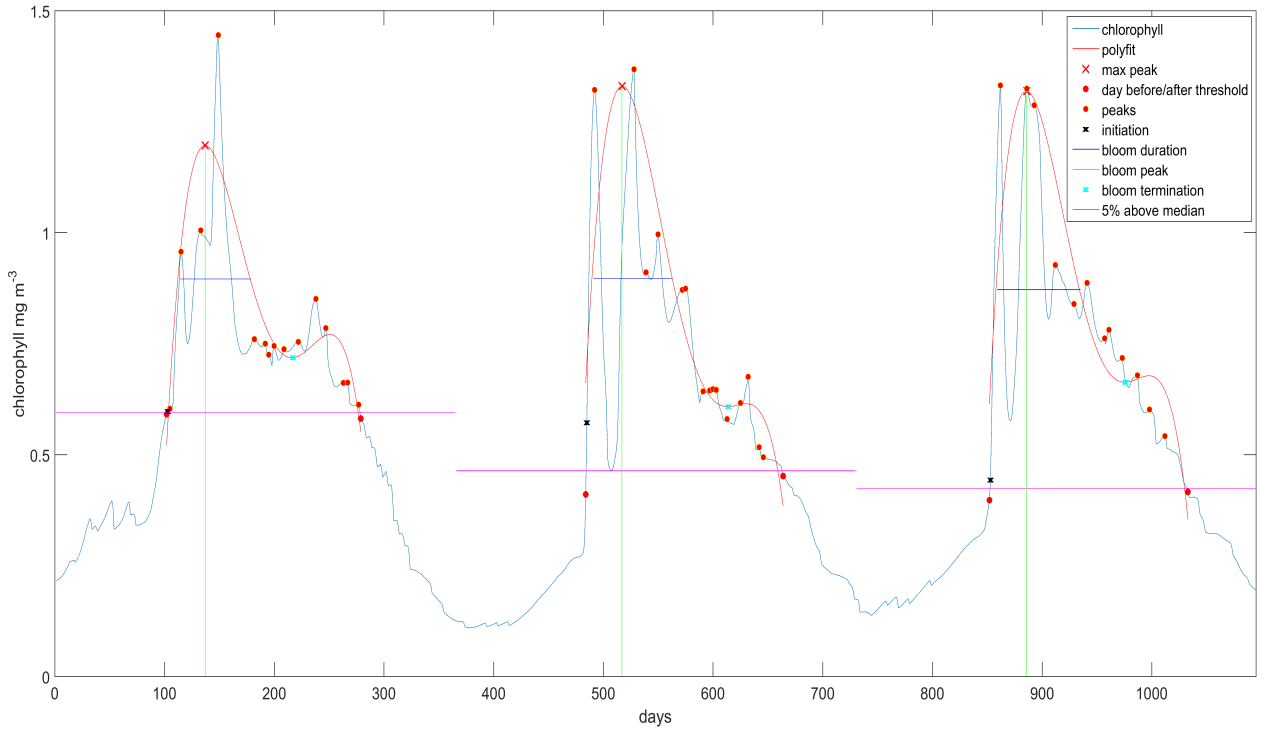


Figure 2.13: Determining phenology using a combination of threshold method and curve fit in the North Atlantic (60°)

2.9 Pilot Study: Altering default functional forms in a simple NPZ model

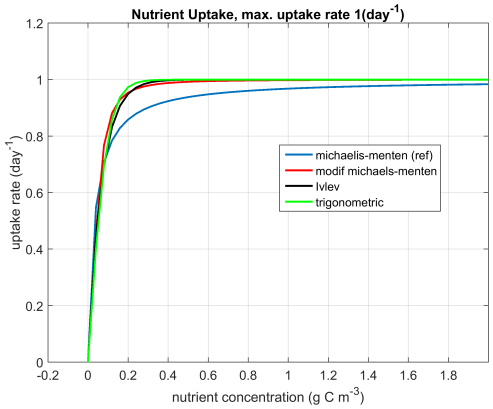
The effect of structural sensitivity is prominent in a simple predator-prey model. Since all complex marine biogeochemical models have common features; the nutrient, phytoplankton, and zooplankton compartments, before moving on to a more complex model, the effect of structural sensitivity in a simple NPZ model may give insights to how it may affect a more complex marine biogeochemical model. Here, structural sensitivity is tested by altering the default functional forms with equations that have similar properties as the default form. In this experiment, simple 0-D NPZ model by Edwards and Brindley (1996), which is based on Steele and Henderson (1981), is used. The model can be represented by coupled ordinary differential equations:

$$
\begin{aligned}
\frac{dN}{dT} &= \text{-uptake} + \text{respiration} + \text{Z excretion} + \text{Z predation excretion} + \text{mixing} \\
&= -\frac{N}{e+N}\frac{a}{b+cP}P + rP + \frac{\beta\lambda P^2}{\mu^2+P^2} + \gamma dZ^2 + k(N_o - N) \\
\frac{dP}{dT} &= \text{uptake} - \text{respiration} - \text{grazing} - \text{sinking} - \text{mixing} \\
&= -\frac{N}{e+N}\frac{a}{b+cP}P - (rP) - \frac{\lambda P^2}{\mu^2+P^2} - (s+k)P \\
\frac{dZ}{dT} &= \text{growth} - \text{higher predation} \\
&= \frac{\alpha\lambda P^2}{\mu^2+P^2} - dZ^2
\end{aligned}
\tag{2.10}
$$

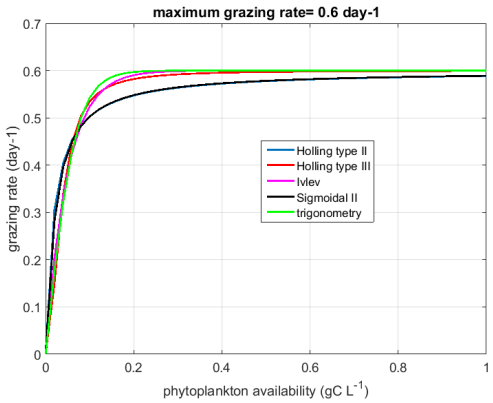
Similar to the default MEDUSA, this model uses rectangular hyperbolic nutrient uptake, $\frac{N}{e+N}$, where e is the half saturation constant. The term $a/(b+cP)$ is the light limited growth, where b and c are the limitations due to light attenuation by water and self-shading of the phytoplankton, respectively. In the original equation, the depth averaged daily phytoplankton growth rate is given by 2.58 $P_{max}/(b+cP)M$, where P_{max} is the maximum phytoplankton growth rate and M is the mixed layer depth. This model uses a constant mixed layer depth of 12.5m. Therefore in this equation, a/b gives the maximum growth rate averaged over the depth of the mixed layer (Edwards and Brindley, 1996). The respiration term is represented by a linear term. Zooplankton grazing is described using Holling type III, where λ is the maximum grazing rate and μ is the half saturation constant, similar to that in MEDUSA. However, not all of the prey is assimilated to the zooplankton, and this efficiency is described using the term α. Parts of the prey that is not assimilated to the zooplankton is described with the parameter β, and this will be regenerated as nutrients, making $\alpha+\beta=1$.

The higher predation of zooplankton is described in the quadratic form dZ^2, where d is the maximum predation rate. The zooplankton loss is regenerated as nutrient, represented by γ, and which is described within the nutrient equation. The exchange of nutrients with the nutrient rich water below the mixed layer is modelled as $k(N_o - N)$, where N_o is the sub mixed-layer nutrient concentration and k is the fraction of the mixed layer which is exchanged daily with the deeper water due to diffusive processes. The last term of the phytoplankton equation is the sinking and mixing. The term s, describes the sinking rate of phytoplankton out of the mixed layer.

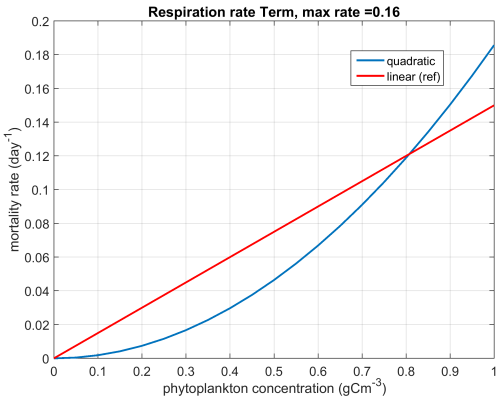
From these equations, it is possible to alter the default functional forms using the equations described in Table 1.1. For zooplankton grazing, we used Holling type II, Ivlev, trigonometric function ($\frac{2}{\pi}\arctan(\frac{\mu}{k})$), and sigmoidal II function, which is similar to the Holling type III ($\frac{P^2}{(\mu+P)(\mu+P)}$) to describe the zooplankton grazing. Apart from Michaelis-Menten function, nutrient uptake, can also be described using,



(a) Nutrient uptake functions



(b) Grazing functions



(c) Respiration functions

Figure 2.14: Uptake, grazing, and respiration functions that have been made equivalent using non-linear least square.

trigonometric, exponential (ivlev), and sigmoidal functions. These functions are all similar in shape and have features such as the shape defining parameter and maximum rates. Following the early work of Steele and Henderson (1992), the quadratic and linear functions are used for respiration. The parameter values and initial conditions are adopted from the original Edwards and Brindley (1996) paper. Therefore, there will be four uptake functions, four grazing, and two respiration (phytoplankton mortality). Similar to the study by Fussmann and Blasius (2005), non-linear least squares is used in order to make the functional forms more equivalent, as seen on Figure 2.14.

In this experiment, only one functional form will be altered at a time, instead of having multiple functional form combinations. The simple model is run for 730 days, with initial conditions of 0.4, 0.1, and 0.05 gCm^{-3} for nutrient, phytoplankton, and zooplankton respectively. Then the overall mean, period, number of peaks, and amplitude for all of the model compartments are used as metrics to quantify the effect of altering functional forms.

2.9.1 Results: alternative model simulations

The model metrics and the results for each of the biogeochemical processes are shown in Table 2.5. In terms of uptake functional forms, the mean concentration and amplitude of nutrients from the alternative uptake functions show significantly lower nutrient concentrations compared to the default run. However, the default run produced fewer peaks throughout the time series, compared to the alternative functions. In terms of phytoplankton and zooplankton, the alternative functions do not produce concentrations that are significantly higher or lower than the default run. Similar to the nutrients, in both phytoplankton and zooplankton the number of peaks in the default run is lower than other alternative functions. Phytoplankton from the default uptake function also shows the lowest concentrations but also has the longest period and highest amplitude. Both nutrient and zooplankton from the default functions produce the longest period and highest amplitude.

Altering the grazing function produces more variability for all the model metrics, and therefore makes the ensemble range larger than other functions. Using the Holling type II function produces the highest mean nutrient concentration, period, and amplitude for all of the model compartments however, the number of peaks produced from perturbing the grazing function is the smallest compared to perturbing other processes (including nutrient uptake and respiration). The lowest mean zooplankton concentration compared to other functional forms (including nutrient uptake and grazing) is produced from the Holling type II grazing function. From Figure 2.15a, the shape of the nutrient peaks from Holling type II is the most distinct compared to other functions, where the concentration decreases rapidly after reaching the peak. The opposite is observed in phytoplankton and zooplankton whereby their concentrations increase more rapidly compared to how they decrease, shown in Figures 2.15b and 2.15c.

The phytoplankton concentration during low concentration periods can go down to almost zero, when Holling type II is used. Although the sigmoidal II equation is more similar to the default Holling type III function, the amplitudes for nutrient and phytoplankton are almost three times higher than the default function, and the period for all the model compartments are almost twice as long than for the default function. This functional form also produced the highest mean phytoplankton concentration. Similar to Holling type II, the phytoplankton concentration can reach close to zero during non-bloom periods when applying sigmoidal II function. Using a trigonometric grazing function, dampens oscillations at all model compartments, similar to when a hyperbolic tangent is used in the simple predator and prey model (Figure 1.6).

In terms of phytoplankton respiration, from Figure 2.14c, the two functional forms are distinctive. Using the quadratic term resulted in the lowest nutrient concentrations, and an increase in phytoplankton and zooplankton concentrations compared to the default model. The period, and number of peaks, for all the

model compartments are reduced, and increased respectively, compared to the default model. All of the amplitudes are also higher than the default function, especially in zooplankton and phytoplankton.

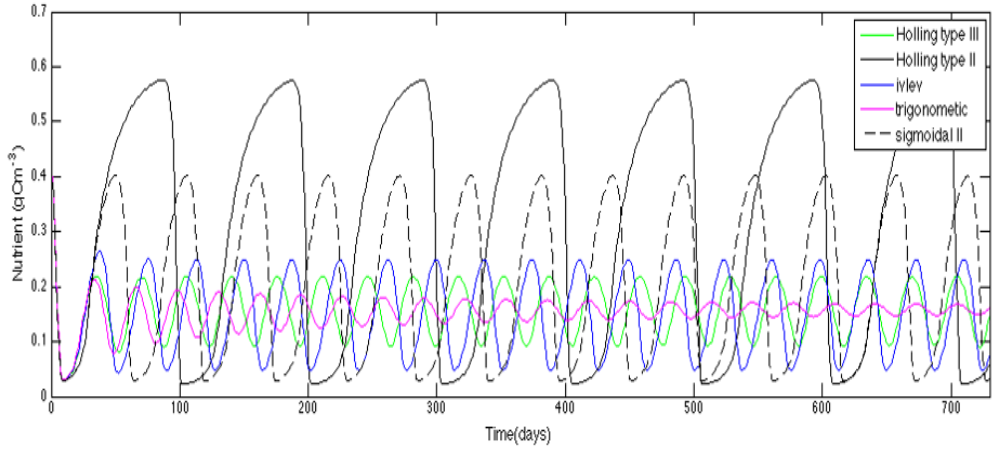
Table 2.5: Mean, period, number of peaks, and amplitude for nutrient, phytoplankton, and zooplankton using simple NPZ model. The units for all of the model compartments are gCm^{-3}. In the sigmoidal forms, as seen on Figure 2.15 the oscillation dampens, therefore there are no amplitude

	Uptake				Grazing				Respiration	Range	Mean
	$\frac{N}{e+N}*$	$\frac{N^2}{e+N^2}$	$1-\exp(eN)$	$\frac{2}{\pi}\arctan(\frac{N}{e})$	$\frac{\lambda P}{\mu+P}$	$\lambda(1-\exp(-\mu P))$	$\lambda\tanh(\mu P)$	$\frac{\lambda P^2}{(\mu+P)(\mu+P)}$	rP^2		
Metrics						Nutrient					
Mean (gCm^{-3})	0.158	0.122	0.139	0.131	0.324	0.150	0.154	0.206	0.084	0.239	0.163
Period (days)	34	30	30	26	100	38	32	55	28	74	41.444
No of peaks	20	27	25	27	7	19	22	13	26	20	20.667
Amplitude	0.062	0.044	0.038	0.025	0.277	0.101		0.187	0.084	0.252	0.102
						Phytoplankton					
Mean (gCm^{-3})	0.064	0.069	0.066	0.065	0.084	0.078	0.065	0.089	0.074	0.025	0.072
Period (days)	35	29	29	27	101	39	33	55	27	74	41.667
No of peaks	21	25	26	28	8	20	23	14	26	20	21.222
Amplitude	0.035	0.030	0.027	0.017	0.220	0.070		0.144	0.073	0.203	0.077
						Zooplankton					
Mean (gCm^{-3})	0.075	0.074	0.073	0.074	0.037	0.069	0.072	0.058	0.077	0.040	0.068
Period (days)	35	29	29	26	101	39	32	55	28	75	41.556
No of peaks	21	25	28	27	7	20	23	13	26	21	21.111
Amplitude	0.011	0.007	0.007	0.004	0.036	0.016		0.025	0.077	0.073	0.023

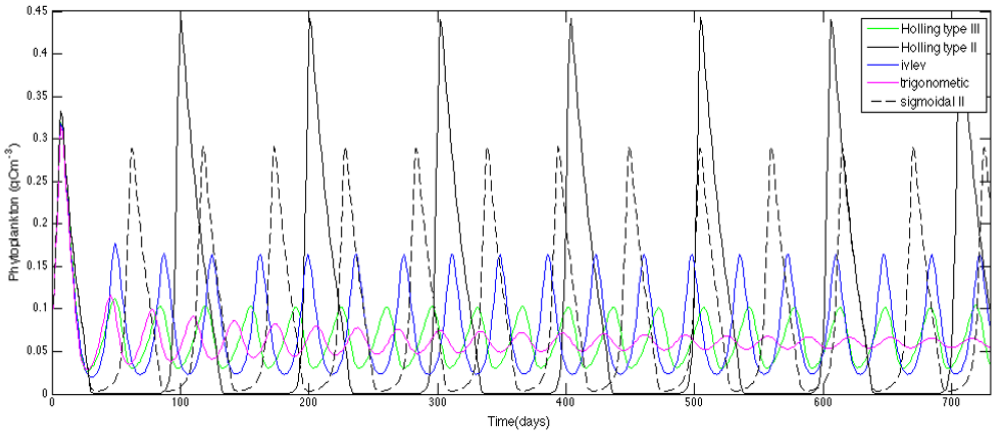
2.9.2 Lessons from a simple model

From Figure 2.14, functions that produced the most deviations compared to other alternative functions are the default uptake function (Michaelis-Menten), Holling type II, and Sigmoidal II for grazing. The default uptake function has a higher uptake rate compared to the other uptake functions when the nutrient concentration is low, however when the nutrients raise above 0.1 gCm^{-3}, the uptake rate is lower than other functions, shown in Figure 2.14a, making the mean nutrient concentration slightly higher compared with other functional forms. This may explain why the phytoplankton concentrations is slightly lower than other alternative uptake functions. Similarly, when using Holling type II or Sigmoidal II functions, lower grazing rate is produced when the phytoplankton concentration is above 0.05 gCm^{-3}, especially for Holling type II, making the overall phytoplankton mean concentration higher than other grazing functions. This can also be applied in the phytoplankton respiration, where the quadratic forms generally produces more phytoplankton, therefore less nutrients, since the linear respiration is always higher than the quadratic rate.

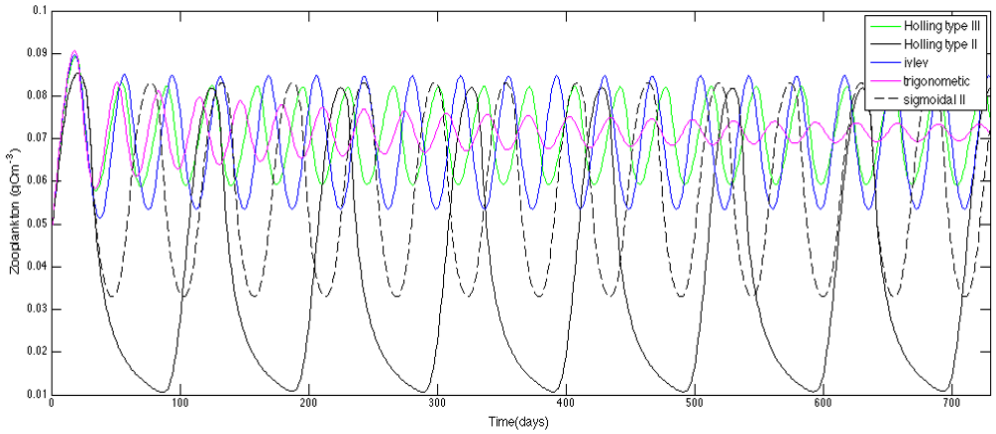
In terms of model dynamics, altering the grazing equation produces the most distinct dynamics for all the model compartments, compared to perturbing other processes. This is because the grazing equation appeared at all the state variable equations (see equation 2.10). Using the type II grazing also causes the phytoplankton concentration to crash (reduce down close to zero), but increases rapidly as the nutrients decreases, which has not been observed in other functional forms. The trigonometric function on the other hand produces nutrient, phytoplankton, and zooplankton concentrations with amplitudes getting smaller and smaller. These patterns have also been observed in the predator and prey model (Figure



(a) Nutrient



(b) Phytoplankton



(c) Zooplankton

Figure 2.15: Time series of nutrient, phytoplankton, and zooplankton concentration from a simple NPZ model when the grazing functions are altered.

1.6(a) and (c)) showing that the effect of structural sensitivity in predator-prey does propagate to the NPZ model. The high amplitudes in Holling type II may be due to the unstable equilibrium and the trajectories settle into a stable limit cycle with large amplitudes (Fussmann and Blasius, 2005).

From these model runs, we can expect that altering the model equations in a simple model can alter the overall mean concentrations, period, and amplitude, which therefore affect the model predictions, such as phytoplankton peak timing and concentrations. This initial experiment shows how altering one functional form at a time can generate a large range, especially when altering the grazing function, making the NPZ model very sensitive to the choice of grazing functions, especially between sigmoidal and hyperbolic functions. These results also concur with Anderson et al. (2010). Furthermore, changing phytoplankton mortality from linear to quadratic doubles the amplitude of phytoplankton and increase the zooplankton amplitude seven-fold, making the NPZ model also sensitive to the change in phytoplankton as well as the zooplankton mortality (Edwards and Yool, 2000; Edwards, 2001). Therefore if all the functional forms are perturbed more range and variations would be produced from the alternative simulations.

In the next chapters, we will explore the structural sensitivity in an operational intermediately complex marine biogeochemical model.

Chapter 3

The effect of structural sensitivity in a marine biogeochemical model [2]

In the previous chapter, the method of generating the ensemble and examining the effect of structural sensitivity in an intermediately complex biogoechemical model have been described. This chapter explores the effect of structural sensitivity by comparing the observations with the ensemble mean and median using statistical metrics, the spread (NRR), and ensemble range. Inter-annual and monthly variabilities are considered, and both biological concentrations and phytoplankton bloom phenology are assessed. The Abyssal Plain station comparisons are discussed in section 3.1, followed by the two oligotrophic stations 3.2, and the two coastal stations 3.3. The effect of structural sensitivity will be summarised and discussed in section 3.4.

Each section that describes different regions will have separate sections for the chlorophyll and DIN concentrations, which describe their statistical metrics interannual, and seasonal means, and the phytoplankton phenology. At each stations, the ensemble is going to be run fully (128 members) and also one at a time, where only one process function is changed, while keeping other functions in their default state, like that in section 2.9. This is done in order to compare the range from altering one process function with the full ensemble, therefore showing how efficient the large ensemble is in producing ensemble range.

As explained previously, we choose DIN and chlorophyll as in situ data for these variables are available at all of the stations, which will be used to compare the ensemble results.

[2]This chapter is based on parts of the paper Anugerahanti, P., Roy, S. and Haines, K. (2018), 'A perturbed biogeochemistry model ensemble evaluated against in situ and satellite observations, Biogeosciences **15**(21), 66856711.

Table 3.1: Error statistics, 10-year mean, and NRR of chlorophyll (mg m^{-3}) and DIN (mmol m^{-3}) concentration at five stations for the default run, ensemble mean, ensemble median, and the ensemble range (ensemble maximum - ensemble minimum). These are calculated from surface to 200 m depth, starting from January 1998 to December 2007. Bias is (model output) – (in situ observation). Bold text indicate the smallest RMSE. At Station L4 error statistics and mean are taken from the surface and starts from January 1999 for chlorophyll and June 2000 for DIN. For station PAP, error statistics are taken from 2002-2004 since in situ data is only available during that time.

		DIN profile				Chlorophyll profile				Surface chlorophyll				Integrated chlorophyll			
Stations		r	RMSE	Bias	Mean	r	RMSE	Bias	Mean	r	RMSE	Bias	Mean	r	RMSE	Bias	Mean
PAP	Ens mean	0.23	3.26	0.61	6.59	0.42	0.32	0.06	0.48	0.45	0.51	0.22	0.66				
		(±0.07)	(±2.57)	(±5.13)	(±5.24)	(±0.37)	(±0.73)	(±0.68)	(±0.75)	(±0.38)	(±0.73)	(±0.68)	(±0.76)				
	Ens median	0.23	**3.16**	0.54	6.38	0.49	**0.29**	0.003	0.42	0.54	**0.46**	0.15	0.60				
	Default run	0.21	3.32	-0.20	5.64	0.28	0.40	0.18	0.59	0.36	0.57	0.30	0.74				
	in situ				5.83				0.42				0.44				
	NRR		1.29				1.19				1.25						
ALOHA	Ens mean	0.77	**1.06**	0.67	1.20	0.22	**0.10**	-0.06	0.06	0.22	**0.05**	-0.01	0.10	0.69	**2.73**	-0.72	3.80
		(±0.03)	(±0.19)	(±0.39)	(±0.39)	(±0.49)	(±0.04)	(±0.11)	(±0.11)	(±0.47)	(±0.09)	(±0.13)	(±0.14)	(±0.60)	(±5.49)	(±7.09)	(±10)
	Ens median	0.77	**1.06**	0.68	1.18	0.14	0.11	-0.07	0.05	0.13	**0.05**	-0.01	0.07	0.56	3.3	-1.17	3.34
	Default run	0.77	1.09	0.61	1.10	0.28	**0.10**	-0.03	0.09	0.27	0.07	0.03	0.11	0.70	4.71	1.25	5.77
	in situ				0.50				0.12				0.08				4.52
	NRR		1.39				1.29				1.07				1.01		
BATS	Ens mean	0.56	1.39	1.16	1.77	0.19	0.33	-0.12	0.05	0.22	0.33	-0.12	0.05	0.39	52.13	-19.39	6.18
		(±0.38)	(±0.84)	(±1.00)	(±1.01)	(±0.37)	(±0.05)	(±0.16)	(±0.16)	(±0.58)	(±0.15)	(±0.05)	(±0.15)	(±0.54)	(±9.40)	(±21)	(±14)
	Ens median	0.55	1.39	1.16	1.77	0.11	0.33	-0.12	0.05	0.06	0.34	-0.12	0.05	0.27	**23.30**	-17.71	4.51
	Default run	0.58	**0.73**	0.62	1.35	0.23	**0.31**	-0.07	0.10	0.28	**0.31**	-0.07	0.09	0.43	48.58	-10.77	13.14
	in situ				0.98				0.17				0.15				23.90
	NRR		1.38				1.39				1.40				1.40		
Cariaco	Ens mean	0.78	**2.97**	0.61	5.39	0.29	**0.83**	-0.02	0.49	0.13	**1.23**	0.02	0.77	0.41	**17.73**	-1.05	11.47
		(±0.08)	(±0.49)	(±2.54)	(±2.54)	(±0.34)	(±0.42)	(±0.93)	(±0.93)	(±0.22)	(±0.33)	(±1.90)	(±0.57)	(±0.40)	(±7.90)	(±17)	(±17)
	Ens median	0.76	3.24	0.51	5.29	0.20	0.88	-0.18	0.32	0.072	1.29	-0.29	0.46	0.29	19.46	-5.51	7.00
	Default run	0.76	3.29	0.59	5.37	0.22	0.87	-0.09	0.42	0.11	1.27	-0.18	0.57	0.34	18.71	-3.86	8.65
	in situ				4.78				0.51				0.76				12.52
	NRR		1.25				1.19				1.21				1.17		
L4	Ens mean	0.70	2.94	1.56	4.52					0.25	1.05	0.42	1.76				
		(±0.14)	(±2.13)	(±4.06)	(±4.06)					(±0.33)	(±1.67)	(±2.61)	(±2.61)				
	Ens median	0.68	3.10	1.73	4.69					0.21	**1.02**	0.27	1.61				
	Default run	0.52	**2.67**	1.12	4.08					0.31	1.13	0.83	2.17				
	in situ				2.96								1.34				
	NRR		1.31								1.21						

3.1 Abyssal Plain

The Abyssal Plain is represented by station PAP located in the North Atlantic. However in situ sampling is limited, with nitrate only measured from mid 2002 to mid 2004 and to a maximum depth of 400 m, and chlorophyll from mid 2003 to mid 2005 with maximum depth of 200 m, described in Table 2.2. Surface chlorophyll is derived from SeaWIFs (8-day averages) and is available for the full 10-year time series as seen in Figure 3.2. As discussed in the previous chapter, section 2.4, some of the MEDUSA original parameters have been changed, in order to accommodate the oligotrophic region. Since the in situ data is only available from mid 2002, a model output is used as an initial condition. Subsections on using MEDUSA parameter instead of the modified parameter and using in situ DIN as initial condition will be shown in sections 3.1.3 and 3.1.4, respectively, in order to examine the effect of changing parameter and initial condition to the ensemble range and phytoplankton phenology.

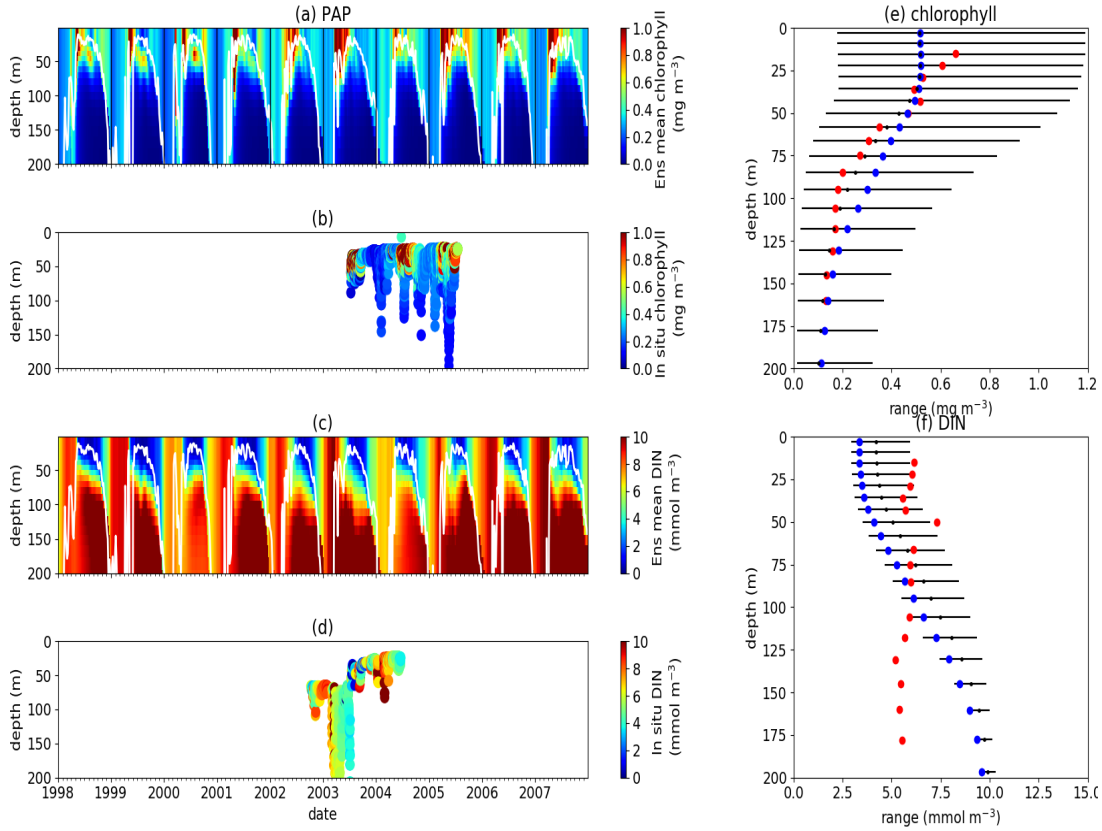


Figure 3.1: Chlorophyll and DIN profiles from ensemble mean ((a) and (c) respectively), in situ observations ((b) and (d) for chlorophyll and DIN respectively), and minimum and maximum range of concentrations at each depth ((e) for chlorophyll and (f) for DIN) at station PAP. The range are obtained by averaging the concentrations from all ensemble members for 10 years at each depths. Black dots in the second column show the mean concentration of the ensemble mean over the time series (from January 1998-December 2007), whilst blue and red dots denote the default run and in situ concentrations, respectively. White solid line in (a) shows mixed layer depth.

3.1.1 DIN and chlorophyll in light-limited region

The PAP data show seasonality in both chlorophyll and DIN profile concentrations, with high DIN during winter (December-April), and a decline in summer, Figure 3.1(b). The averaged DIN profile peaks in February, with a spike of high DIN in September, as shown in Figure 3.6(a). At around 70 m the highest concentrations of chlorophyll occur in May-June, summarised in Figure 3.1(b), similar to that in the surface.

The ensemble

The ensemble mean reproduces the seasonality in averaged DIN in Figure 3.6(a), where both ensemble mean and in situ show peak bloom between March and April, but a secondary peak is observed in July instead of the ensemble mean peak in September. The ensemble mean chlorophyll also has a seasonal cycle but with chlorophyll confined to shallower depths than in situ, summarised in Figure 3.1(a) and (e).

The ensemble mean chlorophyll starts to decline below 50 m, which also corresponds to the decline in the chlorophyll range with depth, shown on Figure 3.1(e). For chlorophyll profile the ensemble median is better than the ensemble mean, showing a higher correlation, a lower RMSE, and lower bias, than the ensemble mean against the in situ, as shown in Table 3.1. In terms of chlorophyll distribution, the ensemble range capture the in situ observation, and the ensemble mean concentrations are often close to the in situ concentration, shown in Figure 3.1(e), resulting an NRR of 1.20.

The ensemble range shows an increase in DIN concentration below 40m, however, in the in situ DIN, there are no clear nitricline, shown on Figure 3.1(f). The DIN vertical profile means fall within the upper limit of the ensemble range at the top $\sim$50 m, and within the ensemble between 50m and 100m. But below 100m, the ensemble mean cannot capture the in situ DIN concentrations, resulting an NRR of 1.25. The highest chlorophyll profile RMSEs ($>$ 0.62 mg m^{-3}) are produced from ensemble members that combine G_1 with $\rho_h\zeta_l$, $\rho_q\zeta_q$, and $\rho_h\zeta_s$, and this also coincides with high chlorophyll profile concentrations ($>$ 0.7 mg m^{-3}). However for DIN the ensemble mean/median RMSEs are higher than in other regions, which is traced to be due to ensemble members that contain the U_tG_2 combination which has a particularly high DIN bias ($>$ 9 mmol m^{-3}), as shown in Figure 3.20(j).

At the surface, the ensemble peak chlorophyll occurs in May, similar to the satellite observation, although peak concentrations are higher than satellite observation. These peaks shows later timings compared to the in situ observation where the chlorophyll peaks occurred in mid-April (Hartman et al., 2010). The decline in surface chlorophyll in the observation has been captured by six ensemble members ($U_t\rho_h\zeta_hG_2$, $U_t\rho_h\zeta_sG_2$, $U_t\rho_l\zeta_qG_2$, $U_t\rho_q\zeta_lG_2$, $U_t\rho_q\zeta_qG_2$, and $U_t\rho_s\zeta_lG_2$), although with weaker correlations ($r =$ -0.14 ($\pm$0.06), $p < 0.05$). An inter-annual decline has been observed in the satellite-derived chlorophyll, ($r =$ -0.21, $p < 0.05$), shown on figure 3.3(a).

At the surface, in years 1998, 1999, and 2001, the satellite-derived chlorophyll is within the inter-quartile range, however in other years, it is well below the ensemble inter-quartile box limits, Figure 3.3(a). The ensemble spread for inter-annual means has NRR of 1.26, and there is also an effect on the overall 10-year ensemble spread, as seen in Figure 3.2 with an NRR=1.29. High surface chlorophyll concentrations with high RMSE ($>$0.8 mg m^{-3}) are seen when combining U_h with $\rho_h\zeta_l$, $\rho_q\zeta_q$, and $\rho_h\zeta_s$, see Figure 3.19(e) and (j), which is consistent with the largest errors in the profile average values. However the low chlorophyll concentrations ($<$ 0.4 mg m^{-3}) that are produced when combining U_t and G_2 in the profile averages, is not reproduced in the surface chlorophyll.

The ensemble range of surface chlorophyll annual mean is 0.7 mg m^{-3}. However if we only allow one process function to change in each ensemble member, keeping the other processes with their default functions, the new 11 member ensemble range reduces to 0.58 mg m^{-3}, still covering 83% of the full

ensemble (128 members). When the original MEDUSA parameters are used, the ensemble fits the inter-annual surface chlorophyll observations slightly better, but the DIN fit gets worse.

Default run

The seasonality for DIN at PAP from the default run mostly follows the ensemble mean and median, although the concentration is slightly lower in ensemble median, Figure 3.6(a). The default run also produces seasonality of chlorophyll profile. The distribution of chlorophyll from the default run is confined in the upper 50m in the summer. During winter, the chlorophyll with a concentration of $\sim$0.5 mg m^{-3} can penetrate down to $\sim$200m. These distributions are similar to the ensemble mean. For chlorophyll distribution in the water column, the default run overestimates the in situ below 75m, and some of the default run chlorophyll concentrations are very close to the in situ, shown in Figure 3.6(e). This produces a higher bias, RMSE, and lower correlation for the chlorophyll profile compared to the ensemble mean and median, as shown in Table 3.1. Default run also shows an increase in DIN concentration below 40m. The default run produces the lowest bias compared to the ensemble mean and median, but the RMSE is higher than the ensemble mean and median.

In terms of the surface chlorophyll the default run in general show higher concentration than the ensemble mean and median, as shown in Figure 3.2(a). The peak chlorophyll in the surface occurs in May, similar to the ensemble mean and median. The default run also overestimates the satellite-derived chlorophyll concentration. At most months, the default run produces higher chlorophyll concentrations than the satellite-derived chlorophyll, shown in Figure 3.4. This overestimation also carried on to the annual mean, although in 1999, the satellite-derived mean is very close to the default run 3.3(a). In the statistical metrics, the bias, RMSE, correlations from the default run are worse than the ensemble mean and median.

3.1.2 Phenology: the North Atlantic Spring bloom

Station PAP is a long-term ocean observatory that aims to observed pCO_2 and nitrate continuously (Körtzinger et al., 2008). This station is also located 350km northeast from the NABE site, where North Atlantic spring bloom was studied (Körtzinger et al., 2008). At station PAP, the initiation from the satellite-derived chlorophyll occurs in April, see Figure 3.8(a). Although typical North Atlantic blooms happen in spring (Behrenfeld and Boss, 2014), most peak blooms at PAP occur in late May-early June, as shown in Table 3.2 and Figure 3.4, which is one month later compared to the in situ observation by Hartman et al. (2010), shown in Figure 1.12. Additionally, a late bloom in September from satellite-derived chlorophyll occurred in 2005, making the mean bloom timing fall in June. The peak height is observed

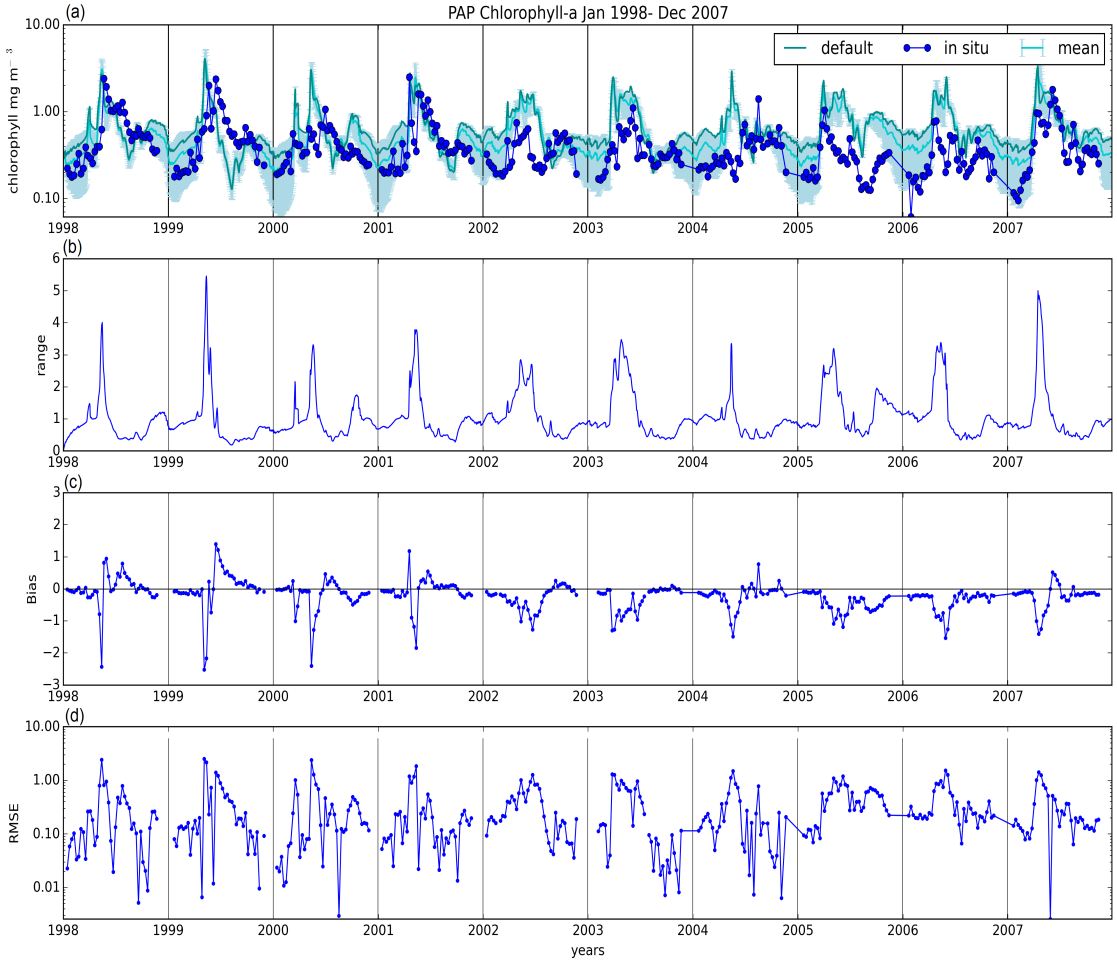


Figure 3.2: (a) SeaWIFs-derived surface chlorophyll-*a* at station PAP overlain with the ensemble 75th and 25th (blue shade), ensemble mean (cyan), and default run (dark cyan). Statistical metrics associated with the ensemble mean's surface chlorophyll such as range, bias, and RMSE are shown on (b), (c), and (d) respectively

to be 1.52 mg m^{-3} with an amplitude of 0.45 mg m^{-3} from the satellite. The duration of the bloom is around three months (95 days), putting the mean termination in July.

The observed initiation time at PAP is within the ensemble inter-quartile range, and the ensemble median is only one week earlier than the observations. However, due to the inter-annual variability, the observed bloom peaks occur about a month later than the ensemble mean and median, although the bloom timing is still within the ensemble range, with NRR value of 1.31. The ensemble mean produced higher peak chlorophyll (2.03 mg m^{-3}) and therefore higher amplitude. This puts the satellite-derived chlorophyll at the lower end of the ensemble range. The termination for ensemble mean and median is two days later and earlier respectively than the observed termination. This puts the satellite observed duration time within the ensemble inter-quartile range and very close to the ensemble mean duration.

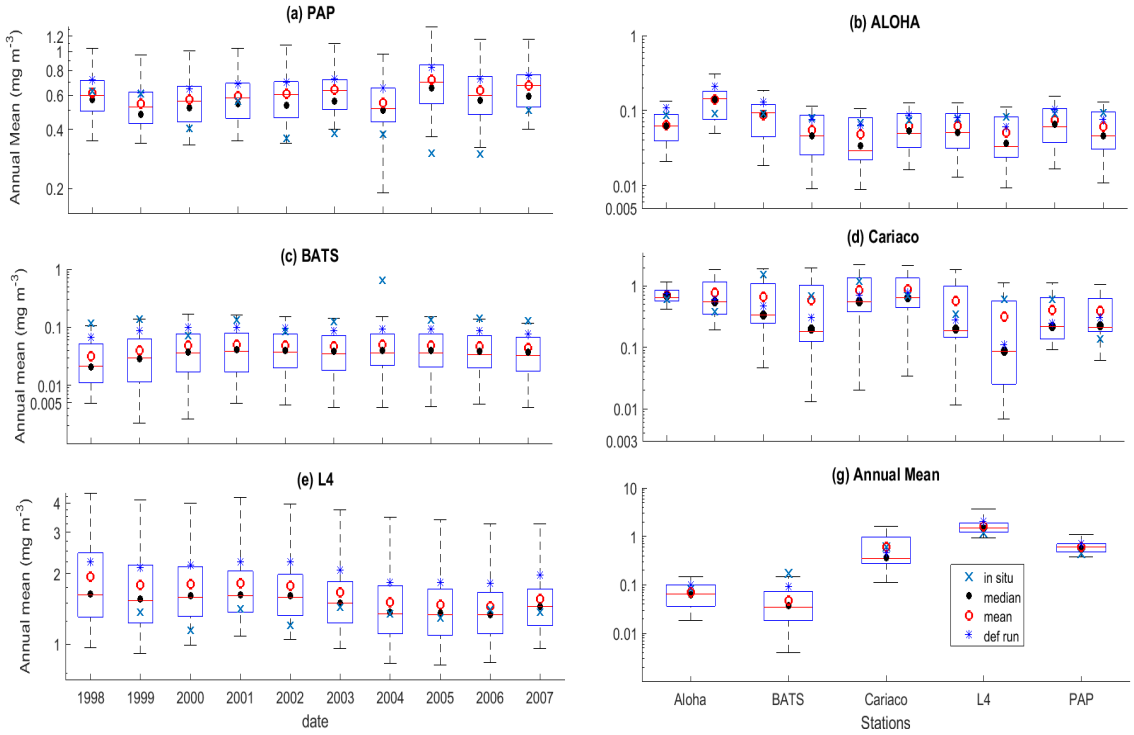


Figure 3.3: Inter-annual mean of surface chlorophyll from all the study sites ((a)-(e)) and the 10-year annual mean (g), all measured in mg m^{-3}. The boxplots show the ensemble annual means. Blue cross is the in situ observation, red open circle, black dot, and blue stars are the ensemble mean, median, and the default run respectively. The blue box is the 75^{th} (top) and 25^{th} (bottom) quartiles. Red line is the median. The whiskers are the ensemble minimum and maximum mean of surface chlorophyll. Annual mean values and NRR are described in Table 3.2.

3.1.3 Original MEDUSA parameters at station PAP

In this subsection, we present the ensemble results from running the 1-D MEDUSA in station PAP, using the original MEDUSA parameter which differ to that using the modified parameter, described in 2.4. From the results, the default overestimates the chlorophyll and show an overall poor match with the in situ when the model is run using MEDUSA parameters. The poor match may be due to the higher uptake and grazing parameter from HadOCC.

Running the ensemble using MEDUSA parameters produces similar seasonality, with the highest mean surface chlorophyll occurring in May. However, all the satellite-derived seasonal and inter-annual surface chlorophyll means are within the ensemble range, summarised in Fig 3.9(a) and (b), unlike that when modified parameters were used. This results in a broader ensemble spread, as indicated by the NRR values, summarised in Table 3.3, that are mostly closer to reliable spreads. The NRR value for the annual mean is broader than the modified parameters. The default run is also closer to the satellite-derived chlorophyll concentration, especially when it declines between July to October.

In terms of DIN, not all of the in situ observations have been captured by the ensemble range. The decline

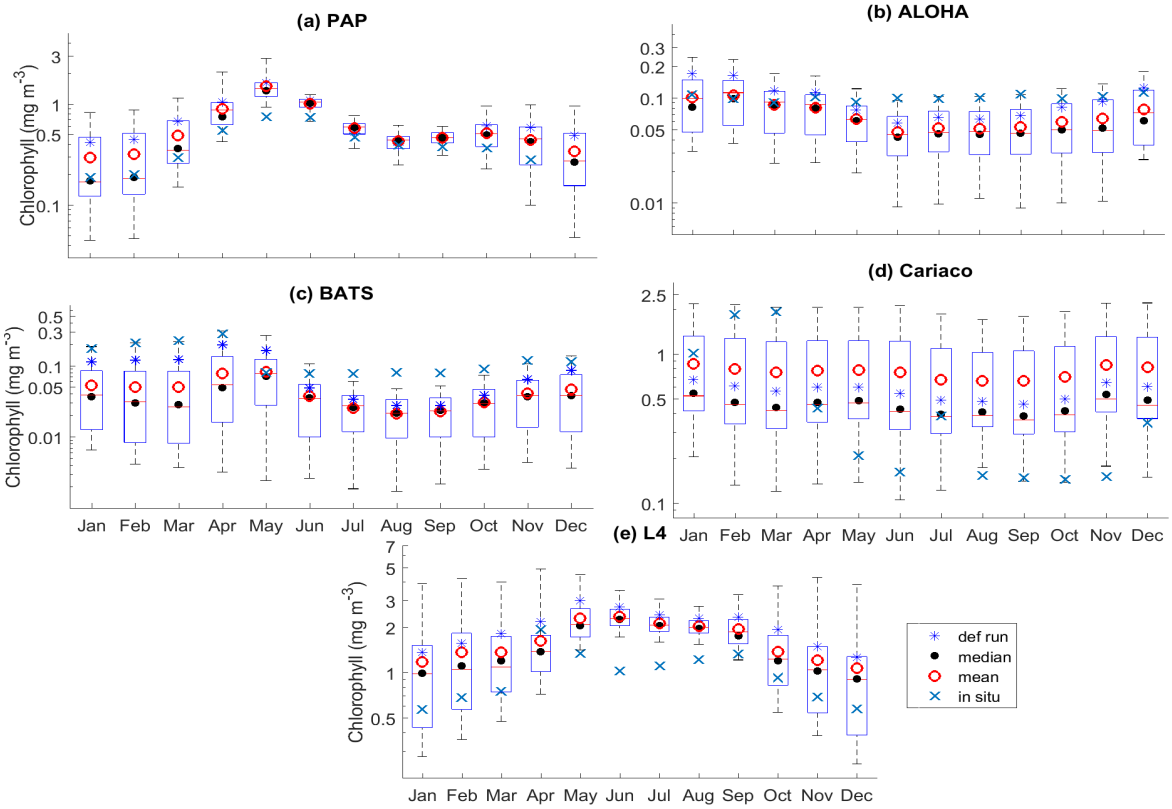


Figure 3.4: 10-year monthly mean surface chlorophyll from all the study sites ((a)-(e)), showing the seasonal dynamics of surface chlorophyll (mg m^{-3}). The boxplots show the ensemble seasonal means. Blue cross is the in situ observation, red open circle, black dot, and blue stars are the ensemble mean, median, and the default run respectively. The blue box is the 75th (top) and 25th (bottom) quartiles. The red line is the median. The whiskers are the ensemble minimum and maximum mean of surface chlorophyll. In station PAP, in situ data for December is not available due to low light and high cloud cover.

of DIN concentration occurring between February to August is not captured and therefore making the ensemble overestimates in situ monthly means. The DIN produced by ensemble only starts to decline between July to September, instead of February to March. This makes the NRR value for DIN profile higher (1.33) than that running the model using the parameters described in Table 1.

Table 3.3 summarised the statistical metrics when running the ensemble using MEDUSA parameters in station PAP. The default run produces better RMSE for DIN, although the correlation coefficient is higher in the ensemble median, similar to the modified parameters run. However, the RMSE and bias for the default run, ensemble mean, and median in DIN are generally higher compared to the modified parameters. Running the ensemble using MEDUSA parameters enhance the correlation coefficient for chlorophyll, both on the surface and over the depths, as well as reducing the RMSEs and bias for the ensemble mean, median, and default run. These improvements, therefore, making the NRR values for chlorophyll profile and surface closer to unity compared to using the modified parameter.

In the original MEDUSA parameter, the maximum uptake is slightly lower (0.53 day^{-1}), and the grazing

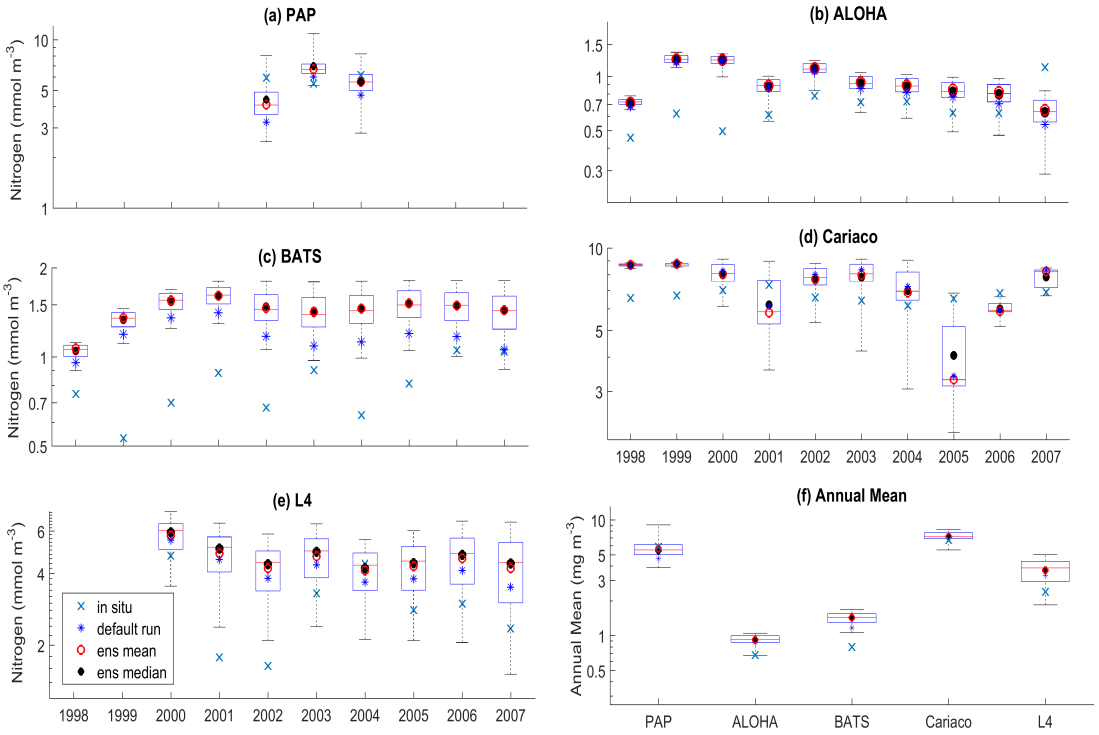


Figure 3.5: Inter-annual variability of DIN averaged over 200 m, from all the study sites ((a)-(e)), and the annual mean (f). Since the in situ data for PAP does not always cover the first 200m, the overall mean DIN concentration from all depth is used instead. For station L4, in situ DIN is only collected on the surface. Blue cross is the in situ observation, red open circle, black dot, and blue stars are the ensemble mean, median, and default run respectively. The blue box is the 75$^{\text{th}}$(top) and 25$^{\text{th}}$(bottom) quartiles. Red line is the median, and the whiskers are the ensmeble minimum and maximum of the averaged DIN. In station L4 and PAP data for DIN is only available from 2000-2007 and 2002-2004 respectively.

rate is twice higher (2.0 day^{-1}), compared to that in the modified parameter. This will cause lower phytoplankton abundance, and higher DIN concentrations as shown in Figure 3.9 and from the 10-year mean shown in Table 3.3, especially in surface chlorophyll, where the mean is reduced to 60%, resulting in smaller bias. However, the effect of structural sensitivity is still quite similar to when using the modified parameter, where higher chlorophyll is produced when G_1 is applied. In terms of DIN, apart from applying G_1, low concentration is produced, when the functional form is combined with $\rho_l\zeta_q, \rho_h\zeta_l, \rho_q\zeta_q$, and $\rho_h\zeta_s$. These functional form combinations coincide with low RMSEs. However, higher DIN concentration > 9.5 mmol m^{-3}, is observed when the ensemble members contained G_2, unlike that in the modified parameter whereby high DIN is only produced when combining U_4 with G_2. In terms of surface chlorophyll, lower concentrations when using functional forms that contain U_4 and U_3 have not been observed. Instead, high chlorophyll concentrations and hence RMSE, both in the surface and profile, is produced when the functional forms that contain $\rho_l\zeta_q$ and combining G_1 with $\rho_h\zeta_l$. These are summarised in Figure 3.10.

When using MEDUSA parameters, the observed peak height is not within the interquartile range, with

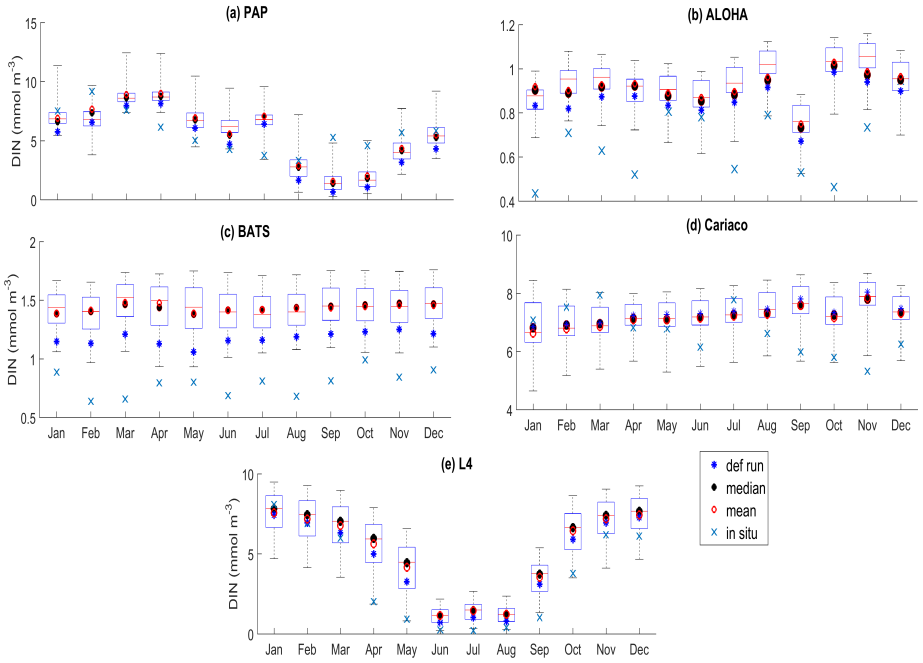


Figure 3.6: 10-year monthly mean of DIN averaged over 200 m from all the study sites ((a)-(e)), showing the seasonal dynamics of DIN (mmol m^{-3}). For station PAP, the DIN shown is the overall profile, and in L4, the in situ DIN concentration is only available at the surface. The boxplot shows the ensemble monthly means. Blue cross is the in situ observation, red open circle, black dot, and blue stars are the ensemble mean, median, and the default run respectively. The blue box is the 75^{th} (top) and 25^{th} (bottom) quartiles. The red line is the median. The whiskers are the ensemble minimum and maximum mean of averaged DIN. In station PAP, the in situ data is only collected from 2002-2004 and L4 from 2000-2007.

the default run, ensemble mean, and median underestimate the peak, shown in Figure 3.11(a). However for amplitude, the observation is within the interquartile range, although the ensemble mean, median, and the default run still underestimating the amplitude, unlike that in the original run. The observation of bloom initiation is still within the interquartile range, with the ensemble median showing similar timing as the observation (both in 20th April). The ensemble captures the observation on other metrics, such as the timing of the bloom peak, duration and termination, shown in Figure 3.11(b) and (c). This might be due to the broader range of ensemble that is produced from using the MEDUSA parameter instead of the modified parameter ensemble. For example, the initiation timing from the MEDUSA parameter ensemble occurs between March to early July, whereas the modified parameter range is only between March to early May. The most striking difference is in termination, where the MEDUSA parameter ensemble produces termination range between late May to November, and the modified parameter only covers termination time between June to October.

Indeed, using the MEDUSA parameters have improved the NRR values, RMSE, and correlation coefficients for surface chlorophyll and profile, but not DIN. The RMSE and bias for DIN are higher compared to using the modified parameters, as well as its NRR, indicating a narrow ensemble spread. Bias in

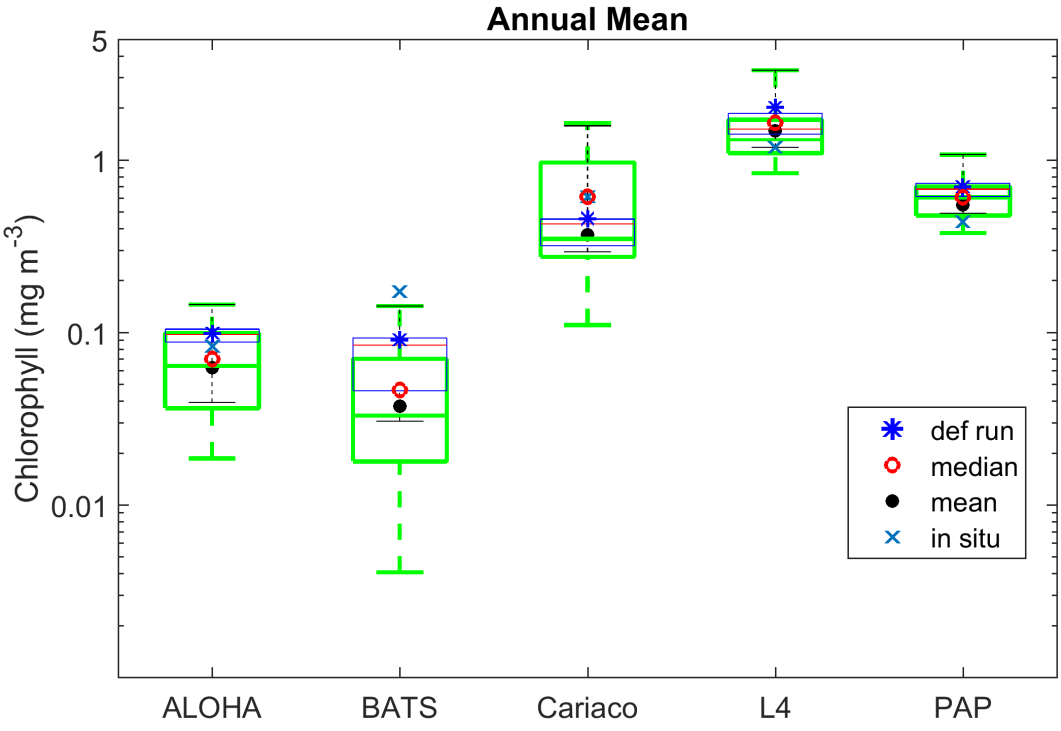


Figure 3.7: Annual mean of surface chlorophyll when changing only one process at a time (blue box), overlain with annual mean of all ensemble members (green box) at five oceanographic stations. Ensemble mean and median plotted in the figure (shown in red open circle and black closed circle), are the from the 128 ensemble members.

chlorophyll is also higher than that in the original run using modified parameters. Nonetheless, G_2 still produced lower chlorophyll concentrations. The mortality functions that produced high chlorophyll and low DIN are also similar to that in the modified parameter.

3.1.4 Station PAP using in situ DIN as initial condition

The in situ DIN is only available from mid-2003 to mid-2005, and are not available at all depths (from 7-400 m) for every in situ measurements, therefore we initialised the DIN using test station data from 50°N 20°W. Here, we averaged all the available DIN at each depth and used it as the initial conditions. The ensemble is run using the modified parameters in 1D MEDUSA. The inter-annual and seasonal mean are shown in Figure 3.12. In terms of chlorophyll, the satellite observation for the seasonal and inter-annual mean are always within the ensemble range, as shown in Figure 3.12(a) and (b) respectively. This results in the NRR being closer to unity in for annual mean (NRR= 1.05), and profiles (NRR= 1.07), although the surface chlorophyll is deemed too wide, summarised in Table 3.4. In terms of DIN, the inter-annual mean has an NRR value of 1.01, close to unity, as the in situ are always within the ensemble range, and there are only three data points available. However, in the seasonal means of DIN, shown

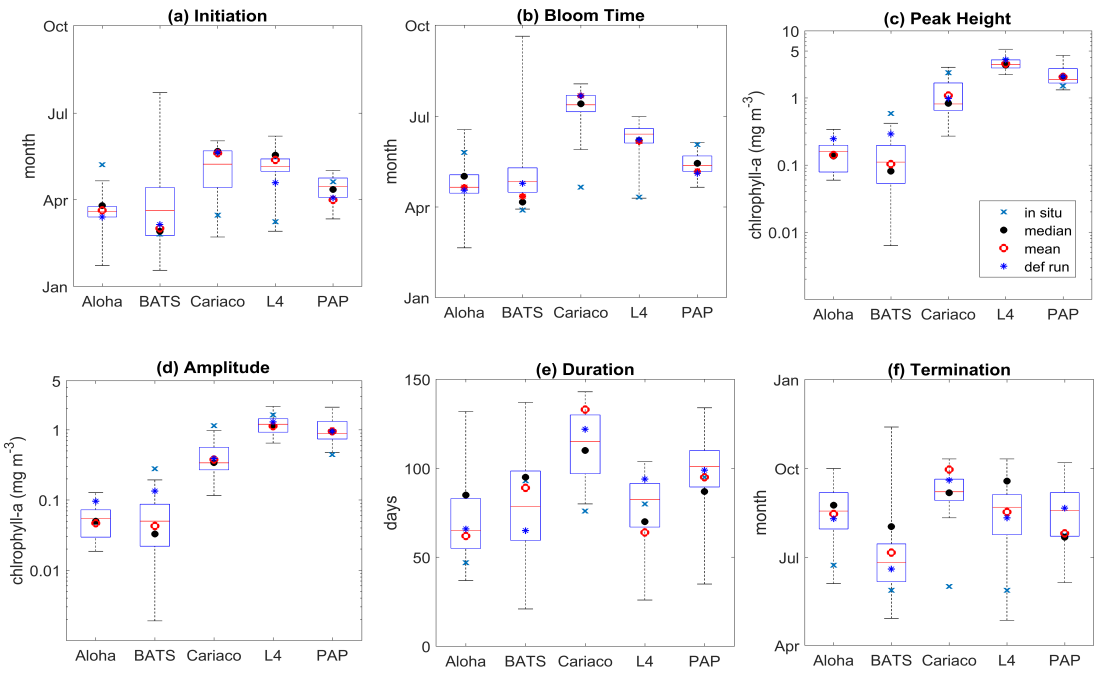


Figure 3.8: Phytoplankton phenology metrics at the five stations. Blue cross is the in situ, red, black, and blue dots are the ensemble mean, median, and the default run respectively. The timings and concentrations are averaged annually from January 1998 to December 2007.

in Figure 3.12(c), during months of high concentrations, the ensemble underestimate this. Unlike using the MEDUSA parameters (Section 3.1.3), when the model is initialised using in situ DIN, the ensemble can capture the decline of in situ DIN concentration between March to June, but not the increase in July. This mismatch of pattern, therefore narrows the ensemble spread to 1.38.

