Towards a mechanistic understanding of the ocean's biological carbon pump

Anna Rufas Blanco
Department of Earth Sciences and Wolfson College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Michaelmas 2021

To my grandma,
and to my parents.

Acknowledgements

First of all, I thank my supervisors Samar Khatiwala and Adrian Martin, for envisioning this project and entrusting it to me. Without their brilliant analytical thinking and support of my ideas, I would have lost focus without realising. In particular, I thank Samar for making available all sorts of academic and personal resources I have needed throughout the past five years and for always trying to squeeze the best out of me.

My thesis project was part of a larger NERC large grant, COMICS, and I thank the team at the National Oceanography Centre Southampton for many useful discussions and advice and for helping me navigate the world of oceanographic flux data. I also thank Wolfson College and the MPLS division, which have made my project financially possible by securing additional funding to support me during the final year of my thesis.

Tinna Jokulsdottir, this genius project would not exist had not you invested all the sweat and blood you did before I took it over. I thank you for your guidance and for welcoming me to Chicago with your lovely family, my first visit ever to the U.S. and which I will never forget. I am very grateful for having been assigned the most supportive office in Earth Sciences, our international 30.16, my home when I was missing my homeland, my family when I was missing my kind. Tin Tin Naing, Annie Cheng, Afsaneh Mohammadzaheri, Naomi Saunders, Mohammed Al Fahmi, Tyler Ambrose (by order of office farewell). I also extend my thanks to Sophy Oliver for her friendship at first sight –as soon as you opened the door of my office for the first time I knew we were going to bond over–, and Brook Keats and Harriet Godwin for their many pieces of advice on how to stay put.

I wish also to thank the following friends, whose good advice and friendship I treasure: Caroline Mengeot, Geraldine Adiku, Giorgios Margaritis, Moisès Sánchez-Fortun, Lidia Carracedo, Adrián Martínez, Vlad Macovei, Sofia Darmaraki, Sinhué Torres-Valdés, José Alfredo Guerrero, Adina Pusok.

Most of all, I wish to thank my mum and dad for being my biggest cheerleaders and for being the last ones in losing faith when all seems lost.

Abstract

Marine particles lock atmospheric CO_2 into organic structures that sink and export carbon into the ocean interior, a process known as the biological carbon pump (BCP). Part of this flux of particulate organic carbon (POC) is attenuated in the mesopelagic ocean as respiration, disaggregation and solubilisation wear out the sinking particles and return the organic carbon into its inorganic form. Sediment traps and imaging systems have revealed the importance that particle material composition and microstructure have in the degradation of mesopelagic POC flux. However, the data collected are scarce and have too many spatiotemporal biases to provide a mechanistic picture of which particle attributes, particle dynamics and biogeochemical factors control it. Here, I present a stochastic model of Lagrangian marine particles that mechanistically interact with each other and their biogeochemical environment as they sink through the water column. A Lagrangian framework tracks particles throughout their life history and allows recording their fractal radius, porosity, density, stickiness and sinking velocity, providing information of particle attributes alongside POC fluxes. The model is applied globally to quantify spatial variations in the vertical attenuation rate of POC flux (Martin's b) and its dependence on proposed controlling factors such as NPP, phytoplankton community composition and seawater temperature. I find that Martin's b is lowest in productive, diatom-rich ecosystems and highest in oligotrophic, picophytoplankton-rich ecosystems. Particle density and ballast provided by the opal sourced by diatoms are the particle attributes that contribute the most to the transfer of POC flux to the mesopelagic, and particle dynamics of coagulation and zooplankton egestion assist in spreading the opal amongst the broader pool of particles. These results emphasise the role of both surface water conditions (NPP and phytoplankton community composition) and mesopelagic conditions (particle transformation dynamics) in controlling the BCP-mediated transfer of carbon to the ocean interior.

Extended abstract

Marine particles are invisible to the naked eye, nevertheless are central to the marine carbon cycle. Carbon enters the oceanic system as atmospheric carbon dioxide (CO_2), which diffuses into seawater and is fixed by phytoplankton cells, forming organic carbon. Phytoplankton sustain a food web that produces a myriad of biogenic marine particles that sink into the ocean interior, forming a rain of marine snow. Marine particles are the vectors that connect the productive surface ocean with the darker ocean interior and the seafloor sediments, carrying particulate organic carbon (POC) that sustains the metabolism of marine life. The transport of large amounts of carbon from the atmosphere into the ocean interior as marine particles is a process known as the biological carbon pump (BCP) and contributes to the oceanic storage of atmospheric carbon. After marine particles are produced, respiration, disaggregation and solubilisation take place, which return POC into its inorganic form, i.e., CO_2, attenuating the flux of POC along depth. The rate at which the flux of POC is degraded in the mesopelagic ocean (100–1000 m) has an impact on atmospheric CO_2 concentrations, making the ocean's BCP a research topic of particular value in the light of climate change mitigation policies and projections of future CO_2 storage in the ocean.

Global observations of POC flux started to be taken more than 30 years ago, when the U.S. launched the Joint Global Ocean Flux Study to investigate POC distribution, cycling and composition. Observations made since then by sediment traps, radionuclides, imaging systems and remote sensing indicate that the amount of POC sequestered in the deep ocean (< 1000 m) is controlled by an amalgam of environmental and ecosystem factors, including seawater temperature, oxygen (O_2) concentration, net primary production (NPP) and composition of the phytoplankton and zooplankton communities. However, oceanographers still do not understand the mechanisms by which those climate-sensitive factors control the global pattern of POC flux degradation in the mesopelagic ocean (100–1000 m). That is due to the intimate interplay between all of those factors, the economic costs and technical challenge of sampling the vast ocean to obtain more observations, mismatches in the values reported by different observational technologies, and the lack of a unifying protocol on the key depths and moment of the year to sample POC flux to ensure intercomparability between ocean locations.

With regard to modelling the BCP, the conventional practice has been designing models that take an Eulerian perspective at POC flux. POC is calculated as a continuum property, i.e., an average concentration of particulate material in a water volume, inherently assuming

that all particles in there are identical. Eulerian models are an excellent approach to simulate rates of change of dissolved tracers, like phosphate, nitrate or dissolved inorganic carbon, which lack a discrete nature. However, when it comes to particulate tracers, like POC, a Lagrangian approach becomes more natural and the only way to track particles individually. Tracking particles allows recording for each one a set of attributes, like size and sinking velocity, which evolve along the particle's trajectory. Thus, ensemble properties, like flux, emerge from the pool of different particles, each with distinct life histories. A simulation as such can provide insights into the controls at the particle level that affect mesopelagic degradation of POC flux..

This thesis presents a stochastic model of Lagrangian marine particles called SLAMS-2.0 (Stochastic, Lagrangian Aggregate Model of Sinking particles, v2.0). I apply it to understand the particle attributes, particle dynamics and biogeochemical controls that affect the decline in POC flux with depth, as characterised by the POC flux attenuation rate (the Martin's b) and the fraction of surface POC flux that reaches the deep ocean (the transfer efficiency, T_{eff}). SLAMS-2.0 follows up the efforts made by Jokulsdottir and Archer (2016) in developing a stochastic particle tracking model for the ocean's BCP. SLAMS-2.0 explicitly represents particle microstructure, composition, stickiness and sinking velocity according to fractal scaling and Stokes