Using the DIN input from the in situ data has enhanced the correlation coefficient in terms of DIN and reduced the RMSE compared to the original run and using MEDUSA parameters. The ensemble range for mean chlorophyll and DIN means are lower, especially for DIN, where the range for the means is now less than half of the original run, summarised in Table 3.4. The low DIN concentration is expected as the in situ DIN, have lower concentration, and therefore reducing the ensemble range. Similar to using the test station data as initial conditions, the RMSE is lower in the ensemble mean and median compared to the default run, as well as the correlation coefficient. This is also similar to the chlorophyll profile and surface, although, in default run, the correlation is significantly lower than the one in Table 3.1. This run also simulated lower chlorophyll means compared to the in situ, especially in the ensemble mean and median, as indicated by the negative bias. The NRR value for the chlorophyll profile is also reduced, from 1.20 to 1.07. However, the mean in situ chlorophyll is underestimated by the ensemble mean and median, summarised in Table 3.4.

Similar to that in section 3.1.3 and using the original run, G_2 still produced lower chlorophyll concentra-

Table 3.2: Surface annual mean and phytoplankton phenology from in situ, ensemble mean, median, and default run. The range and NRR in the bracket are the values for changing the functional form one process at a time (shown on Figure 3.7).

Stations		Annual Mean (mg m^{-3})	Initiation Time	Bloom (mg m^{-3})	Peak Height (mg m^{-3})	Amplitude (mg m^{-3})	Duration	Termination
PAP	Ens mean	0.61	01 Apr	07 May	2.07	0.96	95	26 Jul
	Range	±0.70(0.58)	±51	±45	±2.98	±1.63	±99	±124
	NRR	1.26 (1.37)	1.14	1.31	1.08	1.09	1.42	1.60
	Ens med	0.55	12 Apr	15 May	2.03	0.95	87	22 Jul
	Default run	0.71	03 Apr	05 May	2.1	0.96	99	21 Aug
	in situ	0.44	20 Apr	03 Jun	1.52	0.44	95	24 Jul
ALOHA	Ens mean	0.07	21 Mar	21 Apr	0.14	0.047	62	15 Aug
	Range	±0.13(0.11)	±89	±119	±0.28	±0.11	±95	±119
	NRR	0.84 (1.17)	1.35	1.29	0.97	1.19	1.56	1.28
	Ens med	0.063	26 Mar	02 May	0.14	0.05	85	24 Aug
	Default run	0.10	14 Mar	18 Apr	0.25	0.096	66	10 Aug
	in situ	0.084	08 May	26 May	0.14	0.048	47	23 Jun
BATS	Ens mean	0.047	02 Mar	12 Apr	0.1	0.043	89	06 Jul
	Range	±0.14(0.11)	±187	±174	±0.42	±0.19	±116	±198
	NRR	1.40 (1.39)	1.18	1.17	1.42	1.42	1.08	1.20
	Ens med	0.038	28 Feb	06 Apr	0.08	0.033	95	02 Aug
	Default run	0.091	06 Mar	25 Apr	0.29	0.13	65	19 Jun
	in situ	0.17	25 Feb	29 Mar	0.58	0.27	93	28 May
Cariaco	Ens mean	0.61	20 May	22 Jul	1.09	0.38	133	30 Sep
	Range	±1.53(1.29)	±101	±66	±2.61	±0.86	±63	±61
	NRR	0.78 (0.90)	1.48	1.40	1.39	1.42	1.88	1.55
	Ens med	0.37	22 May	14 Jul	0.83	0.34	110	06 Sep
	Default run	0.46	21 May	22 Jul	0.98	0.39	122	19 Sep
	in situ	0.61	16 Mar	21 Apr	2.39	1.15	76	01 Jun
L4	Ens mean	1.65	13 May	06 Jun	3.25	1.13	64	17 Aug
	Range	±2.48(2.14)	±100	±82	±3.12	±1.50	±78	±167
	NRR	1.00 (1.36)	1.49	1.42	1.32	1.48	1.22	1.19
	Ens med	1.49	18 May	07 Jun	3.09	1.13	70	18 Sep
	Default run	2.03	19 Apr	08 Jun	3.73	1.3	94	11 Aug
	in situ	1.20	09 Mar	11 Apr	3.58	1.64	80	28 May

tions, and combining G_1 with $\rho_h\zeta_l, \rho_q\zeta_q$, and $\rho_h\zeta_s$ produce high chlorophyll profile concentrations which coincide with high RMSE. This has also been observed in the surface, although the RMSE is slightly higher, with the addition of ensemble members that contain $\rho_l\zeta_q$ and G_1 to have high chlorophyll concentrations > 0.6 mg m^{-3}. In DIN, using functional form combinations mentioned above, produce lower DIN concentration. Unlike the original run, we have not observed high DIN when combining U_4 and G_2, and overall, the ensemble produced lower chlorophyll with smaller range (see Table 3.4).

In terms of phenology, using in situ DIN as the initial condition, produces ensemble interquartile range that captures the satellite-derived chlorophyll, instead of overestimating the peak of the bloom, with the default run produces closer peak height (1.20 mg m^{-3}) to the in situ (1.52 mg m^{-3}), compared to the ensemble median and mean. The observed amplitude is also within the ensemble interquartile range, unlike the modified parameter ensemble. In terms of the timing of initiation, and bloom, the ensemble

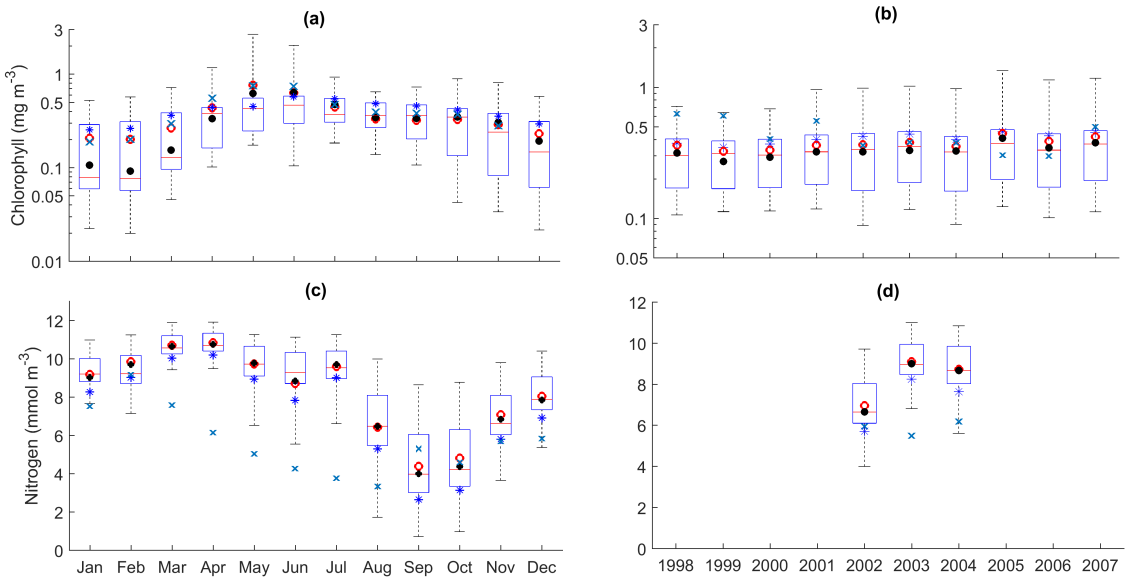


Figure 3.9: Seasonal and inter-annual mean of surface chlorophyll ((a) and (b)) and DIN ((c) and (d)) when running the ensemble MEDUSA using the original MEDUSA parameter. Blue cross is the in situ, red, black, and blue dots are the ensemble mean, median, and the default run respectively.

Table 3.3: Mean, correlation, RMSE, bias, and NRR for station PAP when the ensemble is run using MEDUSA parameters. The numbers in italic are statistical metrics from the ensemble run using the modified parameters.

Station PAP	DIN				chlorophyll				surface chlorophyll			
	r	RMSE	Bias	Mean	r	RMSE	Bias	Mean	r	RMSE	Bias	Mean
Ens mean	0.20	4.09	-3.08	8.99	0.50	**0.25**	0.08	0.32	0.48	0.32	0.04	0.40
	(±0.15)	(±1.82)	(±4.20)	(±4.27)	(±0.68)	(±0.41)	(±0.68)	(±0.75)	(±0.42)	(±0.70)	(±0.86)	(±0.86)
	0.23	*3.26*	*0.61*	*6.59*	*0.42*	*0.32*	*0.06*	*0.48*	*0.45*	*0.51*	*0.22*	*0.66*
Ens median	0.19	4.08	-2.16	8.90	0.60	0.26	0.110	0.29	0.54	**0.32**	0.09	0.36
	0.23	***3.16***	*0.54*	*6.38*	*0.49*	***0.29***	*0.003*	*0.42*	*0.54*	***0.46***	*0.15*	*0.60*
def run	0.19	**3.72**	-2.98	8.07	0.53	**0.25**	0.04	0.36	0.4	0.33	0.01	0.43
	0.21	*3.32*	*-0.20*	*5.64*	*0.28*	*0.40*	*0.18*	*0.59*	*0.36*	*0.57*	*0.30*	*0.74*
In situ				5.83				0.39				0.44
NRR			*1.25*	1.33			*1.20*	1.11			*1.29*	1.02

median can reproduce the observed initiation timing. Similar to the original run, the default run is outside the interquartile range of the initiation. For bloom time, the ensemble mostly produces earlier timing (between mid-April to June) compared to the observation (in 4th June), as seen in Figure 3.11(b), with the ensemble median showing the closest bloom timing to the observation (30th May). However, for termination of the bloom, the observation is within the ensemble interquartile range, but all the mean, median, and default run show later termination time. Compared to the original run, the initiation, bloom, and termination time from this ensemble are ∼3 days narrower than the original run. The observed duration of the bloom is also within the interquartile range, with the mean of the ensemble producing the closest duration, similar to the original run discussed in section 3.1.2.

Overall, using the in situ DIN concentration to initialise the model improves the DIN and chlorophyll RMSE and correlation coefficient, as well as the chlorophyll profile's NRR, but makes the DIN and

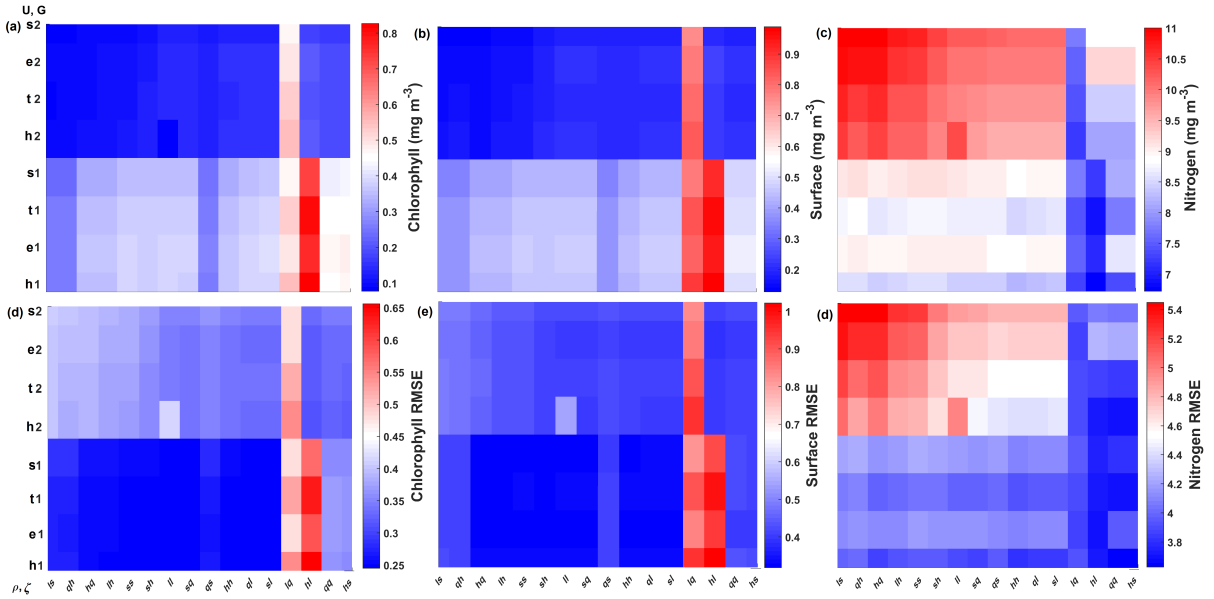


Figure 3.10: Mean chlorophyll profile (a), surface (b), and DIN (c), along with their RMSE ((d), (e), and (f) respectively) in station PAP when the ensemble is run using MEDUSA parameters. X-axis describe the combination of nutrient uptake (U) and grazing (G), and the y-axis describe the combination of phytoplankton (ρ) and zooplankton (ζ) mortality.

Table 3.4: Mean, correlation, RMSE, bias, and NRR for station PAP when the initial condition is the mean in situ DIN, using the modified parameters. The numbers in italic are the statistical metrics from the ensemble run using the modified parameters.

Station PAP	DIN				chlorophyll				surface chlorophyll			
	r	RMSE	Bias	Mean	r	RMSE	Bias	Mean	r	RMSE	Bias	Mean
Ens mean	0.26	2.77	-0.76	5.1	0.5	0.27	-0.08	0.29	0.44	0.32	-0.046	0.4
	(±0.06)	(±0.87)	(±1.85)	(±1.90)	(±0.57)	(±0.37)	(±0.50)	(±0.54)	(±0.31)	(±0.48)	(±0.60)	(±0.59)
	0.23	*3.26*	*0.61*	*6.59*	*0.42*	*0.32*	*0.06*	*0.48*	*0.45*	*0.51*	*0.22*	*0.66*
Ens median	0.26	**2.77**	-0.71	5.04	0.6	**0.26**	-0.11	0.32	0.47	**0.32**	-0.087	0.36
	0.23	***3.16***	*0.54*	*6.38*	*0.49*	***0.29***	*0.003*	*0.42*	*0.54*	***0.46***	*0.15*	*0.60*
def run	0.24	3.14	-1.37	4.44	0.16	0.34	0.043	0.44	0.11	0.45	0.037	0.48
	0.21	*3.32*	*-0.20*	*5.64*	*0.28*	*0.40*	*0.18*	*0.59*	*0.36*	*0.57*	*0.30*	*0.74*
In situ				5.83				0.39				0.44
NRR			*1.25*	1.38			*1.20*	1.07			*1.29*	0.9

surface chlorophyll NRR too narrow and broad, respectively. However, it reduces the ranges of mean DIN, chlorophyll profile, and surface. The general effect of perturbing functional forms are still very similar to that in the original run, apart from DIN where using G_2 and U_4 does not produce high (>9 mmol m^{-3}) DIN concentration. The NRR for DIN is also considerably narrow compared to the original run.

These two exercises show the importance of parameter adjustment and initial condition when running the biogeochemical model at a specific site. In terms of concentrations and seasonality, using the modified parameters and using the in situ DIN as initial condition yield a better match to the in situ chlorophyll profile, and satellite-derived chlorophyll. However, for DIN, the seasonality is better captured, and the ensemble range for both DIN and chlorophyll are broader when using the modified parameter and test

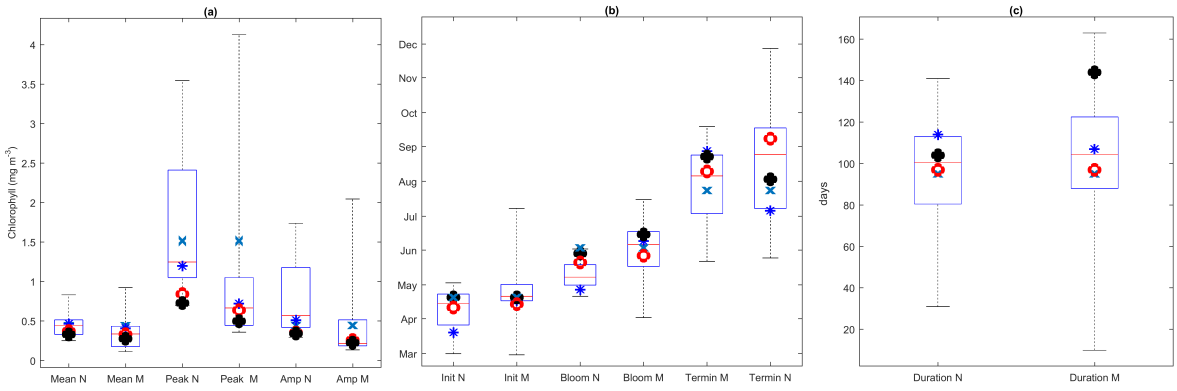


Figure 3.11: Phytoplankton bloom phenology at station PAP using in situ DIN (N) and MEDUSA parameters (M). Annual mean, peak, and amplitude of the bloom is shown on (a), the timing of initiation, bloom peak, and termination are shown in (b), and duration in (c). Blue cross is the in situ, red, black, and blue dots are the ensemble mean, median, and the default run respectively. M denote the spread from the ensemble that use MEDUSA parameters and N denote the ensemble that uses in situ DIN as initial condition.

data as an initial condition.

3.2 Oligotrophic Ocean

3.2.1 Low nutrients in the euphotic zone and DCM

The oligotrohpic region is represented at stations ALOHA and BATS. As discussed in section 1.6, this region is characterised by the scarcity of nutrients at the euphotic zone, but may be plentiful at deeper depths (Dave and Lozier, 2010; Lipschultz, 2001). High DIN levels (> 1.0 mmol m^{-3}) are only found below $\sim$ 150 m depth, shown in Figure 3.14(d) and (j) for ALOHA and BATS respectively. The annual means of the averaged DIN profile in the top 200 m are 0.68 mmol m^{-3} and 0.80 mmol m^{-3} for ALOHA and BATS respectively. In station ALOHA, an increasing trend of inter-annual in situ averaged DIN profile ($r = 0.69$, $p < 0.03$) has been observed, shown on Figure 3.5(b). In the oligotrophic region, seasonality have not been observed in both chlorophyll and DIN. However, there are low chlorophyll (< 0.1 mg m^{-3}) months that have been observed in July-October, as shown in Figure 3.4(b) and (c). Another feature of the oligotrophic ocean is a deep chlorophyll maximum (DCM) that occurs below the mixed layer, when the chlorophyll concentration in the surface is low (Fennel and Boss, 2003; Letelier, Karl, Abbott and Bidigare, 2004). At both stations, the DCM is observed $\sim$ 75m at BATS and $\sim$ 125m at ALOHA and continuously occurs throughout the year.

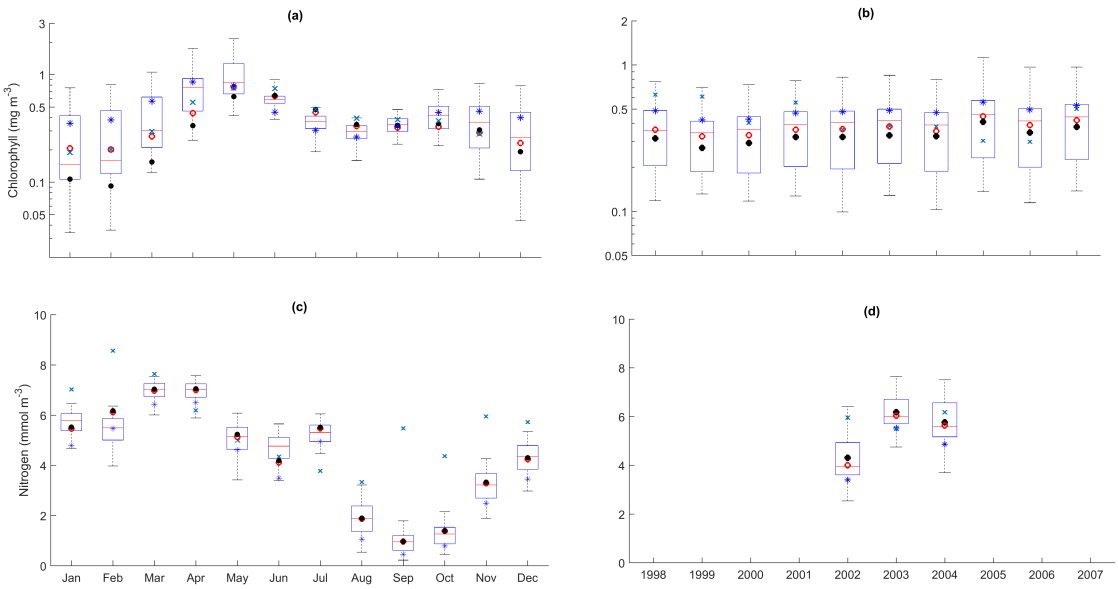


Figure 3.12: Seasonal and inter-annual mean of surface chlorophyll ((a) and (b)) and DIN ((c) and (d)) when running the ensemble using the in situ DIN. Blue cross is the in situ, red, black, and blue dots are the ensemble mean, median, and the default run respectively.

The ensemble

The ensemble mean has reproduced DIN concentration distribution in station ALOHA as seen in Figure 3.14(c) and (d). The ensemble range decreases as the depth increases for DIN, with a high ensemble range found at depths between 3-50 m, Figure 3.14(f) and (l). However, at BATS, the DIN concentration in the top 200 m, Figure 3.14(i), has significantly higher DIN concentration (> 1.0 mmol m^{-3}) than the in situ. This consequently leads to a higher annual mean of DIN, and overestimation in monthly, Figure 3.6(c) and inter-annual variability, Figure 3.5(c) of averaged DIN profile, for all the ensemble members. This higher averaged DIN profile concentration has also been observed at ALOHA, whereby both the ensemble mean and median have annual means of > 0.9 mmol m^{-3}. The increasing trend in DIN ($r = 0.67$, $p < 0.03$) has also been observed in 28.9% of the ensemble members, which uses G_2 as its grazing function.

The DIN concentrations at both oligotrophic stations are mostly overestimated by the ensemble mean/median, and the opposite has been observed in chlorophyll. At BATS all of the ensemble members overestimate chlorophyll profile, surface, and integrated. From Figure 3.18(a) and (b) a low chlorophyll profile means (< 0.08 mg m^{-3}) are produced from ensemble members that combine G_2 with $\rho_l\zeta_l$, $\rho_l\zeta_q$, $\rho_l\zeta_s$, $\rho_q\zeta_h$, and U_e, which coincides with high RMSE, shown in Figure 3.18(e) and (f).

The ensemble always encompass the in situ chlorophyll concentrations at the top 100m in ALOHA, and all depth at BATS, shown on Figure 3.14(k) and (e), although the ensemble mean and median un-

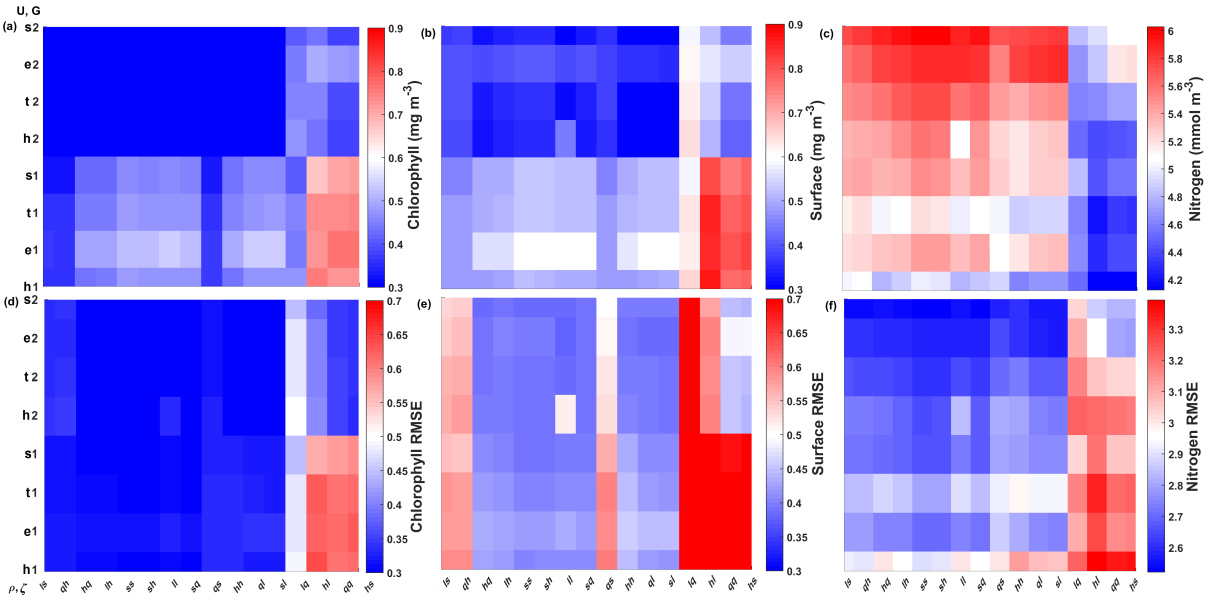


Figure 3.13: Mean chlorophyll profile (a), surface (b), and DIN (c), along with their RMSE ((d), (e), and (f) respectively) in station PAP when the ensemble is run using MEDUSA parameters. X-axis describe the combination of nutrient uptake (U) and grazing (G), and the y-axis describe the combination of phytoplankton (ρ) and zooplankton (ζ) mortality.

derestimate the in situ chlorophyll concentrations. DCM has also been observed at both stations in the ensemble mean. However, the depths at which the DCM is simulated by the ensemble means are slightly shallower, 70-150 m in ALOHA, but at similar depth in BATS. The ensemble at these stations show no continuous DCM, shown on Figure 3.14(a) and (g) for ALOHA and BATS respectively, despite continuous DCMs from the in situ. The depth of the DCM coincides with the higher ensemble range, with the range decreasing with depth after the DCM depth.

High chlorophyll vertical profile means at both stations are produced from ensemble members that combine $U_h G_1$ and $U_t G_1$. High DIN profile means are produced when U_e and U_s are combined with any mortality functions, summarised in Figure 3.20(a) and (b). Combining these uptake functions with G_2 will also increases the mean DIN concentration even further, and therefore increase the RMSE in station BATS. Nevertheless, since overestimation of mean DIN concentration occurs in most of the ensemble members, and therefore the ensemble mean, the NRRs are high for both ALOHA and BATS (NRR=1.38 and 1.40 respectively).

For integrated chlorophyll in the oligotrophic region, the ensemble mean and median have smaller RMSEs and a better correlation coefficient, compared to the default run. The NRR for the integrated chlorophyll is closer to 1 compared to the surface and chlorophyll profiles. From Figure 3.14, in 1999, there is an increase of chlorophyll concentration due to a surge in DIN. This makes the integrated primary production in 1999 overestimated the in situ value, shown in Figure 3.15, although most of the year the in situ is within the interquartile range.

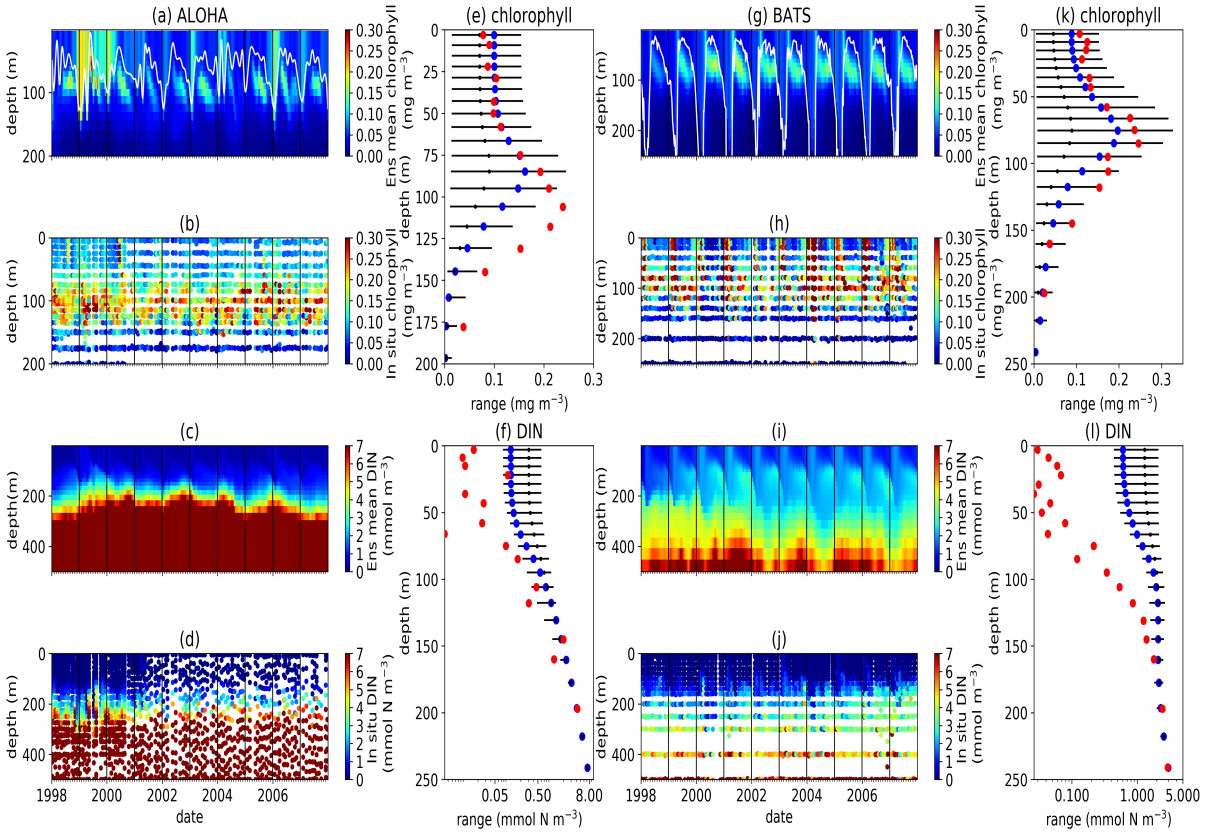


Figure 3.14: Time series (from January 1998-December 2007) of ensemble mean and in situ, and range of chlorophyll and DIN concentrations at oligotrophic stations. Station ALOHA is shown on (a)-(f) and BATS is shown on (g)-(l). White solid line in (a) and (g) represents mixed layer depth. (e), (f), (k), and (l) are the range of chlorophyll ((e) for ALOHA and (k) for BATS) and DIN ((f) for ALOHA and (l) BATS) over the depth. The blue and red dots represent the default run and the in situ mean concentration over the depth. The range is obtained by averaging the chlorophyll and DIN concentrations of each ensemble members over the time series at each depth. Black dots in (e), (f), (k), and (l) are the mean of the ensemble. Ensemble mean chlorophyll profiles (shown on (a) and (g)) and DIN ((c) and (l)) are obtained from all of the ensemble members. in situ chlorophyll are shown in (b) and (h), and DIN are shown in (d) and (j), for ALOHA and BATS respectively.

Default run

The default run also produces deep chlorophyll maxima and overestimates the DIN concentration at the top $\sim$150m at both stations, however, compared to the ensemble mean, the default run always produces lower DIN concentrations. This makes the DIN concentration from the default run closer to the in situ observation, especially in May, June, and December at ALOHA, shown in Figure 3.6. Consequently, bias and correlation of DIN at ALOHA and BATS for the default run is either similar or better than the ensemble mean and median.

In terms of chlorophyll, unlike the ensemble mean and median, the default run is very close to the in situ chlorophyll concentrations at the top 75m, especially at ALOHA. However, the DCM produced the default run is also not as continuous as the in situ chlorophyll for both ALOHA and BATS, although

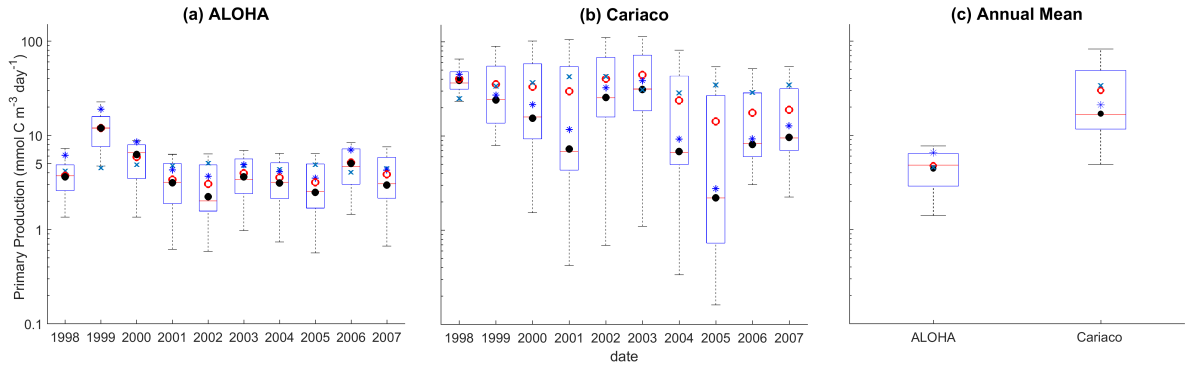


Figure 3.15: Mean integrated primary production averaged over 200m that are available in (a) ALOHA and (b) Cariaco, and (c) the annual mean. Blue cross is the in situ, red, black, and blue dots are the ensemble mean, median, and the default run respectively. The NRR for ALOHA and Cariaco are 1.12 and 0.80 respectively.

the concentration is higher than the ensemble mean and median, Figure 3.14(e) and (k). The depth of the DCM at ALOHA from the default run, occurred in ∼80m, shallower than the in situ, but at BATS the DCM produced by the default run is at a similar depth, and but still underestimates the in situ concentration, shown in Figure 3.14(k). Below 100m, the default run underestimates the chlorophyll concentration at ALOHA, making the bias and correlation for the default run better than the ensemble mean and median, summarised in Table 3.1. However, at BATS, in the top $\sim$ 40m, are overestimated by the ensemble, but below that depth, the default run produces higher chlorophyll, as seen in Figure 3.14(k). This also makes the default run produces better bias, RMSE, and correlation for chlorophyll profile, compared to the ensemble mean and median.

3.2.2 Surface Chlorophyll

At station ALOHA, shown in Figure 3.16, for the surface chlorophyll, the RMSEs from the ensemble mean and median are lower compared to the chlorophyll profile. The surface mean concentrations from the ensemble mean/median are also closer to the in situ concentration, summarised in Table 3.1. Ensemble members with low surface chlorophyll means are consistent with the profile averaged values, although high surface chlorophyll errors also coincide with high surface mean, summarised in Figure 3.19(a) and (f). The low RMSEs for surface chlorophyll at ALOHA from ensemble mean and median are also reflected in the NRR (1.07, see Table 3.1), and the ensemble almost always encompasses the in situ observations, as seen on Figure 3.16. At BATS, most of the ensemble members underestimate the in situ surface chlorophyll, similar to that in the profile. This could be seen from its surface time series where the in situ almost always overestimate the ensemble mean and median, especially during bloom period in 2004, Figure 3.17. This in turns produces high bias and RMSE for ensemble mean and median, as well high NRR.

At ALOHA the range for inter-annual means has low NRR (0.84), lower than the overall time series mean. However, if we only allow one process function to change in each ensemble member, whilst keeping the other processes with their default function, the new ensemble produces higher NRR (1.17), and the in situ annual mean is no longer within the inter-quartile range, as shown on Figure 3.7 and summarised in Table 3.2. In the interannual mean range, at both stations, although most of the ensemble members underestimated the in situ concentrations, most of the year, in situ surface chlorophyll is within the ensemble range. For the seasonal variability, from June to September, the ensemble at both ALOHA and BATS underestimate the in situ concentrations (the ensemble mean goes as low as 0.045 and 0.022 mg m^{-3} for ALOHA and BATS respectively), shown in Figure 3.4(b) and (c). However at months of high chlorophyll, (> 0.1 mg m^{-3}) the ensemble can capture the in situ.

The default run in oligotrophic regions generally produces higher chlorophyll and lower DIN concentrations compared to the ensemble mean and median. Default run also matches better with in situ data as the correlation coefficient, r is higher at both stations. However, at BATS due to the high chlorophyll concentrations, the default run shows better RMSE, correlation, and bias. In the interannual mean at ALOHA, in 2000, 2001, 2003, and 2004, and at BATS in 2002, the in situ is very close to the default run, which may explain the low bias.

3.2.3 Phenology at the oligotrophic regions

The observed initiation times at ALOHA show more inter-annual variability which is common in this station (Lipschultz, 2001) , and may occur in June, August and October, as well as in December and January. This causes the average observed initiation time to end up in May. The chlorophyll at ALOHA, Figure 3.4(b), shows peak highs (> 0.1 mg m^{-3}) in June, August, and September as well as December and January. At BATS, the initiation occurs mostly between January and February, although in 2002, the initiation occurred in October. Bloom peaks generally occur a month later, and the terminations vary between April and May, apart from 2002 when it was in December. The height of the peak and amplitude at ALOHA are 0.14 and 0.05 mg m^{-3} respectively. At BATS these metrics have slightly higher chlorophyll concentrations, of 0.60 and 0.28 mg m^{-3} for peak and amplitude respectively. The duration of the bloom at ALOHA is rather short compared to other stations, ~ 50 days, whereas BATS is ~ 90 days.

At ALOHA Figure 3.8(a) shows the ensemble run having initiation times between late January and April instead, so the observations fall outside the ensemble range, Figure 3.8(b), and the ensemble does not show as strong inter-annual variability as the observation, as seen on Figure 3.16. The ensemble at BATS has the largest range of phenological timings, especially in termination time, and this matches the

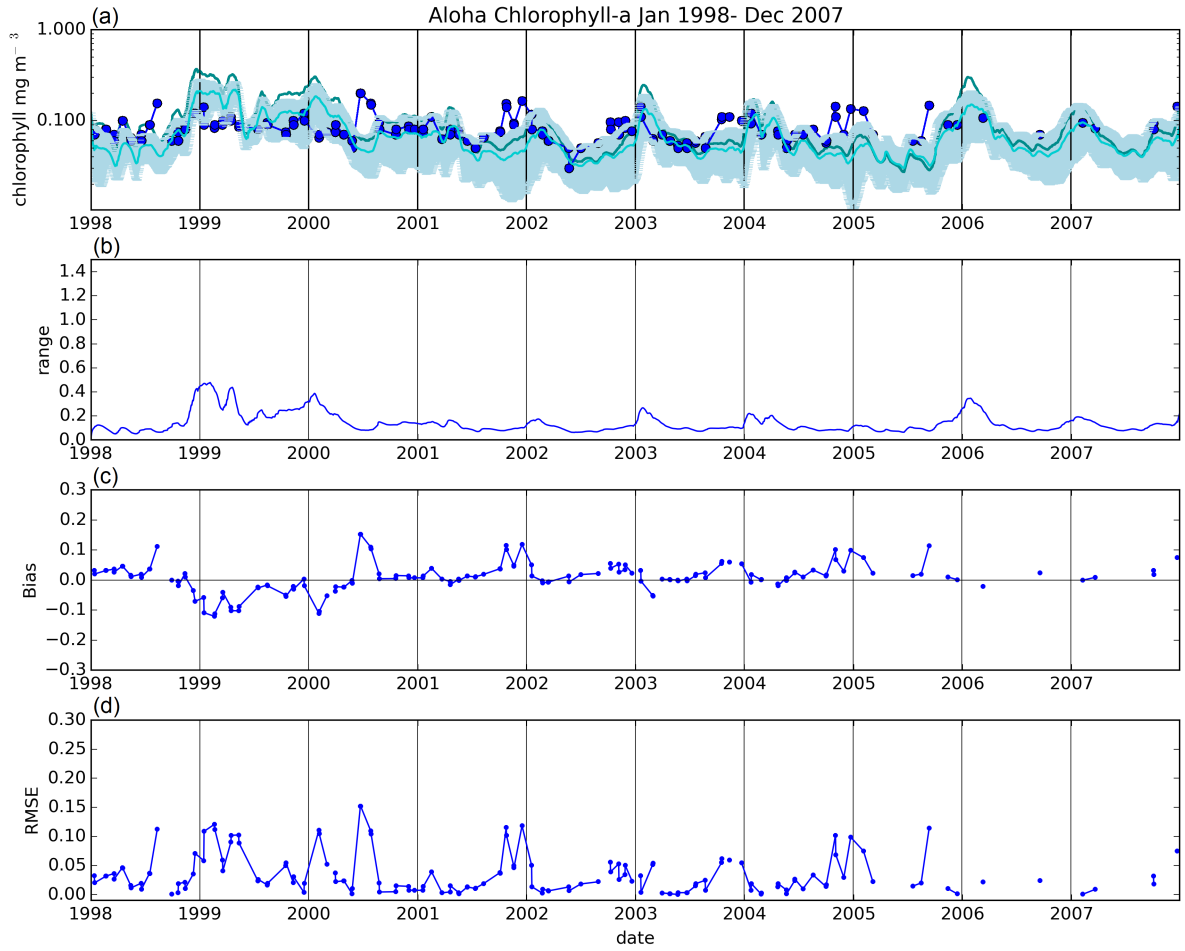


Figure 3.16: (a) in situ surface chlorophyll at station ALOHA overlain with the ensemble 75th and 25th percentile (blue shade), ensemble mean (cyan), and default run (dark cyan). Statistical metrics associated with the ensemble mean's surface chlorophyll such as range, bias, and RMSE are shown on (b), (c), and (d) respectively

observations better. For bloom initiation, the in situ timing is within the inter-quartile range and only three days earlier than the ensemble median. However, since the earliest peak occurs in mid-January in the ensemble, the peak bloom time from the ensemble at BATS are usually later than in situ. Nonetheless, the ensemble estimates of bloom peaks for 30°N, which agrees with a study by Racault et al. (2012), who identify early April as the peak bloom time. Although the ensemble range of peak bloom time at BATS is high (174 days), the NRR =1.17, as the ensemble does not cover all the in situ timings.

At ALOHA ensemble bloom peak and amplitude inter-quartile ranges encompass the observations, with ensemble mean and median being very close to the observation. However, at BATS, the in situ observations for bloom peak and amplitude are outside the ensemble range, consistent with section 3.2.2, where ensemble range underestimates in situ chlorophyll concentration. The observed bloom duration at ALOHA and BATS are within the ensemble range, although the inter-quartile range at ALOHA shows longer bloom durations. For bloom termination, both stations show later termination, with the ensemble

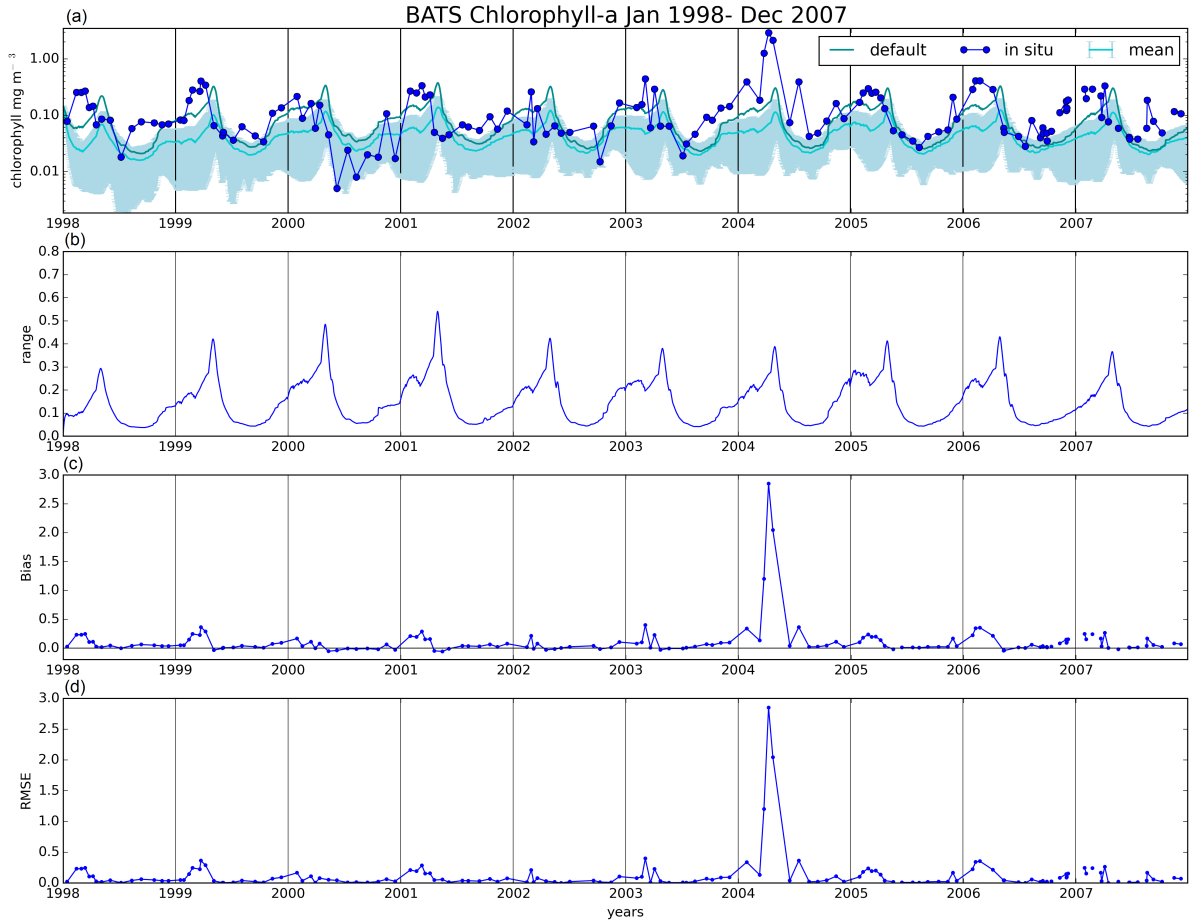


Figure 3.17: (a) in situ surface chlorophyll at station BATS overlain with the ensemble 75th and 25th percentile (blue shade), ensemble mean (cyan), and default run (dark cyan). Statistical metrics associated with the ensemble mean's surface chlorophyll such as range, bias, and RMSE are shown on (b), (c), and (d) respectively.

mean being almost two months late and a month late for ALOHA and BATS respectively. However, at ALOHA, located at 22°N, the ensemble median for termination at the end of August agrees well with the observations from Racault et al. (2012).

3.3 Coastal

Coastal zones are represented by stations Cariaco and L4, with in situ observations showing strong seasonality, in Figure 3.22(b), (d), (g), and (h). In this section, station Cariaco and L4 are going to be separated into two subsections, because the two stations are situated at different regions, Cariaco in the tropical, where nutrients supply to the euphotic zone is driven by upwelling in subsection 3.3.1, and L4 in the subpolar regions where the phytoplankton growth is mostly light-limited in subsection 3.3.2

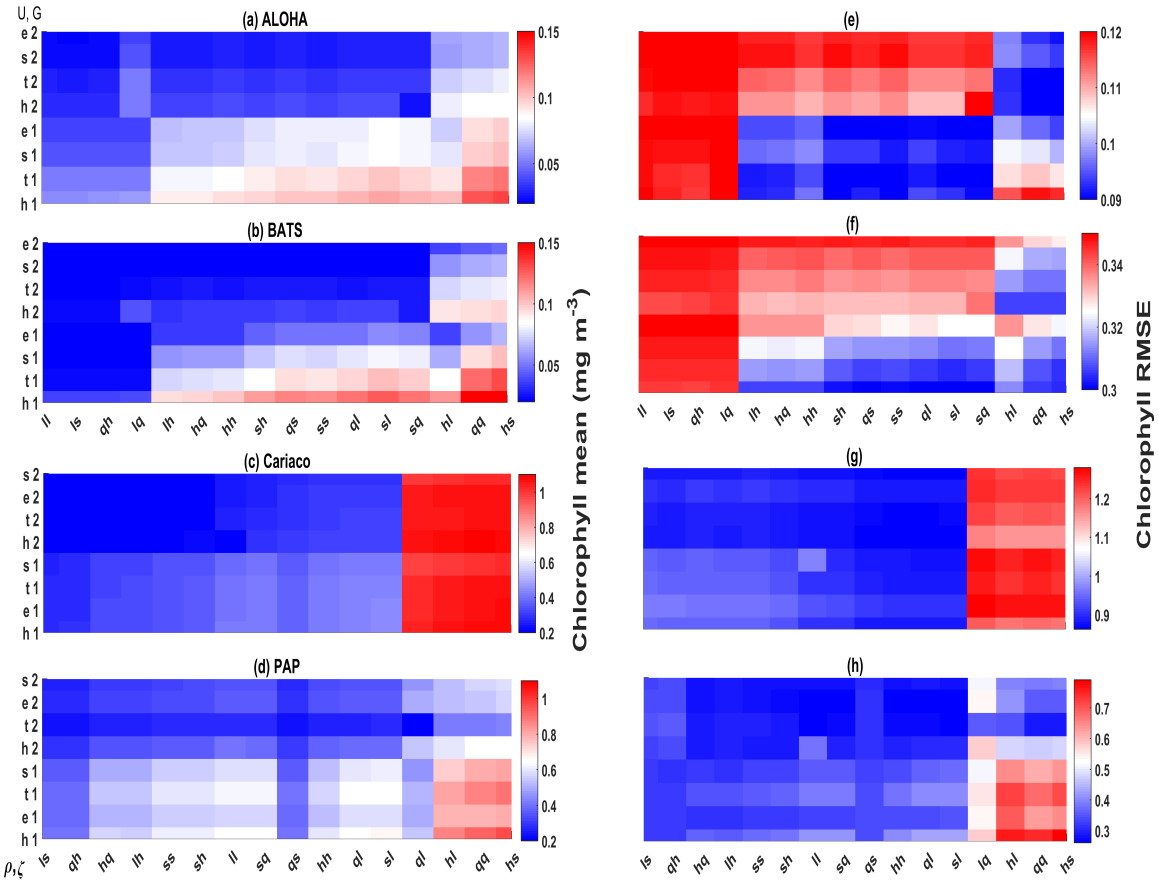


Figure 3.18: Chlorophyll profile 10-year means ((a)-(d)) and its RMSEs ((e)-(h)) at four oceanographic station from all of the ensemble members. Station L4 is not included as chlorophyll data is only taken at the surface. These are arranged by the lowest chlorophyll (top left) mean to the highest (bottom right), depending on the oceanographic regions.

3.3.1 DIN and Chlorophyll variability at Cariaco

The ensemble

At Cariaco, highest mean observed DIN profile concentration (> 7.5 mmol m^{-3}) at the top 200 m is observed in March and July, and lowest (< 5.5 mmol m^{-3}) in November, see Figure 3.6(d), as the upwelling controls the DIN concentration. The in situ profiles at Cariaco show high chlorophyll concentrations (>1 mg m^{-3}) within the upper 30 m between December-February (see Figure 3.22(b)) coinciding with the rise of DIN from depth to $\sim$30 m, seen in Figure 3.22(d), increasing the DIN concentration to $\sim$ 5 mmol m^{-3} from $<$ 1 mmol m^{-3}.

At Cariaco seasonal variability is absent from the ensemble for both chlorophyll and DIN, especially in the surface 3.21, which may be due to the absence of upwelling due to capping of vertical velocity. However the DIN bias have been massively reduced to 0.61 and 0.51 mmol m^{-3} for the ensemble mean and median respectively, summarised on Table 3.1, compared to before the capping (for the ensemble

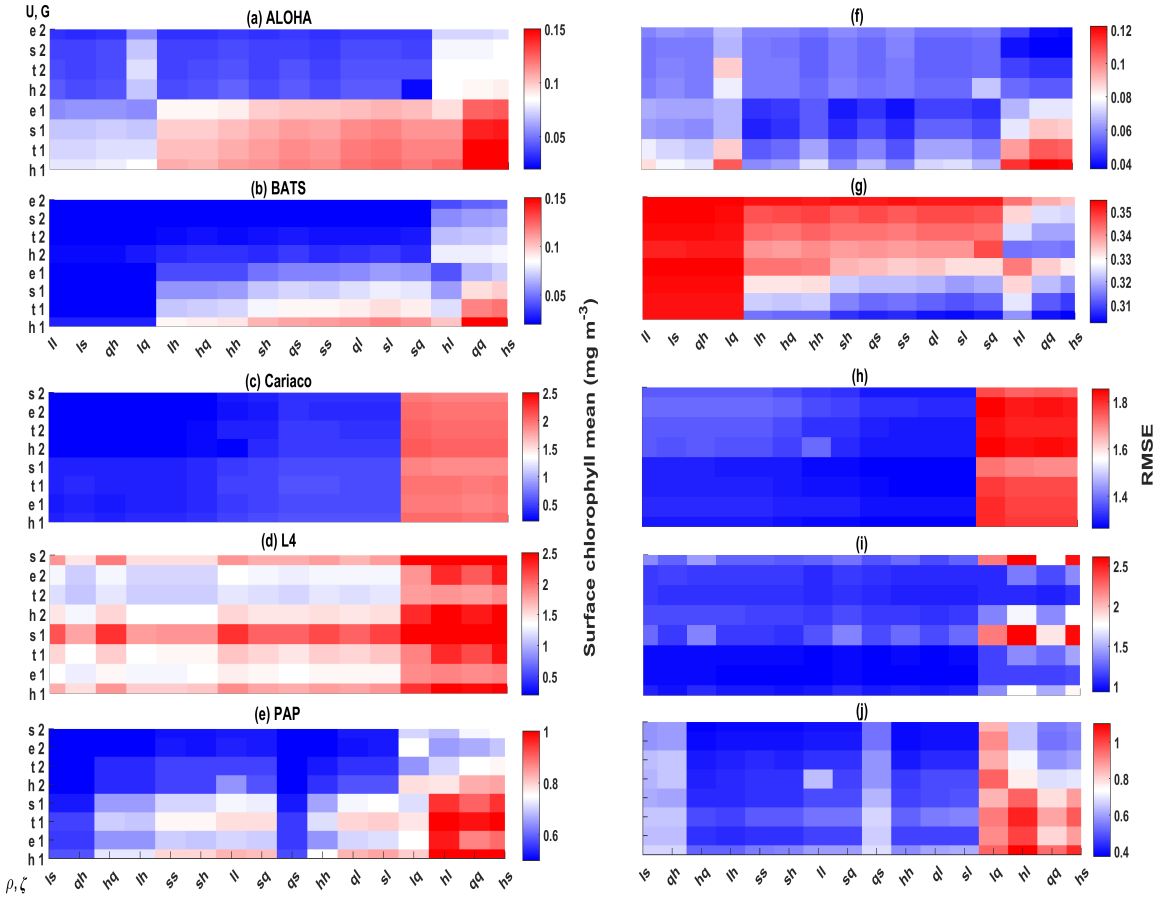


Figure 3.19: 10-year mean and RMSE of surface chlorophyll (mg m^{-3}) at five stations from all ensemble members. The first panel ((a)-(e)) shows surface chlorophyll mean and RMSEs are shown on the second panel ((f)-(j)). Concentrations and RMSEs are arranged by the lowest chlorophyll (top left) mean to the highest (bottom right), depending on the oceanographic regions. For station PAP, the sequence is sorted based on coastal station. The y-axis shows combination of uptake (U_h, U_s, U_e, and U_t) and grazing (G_1 and G_2), and x-axis shows combinations of phytoplankton (ρ) and zooplankton (ζ) mortalities.

mean the bias was 5.28, see Table 2.3). There are only two years simulating downwelling of DIN, in 2001 and between 2005-2006, shown in Figure 3.22(c), but this occurred in most years on the observation, shown in Figure 3.22(d). The chlorophyll concentration is almost constant (above 0.7 mg m^{-3}) in the upper 30 m and at the surface (see Figure 3.21), apart from a decline in concentration to $\sim$ 0.5 mg m^{-3}, followed by a sharp chlorophyll peak in the winter (December-January) in 2007, shown in Figure 3.4(d) and in Figure 3.22(a). Despite this discrepancy, the ensemble has captured the in situ chlorophyll and DIN profiles at all depth. The NRRs for chlorophyll and DIN profiles are 1.19 and 1.25, respectively. The ensemble mean for chlorophyll profile at Cariaco produced the smallest bias, compared to other stations, summarised in Table 3.1. A decline of chlorophyll was noted at Cariaco from 2004 (Taylor et al., 2012), and this is captured by the ensemble mean and median (r = -0.72, $p < 0.05$, r = -0.66, $p <$ 0.05 respectively).

At Cariaco although surface chlorophyll seasonality is not well reproduced, the ensemble range is wide

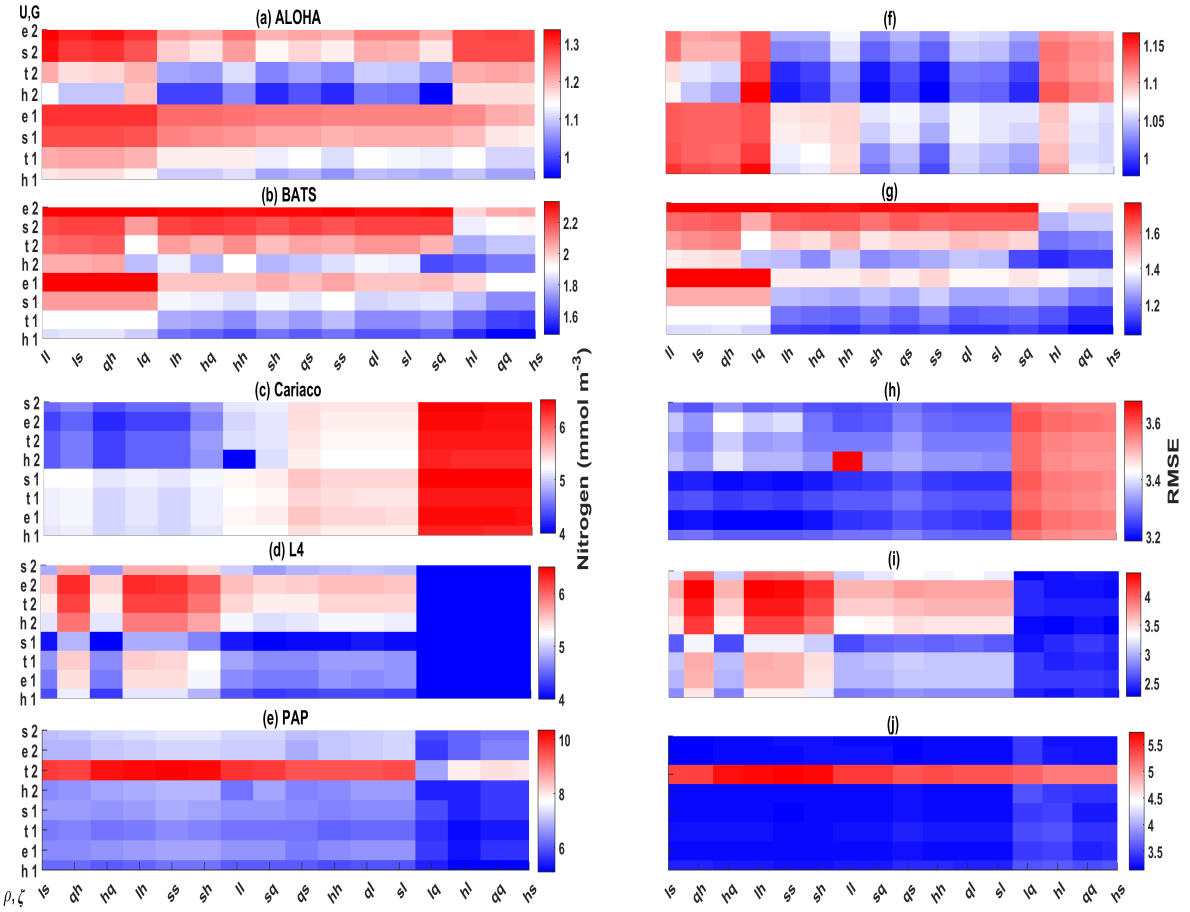


Figure 3.20: 10-year mean and RMSE of DIN (mmol m^{-3}), at five stations from all ensemble members. The first panel ((a)-(e)) shows DIN mean and RMSEs are shown on the second panel ((f)-(j)). Concentrations and RMSEs are arranged by the lowest chlorophyll (top left) mean to the highest (bottom right), depending on the oceanographic regions. For station PAP, the sequence is sorted based on coastal station. The y-axis shows combination of uptake (U_h, U_s, U_e, and U_t) and grazing (G_1 and G_2), and x-axis shows combinations of phytoplankton (ρ) and zooplankton (ζ) mortalities.

so that in situ concentrations mostly fall within it, apart from August and November, summarised in Figure 3.4(d) and 3.21. The annual mean of surface chlorophyll and averaged DIN in the top 200 m are also within the ensemble range, Figure 3.3(f) and 3.5(f), with the NRR 0.78 and 1.15 for chlorophyll and the averaged DIN profile respectively. Weak positive correlations, small RMSEs, and bias are found for ensemble mean surface chlorophyll at Cariaco, summarised on Table 3.1, which has improved the default run. Similar to the oligotrophic stations, the integrated chlorophyll shows better correlation with in situ measurements, compared to both surface and chlorophyll profiles, with results summarised in Table 3.1.

The overestimation of DIN and chlorophyll are produced when the model uses $\rho_l\zeta_q$, $\rho_h\zeta_l$, $\rho_q\zeta_q$, and $\rho_h\zeta_s$ combinations as these functional forms produced high chlorophyll means ($>$ 0.8 mg m^{-3} in Cariaco profile, with higher RMSEs, primarily when U_s is also used. Higher DIN concentrations in Cariaco ($>$5 mmol m^{-3}) with high RMSEs ($>$ 3.4 mmol m^{-3}) are also associated with the same ensemble members, summarised in Fig 3.20(c). From Table 3.1, at Cariaco in situ surface chlorophyll concentrations, are

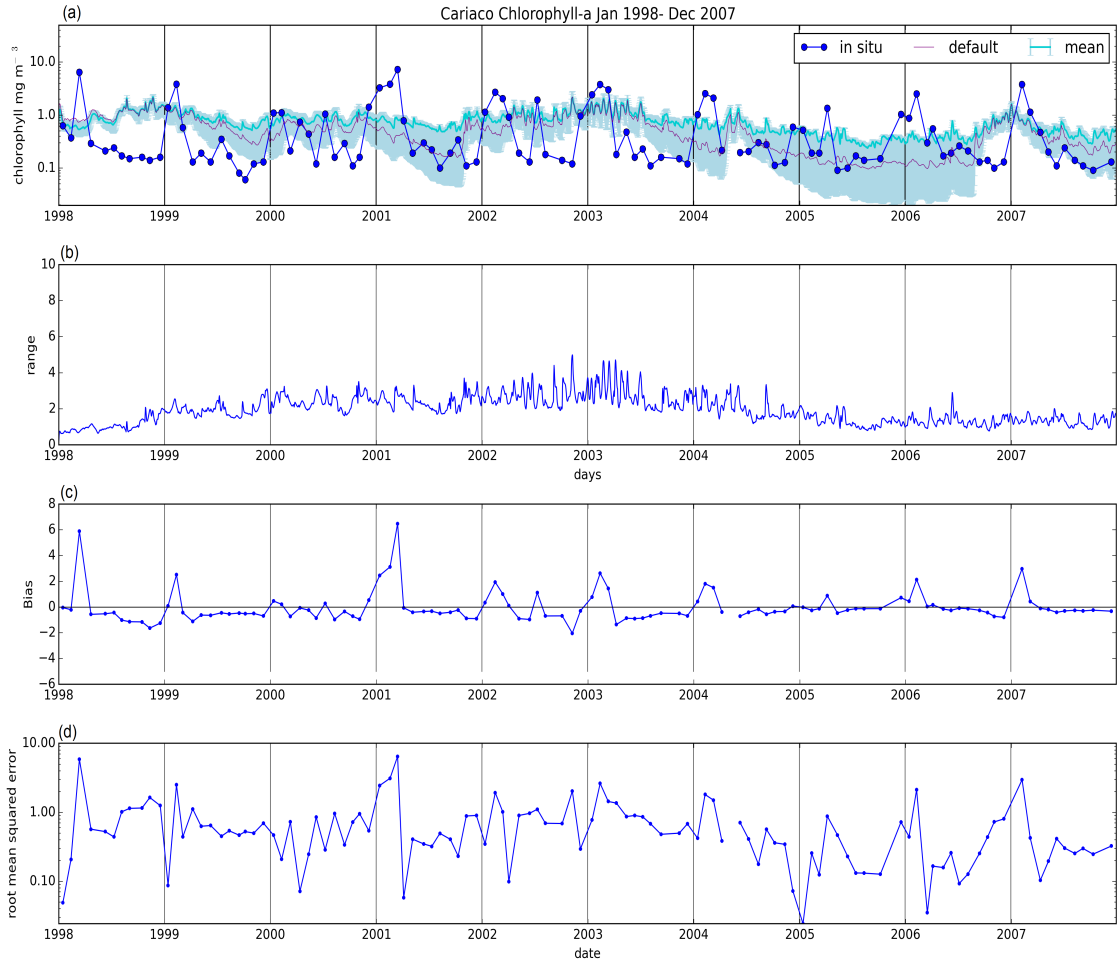


Figure 3.21: (a) in situ surface chlorophyll at station Cariaco overlain with the ensemble 75th and 25th percentile (blue shade), ensemble mean (cyan), and default run (violet). Statistical metrics associated with the ensemble mean's surface chlorophyll such as range, bias, and RMSE are shown on (b), (c), and (d) respectively.

slightly overestimated by the ensemble mean, but most other ensemble outputs are underestimated, except for ensemble members that use the combinations above. Unlike the oligotrophic regions, these high chlorophyll concentrations in the coastal stations coincide with higher RMSE (> 1.7 m^{-3}).

Default Run

At Cariaco, the DIN concentration from the default run is slightly lower than the ensemble mean and median as shown in Figure 3.22(f), especially at the top 75m, but starts to overestimate the ensemble mean below this depth. Default run also shows very close DIN concentration to the in situ observation, making the bias slightly lower than the ensemble mean. For chlorophyll, the default run shows lower concentrations, as shown in Figure 3.22(e), at the top 30m, but then exceeds the ensemble mean concentrations below that depth. Although during times of high chlorophyll, such as in 2003 and 2007, the default shows almost similar concentration as the ensemble mean, shown in Figure 3.5(d), and 3.6(d).

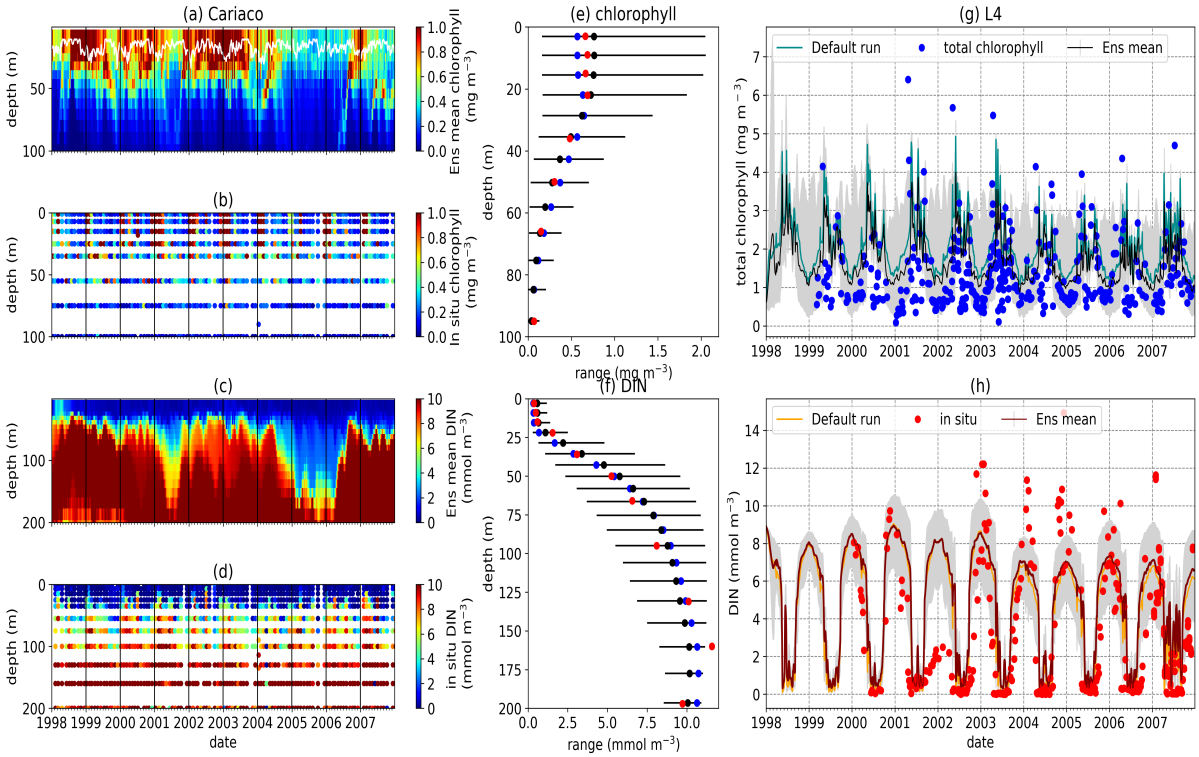


Figure 3.22: Time series of chlorophyll and DIN profile of ensemble mean, their range, and in situ concentrations at the coastal stations Cariaco (a-f) and L4 (g-h) from January 1998-December 2007. (a) and (c) show chlorophyll and DIN ensemble mean at Cariaco respectively. White solid line in (a) is the mixed layer depth. (e) and (f) show ranges of chlorophyll and DIN concentrations at each depth, respectively. The black dots are the mean of the ensemble. These range are obtained form the 10-year mean concentrations at each depth. Since in situ chlorophyll and DIN were taken at the surface in station L4, only surface time series were shown in (g-h). The grey shades on chlorophyll, shown in (g), and DIN, shown in (h) time series show the range of chlorophyll and DIN concentrations. Blue and red dots are in situ concentrations for chlorophyll and DIN respectively.

Although the bias for chlorophyll profile from the default run is higher than the ensemble mean, in surface chlorophyll, the decline is captured by the default run, although not as sharp as the ensemble mean and median (r =-0.35, $p < 0.05$). Overall, in this station, the default run shows poorer results compared to the ensemble mean and median, at most statistical metrics.

3.3.2 DIN and chlorophyll variability in light-limited L4

The ensemble

Similar to in situ Cariaco, the seasonality of DIN is also observed at L4, where high concentrations ($>$ 8 mmol m^{-3}) occurs between November-February, with very low values ($\sim$0.1 mmol m^{-3}), during summer months (Smyth et al., 2010). For chlorophyll sharp peaks of chlorophyll are observed during spring (March-April) and fall (September), coinciding with the sharp decline of DIN between March and July (from $\sim$6 in March to 0.22 mmol m^{-3} in July), shown in Figure 3.6(e), resulting in an annual mean

of 2.40 mmol m^{-3} and 1.20 mg m^{-3} for DIN and chlorophyll respectively. During non-bloom periods, chlorophyll is observed from 0.09-2 mg m^{-3}, with peak concentrations up to 6.4 mg m^{-3}.

At L4 the ensemble shows seasonality although the inter-quartile range often overestimates the surface DIN concentrations, especially during the sharp decline in spring and summer, shown in Figure 3.6(e). In terms of interannual mean, only in 2001 and 2002, the in situ is outside range. As explained previously, there are two bloom events in station L4 (Smyth et al., 2010). However, these are not well represented by the ensemble as it only simulates one peak between May-June, summarised in Figure 3.4(e). However, if only diatom chlorophyll concentration is shown, the two bloom events are captured, especially in the default run shown in Figure 3.23. Similar to the observation from Smyth et al. (2010), spring bloom is dominated by the diatoms, as shown in Figure 3.23(a). However, the non-diatoms in the ensemble tend to bloom in the summer and other times when blooms are not supposed to occur 3.23(b), which may cause the discrepancy in the seasonal variability at L4. The ensemble also shows higher concentration range during non-bloom periods (ensemble range from 0.28-3.13 mg m^{-3}), which is likely caused by the non-diatoms, so that the surface chlorophyll during periods of very low concentration (for example in January 2001 and May 2003) is not fully captured by the ensemble interquartile range.

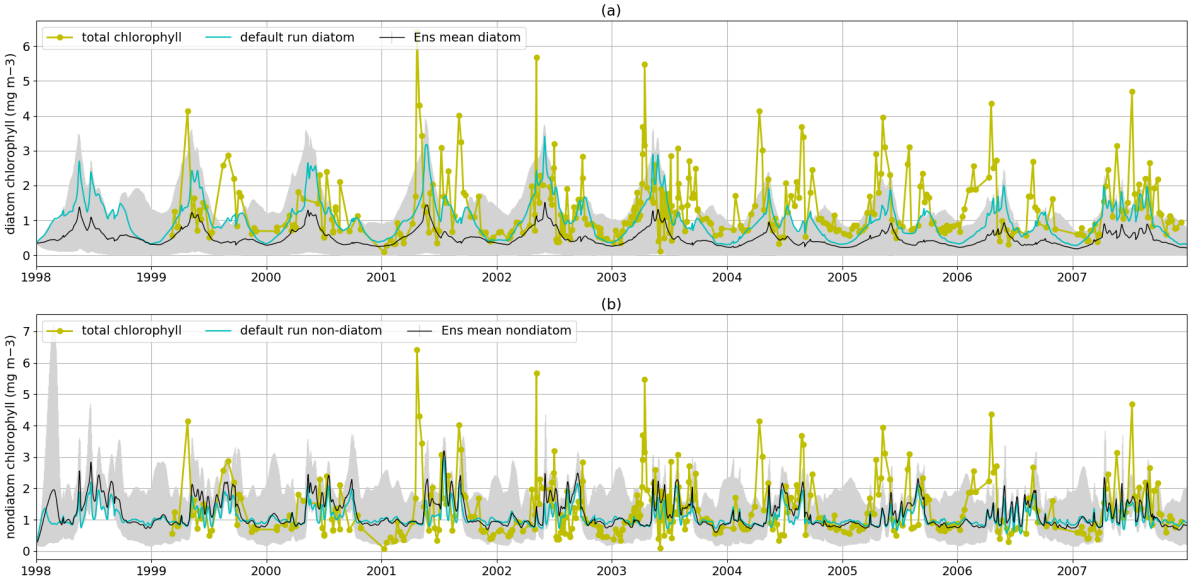


Figure 3.23: (a) In situ surface total chlorophyll at station L4 (yellow) and the ensemble 95th and 5th (grey shade) percentile, ensemble mean (black), and default run (dark cyan) of diatom surface chlorophyll. (b) Similar to (a), but the non diatom

The annual mean of the full ensemble has NRR=1.00 at L4, with the in situ chlorophyll close to the ensemble median (see Figure 3.3(e)). If the ensemble is reduced to single process perturbations the NRR increases to 1.36, and the in situ data is no longer within the ensemble range, shown on Figure 3.7. Similar to Cariaco, the overestimation of DIN and chlorophyll, therefore higher RMSEs, are produced when the model uses $\rho_l \zeta_q$, $\rho_h \zeta_l$, $\rho_q \zeta_q$, and $\rho_h \zeta_s$, summarised in 3.20(e)

Default Run

The default run shows similar seasonality as the ensemble mean, and produces higher total monthly chlorophyll and interannual means than the ensemble mean or median, shown in Figure 3.4(e). Due to the high chlorophyll concentration, the bias and RMSE are high, but, the default run produces a higher correlation coefficient compared to the ensemble mean and median. In terms of DIN, the default run produces lower DIN concentration than the ensemble mean and median, which are closer to the in situ concentration, therefore, producing lower bias and RMSE, as seen on Table 3.1. The default run produces a higher concentration of diatoms than 90% of the ensemble, but for non-diatoms concentrations, the default run produces almost similar concentration as the ensemble mean, shown in Figure 3.23. As discussed previously, the diatom concentration from the default run produces two distinct peaks of chlorophyll 3.23(a).

3.3.3 Phytoplankton Phenology

For coastal stations L4 and Cariaco, the in situ initiation typically happens in mid-March, with peak bloom timing in April for both stations. At Cariaco the mean peak height is 3.5 mg m^{-3}, with mean amplitude 1.15 mg m^{-3}, shown in Figure 3.8(c) and (d). At L4, the mean peak height is slightly higher (3.6 mg m^{-3}), with a higher amplitude (1.64 mg m^{-3}). Both stations have nearly similar bloom duration, of 76 and 80 days for Cariaco and L4 respectively. This makes the termination times for both stations very similar, which happen in June.

The ensemble means show later initiation, with the 75^{th} and 25^{th} spanning mid April to end of May for Cariaco, and between early and mid May for L4. However, the overall ensemble range covers the observed initiation, in Fig 3.8(a). This later timing is also clear in peak bloom times for both stations, shown on Figure 3.8(b), whereby in L4, the inter-quartile range of the bloom occurs mostly in June, and the ensemble range for Cariaco, between the end of May and early August. Consequently, the in situ observations for Cariaco and L4 both falls outside the ensemble range. Figure 3.4(e) shows that the bloom at L4 is simulated by the ensemble one to two months later.

At Cariaco the ensemble mean peak height (1.09 mg m^{-3}) and amplitude (0.38 mg m^{-3}) reach less than half of the in situ values (2.39 mg m^{-3} and 1.15 mg m^{-3}, for peak and amplitude, respectively), which makes the in situ concentration fall outside the inter-quartile range for peak height and amplitude. This underestimate of the peak and bloom amplitude, results in NRR of 1.40 and 1.39 respectively. At L4 chlorophyll peaks are within the inter-quartile range, and amplitudes are within the full ensemble range. The bloom duration at Cariaco is also overestimated (up to 143 bloom days) and this, along with the

late initiation of the bloom, results in a three month late termination. Cariaco is the only station with peak bloom time, duration, and termination outside the ensemble range, due to the lack of chlorophyll seasonality, as noted in section 3.3, also resulting in higher NRR values. At L4 the duration of the bloom is within the ensemble range, however, since the initiation and bloom timing of the inter-quartile range is later than the observation, the inter-quartile range also shows later termination time.

3.4 Summary and Discussion

In this chapter, structural sensitivity associated with the mathematical formulation of the processes in an intermediately complex biogeochemical model have been investigated. This is done by generating its ensemble outputs of chlorophyll and DIN. The ensemble consists of 128 ensemble members, each with different process function combinations. In order to maintain phenomenological similarity, these functions are calibrated using non-linear least squares, while keeping the maximum process rates fixed and using the range of concentrations that have been observed across all of the stations. We have chosen nutrient uptake, zooplankton grazing, and plankton mortalities to vary, as these are the core processes of every marine biogeochemical model, from the simplest to the most complex. Through this approach, we provide a perturbed biology ensemble conditioned upon structural uncertainties in model formulation. Applying structural sensitivity in the 1-D framework has also allowed a large range of process variability to be explored for several different oceanographic regions, and with minimal computational cost. The results are compared with a single default run, and in situ observations at five oceanographic stations. From these assessments, we find that small perturbations in model structure can produce a wide range of results regarding chlorophyll and nutrient concentration as well as phytoplankton phenology. Compared to parametric sensitivity studies in biogeochemical models, studies of structural sensitivity are much more limited.

Our findings reveal that in all regions, the Holling Type II (G_2) grazing function lowers the chlorophyll concentrations especially at low concentrations, which has also been observed by Anderson et al. (2010) . The nutrients respond in the opposite direction with enhanced DIN concentrations. This is expected as at low concentrations, using the G_2 function will graze more phytoplankton, as shown on Figure 2.3(b). Even when fitting for a phytoplankton concentration range similar to oligotrophic regions (0.001-0.5 mmol m^{-3}) is applied, higher grazing rate with G_2 is still apparent in lower concentrations (< 0.2 mmol m^{-3}).

Pairing G_2 with the linear (ρ_l) mortality of phytoplankton, which constantly removes phytoplankton regardless of concentration, will reduce the chlorophyll concentration even further; but the opposite

happens when G_2 is paired with linear zooplankton mortality. Yool et al. (2011) has similarly shown that using a linear mortality causes the biggest changes in phytoplankton concentrations compared to quadratic and sigmoidal forms. In contrast, the default phytoplankton (ρ_h) and sigmoidal zooplankton mortality (ζ_s) produce the highest chlorophyll concentrations in all regions, similar to the experiment from Yool et al. (2011). If we use less than half of the current maximum mortality for both ρ_l and ζ_l, then the deviation in phytoplankton concentrations from the default run is less apparent. For example, mean surface chlorophyll obtained from running an ensemble member with $U_h G_1 \rho_l \zeta_l$ at station ALOHA, using $\mu_{nd}, \mu_d, \mu_{mi}$= 0.04 day^{-1} and μ_{me}= 0.08 day^{-1}, is 0.12 mg m^{-3} (default function is 0.11 mg m^{-3}), up from 0.07 mg m^{-3}. This shows that structural sensitivity to some extent captures the parametric sensitivity as well. However compared to the lower maximum mortality, the current parameter set shows lower error during the fitting process, and in order to be consistent with other functional forms, the current parameter set is used.

For nutrient uptake, the exponential (U_e) and sigmoidal (U_s) forms are inefficient as ensemble members which contain these functions produce low chlorophyll and especially high DIN (DIN) concentrations, as shown on Figure 3.19(a), (b), 3.20(a), and (b). This is very apparent in oligotrophic regions. Even though the functional forms have been optimised, the largest deviations occur when DIN is $<$ 1 mmol N m^{-3}, shown in Figure 2.3(a). This deviation still occurs when the concentration range is reduced to 0.001 - 5 mmol m^{-3}. However, the effect is not as noticeable as when using G_1 or G_2.

Stations that produce high chlorophyll concentrations also have high ensemble range, for example, at Cariaco where chlorophyll concentration is high, despite the discrepancy with in situ seasonality, the ensemble range still covers the in situ concentrations and the chlorophyll profile at Cariaco has an NRR value close to 1. However for annual mean chlorophyll and primary production (Figure 3.3(d) and 3.15(b)) the ensemble spread appears too large. Even in the reduced 11 member ensemble where only one process is changed, the range still covers 80% of the full ensemble range of the surface chlorophyll annual mean, see Table 3.2 and Figure 3.7. This emphasises that perturbing functional forms will produce a large range of model results. In some cases, this reduced range may be statistically more meaningful than the full range. For example, compared with the full ensemble, the reduced ensemble range for Cariaco's annual-mean chlorophyll gives an NRR closer to unity. Therefore, it may be possible to systematically reduce the number of ensemble members, whilst retaining a realistic ensemble range, which will reduce computational costs.

In the depth profiles, most of the stations capture the chlorophyll well, but DIN concentrations are mostly overestimated, especially at the oligotrophic stations. At Cariaco the best match between in situ DIN and chlorophyll with the ensemble range have been observed. However, mismatch between seasonal patterns in the observations and the ensemble has been shown in this station, which is mostly caused

by the physical dynamics. At Cariaco the upwelling of nutrients that feeds the phytoplankton is not well simulated by the vertical velocity, especially after the capping process described in section 2.5.1. This emphasises that despite using the ensemble approach, a coupled-biogeochemical model is only as good as its physical model (Doney, 1999), as the physical components such as mixing and upwelling dictate the seasonal pattern, phytoplankton community structure, and primary production (Sinha et al., 2010). Furthermore, in situ studies in Cariaco have shown that diatoms should be present in this station (Taylor et al., 2012; Pinckney et al., 2015), however all of the ensemble members, including the default run produce no diatoms. A further study in comparing the chlorophyll distribution in the water column between perturbing the physical input and the biogeochemical equations may improve our understanding on how the two perturbations produces alternative simulations.

At most stations, the ensemble mean produced lower RMSE and higher DIN correlations with in situ compared to the default run, as shown in Table 3.1. This suggests that the structural ensemble is also likely to produce a mean field closer to the observation than a single-structure model that has not been specifically tuned to one station. However for chlorophyll concentration the default run has higher correlation coefficient and lower bias than the ensemble mean and median, especially in the oligotrophic regions. This may be because using default function, the model produces higher chlorophyll and lower DIN than the ensemble mean and median, and in the oligotrophic regions the ensemble tend to overestimate DIN and underestimate the chlorophyll. Reducing the number of ensemble members, in a further study, may improve the bias and correlation in ensemble mean and median, as some of the ensemble members contribute to this high bias, especially those which uses ρ_I and G_2.

Even though at some stations, such as BATS, the in situ surface chlorophyll is underestimated by most ensemble members and the ensemble median and mean RMSE is higher for the monthly means (Figure 3.4(c)), the in situ or satellite-derived chlorophyll values (during months of high chlorophyll) are within the ensemble range. For example at PAP, ALOHA, Cariaco and L4 (with some exceptions in summer months) (see Figure 3.4(b)), the in situ and satellite derived chlorophyll are generally within the ensemble range. We further note that, considerable model bias such as lower modelled concentrations of chlorophyll, compared to the in situ data, has been observed for the default 3-D MEDUSA model itself, e.g., in the subtropical gyre (Yool et al., 2011). This may be due to the absence of DIN fixers and picoplankton in MEDUSA, which cause the increase of plankton concentration in the summer (White et al., 2015), or due to the fact that the phytoplankton uptake equation in MEDUSA does not allow phytoplankton to acclimatise in the oligotrophic region through optimum uptake kinetics (Smith et al., 2009; Yool et al., 2011), making this process inefficient.

Apart from the model's state variables such as chlorophyll and nutrient concentrations, we have looked into the model-derived phytoplankton phenology because of its importance to marine ecosystems. For

example the timing of phytoplankton blooms affect the survival of zooplankton and fish larvae, as observed by Cushing (1990). The timing of the blooms has also been shown to control the variability of pCO_2 in the sub-polar region (Bennington et al., 2009).

There are some relatively small differences in the timing of phenological events between the ensemble mean/median, and the default run, ranging from a couple of days to a couple of weeks, as shown in Table 3.2. However the timing of initiation, bloom peak, and termination show wide ensemble inter-quartile ranges for all stations and can lie between $\sim$20 and 100 days earlier than the in situ timing, apart from stations PAP and ALOHA, see Figure 3.8(b). For this reason at most stations the observed phenology metrics fall within the full ensemble range. The ensemble range also mostly encompasses the in situ peak and amplitudes, shown on Figure 3.8(c).

Despite having a reliable spread in the annual mean, stations such as L4 show some mismatch in phytoplankton phenology against observations. In situ initiation, bloom timing, and duration at L4 are earlier than most of the ensemble members, although still lying within the ensemble range. Some ensemble mean timings (termination and peak bloom time) in this station are similar to the satellite observations at this latitude (Racault et al., 2012). When in situ chlorophyll is fitted with a smooth curve, the highest peak mostly occurs during spring (March-April). But model metrics, including ensemble mean and median, are noisy, and peaks mostly fall in the summer (May-July), which makes the in situ timing falls in the lower end of the ensemble range. Moreover, at L4 distinct phytoplankton blooms occur twice a year: first in spring and the second in autumn (Smyth et al., 2010). These blooms are sometimes well simulated, e.g. in Figure 3.22(g) and especially when only diatom bloom is simulated 3.23(a), but are not as distinct as the in situ measurements, due to the variability of the model. Apart from the high non-diatom chlorophyll concentration during non-bloom periods (see Figure 3.3.2(b)), these discrepancies may also be caused by the way zooplankton select their prey in MEDUSA. In a study by Sailley et al. (2014) grazing selection based on total prey concentration can result in rapid nutrient turn-over, which leads to a single bloom peak, but if the selection is based on the stoichiometry of C:N, the nutrients would regenerate more slowly leading to two chlorophyll peaks. However, the difference in peak timing does not affect the duration of the blooms, and the in situ duration is well within the ensemble inter-quartile range. Furthermore, the ensemble also simulated that the spring bloom is dominated with diatoms, which is consistent with the observation (Smyth et al., 2015). More generally, discrepancies in predicting bloom timing by large-scale biogeochemical models are reported in many studies, e.g., Henson et al. (2017) and Kostadinov et al. (2017). Henson et al. (2017) shows that compared with the satellite data, the 3-D MEDUSA 2.0 (Yool et al., 2013) model estimates spring blooms starting $\sim$50 days late, and southern hemisphere subtropical blooms starting $\sim$50 days earlier.

By generating an ensemble of 7 CMIP5 models, Kostadinov et al. (2017) highlighted that the difference

in bloom timing between the model ensemble and satellite-derived chlorophyll is typically >1 month over most of the ocean. This agrees with our study (see, Table 3.2), as most of our ensemble members have earlier bloom initiation dates, and the difference between the ensemble mean and in situ bloom timing, e.g. PAP and L4, are more than one month. Additionally, the whole ensemble range produced by this study provides an uncertainty range for the timing of phytoplankton blooms. The ensemble range almost always encompasses the observed annual mean, peak height, and amplitude. Therefore it may be suitable to use the ensemble model in order to forecast these phenological aspects, by running the ensemble in 3-D. Further, it may also be possible to improve the accuracy of the ensemble range, by systematically removing certain ensemble members in a future study.

In this chapter, we have shown that using a perturbed biogeochemistry ensemble, MEDUSA is able to encompass the in situ and satellite observation of chlorophyll, and reduced the RMSE from the default run, without optimising the parameters. This study also emphasises that ensemble simulations can be designed to represent and estimate the uncertainty in an ocean biogeochemical model, as well as capturing the observations.

Finally, the unresolved biases between in situ observations and sometimes the entire ensemble of results, such as the phytoplankton peak timings at L4, emphasises that the inclusion of some missing processes, such as active prey selection, may be needed to improve the performance of the model (Friedrichs et al., 2007; Kriest et al., 2010; Sailley et al., 2014). Additionally functional forms which describe chemostat experiments, such as the droop function are not as structurally sensitive as the logistic equations (Aldebert et al., 2018), such as Monod and Holling type III, that are used in MEDUSA. Furthermore nutrient uptake adaptation (Smith et al., 2009) may improve the chlorophyll simulation at the oligotrophic stations. We did not include equations that allow such selection or species, as this study tries to ensure that all the equations have similar properties to the default MEDUSA, in order to show that perturbing the structure of the model equations can result in different plankton and nutrient dynamics. Comparing the performance of greater model complexity and the ensemble method is beyond the scope of this study. Furthermore, the mismatch of the phenology between the ensemble and the in situ observation, such as that in station Cariaco, may be largely caused by the physical input, which drives the upwelling and mixing process, therefore controlling the seasonal pattern of the phytoplankton (Doney, 1999; Sinha et al., 2010).

Chapter 4

Reducing the ensemble and generating perturbed physics ensemble

4.1 Introduction

From the previous chapter, MEDUSA has been run with fewer ensemble members whereby in each member, we only change one process function, while keeping the other processes in their default functions. From this experiment, the new ensemble covers at least 80% (observed in station BATS) of the original ensemble range in terms of an annual mean of surface chlorophyll, which shows that in the full 128 members ensemble, some members produce almost similar mean annual means concentrations. Therefore it is possible to reduce the ensemble, in order to minimise the computational time and cost, especially in a fully coupled physics-biogeochemical model in a 3D system, but still retaining the range, and may even improve the bias and RMSE. In this chapter, the method of reducing the ensemble and the selected members will be described in section 4.2–4.3.3.

Additionally, one of the essential inputs in running a marine biogeochemical model, is the physical input, as it determines the distribution and movements of nutrients and therefore affects other model compartments (Doney, 1999; Friedrichs et al., 2006; Sinha et al., 2010). The second part of this chapter will explain the method of generating a perturbed physics ensemble, especially the vertical velocity, which determine the nutrient supply to other biogeochemical compartments. The discrepancy between in situ and modelled chlorophyll have been observed when directly assimilating ocean general circulation model to force the biogeochemical model, due to spurious vertical velocity (Ourmières et al., 2009; Raghukumar et al., 2015; Ford and Barciela, 2017). The method of perturbing the physics is described in section 4.4. Furthermore, the perturbed physics ensemble will be quantified using statistical metrics

and surface interannual means and compared with the reduced perturbed biogeochemistry ensemble. To examine which perturbation (biogeochemistry or physics) produces more spread and variability, in this chapter, we will also explore when both biogeochemistry and physics are perturbed. The reduced ensemble is used to perturbed the biogeochemistry instead of the full 128-member ensemble for this experiment. The results of perturbing the physics, and both biogeochemistry and physics are explained in section 4.5.

4.2 Reducing the ensemble

This method aims to retain the RMSEs between the observation and the ensemble, despite reducing the ensemble. In the previous chapter, we have calculated the RMSEs of chlorophyll and DIN. Here, principal component analysis (PCA) is applied to the RMSEs so we can group ensemble members that are clustered together. Using k-means cluster, and calculating the ensemble members that are closest to each centroid, it is possible to select the ensemble members that best represent the full (128 members) ensemble, which may also retain the RMSE. Before applying PCA, we normalised the RMSEs, due to the difference in units between chlorophyll (mg m^{-3}) and DIN (mmol m^{-3}). This is done by dividing the RMSE of an ensemble member by the standard deviation of the overall RMSEs:

$$\frac{\sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(m_i - f_i\right)^2}}{\sigma} \tag{4.1}$$

where n is the number available days of in situ concentration, m_i, is the modelled concentration, f_i is the observed concentration, and σ is the standard deviation of the full ensemble's RMSE.

After normalising the RMSEs, we then apply PCA to the normalised RMSE (nRMSE) of chlorophyll and DIN, at every station, resulting in ten principal scores. In the scores, there are differences of (1) RMSE between different ensemble members, (2) regions (abyssal plain, oligotrophic, and coastal), (3) DIN, and (4) chlorophyll. The scores are then plotted into a three-component scores plot, showing how each ensemble members are clustered together. We then use k-means cluster, to group the ensemble members that are closer together, identify each cluster member, and select one that is closest to the centroids as the member which best represent a cluster. If a 10 clusters score plot is applied, then ten members will be selected as representative of the full (128) ensemble. The k-means distance between clusters is decided using squared Euclidean distance, whereby each centroid is the mean of the points in that particular cluster. We applied 20, 15, 13, 12, 11, 10, 9, 8, 7, 6 and 5 clusters (regions), as the typical ensemble members in a climate model are between 20 to 10 (Murphy et al., 2007). The k-means clustering is

applied independently to the full ensemble members, and we do not reduce the larger clusters into a smaller one. We also calculate the percentage of range coverage between the full ensemble and selected ensemble members.

In the previous ensemble runs, the model results that deviate the most from the default run is when linear plankton mortality is applied. This is consistent with the observation of Yool et al. (2011). To avoid this discrepancy, if the ensemble member that is the closest to centroids contains linear mortality, the ensemble will be replaced. However, this member must be relatively close to the centroids. If the cluster only has linear mortality, the functional form is not changed. We will explore the effect of replacing linear mortality in section 4.3.2 and 4.3.3.

4.3 Results from reducing the ensemble

The result section will be separated into several subsections. In subsection 4.3.1, we will explore how reducing the ensemble members will affect the ensemble range and the RMSE. Furthermore we will look on which ensemble shows better range coverage of the full ensemble members. The effect of replacing these ensemble members with non-linear mortality is explored in subsection 4.3.2. Finally, the comparison between the ensemble means of the reduced ensembles, and the full ensemble will be explored in section 4.3.3

4.3.1 The range and RMSEs of the new ensembles

When we apply the k-mean cluster to the score plot, fewer clusters will result in regions where ensemble members within that region are far apart. For example, in Figure 4.1(a) and (b), in region six and seven, respectively, the cluster members could be split into two regions instead of one. This results in centroids that do not represent the spread of the PCA and its clusters very well, as the centroid is further away from ensemble members. The squared euclidean distance calculates the centroids from the mean of the points in that cluster (Mathworks), which makes the centroid point further from the ensemble. The Figure also shows that there are no clear individual functional forms that are clustered together, apart from the combination of linear phytoplankton with quadratic zooplankton mortalities, and trigonometric uptake with Holling type II grazing (please refer to Table 4.1 for functional forms associated with the ensemble numbers).

From Figure 4.2(a), the nRMSEs of chlorophyll and DIN from the full ensemble are often lower than the reduced ensemble, especially at non-oligotrophic stations. The average nRMSEs between chlorophyll

Table 4.1: Functional form combinations for Figure 4.1. The first and second letters show the phytoplankton and zooplankton mortalities respectively. Following the mortalities, nutrient uptake function is stated, and the zooplankton grazing function appears last. m h, l, q, and s denotes the hyperbolic, linear, quadratic, and sigmoidal mortality respectively. The default grazing is left blank and holling type II grazing is abbreviated to holII

1	hh default	33	hs default	65	ls default	97	qs default
2	hh default holII	34	hs default holII	66	ls default holII	98	qs default holII
3	hh exponential	35	hs exponential	67	ls exponential	99	qs exponential
4	hh exponential holII	36	hs exponential holII	68	ls exponential holII	100	qs exponential holII
5	hh sigmoidal	37	hs sigmoidal	69	ls sigmoidal	101	qs sigmoidal
6	hh sigmoidal holII	38	hs sigmoidal holII	70	ls sigmoidal holII	102	qs sigmoidal holII
7	hh trigonometric	39	hs trigonometric	71	ls trigonometric	103	qs trigonometric
8	hh trigonometric holII	40	hs trigonometric holII	72	ls trigonometric holII	104	qs trigonometric holII
9	hl default	41	lh default	73	qh default	105	sh default
10	hl default holII	42	lh default holII	74	qh default holII	106	sh default holII
11	hl exponential	43	lh exponential	75	qh exponential	107	sh exponential
12	hl exponential holII	44	lh exponential holII	76	qh exponential holII	108	sh exponential holII
13	hl sigmoidal	45	lh sigmoidal	77	qh sigmoidal	109	sh sigmoidal
14	hl sigmoidal holII	46	lh sigmoidal holII	78	qh sigmoidal holII	110	sh sigmoidal holII
15	hl trigonometric	47	lh trigonometric	79	qh trigonometric	111	sh trigonometric
16	hl trigonometric holII	48	lh trigonometric holII	80	qh trigonometric holII	112	sh trigonometric holII
17	hq default	49	ll default	81	ql default	113	sq default
18	hq default holII	50	ll default holII	82	ql default holII	114	sq default holII
19	hq exponential	51	ll exponential	83	ql exponential	115	sq exponential
20	hq exponential holII	52	ll exponential holII	84	ql exponential holII	116	sq exponential holII
21	hq sigmoidal	53	ll sigmoidal	85	ql sigmoidal	117	sq sigmoidal
22	hq sigmoidal holII	54	ll sigmoidal holII	86	ql sigmoidal holII	118	sq sigmoidal holII
23	hq trigonometric	55	ll trigonometric	87	ql trigonometric	119	sq trigonometric
24	hq trigonometric holII	56	ll trigonometric holII	88	ql trigonometric holII	120	sq trigonometric holII
25	sl default	57	lq default	89	qq default	121	ss default
26	sl default holII	58	lq default holII	90	qq default holII	122	ss default holII
27	sl exponential	59	lq exponential	91	qq exponential	123	ss exponential
28	sl exponential holII	60	lq exponential holII	92	qq exponential holII	124	ss exponential holII
29	sl sigmoidal	61	lq sigmoidal	93	qq sigmoidal	125	ss sigmoidal
30	sl sigmoidal holII	62	lq sigmoidal holII	94	qq sigmoidal holII	126	ss sigmoidal holII
31	sl trigonometric	63	lq trigonometric	95	qq trigonometric	127	ss trigonometric
32	sl trigonometric holII	64	lq trigonometric holII	96	qq trigonometric holII	128	ss trigonometric holII

and DIN from all the oceanographic region, Figure 4.2 (f) shows that 10, 7, and 6 ensemble members, has smaller nRMSEs compare to the full ensemble. This may be due to the low nRMSE in chlorophyll, compare to the larger reduced ensemble. Figure 4.2(c)–(e) show that reducing the ensemble may increase the nRMSEs.

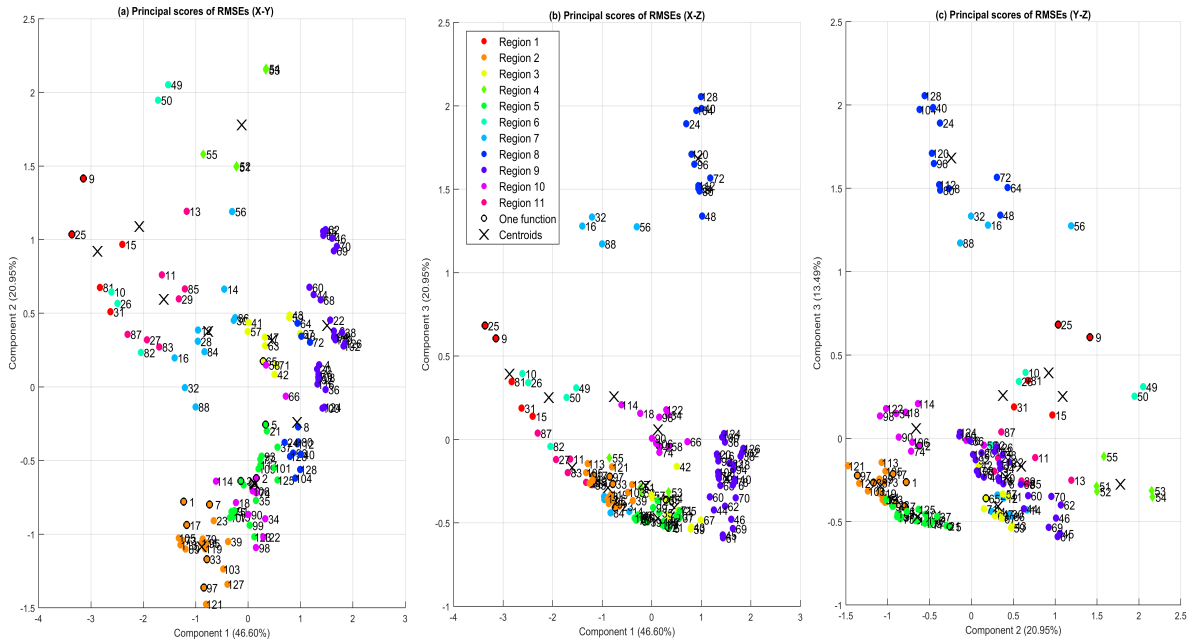


Figure 4.1: Principal component scores for different functional forms in three components, using 11 regions, and the centroids. 81% of the scores are available on the first three component, therefore the plot is separated into three planes, X-Y (a), X-Z (b), and Y-Z (c). The colour denotes the different regions, cross denotes the centroids, and the numbers denotes the ensemble members (please see Table 4.1 for functional forms combination and its corresponding numbers). Data points that has black border are ensemble members which only one process functional forms combination is changed whilst keeping other processes at its default function.

In terms of range, it is expected that the fewer the ensemble members, the lower the coverage. This have been observed for the chlorophyll at the oligotrophic stations, PAP, and L4, summarised in Table 4.2 and Figure 4.3. However, in station Cariaco, the range coverage is always above 90% both chlorophyll profile and surface.

4.3.2 The effect of replacing linear mortalities to the RMSEs and range

It is known that the linear closure term shows the most deviation compared to other functional forms (Yool et al., 2011). The function tends to produce unforced oscillation in simple model (Steele and Henderson, 1992; Edwards and Yool, 2000). In this subsection, ensemble members with plankton linear mortality that are the closest to the centroids are replaced with another ensemble member that does not contain linear mortality. If in a cluster all ensemble members contain linear mortality, then the

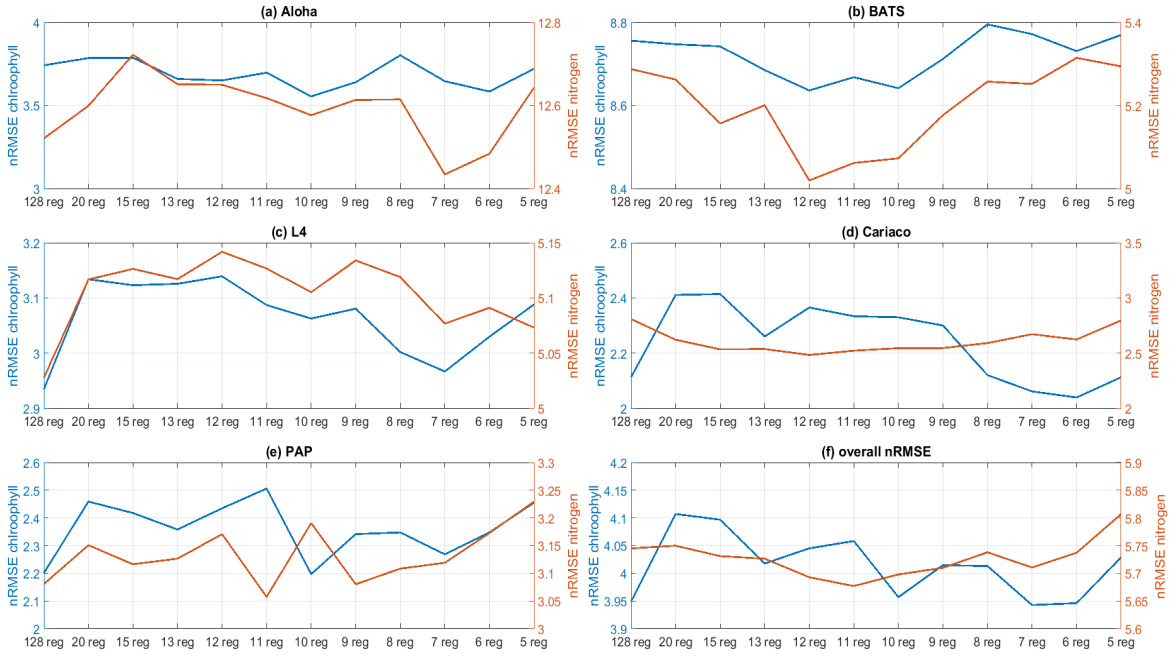


Figure 4.2: Mean nRMSEs from all the stations (a to e) and the averaged DIN (orange) and chlorophyll (blue) mean nRMSE from all the stations (f), from the full ensemble (128 members) to five members that are closest to the centroids.

Table 4.2: Percentage coverage of ensemble range for chlorophyll, DIN profile, and surface chlorophyll, from the full ensemble to 5 ensemble members. L4 does not have in situ chlorophyll profile, therefore excluded from the range calculation.

Stations	Range%										
	20 reg	15 reg	13 reg	12 reg	11 reg	10 reg	9 reg	8 reg	7 reg	6 reg	5 reg
Chlorophyll Profile											
Aloha	89.10	96.26	85.36	91.15	91.15	67.49	60.05	60.05	60.05	60.05	61.69
BATS	77.91	95.69	73.21	92.94	92.94	57.52	52.28	52.28	52.28	52.28	46.60
Cariaco	99.80	99.80	98.62	98.70	98.70	97.53	97.53	94.57	95.95	95.95	96.42
L4											
PAP	87.13	88.83	80.23	88.83	88.83	73.92	63.45	73.99	73.99	73.99	73.99
DIN Profile											
Aloha	84.22	77.42	70.80	62.53	61.89	62.28	61.89	60.79	62.28	41.44	46.61
BATS	91.55	99.95	86.89	97.56	97.56	80.45	80.45	76.39	76.39	76.39	65.67
Cariaco	84.17	83.94	83.71	89.25	89.25	88.78	89.25	89.48	87.35	87.35	69.57
L4	94.46	75.54	73.81	75.45	75.45	73.71	73.71	65.77	65.77	65.77	89.09
PAP	88.32	81.89	76.85	81.89	81.89	75.08	75.86	80.63	80.63	80.63	80.63
Surface Chlorophyll											
Aloha	87.75	92.78	87.05	91.48	91.48	81.04	59.98	61.98	61.98	61.98	58.25
BATS	78.06	97.94	75.26	94.85	94.85	60.68	53.31	53.31	53.31	53.31	46.40
Cariaco	98.47	98.47	98.53	98.64	98.64	98.70	98.70	94.30	94.82	94.82	92.61
L4	95.48	94.48	96.56	92.80	92.80	94.88	94.88	59.57	59.57	59.57	63.99
PAP	88.93	87.54	76.34	83.42	83.42	67.57	67.39	78.55	78.55	78.55	84.54

representative member is kept.

Although the lowest overall nRMSE is shown in 10, 7, and 6 members, shown in Figure 4.2, these members show low range coverage, summarised on Table 4.2. Therefore, in this subsection, we use 15, 11, 10, and 7 members in order to encompass the various ranges and RMSEs. From Figure 4.3(a), it is

often that the range coverage from the replaced linear mortalities (hereafter RLM) ensembles show little difference compared to the original ensemble members. The nRMSEs from RLM ensembles are slightly lower than the reduced ensembles. The biggest difference is observed at 7 ensemble members, whereby the nRMSE reduced from 4.83 to 4.80. Shown in Figure 4.4.

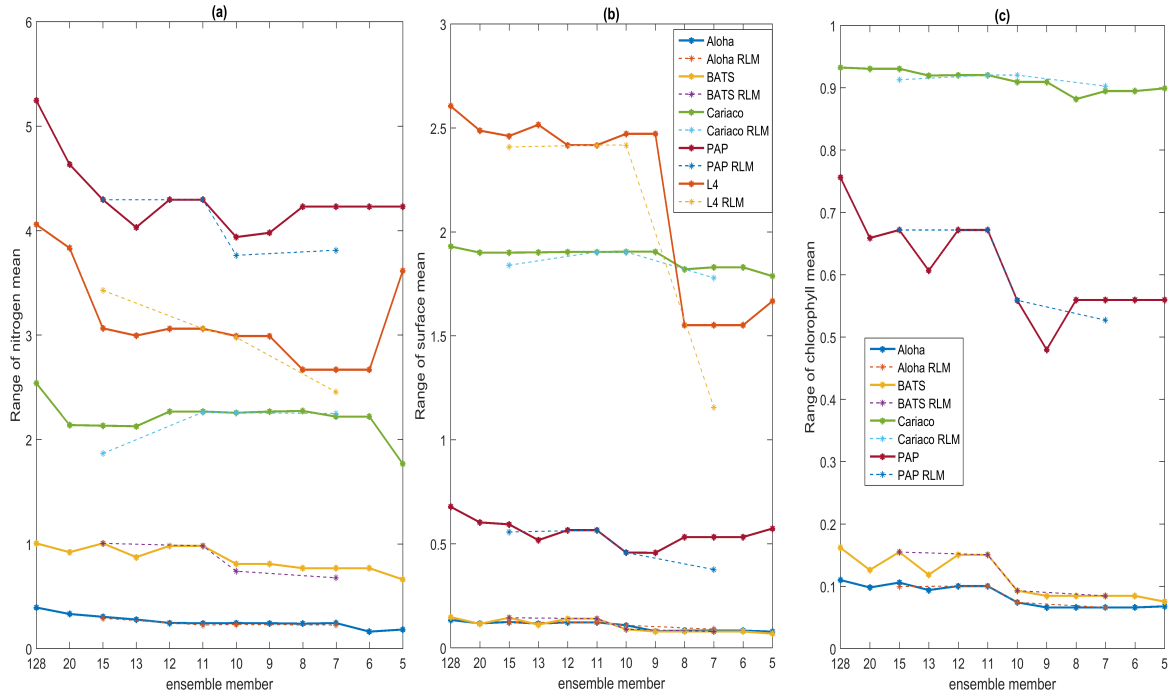


Figure 4.3: Range of DIN profile (a), surface chlorophyll (b), and profile (c) means from the original ensemble members to five ensemble members. The dashed line denotes replacing ensemble members with linear with ensemble members with non-linear mortalities on 15, 11, 10, and 7 ensemble members.

Replacing the linear mortalities lower the chlorophyll nRMSE at most stations, apart from station L4, summarised in Figure 4.5(c). At some stations, replacing the linear mortalities do not affect the nRMSE. For example, at BATS and Cariaco, using 11 and 10 members, respectively, replacing the linear mortality produce the same nRMSEs. However, RLM ensembles do not always produce lower DIN nRMSE. For example, at BATS and Cariaco, higher nRMSE is produced from RLM ensemble below 15, Figure 4.6(b) and (d). RLM ensembles produce a slightly lower range for DIN profile, apart from the 15RLM and 7RLM. At most of the stations, the smaller the ensemble members, the more significant the difference of DIN range between the two ensembles are, summarised in Table 4.3.

Replacing the linear mortality reduces the range at most stations because this functional form deviates the most from the default mortality. However, at L4 15, 11, and 7 RLM resulted in a higher range (from 75% to 82%, see Table 4.3), compared to retaining linear mortality. At PAP, the ranges of RLM ensembles are higher in 15 and 7 ensembles, compared to other ensembles.

For the chlorophyll profile, at most of the stations, there is no significant difference between the ranges

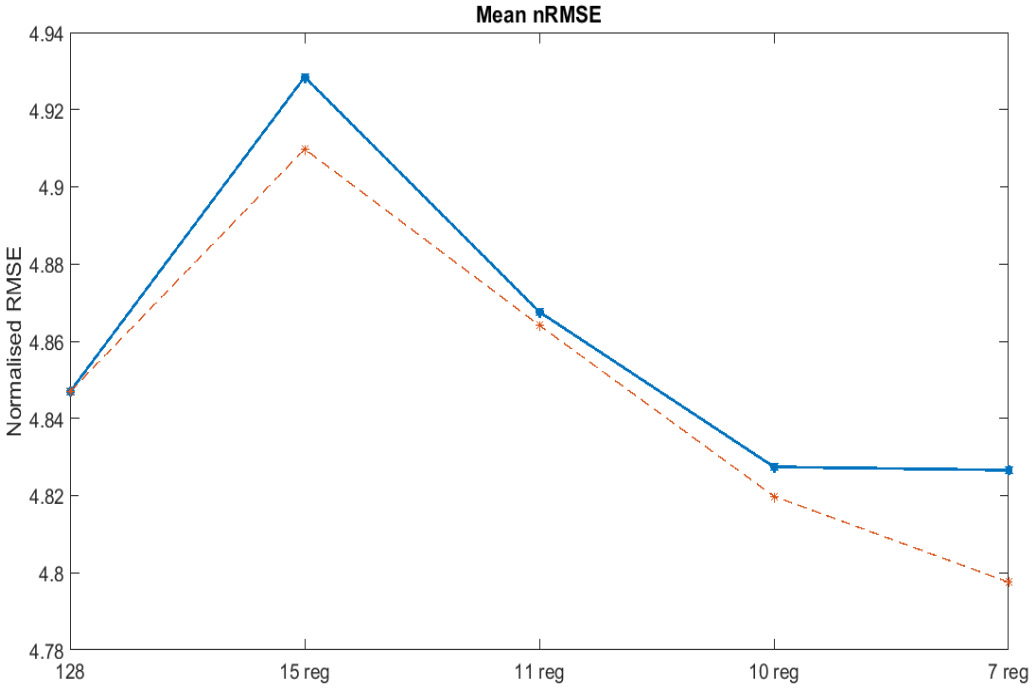


Figure 4.4: Mean nRMSE from all the oceanographic stations. The dashed line denotes reduced ensemble with replaced linear mortality on 15, 11, 10, and 7 ensemble members.

of RLM and reduced ensembles, summarised in Table 4.3. Replacing the linear mortalities have reduced the chlorophyll nRMSEs, both the surface and profile. RLM ensembles produce lower range coverages apart from, 11 RLM which shows no difference. The nRMSEs of chlorophyll profile have been improved when linear mortalities are replaced. However, the range coverage, especially in 15 ensemble members, is reduced.

From these results, when the ensemble is reduced to below 11 members, the range coverages can fall below 70% (Table 4.2) and replacing linear mortalities reduce the nRMSEs, summarised in Figure 4.4 and Table 4.3. The range coverages from RLM ensembles are also lower than the reduced ensembles. RLM ensemble that mostly retains the range is 11 members, Table 4.3. By using 11 RLM, it is possible to have high range coverage and also low nRMSEs.

4.3.3 Other statistical metrics in the reduced ensembles

From the previous subsection, 11 members ensemble shows a good range coverage, and lower nRMSE for chlorophyll even when the linear mortality is replaced. Here we will explore other statistical metrics, apart from range and RMSEs in different regions. In this subsection, along with the 11 members ensemble, we will use the ensemble means of ensembles which show high coverage (15 members) and low nRMSEs (7 regions).

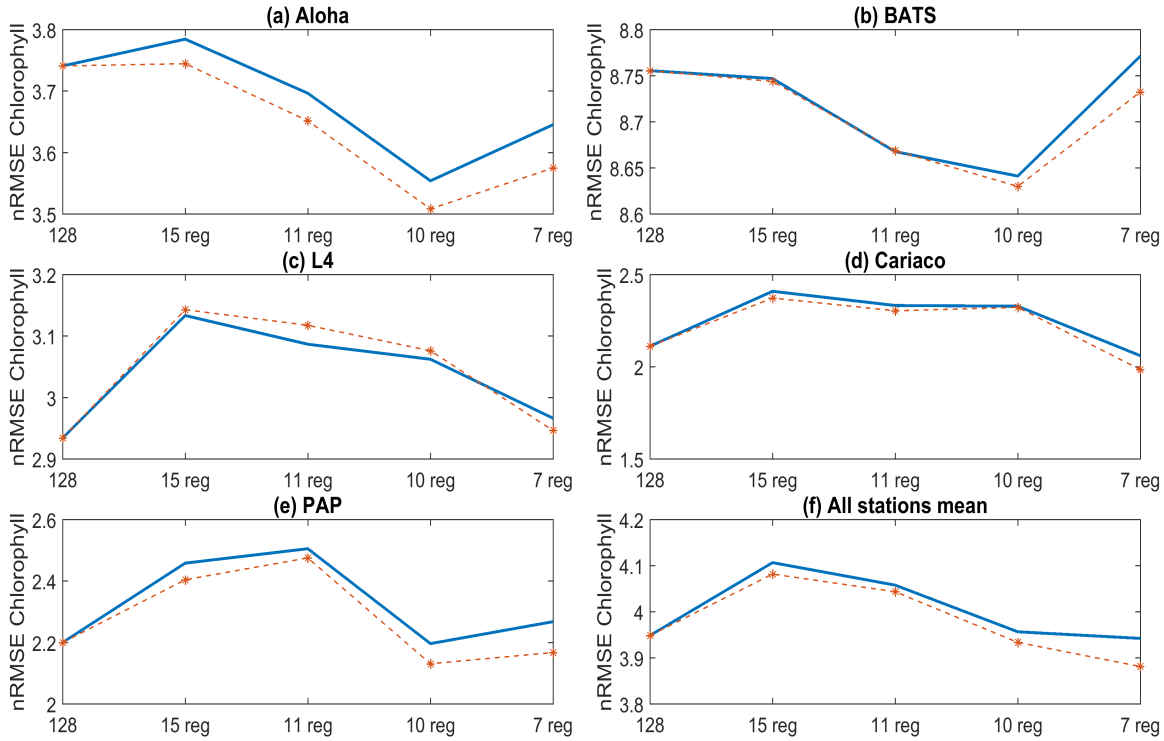


Figure 4.5: Mean nRMSE for chlorophyll from all the oceanographic stations (a to e). Subfigure (f) shows the mean chlorophyll nRMSE from all stations. The orange dashed line denotes reduced ensemble with replaced linear mortality on 15, 11, 10, and 7 ensemble members.

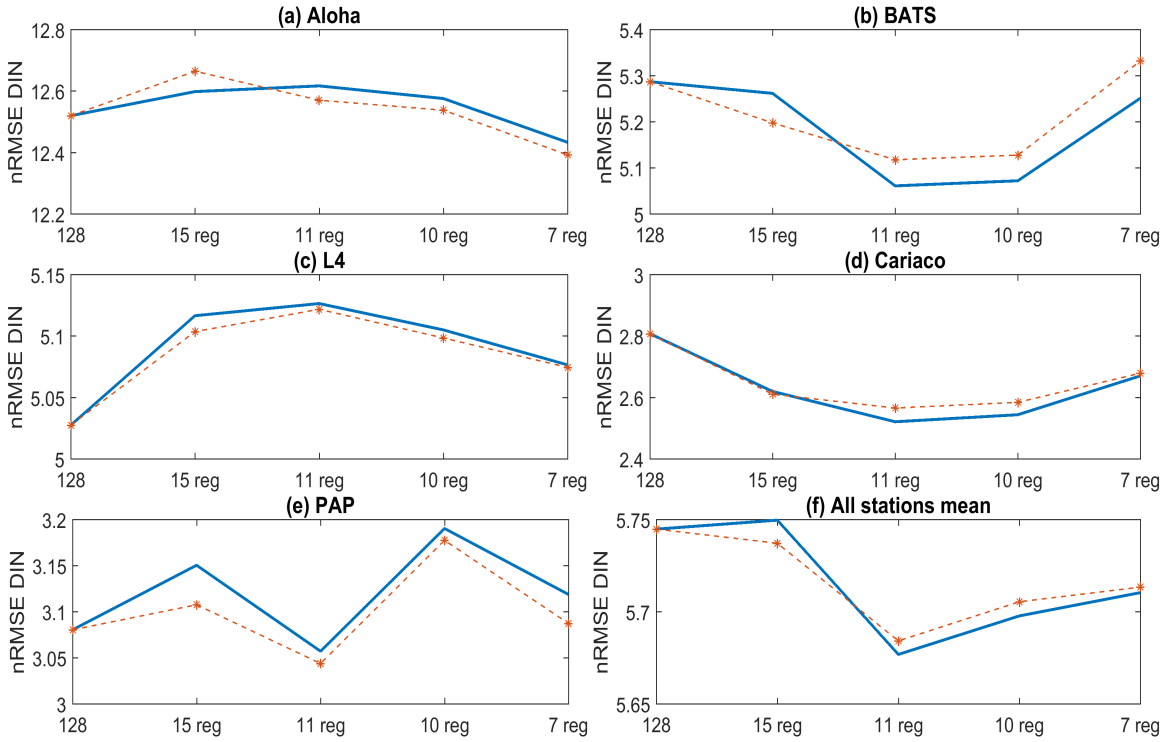


Figure 4.6: Mean nRMSE for DIN from all the oceanographic stations (a to e). Subfigure (f) shows the mean DIN nRMSE from all stations. The orange dashed line denotes reduced ensemble with replaced linear mortality on 15, 11, 10, and 7 ensemble members.

Table 4.3: Percentage of range coverage between original ensemble members that are closest to the centroids and those which other mortalities are used instead of linear mortalities.

Stations	Range%							
	15 RLM	15 reg	11 RLM	11 reg	10 RLM	10 reg	7 RLM	7 reg
Chlorophyll Profile								
Aloha	90.55	96.26	91.15	91.15	67.49	67.49	60.05	60.05
BATS	95.65	95.69	92.94	92.94	57.45	57.52	52.28	52.28
Cariaco	97.90	99.80	98.70	98.70	98.70	97.53	95.95	96.80
L4								
PAP	88.83	88.83	88.83	88.83	73.92	73.92	73.99	69.70
DIN Profile								
Aloha	73.93	77.42	57.37	61.89	58.83	62.28	62.28	57.37
BATS	99.96	99.95	97.56	97.56	73.55	80.45	76.39	67.24
Cariaco	73.48	83.94	89.02	89.25	88.78	88.78	87.35	88.53
L4	81.89	75.54	81.89	75.45	71.76	73.71	65.77	60.51
PAP	84.49	81.89	75.45	81.89	73.41	75.08	80.63	72.68
Surface Chlorophyll								
Aloha	90.74	92.78	91.48	91.48	81.04	81.04	61.98	67.04
BATS	97.87	97.94	94.85	94.85	60.68	60.68	53.31	53.31
Cariaco	95.33	98.47	98.64	98.64	98.64	98.70	94.82	92.18
L4	92.46	94.48	92.80	92.80	92.80	94.88	59.57	44.38
PAP	82.13	87.54	83.42	83.42	67.57	67.57	78.55	55.53

At PAP, correlations from the reduced ensemble means are either similar to the full ensemble (DIN), or worse (surface and profile chlorophyll). From Table 4.4, using 7, 7RLM or 11RLM ensembles produce better or similar bias as the full ensemble in terms of chlorophyll profile and surface, and 7, 11, or 11RLM for DIN at station PAP. None of the reduced ensembles shows better DIN biases compared to the default run.

At ALOHA, the full and reduced ensemble means for DIN correlations do not differ much. At BATS, the correlations from the full ensemble are improved. Most of the reduced ensemble produce higher RMSEs for DIN compare to the full ensemble at the oligotrophic stations. However, using 7 RLM at ALOHA and 11 members at BATS produce better RMSEs compare to the full ensemble. Chlorophyll profile correlation has also been improved by the reduced ensemble at both stations, although the default run still shows a better match. At BATS, the ensemble means from 11, 11 RLM, 15, and 15 RLM produce similar bias as the full ensemble. However, at ALOHA, most of the ensembles produce similar bias as the full ensemble. For the surface chlorophyll, all the reduced ensemble, apart from 7 members, produce better or similar correlations and biases compared to the full ensemble at both stations, summarised in Figure 4.4. These further shows that reducing the ensemble can still retain and even improve the bias and correlations of the ensemble mean. Therefore at the oligotrophic stations, better correlations, lower RMSE, and bias are produced when using 11, 11 RLM, and 7 ensemble members.

At the coastal stations, the correlation between in situ DIN and the ensemble means are always above

70%. At Cariaco, the reduced ensemble shows higher (15 RLM and 11 RLM) or similar DIN correlations compare to the full ensemble means. At L4, the DIN correlation from 11 members is better than the full ensemble but lower than the default run, summarised in Table 4.4. At Cariaco, the reduced ensembles produce worse biases than the full ensemble for both DIN and chlorophyll. However, at L4, all the reduced ensembles produce better biases than the full ensemble, especially 11 members for both DIN and chlorophyll.

Most of the reduced ensembles presented here produce better or similar chlorophyll profile correlation and bias as the full ensemble at both stations. However, reduced ensembles that show higher correlations than the full ensemble produce higher biases and RMSEs than the full ensemble, Table 4.4. From these calculations 7, 11, and 11 RLM for surface chlorophyll produce better statistical metrics.

4.3.4 Chosen ensemble members

From the results in the previous sections, from Table 4.4 the ensemble means from 11 RLM shows better RMSEs and bias. This ensemble also mostly retains the range from the full ensemble. We therefore chose 11 RLM ensemble members to represent the full ensemble. The functional form combinations are as follows:

1. hh default holII ($U_h\ G_2\ \rho_h\zeta_h$)
2. hh trigonometric holII ($U_t\ G_2\ \rho_h\zeta_h$)
3. hq sigmoidal holII ($U_s\ G_2\ \rho_h\zeta_q$)
4. ll exponention holII ($U_e\ G_2\ \rho_l\zeta_l$)
5. ls trigonometric ($U_t\ G_1\ \rho_l\zeta_s$)
6. ql hyperbolic ($U_h\ G_1\ \rho_q\zeta_l$)
7. sl hyperbolic holII ($U_h\ G_2\ \rho_s\zeta_l$)
8. sl exponential holII ($U_e\ G_2\ \rho_s\zeta_l$)
9. sl sigmoidal ($U_s\ G_1\ \rho_s\zeta_l$)
10. sq sigmoidal ($U_s\ G_1\ \rho_s\zeta_q$)
11. sq trigonometric holII($U_t\ G_2\ \rho_s\zeta_q$)

Table 4.4: Error statistic table for 10 to 13 ensemble members, with and without linear mortality included in the ensemble members. RLM denotes ensembles without linear mortalities. Highest correlation, lowest RMSE, and bias are printed in bold.

Station		DIN Profile								Chlorophyll Profile								Surface Chlorophyll							
		Ensemble mean							Def	Ensemble mean							Def	Ensemble mean							Def
		15	15RLM	11	11 RLM	7	7RLM	Full		15	15 RLM	11	11 RLM	7	7RLM	Full		15	15RLM	11	11 RLM	7	7RLM	Full	
PAP	r	0.23	0.23	0.23	0.23	0.23	0.22	0.23	0.21	0.41	0.42	0.40	0.40	0.40	0.36	0.42	0.28	0.38	0.38	0.38	0.39	0.38	0.37	0.45	0.36
	RMSE	3.26	3.27	3.23	3.23	3.37	3.22	3.26	3.32	0.32	0.31	0.32	0.31	0.31	0.32	0.32	0.40	0.51	0.51	0.54	0.48	0.48	0.47	0.51	0.57
	Bias	-0.72	-0.74	-0.68	-0.71	-0.98	-0.71	0.61	0.20	-0.07	-0.07	-0.08	-0.07	-0.05	-0.09	0.06	-0.18	-0.23	-0.23	-0.28	-0.22	-0.19	0.22	0.22	-0.30
	Mean	6.56	6.57	6.11	6.54	6.83	6.55	6.59	5.64	0.49	0.50	0.51	0.49	0.46	0.51	0.48	0.00	0.67	0.67	0.68	0.66	0.64	0.66	0.66	0.74
ALOHA	r	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.24	0.24	0.24	0.26	0.28	0.25	0.22	0.29	0.22	0.22	0.24	0.23	0.23	0.22	0.22	0.27
	RMSE	1.08	1.07	1.07	1.06	1.07	1.05	1.06	1.09	0.10	0.96	0.10	0.10	0.09	0.97	0.10	0.10	0.05	0.05	0.05	0.04	0.05	0.04	0.05	0.07
	Bias	-0.68	-0.68	-0.67	-0.68	-0.68	0.66	0.67	-0.61	0.06	0.06	0.06	0.06	0.05	0.06	-0.06	-0.03	0.01	0.01	0.00	0.01	0.00	0.01	0.01	0.03
	Mean	1.18	1.18	1.17	1.18	1.18	1.16	1.20	1.10	0.06	0.06	0.07	0.06	0.07	0.06	0.06	0.03	0.08	0.08	0.08	0.08	0.09	0.08	0.10	0.11
BATS	r	0.56	0.56	0.57	0.56	0.55	0.56	0.56	0.59	0.22	0.2	0.20	0.22	0.15	0.20	0.19	0.23	0.25	0.25	0.30	0.28	0.13	0.22	0.22	0.28
	RMSE	1.40	1.42	1.29	1.42	1.43	1.43	1.39	1.08	0.32	0.33	0.32	0.32	0.33	0.33	0.33	0.31	0.33	0.33	0.33	0.33	0.34	0.33	0.33	0.31
	Bias	1.17	-1.19	-1.05	-1.19	1.20	-1.19	1.16	-0.73	0.11	0.12	0.11	0.12	0.13	0.13	0.12	0.07	0.12	0.12	0.11	0.12	0.13	0.12	0.12	0.07
	Mean	1.78	1.81	1.66	1.81	1.81	1.80	1.77	1.35	0.06	0.05	0.07	0.05	0.04	0.05	0.05	0.10	0.05	0.05	0.06	0.04	0.04	0.05	0.05	0.09
Cariaco	r	0.78	0.79	0.78	0.79	0.78	0.78	0.78	0.76	0.29	0.28	0.32	0.29	0.29	0.28	0.29	0.22	0.13	0.17	0.16	0.13	0.13	0.12	0.13	0.11
	RMSE	3.00	2.98	2.99	2.96	3.00	2.97	2.97	3.29	0.84	0.84	0.88	0.84	0.85	0.84	0.83	0.87	1.24	1.29	1.30	1.23	1.25	1.24	1.23	1.27
	Bias	0.96	0.93	-0.96	-0.91	-0.70	0.72	0.61	-0.59	0.00	0.02	-0.15	0.02	-0.02	0.01	0.02	0.09	0.06	-0.38	-0.38	0.01	0.10	0.03	0.02	0.18
	Mean	5.74	5.71	5.73	5.68	5.75	5.49	5.93	5.37	0.51	0.51	0.66	0.49	0.53	0.50	0.49	0.42	0.82	1.14	1.14	0.77	0.86	0.78	0.77	0.58
L4	r	0.70	0.70	0.71	0.70	0.70	0.70	0.70	0.72									0.26	0.25	0.30	0.25	0.24	0.24	0.25	0.31
	RMSE	2.83	2.87	2.58	2.87	2.87	2.87	2.94	2.67									1.04	1.03	1.19	1.04	1.03	1.01	1.05	1.28
	Bias	-1.41	-1.46	-0.93	-1.46	-1.45	-1.45	1.56	-1.12									-0.43	-0.40	-0.77	-0.41	-0.34	-0.31	0.42	-0.83
	Mean	4.364	4.4149	3.88	4.41	4.41	4.41	4.52	4.08									1.77	1.74	2.11	1.75	1.68	1.65	1.76	2.17

4.4 Generating the noise

As stated in the previous chapters, the marine biogeochemical model is also sensitive to the physical oceanography inputs, especially the mixed layer depth and vertical velocity. The latter dictates the availability of nutrients in the upper water column ($>$ 200m), which is often poorly known. As shown in Chapter 2, spurious vertical velocity has shown to alter the DIN distribution in the water column. This affects the chlorophyll concentration. Thus, we chose to vary the vertical velocity input of the 1-D model by adding random noise with different strength.

The vertical velocity is obtained from 5-day averaged NEMO-FOAM (Storkey et al., 2010), which has been capped (see chapter 2 section 2.5.1 for further details) to minimise overestimation in nutrients. This is a common problem when directly assimilating physical data into an ocean biogeochemical model (Raghukumar et al., 2015; Ford and Barciela, 2017). The vertical velocity noise (N_{v_d}) is generated by multiplying random noise ($Rand_s$) with the monthly anomaly. The anomaly is calculated by subtracting the 5-day mean vertical velocity (v_d, where d is the date-month-year), from the monthly mean ($\bar{m}$), and then multiplied by a random number ($Rand_s$) generated by python at each depth level. This can be summarised as:

$$N_{v_d} = (\bar{m} - v_d) \times Rand_s \tag{4.2}$$

This will results in 6 perturbed vertical velocity profiles for each month. The random numbers that we use here define the perturbation strength, which can be categorised into, low (random numbers generated between -0.5 and 0.5), medium (between -1 and 1), and high (between -2 and 2). The difference in perturbation strength is applied to examine how these noises can affect the model dynamics. In order to add this noise to the 5-day averaged vertical velocity, we randomise the assignment of the noise to a certain 5-day averaged vertical velocity profile to make the overall noisy vertical velocity more variable. For example, if N_{v_d} is generated from the v_{140199} then this may be added to v_{150600}. Therefore the new noisy vertical velocity (v_{N_d}) for a certain date could be summarised as:

$$v_{N_d} = Rand(N_{v_d}) + v_d \tag{4.3}$$

where $Rand(N_{v_d})$ is the randomise noise that has been assigned to a certain vertical velocity. This is generated again 11 times, both for the noises and the vertical velocity assignments, which forms 11 ensemble members, where each member has different vertical velocity. The ensemble is run using the default MEDUSA functional forms for the perturbed physics. For perturbing both biogeochemistry and

physics, the perturbed biogeochemistry ensemble discussed in 4.2 is run using the perturbed vertical velocity generated using equation 4.3. An example of the vertical velocity profile can be seen in Figure 4.7

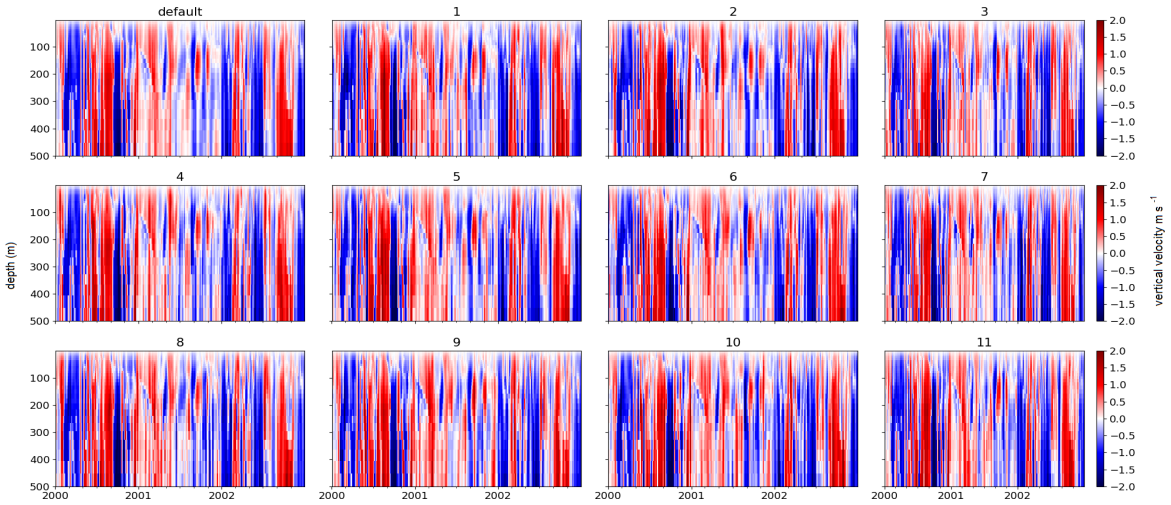


Figure 4.7: Vertical velocity profiles (m s^{-1}) from January 2000 to December 2002 at station ALOHA for the perturbed physics ensemble when random number between -1 and 1 is used. The title of each sub-figure represents the default run and ensemble members.

4.5 Error statistics from different perturbation strength

In this section we will compare the spread of interannual mean of chlorophyll concentrations from perturbing the (i) functional forms (perturbed biogeochemistry ensemble, PBE), (ii) vertical velocities by adding noises (perturbed physics ensemble, PPE), and (iii) both the functional forms and vertical velocities (perturbed biogeochemistry and physics ensemble, PBPE). The different perturbation strengths of the vertical velocity are shortened to PPE-low, PPE-medium, and PPE-high. Similar perturbation strength abbreviations are also applied for the PBPEs. We will also compare the new ensemble mean and median with the default run and observations. For completeness, in this section, we include the default run to all of the ensembles, including the reduced PBE, making all the ensembles having 12 members instead of 11. The abyssal plain is described in section 4.5.1, oligotrophic region in 4.5.2, and coastal stations are described in 4.5.3.

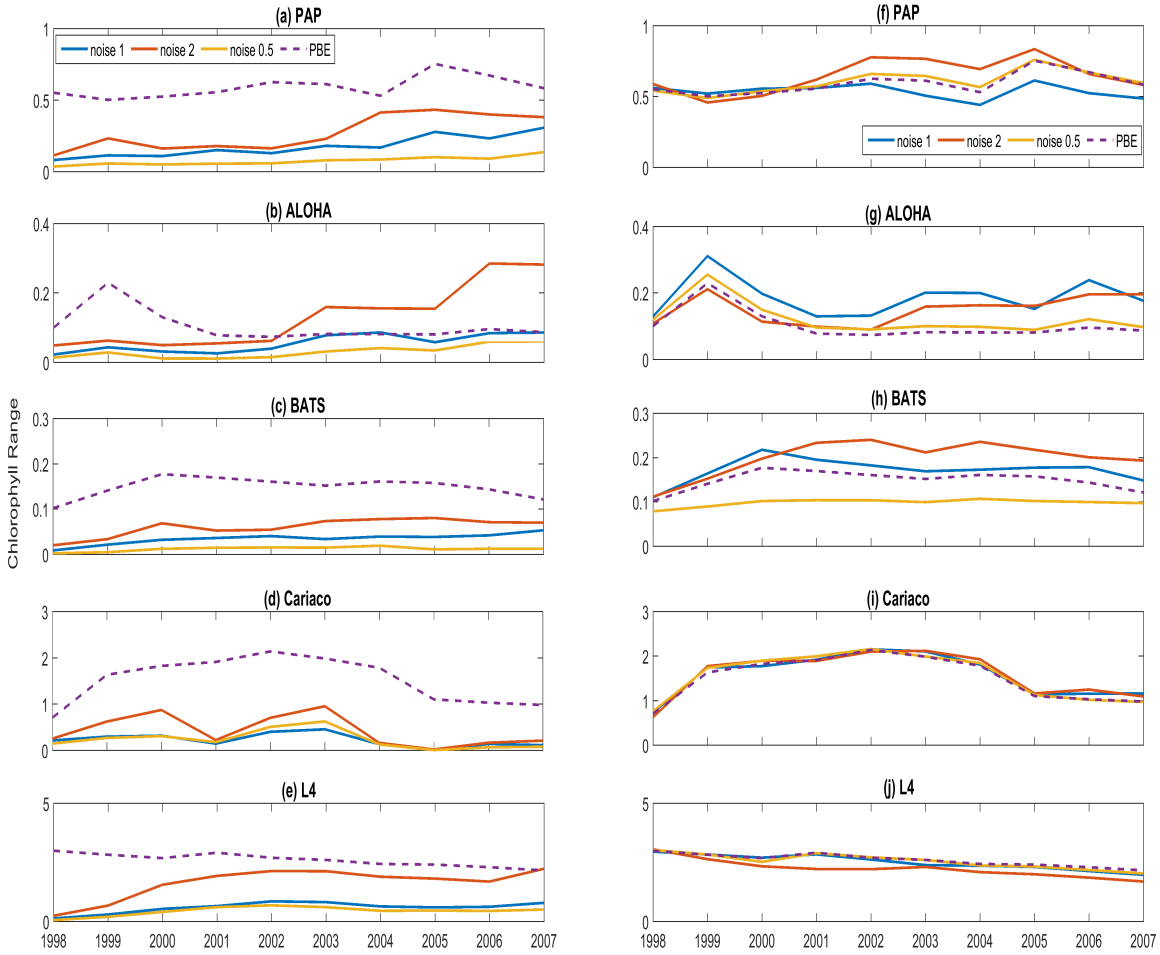


Figure 4.8: Range of surface chlorophyll interannual mean when perturbing the biological equations (a-e) and adding noise to vertical velocities (f-j) at five different oceanographic stations. Perturbing only the functional forms are shown in purple, adding noise between $\pm$ 0.5, 1, and 2 are shown in yellow, orange, and blue respectively. Y-axis shows the year the ensembles are run

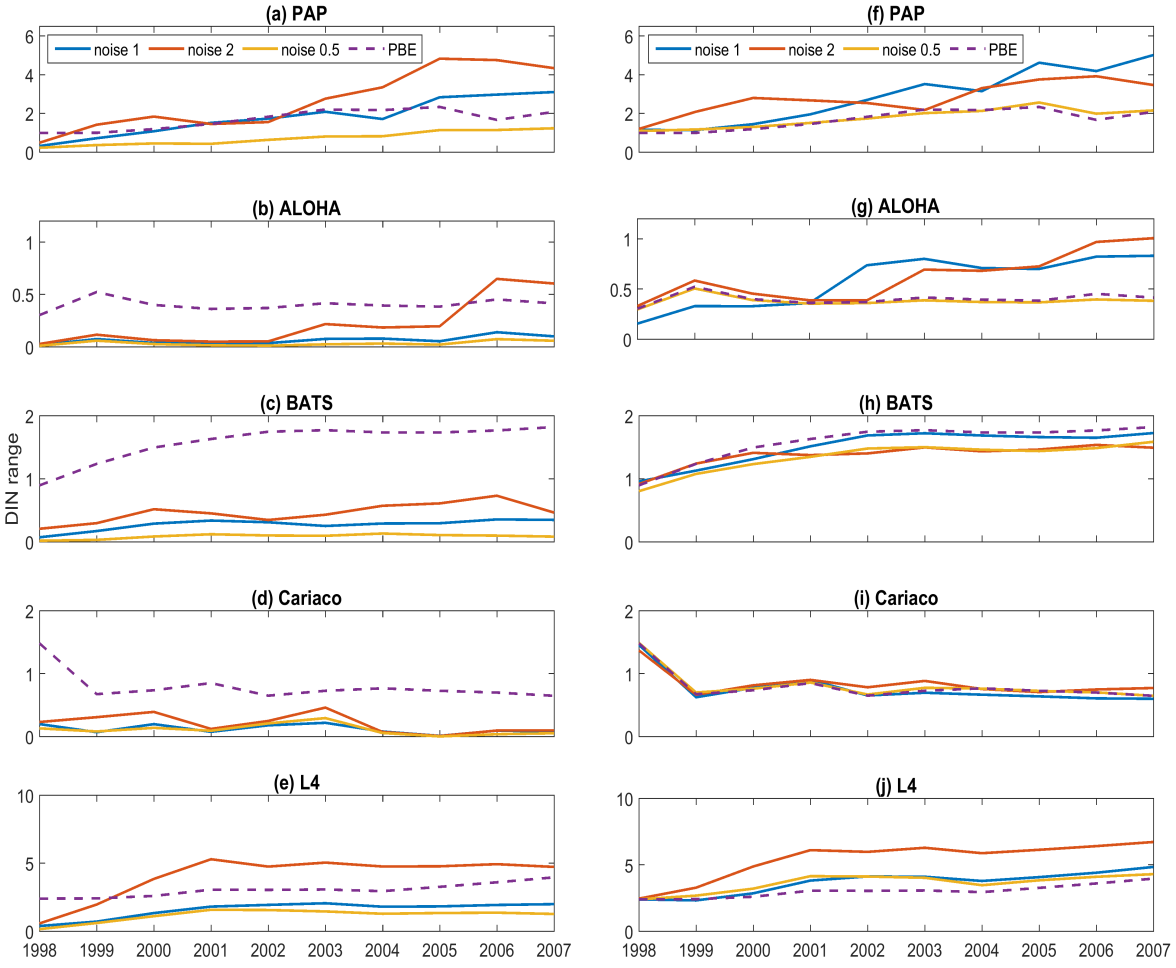


Figure 4.9: Range of surface DIN interannual mean when perturbing the biological equations (a-e) and adding noise to vertical velocities (f-j) at five different oceanographic stations. Perturbing only the functional forms are shown in purple, adding noise between $\pm$ 0.5, 1, and 2 are shown in yellow, orange, and blue respectively. Y-axis shows the year the ensembles are run

Table 4.5: Statistical metrics comparing the default run and the mean and median of the original perturbed biogeochemistry ensemble (reduced to 12 members), only adding to the vertical velocities (PPE), and both perturbing functional forms and adding noises (PBPE) with the observation from PAP station. The noises that are used are random numbers between $\pm$ 0.5, $\pm$ 1, and $\pm$ 2). The numbers that are in bold are the lowest bias and RMSE, and the highest correlation.

		chlorophyll profile (mg m^{-3})				DIN profile (mmol m^{-3})				surface chlorophyll (mg m^{-3})			
Station		median	mean	default	in situ	median	mean	default	in situ	median	mean	default	in situ
PAP	r	0.297	**0.395**	0.279		0.227	0.230	0.206		0.337	0.345	0.363	
	RMSE	0.382	**0.320**	0.404		3.217	3.231	3.323		0.539	0.565	0.566	
	Bias	-0.152	**-0.079**	-0.177		-0.248	-0.682	0.202		-0.270	-0.274	-0.300	
	mean	0.560	0.556	0.594	0.423	6.138	6.005	5.639	5.815	0.715	0.718	0.745	0.447
	Range	0.617				1.995				0.567			
PBPE	r	0.308	0.340	0.279		0.223	0.227	0.206		0.337	0.345	0.363	
noise= 0.5	RMSE	0.362	0.370	0.404		3.197	3.176	3.323		0.539	0.565	0.566	
	Bias	-0.139	-0.132	-0.177		-0.307	-0.208	0.202		-0.270	-0.274	-0.300	
	mean	0.552	0.556	0.594	0.423	6.175	6.145	5.639	5.815	0.714	0.721	0.745	0.447
	Range	0.637				1.960				0.580			
	r	0.272	0.273	0.279		0.208	0.209	0.206		0.369	0.366	0.363	
PPE	RMSE	0.403	0.403	0.404		3.301	3.335	3.323		0.565	0.566	0.566	
	Bias	-0.172	-0.172	-0.177		0.145	0.071	0.202		-0.301	-0.302	-0.300	
	mean	0.595	0.595	0.594	0.423	5.670	5.745	5.639	5.815	0.746	0.747	0.745	0.447
	Range	0.052				1.101				0.069			
PBPE	r	0.303	0.331	0.279		0.230	**0.236**	0.206		0.354	0.351	0.363	
noise= 1	RMSE	0.356	0.362	0.404		3.119	**3.075**	3.323		0.524	0.554	0.566	
	Bias	-0.127	-0.125	-0.177		0.125	**-0.006**	0.202		-0.257	-0.265	-0.300	
	mean	0.547	0.545	0.594	0.423	5.698	5.829	5.639	5.815	0.702	0.709	0.745	0.447
	Range	0.540				**4.085**				0.497			
PPE	r	0.272	0.273	0.279		0.208	0.209	0.206		0.369	0.366	0.363	
	RMSE	0.403	0.403	0.404		3.301	3.335	3.323		0.565	0.566	0.566	
	Bias	-0.172	-0.172	-0.177		0.145	0.071	0.202		-0.301	-0.302	-0.300	
	mean	0.585	0.581	0.594	0.423	5.410	5.450	5.639	5.815	0.730	0.736	0.745	0.447
	Range	0.096				2.737				0.138			
PBPE	r	0.318	0.325	0.279		0.226	0.221	0.206		0.365	0.365	0.363	
noise=2	RMSE	0.340	0.356	0.404		3.220	3.226	3.323		**0.512**	0.550	0.566	
	Bias	-0.117	-0.116	-0.177		-0.161	-0.260	0.202		**-0.250**	-0.268	-0.300	
	mean	0.536	0.535	0.594	0.423	5.982	6.081	5.639	5.815	0.695	0.713	0.745	0.447
	Range	**0.632**				2.825				**0.616**			
PPE	r	0.266	0.273	0.279		0.209	0.206	0.206		0.374	**0.385**	0.363	
	RMSE	0.391	0.385	0.404		3.284	3.316	3.323		0.552	0.550	0.566	
	Bias	-0.153	-0.147	-0.177		0.161	0.120	0.202		-0.290	-0.293	-0.300	
	mean	0.575	0.568	0.594	0.423	5.647	5.688	5.639	5.815	0.735	0.738	0.745	0.447
	Range	0.225				3.657				0.228			

4.5.1 Abyssal Plain

Statistical metrics

At PAP, after the addition of default run into the ensemble, the PBE mean and median shows higher correlations, lower bias and RMSEs for surface and chlorophyll profile compared to the default run. For chlorophyll profile, the PPE shows lower bias and RMSE compared to only running the default form using unperturbed vertical velocities, and the bias reduces further as the noise increases. PPE-high produces ensemble mean with the highest correlation, although the RMSE and bias are also high. However, chlorophyll ranges from PPEs, regardless of the noise, are always lower than the PBE.

PBPE medium and high produce higher ranges than PBE. The highest range for both chlorophyll surface and profile is produced by PBPE-high. In terms of the surface chlorophyll, PBPE-high also produces the

lowest bias and RMSE, for both the ensemble mean and median, summarised in Table 4.5. PBPE-high also produces the highest range of surface chlorophyll. Although the DIN RMSEs from the ensemble mean and median from the reduced PBE are lower than the default run, the biases are still higher than the default. When the vertical velocity is perturbed, in PBPE-medium, the lowest DIN bias and RMSE are produced.

Interannual mean range of DIN and chlorophyll

From Figure 4.9(a), a sharp increase in the interannual mean range have been observed at PAP. Although there is no significant increase in interannual DIN concentration from the default run, some of the PPE members show a significant sharp increase in interannual DIN concentration (e.g. in PPE-high, r =0.89, $p < 0.05$) and a significant decrease over the years ($r = -0.81, p < 0.05$). This resulted in the strong increase in interannual mean range of DIN, especially in PPE-medium and high ($r > 0.93$, $p < 0.05$), shown in Figure 4.10. The figure also shows that the higher the noise, the larger the interannual spread it garners, Figure 4.9(a). Compared to the PBE range, the PPE-high and medium produce a higher range, especially after 1999. The surface chlorophyll also follows the increasing trend of surface DIN ($r > 0.85$, $p < 0.05$ for all the ensembles), with the PPE-high garnering the largest range.

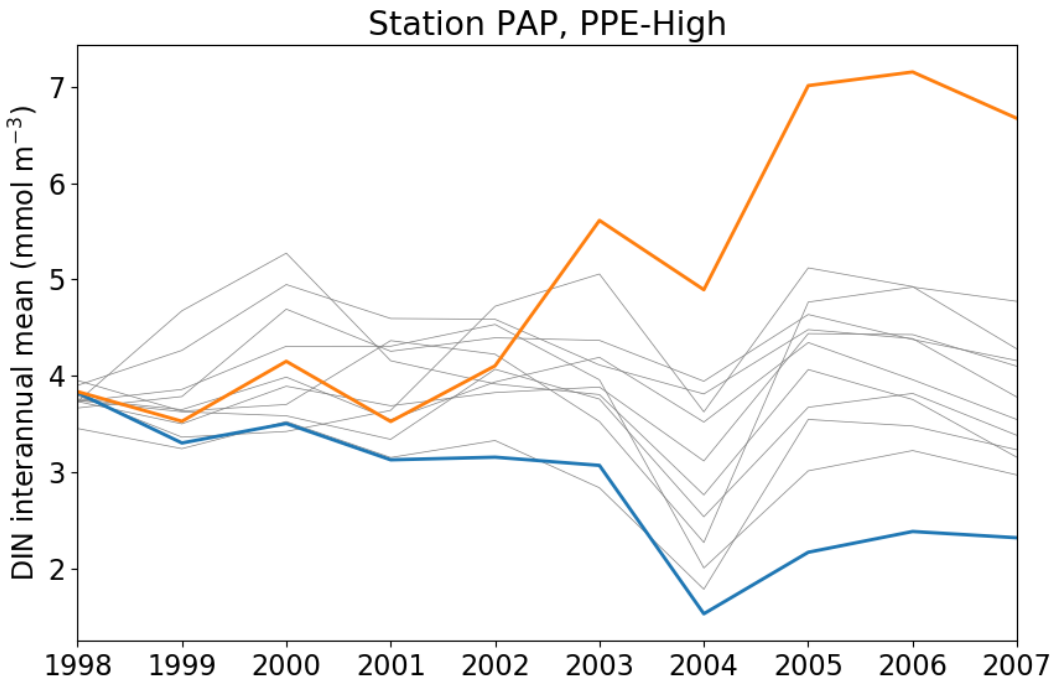


Figure 4.10: DIN interannual mean from perturbing the vertical velocity using the highest perturbation strength (between -2 to 2), calculated from January 1998 to December 2007. The grey lines show ensemble members that produce insignificant trend, whereas the orange and blue line show ensemble members that produce significant increase and decrease respectively. Y-axis shows the year the ensembles are run

The lowest chlorophyll range for all the perturbed physics ensembles have been produced on the first year, 1998, for chlorophyll. The ensemble members from PPE-high can produce up to 12.9% higher and 16.5% lower annual mean chlorophyll than the default run. However, the chlorophyll range of the PPE ensembles is still lower than the original reduced ensemble.

Similar to the PPEs, a significant increase in annual DIN mean over the years in some of the PBE members ($r > 0.47, p < 0.05$), however, there is no significant decrease. This makes the PBE range also increases ($r = 0.81,\ p > 0.05$). Now that both PBE and PPE show an increase in DIN range and concentration, the PBPE ranges also show an increase over the years, regardless of the perturbation strength ($r > 0.86,\ p < 0.05$). However, the highest overall DIN range is garnered from PBPE-medium (2.90 mmol m^{-3}) instead of PBPE-high (2.80 mmol m^{-3}), Figure 4.9(f).

The interannual mean of surface chlorophyll from some of the PBE members also increases significantly over the year ($r > 0.64, p < 0.05$), but the range shows no significant increase. However, the range from PBPE-low also shows a significant increase in range over the years ($r = 0.65, p < 0.05$), but using higher perturbation strength, the PBPEs show no significant trend, Figure 4.8(f). The spreads also do not increase as the noise increases. These might be caused by none of the PBE and PBPE members showing a decrease in the interannual mean. Larger range is produced from PBPE-low (0.61 mg m^{-3}) than PBPE-medium (0.53 mg m^{-3}). But PBPE-medium produces lower ensemble range compared to the original ensemble (0.59), especially from 2002 to 2007. Similar to PPE, the highest range of surface chlorophyll annual mean is produced from PBPE-high and the ensemble members can produce up to 51.9% higher and 37.8% lower annual mean than the default run.

In the PBE, in 2002, 2003, 2005, and 2006 in situ surface chlorophyll concentrations are overestimated by the ensemble, and in 1999, the observed chlorophyll is close to the ensemble mean, median, and default. PPE produces ensemble mean and median close to the reduced ensemble's, Figure 4.11(a) to (c), making only at 1999 the in situ is within the ensemble range, for all the three ensembles. For the PBPE, years when in situ observations are within the ensemble range are similar to the reduced ensembles, apart from in 2004, where the in situ is within PBPE-high range, Figure 4.11(f)..

4.5.2 Oligotrophic

Statistical Metrics

In the oligotrophic stations, the default run often produces better correlation coefficient and bias, especially for chlorophyll profile and surface. At ALOHA, in the reduced ensemble, the chlorophyll profile RMSEs from the ensemble mean and median are lower than the default run. Unlike at PAP, the PBPEs

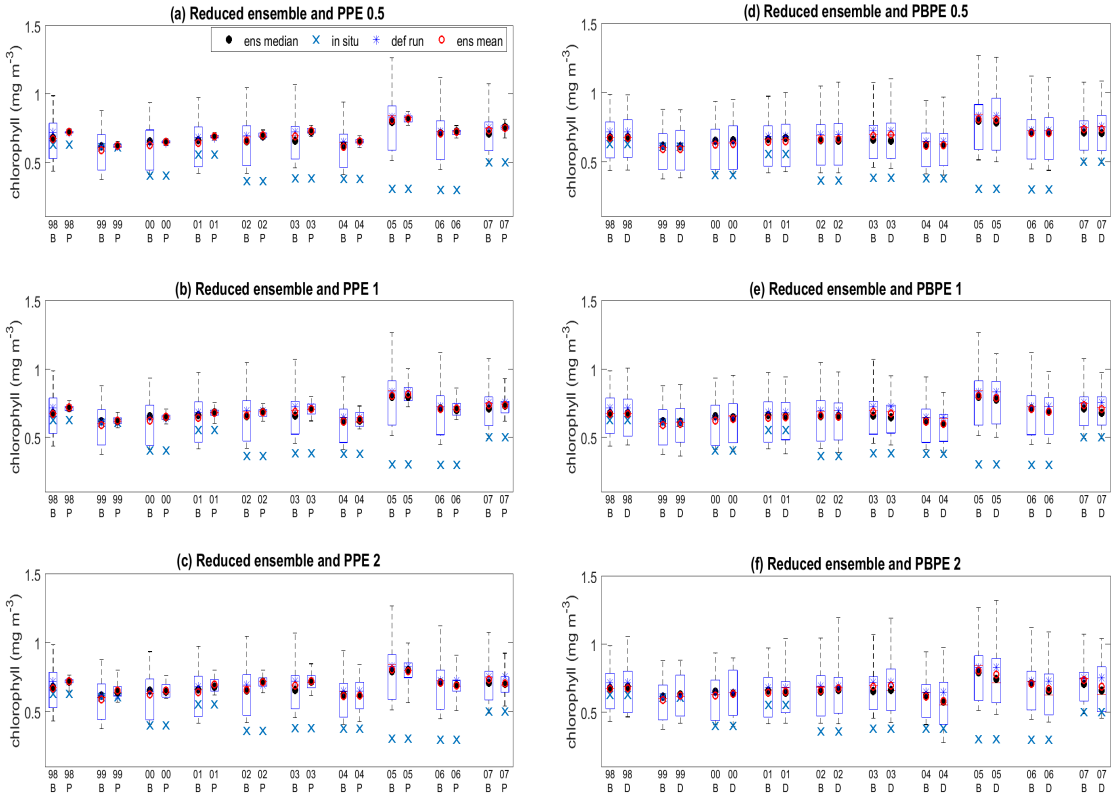


Figure 4.11: Interannual surface mean at station PAP from January 1998 to December 2007. B, P, and D denote when the ensemble that results from perturbing the functional forms, vertical velocity, and both functional forms and vertical velocities respectively. (a), (b), and (c) shows the perturbed biology ensemble compared with adding noises between $\pm$ 0.5, 1, and 2, respectively, to the vertical velocities. (d), (e), and (f) compares perturbed biology ensemble with both perturbing functional forms and adding noises between $\pm$ 0.5, 1, and 2, respectively, to the vertical velocities. Red open circle, black dot, blue star, and blue cross show ensemble mean, median, default run, and in situ observation of chlorophyll concentrations.

do not reduce the RMSE, although PBPE-medium produced the highest chlorophyll range compared to other perturbations. As the default run already produces low bias and high correlation coefficient, PPE-medium ensemble mean shows the highest correlation, and PPE-high ensemble mean shows the lowest bias for the chlorophyll profile.

At ALOHA, for surface chlorophyll, the highest correlation is produced from PPE-high shown in Table 4.6. PBPE-high produces the lowest RMSE and bias. However, PBPE-medium produces the highest surface chlorophyll range. For the integrated chlorophyll, the PBE median still produces the highest correlation and lowest bias. For RMSEs, the PBPE-high produces the lowest RMSEs of the integrated chlorophyll. Similar to the surface chlorophyll, the highest range is produced from the PBPE-high.

The default run has also produced low bias for DIN compared to the full ensemble at ALOHA. Perturbing physics, like that in PPE-medium, produces better bias than the default run, summarised on Table 4.6. The PBPE-low produces better RMSE and DIN. Despite producing higher RMSE and bias than the PBE,

the highest correlation coefficient is produced from PBPE-high, summarised on Table 4.6. Although producing high bias and RMSE, the largest spread for DIN is produced from PPE-high.

At station BATS, chlorophyll profile, surface, and integrated from the default run produced low RMSE and bias compared to the full ensemble. Lower bias and RMSE are observed from the ensemble mean of PPE-high for all the chlorophyll metrics. Adding noise to the PBE would increase the range, and therefore the highest range for chlorophyll profile, surface, integrated, and DIN profile is produced from PBPE-high as seen in Table 4.6. The default run produces low DIN bias, RMSE, and high correlation. The RMSEs and biases increase with the perturbation strength. However, PPE-high produces the highest correlation. This makes the smallest RMSE and bias for DIN profile are still produced from the default run.

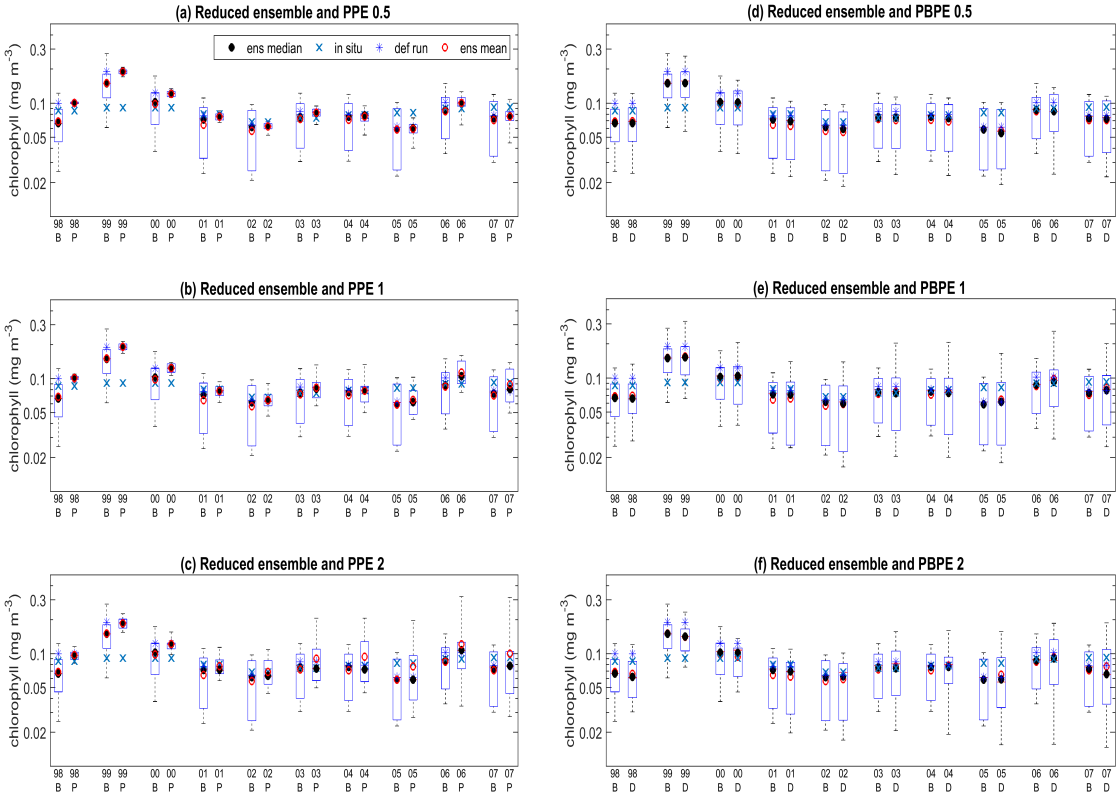


Figure 4.12: Interannual surface mean at station ALOHA from January 1998 to December 2007. B, P and, D denote when the ensemble that results from perturbing the functional forms, vertical velocity, and both functional forms and vertical velocities. (a), (b), and (c) shows the perturbed biology ensemble compared with adding noises between $\pm$ 0.5, 1, and 2, respectively, to the vertical velocities. (d), (e), and (f) compares perturbed biology ensemble with both perturbing functional forms and adding noises between $\pm$ 0.5, 1, and 2, respectively, to the vertical velocities. Red open circle, black dot, blue star, and blue cross show ensemble mean, median, default run, and in situ observation of chlorophyll concentrations.

Table 4.6: Statistical metrics comparing the default run and the mean and median of the original perturbed biogeochemistry ensemble (reduced to 12 members), only adding to the vertical velocities (PPE), and both perturbing functional forms and adding noises (PBPE) with the observation from stations ALOHA and BATS. The noises that are used are random numbers between $\pm$ 0.5, $\pm$ 1, and $\pm$ 2). The numbers that are in bold are the lowest bias and RMSE, and the highest correlation.

Station		chlorophyll profile (mg m^{-3})				DIN profile (mmol m^{-3})				surface chlorophyll (mg m^{-3})				integrated chlorophyll (mg m^{-2})			
		median	mean	default	in situ	median	mean	default	in situ	median	mean	default	in situ	median	mean	default	in situ
ALOHA	r	0.282	0.277	0.309		0.776	0.775	0.772		0.266	0.244	0.288		**0.752**	0.729	0.727	
	RMSE	**0.092**	**0.092**	0.093		1.056	1.049	1.068		0.046	0.046	0.065		2.581	2.664	4.152	
	Bias	0.052	0.053	0.032		-0.654	-0.650	f-0.590		-0.003	-0.001	-0.020		**0.030**	0.191	-1.000	
	mean	0.071	0.069	0.090	0.123	1.154	1.150	1.091	0.501	0.088	0.085	0.105	0.084	4.485	4.324	5.515	4.515
	Range	0.095				0.232				0.116				9.674			
PBPE	r	0.273	0.272	0.309		0.777	0.776	0.772		0.248	0.243	0.288		0.733	0.723	0.727	
noise= 0.5	RMSE	0.093	0.093	0.093		1.046	**1.032**	1.068		0.047	0.046	0.065		2.619	2.678	4.152	
	Bias	0.053	0.054	0.032		-0.648	-0.639	-0.590		-0.002	0.000	-0.020		0.130	0.249	-1.000	
	mean	0.069	0.068	0.090	0.123	1.148	1.139	1.091	0.501	0.087	0.084	0.105	0.084	4.385	4.266	5.515	4.515
	Range	0.092				0.358				0.109				9.920			
PPE	r	0.309	0.310	0.309		0.766	0.766	0.772		0.290	0.291	0.288		0.723	0.723	0.727	
	RMSE	0.099	0.099	0.093		1.418	1.419	1.068		0.081	0.080	0.065		4.115	4.098	4.152	
	Bias	0.027	0.027	0.032		-0.836	-0.835	-0.590		-0.025	0.025	-0.020		-0.929	-0.902	-1.000	
	mean	0.089	0.089	0.090	0.123	1.080	1.079	1.091	0.501	0.103	0.103	0.105	0.084	5.827	5.821	5.515	4.515
	Range	0.022				0.296				0.023				7.841			
PBPE	r	0.255	0.264	0.310		0.776	0.779	0.772		0.240	0.260	0.288		0.729	0.740	0.727	
noise= 1	RMSE	0.094	0.093	0.093		1.043	1.059	1.068		0.047	0.047	0.065		2.660	2.729	4.152	
	Bias	0.054	0.051	0.032		-0.640	-0.665	-0.590		-0.003	0.004	-0.020		0.037	-0.058	-1.000	
	mean	0.069	0.072	0.090	0.123	1.140	1.166	1.091	0.501	0.088	0.088	0.105	0.084	4.478	4.573	5.515	4.515
	Range	**0.133**				0.427				**0.166**				**20.080**			
PPE	r	**0.315**	0.314	0.309		0.772	0.775	0.772		0.295	0.300	0.288		0.728	0.736	0.727	
	RMSE	0.093	0.093	0.093		1.063	1.077	1.068		0.065	0.065	0.065		4.150	4.182	4.152	
	Bias	0.034	0.032	0.032		**-0.589**	-0.607	-0.590		-0.020	0.021	-0.020		-1.025	-1.122	-1.000	
	mean	0.090	0.091	0.090	0.123	1.360	1.379	1.363	0.501	0.105	0.106	0.105	0.084	5.540	5.637	5.515	4.515
	Range	0.034				0.471				0.049				11.545			
PBPE	r	0.260	0.263	0.310		0.774	**0.788**	0.772		0.228	0.250	0.288		0.735	0.753	0.727	
noise= 2	RMSE	0.094	0.093	0.093		1.033	1.090	1.068		0.044	**0.043**	0.065		**2.532**	2.429	4.152	
	Bias	0.055	0.054	0.032		-0.625	-0.699	-0.590		-0.001	**0.000**	-0.020		0.189	0.185	-1.000	
	mean	0.067	0.068	0.098	0.123	1.125	1.199	1.091	0.501	0.085	0.084	0.105	0.084	4.326	4.330	5.945	4.515
	Range	0.069				0.942				0.084				15.920			
PPE	r	0.297	0.294	0.309		0.767	0.780	0.772		0.281	**0.305**	0.288		0.692	0.731	0.727	
	RMSE	0.098	0.097	0.093		1.416	1.487	1.068		0.078	0.079	0.065		5.103	5.164	4.152	
	Bias	0.030	**0.024**	0.032		-0.823	-0.909	-0.590		-0.023	0.031	-0.020		-1.239	-1.740	-1.000	
	mean	0.093	0.099	0.090	0.123	1.324	1.409	1.091	0.501	0.108	0.117	0.105	0.084	5.754	6.255	5.515	4.515
	Range	0.052				**0.953**				0.117				6.255			
BATS	r	0.190	0.205	0.228		0.521	0.536	0.561		0.230	0.264	0.289		0.421	0.422	0.436	
	RMSE	0.324	0.320	0.306		1.436	1.410	**1.181**		0.331	0.327	0.311		51.629	51.018	48.330	
	Bias	0.119	0.109	0.066		-1.194	-1.169	**-0.849**		0.110	0.103	0.066		16.837	15.596	9.901	
	mean	0.054	0.064	0.107	0.173	1.809	1.784	1.464	0.615	0.055	0.063	0.099	0.166	7.065	8.307	14.002	23.902
	Range	0.165				0.909				0.156				14.385			
PBPE	r	0.221	0.205	0.228		0.528	0.538	0.561		0.280	0.267	0.289		0.461	0.424	0.436	
noise=0.5	RMSE	0.323	0.319	0.306		1.445	1.422	1.181		0.329	0.326	0.311		51.344	50.900	48.330	
	Bias	0.117	0.107	0.066		-1.207	-1.184	-0.849		0.109	0.101	0.066		16.657	15.381	9.901	
	mean	0.056	0.066	0.107	0.173	1.822	1.799	1.464	0.615	0.056	0.065	0.099	0.166	7.245	8.521	14.002	23.902
	Range	0.163				0.875				0.155				15.136			
PPE	r	0.228	0.230	0.228		0.562	0.563	0.561		0.290	0.293	0.289		0.435	0.437	0.436	
	RMSE	0.305	0.305	0.306		1.188	1.195	1.181		0.311	0.311	0.311		48.314	48.252	48.330	
	Bias	0.066	0.065	0.066		-0.859	-0.867	-0.849		0.066	0.065	0.066		9.818	9.716	9.901	
	mean	0.107	0.108	0.107	0.173	1.473	1.481	1.464	0.615	0.100	0.101	0.099	0.166	14.084	14.186	14.002	23.902
	Range	0.013				0.182				0.012				2.345			
PBPE	r	0.148	0.202	0.228		0.537	0.540	0.561		0.171	0.269	0.289		0.373	0.424	0.436	
noise=1	RMSE	0.328	0.319	0.306		1.444	1.421	1.181		0.334	0.326	0.311		52.256	50.887	48.330	
	Bias	0.125	0.107	0.066		-1.209	-1.183	-0.849		0.117	0.101	0.066		17.690	15.326	9.901	
	mean	0.048	0.066	0.107	0.173	1.824	1.798	1.464	0.615	0.049	0.065	0.099	0.166	6.212	8.576	14.002	23.902
	Range	0.189				0.728				0.180				15.423			
PPE	r	0.233	0.233	0.228		0.565	0.565	0.561		0.299	0.298	0.289		0.444	0.443	0.436	
	RMSE	0.305	0.305	0.306		1.223	1.205	1.181		0.310	0.310	0.311		48.163	48.224	48.330	
	Bias	0.064	0.065	0.066		-0.904	-0.884	-0.849		0.064	0.065	0.066		9.562	9.726	9.901	
	mean	0.109	0.108	0.107	0.173	1.518	1.498	1.464	0.614	0.102	0.101	0.099	0.166	14.340	14.176	14.002	23.902
	Range	0.034				0.561				0.032				7.468			
PBPE	r	0.200	0.193	0.228		0.544	0.543	0.561		0.302	0.258	0.289		0.448	0.423	0.436	
noise= 2	RMSE	0.322	0.317	0.306		1.434	1.486	1.181		0.327	0.323	0.311		51.133	50.360	48.330	
	Bias	0.114	0.099	0.066		-1.196	-1.255	-0.849		0.108	0.093	0.066		16.259	14.265	9.901	
	mean	0.059	0.074	0.107	0.173	1.812	1.871	1.464	0.616	0.058	0.072	0.099	0.166	7.643	9.637	14.002	23.902
	Range	**0.210**				**1.099**				**0.207**				**21.633**			
PPE	r	**0.243**	0.238	0.228		0.570	**0.570**	0.561		**0.318**	0.310	0.289		**0.457**	0.450	0.436	
	RMSE	**0.303**	0.303	0.306		1.217	1.266	1.181		**0.308**	0.308	0.311		**47.813**	47.846	48.330	
	Bias	0.061	**0.059**	0.066		-0.895	-0.952	-0.849		0.061	**0.060**	0.066		9.128	**8.927**	9.901	
	mean	0.112	0.114	0.107	0.173	1.511	1.567	1.464	0.615	0.104	0.106	0.099	0.166	14.774	14.975	14.002	23.902
	Range	0.114				1.014				0.059				9.611			

Interannual ranges of DIN and chlorophyll

Similar to the original PBE run, the ranges produced by the ensembles at the oligotrophic stations are the lowest. For the PPEs, the ranges of surface DIN show a strong increase over the time series for PPE-medium and high at both stations ($r = 0.72$ and $r = 0.82$, $p < 0.05$ respectively for ALOHA and $r = 0.74$ and $r = 0.75$, p< 0.05, respectively for BATS). This may be because two of the PPE-high members show a significant increase in DIN interannual mean at ALOHA ($r > 0.73$, $p < 0.05$), while the rest of the PPE-high members show no significant trend, shown in Figure 4.13. Similarly, two of the PPE-medium members also show increasing DIN interannual mean, and the rest are decreasing, but these trends of increasing and decreasing DIN interannual means are insignificant. At BATS, an ensemble member from PPE-high shows a significant increase in DIN concentration over the year ($r = 0.59, p < 0.05$), but the rest of the members show no significant trend. These lead to an increase in the surface DIN interannual range.

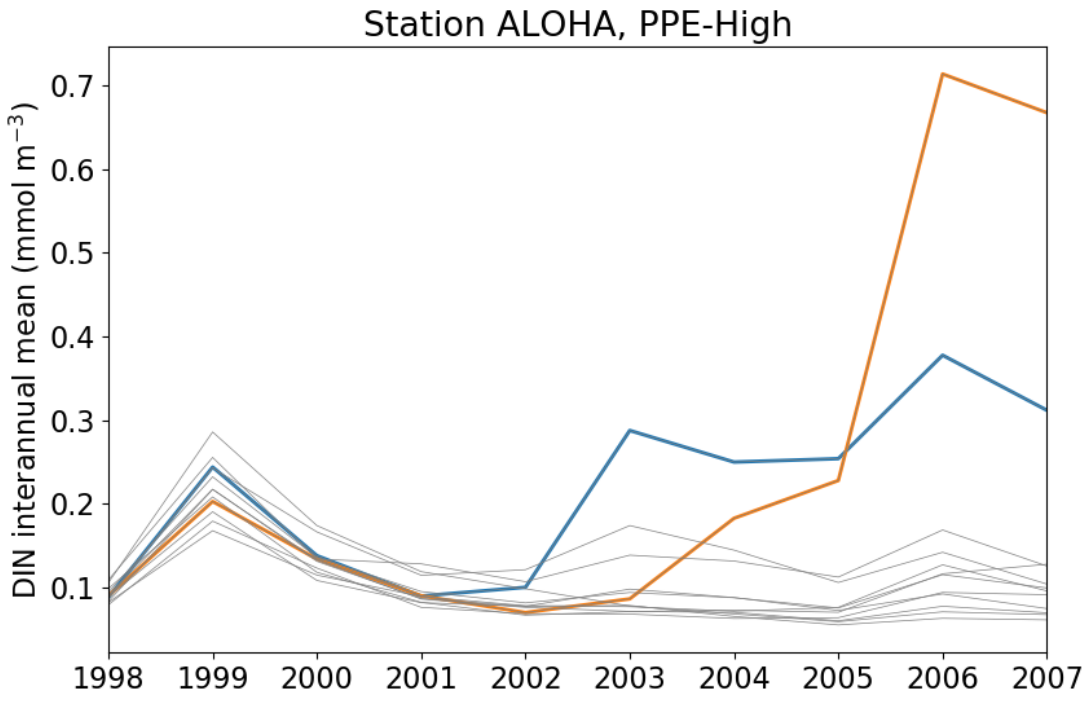


Figure 4.13: DIN interannual mean calculated from January 1998 to December 2007, from perturbing the vertical velocity using the highest perturbation strength (between -2 to 2). The grey lines show ensemble members that produce insignificant trend, whereas the orange and blue line show ensemble members that produce significant increase

The chlorophyll interannual mean range for PPE also shows an increase throughout the time series at both stations ($r = 0.80$, $p < 0.05$ for all of the ensembles at ALOHA and r =0.87 and $r = 0.79$, $p <$ 0.05 for the PPE-medium and high, respectively at BATS). However, there is no significant trend in PBE

surface chlorophyll range. This increasing trend from PPE-high, Figure 4.8(b), exceeds the range of the reduced ensemble from 2003 to 2007, making the PPE-high garnering larger spread than the PBE. The ensemble member from PPE-high at ALOHA can produce 70% higher and 6.53% lower surface chlorophyll annual mean than the default run, meaning that perturbing the vertical velocity produced ensemble members with mostly high DIN concentration at ALOHA. Similar to ALOHA, at BATS, the higher the noises of the PPE, the higher the range produced, Figure 4.8(c), however, the PPE-high range is not as high as the PBE. From the PPE-high at BATS, the ensemble member can produce up to 41% higher or 21% lower chlorophyll annual mean than the default run.

The PBE interannual mean range for DIN at BATS shows a significant increase ($r = 0.83, p < 0.05$), but not in ALOHA. The DIN ranges from PBPEs increase with the perturbation strength, for example the PBPE-medium and PBPE-high at ALOHA ($r = 0.87$ & $r = 0.90$, $p < 0.05$, respectively for PBPE medium and high). PBPE-high also produces the largest spread, Figure 4.9(g). At BATS since the PBE already shows an increasing trend, all the PBPE ensembles follow this trend ($r > 0.79$, $p < 0.05$). PBPE-low at BATS produces the highest overall surface DIN annual mean range (1.55 mmol m^{-3}), followed by PBPE medium and PBPE high (1.51 and 1.38 mmol m^{-3}, respectively), summarised in Figure 4.9.

In contrast to the DIN range, there is no clear trend of the interannual mean range of the surface chlorophyll from PBPE at both stations, apart from PBPE-high (r= 0.674, $p < 0.05$) at ALOHA. PBPE-high and medium exceed the ensemble spread produced from PBE at both stations, summarised in Figure 4.8(g) and (h). The ensemble members from PBPE-high can produce up to 114.6% higher and 91% lower than surface chlorophyll annual mean than the default run at BATS. Using PBPE-high at ALOHA, the ensemble members can produce 20% higher and 70% lower surface chlorophyll annual means than the default run. However, PBPE-medium can garner 104.4% higher and 70% lower surface annual mean than the default run.

At ALOHA the in situ interannual means are always within the ensemble range of the PBE. The observed chlorophyll concentrations are mostly higher than the ensemble mean or median. Since default run produces higher chlorophyll concentrations and therefore closer to the in situ, increasing the noise also increases the range, and therefore it captures most of the in situ interannual means, Figure 4.12(b) and (c). PBPEs captured the in situ chlorophyll concentrations regardless of the noise strength, Figure 4.12(d). As explained in the previous paragraph, PBPE-high produces less range compared to only adding PBPE-medium, as the latter produces higher upper adjacent than PBPE high.

At BATS in situ is within the range only in a few years, Figure 4.14(a). Most of the in situ concentrations are underestimated by the original PBE. The PPE-low produced very small range, which in turns makes the interannual means of surface in situ chlorophyll never within the range. However, PPE-high produces

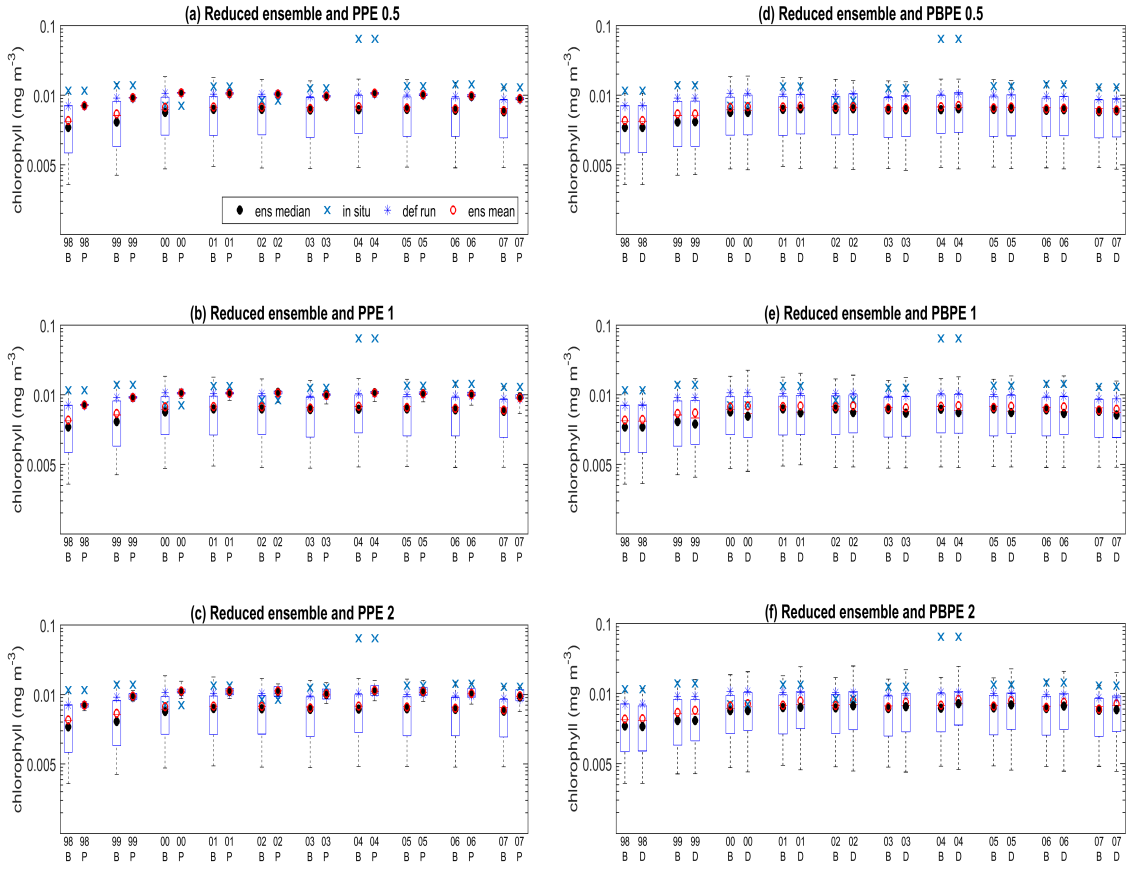


Figure 4.14: Interannual surface mean at station BATS from January 1998 to December 2007. B, P, and D denote when the ensemble that results from perturbing the functional forms, vertical velocity, and both functional forms and vertical velocities. (a), (b), and (c) shows the perturbed biology ensemble compared with adding noises between $\pm$ 0.5, 1, and 2, respectively, to the vertical velocities. (d), (e), and (f) compares perturbed biology ensemble with both perturbing functional forms and adding noises between $\pm$ 0.5, 1, and 2, respectively, to the vertical velocities. Red open circle, black dot, blue star, and blue cross show ensemble mean, median, default run, and in situ observation of chlorophyll concentrations.

enough spread in 2003, 2005, 2006, and 2007, to encompass the in situ concentrations, Figure 4.14(c). The ensemble range has been improved when both biogeochemistry and physics are perturbed, especially in PBPE-medium and high. In the original PBE, in 1998 and 1999, in situ interannual means are not within the ensemble range, and this has been improved by the PBPE, Figure 4.14(e) and (f).

4.5.3 Coastal

Statistical metrics

Table 4.7 summarised the statistical metrics for all the ensemble produced from varying functional forms, adding noises to the vertical velocities, and applying both to the model for the coastal stations. At Cariaco, the default run shows lower bias for both chlorophyll profile and surface, compared to the

Table 4.7: Statistical metrics comparing the default run and the mean and median of the original perturbed biogeochemistry ensemble (reduced to 12 members), only adding to the vertical velocities (w), and both perturbing functional forms and adding noises (PBPE) with the observation from the coastal stations (Cariaco and L4). The noises that are used are random numbers between ± 0.5, ± 1, and ± 2). The numbers that are in bold are the lowest bias and RMSE, and the highest correlation.

		chlorophyll profile (mg m^{-3})				DIN profile (mmol m^{-3})				surface chlorophyll (mg m^{-3})				integrated chlorophyll (mg m^{-2})			
Station		median	mean	default	in situ	median	mean	default	in situ	median	mean	default	in situ	median	mean	default	in situ
Cariaco	r	0.090	**0.157**	0.112		0.757	0.785	0.762		0.230	**0.316**	0.223		0.314	**0.437**	0.336	
	RMSE	1.278	1.276	1.266		3.203	2.970	3.291		0.865	0.865	0.865		18.983	17.812	18.712	
	Bias	0.210	-0.313	0.180		-0.521	-0.879	-0.589		0.141	-0.123	0.087		4.324	-3.288	3.862	
	mean	0.545	1.068	0.576	0.755	5.298	5.657	5.367	4.777	0.367	0.631	0.421	0.508	8.195	15.808	8.657	12.519
	Range	1.905				2.264				0.921				16.396			
PBPE	r	0.101	0.155	0.112		0.750	**0.786**	0.762		0.241	0.315	0.223		0.319	0.433	0.336	
noise= 0.5	RMSE	1.286	1.277	1.266		3.288	2.970	3.291		0.871	0.866	0.865		18.932	17.859	18.712	
	Bias	0.141	-0.315	0.180		-0.652	-0.886	-0.589		0.109	-0.124	0.087		3.222	-3.336	3.862	
	mean	0.614	1.071	0.576	0.755	5.429	5.663	5.367	4.777	0.399	0.632	0.421	0.508	9.298	15.855	8.657	12.519
	Range	1.924				2.299				0.934				16.417			
PPE	r	0.107	0.106	0.112		0.762	0.762	0.762		0.220	0.219	0.223		0.332	0.330	0.336	
	RMSE	1.267	1.267	1.266		3.254	3.218	3.291		0.865	0.864	0.865		18.757	18.782	18.712	
	Bias	0.187	0.190	0.180		-0.523	-0.464	-0.589		0.091	0.093	0.087		3.984	4.023	3.862	
	mean	0.614	1.071	0.576	0.755	5.429	5.663	5.367	4.777	0.399	0.632	0.421	0.508	9.298	15.855	8.657	12.519
	Range	0.237				1.547				0.113				0.684			
PBPE	r	0.073	0.151	0.112		0.755	0.784	0.762		0.222	0.311	0.223		0.287	0.424	0.336	
noise= 1	RMSE	1.286	1.264	1.266		3.074	2.842	3.291		0.869	0.857	0.865		19.172	17.798	18.712	
	Bias	0.193	-0.272	0.180		-0.329	-0.593	-0.589		0.133	-0.103	0.087		4.058	-2.679	3.862	
	mean	0.563	1.028	0.576	0.755	5.106	5.370	5.367	4.777	0.375	0.611	0.421	0.508	8.462	15.198	8.657	12.519
	Range	**1.969**				3.030				0.963				**20.256**			
PPE	r	0.093	0.092	0.112		0.755	0.755	0.762		0.205	0.202	0.223		0.311	0.307	0.336	
	RMSE	1.274	1.274	1.266		3.120	3.084	3.291		0.867	0.867	0.865		19.033	19.056	18.712	
	Bias	0.232	0.229	0.180		-0.106	-0.048	-0.589		0.114	0.112	0.087		4.640	4.605	3.862	
	mean	0.614	1.071	0.576	0.755	5.429	5.663	5.367	4.777	0.399	0.632	0.421	0.508	9.298	15.855	8.657	12.519
	Range	0.126				1.102				0.063				0.826			
PBPE	r	0.071	0.138	0.112		0.754	0.776	0.762		0.227	0.309	0.223		0.304	0.435	0.336	
noise=2	RMSE	1.284	**1.254**	1.266		2.988	**2.729**	3.291		0.867	**0.846**	0.865		18.956	**17.560**	18.712	
	Bias	**0.173**	-0.233	0.180		-0.082	**-0.035**	-0.589		0.134	**-0.072**	0.087		3.931	**-1.955**	3.862	
	mean	0.582	0.989	0.576	0.755	4.859	4.812	5.367	4.777	0.374	0.580	0.421	0.508	8.588	14.475	8.657	12.519
	Range	1.963				**4.116**				1.010				19.420			
PPE	r	0.074	0.068	0.112		0.715	0.746	0.762		0.185	0.187	0.223		0.292	0.294	0.336	
	RMSE	1.288	1.284	1.266		3.212	3.011	3.291		0.873	0.869	0.865		19.322	19.230	18.712	
	Bias	0.274	0.254	0.180		0.728	0.670	-0.589		0.143	0.130	0.087		5.363	5.096	3.862	
	mean	0.614	1.071	0.576	0.755	5.429	5.663	5.367	4.777	0.399	0.632	0.421	0.508	9.298	15.855	8.657	12.519
	Range	0.435				3.483				0.219				2.410			
L4	r					0.713	0.714	**0.718**		0.234	0.274	0.309					
	RMSE					2.721	2.615	2.668		1.078	1.143	1.276					
	Bias					-1.221	-1.017	-1.123		-0.462	-0.673	-0.834					
	mean					4.176	3.972	4.078	2.955	1.802	2.013	2.173	1.340				
	Range					3.063				**2.418**							
PBPE	r					0.710	0.713	0.718		0.246	0.268	0.309					
noise= 0.5	RMSE					2.689	2.634	2.668		1.065	1.152	1.276					
	Bias					-1.145	-1.057	-1.123		-0.455	-0.682	-0.834					
	mean					4.100	4.012	4.078	2.955	1.795	2.022	2.173	1.340				
	Range					3.603				2.364							
PPE	r					0.718	0.716	0.718		0.310	0.308	0.309					
	RMSE					2.654	2.670	2.668		1.265	1.279	1.276					
	Bias					-1.092	-1.114	-1.123		-0.820	-0.835	-0.834					
	mean					4.047	4.069	4.078	2.955	2.179	2.192	2.194	1.340				
	Range					1.237				0.592							
PBPE	r					0.698	0.715	0.718		0.237	0.274	0.309					
noise= 1	RMSE					2.642	**2.553**	2.668		1.042	1.110	1.276					
	Bias					-0.936	**-0.860**	-1.123		-0.411	-0.624	-0.834					
	mean					3.891	3.815	4.078	2.955	1.751	1.964	2.173	1.340				
	Range					3.740				2.356							
PPE	r					0.711	0.717	0.718		**0.314**	0.312	0.309					
	RMSE					2.563	2.571	2.668		1.170	1.204	1.276					
	Bias					-0.830	-0.905	-1.123		-0.703	-0.744	-0.834					
	mean					3.785	3.860	4.078	2.955	2.042	2.084	2.173	1.340				
	Range					1.635				0.755							
PBPE	r					0.707	0.706	0.718		0.228	0.254	0.309					
noise=2	RMSE					2.468	2.597	2.668		**1.037**	1.094	1.276					
	Bias					-0.303	-0.896	-1.123		**-0.399**	-0.576	-0.834					
	mean					3.258	3.851	4.078	2.955	1.739	1.916	2.173	1.340				
	Range					**5.670**				1.969							
PPE	r					0.700	0.711	0.718		0.287	0.298	0.309					
	RMSE					2.577	2.587	2.668		1.169	1.237	1.276					
	Bias					-0.749	-0.892	-1.123		-0.669	-0.769	-0.834					
	mean					3.704	3.847	4.078	2.955	2.009	2.108	2.173	1.340				
	Range					4.164				2.031							

reduced PBE mean and median. However, the PBE mean produces the highest correlation compared to other ensembles. For the chlorophyll profile, PBPE-medium produces the highest range. The PBPE-high ensemble mean and median produce the lowest RMSE and bias, respectively.

For surface chlorophyll at Cariaco, the highest range is recorded from PBPE-high, and its ensemble mean produced the lowest RMSE and bias. The PPEs produce higher bias of chlorophyll compared to the default run. The biases of these ensemble means and medians get higher as the perturbation strength increases. For integrated chlorophyll, PBPE-medium produced the lowest RMSE and bias, compared to other ensembles and also produced the highest range.

For the DIN profile at Cariaco, the original PBE mean already has a higher correlation coefficient compared to the default run. However, its bias is higher than the default and ensemble median. PBPE-high ensemble means produced the lowest RMSE and bias. Additionally, the highest range is produced from PBPE-high. Similar to the previous stations, the range of DIN from the PPE increases with the perturbation strength. Furthermore, PPE-high produces a higher range than PBE.

At L4, the in situ observation is only available in the surface. Compared to the default run, the ensemble mean and median of the reduced PBE shows lower correlations, but better RMSEs and bias for both DIN and chlorophyll. Unlike Cariaco, the range of surface chlorophyll is at its highest from PBE. However, the ensemble median from PBPE-high produced the lowest RMSE and bias, summarised in Table 4.7. The PPE-medium produces better the correlation between the in situ and ensemble mean and median, compared to the default run.

In terms of surface DIN, only perturbing the vertical velocity worsen the ensemble mean and median correlations, compared to default run, which produces the highest correlation. The PBPE-medium produces the lowest RMSE and bias compared to other ensembles. Similar to Cariaco, the highest DIN range is produced from PBPE-medium.

Interannual ranges of DIN and chlorophyll

Similar to the other stations, the DIN interannual mean range from the PPE-medium and PPE-high at L4 increases throughout the time series (r =0.80 and 0.73, for PPE-medium and high, respectively $p <$ 0.05), despite the PBE range showing a decline over the years, shown in Figure 4.9(e). The range also increases as the noise gets stronger, with PPE-high producing higher DIN annual mean than PBE (range 4.08 mmol m^{-3} and 3.95 mg m^{-3} for PPE-high and PBE, respectively). At L4, for both the PBE-medium and high most of the ensemble members show a significant decrease in the DIN interannual mean. However, some members show an insignificantly weak increase in surface interannual DIN range.

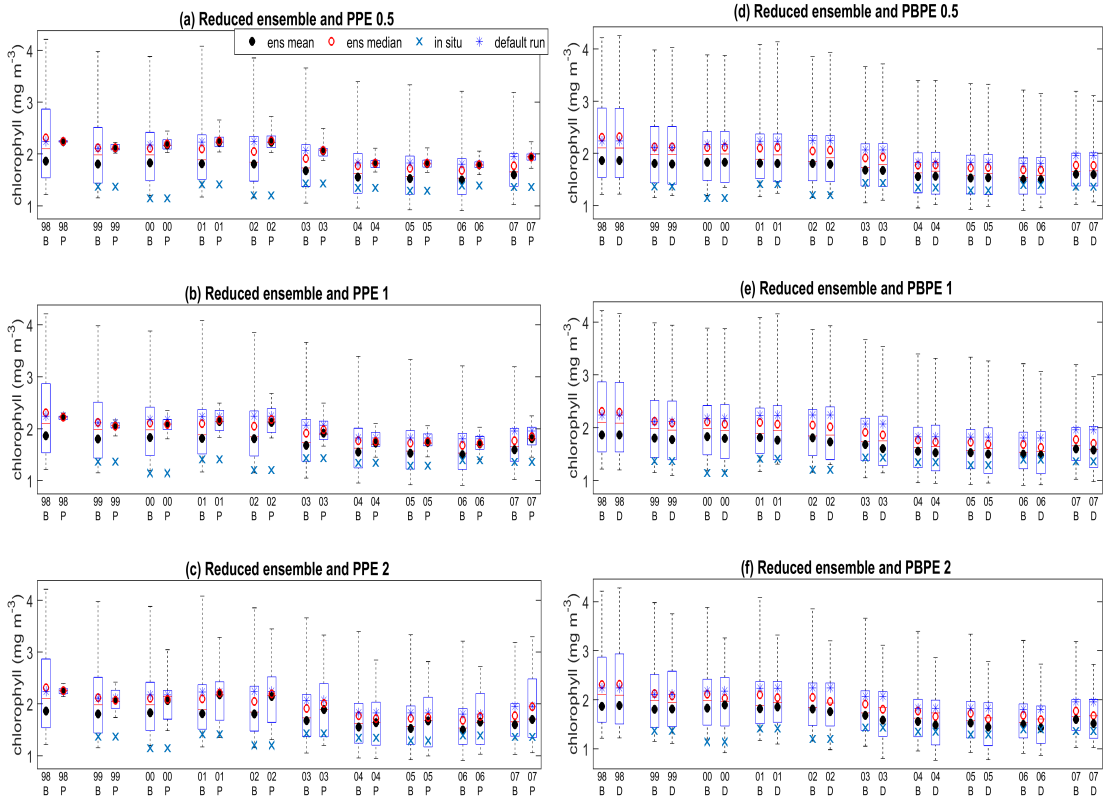


Figure 4.15: Interannual surface mean at station L4 from January 1998 to December 2007. B, P, and D denote when the ensemble that results from perturbing the functional forms, vertical velocity, and both functional forms and vertical velocities. (a), (b), and (c) shows the perturbed biology ensemble compared with adding noises between $\pm$ 0.5, 1, and 2, respectively, to the vertical velocities. (d), (e), and (f) compares perturbed biology ensemble with both perturbing functional forms and adding noises between $\pm$ 0.5, 1, and 2, respectively, to the vertical velocities. Red open circle, black dot, blue star, and blue cross show ensemble mean, median, default run, and in situ observation of chlorophyll concentrations.

These members may cause an increasing range of interannual DIN in L4. The PPE DIN interannual range in Cariaco has no significant trend, and the overall annual mean range increases with noise. In terms of chlorophyll interannual mean, the higher the PPE noise, the larger the ensemble spread at both Cariaco and L4. However, at Cariaco, all PPE ranges have no significant trend, shown in Figure 4.8d. At L4, PPE-high can produce up to 42.9% higher and 31.6% lower annual mean than the default run. Similarly, at Cariaco, the ensemble members from PPE-high can produce up to 33.4% higher and 43.6% lower surface chlorophyll annual mean than the default run.

At L4, the DIN annual mean range of PBPE increases with noise, (ranges for PBPE-low, medium, and high are 3.64, 3.69, and 5.43, respectively). The interannual ranges of DIN from PBPEs also increase throughout the time series, ($r > 0.77$, $p < 0.05$, for all of the ensembles). This increasing interannual DIN range has been observed from the PBE ($r = 0.94$, $p < 0.05$), Figure 4.9(j). Similar to perturbing the physics, the increase in range over the year are caused by some ensemble members that show increasing DIN interannual mean concentration while most of the members show significant depletion, shown in

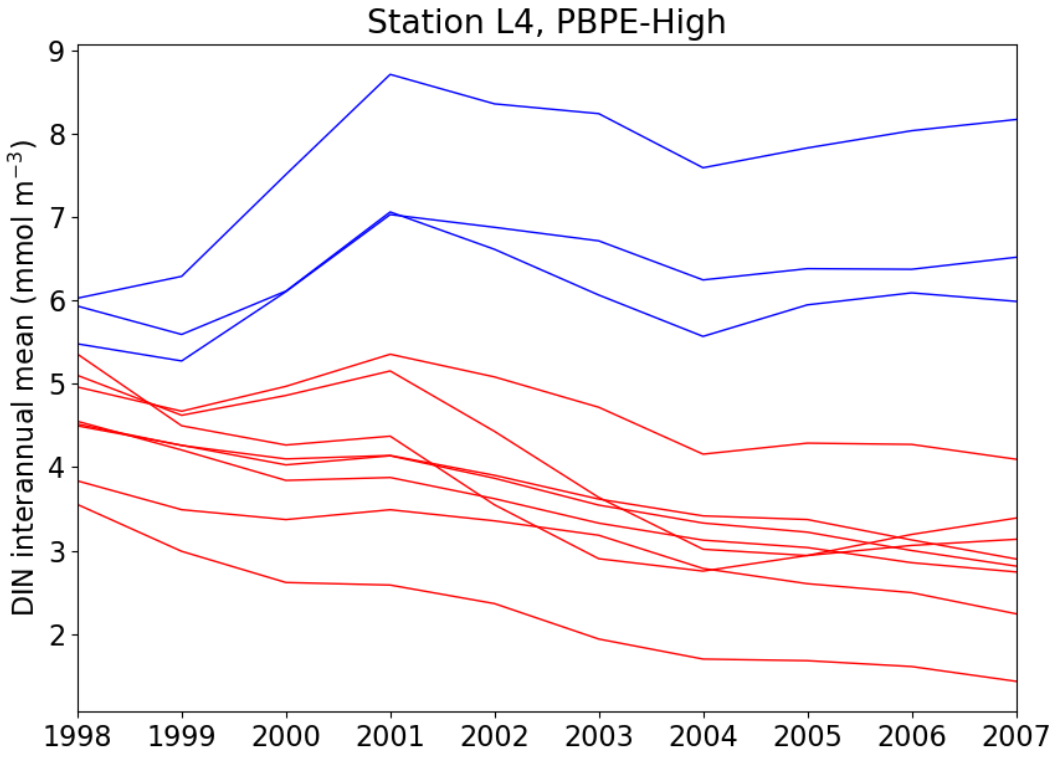


Figure 4.16: DIN interannual mean calculated from January 1998 to December 2007, from perturbing both the vertical velocity and biogeochemistry using the highest perturbation strength (between -2 to 2). The red lines show ensemble members that produce decreasing DIN interannual mean, whilst blue lines show ensemble members that produce insignificant increase in DIN interannual mean.

Figure 4.16. At Cariaco, the range of PBPE for DIN does not increase with noise, nor with the time series, shown in Figure 4.9(i). In chlorophyll, despite using the highest perturbation strength, no meaningful trend of the range over the time series is produces at Cariaco.

Chlorophyll annual mean ranges from PBPEs do not always increase as the perturbation strength increases at L4. In Figure 4.8(j), all of the PBPE range also produced decreasing trend ($r < -0.93, p < 0.05$ for all the PBPEs), similar to PBE ($r = -0.95, p < 0.05$). Most of the ensemble members in PBPE show a strong decreasing trend in the surface chlorophyll interannual mean $r < -0.67, p < 0.05$, and although one of the members show a weak, insignificant increase, this does not affect the interannual mean range. The smallest range at L4 have been observed from PBPE-high, followed PBPE-medium, and PBPE-low. At Cariaco, the surface chlorophyll interannual range from PBPEs do not show any significant trend. At both coastal stations, the chlorophyll interannual mean ranges from perturbing both physics and biogeochemistry show opposite trend to DIN, summarised in Figures 4.9(i), (j), 4.8(i), and (j). The ensemble members from PBPE-high can produce up to 50.8% higher and 48.3% lower annual mean than the default run, but PBPE-low can produce up to 76.5% higher surface chlorophyll annual mean than the

default run (but only 43% lower). At Cariaco, the ensemble members from PBPE using the highest noise can produce a lot higher annual mean than the default run (up to 262.6% higher and 77.6% lower surface chlorophyll annual mean).

In station L4, when the functional forms are perturbed, apart from the year 2000 the in situ interannual mean is within the ensemble range, and are often within the interquartile range, Figure 4.15. Since the default run has mostly higher chlorophyll concentrations compared to the ensemble mean and median, the ensemble spread from PPE-small overestimated the in situ observation throughout the time series 4.15(a). However, increasing the noise to medium makes the in situ observation is within the ensemble interquartile range from 2003, Figure 4.15(c). The interquartile ranges in 2003, 2005, 2006, and 2007 are higher than PBE-high. The PBPE-high discussed above, and shown in Figure 4.8(j) shows the lowest range. This low range results in the in situ observation always within the ensemble range, Figure 4.15(f).

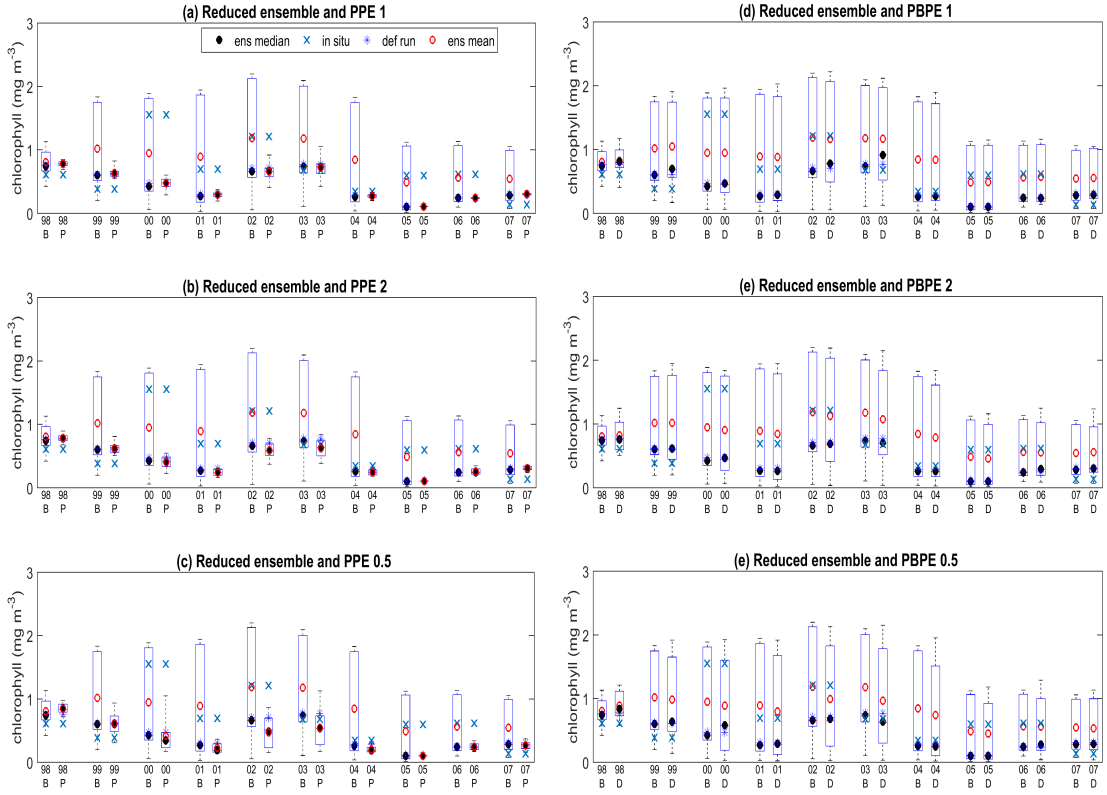


Figure 4.17: Interannual surface mean at station Cariaco from January 1998 to December 2007. B, P, and D denote when the ensemble that results from perturbing the functional forms, vertical velocity, and both functional forms and vertical velocities. (a), (b), and (c) shows the perturbed biology ensemble compared with adding noises between $\pm$ 0.5, 1, and 2, respectively, to the vertical velocities. (d), (e), and (f) compares perturbed biology ensemble with both perturbing functional forms and adding noises between $\pm$ 0.5, 1, and 2, respectively, to the vertical velocities. Red open circle, black dot, blue star, and blue cross show ensemble mean, median, default run, and in situ observation of chlorophyll concentrations.

At Cariaco, in situ surface chlorophyll interannual mean is always within the PBE range, Figure 4.17. Unlike that in L4, the default run is lower than the PBE mean. This makes perturbing the vertical veloc-

ities mostly underestimates, apart from in 2007, the in situ concentrations, especially in the year 2005, whereby the ensemble ranges are at their narrowest compared to the other years for all the ensembles produced from PPE, Figure4.17(a) to (c). Following PBE, the PBPEs the in situ interannual means are within the ensemble range. In situ concentrations that are within the PBE interquartile range stay within the PBPEs ranges, regardless of the perturbation strength.

4.6 Summary

We have explored two sources of uncertainties in a biogeochemical model; uncertainty in the choice of process equations, and the physical inputs. Here, the vertical velocity is chosen as the physical input representation, as it controls the availability of nutrients to the phytoplankton and therefore may affect the bias in anthropogenic CO_2 uptake (Doney et al., 2004). From the model runs, perturbing the vertical velocity using three different perturbation strengths show that generally, the higher the perturbation strength, the larger the range it would garner.

At most of the stations, perturbing the vertical velocity using medium and high noises produce an increasing trend of interannual DIN and hence the chlorophyll ranges. A few of the PPE members that produced opposing trend than the majority of the members may be the cause for the increase. For example, the majority of the PPE-high members in L4 show a weak increase ($r < 0.5$) but other members show a slight decrease in interannual means ($r >$ -0.5). Therefore, the range over the time series increases. At PAP, the increase in interannual range is caused by two PPE-high members that show opposing trends (sharp increase and decline), while other ensemble members show insignificant trends.

The ensemble means from the PPEs also produce better correlations and bias for chlorophyll than PBE. At oligotrophic stations, PPEs have produced better correlations and biases than the default run (e.g. chlorophyll profile and surface correlations from PPE-high at BATS are 0.243, and 0.318 mg m^{-3}, respectively, and the default run is 0.228 and 0.289 mg m^{-3} for chlorophyll profile and surface respectively). In the interannual means of surface chlorophyll, only the PPEs produce less spread than PBE and therefore does not garner enough range to encompass the in situ concentrations. At stations where the default run produces smaller bias than the PBE mean, the PPE-high produces enough spread for the ensemble to encompass the in situ concentrations. The PPE-high tends to produce a wider spread towards the end of the time series. Consequently, the in situ concentrations are almost always within the range at the end of the year.

Compared to the reduced PBE, the PPEs have not produced as much chlorophyll spread. However, in PPE-medium and PPE-high, the DIN profile spreads are larger than the PBE (e.g. at PAP, the range of

DIN profile from PBE is 1.99 mmol m^{-3}, and PPE-medium is 2.74 mmol m^{-3}). Perturbing the vertical velocity changes the nutrient availability in the upper water column in different ensembles and perturbing the functional forms change the interaction strength between nutrient, phytoplankton, and zooplankton compartments interact. Therefore PBE produces larger spread than PPE. For the statistical metrics, at most of the stations, the RMSE of DIN and chlorophylls from PBE means are better than perturbing the vertical velocity, although the correlations are improved. Only PPE-high produces large enough range to encompass the in situ surface chlorophyll interannual mean. Since the PPE-high spread increases over the year, the ensemble sometimes produced bigger spread than the PBE towards the end of the year.

Generally, the pattern of the interannual mean range from PBPE follows the PBE. Similar to only perturbing the vertical velocity, PBPE-high often produces the largest spread. Ensemble means from PBPEs generally produce better bias, r, and RMSE at the coastal and abyssal plain stations, Table 4.5 and 4.7, compared to the PBE. Since the pattern of PBPE is similar to PBE, the in situ surface chlorophyll annual means are often within the ensemble spread if the perturbed biogeochemistry ensemble has already encompassed the in situ and produced larger spread than PBE. Therefore according to these results, perturbing both can produce larger spread and also slightly better statistical metrics, without having high computational cost. At most of the stations, we have observed an increasing trend over the year from PBPE. However, the interannual chlorophyll range does not always follow similar trend as DIN (e.g. Figure 4.8(f) and 4.9(f)). This may be due to perturbing the biogeochemistry has a stronger effect on chlorophyll. Additionally at the coastal stations and BATS, an opposing trend have been observed from the two interannual means, (e.g. Figure 4.9(i) with 4.8(i) and 4.9(j) with 4.8(j)).

To explore further about the three different perturbations and the opposing trend of DIN and chlorophyll, we need to investigate the characteristic signatures of PBE, PPE, and PBPE. According to an earlier study, the physical input (Friedrichs et al., 2006) profoundly influences phytoplankton and DIN distribution in the water column. In the next chapter, the effect of perturbing the biogeochemistry in DIN, chlorophyll, and zooplankton distributions are going to be quantified. Furthermore, the characteristic signatures will be explored in the next chapter.

Perturbing with low noise does not show much difference. However, using high perturbation strength often produces a steady increase in DIN concentration. Therefore, in the next chapter, the medium noise is chosen for PPE and PBPE.

Chapter 5

Characteristic signatur in perturbed biogeochemistry and physics ensembles [3]

5.1 Introduction

As discussed in the previous chapters, ocean biogeochemical models have been developed to understand how the ocean ecosystem responds to the changes in both the physics and the biogeochemistry (Doney et al., 2012; Yool et al., 2013; Butenschön et al., 2016). Key uncertainties that affect ocean biogeochemical models include physical processes, with vertical mixing and upwelling of nutrients often poorly known (Doney, 1999; Sinha et al., 2010; Friedrichs et al., 2006), and the various choices for the mathematical formulations of the biological process for nutrient uptake, zooplankton grazing, and plankton mortality (Gentleman et al., 2003; Anderson et al., 2010; Adamson and Morozov, 2013). These processes are described by functional forms relating them to concentrations of plankton and nutrients, as well as changes in temperature and light.

Choosing different physical environments can strongly affect simulations of chlorophyll distribution through the water column (Friedrichs et al., 2006), as well as regional distributions of phytoplankton functional types in the surface ocean (Sinha et al., 2010). Additionally, varying the physical ocean models can produce $\pm 25\%$ variations in future (2100) anthropogenic CO_2 uptake (Doney et al., 2004). When the structure of an ocean biogeochemical model, especially the choice of the functional forms representing biogeochemical processes, strongly determine the model dynamics (Edwards and Yool, 2000; Fussmann and Blasius, 2005). For example, when the grazing function alone is altered from Holling type II to Holling type III (both of which are common in the literature) three times higher phytoplankton

[3] This chapter is based on parts of the paper Anugerahanti, P., Roy, S. and Haines, K. (2019), Perturbed Biology and Physics signatures in an ocean biogeochemical model ensemble, Frontiers in Marine Science. (In review).

concentrations have been produced (Anderson et al., 2010). Impacts of altering mortality are shown in both uncoupled NPZ models (Steele and Henderson, 1992; Edwards and Yool, 2000) and coupled ocean biogeochemical models (Yool et al., 2011). As shown by Yool et al. (2011), choosing a linear mortality in intermediately complex ocean biogeochemical model, can double the diatom biomass in high latitudes, compared to other functions. So the uncertainties arising from both physical and biogeochemical formulations contribute to the discrepancies between the models and observations (Anderson, 2010; Allen et al., 2010).

One way of accounting for multiple sources of uncertainty in ocean biogeochemical model simulations is to move away from deterministic simulations towards ensemble results which can be designed to deliver a probability distribution of outcomes. Perturbed physics ensembles have been used to estimate the uncertainties of climate projections (Tinker et al., 2015; Subramanian and Palmer, 2017) or forecasting the climate probabilistically (Tebaldi and Knutti, 2007; Murphy et al., 2007). Ensembles are also regularly used to quantify uncertainties in data assimilation applications (Anderson, 2001; Moradkhani and Meskele, 2010; Roy et al., 2012) to allow weighting of model results compared with new observations. From the previous chapter, differences between the range produced by the PBE and PPE in surface DIN and chlorophyll interannual range and other statistical metrics have been observed.

This chapter explores the variability that may arise in an intermediately complex 1-D ocean biogeochemical model by generating ensembles from perturbing (i) PBE, and (ii) PPE, and (iii) PBPE, especially within the water column. We will quantify the variability generated by the perturbed ensembles, identify, and distinguish the characteristics from different biological and physical perturbations based on several biogeochemical property metrics. From these characteristics, we can explore how the different perturbations may affect the model dynamics. Since the ocean biogeochemical model behaviour varies across different biogeographical provinces (Kriest et al., 2012), similar to the previous chapter, the model is run at five monitored ocean sites that span from coastal to oligotrophic regions.

This chapter is organised as follows: Brief descriptions of the metrics are explained in section 5.2. The basic diagnostics of the ensemble which relates to the bulk properties followed by the effect of perturbations in deep chlorophyll maxima are discussed in section 5.3. The different characteristic signatures of the PBE and PPE are described and discussed in section 5.4.

5.2 Ensemble metrics

Generating the three ensembles has been described in Chapter 4, and is summarised in Figure 5.1. This chapter will highlight the key properties of the model ensembles which we use to compare with obser-

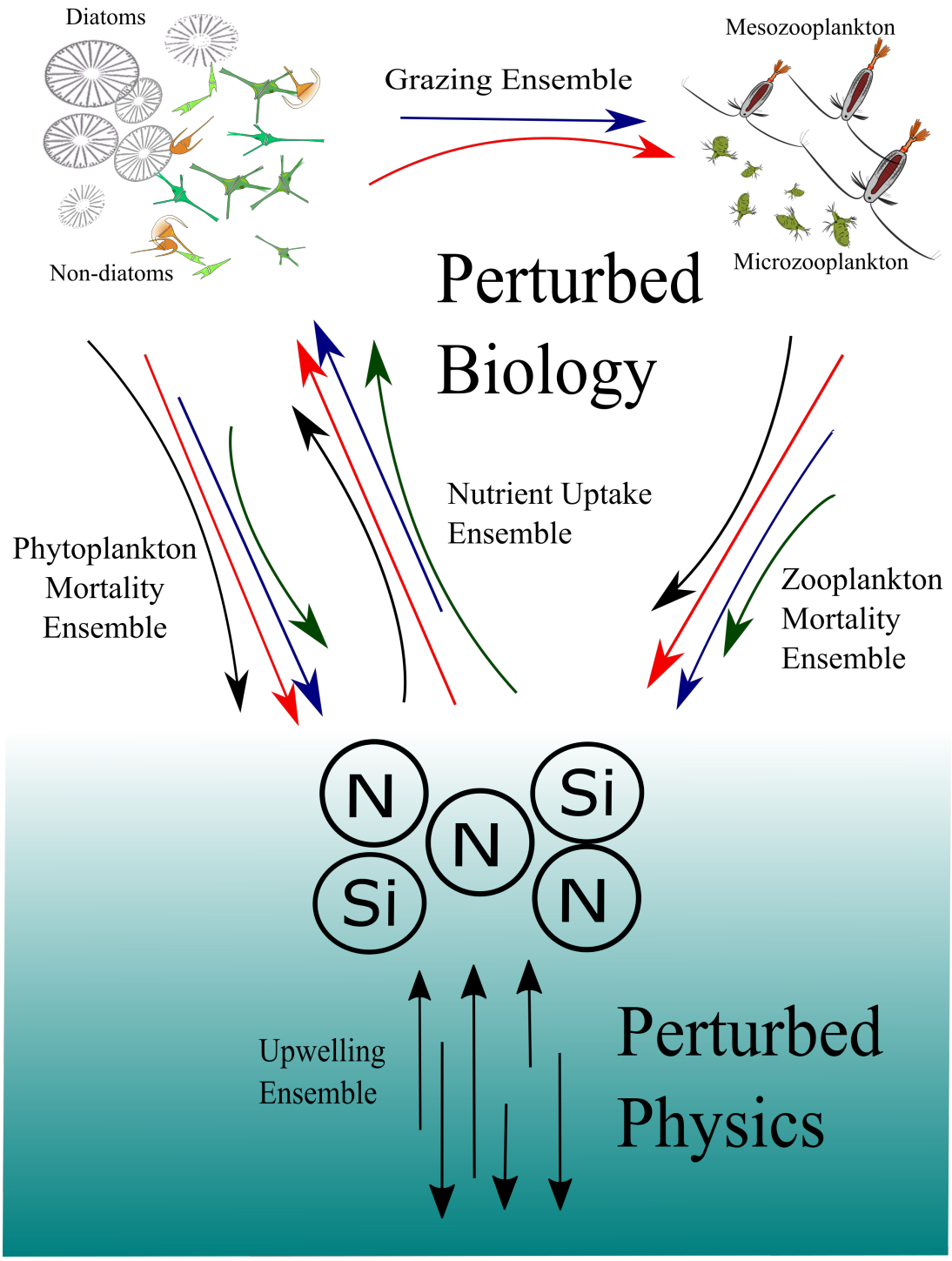


Figure 5.1: Schematic diagram showing how the ensembles are generated. The arrows in the top part represent the different functional forms which describe the key biogeochemical processes. These functions are similar in shape and have been tuned to enhance their similarities using non-linear least squares. We have chosen 4 functional forms for nutrient uptake; (hyperbolic (U_h), sigmoidal (U_s), exponential (U_e), and trigonometric (U_t)). We use 4 forms for zooplankton (ζ) and phytoplankton mortalities; (hyperbolic (ρ_h), linear (ρ_l), quadratic (ρ_q), sigmoidal (ρ_s)), and two functional forms for zooplankton grazing (Holling type III (G_1) and Holling type II (G_2)), resulting in 128 functional combinations including the default run. These are reduced to 12 to minimise the computational cost, as described in the supplementary material. The arrows at the bottom represent varying vertical velocities which generate the PPE. The PBPE is the combination of the two.

vations at the 5 oceanographic stations. The spread of the annual means of dissolved inorganic nitrogen (DIN mmol m^{-3}), chlorophyll (mg m^{-3}), and zooplankton (mmol m^{-3}) concentrations are the basic diagnostics throughout the water column. At the oligotrophic stations deep chlorophyll maxima (DCM) is a common feature that occurs below the mixed layer when surface chlorophyll concentration is low (Fennel and Boss, 2003; Letelier, Karl, Abbott and Bidigare, 2004). The DCM evolution is explored by its maximum depth and concentration over the winter (December-January-February), spring (March-April-May), summer (June-July-August), and fall (September-October-November). The range of DCM depth, timing of maximum depth, and concentration are examined for both the PPE and PBE, and whether the in situ observations of these quantities lie within the ensemble ranges. Additionally at stations which are situated in the North Atlantic, both diatoms and non-diatoms are simulated by the model. As suggested in Chapter 3 section 3.4, at stations L4 and PAP, we will examine how perturbing the vertical velocity and process equation may affect the distribution of dominant phytoplankton type in the water column. This is done by comparing the concentration of diatom and non-diatom chlorophyll at each depth for each day of the simulation. The dominant PFT is selected when the chlorophyll concentration has exceeded 60% of the total chlorophyll (e.g. if the concentration of diatom has reached 60% of the total chlorophyll, then diatom is the dominant PFT).

We also examine the correlated variability of DIN, chlorophyll, and zooplankton anomaly concentrations across different ensemble members, during bloom and non bloom-periods (e.g. bloom period at PAP occurred between May and July as shown in Chapter 2 and 3). The anomaly is calculated by subtracting the concentrations of phytoplankton, zooplankton, and DIN from the concentrations of the ensemble mean, then the correlation is calculated between the compartments from the mean of each of the ensemble members. Also the phytoplankton fractions in nitrogen is shown. From these metrics, it is possible to determine signatures of the processes which have been changed within the ensembles. The anomaly correlations along with these fractional metrics give an indication of the processes involved in the temporal changes seen from the in situ observations, suggesting it may be possible to infer which processes (physical, biological, or both) may be responsible.

5.3 The effect of perturbing biogeochemistry, physics, and both

5.3.1 The range of DIN, chlorophyll, and zooplankton concentration at depth

From Chapter 3, we have seen that the PBE mostly encompasses the in situ surface chlorophyll concentrations at all of the stations. However, in the oligotrophic stations, the DIN concentrations are overestimated by the ensemble. Although the ensemble has been reduced, the in situ chlorophyll at all of the

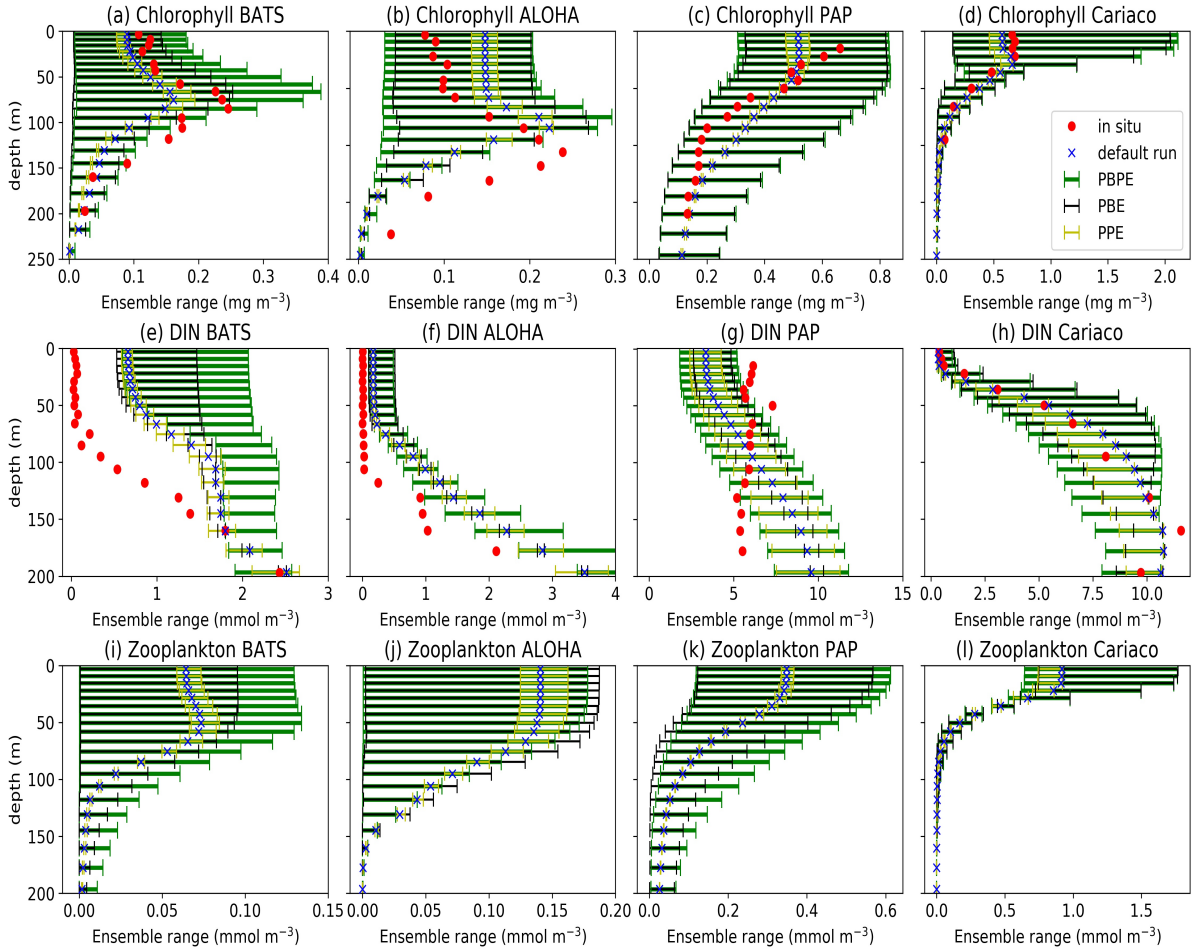


Figure 5.2: Ensemble range of annual mean of chlorophyll (a to d), DIN (e to h), and zooplankton (i to l) profiles between 1998 to 2007 at BATS (a, e, and i), ALOHA (b, f, and j), PAP (c, g, and k), and Cariaco (d, h, and l). Blue cross, red, yellow, black, and green dots denote the mean concentrations from the default run, in situ, the ensemble mean of PPE, PBE, and perturbing both biogeochemistry and physics. For station PAP, the annual mean is taken form 2004 for DIN and 2003 for chlorophyll. Station L4 profiles are not shown because in situ data is only available in the surface.

stations are still within the ensemble range. In terms of DIN, the ensemble at the oligotrophic stations still produces ranges that overestimates the in situ concentrations in most depth. For other stations, the ensemble ranges can still capture the in situ chlorophyll and DIN, similar to the full ensemble.

Perturbations to the vertical velocity used for the PPE, produce relatively little spread in the bulk properties (e.g. the total concentration of DIN, chlorophyll, and zooplankton) at each depth. The immediate impact should be seen in the concentrations of DIN at upper ocean levels but in the biologically active top ($\sim$75m) the PPE range does not change significantly although it widens at depth (Figure 5.2) showing that the vertical velocity variations are having an impact below the biologically active layers. However this does not have a big impact on bulk properties (Figure 5.2). These bulk properties have also been seen to be fairly insensitive when different ocean general circulation models have been coupled with the same ocean biogeochemical model (e.g., Sinha et al. (2010)).

The spreads generated by perturbing both physics and biology (i.e., PBPE) are mostly only slightly wider than for PBE alone, at least in the biologically active zone. Below this layer, the ensemble from PPE, especially in the oligotrophic stations produces larger spread (Figure 5.2(e) and (f)). However at Cariaco and PAP, the PBPE produces slightly larger DIN spread than PPE and PBE, even below the biologically active depths (Figure 5.2(g) and (h)). The spreads generated by perturbed physics alone are therefore mostly insufficient to encompass the in situ observations, especially for chlorophyll. In contrast the observed concentrations of chlorophyll at all five stations, from surface to deep water, are mostly within the PBE range (Figure 5.2(a)-(d)), suggesting that the full range of biological production through a strong nutrient gradient can be obtained by perturbing the biological processes equations.

Discrepancy between observations and ensemble simulation

If we consider individual stations, as discussed in Chapter 3, at Cariaco good agreement is found for DIN between the in situ and all the ensembles; PBE, PPE, and PBPE. However, at the oligotrophic stations BATS and ALOHA, a large mismatch between the observed and modelled DIN is apparent (Figure 5.2(e)–(h)). Similar to the results in Chapter 3, the observed DIN in the top 150m (Figure 5.5(e)–(f)) are beyond the ranges produced by either PBE or PBPE. This discrepancy in the oligotrophic stations indicates that uptake process is insufficient at low nutrient concentrations, leaving high DIN concentrations along with an underestimation of chlorophyll, (Yool et al., 2011; Cox and Kwiatkowski, 2013; Kwiatkowski et al., 2014) similar to the results in Chapter 3. Furthermore, at ALOHA the range produced by the PBE and PBPE is not wide enough to encompass the in situ maximum chlorophyll depth ($\sim$ 110m). The PBPE and PBE produce maximum ranges at $\sim$ 80m, therefore at greater depths, the light availability is even lower, and therefore low phytoplankton growth rate is simulated, despite higher DIN

concentrations at this depth. Meanwhile, at BATS, the oligotrophic conditions have not been simulated well by the ensemble, where at the top $\sim$ 75m, the DIN concentration is a magnitude higher than the in situ. However, the in situ chlorophyll concentration at BATS is within the ensemble range. This may be due to the 1-D model not simulating the input temperature well, making the nutrient uptake at BATS done inefficiently.

Compared to the previous run in Chapter 3, the in situ DIN concentrations at station PAP that are within the full ensemble (at the top $\sim$40 m, Figure 3.1) are now outside the PBE range, shown in Figure 5.2(g). The observed DIN profile does not show an increase in DIN concentrations with depth. However, the model ensembles do show nutrients increasing with depth, such that DIN concentration is underestimated near the surface (at $<$75m), and overestimated at depth $>$120m (Figure 5.2(g)). From Figure 5.3, the DIN concentrations are not sampled evenly at different depths. This might be due to the quality-controlled DIN concentrations only being available at certain times and depths, and lateral advection that occurred between 2003 and 2004, which increases the DIN concentration below $\sim$150m (Hartman et al., 2015). The in situ DIN profile shown in Figure 5.2(g) shows little variation with depth, unlike the ensemble simulation. However, the chlorophyll, shown in Figure 5.4, has more samples that are quality controlled at different depths and times.

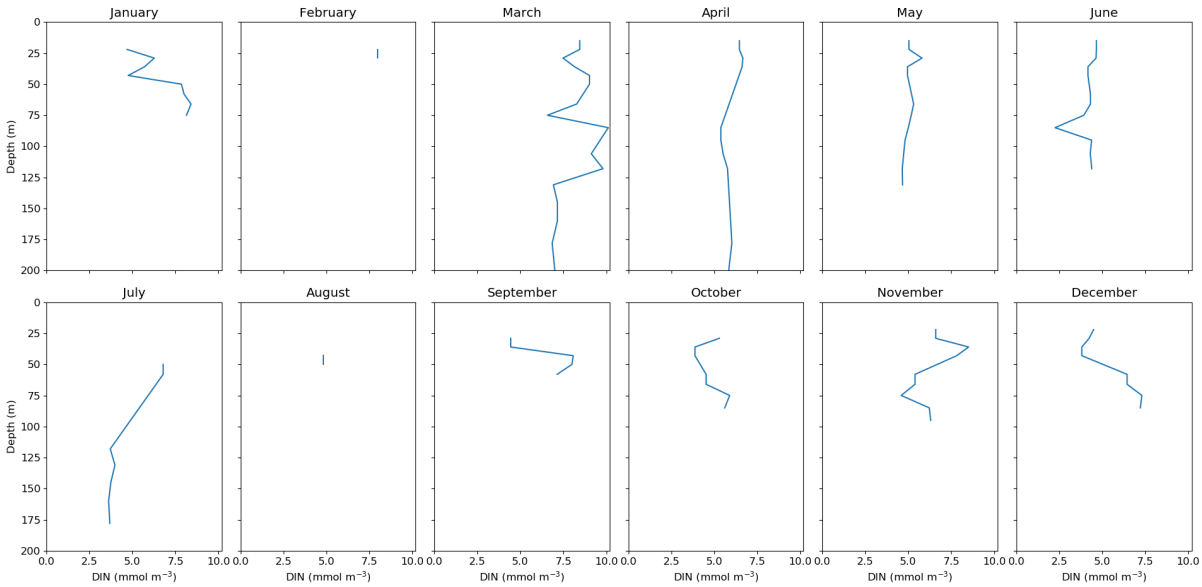


Figure 5.3: Monthly mean of in situ DIN (nitrate plus nitrite) at station PAP. The samples are taken between 2002 to 2004, in a sensor frame at 30m, which samples within deep chlorophyll maxima .

5.3.2 DCM at the oligotrophic stations

One of the characteristics of the oligotrophic regions is its DCM that occurs during months of low surface chlorophyll. The PBE and PPE members differ considerably in terms of DCM generations (e.g.

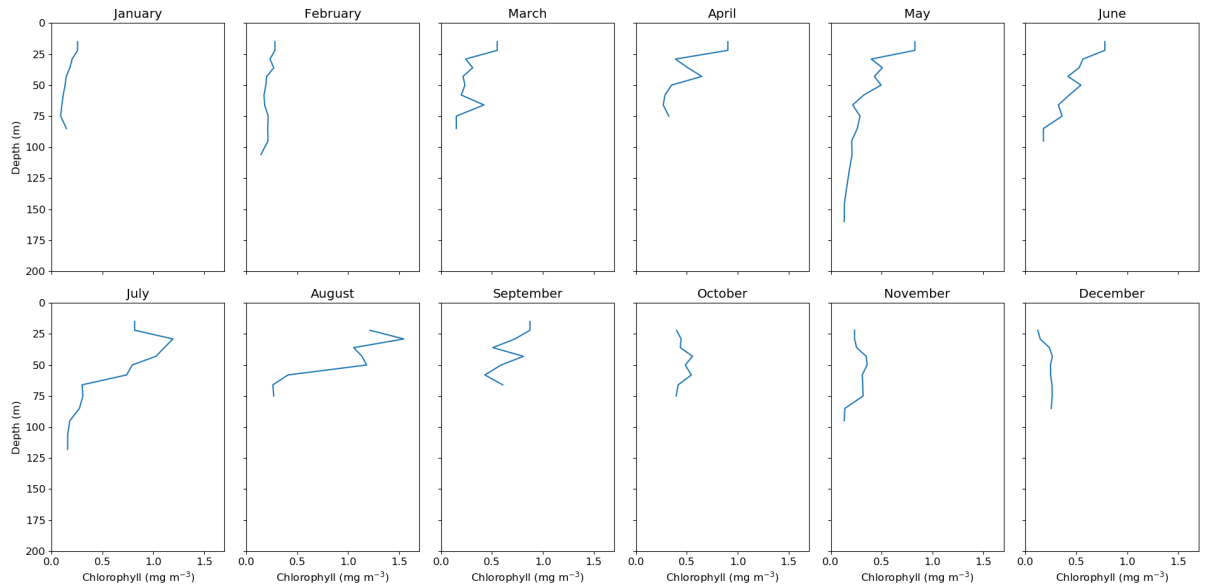


Figure 5.4: Monthly mean of in situ chlorophyll-α derived from the fluorescence at station PAP. The samples are taken between 2003 to 2005, in a sensor frame at 30m, which samples within deep chlorophyll maxima .

Figure 5.5 which shows chlorophyll distributions from 4 different members at BATS and ALOHA). The DCM is seen in all members for part of each year with considerable variability in maximum chlorophyll concentration and depth. In the observations, the deepest DCM always occurs in the summer and is shallowest in winter (Mignot et al., 2014). Similar to the range in the concentrations, the range of DCM depths from PBE is larger than PPE, with the observed DCM depths generally within the ensemble range with shallower DCM depths observed in BATS (e.g. the deepest DCM depths at ALOHA, PBE range 40-118m, PPE range 88-102m and in situ depth=114m, and at BATS PBE range 69-103m, PPE range 75-85m, and in situ depth = 93m, see Table 5.1). Both PBE and PPE members reproduce the observed shallowest DCM timings at both stations, however most of the PBE and all of the PPE members have the deepest DCM later in the autumn. The PBPE members mostly follows the pattern and timings of PBE, although the DCM depth range is slightly wider (e.g. at ALOHA, PBPE range= 66-114m, and at BATS PBPE range= 66-103m). The chlorophyll patterns from PPE members remain similar to the default run (e.g. all the PPE members showing the deepest DCM in fall at both stations, and the shallowest depth for the PPE members at BATS don't differ much see Table 5.1) and those PPE members that show weaker, or stronger negative vertical velocity than the default produce a slightly less prominent DCM.

Although the nutrient availability for the PBE is unchanged, the chlorophyll distribution and continuity of DCM depends on processes such as nutrient uptake, zooplankton grazing, and phytoplankton mortalities strength, and therefore it differs across ensemble members (e.g Figure 5.5(b) and (g)). Since DCMs occur at depth where the phytoplankton growth rate is in balance with the loss (Fennel and Boss, 2003; Cullen, 2014), variations in DCM depths, pattern, and continuity in PBE are therefore due to different loss and

Table 5.1: Maximum and minimum DCM depths and its timing of occurrence (see 5.2 for the months associated with the seasons) for PBE, PPE, and PBPE members. The DCM is obtained from the mean monthly chlorophyll and averaging the depth of the chlorophyll maximum between the three months to determine the timing of the deepest or shallowest DCM

	Stations							
	ALOHA				BATS			
	Minimum depth (m)	Season	Maximum depth (m)	Season	Minimum depth (m)	Season	Maximum depth (m)	Season
in situ	91.67	Winter	114.00	Summer	47.00	Winter	92.67	Summer
ensemble member				PBE				
default	67.67	Winter	91.67	Fall	34.33	Winter	78.33	Fall
$U_h G_2 \rho_h \zeta_h$	33.67	Winter	114.00	Fall	41.00	Winter	85.00	Summer
$U_t G_2 \rho_h \zeta_h$	33.67	Winter	118.33	Fall	38.67	Winter	95.33	Summer
$U_s G_2 \rho_h \zeta_q$	27.33	Spring	118.33	Fall	48.33	Spring	103.00	Summer
$U_e G_2 \rho_l \zeta_l$	3.00	Fall	39.67	Summer	43.00	Summer	88.67	Winter
$U_t G_1 \rho_l \zeta_s$	27.00	Winter	75.00	Summer	50.33	Summer	74.33	Winter
$U_h G_1 \rho_q \zeta_l$	67.67	Winter	88.33	Fall	27.00	Spring	78.33	Fall
$U_h G_2 \rho_s \zeta_l$	76.00	Spring	102.33	Fall	41.33	Spring	88.33	Summer
$U_e G_2 \rho_s \zeta_l$	30.33	Winter	98.67	Fall	38.67	Spring	85.00	Fall
$U_h G_1 \rho_s \zeta_l$	27.00	Winter	88.33	Fall	21.33	Winter	69.00	Fall
$U_h G_1 \rho_s \zeta_q$	27.00	Winter	91.67	Fall	23.33	Winter	69.00	Fall
$U_t G_1 \rho_s \zeta_q$	66.33	Spring	91.67	Fall	32.33	Winter	78.33	Fall
				PPE				
1	66.33	Spring	88.33	Fall	34.33	Winter	78.33	Fall
2	64.33	Winter	88.33	Fall	34.33	Winter	78.33	Fall
3	67.67	Winter	91.67	Fall	34.33	Winter	78.33	Fall
4	67.67	Winter	91.67	Fall	34.33	Winter	75.00	Fall
5	69.33	Spring	95.00	Fall	34.33	Winter	78.33	Fall
6	71.00	Winter	95.00	Fall	34.33	Winter	78.33	Fall
7	64.33	Winter	88.33	Fall	36.00	Spring	85.00	Fall
8	71.00	Winter	102.33	Fall	34.33	Winter	78.33	Fall
9	67.67	Winter	91.67	Fall	34.33	Winter	78.33	Fall
10	67.67	Winter	88.33	Fall	34.33	Winter	78.33	Fall
11	40.33	Winter	98.67	Fall	34.33	Winter	78.33	Fall
				PBPE				
$U_h G_2 \rho_h \zeta_h$	33.67	Winter	114.00	Fall	41.00	Winter	82.00	Summer
$U_t G_2 \rho_h \zeta_h$	29.67	Spring	85.00	Summer	38.67	Winter	95.33	Summer
$U_s G_2 \rho_h \zeta_q$	11.67	Spring	114.00	Fall	45.33	Winter	103.00	Summer
$U_e G_2 \rho_l \zeta_l$	3.00	Fall	66.33	Summer	43.00	Summer	88.67	Winter
$U_t G_1 \rho_l \zeta_s$	14.00	Spring	75.00	Summer	53.00	Summer	75.33	Fall
$U_h G_1 \rho_q \zeta_l$	64.33	Winter	81.67	Fall	24.67	Spring	75.00	Fall
$U_h G_2 \rho_s \zeta_l$	76.00	Spring	102.33	Fall	39.67	Winter	88.33	Summer
$U_e G_2 \rho_s \zeta_l$	33.67	Winter	95.00	Fall	38.67	Spring	85.00	Fall
$U_h G_1 \rho_s \zeta_l$	30.33	Winter	88.33	Fall	21.33	Winter	69.00	Fall
$U_h G_1 \rho_s \zeta_q$	38.00	Winter	88.33	Fall	21.33	Winter	66.00	Fall
$U_t G_1 \rho_s \zeta_q$	67.67	Winter	91.67	Fall	9.00	Winter	75.00	Fall

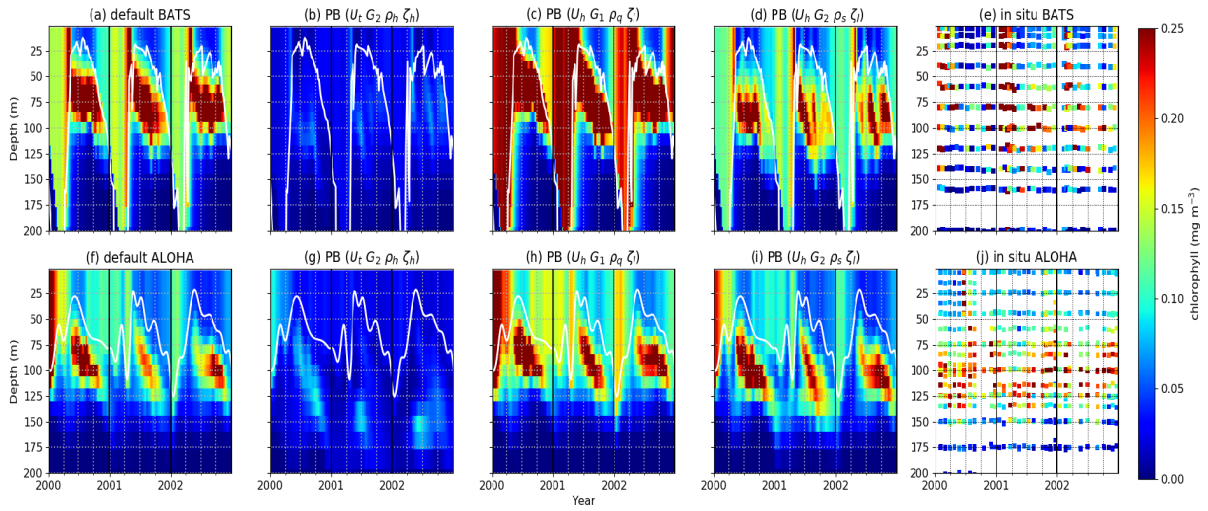


Figure 5.5: Chlorophyll distribution in the water column from 1st January 2000 to 31st December 2002 at station BATS (a to e) and ALOHA (f to j). White solid lines are the mixed layer depth. The default run is shown in (a) and (f) for BATS and ALOHA respectively. Different ensemble members from perturbing the biogeochemistry with their functional forms combinations are shown in (b), (c), and (d), for ALOHA, and (g), (h), and (i) for BATS. (e) and (j) are observed chlorophyll from BATS and ALOHA, respectively.

growth rates throughout euphotic depths. In oligotrophic regions, the nutrient supply is low in the top $\sim$ 150m (see Figure 5.2(e) and (f)) so that PBE members with high grazing and mortality rates under low nutrient and chlorophyll conditions (e.g. Holling type II (G_2), hyperbolic (ρ_h), and linear phytoplankton mortality (ρ_l), see Figure 2.3 for the grazing and mortality curves), have higher losses in the top 75m. At greater depths, nutrient is plentiful allowing phytoplankton growth to exceed the loss rate, giving a deeper DCM. Therefore, a slightly larger maximum DCM depth range in PBPE may be caused by the additional net upwelling from perturbing the physics which brings the maximum depth and for members with more downwelling, a deeper maximum DCM depth.

When the mixed layer depth becomes deeper, a balance cannot be achieved as light becomes a limiting factor and therefore the DCM concentration is not prominent below the mixed layer depths (see Figure 5.5(b) and (g)). Furthermore, the nutricline depth at BATS is shallower than ALOHA ($\sim$ 90m and $\sim$ 150m for BATS and ALOHA respectively, see Figure 5.2(e) and (f)), which makes the DCM depth at BATS shallower than ALOHA, as summarised in Table 5.1. These results suggest that perturbing the biogeochemistry can result in much greater variability in the evolution and continuity of the DCM, compared to perturbing the physics alone. This agrees with an earlier study which shows that in a 1-D biogeochemical model without data assimilation, the chlorophyll and zooplankton distribution in the water column is more sensitive to varying the model structure (in terms of complexity), compared to varying the physical forcings (Friedrichs et al., 2006). Furthermore, when perturbing both physics and biogeochemistry, the effect of perturbing the latter predominantly determines the ensemble spread and

chlorophyll distribution (see Appendix B for PPE and PBPE DCM distributions).

5.3.3 Dominant PFT in the water column

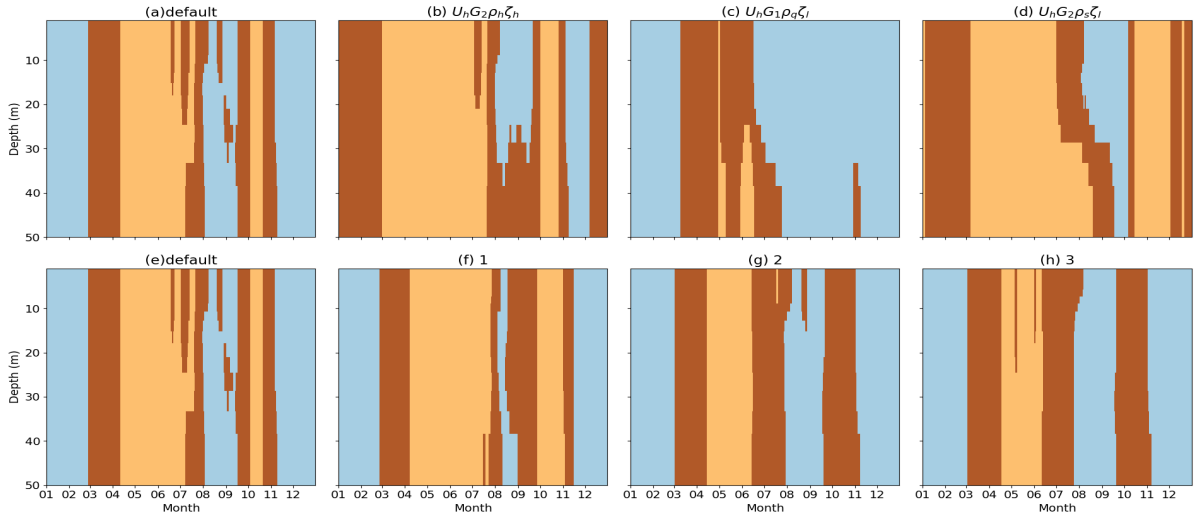


Figure 5.6: Distribution of dominant phytoplankton type at station L4 in 2003. The PBE results are shown in (a) to (e) and PPE in (e) to (h). Light blue and light brown denotes non diatoms and diatoms, respectively as dominant (concentration is larger than 60% of the total chlorophyll) phytoplankton type in the water column. Brown denotes that none of the PFT reached 60% of the total concentration.

As discussed previously, it has been shown that the regional distribution of dominant PFT is affected when the physical oceanography input is altered (Sinha et al., 2010) as well as when the process equation is perturbed (Anderson et al., 2010). Although our study cannot perform the global distribution study of phytoplankton group, it is possible to see the water column distribution of phytoplankton. The range of diatoms and non-diatoms chlorophyll concentrations from PPE is not as large as the PBE (e.g. diatom chlorophyll ranges from PPE and PBE at PAP are 0.058 mg m^{-3} and 0.555 mg m^{-3}, respectively, and diatom chlorophyll ranges from PPE and PBE at L4 are 0.127 mg m^{-3} and 1.176 mg m^{-3}). These results show that the PBE still produces a broader range than the PPE, even in the bulk properties of the different PFTs. Furthermore the 1-D model can reproduce the occurrence of diatoms in these regions and as the dominant phytoplankton type at L4, especially during bloom timing, as seen in Figure 5.6.

As described in the method section, we define the dominant PFT as the chlorophyll concentration of the PFT that exceeds 60% of the total chlorophyll (diatoms + non-diatoms) concentration. At L4, depth distributions of the dominant PFT varies between PPE members, as shown in Figure 5.6(e) to (h). Similarly, when the biogeochemical equations are perturbed, the depth distribution of PFT also varies even more between members, summarised in Figure 5.6(a) to (e). Both the PPE and PBE agree that between April and July, the dominant PFT is diatoms, which coincides with the bloom timing, and during autumn (August to October), the dominant PFT is the non-diatoms. However, one of the PBE members shown

here does not show dominant diatoms between April to July, instead it occurs in early May throughout the water column, and confined between ∼25m to 50m in June, shown in Figure 5.6(c). All the PPE members, on the other hand, always show continuous diatom domination between April to June. There are also instances where the dominant non-diatoms only occurs at a certain depth. For example, during fall, non-diatom domination occurs at the top ∼30m, but below this depth, there is no dominant phytoplankton (Figure 5.6(b)). This also occurs in PPE; in September, at the top ∼15m, there are no dominant PFT in August, but below that depth, non-diatoms dominate (Figure 5.6(g)).

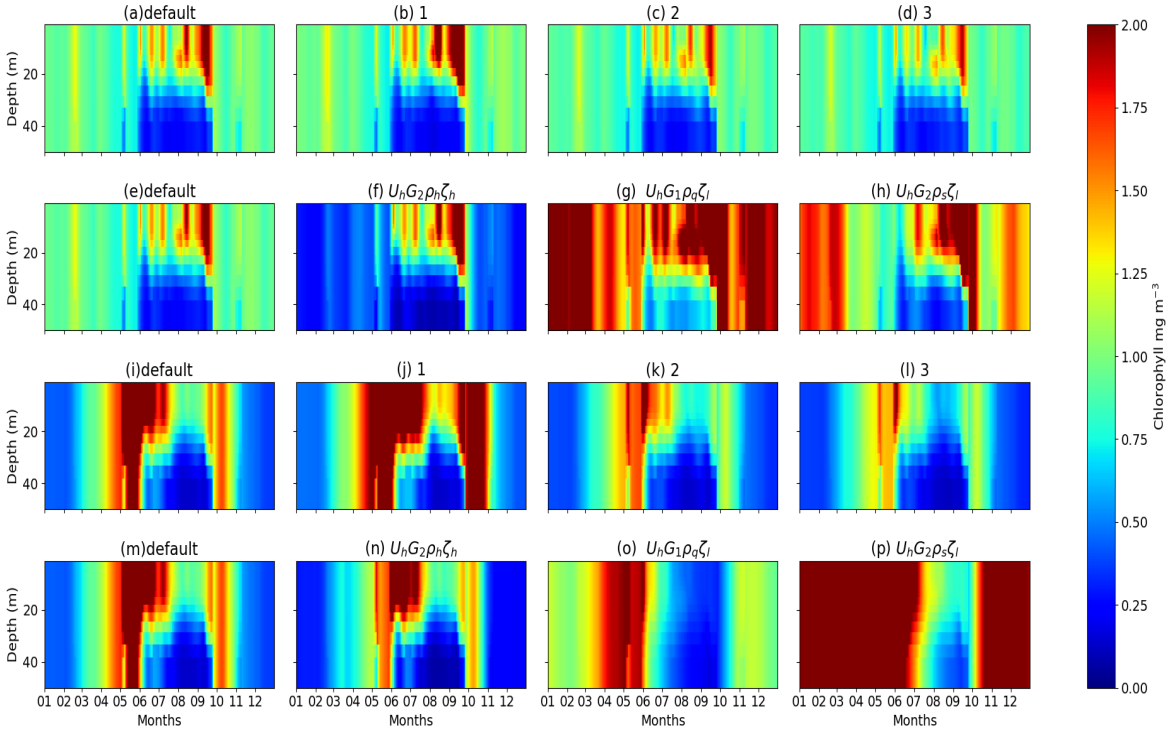


Figure 5.7: Chlorophyll profiles of non-diatom (a to h) and diatoms (i to p) at station L4 form the PPE (a to d for non-diatoms and i to l for diatoms) and PBE (e to h for non-diatoms and m to p for diatoms) members in 2003. Default run chlorophyll distributions are shown in (a), (e), (i), and (m).

Upon further investigations, the distribution of non-diatoms do not differ much within different PPE members, compared to the diatoms (see Figure 5.7(a) to (d)). This result agrees with that in Sinha et al. (2010) and Holt et al. (2014), where both studies show that the mixing intensities affects the diatoms concentrations and therefore its distributions. In PBE, the distributions of both non-diatoms and diatoms vary across different members, compared to PPE, similar to that in the DCM distributions shown previously.

Similar to L4, the distribution and timing of dominant PFTs at PAP vary across different PBE members. In the default run, the diatom domination only occurs between May and June and is confined within ∼70 to 100m, which coincides with the bloom timing at station PAP. In other PBE members shown in Figure 5.8(b) and (d), the diatom domination can occur from May until August, which coincides with the

bloom time. However, in Figure 5.6(c), the diatom dominance only occurred between June and July and was confined within ∼50 to 80m. The depth distribution and the longevity of the dominant PFTs also varies across different PBE members, for example, between ∼140-150m the diatom domination occurs longer than in the shallower depths (Figure 5.8(b) and (d)). Additionally, in Figure 5.8(c), unlike other members shown here, the functional form combinations produces non-diatom as the dominant species between May to July, below ∼120m.

The distribution of dominant PFT between the PPE members, do not vary as much as in PBE, as shown in Figure 5.8(e) to (g), where most of its ensemble members show similar dominant PFT distribution as the default run. However, in Figure 5.8(h), diatom domination occurs from the surface to ∼110m.

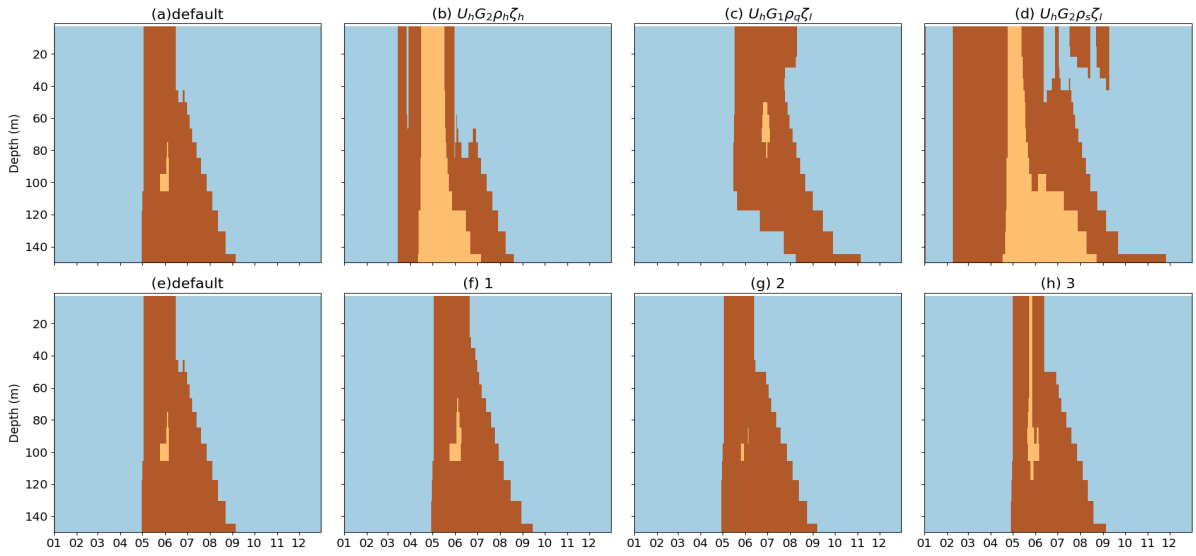


Figure 5.8: Distribution of dominant phytoplankton type at station PAP in 2003, form PBE (a to e) and PPE (e to h). Light blue and light brown denotes non diatoms and diatoms, respectively as dominant (concentration is larger than 60% of the total chlorophyll) phytoplankton type in the water column. Brown denotes that none of the PFT reached 60% of the total concentration.

Between PPE members, the difference in the dominant phytoplankton type depth distribution and occurrence may be due to the difference in mixing strength between the vertical velocity ensemble members. This may also explain the discrepancy between L4 and PAP, which may be caused by the difference in the magnitude of vertical velocity. At L4 and PAP, the mean default vertical velocities within top 50m from 1998 to 2007 is -0.0031 m s^{-1} and 0.00091 m s^{-1}, respectively. Therefore when the noise is added, the perturbation is stronger in L4, which may make the discrepancy in the distribution of PFT between PPE members more noticeable.

For PBE members, as explained in the previous subsections, these differences in dominant phytoplankton type distribution may be due to the difference in the interaction strength between the various functional forms combinations. For example when the mortality function is altered (Figure 5.6 (a) and (b) as well

as 5.6 (a) and (b)), the distribution of diatoms can change drastically, because using G_2 would result in a lower grazing rate when the concentration of phytoplankton is low, so that during winter to early spring the concentration of non-diatoms in ensemble member $U_h G_2 \rho_h \zeta_h$ is lower than the default run. Similarly these lower concentrations also occurs in diatoms, although not as distinctive as the non-diatoms (Figure 5.7 (m) and (n) and 5.9) (m) and (n)). The discrepancy between diatoms and non-diatoms may be due to non-diatoms being grazed by both microzooplankton and mesozooplankton, and diatoms only being grazed by the mesozooplankton. This further shows that perturbing the biogeochemistry provides larger variability even for the dominant PFT distributions.

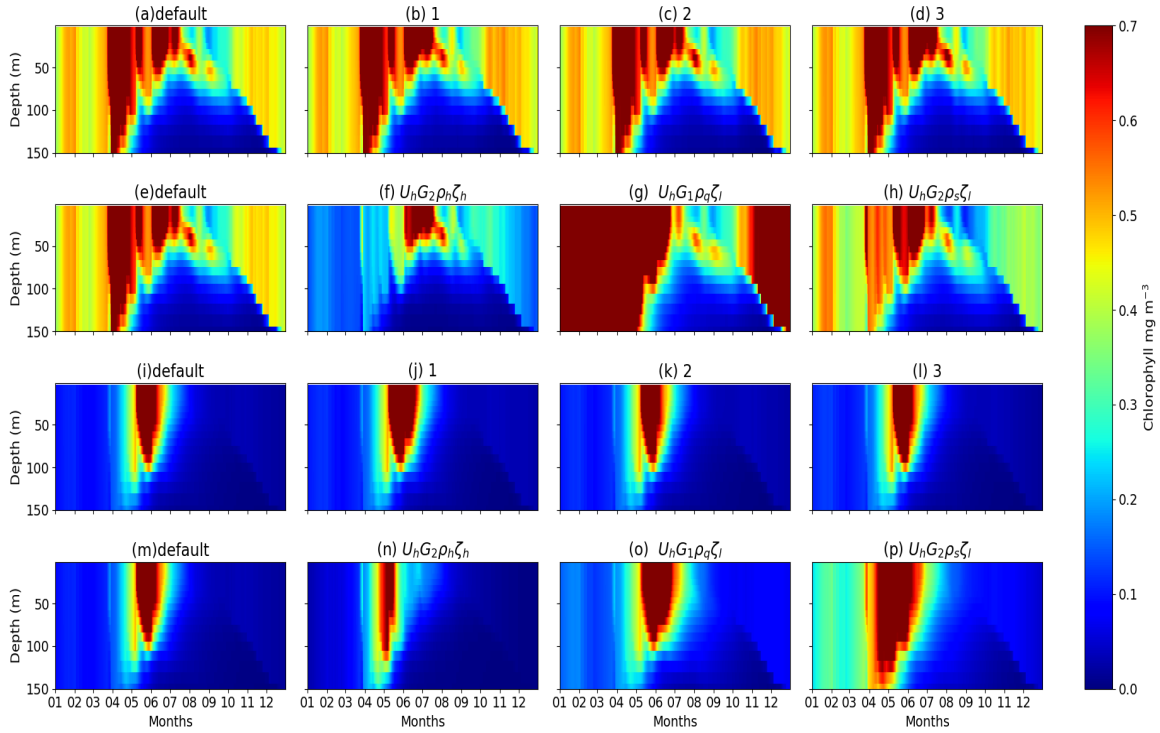


Figure 5.9: Chlorophyll profiles of non-diatom (a to h) and diatoms (i to p) at station PAP form the PPE (a to d for non-diatoms and i to l for diatoms) and PBE (e to h for non-diatoms and m to p for diatoms) members in 2003. Default run chlorophyll distributions are shown in (a), (e), (i), and (m).

When both the biogeochemical process equations and vertical velocity are perturbed together, the PBPE members produce almost similar results to the PBE (see Appendix B for the dominant PFT distributions). Therefore, the difference in PFT distributions shown in Sinha et al. (2010) may vary more sharply across different ensemble members when applying PBE to a 3-D ocean biogeochemical model.

5.4 Signatures and Characteristics in PBE and PPE

5.4.1 Correlations between nutrient, phytoplankton, and zooplankton anomalies

In this subsection, we calculate the anomaly correlations between the different model compartments, such as non-diatoms, diatoms, DIN, silica, mesozooplankton, and microzooplankton. This will give some insights on how different patterns of the ensemble observed in section 5.3 are generated. We choose station PAP and L4 as both stations produces diatoms and therefore it is possible to fully compare between different PFTs and nutrients.

Anti-correlations in PBE and PBPE

From Figures 5.10(a) and (b), perturbing the biogeochemistry produces opposing pattern between DIN and non-diatoms, as well as silica and diatoms, resulting in anti-correlations between nutrients and phytoplankton, as well as phytoplankton and zooplankton. When both the biogeochemistry and physics are perturbed, anti-correlations between model compartments are also observed. During the bloom period, the anti-correlations are often stronger than the non-bloom period, summarised in Table 5.2, for both ensembles. Between the two types of zooplanktons, stronger positive correlations are observed during the bloom period compared to non-bloom period, showing how the two zooplankton types have similar dynamics. However, there are no strong correlations between the two phytoplankton types; at PAP anti-correlations have been shown between diatoms and non-diatoms, but at L4, during the bloom period, weak positive correlations have been produced.

The anti-correlations between nutrients, phytoplankton, and zooplankton occur because perturbing the biogeochemistry changes the way biogeochemical compartments interact with each other, without altering the nutrient concentration in the euphotic zone. For example, one functional forms combination that produces a higher concentration of phytoplankton would result in low nutrients or low zooplankton. Additionally, functional forms that produce high grazing would result in high nutrient and zooplankton concentration. These trade-offs produce the ensemble spread from perturbing the biogeochemistry.

At both stations, negative correlations have been observed between diatoms and non-diatoms, in PBE and PBPE. As we can see from Figures 5.7 and 5.9, non-diatoms usually occur when diatoms concentrations are low, such as during early winter and summer, especially when the ensemble uses G_1 and ζ_l. This makes the correlations between diatoms and non-diatoms negative. However, between zooplankton types, the correlations are positive. This may be because both mesozooplankton and microzooplankton consume non-diatoms and detritus, which may result in similar dynamics. For example, at PAP, there

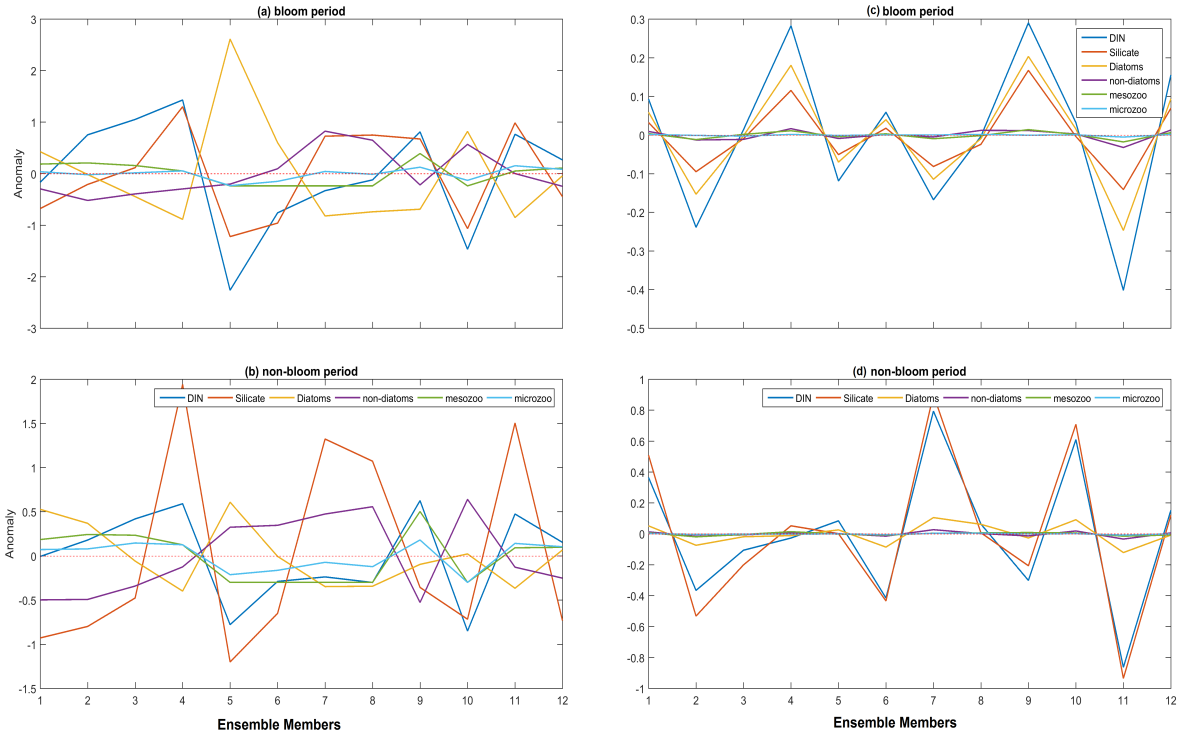


Figure 5.10: Anomalies of different PBE and PPE members in 2005. The lines represented in these panels show surface DIN, silicate, diatom, non-diatom, mesozooplankton, and microzooplankton during bloom and non-bloom periods at L4. Left and right panels show the PBE and PPE anomalies, respectively. The y-axis denotes the ensemble members (see Table 5.2 for the correlation coefficient)

is only one prominent bloom simulated by the ensemble, which makes the correlation between the two zooplankton types stronger during bloom compared to non-bloom period. However at L4, the bloom timings of diatom and non-diatom are different (see Figure 3.23 in Chapter 3 and Smyth et al. (2010)), which makes the correlations less strong during spring bloom (dominated by diatoms), compared to the non-bloom period, where both zooplankton concentrations decrease. The difference in correlations between phytoplankton types may occur due to different half-saturation constants and nutrient types that are consumed by diatoms and non-diatoms.

The results from Table 5.2 also further emphasises that when both biogeochemistry and physics are perturbed, the correlations between anomalies are more similar to PBE than PPE. However, the PBPE often produces weaker correlations compared to the PBE. This might be due to the increasing or decreasing nutrient supply which enhances or reduces the phytoplankton, and therefore zooplankton concentrations, which may weaken the strength of the correlations.

Table 5.2: Correlation table between the anomalies of DIN, silicate (sil), diatom (diat), non-diatoms (non-diat), mesozooplankton (mesozoo), and microzooplankton (micorozo) at station PAP and L4. The correlation are calculated from 1999-2003, for bloom period, and non-bloom period. The noise that is used for PBE and PPE is between -1 and 1. The correlations that are written in bold denotes the bloom period and the non bloom period is written next to it. Italicised correlations are non significant.

Station PAP	Din	Sil	Diat	Non Diat	Mesozoo	Microzoo
PBE						
DIN	1					
Sil	(**-0.24**, *-0.11*)	1				
Diat	(***0.029***, *-0.19*)	(**-0.89**, -0.74)	1			
Non Diat	(**-0.7**, 0.41)	(**0.76**, *0.2*)	(**-0.70**, -0.59)	1		
Mesozoo	(**0.76**, 0.68)	(**-0.055**, *-0.14*)	(***-0.18***, -0.40)	(**-0.44**, 0.58)	1	
Microzoo	(**0.6**, 0.54)	(**-0.21**, -0.48)	(**-0.21**, 0.48)	(**0.12**, -0.65)	(**0.64**, 0.48)	1
PPE						
DIN	1					
Sil	(**0.81**, 0.95)	1				
Diat	(**0.89**, 0.90)	(**0.92**, 0.91)	1			
Non Diat	(**0.88**, 0.91)	(**0.75**, 0.85)	(**0.91**, 0.85)	1		
Mesozoo	(**0.8**, 0.93)	(**0.53**, 0.86)	(**0.74**, 0.89)	(**0.81**, 0.95)	1	
Microzoo	(**0.76**, 0.90)	(**0.78**, 0.81)	(**0.91**, 0.85)	(**0.93**, 0.98)	(**0.68**, 0.97)	1
PBPE						
DIN	1					
Sil	(***0.099***, *0.05*)	1				
Diat	(***-0.16***, *-0.26*)	(**-0.85**, -0.62)	1			
Non Diat	(**-0.40**, *0.047*)	(**0.77**, 0.42)	(**-0.69**, -0.44)	1		
Mesozoo	(**0.67**, 0.67)	(**-0.055**, *-0.061*)	(***-0.21***, -0.47)	(**-0.43**, *0.095*)	1	
Microzoo	(**0.53**, 0.32)	(***-0.18***, -0.44)	(***-0.12***, *0.12*)	(**-0.62**, *-0.13*)	(**0.64**, 0.39)	1
Station L4	Din	Sil	Diat	Non Diat	Mesozoo	Microzoo
PBE						
DIN	1					
Sil	(**0.61**, *0.23*)	1				
Diat	(**-0.67**, *-0.23*)	(**-0.86**, -0.84)	1			
Non Diat	(**-0.56**, -0.65)	(***-0.15***, *0.35*)	(**-0.18**, -0.42)	1		
Mesozoo	(**0.81**, 0.81)	(**0.30**, -0.22)	(**-0.31**, *-0.22*)	(**-0.73**, -0.88)	1	
Microzoo	(**0.73**, 0.9)	(**0.73**, *0.04*)	(**-0.81**, *-0.001*)	(**-0.11**, -0.79)	(**0.63**, 0.92)	1
PPE						
DIN	1					
Sil	(**0.95** ,0.99)	1				
Diat	(**0.99** ,0.97)	(**0.93** ,0.95)	1			
Non Diat	(**0.8** ,0.98)	(**0.78** ,0.97)	(**0.77** ,0.91)	1		
Mesozoo	(**0.99** ,0.98)	(**0.92** ,0.98)	(**0.99** ,0.97)	(**0.79** ,0.96)	1	
Microzoo	(***-0.0042***, 0.96)	(***0.02***, 0.74)	(***-0.046***, 0.7)	(**0.58**, 0.82)	(***0.0043***, 0.67)	1
PBPE						
DIN	1					
Sil	(**0.51**, 0.35)	1				
Diat	(**-0.63**, -0.32)	(**-0.84**, -0.82)	1			
Non Diat	(**-0.45**, -0.38)	(**0.33**, -0.4)	(**-0.28**, *-0.24*)	1		
Mesozoo	(**0.73**, 0.7)	(***0.12***, *-0.23*)	(***-0.28***, *-0.10*)	(**-0.75**, -0.84)	1	
Microzoo	(**0.82**, 0.87)	(**0.68**, *0.10*)	(**-0.78**, *-0.05*)	(***0.19***, -0.61)	(**0.71**, 0.88)	1

Positive correlations in PPE

For PPE, shown in Figure 5.10(c) and (d), it is clear that the anomaly patterns of nutrients, phytoplankton, and zooplankton follow each other, therefore producing strong positive correlations, summarised in Table 5.2. The correlation strength is higher during the non-bloom period, opposite to the PBE anomaly. Strong positive correlations occur because perturbing the vertical velocity also perturbed the availability of nutrients in the euphotic zone, as summarised in Figure 5.1. Therefore if the overall vertical velocity in ensemble results is an upwelling, the nutrient in the euphotic zone increases, which results in the rise of phytoplankton and therefore zooplankton concentrations, making the correlations mostly positive. The difference between how much nutrients is available in the euphotic zone garners the PPE spreads.

The stronger correlation during non-bloom period occurs because, during the bloom period the nutrient concentration is low as the phytoplankton uptake rate is high. During this period, the zooplankton grazing rate is also high. The combination of high grazing and low nutrients would result in the reduction of phytoplankton concentrations. This makes the nutrients, zooplankton, and phytoplankton anomaly correlations more 'negative' during the bloom period compared to non-bloom period. Distinctive correlations during the two periods are shown between microzooplankton and DIN, silicon, diatoms, and mesozooplankton, (shown in Figure 5.2). This may occur because during bloom period, the decline in nutrients coincides with the rise in microzooplankton. The main diet of microzooplankton is non-diatom, therefore the during the diatom bloom, the microzooplankton coincides with the increasing diatoms and mesozooplankton at the later part of the bloom.

5.4.2 Ratios and distribution of nitrogen between different compartments

In section 5.3.1, at the oligotrophic stations, a poor match between in situ DIN and chlorophyll have been observed. Furthermore, from the previous section, from the anomaly, the correlation between nutrients and phytoplankton types are negative, as the interaction between these compartments vary across different ensemble members. In this subsection, the ratio between Chl:DIN, Zoop:Chl, phytoplankton nitrogen and DIN+phytoplankton nitrogen (P:PD) are going to be explored. This is done to ensure whether the ratios from in situ observation have been represented well by the ensemble at different stations. Furthermore, calculations of these ratios are done to ensure that perturbing the biogeochemistry has altered the interaction strength between different model compartments by looking at the proportions between the food and consumer ratio and also how the model currency (nitrogen) is distributed to other compartments.

Surface ratios

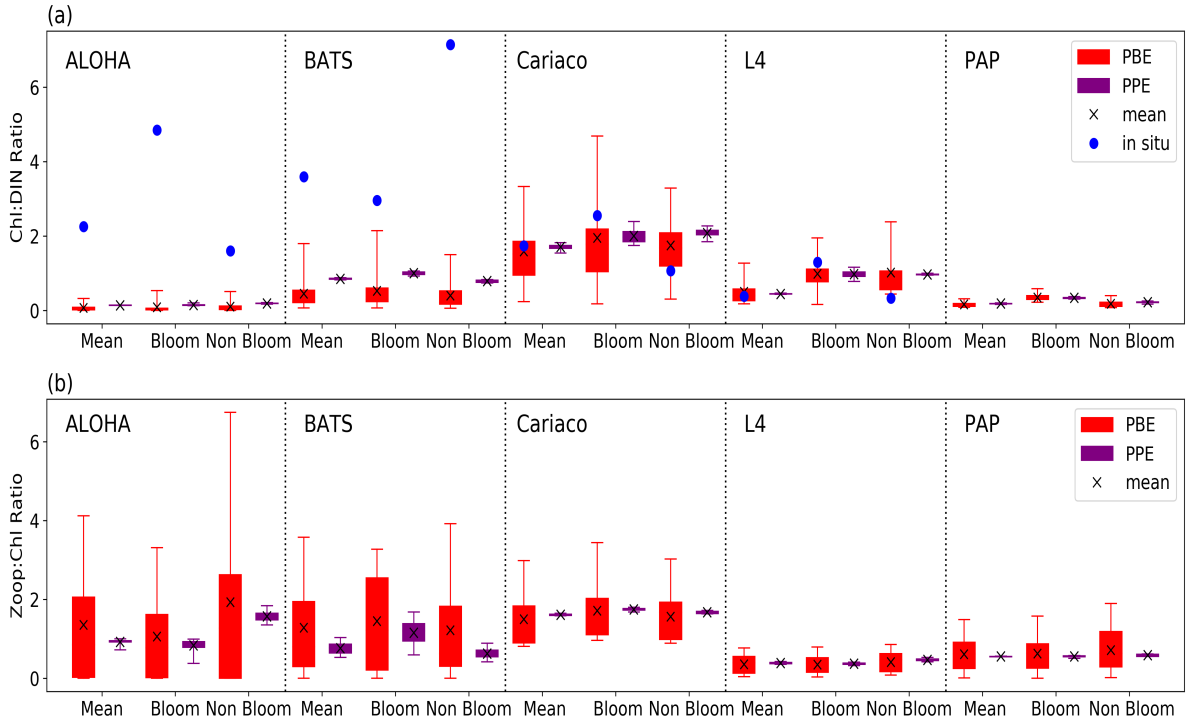


Figure 5.11: Boxplots showing the DIN:Chl and Zoop:Chl between the PPE and PBE calculated from the surface annual mean (from 1st January 1998 to 31st December 2007) and during bloom and non-bloom period. The mean ratio is shown in black cross. The means between PBE and PPE do not show much difference in the coastal and abyssal plain stations.

Figure 5.11 shows the ratios of Chl:DIN and Zoop:Chl from the ensembles in boxplots at all of the stations at the surface. The mean ensemble ratios of Chl:DIN are similar between PBE and PPE at all the stations apart from BATS, where PPE mostly shows a slightly larger mean ratio, Fig. 5.11(a). This may be due to the U_h used in PPE showing relatively high uptake of DIN at the oligotrophic condition, making the Chl to DIN ratio higher in PPE. For Zoop:Chl, the PPE mean ratio is lower than PBE, especially at the oligotrophic regions. The G_1 function used in PPE shows low grazing rate at low phytoplankton concentrations (such as in the oligotrophic stations). This results in lower PPE Zoop:Chl means at oligotrophic stations.

In terms of the range, PBE produces a wider range, Fig. 5.11. The highest Chl:DIN is often similar to the lowest Zoop:Chl. The functional forms that show the maximum ratio varies with the region. At the oligotrophic region, an ensemble member that combines G_1 and ρ_q would yield low grazing and low phytoplankton mortality rate when plankton concentration is low. Combining it with ζ_l would further enhance the chlorophyll concentration, resulting in a maximum Chl:DIN. Consequently, the same ensemble member produces minimum Zoop:Chl. Combining low uptake rate U_s with high grazing rate G_2, then combining these functions with relatively high phytoplankton mortality ρ_h at low concentration,

would result in low Chl:DIN ratio, but high Zoop:Chl ratio.

At the coastal stations, nutrients and therefore phytoplankton are more plentiful. However there are some differences between Cariaco and L4, for example, the chlorophyll mean at L4 is higher compared to Cariaco (ensemble mean= 1.96 and 1.05 mg m^{-3} respectively), and the opposite for DIN (ensemble mean= 0.628 and 4.531 mmol m^{-3} for L4 and Cariaco, respectively) and zooplankton (mean=1.156 mmol m^{-3} and 0.548 mmol m^{-3} for Cariaco and L4, respectively). Furthermore, the seasonality at Cariaco ensemble is not as strong as that in the light-limited L4. Therefore these two stations would have different functional form combinations for the maximum and minimum ratios. At higher phytoplankton concentrations (>1.5 mg m^{-3}), G_2 produced a lower grazing rate. Combining this function with low phytoplankton mortality rate such as ρ_s would yield high Chl:DIN ratio. The minimum Chl:DIN is produced when the ensemble contains ζ_q, as low zooplankton mortality would produce low chlorophyll concentrations. Similar annual maximum and minimum Chl:DIN have been observed for both coastal stations. However, at L4 during the non-bloom period, where phytoplankton is low, using ρ_s and G_1 would result in low mortality and grazing rate, respectively, resulting in maximum Chl:DIN. Changing the functions to ρ_h and G_2 produces minimum Chl:DIN during the non-bloom period.

At station PAP, similar to L4, different ensemble members produce maximum Chl:DIN in different periods. During non-bloom and in the annual means, the maximum Chl:DIN is produced from similar ensemble members as the oligotrophic stations. However, during bloom time, the ensemble member is similar to that in the coastal stations, as the chlorophyll concentrations are high. The minimum Chl:DIN ensemble member is similar to the coastal station for all the seasons and annual mean. In terms of Zoop:Chl, the maximum and minimum ratios are similar to station Cariaco. These discussions further show that the changes in Zoop:Chl or Chl:DIN can be traced back to determine which functional forms are responsible for the changes. By using the ensemble approach, this method may be implemented when investigating the change in the observed ratios.

Although the ranges obtained by the PBE are large, at the oligotrophic stations, the in situ show almost a magnitude higher Chl:DIN ratio. The underestimation of the ratio further shows the overestimation of DIN at both stations, and therefore the nutrient uptake inefficiency in MEDUSA, despite perturbing the functional forms. At in situ ALOHA, the highest Chl:DIN ratio is observed during the bloom period, but at BATS the highest ratio is produced during the non-bloom period. These differences may occur due to during bloom period both DIN and chlorophyll are at their highest concentrations, but during the non-bloom period the mean in situ chlorophyll is higher than DIN (0.05 mg m^{-3} and 0.008 mmol m^{-3} for in situ ALOHA). This relatively high chlorophyll is due to the occurrence of diazotroph or nanophytoplankton in the summer where the nutrient is low (White et al., 2015). At the coastal stations the Chl:DIN ratio from the in situ observations are within the ensemble range, especially in the annual

means. During the bloom period, both in situ ratio shows lower Chl:DIN ratio compared to non-bloom, even at Cariaco, where the ensemble does not produce distinct seasonality compared to the observation. This approach may be used as new metrics to show how well ensemble results of a model could represent the in situ conditions of a particular biogeographical region.

At some of the stations, the zooplankton to chlorophyll ratios do not show the typical bloom and non-bloom ratio. During the bloom timing, the zooplankton:chlorophyll ratio should be small, as the chlorophyll is at its highest concentration, and zooplankton concentration has not caught up with the rise of chlorophyll, and the opposite during non-bloom period. This pattern is apparent at BATS, L4, and PAP, but not at ALOHA and Cariaco, where the opposite has been observed Fig.5.11(b). This may occur because, at ALOHA and Cariaco, the bloom and non-bloom periods produced by the model are not very distinct.

Phytoplankton to total nitrogen ratios in the surface

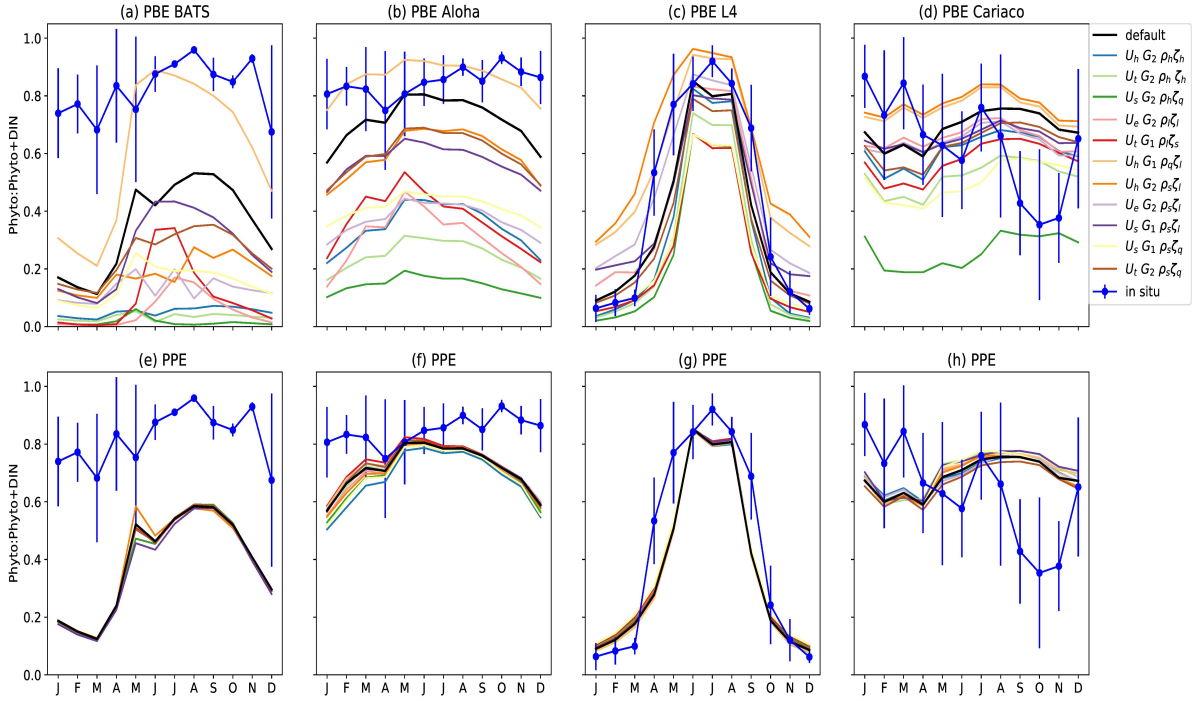


Figure 5.12: Monthly averaged phytoplankton nitrogen proportion from phytoplankton + DIN (P:PD) at four oceanographic stations on the surface. The top and bottom panel show the phytoplankton proportion from PBE and PPE, respectively. Each lines represent the ensemble members. The observation is shown in blue and the error bars are derived from the standard deviations of the monthly P:PD. The nitrogen from phytoplankton is calculated using the chlorophyll to nitrogen ratio from the default run, and the calculation is described in Yool et al. (2011).

Similar to Figure 5.11, Figure 5.12(a) and (b) emphasises that the proportion of nitrogen in phytoplank-

ton is seen to vary sharply from member to member in the PBE, especially at oligotrophic stations where the ordering of members is also seen to change. This occurs because the strength of mortality and grazing varies depending on the steepness of the curve of the functional forms, which dictates how each ensemble member distributes nitrogen at different times. For example in Figure 5.12(a), between ensemble members with trigonometric uptake, Holling type III grazing, linear phytoplankton, and sigmoidal zooplankton mortality ($U_t G_1 \rho_l \zeta_s$) and sigmoidal uptake, Holling type III, sigmoidal phytoplankton, and quadratic zooplankton mortality ($U_s G_1 \rho_s \zeta_q$), overlaps in June. This is because sigmoidal mortality shows a lower rate at low phytoplankton concentrations, but the rate increases rapidly when the concentration increases, and linear phytoplankton mortality always produces high mortality rate, making the phytoplankton proportion from $U_t G_1 \rho_l \zeta_s$ decreases more rapidly at low phytoplankton proportions. The crossing of members at station BATS may be caused by the DIN concentration that coincides with when the uptake rate functional forms cross with each other, meaning that some functional forms would have a more rapid increase in uptake rate than others (see Figure 2.3 in Chapter 2).

In contrast, the PPE shows very little variability in nitrogen fractions across the whole ensemble, summarised in Figure 5.12(e) and (f). At the oligotrophic stations, the observed phytoplankton nitrogen fractions are always > 0.7, indicating that most of the time, this region is nutrient-limited. Similar to Figure 5.11, the ensemble underestimated the phytoplankton fraction, despite generating a large range. However there are two ensemble members can match the phytoplankton fractions: the default run and $U_h G_1 \rho_h \zeta_l$ (see Figure. 5.12 (b)). This is because the hyperbolic uptake function has a higher nutrient uptake rate at low nutrient concentrations, compared to other functional forms.

Furthermore, the underestimation of the phytoplankton fraction may be due to the half-saturation constant being too high for oligotrophic regions. An earlier experiment shows the half-saturation constant for a rectangular hyperbolic formula at the oligotrophic region is $\sim$0.24 mmol m^{-3} (Laws, 2013), making the nutrient uptake functional form for default MEDUSA (0.5 mmol m^{-3} for DIN) is inefficient in nutrient-limited regions. This may also explain the overestimation of DIN in the oligotrophic regions.

The spread obtained by PBE at the Coastal region, Figure 5.12(c) and (d), is enough to encompass the in situ observations and its standard deviations, especially in seasonally strong regions such as L4 (Figure 5.12(c))). In this station, light is the limiting factor for most of the month, which also controls the nutrient uptake rate, which makes the PBE and PPE follow the pattern of the observation, even when nutrients becomes limiting in the summer. In different months, the in situ phytoplankton fraction falls on different ensemble members of PBE, (e.g. from October-March, the in situ phytoplankton fraction generally falls within ensemble members that produces overall low phytoplankton growth rate at low concentrations (such as $U_t G_2 \rho_l \zeta_s$ and $U_h G_2 \rho_h \zeta_h$) and from April to September, falls within members that shows high phytoplankton growth rate and high zooplankton loss (such as $U_h G_1 \rho_s \zeta_l$ and $U_h G_2 \rho_q \zeta_l$). This follows the

North Atlantic bloom studies, where the phytoplankton proportions and growth rates change over the year (Behrenfeld et al., 2013; Behrenfeld and Boss, 2014; Roy et al., 2012), as it is controlled by nutrients, light, and mixed layer conditions, e.g. in the summer, the growth rate of phytoplankton is in equilibrium with loss rate as nutrient depletes and grazing rate is high (Behrenfeld et al., 2013; Behrenfeld and Boss, 2014).

These results suggest that the PBE generates enough spread to encompass the uncertainty within the observed phytoplankton fraction if the region is seasonally strong, and also explains the variations of growth and loss rate in phytoplankton. We can also see that none of the single PBE members fully captured the observation throughout the year, therefore using a single set of functional forms, is not sufficient to capture in situ and its uncertainty. This also emphasises that the PPE, cannot capture the variability in phytoplankton fractions. As discussed above, ensemble members that captured the in situ fractions are time-dependent, as the ensemble members behave differently depending on the concentration of nutrients, phytoplankton, and zooplankton, especially in seasonally strong regions. Therefore, for further applications such as a probabilistic modelling approach the weighting of uptake, grazing, and mortality process functions can be dependent on time.

5.5 Summary

We have run three different ensembles from 1-D MEDUSA, which are generated from perturbing the biology (PBE) and physics (PPE). The ensemble spread, chlorophyll distributions, and characteristics of the ensembles are explored. The PBE and PBPE generally produce a larger spread of the chlorophyll annual means compared to PPE and encompass the in situ concentrations. Below the biologically active region, the PPE produces larger spread than PBE and PBPE, as below this depth there are less biological activities and nutrient supply is dependent on the PPE. For the chlorophyll distributions, we used the time evolution of DCM as the ensemble metrics and it shows that across different ensemble members, the PBE and PBPE produces a larger spread of DCM depths compared to PPE, with different chlorophyll patterns. This is because the PBE produces more variable loss and growth rate of phytoplankton at different nutrient supply. This makes perturbing the biogeochemistry produces a stronger effect than physics when both factors are perturbed. Perturbing the biogeochemistry can also affect the distribution of diatoms and non-diatoms in the water column. Additionally, at a station where vertical velocity is strong, such as L4, perturbing the vertical velocity can change the distribution of the dominant phytoplankton type in the water column.

The characteristics between PBE and PPE are examined using the correlations between the ensemble

members and the proportion of phytoplankton nitrogen. Within the PPE members, the biogeochemical compartments tend to vary together with strongly positive correlations with nutrient supply, whereas, in PBE, the nutrients and phytoplankton have negative correlations, as perturbing the biogeochemistry changes the interaction strength of phytoplankton, zooplankton, and nutrients at different concentrations. Across different members, the PBE produces a larger spread of phytoplankton nitrogen proportions at different times and therefore larger ensemble spread than the PPE. This shows that the PBE produces a better representation of uncertainty in biogeochemical variables. The metric also shows that where phytoplankton is light-limited, the functional forms generally perform better, as the observed phytoplankton proportions show similar patterns to the PBE and PPE. Ensemble members that captured the in situ fractions are time-dependent, as the ensemble members behave differently depending on the concentration of nutrient, phytoplankton, and zooplankton, especially in seasonally strong regions. Therefore, for applications in probabilistic modelling the weighting of uptake, grazing, and mortality process functions is time dependent.

Chapter 6

General Discussion

6.1 Summary

In this thesis, we have presented a method in generating a perturbed biogeochemisty ensemble, generated by altering the biogeochemical process equations, which have been tuned to the default functional forms. It has been shown that altering the model process equation can alter model dynamics in a predator and prey model (Fussmann and Blasius, 2005; Flora et al., 2011; Adamson and Morozov, 2013) more than varying the parameters. We implement this method first in a 0D NPZ model generated by Edwards and Brindley (1996). The results show that altering the biogeochemical process formulations can change the model dynamics of the NPZ model, especially when the grazing function is altered from sigmoidal to hyperbolic-type functions. Sigmoidal grazing functions (such as Holling type III and Sigmoidal II) produce the highest phytoplankton concentrations, with the longest period of oscillations and significantly higher amplitudes because after tuning, the maximum grazing rate is lower than other hyperbolic functions and lower grazing rate when phytoplankton concentration is > 0.1 gC L^{-1}.

Further, we examine whether structural sensitivity also applies to a more complex model. Indeed, in an earlier study by Anderson et al. (2010); Yool et al. (2011), altering the grazing and mortality function can produce different bulk properties of chlorophyll. However, these studies do not thoroughly examine when core biogeochemical processes, such as phytoplankton nutrient uptake, zooplankton grazing, and plankton mortality, are perturbed together, or individually. In this study, we choose MEDUSA, an intermediately complex operational biogeochemical model which is also the ocean biogeochemical component of the UK-ESM1. MEDUSA is essentially an NPZD model, with three nutrient compartments (DIN, Silica, and Iron), two phytoplankton (non-diatoms and diatoms), and two zooplankton partitioned by size (microzooplankton and mesozooplankton), shown in Figures 2.1a and 2.1b. Therefore the un-

derlying biogeochemical process is similar to that of an NPZD model, such as phytoplankton nutrient uptake, zooplankton grazing, and plankton mortality.

The ensemble is generated by altering the key biogeochemical processes default functions of MEDUSA with alternative functional forms; four nutrient uptakes, two zooplankton grazings, and four phytoplankton and zooplankton mortalities equations, totalling in 128 ensemble members, which have been described in Chapter 2. These functional forms are chosen as they have been used in other biological models and also are phenomenologically similar. For example, in the default MEDUSA, zooplankton grazing is modelled using the Holling Type III function, which has a saturating response, with a maximum grazing rate and a half-saturation coefficient. Another common functional form that is used in a biogeochemical model is the Holling type II, which has a similar saturating curve. In order to minimise computational cost, we used a 1-D version of the MEDUSA model that is provided by the software MarMOT. The 1-D run is then compared with five different stations which represent various biogeographical regions; ALOHA and BATS for oligotrophic (or trade winds), Cariaco and L4 for coastal, and PAP for the abyssal plain region at the North Atlantic, which can be grouped with the westerlies (Longhurst et al., 1995) or the subarctic gyre (Ducklow, 2003).

Using statistical metrics (such as range, correlation, RMSE, and bias) to compare chlorophyll and DIN from the model output with in situ observations as well as phytoplankton phenology to examine the model dynamics, the ensemble has shown to capture most of the in situ and satellite-derived chlorophyll concentration, which are explained in Chapter 3. The ensemble is also likely to produce a mean field closer to the observation than a single structure model that has not been tuned to one station. The widest and smallest range are observed in the coastal and oligotrophic stations, respectively. Among the five stations, station ALOHA has shown the most reliable spread (for the surface and integrated chlorophyll and bloom peak height), as well as station L4 (for inter-annual mean) according to the NRR. Changing Holling Type III grazing to Holling Type II lowers the chlorophyll concentrations, especially at the oligotrophic stations, with the nutrients responding in the opposite direction. Pairing Holling Type II with linear phytoplankton mortality will reduce the chlorophyll concentration even further. For the nutrient uptake, exponential and sigmoidal forms show inefficient phytoplankton uptake, which results in low chlorophyll concentrations and high nutrient concentration. From this study, we can see that some of the sensitivity that have been observed in the simple NPZ model has been carried on to the more complex model, altering phytoplankton mortality in intermediately complex model affecting the chlorophyll dynamics more.

To examine how structural sensitivity affects model dynamics, we use phytoplankton phenology as a model metric. The difference of the phenological timing between the ensemble mean/median and the default run ranges from a couple of days to weeks. The timing of initiation, bloom peak, and termination

also show wide ensemble inter-quartile ranges for all stations which can lie between ∼20 and 100 days earlier than the in situ timing, making the observed phenology metrics fall within the full ensemble range. The in situ peak and annual means are also mostly captured by the ensemble. However, at the coastal stations, the phenology timings are not well captured. These discrepancies may be due to the poor representation of the physical environment at Cariaco, where the ensemble does not simulate upwellings of nutrients. Furthermore, the way zooplankton select their prey in the model, does not take account of the stoichiometry of C:N, making the double peaks in L4 not simulated well (Sailley et al., 2014). From these results, given that the physical oceanography aspect is well modelled, and the equations are theoretically justified, it is possible to use the ensemble approach to forecast these phenological aspects. Furthermore, it may also be possible to improve the accuracy of the ensemble spread by systematically removing ensemble members, which also reduces the computational cost.

We then systematically reduced the ensemble by using PCA and k-means cluster to select the members that are most representative of the full ensemble, in Chapter 4. The 12 members ensemble, including the default run, have been selected, and the reduced ensemble has retained most of the spread, RMSEs and biases from the original run. A perturbed physics ensemble has also been generated, to compare the spread between the reduced ensemble with the perturbed physics ensemble (PPE), where vertical velocity is varied using different noise strength. This is done because the physical input to the biogeochemical model has been shown to affect the chlorophyll distribution more than perturbing the model complexity (Friedrichs et al., 2006; Raghukumar et al., 2015). Additionally, to see which perturbation affects the model results more, the reduced ensemble is run using the perturbed vertical velocity ensemble, resulting in a perturbed biology physics ensemble (PBPE).

When the physical oceanography is perturbed, the higher the perturbation strength, the larger the range it would garner as vertical velocity controls the availability of nutrients for phytoplankton. Using the highest perturbation strength, some ensemble members produced an increasing DIN concentration over the time series due to a more intense upwelling. Additionally, this increase in surface DIN and chlorophyll concentration has also been shown in the spin-up run in Chapter 2, which suggests that the vertical velocity has a tendency to increase the surface DIN over the time series. However, perturbing only vertical velocity produces less range compared to perturbing the biogeochemistry, and therefore not garnering enough spread to capture to the in situ concentration. Furthermore, when both the biogeochemistry and vertical velocity are perturbed together, the spread becomes larger than only perturbing the biogeochemistry, although the pattern of the chlorophyll follows that of the perturbed biogeochemistry ensemble (PBE). Perturbing both biogeochemistry and physics also produce better bias, correlation, and RMSE at stations where high chlorophyll concentrations are expected. In the PBE and PBPE, opposing trends of interannual range between DIN and chlorophyll have been observed.

Upon further investigation in Chapter 5, the narrow DIN range produce by PPE is observed only at the biologically active region (up to ∼75m), but below this depth, the PPE range for DIN is larger than PBE and PBPE. But for chlorophyll and zooplankton, the PBE and PBPE ranges are still larger than PPE. To investigate which perturbation affects chlorophyll distribution, the time evolution of deep chlorophyll maxima (DCM) that has been observed in oligotrophic regions such as BATS and ALOHA, is used as an ensemble metric. The PBE members produced various chlorophyll depth distribution and consequently more variable maximum DCM depth (see Figure 5.5 and Table 5.1) compared to perturbing the physics, with the PBPE producing a similar pattern to the PBE. In order to understand how the perturbed biogeochemistry garners its range, the characteristics of both PBE and PPE ensembles are examined using the anomalies between the NPZ compartments.

Within the PPE members, the biogeochemical compartments (nutrient, zooplankton, and phytoplankton) tend to vary together with strongly positive correlations especially during the non-bloom period, whereas for PBE the correlation between nutrients and phytoplankton is negative, especially during the bloom period. This is because perturbing the biogeochemistry changes the interaction strength of phytoplankton, zooplankton, and nutrients, but perturbing the physics only changes how much nutrients are available in the euphotic zone. Furthermore, during the bloom period, the nutrient supply is utilised more by phytoplankton, making the anti-correlation for PBE stronger, but correlation for PPE weaker. This further shows that PBE produces a better representation of uncertainty in biogeochemical variables. Nevertheless, the ensemble generally performs better in light-limited region, as it shows a similar pattern to the in situ, and always captures the in situ phytoplankton nitrogen proportion. Because ensemble members are behaving differently depending on the availability of light, the concentration of nutrients, phytoplankton and zooplankton, ensemble members that captured the in situ fractions are time-dependent, especially at light-limited regions.

In the next sections, the effect of altering the functional forms and the ensemble spread between the two different perturbations will be discussed in the next two sections.

6.2 Altering the functional forms

Altering the functional forms of MEDUSA such as zooplankton grazing and mortality have been done previously (Steele and Henderson, 1992; Edwards and Yool, 2000; Anderson et al., 2010). Our study includes perturbing both the mortality and grazing, as well as the nutrient uptake function. Generally functions whose curves show the most distinctive shapes compared to the default run formulation would deviate the most. For example, as discussed in Chapter 3, using sigmoidal uptake at the oligotrophic sta-

tions produces less chlorophyll (for example at ALOHA, the mean surface chlorophyll is 0.085 mg m^{-3} from U_s) compared to other functional forms (mean surface chlorophyll from U_h, U_t, and U_e are > 0.098 mg m^{-3} at station ALOHA), because the sigmoidal function produces a lower uptake rate compared to the default uptake at low nutrient concentration and other uptake functions. However, phytoplankton uptake functions that are used here are not sufficient for the oligotrophic regions, as in situ nutrients and chlorophyll are almost always overestimated and underestimated, respectively, by the ensemble. This is because the phytoplankton in the oligotrophic regions are more adapted to low nutrients, and therefore having a lower half saturation constant than MEDUSA (Laws, 2013). In order to capture the nutrient uptake process better, it is also possible to use optimal uptake kinetics, where phytoplankton can acclimatise to the DIN concentration in the ocean (Smith et al., 2009, 2016a). However, using our current method of optimisation, it is impossible to apply optimal uptake kinetics into the ensemble, as the half saturation constant is dependant on other parameters such as maximum nutrient affinity.

Although altering the phytoplankton nutrient uptake can affect the chlorophyll and DIN concentrations, altering the zooplankton mortality to linear affects the model dynamics and results even more, where the chlorophyll concentration deviates more from the default run compared to using the quadratic function (for example chlorophyll concentration for linear, quadratic, and default are 1.02, 0.68, and 0.70 mg m^{-3} respectively at station PAP). These large deviations have also been produced in the previous study, which examines how changing zooplankton mortality functional forms can affect an NPZ model dynamics (Steele and Henderson, 1992; Edwards and Yool, 2000). As discussed in Chapter 3, when linear phytoplankton mortality is used, low chlorophyll concentrations have also been observed (for example at station PAP phytoplankton linear, quadratic, and default mortality chlorophyll concentrations are 0.52, 0.78, and 0.70 mg m^{-3}, respectively). We have also shown that to some extent, structural sensitivity captures the parametric sensitivity as well, because when the mortality rate is halved from the optimised parameter (summarised in Table 2.1), the chlorophyll concentration does not differ much compared to the default function.

Linear is the most distinctive of all the functional forms that were used to describe plankton mortality, and it produces oscillations in the NPZ model (Steele and Henderson, 1992). As shown in both the MEDUSA and NPZ model, the linear function does show some discrepancy compared to the default run. However, in Chapter 3, when the maximum mortality rate for the linear function is halved, the mean chlorophyll becomes closer to the default run. This shows that changing the parameter value can make the ensemble member result more similar to the default run, and therefore 'hiding the underlying model problems' (Anderson and Mitra, 2010). However, using the current parameter, linear mortality increases the range, which has been shown in Chapter 4. When the principal component scores are grouped using k-means cluster, there are a few clusters where all members contain linear zooplankton mortality, which made

some of the ensemble members in the reduced ensemble still contain some linear mortality functions. If all ensemble members that contain linear functions are removed, the spread becomes ∼20-25% narrower (for example, the full ensemble range of mean surface chlorophyll at ALOHA= 0.127 mg m^{-3}, no linear ensemble= 0.098 mg m^{-3}).

It has been emphasised that the correct formulation is essential in a deterministic model (Flynn and Mitra, 2016; Anderson, 2019). Yet, this study shows that we are not supposed to lean on only one set of functional forms. As seen in Chapter 5, none of the single functional forms can capture all the biological details at a different place and time. Therefore we cannot rely on one functional form (regional versus temporal modelling), which is why we have not given preference to one set of functional forms. However, it is still important for the equations that are used in the ensemble to at least correctly capture the mechanism of grazing or mortality, since simplicity in the equation is a virtue, but only when it is meaningful and can be applied to further our understanding of ocean ecosystem (Anderson, 2019). Furthermore, perturbing the equation is not enough. As stated previously in Chapter 1, limit cycle oscillation is more common in the parameter space to occur in the linear, compared to quadratic, hyperbolic, or sigmoidal mortality (Edwards and Yool, 2000). Therefore before generating a perturbed biogeochemistry ensemble, it is important to taken into account the parameter values for linear phytoplankton and zooplankton mortalities.

Another process that affects the model dynamics in MEDUSA when perturbed is the zooplankton grazing, where Holling type II (Class I, G_2) shows significantly lower chlorophyll concentration than the sigmoid zooplankton grazing, especially in the oligotrophic regions, where the chlorophyll at station ALOHA from Holling type II is less than half of that in Holling type III (Class III, G_1). Our study concurs with Anderson et al. (2010), which has shown that altering zooplankton grazing between two classes in a more complex model have altered both the bulk properties of chlorophyll as well as the distribution of the phytoplankton type. These classes have been discussed in Gentleman et al. (2003), in which various zooplankton grazing functional forms are partitioned into three different classes. In this study we used two classes: class I, which describe no switching of prey, and class II, where the prey switching is done passively due to density dependent behaviours. Because of these different descriptions of formulations, it is expected that the two grazing functions would produce different model results. However, unlike the previous study by Anderson et al. (2010), where using G_1 grazing would result in lower diatoms during spring time than G_2. In our study, using G_1 produces a higher concentration during spring time than G_2 (surface diatoms chlorophyll means for G_1 and G_2 are 1.16 and 1.09 mg m^{-3}, respectively, at station L4). The two grazing functions also produced two different bloom timing for the total chlorophyll concentrations, with G_1 generally producing an earlier bloom than G_2 (see Figures 5.7 and 5.9 subfigures (e),(f),(m),(n)), which shows that the model dynamics have changed. However both studies agrees that

the total chlorophyll from G_1 is higher than G_2, and mesozooplankton concentration from G_1 is higher than G_2 whilst the opposite is produced for microzooplankton. These differences between the two studies may be caused by the parameter assignment in Anderson et al. (2010) and MEDUSA, and the fact that there are three different phytoplankton groups in PlankTOM5.2, and only two in MEDUSA-1.0. Furthermore, this study only changes the half saturation constant in order to ensure the similarities between functional forms and is run in 1-D with only two stations, located in the NADR and NECS, that produces the diatom. In Anderson et al. (2010), all parameters are kept the same for all functional forms and the model is run in 3-D, making the mean concentration of diatoms and mixed phytoplankton more disperse, as there are more regions that are averaged.

It has been demonstrated by Gentleman et al. (2003), that changing parameters can influence the assumed dynamics, meaning that sensitivity to parameter values indicates sensitivity to behavioural assumptions that are unsupported. Indeed, equations such as Michaelis-Menten and the sigmoidal are not theoretically justified to describe multiple grazing uptake (Flynn and Mitra, 2016). In Disk equation, the handling time h, and successful attack rate a, are explicitly represented in the equation and assumes that the zooplankton attack and handle only one resource at a time, and that these rates are resource dependent. Meanwhile in the Michaelis-Menten type equation, instead of attack and handing rates, there are maximum grazing, half saturation constant, and the weight of the prey preference. This makes the Michaelis-Menten type equation overparameterised (Gentleman, 2002) and also makes the handling rate similar for all prey types, because in Michaelis-Menten type grazing equation the maximum grazing rate, g_{max} is equal to $\frac{1}{h}$. Additionally the passive selection in sigmoidal or even the Disk equation itself will lead to nonoptimal intake when the multiple resources are of a different quality (Gentleman et al., 2003), making all the common biogeochemical models susceptible to structural sensitivity. Again, this shows that we are not supposed to depend on one set of functional forms, especially for grazing where changing the process equation affects the biogeochemical model dynamics the most.

Indeed, making the grazing function more realistic has been proven to improve model-data misfits (Sailley et al., 2014), however in this thesis, we look for equations that have a similar curve and parameters, and also examine how structural sensitivity can affect the model dynamics. Therefore, from our ensemble results, similar to the pilot study using a simple NPZ model, it is clear that MEDUSA is most structurally sensitive to the plankton mortalities and the grazing function, especially when two different classes are used alternatively. This is perhaps due to the similar reason in the pilot study where these equations are present in almost all of the model compartments, and how G_2 and linear mortality have significantly different rates, especially when phytoplankton abundance is low, compared to other functional forms. Varying the parameter values that control the rates of the NPZ processes may also garner better ensemble range that captures in situ observation, especially at the oligotrophic stations, as

parameter values are not portable (Ward et al., 2013). In this study, the parameters are chosen in order to ensure the phenomenological similarity to the default MEDUSA run and emphasise how functional forms with similar shape of function can garner a large range. It is also expected in this study that the ensemble will generate a large range, because although these functions are generated from experiments (Laws, 2013; Gentleman et al., 2003), these functional forms are simplified, and therefore more prone to structural sensitivity than functions that explicitly represent a chemostat-like environment (Aldebert et al., 2018). However, the ensemble range garnered from these uncertainties have shown to capture the in situ observations, especially at the coastal and light limited stations, where the standard deviation has also been mostly captured by the ensemble (see Figure 5.12). This shows that without altering much of the parameter values, the ensemble can capture the in situ observations. Furthermore, implementing a more accurate zooplankton grazing function where prey selectivity is done actively, such as the one that is described in Sailley et al. (2014), would introduce more parameter uncertainty, as these equations have more parameters than the classic Holling type II and III grazing function.

Although some of the process functional forms in MEDUSA are questionable (Kwiatkowski et al., 2014), as discussed in Chapter 5, altering the process equations has produced a large spread because it also alters the distribution of nitrogen, which is the model currency, to other compartments, namely phytoplankton, zooplankton, and DIN. At the coastal stations, the standard deviation of the observations are also within the ensemble spread, as shown in Figure 5.12. It also suggests that using a single set of functional forms is not sufficient to capture in situ and its uncertainty. In the next section we will further discuss the ensemble spread, both when perturbing the biogeochemical process equation, as well as physical input.

6.3 Ensemble spread: perturbed biogeochemistry and physics

In this study, we are not looking for the best set of functional forms that can reproduce the in situ observations, but we take into account the whole ensemble range. Perturbing the biogeochemistry has been shown to encompass the in situ observation, both the phytoplankton bloom phenology and chlorophyll concentration. We use the normalised RMSE ratio (NRR) which indicates how large the spread of the ensemble is. The NRR from the full ensemble indicates a narrow ensemble apart from the surface and integrated chlorophyll at station ALOHA, and some of the phenology metrics at various stations (see Table 3.2). This shows that although the in situ are captured, the ensemble might produce more members which show lower concentrations than in situ, therefore making the ensemble too narrow. From Chapter 3, it has been shown that the default run is mostly outside the interquartile range of the ensemble, or close to the upper quartiles of chlorophyll or lower quartiles of DIN, in seasonal, interannual, annual, and the depth profiles, instead of the ensemble mean. This is because the default run produces higher chloro-

phyll and lower DIN, especially in the oligotrophic regions. Furthermore, from Table 3.1, the bias from the default run is often lower than the ensemble mean and median, which makes the ensemble narrow, according to the NRR. However these metrics depend on the ratio of the time averaged RMSE of the ensemble mean to the mean RMSE of the ensemble members, making the NRR contain the bias information from the ensemble members. If the bias is removed from the NRR, the NRR is still narrow. For example at station L4, the NRR for surface chlorophyll is 1.21, but when the model bias is removed from the ensemble, NRR becomes lower (1.15), and adding errors narrows the ensemble even more (NRR= 1.30). Therefore the NRR itself is not a very good metric to measure whether the ensemble has a good spread to encompass the in situ, especially if the model contains high bias, such as the 1-D MEDUSA.

This also further shows that some of the ensemble members produce almost similar results, for example ensemble members that combine sigmoidal uptake, with any nonlinear zooplankton mortality, and G_2 would produce similar mean surface chlorophyll results (0.033 mg m^{-3} at station ALOHA), despite the different phytoplankton mortality. Almost similar ensemble members results are expected as changing some functional forms, such as zooplankton grazing and altering plankton mortality to linear can affect the model dynamics more than other functions, which makes altering other functions, such as changing phytoplankton mortality from default to quadratic, will not affect the model results significantly. This has been shown on the scores plot in Chapter 4, Figure 4.1, where scores of RMSEs from different members are clustered together. Therefore it is necessary to reduce the ensemble, by selecting ensemble members that are representative of each cluster. Furthermore a 3-D marine biogeochemical model will require more computational power than the current 1-D model to run, making this step necessary in order to apply the PBE to a 3-D model. The reduced PBE have shown to mostly retain the properties of the full ensemble, and even produced better correlations, RMSE, and bias, compared to the full ensemble. However, in terms of the ensemble range, the reduced PBE has a slightly lower range compared to the full ensemble, especially at station PAP. This lower range causes the ensemble to not capture some of the in situ concentrations, (see Figure 4.11).

Compared to perturbing the vertical velocity - where essentially the physical oceanography input is perturbed and the model is run using the default functional forms, the reduced ensemble still mostly produced a higher range - regardless of the perturbation strength, especially in chlorophyll. The PPE members mostly follow the pattern of the default run, with a slight change in the NPZ concentrations. The default model uses functional forms that are more optimal for phytoplankton growth, which makes the default run often produces better bias and correlations at oligotrophic stations, but at regions where chlorophyll concentration is high (such as the coastal and abyssal plain stations), the chlorophyll bias is higher in the default run than the ensemble mean and median. As the default run is showing more optimal nutrient uptake, on its own, the DIN biases in the default run are lower than the ensemble mean,

making the correlation and bias of the default run often improved by the ensemble mean from perturbing the physics using small or medium noise. With a high perturbation strength, at certain stations, the DIN range is higher than perturbing the biogeochemistry (see Figure 4.9(a), (b), and (e)). The high DIN range is produced because the higher the perturbation strength, the higher the vertical velocity range. The range of DIN concentrations increase strongly at most of the stations over the years. These are caused by some ensemble members producing an overall increase in the DIN interannual mean due to the overall upwelling from vertical velocity, whilst the majority of the ensemble members show insignificant increase in DIN interannual mean (or the majority of the ensemble members show an insignificant decrease in DIN whilst some show a significant increase), as shown in Figures 4.10 and 4.13. This in turns produce an increasing trend of DIN range, especially when the perturbation strength is high.

The vertical velocity controls the nutrient supply, and as shown in section 2.5.3, the original NEMO-FOAM output has been shown to mostly increase nutrient concentration throughout the time series, apart from station BATS and L4. Even when the model is spun up for 50 years, shown in Figure 2.9, the DIN concentration increases throughout the time series, meaning that the vertical velocity may have a net upwelling. Therefore, when high noise is added to the vertical velocity, at stations which show an increase in nutrients with time would also produce similar results, which may explain why some members show an increase in DIN. This increase in nutrients may also be caused by the inefficient phytoplankton nutrient uptake, as the maximum rate is similar across all uptake functional forms. Furthermore, in a 1-D model, there is no horizontal velocity, which may also make the nutrients increase on the upper water column. Due to this increase in nutrients, using a high perturbation strength would make the simulation unrealistic.

These nutrient enhancements also lead to an increase in chlorophyll mean towards the end of the time series. This means the PPE often only covers the in situ concentration for the last few years of the time series, especially when the perturbation strength is high (between -2 and 2). However using small (between -0.5 to 0.5) perturbations do not produce much change, and therefore rarely capture the in situ concentration, which leads us to use the medium noise (between -1 and 1) for the PPE in Chapter 5. Meanwhile perturbing the biogeochemistry produces a consistently large range throughout the time series. This further show that perturbing the physics, despite various perturbation strengths, does not garner a large enough range to cover the in situ concentration at different times and regions. Due to the ‘stronger’ effect of perturbing the biogeochemistry, when vertical velocity is perturbed together with the biogeochemical equations, the effect of perturbing the biogeochemistry is more dominant, where the pattern of DIN, chlorophyll, and zooplankton concentration, as well as distributions follow the PBE. This is somewhat concurrent to the earlier study by Friedrichs et al. (2006) and Sinha et al. (2010), where without data assimilation, changing the physical oceanography inputs, which includes the vertical

velocity and mixed layer depth, do not change much of the chlorophyll and zooplankton. However, below the biologically active depth (below 75m-100m), the DIN range is higher when the vertical velocity is perturbed.

Although the PBE DIN range is high on the surface, most of the ensemble members overestimate the DIN concentrations in the water column, as discussed in the previous subsection. This overestimation caused an underestimation in the Chl:DIN ratio (Figure 5.11) and also of the nitrogen proportions in the water column (Figure 5.12(a) and (b)), where the in situ nitrogen proportion can go up to 1, and the highest ensemble member can only produce up to 0.85. High phytoplankton nitrogen proportion means that the nutrient uptake is very efficient in that all of the DIN is transferred into phytoplankton nitrogen. This further shows inefficient the nutrient uptake process is in MEDUSA. We can also use this proportion as a metric for optimum DIN uptake in the ocean.

Despite the underestimation of phytoplankton nitrogen proportion, some of the PBE members can capture the observed nitrogen proportions, as well as its standard deviations, especially during the non-bloom period. This shows an improvement to the model, as the default MEDUSA run underestimated the background observed chlorophyll (Yool et al., 2011, 2013). Furthermore, although the in situ nitrogen proportions at all of the stations produce large temporal variability, the PBE spread can mostly encompass the variability, particularly in light-limited regions. At station Cariaco, despite the poor match between the in situ and modelled chlorophyll concentrations, the PBE spread is so large that the observed nitrogen proportions are generally within the range, showing that even with some mismatch in the phytoplankton phenology, the PBE spread can still capture the observation. However, our PPE cannot produce as much spread as the PBE. These results suggest that at non-oligotrophic and light limiting regions, PBE may be able to provide better uncertainty representation in the observation. This may be helpful in applications such as probabilistic modelling and data assimilation.

In Sinha et al. (2010), the distribution of phytoplankton groups differ between two different ocean general circulation model, shown in Figure 1.4. These differences might be caused by the difference in mixing strength: OCCAM with its higher mixing produces more diatoms in the North Atlantic, North Pacific, and Southern Oceans compared to NEMO. In Chapter 5, we have also demonstrated that at station L4, when the vertical velocity is perturbed, the distribution of dominant phytoplankton types throughout the water column differ between different ensemble members, and these differences are comparable to PBE members. However, at PAP where the phytoplankton concentrations are half of that in L4, although the distribution of dominant phytoplankton types are affected by the vertical velocity (see 5.8(a) to (d)), the differences between PPE members are not as noticeable as the PBE. As discussed in Chapter 5, these differences between the two stations may be caused by the magnitude of vertical velocity. The distribution of dominant phytoplankton type shows that both the vertical velocity and the interaction strength

between NPZ compartments can change the distribution of dominant phytoplankton type. However the discrepancy between the PPE members may depend on the concentration of the phytoplankton, or the magnitude of the vertical velocity. This metric shows the importance of the vertical velocity in controlling the phytoplankton distribution, although perturbing the biogeochemistry still shows a larger effect on both concentration as well as phytoplankton types distribution in the water column, which agrees with the previous studies (Anderson et al., 2010).

In Chapter 1, an earlier study has shown that phytoplankton production is affected by sub-mesoscale processes (Lévy et al., 2001). In order to enhance productivity in the oligotrophic gyre, eddy-induced nutrient pumping is needed. High resolution is needed in order to simulate eddy-induced nutrient pumping in the physical oceanography model (McGillicuddy et al., 2003). In this model, the vertical velocity is obtained from NEMO output and NEMO-FOAM with 1/4° resolution with 63 depth levels. Running PPE by perturbing the vertical velocity, using similar method described in Chapter 4, in higher resolution (e.g 1/12° with 75 depth levels) may produce different results as sub-mesoscale processes, such as eddies and fronts, can be represented by the physical oceanography model, which can be applied in the 3-D MEDUSA-2.0.

Chapter 7

Conclusion and Future Direction

It is known that the marine biogeochemical model has many sources of uncertainty, such as the model parameters, physical oceanography input, and the model structure, such as its equations and complexity. Our study highlights the importance of conducting structural sensitivity analyses, resulting from altering the model biogeochemical equations, in addition to parameter sensitivity analyses, by including mathematical functions that and capture sufficient information of the key biogeochemical processes known from experimental studies. Since the deterministic functions alone cannot capture all the details of the biogeochemical processes at different regions (Anderson et al., 2010), we have also introduced a method whereby instead of having only one deterministic default model output, we use an ensemble generating a range of possible outcomes arising from alternative model structures. Using this method, the ensemble of perturbations generally encompass the in situ observations, especially at stations which are light limited. This is because perturbing the biogeochemical process functions altered the interaction strength between the model compartments, which are shown in how the nitrogen is distributed differently within the model compartments, for different ensemble members. Our study also emphasises that a single set of functional forms is not sufficient to capture the in situ and its uncertainty.

The ensemble is further reduced systematically, while still retaining the statistical metrics and ensemble range. Reducing the ensemble is done to minimise the computational time further and also improving the bias. We have also generated a perturbed physics ensemble, where the vertical velocity that controls the availability of nutrients to phytoplankton are perturbed. The reduced ensemble has retained the range of the full ensemble and also improved the correlations and bias. Compared to reduced PBE, perturbing the physics does not generate as large spread and cannot always encompass the in situ observations. Although the chlorophyll distribution is affected by the physics, in the perturbed physics ensemble, the variability in the chlorophyll distribution is not as noticeable as perturbing the biogeochemistry. We have also examined the characteristics of both ensembles. The perturbed physics generates its range by

altering the nutrient supply to the phytoplankton, which makes the biogeochemical compartments within the PPE members vary together with strongly positive correlations with nutrient supply. The opposite is observed form PBE, where nutrients and phytoplankton compartments have negative correlations because perturbing the biogeochemistry alter the interaction strength of phytoplankton, zooplankton, and nutrients at different concentrations. We have also observed that, because of the way PBE alters the interaction strength of the model compartments, the ensemble members that captured the in situ are also time-dependent.

From these conclusions, our study shows promise that an ensemble of a single biogeochemical model resulting from perturbing the biogeochemical processes may produce meaningful prediction range of its state variables. Therefore it is possible to apply this method on other operational models, and also data assimilation with satellite-derived chlorophyll, which will be explained further in the next section.

7.1 Limitations and Future Directions

We have shown how perturbing the functional forms, and vertical velocity (physics) affect the model dynamics: from the bulk concentrations of chlorophyll and DIN, as well as the distribution of dominant phytoplankton types in the water column. However, there are some limitations of the study and further improvement that can be made to explore the effect of structural sensitivity and to utilise this uncertainty even further.

7.1.1 3D Perturbed biogeochemistry ensemble

This study is restricted on a 1-D framework, focusing on chlorophyll and DIN at five oceanographic stations. There are other regions we have not explored, such as the monsoon and polar regions, where drastic changes will occur due to climate change. We can explore the structural sensitivity in more regions by running the PBE on a 3-D framework, which can be further classified using the Longhurst biogeographical provinces (Longhurst et al., 1995). Furthermore, we have not assessed the effect of structural sensitivity on PFT distribution in the global ocean in this study, and this can be done by running the PBE on a 3-D framework. Applying PBE to the more recent model, MEDUSA-2.0 will also allow us to explore the effect of structural sensitivity on other variables that are not available in MEDUSA 1.0, such as the primary production, CO_2, and alkalinity. This can be taken further to show the uncertainties in the projections of CO_2 and alkalinity in the ocean under different climate scenarios, which is one of the objectives of the model inter-comparison project (Kwiatkowski et al., 2014).

Running the marine biogeochemical model in a 3-D framework may also produce a better representation of the DIN distributions in the water column, due to the explicit horizontal advection, eddies, and fronts that are available in 3-D models. This can improve the DIN and chlorophyll distributions and phenology in regions where the intrusion of deep nutrient-rich waters occurs below the euphotic layer, such as Cariaco (Walsh et al., 1999). However, there are some challenges when it comes to running the ensemble on a 3-D framework. Firstly, we have to take into account of the GCM resolution as the 3-D model will be run together with a GCM (such as NEMO). Eddy-induced nutrient pumping can be simulated in a super fine resolution (Popova et al., 2006), but this will make the computational cost expensive. Therefore a suitable resolution has to be decided before running the 3-D PBE so we can reach a balance between computational cost and capturing important physical processes. Another extension of this study is to compare the default deterministic 3-D MEDUSA-2.0 run in finer resolution with 3-D PBE MEDUSA-2.0 run in standard 1/4° resolution, and which model run can reproduce better in situ observation using similar metrics as that in Chapter 3.

In this thesis, we only perturbed one physical parameter, which is the vertical velocity. There are other essential physics variables that can affect phytoplankton growth, such as the MLD (both the strength and depth), downwelling shortwave radiation, and temperature. Perturbing these physical oceanography variables on a 3-D framework would thoroughly assess the effect of perturbing the physics to biogeochemical model variables, particularly the distribution of chlorophyll, different PFTs, and nutrients.

Apart from the physical input, to initialise the biogeochemical fields, such as DIN, Oxygen, and silicate, we need observation data. We can obtain this data from World Ocean Atlas and model outputs, similar to the MEDUSA-2.0 original run (Yool et al., 2013). Varying initial conditions can also affect the model results, as shown in Chapter 3 at station PAP. The model sensitivity with initial conditions can be explored further using other in situ data such as the Bio-Argo float. Comparing the ensemble with observations can also be done using the World Ocean Atlas and also satellite-derived chlorophyll. To compare the global PFT distribution, we can utilise satellite-derived products (e.g. Hirata et al. (2011) and Roy et al. (2011)) to directly compare the distribution of diatoms and non-diatoms between the ensemble, default run, and observations. Furthermore, we can also examine how structural sensitivity can affect the emergent relations between phytoplankton community and chlorophyll using the method of De Mora et al. (2016). However, in the current 1-D experiment, some of the PBE members do not produce diatoms even in L4 or PAP.

Apart from using satellite-derived chlorophyll for direct comparison with 3-D PBE, it is possible to utilise satellite-derived chlorophyll for data assimilation.

7.1.2 Data Assimilation

As discussed in Chapter 1 section 1.5, it is possible to utilise the uncertainty estimate from a model, and improve its skill. There are two main methods in data assimilations, and one of them is to use the model trajectory that best fits the data. The PBE and PBPE have provided sufficient error estimates, and their ensemble means and medians have mostly reduced the RMSE of the default run. Therefore it is possible to apply the trajectory method using the PBE and PBPE as the different trajectories. Previous data assimilation studies have used ocean colour (Ford et al., 2012; Ford and Barciela, 2017) and both ocean colour with in situ data (Ciavatta et al., 2014) to be assimilated with the biogeochemical model trajectories. In these studies, the trajectories are generated by varying the model parameters. The PBE and PBPE have shown to produce better uncertainty estimates than varying the vertical velocity, which may be useful when applied to data assimilation.

One of the advantages of perturbing the biogeochemistry is that we do not have to optimise the parameter values too much. However, we can also combine both the structural as well as parametric uncertainty as described in Aldebert and Stouffer (2018), where Bayesian statistic is used to vary the parameter values. This approach may provide even better uncertainty estimates compared to only varying the model process equation.

In an earlier study, the PISCES model has been used to study the efficiency of iron fertilisation. However, there are still some significant uncertainties that need to be explored further using observational data and models (Aumont and Bopp, 2006). A study by Watson et al. (2008), suggests that to reduce the uncertainties in the effects of iron fertilisation, assimilating in situ data and remote sensing products with model is required. In our current study, we have not explored the iron compartment of MEDUSA. A further study may include how structural sensitivity may affect the iron dynamics and assimilate in situ iron data with the model. Data assimilation and a high-resolution ensemble NEMO-MEDUSA can be used to plan the locations and timing of future in situ iron fertilisation, as well as quantifying the uncertainties.

Optimised vs unoptimised runs in 1-D model

It has been shown in a previous study that in a 1-D marine biogeochemical model, after objectively optimising the model, changing the physical forcings can produce greater changes in model dynamics such as the chlorophyll and zooplankton distribution (Friedrichs et al., 2006). It is unknown whether the ensemble will show a larger range and more variable chlorophyll and zooplankton distributions if we optimised each of the reduced PBE members, instead of only optimising the default model. Then, we

can determine whether perturbing the physical input can produce a larger range than the optimised PBE. Although more laborious, this effort may reduce the current model bias and also explore whether ocean biogeochemical model is more sensitive to the choice of functional forms or the physical model input after data assimilation.

To extend this experiment, we can include other functional forms that we have not currently chosen, such as the Disk equation, or the droop cell quota, which is less structurally sensitive compared to the Monod equation in MEDUSA (Aldebert et al., 2018). Another possible addition is to use a different physical oceanography model output (such as the GOTM model, which is made for 1-D models (Butenschön et al., 2016)) to see how different physical oceanography models can affect the biogeochemical model results. Similar metrics that are in Chapter 2, can be used to compare the range produced from altering the process equations and physical oceanography input.

7.1.3 Bloom Experiment

In Chapter 2 section 2.8, we have discussed the two hypotheses on the bloom initiation. In the disturbance recovery hypothesis, the bloom initiation occurred in early winter (Behrenfeld et al., 2013), but in the critical depth hypothesis, the initiation happens in early spring (Sverdrup, 1953). In PBE, the bloom phenology is represented probabilistically. This means that we obtain the initiation and termination of the bloom as a range. Using the ensemble approach means that the model can produce earlier or later timings than the observation, which may be useful to investigate the disturbance recovery hypothesis described in Behrenfeld et al. (2013).

The bloom study by Behrenfeld and Boss (2014) has not taken into account the species succession and selective grazing, which may lead to a bloom initiation (Behrenfeld et al., 2013; Behrenfeld and Boss, 2014). However, in situ studies have reported that the phytoplankton bloom occurs due to the increase in phytoplankton group (Steinberg et al., 2001; Smyth et al., 2015; Pinckney et al., 2015). Our study has shown that during the bloom period at the North Atlantic stations, the PFT dominance has changed from smaller non-diatoms to diatoms, as shown in Chapter 5. This may allow us to investigate further how species succession during bloom initiation occurs. Furthermore, a study has also shown that by adding disturbance into the model simulation, the diversity index of phytoplankton increases (Smith et al., 2016). To investigate the biogeochemical aspect of the bloom initiation, we can apply the ensemble method into a model with more PFT representations than MEDUSA, such as PlankTOM (Le Quere et al., 2005). It is also possible to investigate different grazing selection by applying PBE because the ensemble allows us to apply different selective grazing types in the model. In this thesis, have included Class 1 (no prey switching) and 2 (passive switching). Additionally, we can include active switching models (Class 3,

described in Gentleman et al. (2003)), to extend the grazing strategies. This phytoplankton bloom study can also be extended to examine the effect of structural sensitivity in zooplankton, which we have not explored thoroughly in this thesis.

Increasingly, models are used as tools to support decision making. Using an ensemble-based model can be useful in providing alternative simulations from the 'default' model. When devising a climate-adaptation plan, we can use the ensemble-based model so that various method of mitigations can be developed for the range of possible model outcome. As PBE shows a large range of outcomes, altering functional forms to generate ensembles in other biological models may be used as decision-making tools.

Appendix A

Running MarMOT 1.1-alpha

The simulation in MarMOT is controlled by input tables or item tables that contain one or more instances of a particular input item, such as the model parameter and physical forcings. When simulating the model in various different regions, the input tables are combined according to the case table, which defines a simulation case, characterised by specific combinations of input data for a specific region or ensemble. The simulation case is identified by a site name, if the model is run in different regions, or ensemble numbers, if the input parameters are varied. The MarMOT runs the biogeochemical model according to the specifications from the input case table. The model output can be selected according to user requirements and is provided in a file containing the output tables. These input tables will be explained in more detail in section A.

In order to run the model, a set up script containing the "experiment control table" is used. This file contains the location of input files that are going to be used and the output files, which are the model results on the surface (`outtdayf`) and throughout the water column (`outktdayf`), as well as the variables files (`outtdayvarf` and `outdktdayvarf`), where model output we want to see in the table can be specified. The input files, which are written in ASCII format, consist of:

- Parameter set item table, where MEDUSA parameters, simulation options, site information, and the time period of model run is defined. The input command to write the script is written in `qcr` font.
  - `model3f` = Lists of MEDUSA parameters and values.
  - `optionf` = Contains the code for simulation options; which models are used, both the ecosystem model number and also the photosynthesis light limitation and attenuation model, and advection and diffusion schemes. MEDUSA's model code is number three. The recommended advection scheme is monotonic upstream scheme for conservative laws (MUSCL, number two), and diffusion scheme is implicit (number one). In this study, we use the recommended advection schemes (`model=3`, `advection= 2`, `diffusion=1`).
  - `taxisf` = Time axis and the year length in days. When running the model in 1998 daily, we can specify the base year (`baseyear`) as 1998, and year length (`yearlen`) of 365 days.
  - `timeperiodf` = Time period of the model run, where start and finish times are defined. Therefore in this script, if the model starts in 1998 and is run for 10 years, the start year (`startyear`) is 1998, the start of the day (`startday`) is 1, the last calendar year of the model run (`finyear`) is 2007, and the last day of the model run (`finday`) in year 2007 will be 365.
  - `environf` = Site information, including maximum depth, both in level (`kmax`) and meters (`maxdepth`), latitude, and longitude, both in degrees (°N and °S)and NEMO coordinates (`j`

and i). The site is identified as jAAAiBBB, where AAA is the *j*-coordinate and BBB is the *i*-coordinate on the NEMO grid.

- Gridded domain items, where depth level and initial conditions are defined. For this gridded domain items there will be two files, the first one is where the total depth level or initial conditions are specified, and the second one is the input files for each day or both depth and day, and is saved in '.dat' format.
  - zlevelf = Depth levels and real depth in meters. The first file contains the start of depth level, k1, depth increments kstep, and total depth nk. The second file will contain the actual depth levels k, which could range from 1 to 75, and its real depth in meters zbot.
  - initf = Initial condition for biogeochemical tracers at each depth level. The first file contains the name of the site, specified in environ, the k1, kstep and nk. The second file contains all initial conditions for all the state variables, for example DIN, silicate, iron, phytoplankton, chlorophyll, and zooplankton concentrations including the site, k, i, j.
  - ftf = Input for physical forcing, such as mixed layer depth (mld) and solar radiance (sol), and iron (Fe) time series. The first one is the name of the site, along with the start tmin and end of the run tmax, as well as the timestep nt. The second one is the actual input data, that has been specified for each day.
  - fktf = Input for physical forcing at each depth level, such as vertical velocity (w), vertical eddy diffusivity (vdc), and temperature (temp). Similar to the ftf, the first file contains the site name, tmin, tmax, nt, kstep, and nk. The second one is the physical forcing inputs for each day and depth.

- Input case table, where the sites are identified.
  - casef Site identification and number if running at multiple sites

Examples of these input files are described in the next section. After locating the directory of the script, the model is then run by calling this script after typing marmot in the terminal, and a user interface will appear. The interface will confirm all the input and output tables that have been specified in the script (e.g. test3). These are summarised in Figure A.1. A template of input files stated above can be generated in the MarMOT interface when no script is called. The output data that is included in the outtdayf and outktdayf includes the case variables (if run in multiple sites), time, depth, profile forcing variables (such as w and vdc), and the biogeochemical tracers (such as chlorophyll, DIN, silica, and zooplankton concentrations).

The MarMOT runs the biogeochemical model according to the specification from the input case table, where the sites are identified. In our run, we identify each stations separately. Here we use station ALOHA and the default run as an example.

A.1 Experimental control table script

Experiment control table describes the location of the input tables, and is describe below:

```
NAME VALUE
model3f in16/param_phzh # the parameter files
optionf in_aloha/option_aloha # option file
taxisf in/taxis_base1998 # base year and year length
```

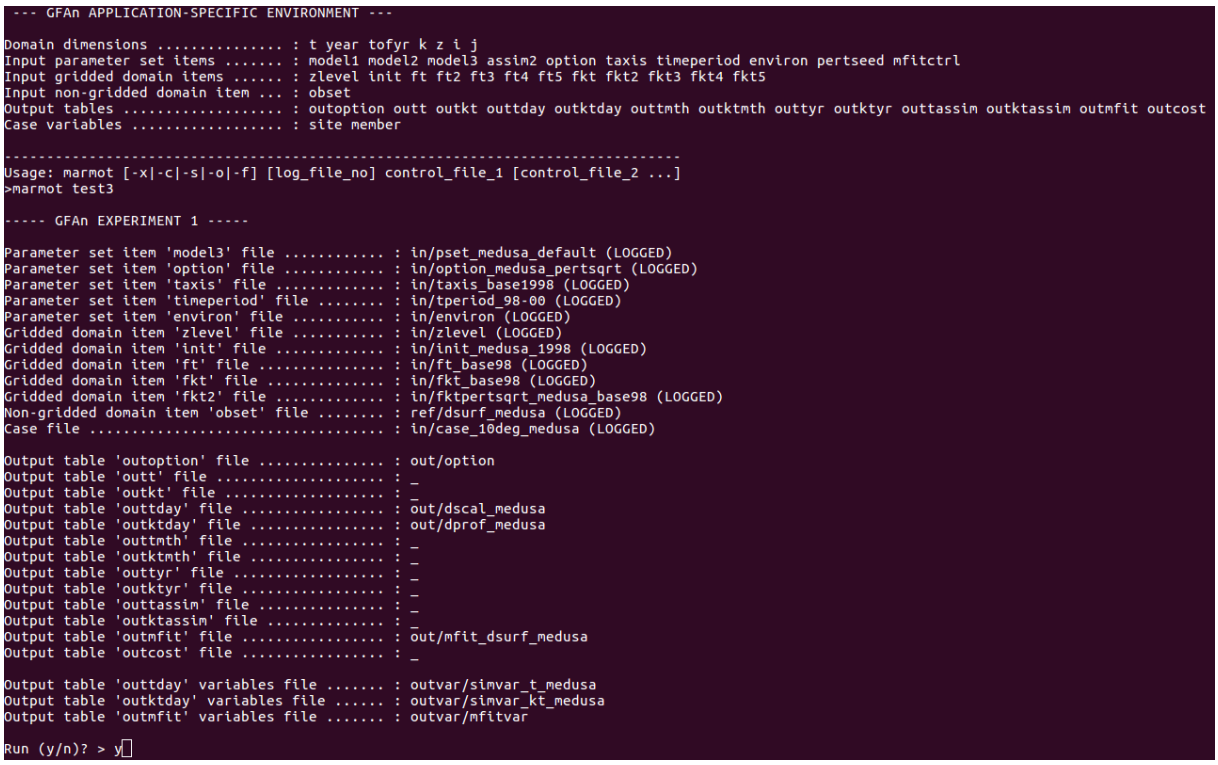


Figure A.1: Screenshot from MarMOT user interface. Here, we call the script 'test3' which describes the locations of all the input tables, such as the parameter table (pset_medusa_default), site informations (environ), and the output files on the surface (dscal_medusa) and throughout the depth (dprof_medusa). Other gridded domain such as fkt2 is used when a perturbation term used to represents the effect of horizontal flux divergence, which is not used in this study.

```
timeperiodf in_aloha/tperiod_98 # start and finish times
environf in_aloha/environ_aloha # site information
zlevelf in_aloha/zlevel # depth levels
initf in_aloha/init_1998_aloha_3 # initial conditions
ftf in_aloha/aloha_base98 # physical forcing input
fktf in_aloha/fkt_aloha_1998_2 # physical forcing for each levels
casef in_aloha/case_aloha # site identification
outoptionf out_aloha3/option_aloha2 # simulation options
outtdayf out_aloha3/dscal_aloha3_def # output file for surface
outtdayvarf outvar/simvar_t_medusa # output variables for surface
outktdayf out_aloha3/dprof_aloha3_def # output file for profiles
outktdayvarf outvar/simvar_kt_medusa # output variables for surface
```

A.2 MEDUSA parameters (model3f)

```
NAME VALUE

rcchlminndia 20 # minimum C to Chl ratio for non-diatoms (gC/gChl) [=1/xthetam]
frcchlmindiat 1 # minimum C to Chl ratio for diatoms relative to non-diatoms
[=xthetam/xthetamd]
rsinpdiatmin 0.2 # minimum diatom Si:N ratio [=xsin0]
rsinpdiatmax 5 # maximum diatom Si:N ratio [=1/xnsi0]
rcnphy 6.625 # phytoplankton C:N ratio (molC/molN) [=xthetap]
rcnzmi 6.625 # microzoo.  C:N ratio (molC/molN) [=xthetazmi]
rcnzme 6.625 # mesozoo.  C:N ratio (molC/molN) [=xthetazme]
rcndet 6.625 # detritus C:N ratio (molC/molN) [=xthetad]
rfen 30e-6 # phytoplankton Fe:N ratio (molFe/molN) [=xrfn]
alphachlndia 15 # chl.  specific initial slope of P-I curve for non-diatoms (gC/
gChl/ (W/m2)/d) [=xaln]
falphachldiat 0.75 # chl.  specific initial slope of P-I curve for diatoms rel-
ative to non-diatoms [=xald/xaln]
photmax0ndia 0.8 # maximum growth rate for non-diatoms at 0 degC (/d) [=xvpn]
fphotmax0diat 1.0 # maximum growth rate for diatoms relative to non-diatoms at
0 degC [=xvpd/xvpn]
ursininf 1.5 # hypothetical growth ratio at infinite Si:N ratio [=xuif]
kdinndia 0.5 # half-sat.  conc.  for DIN uptake by non-diatoms (mmolN/m3 ) [=xnln]
fkdindiat 1.5 # half-sat.  conc.  for DIN uptake by diatoms relative to non-
diatoms [=xnld/xnln]
ksidiat 0.75 # half-sat.  conc.  for Si uptake by diatoms (mmolSi/m3 ) [=xsld]
kfendia 0.00033 # half-sat.  conc.  for Fe uptake by non-diatoms (mmolFe/m3 )
[=xfln]
fkfediat 2.03 # half-sat.  conc.  for Fe uptake by diatoms relative to non-diatoms
[=xfld/xfln]
gmaxzmi 1 # microzoo.  maximum grazing rate (/d) [=xgmi]
fgmaxzme 1 # mesozoo.  maximum grazing rate relative to microzoo [=xgme/xgmi]
```

```
kgzmi 0.8 # microzoo. grazing half-sat. conc. (mmolN/m3 ) [=xkmi]
fkgzme 0.375 # mesozoo grazing half-sat. conc. relative to microzoo. [=xkme/xkmi]
fgmessy 0.2 # micro/mesozoo. grazing inefficiency [=xphi]
betan 0.69 # micro/mesozoo. N assimilation efficiency [=xbetan]
offbetac 0 # C assim. efficiency offset as frac. of max. [=(xbetac-xbetan)/(1-
xbetan) OR (xbetac-xbetan)/xbetan]
kc 0.8 # micro/mesozoo. net C growth efficiency [=xkc]
prfzmilive 0.75 # grazing preference of microzoo. for live food (=non-diatoms)
[=xpmipn]
prfzmelive 0.85 # grazing preference of mesozoo. for live food [=xpmepn+xpmepd
+xpmezmi]
cprfzmep 0.5882 # grazing preference of mesozoo. for phyto. given live food
[=(xpmepn+xzmepd) / (xpmepn+xpmepd+xpmezmi)]
cprfzmepndia 0.3 # grazing preference of mesozoo. for non-diatoms given phy-
toplankton [=xpmepn/ (xpmepn+xpmepd)]
metapndia 0.02 # non-diatom metabolic loss rate (/d) [=xmetapn]
fmetapdiat 1 # diatom metabolic loss rate relative to non-diatoms [=xmetapd/xmetapn]
fmetazmi 1 # microzoo. metabolic loss rate relative to non-diatom phyto. [=
xmetazmi/xmetapn]
fmetazme 1 # mesozoo. metabolic loss rate relative to microzoo. [=xmetazme
/ xmetazmi]
mortmaxpndia 0.1 # non-diatom maximum mortality rate (/d) [=xmpn]
kmortpndia 0.5 # non-diatom mortality half-sat. conc. (mmolN/m3 ) [=xkphn]
fmortmaxpdiat 1 # diatom maximum mortality rate relative to non-diatoms [=xmpd/xmpn]
fkmortpdiat 1 # diatom mortality half-sat. conc. relative to non-diatoms [=xkphd
/ xkphn]
fmortmaxzmi 1 # microzoo. maximum mortality rate relative to non-diatom phyto.
[=xmzmi/xmpn]
fkmortzmi 1 # microzoo. mortality half-sat. conc. relative to non-diatom phyto.
[=xkzmi/xkphn]
fmortmaxzme 2 # mesozoo. maximum mortality rate relative to microzoo. [=xmzme
/xmzmi]
fkmortzme 1.5 # mesozoo. mortality half-sat. conc. relative to microzoo. [=xkzme
/xkzmi]
remin0 0.016 # detrital nitrogen remineralisation rate at 0 degC (/d) [=xmd]
rfedust 0.014 # soluble iron fraction of aeolian dust deposition [=xfe sol]
sedfe 0 # iron input from sediment (mmolFe/m2/d) [=xfe sed]
ligand 1e-3 # total ligand concentration (mmol/m3 ) [=xLgT*1e-3]
kfelig 1e5 # dissociation constant for (Fe + ligand) (/(mmolFe/m3 )) [=xk FeL*1e3]
scavfe 1e-3 # scavenging rate of "free" iron (/d) [=xk sc Fe]
ffastmortdiat 0.75 # fast-sinking fraction of diatom mortality losses [=xfd-
frac1]
ffastmortzme 1 # fast-sinking fraction of mesozoo. mortality losses [=xfdfrac2]
```

```
rco3fast90 0.02 # polar CaCO3 to organic C ratio [=xcaco3a]
rco3fast0 0.1 # equatorial CaCO3 to organic C ratio [=xcaco3b]
rprotca 0.07 # calcium carbonate protection ratio (gC/gC) [=xprotca]
rprotsi 0.026 # biogenic silicon protection ratio (gC/gSi) [=xprotsi]
lfastc 188 # excess organic C dissolution length scale (m) [=xfastc]
lfastca 3500 # calcium carbonate dissolution length scale (m) [=xfastca]
lfastsi 2000 # biogenic silicon dissolution length scale (m) [=xfastsi]
sidiss 0.006 # diatom frustule dissolution rate (/d) [=xsdiss]
sinkdet 3 # detritus gravitational sinking rate (m/d) [=vsed*86400]
attenwaterg 0.0232 # downwelling attenuation due to water in blue-green band
(/m) [=xkg0]
attenwaterr 0.225 # downwelling attenuation due to water in red band (/m) [=xkr0]
attenupigg 0.074 # downwelling attenuation due to 1 mg/m3 of pigment in blue-
green band (/m) [=xkgp]
attenupigr 0.037 # downwelling attenuation due to 1 mg/m3 of pigment in red band
(/m) [=xkrp]
attenexpigg 0.629 # exposant for pigment (mg/m3 ) absorption in blue-green band
[=xlg]
attenexpigr 0.674 # exposant for pigment (mg/m3 ) absorption in red band [=xlr]
rchlpig 0.7 # chl.  to total pigment ratio [=rpig]
photmaxtdep 1 # temperature-dependent phytoplankton max.  growth (0:  off, 1:
on) [=jphy]
mortfnpndia 3 # non-diatom mortality functional form (1:  linear, 2:  quadratic,
3:  hyperbolic, 4:  sigmoid) [=jmpn]
mortfnpdiat 3 # diatom mortality functional form (1:  linear, 2:  quadratic,
3:  hyperbolic, 4:  sigmoid) [=jmpd]
mortfnzmi 3 # microzoo mortality functional form (1:  linear, 2:  quadratic,
3:  hyperbolic, 4:  sigmoid) [=jmzmi]
mortfnzme 3 # mesozoo mortality functional form (1:  linear, 2:  quadratic, 3:
hyperbolic, 4:  sigmoid) [=jmzme]
remintdep 1 # temperature-dependent detrital remineralisation (0:  off, 1:  on)
[=jmd]
photopt 1 # implementation of photosynthesis (0:  external, 1:  internal) [=jphot]
dsinkopt 1 # implementation of detrital sinking (0:  external, 1:  internal)
[=jsink]
dustopt 0 # conversion of dust to soluble iron (0:  external, 1:  internal) [=jdust]
logopt 0 # log output to stdout (0:  off, 1:  on) [=lwp]
```

A.3 Option

Contains the information about which model is used, the advection, and diffusion schemes

```
NAME VALUE

model 3 # ecosystem model number
advection 2 # advection scheme (0-2)
diffusion 1 # diffusion scheme (0-1)
```

0 means no advection or diffusion scheme, 1 is implicit in diffusion scheme and upstream scheme in advection. Model 1 is an NPZD model by Oschlies and Garçon (1999) and 2 is the HadOCC model by Palmer and Totterdell (2001).

A.4 Taxis and Timeperiod input files

The taxis file defines the time axis origin, year length if fixed and whether to use periodic annual forcing. The default axis starts at year 1 with periodic 365 day forcing. The file is describe below

```
NAME VALUE

baseyear 1998 # first calendar year of time axis
yearlen 365 # year length if fixed (days)
```

The timeperiod file specifies the start and finish times for the simualtion.

```
NAME VALUE


startyear 1998 # first calendar year of model run
startday 1 # first day of model run in year
finyear 2007 # last calendar year of model run
finday 365 # last day of model run in yea
```

A.5 Environment parameter set item

This table give time invariant values that are typically dependent on the station or site, including maximum depth, latitude, and any constant forcing parameter, but in this study we use set the forcing to zero.

site	kmax	maxdepth	j	i	lat	lon	nom_lat	nom_lon
j518i593	58	4786.8	518	593	22.86	-158	22.75	-158

A.6 Vertical grid item

gridded domain item referenced to the depth level, k, specifying up to 500 depth levels. Omitting this item causes the simulation to be run for a single level stretching from the surface to the deepest depth. There are two files, `zlevel` and `zlevel.dat`. The former is describe below:

```
NAME VALUE

k1 1
kstep 1
nk 37
```

The gridded file for 37 levels is described below: `k zbot`

```
1 6.06386
2 12.2693
3 18.6396
4 25.2019
5 31.9875
6 39.0331
7 46.3811
8 54.0804
9 62.1881
10 70.7699
11 79.9018
12 89.6715
13 100.18
14 111.543
15 123.893
16 137.382
17 152.183
18 168.492
19 186.529
20 206.544
21 228.812
22 253.642
23 281.371
24 312.367
25 347.026
26 385.768
27 429.036
28 477.285
29 530.972
30 590.55
31 656.452
32 729.078
33 808.779
34 895.848
35 990.502
36 1092.87
37 1203.01
```

A.7 Initial conditions

Gridded domain item referenced to the depth level, defines initial profiles of primary tracers and or composition ratios. The content first file, `init_aloha_3` are:

```
site k1 kstep nk
j518i593 1 1 37
```

The example for the second file is shown in Figure A.2

A.8 Physical forcing inputs

There are two different physical forcing inputs: scalar and profile forcing items. The former is referenced to the time dimension and the latter is referenced to the t and k dimensions. For the ensemble runs in this study, the scalar input includes mixed layer depth (`mld`), solar radiation (`sol`), and iron (`fe`), along with the site name and time, `t`. The profile inputs consist of vertical velocity (`w`), temperature `temp`, and vertical diffusivity coefficient (`vdc`). The content of the first files, both for scalar (`aloha_base_98`) and profile (`fkt_aloha_3`) are similar to the initial condition file. The example of the second scalar file is described below:

```
site t sol mld fe
j518i593 2.5 140.71327 47.72704178 0.0004
j518i593 7.5 146.21756 49.65932862 0.0004
j518i593 12.5 156.7739 54.49843682 0.0004
j518i593 17.5 153.771 55.7928273 0.0004
j518i593 22.5 156.3096 55.55119066 0.0004
j518i593 27.5 170.19339 55.12459587 0.0004
j518i593 32.5 178.98274 56.73110492 0.0004
j518i593 37.5 189.388 57.32757672 0.0004
j518i593 42.5 189.94858 57.89794809 0.0004
j518i593 47.5 189.65088 57.3974245 0.0004
j518i593 52.5 189.26494 57.21076073 0.0004
j518i593 57.5 203.13962 51.07504236 0.0004
j518i593 62.5 223.67441 39.96556252 0.0004
j518i593 67.5 230.46655 30.23758777 0.0004
j518i593 72.5 238.89502 22.66889391 0.0004
j518i593 77.5 242.10495 20.22570286 0.0004
j518i593 82.5 215.51154 27.10937347 0.0004
j518i593 87.5 218.82249 41.94496904 0.0004
j518i593 92.5 245.53812 56.62015939 0.0004
j518i593 97.5 247.20531 75.53160781 0.0004
j518i593 102.5 250.83928 87.60096881 0.0004
j518i593 107.5 258.49155 85.57904512 0.0004
j518i593 112.5 258.11212 79.18016596 0.0004
```

As described in Chapter 2 the iron deposition value is obtained from (Mahowald et al., 2009) and is kept constant. However this is not simulated by the model at all stations. For the profile forcing item, the example is shown below. Note that this is only the first five-day averaged input:

```
site t k vdc w temp
j518i593 2.5 1 864000 -0.05286443 24.06920242
j518i593 2.5 2 864000 -0.085447322 24.06922054
j518i593 2.5 3 849650.9354 -0.129893747 24.06923771
j518i593 2.5 4 815080.3802 -0.158311353 24.06925297
j518i593 2.5 5 790592.9672 -0.161882514 24.06926727
j518i593 2.5 6 652346.981 -0.151942787 24.06927681
```

site	k	i	j	din	pndia	pndiat	zmi	zme	det	dsi	dfe	chlndia	chlndiat	pdiatsi	year	tofyr	rsinpdiat	rcchlndia	rcchldiat
j518i593	1	518	593	0.0222	0.3090	0.0459	0.0474	0.0524	0.0028	1.3116	0.0007	0.1063	0.0184	0.1369	1998	2.5	2.9792	228.1204	198.288
j518i593	2	518	593	0.0000	0.3145	0.0459	0.0474	0.0524	0.0031	1.1833	0.0007	0.1081	0.0184	0.1369	1998	2.5	2.9792	228.1195	198.287
j518i593	3	518	593	0.0025	0.3199	0.0459	0.0474	0.0524	0.0032	1.3100	0.0007	0.1100	0.0184	0.1369	1998	2.5	2.9792	228.1173	198.285
j518i593	4	518	593	0.0100	0.3781	0.0459	0.0474	0.0524	0.0035	1.2920	0.0007	0.1300	0.0184	0.1369	1998	2.5	2.9790	228.0956	198.273
j518i593	5	518	593	0.0100	0.3084	0.0462	0.0478	0.0530	0.0038	1.2740	0.0007	0.1075	0.0187	0.1366	1998	2.5	2.9591	224.9969	196.396
j518i593	6	518	593	0.0200	0.2449	0.0672	0.0815	0.0934	0.0042	1.2845	0.0006	0.1400	0.0386	0.1423	1998	2.5	2.1183	137.2097	138.529
j518i593	7	518	593	0.0200	0.1246	0.1296	0.1811	0.2100	0.0044	1.2154	0.0006	0.1300	0.1091	0.1680	1998	2.5	1.2969	75.2003	94.490
j518i593	8	518	593	0.0100	0.0933	0.1849	0.2725	0.3080	0.0047	1.2600	0.0006	0.1550	0.2171	0.2290	1998	2.5	1.2387	47.2331	67.781
j518i593	9	518	593	0.0175	0.0694	0.1636	0.2411	0.2980	0.0049	1.2743	0.0005	0.1500	0.2476	0.2747	1998	2.5	1.6784	36.2850	52.580
j518i593	10	518	593	0.0175	0.0684	0.1052	0.1505	0.2022	0.0056	1.2600	0.0005	0.1700	0.1849	0.2197	1998	2.5	2.0892	31.5702	45.255
j518i593	11	518	593	0.0240	0.0717	0.0552	0.0757	0.0967	0.0137	1.3712	0.0004	0.1900	0.1072	0.1326	1998	2.5	2.4021	29.5793	40.992
j518i593	12	518	593	0.0367	0.0765	0.0228	0.0302	0.0322	0.0878	1.3871	0.0004	0.2100	0.0485	0.0600	1998	2.5	2.6314	28.5623	37.370
j518i593	13	518	593	0.0420	0.0747	0.0061	0.0078	0.0062	0.2996	1.5047	0.0004	0.2100	0.0144	0.0172	1998	2.5	2.8347	27.8886	33.535
j518i593	14	518	593	0.0833	0.0700	0.0007	0.0009	0.0006	0.4785	1.5176	0.0004	0.2017	0.0020	0.0022	1998	2.5	3.0538	27.2247	28.706
j518i593	15	518	593	0.2500	0.0620	0.0000	0.0651	0.0091	0.4369	1.6460	0.0004	0.1867	0.0002	0.0002	1998	2.5	3.1630	26.0473	24.676
j518i593	16	518	593	0.3150	0.0462	0.0094	0.0138	0.0032	0.3541	1.8244	0.0004	0.1475	0.0267	0.0104	1998	2.5	2.2219	24.5900	131.400
j518i593	17	518	593	0.5743	0.0476	0.0009	0.0017	0.0007	0.2796	2.0761	0.0004	0.0867	0.0029	0.0008	1998	2.5	2.4658	43.1177	64.403
j518i593	18	518	593	0.5200	0.0259	0.0000	0.0001	0.0000	0.2182	1.8200	0.0004	0.0569	0.0001	0.0000	1998	2.5	2.4894	35.7535	62.367
j518i593	19	518	593	1.5280	0.0123	0.0000	0.0000	0.0000	0.1701	2.4529	0.0004	0.0271	0.0000	0.0000	1998	2.5	2.5111	35.4065	62.292
j518i593	20	518	593	1.8614	0.0113	0.0000	0.0000	0.0000	0.1326	3.1527	0.0004	0.0250	0.0000	0.0000	1998	2.5	2.5383	35.3729	62.961
j518i593	21	518	593	2.8620	0.0043	0.0000	0.0000	0.0000	0.1027	3.8161	0.0004	0.0095	0.0000	0.0000	1998	2.5	2.5520	35.3021	64.127
j518i593	22	518	593	4.3210	0.0000	0.0000	0.0000	0.0000	0.0782	4.8897	0.0004	-0.0001	0.0000	0.0000	1998	2.5	2.5464	35.1062	64.668
j518i593	23	518	593	5.7843	0.0000	0.0000	0.0000	0.0000	0.0582	6.6735	0.0004	-0.0001	0.0000	0.0000	1998	2.5	2.5263	34.7699	63.932
j518i593	24	518	593	8.4000	0.0000	0.0000	0.0000	0.0000	0.0418	9.0680	0.0004	0.0000	0.0000	0.0000	1998	2.5	2.4966	34.4332	62.482
j518i593	25	518	593	12.4717	0.0000	0.0000	0.0000	0.0000	0.0289	11.8274	0.0004	0.0000	0.0000	0.0000	1998	2.5	2.4578	34.2907	61.045
j518i593	26	518	593	14.8011	0.0000	0.0000	0.0000	0.0000	0.0190	15.2100	0.0004	0.0000	0.0000	0.0000	1998	2.5	2.4075	34.5007	59.867
j518i593	27	518	593	18.3544	0.0000	0.0000	0.0000	0.0000	0.0121	23.3012	0.0004	0.0000	0.0000	0.0000	1998	2.5	2.3399	35.2228	58.971
j518i593	28	518	593	22.6136	0.0000	0.0000	0.0000	0.0000	0.0074	33.5617	0.0004	0.0000	0.0000	0.0000	1998	2.5	2.2414	36.7891	58.424
j518i593	29	518	593	26.3954	0.0000	0.0000	0.0000	0.0000	0.0045	46.3061	0.0005	0.0000	0.0000	0.0000	1998	2.5	2.0874	40.1535	58.553
j518i593	30	518	593	31.8607	0.0000	0.0000	0.0000	0.0000	0.0027	60.8669	0.0005	0.0000	0.0000	0.0000	1998	2.5	1.8299	48.4068	59.999
j518i593	31	518	593	36.5140	0.0000	0.0000	0.0000	0.0000	0.0016	71.7862	0.0005	0.0000	0.0000	0.0000	1998	2.5	1.4362	75.0852	64.718
j518i593	32	518	593	39.8237	0.0000	0.0000	0.0000	0.0000	0.0009	86.3769	0.0005	0.0000	0.0000	0.0000	1998	2.5	1.1017	231.1379	76.528
j518i593	33	518	593	41.0785	0.0000	0.0000	0.0000	0.0000	0.0005	94.9205	0.0005	0.0000	0.0000	0.0000	1998	2.5	1.5423	73.0766	87.066
j518i593	34	518	593	41.5292	0.0000	0.0000	0.0000	0.0000	0.0003	101.8024	0.0005	0.0000	0.0000	0.0000	1998	2.5	2.5046	51.5611	65.702
j518i593	35	518	593	41.6030	0.0000	0.0000	0.0000	0.0000	0.0002	105.7022	0.0005	0.0000	0.0000	0.0000	1998	2.5	2.5862	50.0916	61.295
j518i593	36	518	593	41.4810	0.0000	0.0000	0.0000	0.0000	0.0001	110.0179	0.0005	0.0000	0.0000	0.0000	1998	2.5	2.5419	49.8057	60.746
j518i593	37	518	593	41.0500	0.0000	0.0000	0.0000	0.0000	0.0000	119.4127	0.0008	0.0000	0.0000	0.0000	1998	2.5	2.4875	50.0267	61.039

Figure A.2: MarMOT bioegochemical initial conditions at station ALOHA. The abbreviations will be explain in section A.9.

```
j518i593 2.5 7 473786.8629 -0.08423283 24.06927776
j518i593 2.5 8 9.604214999 0.021365698 24.06927299
j518i593 2.5 9 8.069029043 0.124905545 24.06926155
j518i593 2.5 10 1.006594225 0.203718756 24.06923294
j518i593 2.5 11 0.880308071 0.23515804 24.06899071
j518i593 2.5 12 0.864 0.248250412 23.72561264
j518i593 2.5 13 0.864 0.352519538 23.05021095
j518i593 2.5 14 0.864 0.195539412 22.44806957
j518i593 2.5 15 0.864 0.081666453 21.81431198
j518i593 2.5 16 0.864 0.073316538 21.12526035
j518i593 2.5 17 0.864 0.087504472 20.41242981
j518i593 2.5 18 0.864 0.100361777 19.66542912
j518i593 2.5 19 0.864 0.090112173 18.76627064
j518i593 2.5 20 0.864 0.060468195 17.60260677
j518i593 2.5 21 0.864 0.020820152 16.13798523
j518i593 2.5 22 0.864 -0.009596932 14.4914546
j518i593 2.5 23 0.864 -0.026308592 12.81142092
j518i593 2.5 24 0.864 0.000615318 11.22957325
j518i593 2.5 25 0.864 0.01186705 9.798433304
j518i593 2.5 26 0.864 0.051218655 8.544034004
j518i593 2.5 27 0.864 0.071805221 7.498681068
j518i593 2.5 28 0.864 0.098379366 6.645936489
j518i593 2.5 29 0.864 0.132428608 5.969537258
j518i593 2.5 30 0.864 0.335099775 5.439800501
j518i593 2.5 31 0.864 0.401625174 4.997122049
j518i593 2.5 32 0.864 0.530076729 4.61379385
j518i593 2.5 33 0.864 0.621780285 4.266488552
j518i593 2.5 34 0.864 0.667716526 3.947635889
j518i593 2.5 35 0.864 0.701440061 3.651248097
j518i593 2.5 36 0.864 0.715424639 3.370245814
j518i593 2.5 37 0.864 0.703022569 3.100137234
```

A.9 Output files

There are several output files from MarMOT that are specified in the experiment control table. The first one is the simulation options, which summarises the option (the choice of biogeochemical model, advection, and diffusion schemes) and environment parameter set item. The 'simvar' files are the summary of output variables from the simulation, with `simvar_t_medusa` and `simvar_kt_medusa` containing the scalar and profile variables. Both files contain similar information, however in the scalar file, primary production, solar radiation, and soluble iron are simulated, but the depth level is not. The example for `simvar_t_medusa` is shown below:

```
NAME VALUE


t 1 # time from t axis origin (days)
tmid 0 # time from t axis origin at time step mid-point (days)
year 0 # calendar year
tofyr 0 # time since start of calendar year (days)
z 1 # depth at level mid-point (m)
```

```
zbot 0 # depth at bottom of level (m)
sol 1 # surface solar irradiance forcing (W/m2)
fedust 1 # aeolian input of soluble iron (mmolFe/m2/d)


vdc 1 # vertical diffusion coefficient forcing (m2/d, depth zbot)
w 1 # vertical velocity forcing (m/d, upward, depth zbot)
temp 1 # temperature forcing (degC)
din 1 # dissolved inorganic nitrogen (mmolN/m3 )
pndia 1 # non-diatom phytoplankton nitrogen (mmolN/m3 )
pdiat 1 # diatom nitrogen (mmolN/m3 )
zmi 1 # microzooplankton nitrogen (mmolN/m3 )
zme 1 # mesozooplankton nitrogen (mmolN/m3 )
det 1 # detrital nitrogen (mmolN/m3 )
dsi 1 # dissolved silicon (silicic acid) (mmolSi/m3 )
dfe 1 # dissolved iron (mmolFe/m3 )
rcchlndia 0 # non-diatom C:Chl ratio (gC/gChl)
rcchldiat 0 # diatom C:Chl ratio (gC/gChl)
rsinpdiat 0 # diatom Si:N ratio
chlndia 1 # non-diatom chlorophyll (mg/m3 )
chldiat 1 # diatom chlorophyll (mg/m3 )
pdiatsi 1 # diatom silicon (mmolSi/m3 )
pprod 1 # level mean primary production (mmolC/m3 /d)
sinkflxc 0 # sinking particle flux of carbon (mmolC/m2/d, downward, depth
zbot
```

Variable names that are set to 0 means that the variable is not shown in the output.

Another output files are `dprof_aloha3_def` and `dscal_aloha3_def`. The example files are shown in Figures A.3 and A.4, respectively.

casenum	site	member	t	k	z	vdc	w	temp	din	pndia	pdiat	zmi	zme	det	dsi	dfe	chlndia	chldiat	pdiatsi
1	j518i593	_	0.5	1	3.03193	864000	-0.05286	24.0692	0.017278	0.245656	0	0.101223	0.11418	0.008759	1.2667	0.000654	0.120835	0	0
1	j518i593	_	0.5	2	9.16658	864000	-0.08545	24.0692	0.016815	0.245767	0	0.101225	0.114183	0.008767	1.26403	0.000654	0.120876	0	0
1	j518i593	_	0.5	3	15.4545	849651	-0.12989	24.0692	0.016868	0.245875	0	0.10123	0.114189	0.008768	1.26667	0.000655	0.120917	0	0
1	j518i593	_	0.5	4	21.9207	815080	-0.15831	24.0693	0.017026	0.247076	0	0.101238	0.114198	0.008774	1.26629	0.000656	0.121342	0	0
1	j518i593	_	0.5	5	28.5947	790593	-0.16188	24.0693	0.017027	0.245607	0	0.10126	0.114225	0.008782	1.26592	0.000655	0.120869	0	0
1	j518i593	_	0.5	6	35.5103	652347	-0.15194	24.0693	0.017237	0.244264	0	0.101978	0.115085	0.00879	1.26613	0.000654	0.121559	0	0
1	j518i593	_	0.5	7	42.7071	473787	-0.08423	24.0693	0.017239	0.241732	0	0.104074	0.117539	0.008795	1.26469	0.000653	0.121351	0	0
1	j518i593	_	0.5	8	50.2308	9.60421	0.021366	24.0693	0.017031	0.241057	0	0.106001	0.119607	0.0088	1.26562	0.000652	0.121884	0	0
1	j518i593	_	0.5	9	58.1343	8.06903	0.124906	24.0693	0.025632	0.079889	0	0.216176	0.272864	0.009145	1.27294	0.000512	0.147114	0	0
1	j518i593	_	0.5	10	66.479	1.00659	0.203719	24.0692	0.023841	0.06777	0	0.148331	0.19961	0.00762	1.26246	0.000459	0.167566	0	0
1	j518i593	_	0.5	11	75.3358	0.880308	0.235158	24.069	0.029115	0.070655	0	0.073825	0.094171	0.011967	1.37078	0.00041	0.189211	0	0
1	j518i593	_	0.5	12	84.7867	0.864	0.24825	23.7256	0.049556	0.075441	0	0.029528	0.031447	0.067328	1.38907	0.00041	0.209168	0	0
1	j518i593	_	0.5	13	94.9257	0.864	0.35252	23.0502	0.079583	0.07368	0	0.007603	0.006057	0.233345	1.50441	0.00041	0.209514	0	0
1	j518i593	_	0.5	14	105.862	0.864	0.195539	22.4481	0.143397	0.069068	0	0.001645	0.000746	0.392211	1.51867	0.00041	0.201322	0	0
1	j518i593	_	0.5	15	117.718	0.864	0.081667	21.8143	0.304416	0.061023	0	0.063679	0.009059	0.387443	1.64695	0.000409	0.186038	0	0
1	j518i593	_	0.5	16	130.637	0.864	0.073317	21.1253	0.358817	0.045724	0	0.013648	0.00321	0.32329	1.82553	0.000408	0.147357	0	0
1	j518i593	_	0.5	17	144.783	0.864	0.087505	20.4124	0.60591	0.046929	0	0.001659	0.000668	0.257662	2.07437	0.000407	0.086695	0	0
1	j518i593	_	0.5	18	160.337	0.864	0.100362	19.6654	0.547882	0.025588	0	6.34E-05	4.22E-05	0.201828	1.82233	0.000406	0.056916	0	0
1	j518i593	_	0.5	19	177.51	0.864	0.090112	18.7663	1.54581	0.012137	0	-1.87E-05	-1.13E-05	0.157801	2.45476	0.000405	0.027273	0	0
1	j518i593	_	0.5	20	196.537	0.864	0.060468	17.6026	1.87648	0.011139	0	-1.93E-05	-1.29E-05	0.123585	3.15366	0.000404	0.025057	0	0
1	j518i593	_	0.5	21	217.678	0.864	0.02082	16.138	2.87277	0.00422	0	-1.84E-05	-1.26E-05	0.096333	3.81712	0.000403	0.009512	0	0
1	j518i593	_	0.5	22	241.227	0.864	-0.0096	14.4915	4.32753	-2.61E-05	0	-1.65E-05	-1.14E-05	0.074	4.89022	0.000402	-5.91E-05	0	0
1	j518i593	_	0.5	23	267.506	0.864	-0.02631	12.8114	5.78861	-2.46E-05	0	-1.38E-05	-9.38E-06	0.05546	6.67306	0.000402	-5.63E-05	0	0
1	j518i593	_	0.5	24	296.869	0.864	0.000615	11.2296	8.40278	-1.95E-05	0	-1.08E-05	-7.15E-06	0.040167	9.06747	0.000401	-4.51E-05	0	0
1	j518i593	_	0.5	25	329.697	0.864	0.011867	9.79843	12.473	-1.47E-05	0	-7.79E-06	-5.09E-06	0.027895	11.8276	0.0004	-3.41E-05	0	0
1	j518i593	_	0.5	26	366.397	0.864	0.051219	8.54403	14.8039	-1.05E-05	0	-5.27E-06	-3.43E-06	0.018502	15.2139	0.000413	-2.42E-05	0	0
1	j518i593	_	0.5	27	407.402	0.864	0.071805	7.49868	18.3582	-7.01E-06	0	-3.32E-06	-2.18E-06	0.011753	23.3079	0.000427	-1.58E-05	0	0
1	j518i593	_	0.5	28	453.161	0.864	0.098379	6.64594	22.6169	-4.35E-06	0	-1.93E-06	-1.29E-06	0.007229	33.5722	0.00044	-9.40E-06	0	0
1	j518i593	_	0.5	29	504.129	0.864	0.132429	5.96954	26.4016	-2.43E-06	0	-9.98E-07	-6.91E-07	0.004363	46.3225	0.000452	-4.81E-06	0	0
1	j518i593	_	0.5	30	560.761	0.864	0.3351	5.4398	31.871	-1.15E-06	0	-4.36E-07	-3.22E-07	0.002606	60.8862	0.000464	-1.89E-06	0	0
1	j518i593	_	0.5	31	623.501	0.864	0.401625	4.99712	36.5256	-4.06E-07	0	-1.37E-07	-1.19E-07	0.001543	71.8295	0.000476	-4.29E-07	0	0
1	j518i593	_	0.5	32	692.765	0.864	0.530077	4.61379	39.8298	-6.84E-08	0	-2.58E-08	-2.79E-08	0.000904	86.4094	0.000488	-2.36E-08	0	0
1	j518i593	_	0.5	33	768.929	0.864	0.62178	4.26649	41.0814	-8.93E-09	0	-5.08E-09	-3.78E-09	0.00052	94.9494	0.0005	-9.74E-09	0	0
1	j518i593	_	0.5	34	852.314	0.864	0.667717	3.94764	41.5297	-6.57E-09	0	-3.22E-09	-1.31E-09	0.000292	101.819	0.00049	-1.01E-08	0	0
1	j518i593	_	0.5	35	943.175	0.864	0.70144	3.65125	41.603	-6.18E-09	0	-2.77E-09	-1.07E-09	0.000159	105.713	0.00048	-9.82E-09	0	0
1	j518i593	_	0.5	36	1041.69	0.864	0.715425	3.37025	41.4791	-5.75E-09	0	-2.55E-09	-9.90E-10	8.36E-05	110.063	0.000471	-9.18E-09	0	0
1	j518i593	_	0.5	37	1147.94	0	0	3.10014	41.0501	-5.43E-09	0	-2.38E-09	-9.32E-10	2.78E-05	119.412	0.00076	-8.64E-09	0	0
1	j518i593	_	1.5	1	3.03193	864000	-0.05286	24.0692	0.026648	0.227071	0	0.10245	0.114754	0.015818	1.26588	0.000651	0.112127	0	0

Figure A.3: MarMOT-MEDUSA bioegochemical output for each depth level on the first day of the run at station ALOHA. The abbreviations have been explained above.

casenum	site	member	t	z	sol	fedust	vdc	w	temp	din	pndia	pdiat	zmi	zme	det	dsi	dfe	chlndia	chldiat	pdiatsi	pprod
1	j518i593	_	0.5	3.03193	140.713	0	864000	-0.05286	24.0692	0.017278	0.245656	0	0.101223	0.11418	0.008759	1.2667	0.000654	0.120835	0	0	0.097335
1	j518i593	_	1.5	3.03193	140.713	0	864000	-0.05286	24.0692	0.026648	0.227071	0	0.10245	0.114754	0.015818	1.26588	0.000651	0.112127	0	0	0.13588
1	j518i593	_	2.5	3.03193	140.851	0	864000	-0.05058	24.0627	0.035115	0.214564	0	0.101532	0.113159	0.019746	1.26598	0.000649	0.104664	0	0	0.16609
1	j518i593	_	3.5	3.03193	141.814	0	864000	-0.03456	24.0175	0.042359	0.205653	0	0.100022	0.111034	0.021791	1.26607	0.000646	0.098589	0	0	0.188491
1	j518i593	_	4.5	3.03193	142.915	0	864000	-0.01625	23.9658	0.048414	0.199385	0	0.098164	0.108573	0.022725	1.26615	0.000644	0.093661	0	0	0.205087
1	j518i593	_	5.5	3.03193	144.016	0	864000	0.002059	23.9141	0.053402	0.195058	0	0.096115	0.105911	0.023022	1.26623	0.000642	0.089665	0	0	0.217501
1	j518i593	_	6.5	3.03193	145.117	0	864000	0.020367	23.8624	0.057464	0.192164	0	0.093977	0.103142	0.02297	1.26632	0.00064	0.08642	0	0	0.226894
1	j518i593	_	7.5	3.03193	146.344	0	864000	0.038459	23.8118	0.064686	0.181604	0	0.092628	0.103408	0.022364	1.26731	0.000627	0.08619	0	0	0.241896
1	j518i593	_	8.5	3.03193	148.329	0	864000	0.055263	23.7685	0.072198	0.170895	0	0.09095	0.103587	0.021398	1.26876	0.000612	0.086901	0	0	0.258924
1	j518i593	_	9.5	3.03193	150.44	0	864000	0.071851	23.7264	0.074183	0.171792	0	0.088156	0.099741	0.02091	1.26934	0.000611	0.085035	0	0	0.264609
1	j518i593	_	10.5	3.03193	152.551	0	864000	0.088439	23.6843	0.075648	0.172791	0	0.085457	0.096013	0.02051	1.27012	0.00061	0.08359	0	0	0.269118
1	j518i593	_	11.5	3.03193	154.663	0	864000	0.105027	23.6421	0.076701	0.173842	0	0.082848	0.092402	0.020181	1.2711	0.000608	0.082476	0	0	0.27271
1	j518i593	_	12.5	3.03193	156.435	0	864000	0.120422	23.6028	0.081803	0.166698	0	0.076451	0.084744	0.019202	1.282	0.000591	0.085985	0	0	0.281934
1	j518i593	_	13.5	3.03193	156.173	0	864000	0.12866	23.5801	0.087671	0.158627	0	0.069393	0.076466	0.017777	1.29526	0.000571	0.090383	0	0	0.292627
1	j518i593	_	14.5	3.03193	155.573	0	864000	0.135705	23.5603	0.087939	0.161249	0	0.06748	0.07376	0.017349	1.29616	0.00057	0.089172	0	0	0.294475
1	j518i593	_	15.5	3.03193	154.972	0	864000	0.14275	23.5404	0.088433	0.163525	0	0.06558	0.071076	0.017088	1.29741	0.000569	0.088154	0	0	0.29654
1	j518i593	_	16.5	3.03193	154.372	0	864000	0.149796	23.5205	0.089169	0.165528	0	0.063709	0.068436	0.016935	1.29898	0.000568	0.087314	0	0	0.299005
1	j518i593	_	17.5	3.03193	153.91	0	864000	0.158032	23.502	0.098896	0.16374	0	0.05987	0.063542	0.016578	1.30878	0.000561	0.088719	0	0	0.321879
1	j518i593	_	18.5	3.03193	154.279	0	864000	0.173416	23.4914	0.110328	0.162299	0	0.05575	0.058341	0.016052	1.32066	0.000552	0.091361	0	0	0.354031
1	j518i593	_	19.5	3.03193	154.786	0	864000	0.189991	23.4821	0.110531	0.165729	0	0.054288	0.056211	0.015967	1.32268	0.000551	0.091669	0	0	0.359811
1	j518i593	_	20.5	3.03193	155.294	0	864000	0.206567	23.4728	0.110575	0.169056	0	0.052944	0.054225	0.01602	1.32465	0.00055	0.091944	0	0	0.364947
1	j518i593	_	21.5	3.03193	155.802	0	864000	0.223142	23.4636	0.110538	0.172257	0	0.051707	0.052368	0.016169	1.32659	0.000549	0.092181	0	0	0.369599
1	j518i593	_	22.5	3.03193	156.593	0	864000	0.239628	23.4547	0.11047	0.175324	0	0.050565	0.050626	0.016383	1.32852	0.000548	0.09238	0	0	0.374147
1	j518i593	_	23.5	3.03193	159.086	0	864000	0.255582	23.448	0.110291	0.178337	0	0.049512	0.04899	0.016643	1.3304	0.000546	0.09256	0	0	0.379794
1	j518i593	_	24.5	3.03193	161.863	0	864000	0.271448	23.4418	0.109918	0.181358	0	0.048545	0.047452	0.016939	1.33223	0.000545	0.092729	0	0	0.385124
1	j518i593	_	25.5	3.03193	164.64	0	864000	0.287313	23.4355	0.109376	0.184366	0	0.047656	0.046004	0.017264	1.33399	0.000545	0.092878	0	0	0.389883
1	j518i593	_	26.5	3.03193	167.417	0	864000	0.303178	23.4292	0.108692	0.187338	0	0.046834	0.044638	0.017611	1.33568	0.000544	0.092997	0	0	0.394082
1	j518i593	_	27.5	3.03193	170.066	0	864000	0.314234	23.4219	0.107895	0.190253	0	0.046075	0.043347	0.017974	1.33731	0.000543	0.093081	0	0	0.397623
1	j518i593	_	28.5	3.03193	171.951	0	864000	0.296437	23.4081	0.107099	0.193046	0	0.045367	0.042121	0.018348	1.3389	0.000542	0.093119	0	0	0.400104
1	j518i593	_	29.5	3.03193	173.709	0	864000	0.273831	23.3933	0.106403	0.195672	0	0.044703	0.040949	0.018725	1.34048	0.000541	0.09311	0	0	0.402409
1	j518i593	_	30.5	3.03193	175.467	0	864000	0.251225	23.3785	0.105815	0.198143	0	0.044077	0.039829	0.019097	1.34208	0.00054	0.093066	0	0	0.404732
1	j518i593	_	31.5	3.03193	177.225	0	864000	0.228619	23.3637	0.105334	0.200474	0	0.043487	0.038755	0.01946	1.34367	0.00054	0.092994	0	0	0.40712
1	j518i593	_	32.5	3.03193	179.023	0	864000	0.208107	23.3477	0.10495	0.202682	0	0.042929	0.037724	0.019813	1.34528	0.000539	0.092906	0	0	0.409612
1	j518i593	_	33.5	3.03193	181.064	0	864000	0.200159	23.3247	0.104534	0.204827	0	0.042406	0.036741	0.020159	1.34683	0.000538	0.092809	0	0	0.411969
1	j518i593	_	34.5	3.03193	183.145	0	864000	0.194305	23.3005	0.103961	0.206949	0	0.041922	0.03581	0.020503	1.34828	0.000537	0.092702	0	0	0.41375
1	j518i593	_	35.5	3.03193	185.226	0	864000	0.188451	23.2763	0.103235	0.209032	0	0.041475	0.034927	0.020848	1.3496	0.000537	0.092571	0	0	0.4149
1	j518i593	_	36.5	3.03193	187.307	0	864000	0.182597	23.2521	0.102371	0.211058	0	0.041064	0.03409	0.021191	1.3508	0.000536	0.092407	0	0	0.415416
1	j518i593	_	37.5	3.03193	189.142	0	864000	0.18138	23.2296	0.100853	0.21335	0	0.0407	0.033311	0.021529	1.35174	0.000536	0.092284	0	0	0.413971

Figure A.4: MarMOT-MEDUSA bioegochemical output for each day at station ALOHA. The abbreviations have been explained above.

Appendix B

Phytoplankton Distribution

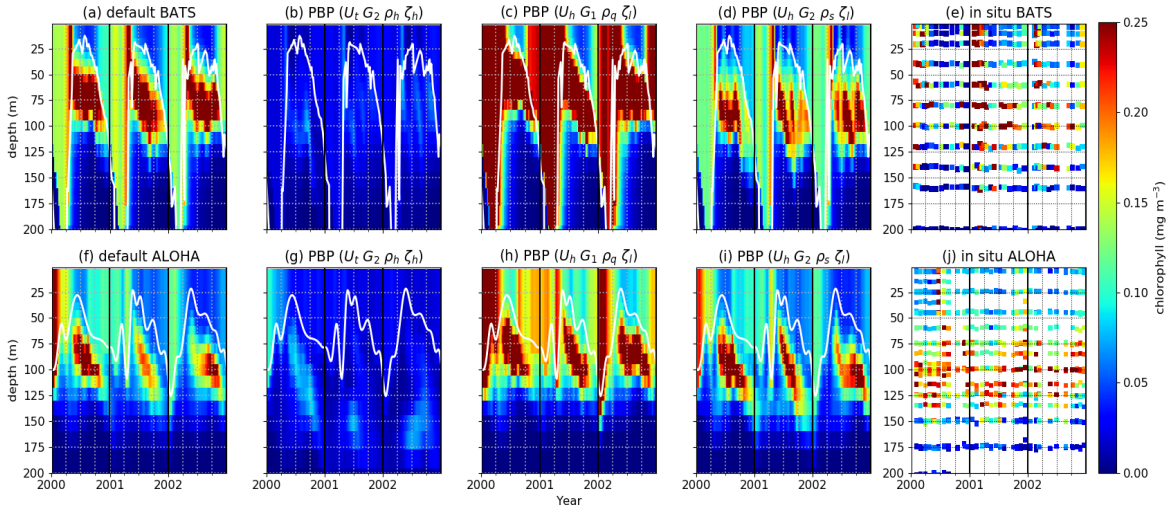


Figure B.1: Chlorophyll distribution in the water column from 1st January 2000 to 31st December 2002 at station BATS (a to e) and ALOHA (f to j) from selected PBPE members. White solid lines are the mixed layer depth. The default run is shown in (a) and (f) for BATS and ALOHA respectively. Different ensemble members from perturbing the biogeochemistry with their functional forms combinations are shown in (b), (c), and (d), for ALOHA, and (g), (h), and (i) for BATS. (e) and (j) are observed chlorophyll from BATS and ALOHA, respectively.

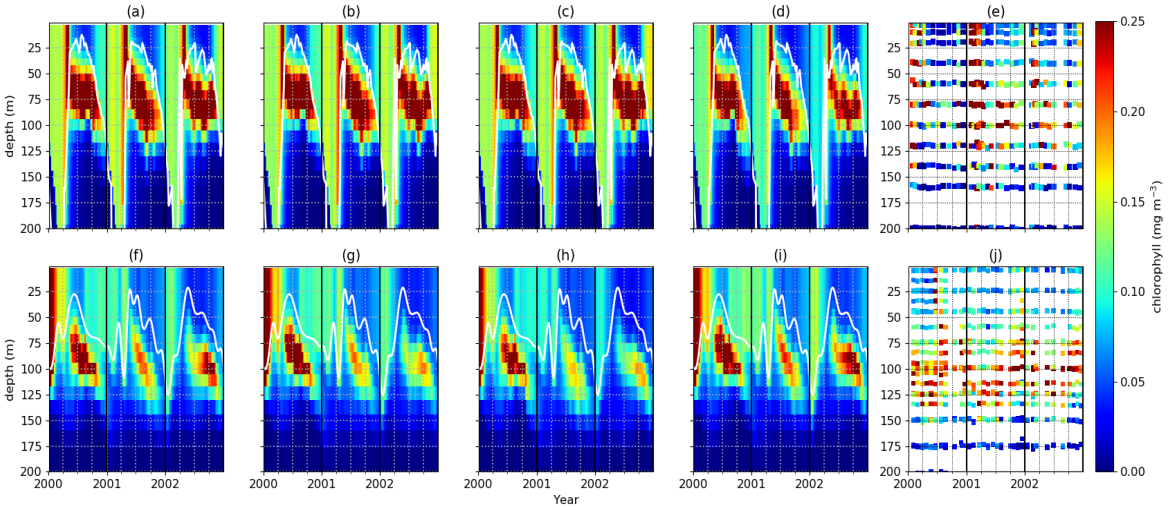


Figure B.2: Similar to Figure B.1, but from PPE members

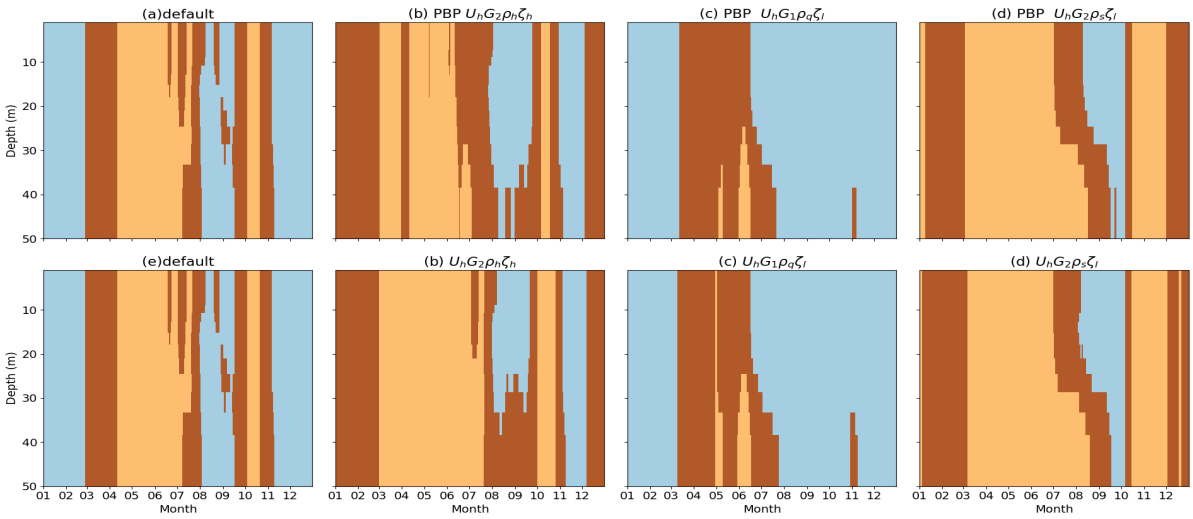


Figure B.3: Distribution of dominant phytoplankton type at station L4 in 2003, form PBPE (a to e) and PBE (e to h). Light blue and cream denotes non diatoms and diatoms, respectively as dominant (concentration is larger than 60% of the mean total chlorophyll) phytoplankton type in the water column. Brown denotes that none of the PFT reached 60% of the mean total concentration.

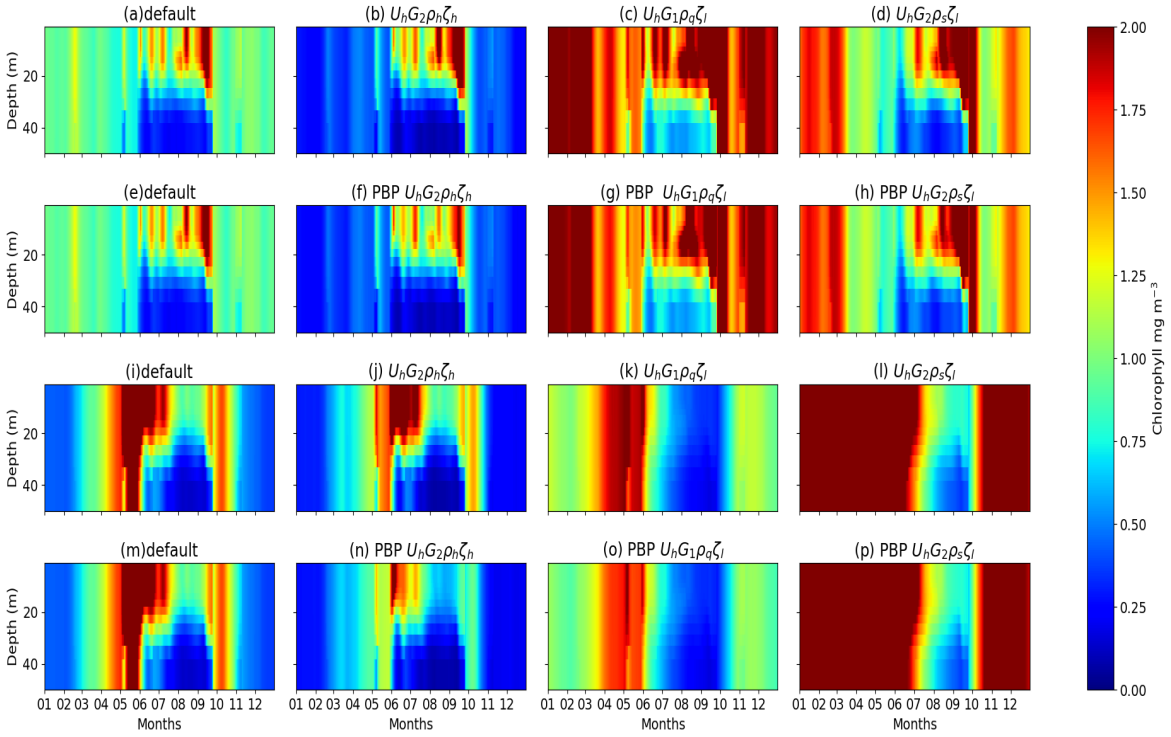


Figure B.4: Chlorophyll profiles of diatom (a to h) and non-diatoms (i to p) at station L4 form the PBE (a to d for non-diatoms and i to l for diatoms) and PBPE (e to h for non-diatoms and m to p for diatoms) members in 2003. Default run chlorophyll distributions are shown in (a), (e), (i), and (m).

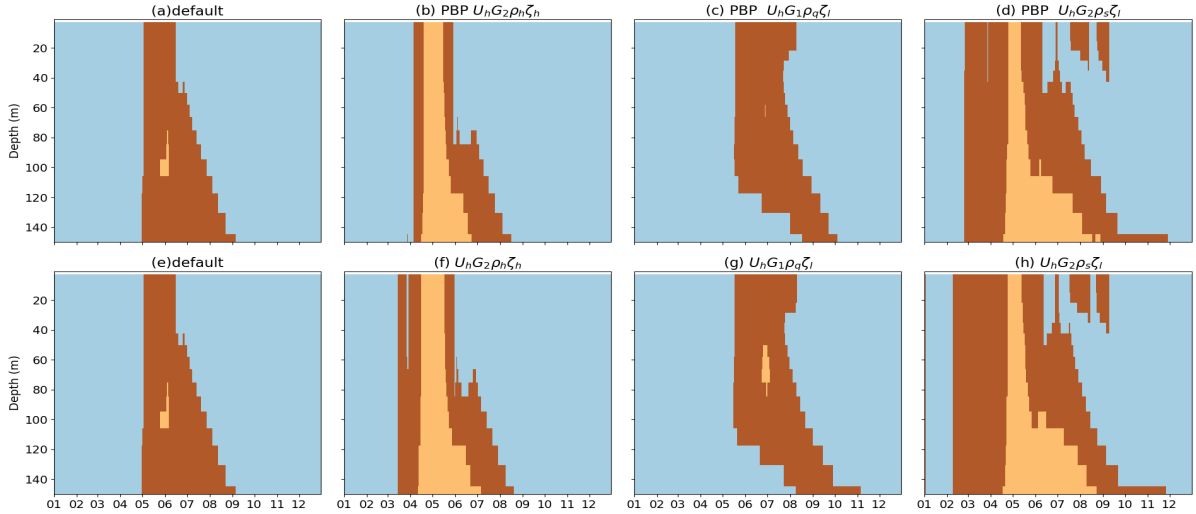


Figure B.5: Similar to B.3, but at PAP.

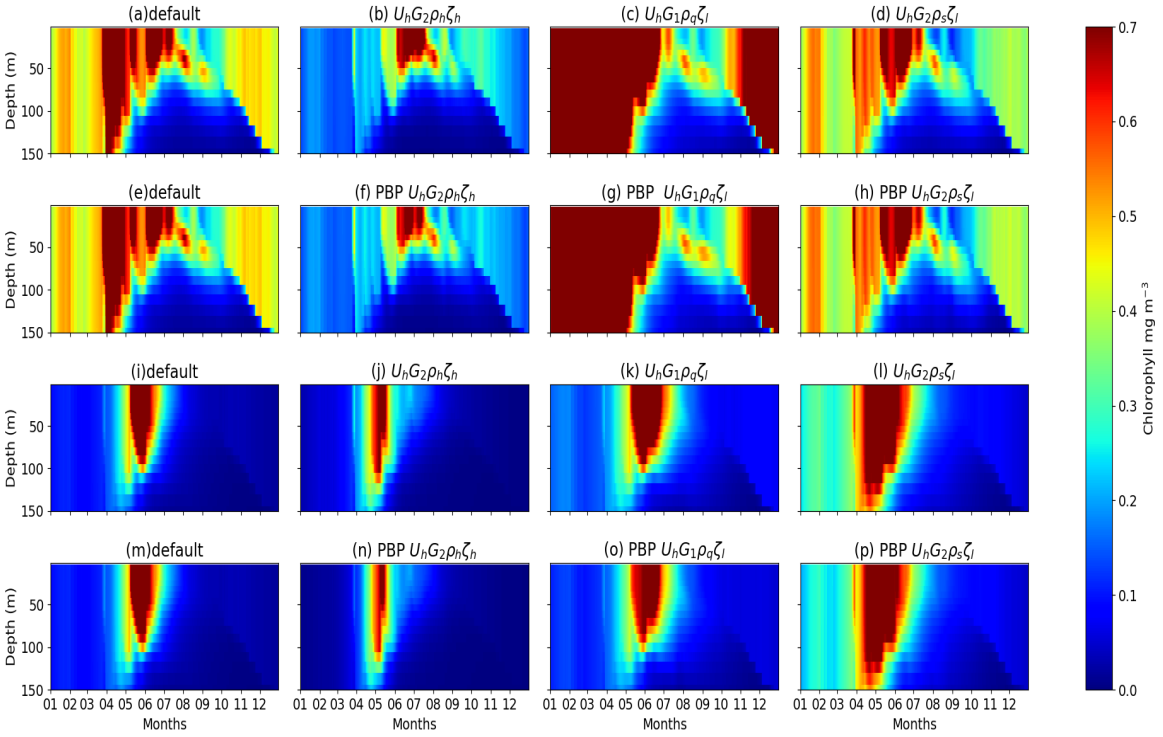


Figure B.6: Similar to B.4, but at PAP

Bibliography

Abramowitz, G., Herger, N., Gutmann, E., Hammerling, D., Knutti, R., Leduc, M., Lorenz, R., Pincus, R. and Schmidt, G. A. (2019), 'ESD Reviews: Model dependence in multi-model climate ensembles: Weighting, sub-selection and out-of-sample testing'.

Adamson, M. W. (2015), Using partially specified models to detect and quantify structural sensitivity in biological systems, PhD thesis, University of Leicester.

Adamson, M. W. and Morozov, A. Y. (2013), 'When can we trust our model predictions? Unearthing structural sensitivity in biological systems', *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **469**(2149), 20120500.

Aiken, J., Rees, N., Hooker, S., Holligan, P., Bale, A., Robins, D., Moore, G., Harris, R. and Pilgrim, D. (2000), 'The Atlantic Meridional Transect: Overview and synthesis of data', *Progress in Oceanography* **45**(3-4), 257–312.

Aldebert, C., Kooi, B. W., Nerini, D. and Poggiale, J. C. (2018), 'Is structural sensitivity a problem of oversimplified biological models? Insights from nested Dynamic Energy Budget models', *Journal of Theoretical Biology* **448**, 1–8.

Aldebert, C., Nerini, D., Gauduchon, M. and Poggiale, J. C. (2016), 'Does structural sensitivity alter complexitystability relationships?', *Ecological Complexity* **28**, 104–112.

Aldebert, C. and Stouffer, D. B. (2018), 'Community dynamics and sensitivity to model structure: Towards a probabilistic view of process-based model predictions', *Journal of the Royal Society Interface* **15**(149).

Allen, J. I., Aiken, J., Anderson, T. R., Buitenhuis, E., Cornell, S., Geider, R. J., Haines, K., Hirata, T., Holt, J., Le Quéré, C., Hardman-Mountford, N., Ross, O. N., Sinha, B. and While, J. (2010), 'Marine ecosystem models for earth systems applications: The MarQUEST experience', *Journal of Marine Systems* **81**(1-2), 19–33.

Anderson, D., Sheinbaum, J. and Haines, K. (1996), 'Data Assimilation in Ocean models', *Rep. Prog. Phys.* **59**, 1665–1735.

Anderson, J. L. (2001), 'An Ensemble Adjustment Kalman Filter for Data Assimilation', *Monthly Weather Review* **129**(12), 2884–2903.

Anderson, T. (2019), 'The beauty of equations', *Ocean Challenge* pp. 20–22.

Anderson, T. R. (2005), 'Plankton functional type modelling: Running before we can walk?', *Journal of Plankton Research* **27**(11), 1073–1081.

Anderson, T. R. (2010), 'Progress in marine ecosystem modelling and the "unreasonable effectiveness of mathematics"', *Journal of Marine Systems* **81**(1-2), 4–11.

Anderson, T. R., Gentleman, W. C. and Sinha, B. (2010), 'Influence of grazing formulations on the emergent properties of a complex ecosystem model in a global ocean general circulation model', *Progress in Oceanography* **87**(1-4), 201–213.

Anderson, T. R., Gentleman, W. C. and Yool, A. (2015), 'EMPOWER-1.0: An Efficient Model of Planktonic ecOsystems WrittEn in R', *Geoscientific Model Development* **8**(7), 2231–2262.

Anderson, T. R. and Mitra, A. (2010), 'Dysfunctionality in ecosystem models: An underrated pitfall?', *Progress in Oceanography* **84**(February), 66–68.

Anderson, T. R. and Pondaven, P. (2003), 'Non-redfield carbon and nitrogen cycling in the Sargasso Sea: Pelagic imbalances and export flux', *Deep-Sea Research Part I: Oceanographic Research Papers* **50**(5), 573–591.

Auber, A., Travers-Trolet, M., Villanueva, M. C. and Ernande, B. (2015), 'Regime shift in an exploited fish community related to natural climate oscillations', *PLoS ONE* **10**(7), 1–18.

Aumont, O. and Bopp, L. (2006), 'Globalizing results from ocean in situ iron fertilization studies', *Global Biogeochemical Cycles* **20**(2), 1–15.

Baird, M. E. and Suthers, I. M. (2007), 'A size-resolved pelagic ecosystem model', *Ecological Modelling* **203**(3-4), 185–203.

Barber, R. T. (1988), Ocean Basin Ecosystem, *in* J. J. Alberts and L. R. Pomeroy, eds, 'Concepts of Ecosystem Ecology: A Comparative View', Vol. 67, Springer-Verlag, New York, chapter 9, pp. 171–193.

Bates, N. R. (2001), 'Interannual variability of oceanic CO and biogeochemical properties in the Western North Atlantic subtropical gyre', **48**.

Behrenfeld, M. J. and Boss, E. S. (2014), 'Resurrecting the ecological underpinnings of ocean plankton blooms.', *Annual review of marine science* **6**(September 2013), 167–94.

Behrenfeld, M. J., Doney, S. C., Lima, I., Boss, E. S. and Siegel, D. A. (2013), 'Annual cycles of ecological disturbance and recovery underlying the subarctic Atlantic spring plankton bloom', *Global Biogeochemical Cycles* **27**(4), 1291–1293.

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. a., McClain, C. R., Sarmiento, J. L., Feldman, G. C., Milligan, A. J., Falkowski, P. G., Letelier, R. M. and Boss, E. S. (2006), 'Climate-driven trends in contemporary ocean productivity.', *Nature* **444**(7120), 752–755.

Bennington, V., Mckinley, G. A., Dutkiewicz, S. and Ullman, D. (2009), 'What does chlorophyll variability tell us about export and air-sea CO 2 flux variability in the North Atlantic ?', **23**(July), 1–11.

Blondeau-Patissier, D., Gower, J. F., Dekker, A. G., Phinn, S. R. and Brando, V. E. (2014), 'A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans', *Progress in Oceanography* **123**, 23–144.

Bopp, L., Aumont, O., Cadule, P., Alvain, S. and Gehlen, M. (2005), 'Response of diatoms distribution to global warming and potential implications: A global model study', *Geophysical Research Letters* **32**(19), 1–4.

Bopp, L., Monfray, P., Aumont, O., Dufresne, J.-L., Le Treut, H., Madec, G., Terray, L. and Orr, J. C. (2001), 'Potential impact of climate change on marine export production', *Global Biogeochemical Cycles* **15**(1), 81–99.

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J. and Vichi, M. (2013), 'Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models', *Biogeosciences* **10**(10), 6225–6245.

Bower, A. S. and von Appen, W.-J. (2008), 'Interannual Variability in the Pathways of the North Atlantic Current over the Mid-Atlantic Ridge and the Impact of Topography', *Journal of Physical Oceanography* **38**(1), 104–120.

Brody, S. R., Lozier, M. S. and Dunne, J. P. (2013), 'A comparison of methods to determine phytoplankton bloom initiation', *Journal of Geophysical Research: Oceans* **118**(5), 2345–2357.

Burrows, M. T., Schoeman, D. S., Buckley, L. B., Moore, P., Poloczanska, E. S., Brander, K. M., Brown, C., Bruno, J. F., Duarte, C. M., Halpern, B. S., Holding, J., Kappel, C. V., Kiessling, W., O'Connor, M. I., Pandolfi, J. M., Parmesan, C., Schwing, F. B., Sydeman, W. J. and Richardson, A. J. (2011), 'The Pace of Shifting Climate in Marine and Terrestrial Ecosystems', *Science* **334**(6056), 652–655.

Butenschön, M., Clark, J., Aldridge, J. N., Icarus Allen, J., Artioli, Y., Blackford, J., Bruggeman, J., Cazenave, P., Ciavatta, S., Kay, S., Lessin, G., Van Leeuwen, S., Van Der Molen, J., De Mora, L., Polimene, L., Sailley, S., Stephens, N. and Torres, R. (2016), 'ERSEM 15.06: A generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels', *Geoscientific Model Development* **9**(4), 1293–1339.

Ciavatta, S., Torres, R., Martinez-Vicente, V., Smyth, T., Dall'Olmo, G., Polimene, L. and Allen, J. I. (2014), 'Assimilation of remotely-sensed optical properties to improve marine biogeochemistry modelling', *Progress in Oceanography* **127**, 74–95.

Cole, H. (2013), The natural variability and climate change response in phytoplankton phenology, PhD thesis, Univeristy of Southampton.

Cole, H., Henson, S., Martin, A. and Yool, A. (2012), 'Mind the gap: The impact of missing data on the calculation of phytoplankton phenology metrics', *Journal of Geophysical Research: Oceans* **117**(8), 2–9.

Corno, G., Karl, D. M., Church, M. J., Letelier, R. M., Lukas, R., Bidigare, R. R. and Abbott, M. R. (2007), 'Impact of climate forcing on ecosystem processes in the North Pacific Subtropical Gyre', *Journal of Geophysical Research: Oceans* **112**(4), 1–14.

Cossarini, G., Lermusiaux, P. F. J. and Solidoro, C. (2009), 'Lagoon of venice ecosystem: Seasonal dynamics and environmental guidance with uncertainty analyses and error subspace data assimilation', *Journal of Geophysical Research: Oceans* **114**(6).

Cox, G. M., Gibbons, J. M., Wood, A. T. A., Craigon, J., Ramsden, S. J. and Crout, N. M. J. (2006), 'Towards the systematic simplification of mechanistic models', *Ecological Modelling* **198**(1-2), 240–246.

Cox, P. M. and Kwiatkowski, L. (2013), Assessment of the iMarNet Ocean Biogeochemical Models, Technical report, University of Exeter, Exeter.

Crook, J. A., Jackson, L. S. and Forster, P. M. (2016), 'Can increasing albedo of existing ship wakes reduce climate change?', *Journal of Geophysical Research* **121**(4), 1549–1558.

Cullen, J. J. (2014), 'Subsurface Chlorophyll Maximum Layers: Enduring Enigma or Mystery Solved?', *Annual Review of Marine Science* **7**(1), 207–239.

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.

Dave, A. C. and Lozier, M. S. (2010), 'Local stratification control of marine productivity in the subtropical North Pacific', *Journal of Geophysical Research: Oceans* **115**(12), 1–16.

De Mora, L., Butenschön, M. and Allen, J. I. (2016), 'The assessment of a global marine ecosystem model on the basis of emergent properties and ecosystem function: A case study with ERSEM', *Geoscientific Model Development* **9**(1), 59–76.

Denman, K. L. (2003), 'Modelling planktonic ecosystems: parameterizing complexity', *Progress in Oceanography* **57**(3-4), 429–452.

Dickey, T. D. and Bidigare, R. R. (2005), 'Interdisciplinary oceanographic observations: the wave of the future', *Scientia Marina* **69**(S1), 23–42.

Doney, S. C. (1999), 'Major challenges confronting marine biogeochemical modeling', *Global Biogeochemical Cycles* **13**(3), 705–714.

Doney, S. C., Fabry, V. J., Feely, R. A. and Kleypas, J. A. (2009), 'Ocean acidification: the other CO2 problem.', *Annual review of marine science* **1**, 169–92.

Doney, S. C., Lindsay, K., Caldeira, K., Campin, J., Drange, H., Dutay, J., Follows, M., Gao, Y., Gnanadesikan, A., Gruber, N., Ishida, A. and Joos, F. (2004), 'Evaluating global ocean carbon models : The importance of realistic physics', *Global Biogeochemical Cycles* **18**(GB3017).

Doney, S. C., Ruckelshaus, M., Emmett Duffy, J., Barry, J. P., Chan, F., English, C. A., Galindo, H. M., Grebmeier, J. M., Hollowed, A. B., Knowlton, N., Polovina, J., Rabalais, N. N., Sydeman, W. J. and Talley, L. D. (2012), 'Climate Change Impacts on Marine Ecosystems', *Annual Review of Marine Science* **4**(1), 11–37.

Droop, M. R. (1973), 'Some thoughts on nutrinet limitation in algae', *Journal of Phycology* **9**, 264–272.

Droop, M. R. (1983), '25 years of algal growth kinetics', *Botanica Marina* (26), 12–99.

Ducklow, H. W. (2003), 'Biogeochemical Provinces: Towards a JGOFS Synthesis', *Ocean Biogeochemistry* pp. 3–17.

Dutkiewicz, S., Follows, M. J. and Bragg, J. G. (2009), 'Modeling the coupling of ocean ecology and biogeochemistry', *Global Biogeochemical Cycles* **23**(4), 1–15.

Edwards, A. M. (2001), 'Adding detritus to a nutrient-phytoplankton-zooplankton model: A dynamical-systems approach', *Journal of Plankton Research* **23**(4), 389–413.

Edwards, A. M. and Brindley, J. (1996), 'Oscillatory behaviour in a three-component plankton population model', *Dynamics and Stability of Systems* **11**(4), 347–370.

Edwards, A. and Yool, A. (2000), 'The role of higher predation in plankton population models', *Journal of Plankton Research* **22**(6), 1085–1112.

Edwards, M., Richardson, A. J. and Martin Edwards & Anthony J. Richardson (2004), 'Impact of climate change on marine pelagic phenology and trophic mismatch', *Nature* **430**(7002), 881–884.

Englund, G. and Leonardsson, K. (2008), 'Scaling up the functional response for spatially heterogeneous systems', *Ecology Letters* **11**(5), 440–449.

Eppley, R. W. (1972), 'Temperature and phytoplankton growth in the sea', *Fishery bulletin* **70**, 1063–1085.

Fasham, M. J. R. (1995), 'Variations in the seasonal cycle of biological production in subarctic oceans: A model sensitivity analysis', *Deep-Sea Research Part I* **42**(7), 1111–1149.

Fasham, M. J. R., Ducklow, H. W. and McKelvie, S. M. (1990), 'A nitrogen-based model of plankton dynamics in the ocean mixed layer', *Journal of Marine Research* **48**(3), 591–639.

Fasham, M. J. R., Sarmiento, J. L., Slater, R. D., Ducklow, H. W. and Williams, R. (1993), 'Ecosystem behaviour at Bermuda station"S" and ocean weather station "India": A general circulation moel and observational anlysis', *Global Biogeochemical Cycles* **7**(2), 379–415.

Fennel, K. and Boss, E. (2003), 'Subsurface maxima of phytoplankton and chlorophyll: Steady-state solutions from a simple model', *Limnology and Oceanography* **48**(4), 1521–1534.

Fennel, W. and Neumann, T. (2014), *Introduction to the Modelling of Marine Ecosystem*, second edn, Elsevier Science.

Fleming, R. H. (1939), 'The control of diatom populations by grazing', *ICES Journal of Marine Science* **14**(2), 210–227.

Flora, C., David, N., Mathias, G., Andrew, M. and Jean-Christophe Poggiale, P. (2011), 'Structural sensitivity of biological models revisited', *Journal of Theoretical Biology* **283**(1), 82–91.

Flynn, K. J. (2018), *Dynamic Ecology - an introduction to the art of simulating trophic dynamics*.

Flynn, K. J. and Mitra, A. (2016), 'Why Plankton Modelers Should Reconsider Using Rectangular Hyperbolic (Michaelis-Menten, Monod) Descriptions of Predator-Prey Interactions', *Frontiers in Marine Science* **3**(September).

Follows, M. J. and Dutkiewicz, S. (2011), 'Modeling diverse communities of marine microbes.', *Annual review of marine science* **3**, 427–451.

Follows, M. J., Dutkiewicz, S., Grant, S. and Chisholm, S. W. (2007), 'Emergent Biogeography of Microbial', *Science* (C), 1843–1847.

Ford, D. A., Edwards, K. P., Lea, D., Barciela, R. M., Martin, M. J. and Demaria, J. (2012), 'Assimilating GlobColour ocean colour data into a pre-operational physical-biogeochemical model', *Ocean Science* **8**, 751–771.

Ford, D. and Barciela, R. (2017), 'Global marine biogeochemical reanalyses assimilating two different sets of merged ocean colour products', *Remote Sensing of Environment* **203**, 40–54.

Franks, P. J. S. (2002), 'NPZ models of plankton dynamics: Their construction, coupling to physics, and application', *Journal of Oceanography* **58**(2), 379–387.

Friedrichs, M. A., Dusenberry, J. A., Anderson, L. A., Armstrong, R. A., Chai, F., Christian, J. R., Doney, S. C., Dunne, J., Fujii, M., Hood, R., McGillicuddy, D. J., Moore, J. K., Schartau, M., Spitz, Y. H. and Wiggert, J. D. (2007), 'Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups', *Journal of Geophysical Research: Oceans* **112**(8), 1–22.

Friedrichs, M. A. M., Hood, R. R. and Wiggert, J. D. (2006), 'Ecosystem model complexity versus physical forcing: Quantification of their relative impact with assimilated Arabian Sea data', *Deep-Sea Research Part II: Topical Studies in Oceanography* **53**(5-7), 576–600.

Furnas, M. J. (1990), 'In situ growth-rates of marine-phytoplankton: approaches to measurement, community and species growth rates', *Journal of Plankton Research* **12**(6), 1117–1151.

Fussmann, G. F. and Blasius, B. (2005), 'Community response to enrichment is highly sensitive to model structure.', *Biology letters* **1**(1), 9–12.

Gehlen, M., Barciela, R., Bertino, L., Brasseur, P., Butenschön, M., Chai, F., Crise, A., Drillet, Y., Ford, D., Lavoie, D., Lehodey, P., Perruche, C., Samuelsen, A. and Simon, E. (2015), 'Building the capacity for forecasting marine biogeochemistry and ecosystems: recent advances and future developments', *Journal of Operational Oceanography* **8**(sup1), s168–s187.

Gentleman, W. (2002), 'A chronology of plankton dynamics in silico: How computer models have been used to study marine ecosystems', *Hydrobiologia* **480**, 69–85.

Gentleman, W., Leising, A., Frost, B., Strom, S. and Murray, J. (2003), 'Functional responses for zooplankton feeding on multiple resources: A review of assumptions and biological dynamics', *Deep-Sea Research Part II: Topical Studies in Oceanography* **50**(22-26), 2847–2875.

Greve, W., Prinage, S., Zidowitz, H., Nast, J. and Reiners, F. (2005), 'On the phenology of North Sea ichthyoplankton', *ICES Journal of Marine Science* **62**(7), 1216–1223.

Griffies, S. M., Böning, C., Bryan, F. O., Chassignet, E. P., Gerdes, R., Hasumi, H., Hirst, A., Treguier, A.-M. and Webb, D. (2000), 'Developments in ocean climate modelling', *Ocean Modelling* **2**(3-4), 123–192.

Gross, T. and Feudel, U. (2006), 'Generalized models as a universal approach to the analysis of nonlinear dynamical systems', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* **73**(1), 1–14.

Halloran, P. R., Bell, T. G. and Totterdell, I. J. (2010), 'Can we trust empirical marine DMS parameterisations within projections of future climate?', *Biogeosciences* **7**(5), 1645–1656.

Harrison, W. G., Head, E. J., Horne, E. P., Irwin, B., Li, W. K., Longhurst, A. R., Paranjape, M. A. and Platt, T. (1993), 'The western North Atlantic bloom experiment', *Deep-Sea Research Part II* **40**(1-2), 279–305.

Hartman, S. E., Jiang, Z. P., Turk, D., Lampitt, R. S., Frigstad, H., Ostle, C. and Schuster, U. (2015), 'Biogeochemical variations at the Porcupine Abyssal Plain sustained Observatory in the northeast Atlantic Ocean, from weekly to inter-annual timescales', *Biogeosciences* **12**(3), 845–853.

Hartman, S. E., Larkin, K. E., Lampitt, R. S., Lankhorst, M. and Hydes, D. J. (2010), 'Seasonal and inter-annual biogeochemical variations in the Porcupine Abyssal Plain 2003-2005 associated with winter mixing and surface circulation', *Deep-Sea Research Part II: Topical Studies in Oceanography* **57**(15), 1303–1312.

Hawaii Ocean Time Series (2017), 'Analytical methods and results'. `http://hahana.soest.hawaii.edu/hot/methods/inuts.html`, Last Accessed 1 June 2017.

Hawkins, S. J., Vale, M., Firth, L. B., Burrows, M. T., Mieszkowska, N. and Frost, M. (2013), 'Sustained Observation of Marine Biodiversity and Ecosystems', *Oceanography: Open Access* **01**(01), 1–4.

Heinrich, A. K. (1962), 'The Life Histories of Plankton Animals and Seasonal Cycles of Plankton Communities in the Oceans', *ICES Journal of Marine Science* **27**(1), 15–24.

Heisler, J., Glibert, P. M., Burkholder, J. M., Anderson, D. M., Cochlan, W., Dennison, W. C., Dortch, Q., Gobler, C. J., Heil, C. A., Humphries, E., Lewitus, A., Magnien, R., Marshall, H. G., Sellner, K., Stockwell, D. A., Stoecker, D. K. and Suddleson, M. (2008), 'Eutrophication and harmful algal blooms: A scientific consensus', *Harmful Algae* **8**(1), 3–13.

Hemmings, J. C. P. and Challenor, P. G. (2012), 'Addressing the impact of environmental uncertainty in plankton model calibration with a dedicated software system: The marine model optimization testbed (MarMOT 1.1 alpha)', *Geoscientific Model Development* **5**(2), 471–498.

Hemmings, J. C. P., Challenor, P. G. and Yool, A. (2015), 'Mechanistic site-based emulation of a global ocean biogeochemical model (MEDUSA 1.0) for parametric analysis and calibration: An application of the Marine Model Optimization Testbed (MarMOT 1.1)', *Geoscientific Model Development* **8**(3), 697–731.

Hemsley, V. S., Smyth, T. J., Martin, A. P., Frajka-Williams, E., Thompson, A. F., Damerell, G. and Painter, S. C. (2015), 'Estimating Oceanic Primary Production Using Vertical Irradiance and Chlorophyll Profiles from Ocean Gliders in the North Atlantic', *Environmental Science and Technology* **49**(19), 11612–11621.

Henson, S. A., Robinson, I., Allen, J. T. and Waniek, J. J. (2006), 'Effect of meteorological conditions on interannual variability in timing and magnitude of the spring bloom in the Irminger Basin, North Atlantic', *Deep-Sea Research Part I: Oceanographic Research Papers* **53**(10), 1601–1615.

Henson, S. A., Yool, A., Cole, H. S., Hopkins, J. and Martin, A. P. (2017), 'Detection of climate change-driven trends in phytoplankton phenology', *Global Change Biology* **24**, 1–11.

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J., Aiken, J., Barlow, R., Suzuki, K., Isada, T., Howell, E., Hashioka, T., Noguchi-Aita, M. and Yamanaka, Y. (2011), 'Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types', *Biogeosciences* **8**(2), 311–327.

Holling, C. S. (1959), 'The components of predation revealed by a study of small mammal predation on the European pine sawfly', *Canadian Entomologist* **91**, 293–320.

Holt, J., Icarus Allen, J., Anderson, T. R., Brewin, R., Butenschön, M., Harle, J., Huse, G., Lehodey, P., Lindemann, C., Memery, L., Salihoglu, B., Senina, I. and Yool, A. (2014), 'Challenges in integrative approaches to modelling the marine ecosystems of the North Atlantic: Physics to fish and coasts to ocean', *Progress in Oceanography* **129**(PB), 285–313.

Huisman, J., Pham Thi, N. N., Karl, D. M. and Sommeijer, B. (2006), 'Reduced mixing generates oscillations and chaos in the oceanic deep chlorophyll maximum.', *Nature* **439**(7074), 322–325.

Hyder, K., Rossberg, A. G., Allen, J. I., Austen, M. C., Barciela, R. M., Bannister, H. J., Blackwell, P. G., Blanchard, J. L., Burrows, M. T., Defriez, E., Dorrington, T., Edwards, K. P., Garcia-Carreras, B., Heath, M. R., Hembury, D. J., Heymans, J. J., Holt, J., Houle, J. E., Jennings, S., Mackinson, S., Malcolm, S. J., McPike, R., Mee, L., Mills, D. K., Montgomery, C., Pearson, D., Pinnegar, J. K., Pollicino, M., Popova, E. E., Rae, L., Rogers, S. I., Speirs, D., Spence, M. A., Thorpe, R., Turner, R. K., van der Molen, J., Yool, A. and Paterson, D. M. (2015), 'Making modelling count - increasing

the contribution of shelf-seas community and ecosystem models to policy development and management', *Marine Policy* **61**, 291–302.

Ivlev, V. (1961), *Experimental Ecology of The Feeding of The Fishes*, Yale University Press, New Haven, CT.

Jassby, A. D. and Platt, T. (1976), 'Mathematical Formulation of the Relationship Between Photosynthesis and Light for Phytoplankton', *Limnology and Oceanography* **21**(4), 540–547.

Ji, R., Edwards, M., MacKas, D. L., Runge, J. A. and Thomas, A. C. (2010), 'Marine plankton phenology and life history in a changing climate: Current research and future directions', *Journal of Plankton Research* **32**(10), 1355–1368.

Kahru, M., Gille, S. T., Murtugudde, R., Strutton, P. G., Manzano-Sarabia, M., Wang, H. and Mitchell, B. G. (2010), 'Global correlations between winds and ocean chlorophyll', *Journal of Geophysical Research: Oceans* **115**(12), 1–11.

Kane, A., Moulin, C., Thiria, S., Bopp, L., Berrada, M., Tagliabue, A., Crépon, M., Aumont, O. and Badran, F. (2011), 'Improving the parameters of a global ocean biogeochemical model via variational assimilation of in situ data at five time series stations', *Journal of Geophysical Research: Oceans* **116**(6), 1–14.

Karl, D. M. and Lukas, R. (1996), 'The Hawaii Ocean Time-series (HOT) program: Background, rationale and field implementation', *Deep Sea Research Part II: Topical Studies in Oceanography* **43**(2-3), 129–156.

Körtzinger, A., Send, U., Lampitt, R. S., Hartman, S., Wallace, D. W., Karstensen, J., Villagarcia, M. G., Llinás, O. and DeGrandpre, M. D. (2008), 'The seasonal pCO2 cycle at 49N/16.5W in the northeastern Atlantic Ocean and what it tells us about biological productivity', *Journal of Geophysical Research: Oceans* **113**(4), 1–15.

Kostadinov, T. S., Cabré, A., Vedantham, H., Marinov, I., Bracher, A., Brewin, R. J., Bricaud, A., Hirata, T., Hirawake, T., Hardman-Mountford, N. J., Mouw, C., Roy, S. and Uitz, J. (2017), 'Inter-comparison of phytoplankton functional type phenology metrics derived from ocean color algorithms and Earth System Models', *Remote Sensing of Environment* **190**, 162–177.

Kriest, I., Khatiwala, S. and Oschlies, A. (2010), 'Towards an assessment of simple global marine biogeochemical models of different complexity', *Progress in Oceanography* **86**(3-4), 337–360.

Kriest, I., Oschlies, A. and Khatiwala, S. (2012), 'Sensitivity analysis of simple global marine biogeochemical models', *Global Biogeochemical Cycles* **26**(2), 1–15.

Kristiansen, T., Drinkwater, K. F., Lough, R. G. and Sundby, S. (2011), 'Recruitment variability in North Atlantic cod and match-mismatch dynamics', *PLoS ONE* **6**(3).

Kwiatkowski, L., Yool, A., Allen, J. I., Anderson, T. R., Barciela, R., Buitenhuis, E. T., Butenschön, M., Enright, C., Halloran, P. R., Le Quere, C., De Mora, L., Racault, M. F., Sinha, B., Totterdell, I. J. and Cox, P. M. (2014), 'IMarNet: An ocean biogeochemistry model intercomparison project within a common physical ocean modelling framework', *Biogeosciences* **11**(24), 7291–7304.

Lacroix, G. and Gregoire, M. (2002), 'Revisited ecosystem model (MODECOGeL) of the Ligurian Sea: seasonal and interannual variability due to atmospheric forcing', *Journal of Marine Systems* **37**(4), 229–258.

Laffoley, D. and Baxter, J. (2016), Explaining Ocean Warming: Causes, scale, effects and consequences, Technical Report September, IUCN, Gland, Switzerland.

Laws, E. A. (2013), 'Evaluation of In Situ Phytoplankton Growth Rates: A Synthesis of Data from Varied Approaches', *Annual Review of Marine Science* **5**(1), 247–268.

Le Quere, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Leticia Cotrim Da Cunha, R. G., Giraud, X., Klaas, C., Kohfeld, K. E., Legendre, L., Manizza, M., Platt, T., Rivkin, R. B., Sathyendranath, S., Uitz, J., Watson, A. J. and Wolf-Gladrow, D. (2005), 'Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models', *Global Change in Biology* **11**(11), 2016–2040.

Letelier, R. M., Bidigare, R. R., Hebel, D. V., Ondrusek, M., Winn, C. D. and Karl, D. M. (1993), 'Temporal variability of phytoplankton community structure based on pigment analysis', **38**(7), 1420–1437.

Letelier, R. M., Karl, D. M., Abbott, M. R. and Bidigare, R. R. (2004), 'Light driven seasonal patterns of chlorophyll and nitrate in the lower euphotic zone of the North Pacific Subtropical Gyre', *Limnol. Oceanogr* **49**(2), 508–519.

Letelier, R. M., Karl, D. M., Abbott, M. R., Flament, P., Freilich, M., Lukas, R. and Strub, T. (2004), 'Role of late winter mesoscale events in the biogeochemical variability of the upper water column of the North Pacific Subtropical Gyre', *Journal of Geophysical Research: Oceans* **105**(C12), 28723–28739.

Levin, S. and Lubchenco, J. (2008), 'Resilience, robustness, and marine ecosystem-based management', *BioScience* **58**(1), 27–32.

Lévy, M., Klein, P. and Treguier, A.-M. (2001), 'Impact of sub-mesoscale physics on production and subduction of phytoplankton in an oligotrophic regime', *Journal of Marine Research* **59**(4), 535–565.

Lipschultz, F. (2001), 'A time-series assessment of the nitrogen cycle at BATS', **48**, 1897–1924.

Lomas, M. W., Steinberg, D. K., Dickey, T., Carlson, C. A., Nelson, N. B., Condon, R. H. and Bates, N. R. (2010), 'Increased ocean carbon export in the Sargasso Sea linked to climate variability is countered by its enhanced mesopelagic attenuation', *Biogeosciences* **7**(1), 57–70.

Longhurst, A. R. (2007), *Ecological geography of the sea*, 2nd edn, Academic Press.

Longhurst, A., Sathyendranath, S., Platt, T. and Caverhill, C. (1995), 'An estimate of global primary production in the ocean from satellite radiometer data', *Journal of Plankton Research* **17**(6), 1245–1271.

Lorenzoni, L., Taylor, G. T., Benitez-Nelson, C., Hansell, D. A., Montes, E., Masserini, R., Fanning, K., Varela, R., Astor, Y., Guzmán, L. and Muller-Karger, F. E. (2013), 'Spatial and seasonal variability of dissolved organic matter in the Cariaco Basin', *Journal of Geophysical Research: Biogeosciences* **118**(2), 951–962.

Mackas, D., Greve, W., Edwards, M., Chiba, S., Tadokoro, K., Eloire, D., Mazzocchi, M., Batten, S., Richardson, A., Johnson, C., Head, E., Conversi, A. and Peluso, T. (2012), 'Changing zooplankton seasonality in a changing ocean: Comparing time series of zooplankton phenology', *Progress in Oceanography* **97-100**, 31–62.

Mahowald, N. M., Baker, A. R., Bergametti, G., Brooks, N., Duce, R. A., Jickells, T. D., Kubilay, N., Prospero, J. M. and Tegen, I. (2005), 'Atmospheric global dust cycle and iron inputs to the ocean', *Global Biogeochemical Cycles* **19**(4).

Mahowald, N. M., Engelstaedter, S., Luo, C., Sealy, A., Artaxo, P., Benitez-Nelson, C., Bonnet, S., Chen, Y., Chuang, P. Y., Cohen, D. D., Dulac, F., Herut, B., Johansen, A. M., Kubilay, N., Losno, R., Maenhaut, W., Paytan, A., Prospero, J. M., Shank, L. M. and Siefert, R. L. (2009), 'Atmospheric iron deposition: global distribution, variability, and human perturbations.', *Annual review of marine science* pp. 245–278.

Mann, D. G. (1999), 'The species concept in diatoms', *Phycologia* **38**(December), 437–495.

Martin, J. H. and Fitzwarer, S. E. (1988), 'Iron deficiency limits phytoplankton growth in the north-east Pacific subarctic', *Nature* **331**, 341–343.

Martin-Jezequel, V., Hildebrand, M. and Brzezinski, M. A. (2000), 'SILICON METABOLISM IN DIATOMS: IMPLICATIONS FOR GROWTH', *Journal of Phycology* **36**(5), 821–840.

Martinez, E., Antoine, D., D'Ortenzio, F. and Gentili, B. (2009), 'Climate-driven basin-scale decadal oscillations of oceanic phytoplankton', *Science* **326**(5957), 1253–1256.

McDonald, C., Bennington, V., Urban, N. and McKinley, G. (2012), '1-D test-bed calibration of a 3-D Lake Superior biogeochemical model', *Ecological Modelling* **225**, 115–126.

McGillicuddy, D. J., Anderson, L. A., Doney, S. C. and Maltrud, M. E. (2003), 'Eddy-driven sources and sinks of nutrients in the upper ocean: Results from a 0.1 resolution model of the North Atlantic', *Global Biogeochemical Cycles* **17**(2), n/a–n/a.

Mignot, A., Claustre, H., Uitz, J., Poteau, A., Ortenzio, F. D. and Xing, X. (2014), 'Understanding the seasonal dynamics and the deep chlorophyll maximum in oligotrophic', *AGU. global biogeochemical cycles* (1), 856–876.

Mitra, A. (2009), 'Are closure terms appropriate or necessary descriptors of zooplankton loss in nutrient-phytoplankton-zooplankton type models?', *Ecological Modelling* **220**(5), 611–620.

Molinero, J. C., Reygondeau, G. and Bonnet, D. (2013), 'Climate variance influence on the non-stationary plankton dynamics', *Marine Environmental Research* **89**, 91–96.

Mongin, M., Nelson, D. M., Pondaven, P. and Treguer, P. (2006), 'Simulation of upper-ocean biogeochemistry with a flexible-composition phytoplankton model: C, N and Si cycling and Fe limitation in the Southern Ocean', *Deep-Sea Research Part I* **53**(5-7), 601–619.

Moradkhani, H., Hsu, K., Hong, Y. and Sorooshian, S. (2006), 'Investigating the impact of remotely sensed precipitation and hydrologic model uncertainties on the ensemble streamflow forecasting', *Geophysical Research Letters* **33**(12), 1–5.

Moradkhani, H. and Meskele, T. T. (2010), Probabilistic Assessment of the Satellite Retrieval Error Translation to Hydrologic Response, *in* M. Gebremichael and F. Hossain, eds, 'Satellite Rainfall Applications for Surface Hydrology', Springer, Netherlands, chapter II, pp. 235–241.

Muller-Karger, F., Varela, R., Thunell, R., Scranton, M., Bohrer, R., Taylor, G., Capelo, J., Astor, Y., Tappa, E., Ho, T.-Y. and john Walsh (2001), 'Annual cycle of primary production in the Cariaco Basin: Response to upwelling and implications for vertical export', *Journal of Geophysical Research* **106**(15), 4527–4542.

Murphy, J. M., Booth, B. B. B., Collins, M., Harris, G. R., Sexton, D. M. H. and Webb, M. J. (2007), 'A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles.', *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* **365**(1857), 1993–2028.

Murray, A. G. and Parslow, J. S. (1999), 'The analysis of alternative formulations in a simple model of a coastal ecosystem', *Ecological Modelling* **119**, 149–166.

Najjar, R. G., Jin, X., Louanchi, F., Aumont, O., Caldeira, K., Doney, S. C., Dutay, J. C., Follows, M., Gruber, N., Joos, F., Lindsay, K., Maier-Reimer, E., Matear, R. J., Matsumoto, K., Monfray, P., Mouchet, A., Orr, J. C., Plattner, G. K., Sarmiento, J. L., Schlitzer, R., Slater, R. D., Weirig, M. F., Yamanaka, Y. and Yool, A. (2007), 'Impact of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean: Results from Phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2)', *Global Biogeochemical Cycles* **21**(3).

Oschlies, A. and Garçon, V. (1999), 'An eddy-permitting coupled physical-biological model of the North Atlantic. 1. Sensitivity to advection numerics and mixed layer physics', *Global Biogeochemical Cycles* **13**(1), 135–160.

Oschlies, A. and Schartau, M. (2005), 'Basin-scale performance of a locally optimized marine ecosystem model', *Journal of Marine Systems* **63**(2), 335–358.

Ourmières, Y., Brasseur, P., Lévy, M., Brankart, J.-m. and Verron, J. (2009), 'On the key role of nutrient data to constrain a coupled physical biogeochemical assimilative model of the North Atlantic Ocean', *Journal of Marine Systems* **75**(1-2), 100–115.

Palmer, J. and Totterdell, I. (2001), 'Production and export in a global ecosystem model', *Deep-Sea Research I* **48**, 1169–1198.

Palter, J. B., Lozier, M. S. and Barber, R. T. (2005), 'The effect of advection on the nutrient reservoir in the North Atlantic subtropical gyre', *Nature* **437**(7059), 687–692.

Parekh, P., Follows, M. J. and Boyle, E. A. (2005), 'Decoupling of iron and phosphate in the global ocean', *Global Biogeochemical Cycles* **19**(2), 1–16.

Parsons, M. L. and Dortch, Q. (2002), 'Sedimentological evidence of an increase in Pseudo-nitzschia (Bacillariophyceae) abundance in response to coastal eutrophication', *Limnology and Oceanography* **47**(2), 551–558.

Pinckney, J. L., Benitez-Nelson, C. R., Thunell, R. C., Muller-Karger, F., Lorenzoni, L., Troccoli, L. and 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.

Pingree, R. D., Maddock, L. and Butler, E. I. (1977), 'The Influence of Biological Activity and Physical

Stability in Determining the Chemical Distributions of Inorganic Phosphate, Silicate, and Nitrate', (57), 1065–1073.

Platt, T., Fuentes-Yaco, C. and Frank, K. T. (2003), 'Spring algal bloom and larval fish survival', *Nature* **423**(May), 398–399.

Platt, T. and Sathyendranath, S. (2008), 'Ecological indicators for the pelagic zone of the ocean from remote sensing', *Remote Sensing of Environment* **112**(8), 3426–3436.

Platt, T., White, G. N., Zhai, L., Sathyendranath, S. and Roy, S. (2009), 'The phenology of phytoplankton blooms: Ecosystem indicators from remote sensing', *Ecological Modelling* **220**(21), 3057–3069.

Popova, E. E., Coward, A. C., Nurser, G. A., de Cuevas, B., Fasham, M. J. R. and Anderson, T. R. (2006), 'Mechanisms controlling primary and new production in a global ecosystem model - Part I: Validation of the biological simulation', *Ocean Science* **2**(2), 249–266.

Prowe, A. E., Pahlow, M., Dutkiewicz, S., Follows, M. and Oschlies, A. (2012), 'Top-down control of marine phytoplankton diversity in a global ecosystem model', *Progress in Oceanography* **101**(1), 1–13.

Racault, M.-F., Le Quere, C., Buitenhuis, E., Sathyendranath, S. and Platt, T. (2012), 'Phytoplankton phenology in the global ocean', *Ecological Indicators* **14**(1), 152–163.

Raghukumar, K., Edwards, C. A., Goebel, N. L., Broquet, G., Veneziani, M., Moore, A. M. and Zehr, J. P. (2015), 'Progress in Oceanography Impact of assimilating physical oceanographic data on modeled ecosystem dynamics in the California Current System', *Progress in Oceanography* **138**, 546–558.

Raick, C., Soetaert, K. and Grégoire, M. (2006), 'Model complexity and performance: How far can we simplify?', *Progress in Oceanography* **70**(1), 27–57.

Raymont, J. (1980), *Plankton and Productivity in the Oceans*, Vol. 1, 2nd edn, Pergamon Press, London.

Redfield, A. (1934), On the proportions of organic derivatives in seawater and their relation to the composition of plankton, *in* R. Daniel, ed., 'James Johnstone Memorial Volume', University Press of Liverpool, pp. 176–192.

Reygondeau, G., Molinero, J. C., Coombs, S., MacKenzie, B. R. and Bonnet, D. (2015), 'Progressive changes in the Western English Channel foster a reorganization in the plankton food web', *Progress in Oceanography* **137**, 524–532.

Riley, G. A. (1946), 'Factors Controlling the Phytoplankton Populations on Georges Bank', *Journal of Marine Research* **6**, 33–47.

Riley, G. A., Stommel, H. A. and Bumpus, D. F. (1949), *Quantitative ecology of the plankton of the western North Atlantic.*, bulletin o edn, Bingham Oceanographic Laboratory, New Haven, CT.

Robinson, C. L. K., Ware, D. M. and Parsons, T. R. (1993), 'Simulated annual plankton production in the northeastern Pacific Coastal upwelling Domain', *Journal of Plankton Research* **15**(2), 161–183.

Robson, B. J., Arhonditsis, G. B., Baird, M. E., Brebion, J., Edwards, K. F., Geoffroy, L., Hébert, M. P., van Dongen-Vogels, V., Jones, E. M., Kruk, C., Mongin, M., Shimoda, Y., Skerratt, J. H., Trevathan-Tackett, S. M., Wild-Allen, K., Kong, X. and Steven, A. (2018), 'Towards evidence-based parameter values and priors for aquatic ecosystem modelling', *Environmental Modelling and Software* **100**, 74–81.

Roy, S., Broomhead, D. S., Platt, T., Sathyendranath, S. and Ciavatta, S. (2012), 'Sequential variations of phytoplankton growth and mortality in an NPZ model: A remote-sensing-based assessment', *Journal of Marine Systems* **92**(1), 16–29.

Roy, S. and Chattopadhyay, J. (2007), 'Enrichment and stability: A phenomenological coupling of energy value and carrying capacity', *BioSystems* .

Roy, S., Platt, T. and Sathyendranath, S. (2011), 'Modelling the time-evolution of phytoplankton size spectra from satellite remote sensing', *ICES Journal of Marine Science* **68**(4), 719–728.

Ruan, S. G. (2001), 'Oscillations in plankton models with nutrient recycling', *Journal of Theoretical Biology* **208**(1), 15–26.

Ryabchenko, V. A., Fasham, M. J. R., Kagan, B. A. and Popova, E. E. (1997), 'What causes short-term oscillations in ecosystem models of the ocean mixed layer?', *Journal of Marine Systems* **13**(1-4), 33–50.

Ryan, J. P., Fischer, A. M., Kudela, R. M., Gower, J. F., King, S. A., Marin, R. and Chavez, F. P. (2009), 'Influences of upwelling and downwelling winds on red tide bloom dynamics in Monterey Bay, California', *Continental Shelf Research* **29**(5-6), 785–795.

Saba, V. S., Friedrichs, M. A., Carr, M. E., Antoine, D., Armstrong, R. A., Asanuma, I., Aumont, O., Bates, N. R., Behrenfeld, M. J., Bennington, V., Bopp, L., Bruggeman, J., Buitenhuis, E. T., Church, M. J., Ciotti, A. M., Doney, S. C., Dowell, M., Dunne, J., Dutkiewicz, S., Gregg, W., Hoepffner, N., Hyde, K. J., Ishizaka, J., Kameda, T., Karl, D. M., Lima, I., Lomas, M. W., Marra, J., McKinley, G. A., Melin, F., Moore, J. K., Morel, A., O'Reilly, J., Salihoglu, B., Scardi, M., Smyth, T. J., Tang, S., Tjiputra, J., Uitz, J., Vichi, M., Waters, K., Westberry, T. K. and Yool, A. (2010), 'Challenges of modeling depth-integrated marine primary productivity over multiple decades: A case study at BATS and HOT', *Global Biogeochemical Cycles* **24**(3), 1–21.

Sailley, S. F., Polimene, L., Mitra, A., Atkinson, A. and Allen, J. I. (2014), 'Impact of zooplankton food selectivity on plankton dynamics and nutrient cycling', *Journal of Plankton Research* **37**(3), 519–529.

Sapiano, M. R. P., Brown, C. W., Schollaert Uz, S. and Vargas, M. (2012), 'Establishing a global climatology of marine phytoplankton phenological characteristics', *Journal of Geophysical Research: Oceans* **117**(8), 1–16.

Sarmiento, J. L., Slater, R., Barber, R., Bopp, L., Doney, S. C., Hirst, A. C., Kleypas, J., Matear, R., Mikolajewicz, U., Monfray, P., Soldatov, V., Spall, S. A. and Stouffer, R. (2004), 'Response of ocean ecosystems to climate warming', *Global Biogeochemical Cycles* **18**(3), n/a–n/a.

Sharples, J., Ross, O. N., Scott, B. E., Greenstreet, S. P. and Fraser, H. (2006), 'Inter-annual variability in the timing of stratification and the spring bloom in the North-western North Sea', *Continental Shelf Research* **26**(6), 733–751.

Siegel, D. A., Behrenfeld, M. J., Maritorena, S., McClain, C. R., Antoine, D., Bailey, S. W., Bontempi, P. S., Boss, E. S., Dierssen, H. M., Doney, S. C., Eplee, R. E., Evans, R. H., Feldman, G. C., Fields, E., Franz, B. A., Kuring, N. A., Mengelt, C., Nelson, N. B., Patt, F. S., Robinson, W. D., Sarmiento, J. L., Swan, C. M., Werdell, P. J., Westberry, T. K., Wilding, J. G. and Yoder, J. A. (2013), 'Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission', *Remote Sensing of Environment* **135**, 77–91.

Siegel, D. A., Doney, S. C. and Yoder, J. A. (2002), 'The North Atlantic spring phytoplankton bloom and Sverdrup's critical depth hypothesis.', *Science* **296**(5568), 730–3.

Sinha, B., Buitenhuis, E. T., Le Quere, C. and Anderson, T. R. (2010), 'Progress in Oceanography Comparison of the emergent behavior of a complex ecosystem model in two ocean general circulation models', *Progress in Oceanography* **84**(3-4), 204–224.

Smith, H. (1936), 'A DISCRIMINANT FUNCTION FOR PLANT SELECTION', *Annals of Eugenics* **7**(3), 240–250.

Smith, S. L., Pahlow, M., Merico, A., Acevedo-Trejos, E., Sasai, Y., Yoshikawa, C., Sasaoka, K., Fujiki, T., Matsumoto, K. and Honda, M. C. (2016a), 'Flexible phytoplankton functional type (FlexPFT) model: Size-scaling of traits and optimal growth', *Journal of Plankton Research* **38**(4), 977–992.

Smith, S. L., Vallina, S. M. and Merico, A. (2016), 'Phytoplankton size-diversity mediates an emergent trade-off in ecosystem functioning for rare versus frequent disturbances', *Scientific Reports* **6**(October), 1–11.

Smith, S., Yamanaka, Y., Pahlow, M. and Oschlies, A. (2009), 'Optimal uptake kinetics: physiological acclimation explains the pattern of nitrate uptake by phytoplankton in the ocean', *Marine Ecology Progress Series* **384**, 1–12.

Smyth, T. J., Atkinson, A., Widdicombe, S., Allen, J. I., Queiros, A., Sims, D. and Brange, M. (2015), 'The Western Channel Observatory', *Progress in Oceanography* **137**, 335–341.

Smyth, T. J., Fishwick, J. R., Al-moosawi, L., Cummings, D. G. and Harris, C. (2010), 'A broad spatio-temporal view of the Western English Channel observatory', **32**(September), 585–601.

Spence, M. A., Blackwell, P. G., Blanchard, J. L. and Shuter, B. (2016), 'Parameter uncertainty of a dynamic multispecies size spectrum model', *Canadian Journal of Fisheries and Aquatic Sciences* **73**(4), 589–597.

Stammer, D., Balmaseda, M., Heimbach, P., Köhl, A. and Weaver, A. (2015), 'Ocean Data Assimilation in Support of Climate Applications: Status and Perspectives', *Annual Review of Marine Science* **8**(1), 491–518.

Steele, J. H. and Henderson, E. W. (1981), 'A Simple Plankton Model', *The American Naturalist* **117**(5), 676–691.

Steele, J. H. and Henderson, E. W. (1992), 'The role of predation in planckton models', *Journal of Plankton Research* **14**(1), 157–172.

Steinacher, M., Joos, F., Frölicher, T. L., Bopp, L., Cadule, P., Doney, S. C., Gehlen, M., Schneider, B. and Segschneider, J. (2009), 'Projected 21st century decrease in marine productivity: a multi-model analysis', *Biogeosciences Discussions* **6**, 7933–7981.

Steinberg, D. K., Carlson, C. A., Bates, N. R., Johnson, R. J., Michaels, A. F. and Knap, A. H. (2001), 'Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry', **48**, 1405–1447.

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., Hare, J. A., Hollowed, A. B., Lehodey, P., Levin, S. A., Link, J. S., Rose, K. A., Rykaczewski, R. R., Sarmiento, J. L., Stouffer, R. J., Schwing, F. B., Vecchi, G. A. and Werner, F. E. (2011), 'On the use of IPCC-class models to assess the impact of climate on Living Marine Resources', *Progress in Oceanography* **88**(1-4), 1–27.

Storkey, D., Blockley, E., Furner, R., Guiavarc'h, C., Lea, D., Martin, M., Barciela, R. M., Hines, A., Hyder, P. and Siddorn, J. (2010), 'Forecasting the ocean state using NEMO:The new FOAM system'.

Subramanian, A. C. and Palmer, T. N. (2017), 'Ensemble superparameterization versus stochastic parameterization: A comparison of model uncertainty representation in tropical weather prediction', *Journal of Advances in Modeling Earth Systems* **9**(2), 1231–1250.

Sverdrup, H. U. (1953), 'On conditions for the vernal blooming of phytoplankton', *ICES Journal of Marine Science* **18**(3), 287–295.

Taylor, G. T., Muller-karger, F. E., Thunell, R. C., Scranton, M. I., Astor, Y. and Varela, R. (2012), 'Ecosystem responses in the southern Caribbean Sea to global climate change', **109**(47), 19315–19320.

Tebaldi, C. and Knutti, R. (2007), 'The use of the multi-model ensemble in probabilistic climate projections', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **365**(1857), 2053–2075.

Thomalla, S. J., Fauchereau, N., Swart, S. and Monteiro, P. M. S. (2011), 'Regional scale characteristics of the seasonal cycle of chlorophyll in the Southern Ocean', *Biogeosciences* **8**(10), 2849–2866.

Tinker, J., Lowe, J., Holt, J., Pardaens, A. and Wiltshire, A. (2015), 'Validation of an ensemble modelling system for climate projections for the northwest European shelf seas', *Progress in Oceanography* **138**, 211–237.

Tinker, J., Lowe, J., Pardaens, A., Holt, J. and Barciela, R. (2016), 'Uncertainty in climate projections for the 21st century northwest European shelf seas', *Progress in Oceanography* **148**, 56–73.

Vargas, M., Brown, C. W. and Sapiano, M. R. P. (2009), 'Phenology of marine phytoplankton from satellite ocean color measurements', *Geophysical Research Letters* **36**(1), 2–6.

Villareal, T. A., Brown, C. G., Brzezinski, M. A., Krause, J. W. and Wilson, C. (2012), 'Summer Diatom Blooms in the North Pacific Subtropical Gyre : 2008 2009', **7**(4).

Walsh, J., Dieterie, D. A., Mfiller-karger, F. E., Bohrer, R., Bissett, W. P., Varela, R. J., Aparicio, R., Diaz, R., Thunell, R., Taylor, G. T., Scranton, M. I., Fanning, K. A. and Peltzer, E. T. (1999), 'Simulation of carbon-nitrogen cycling during spring upwelling in the Cariaco Basin', *Journal of Geophysical Research* **104**(1998), 7807–7825.

Ward, B. A., Friedrichs, M. A. M., Anderson, T. R. and Oschlies, A. (2010), 'Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models', *Journal of Marine Systems* **81**(1-2), 34–43.

Ward, B. A., Schartau, M., Oschlies, A., Martin, A. P., Follows, M. J. and Anderson, T. R. (2013), 'When is a biogeochemical model too complex? Objective model reduction and selection for North Atlantic time-series sites', *Progress in Oceanography* **116**, 49–65.

Watson, A. J., Boyd, P. W., Turner, S. M., Jickells, T. D. and Liss, P. S. (2008), 'Designing the next generation of ocean iron fertilization experiments', *Marine Ecology Progress Series* **364**, 303–309.

White, A., Letelier, R. M., Whitmire, A. L., Barone, B., Bidigare, R. R., Church, M. J. and Karl, D. M. (2015), 'Phenology of particle size distributions and primary productivity in the North Pacific subtropical gyre (Station ALOHA)', *Journal of Geophysical Research Oceans* pp. 7381–7399.

White, M. A., de Beurs, K. M., Didan, K., Inouye, D. W., Richardson, A. D., Jensen, O. P., O'Keefe, J., Zhang, G., Nemani, R. R., van Leeuwen, W. J. D., Brown, J. F., de Wit, A., Schaepman, M., Lin, X., Dettinger, M., Bailey, A. S., Kimball, J., Schwartz, M. D., Baldocchi, D. D., Lee, J. T. and Lauenroth, W. K. (2009), 'Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982-2006', *Global Change Biology* **15**(10), 2335–2359.

Widdicombe, C. E., Eloire, D., Harbour, D., Harris, R. P. and Somerfield, P. J. (2010), 'Long-term phytoplankton community dynamics in the Western English Channel', *Journal of Plankton Research* **32**(5), 643–655.

Wood, S. and Thomas, M. (1999), 'Super-sensitivity to structure in biological models', *The Royal Society* **266**(1419), 565–570.

Yool, A., Popova, E. E. and Anderson, T. R. (2011), 'MEDUSA-1.0: A new intermediate complexity plankton ecosystem model for the global domain', *Geoscientific Model Development* **4**(2), 381–417.

Yool, A., Popova, E. E. and Anderson, T. R. (2013), 'MEDUSA-2.0: An intermediate complexity biogeochemical model of the marine carbon cycle for climate change and ocean acidification studies', *Geoscientific Model Development* **6**(5), 1767–1811.

Yool, A. and Tyrrell, T. (2003), 'The role of diatoms in regulating the ocean's silicate cycle', *Global Biogeochemical Cycles* **17**(4), 1–24.

Zhai, L., Platt, T., Tang, C., Sathyendranath, S. and Hernandez Walls, R. (2011), 'Phytoplankton phenology on the Scotian Shelf', *ICES Journal of Marine Science* **68**(4), 781–791.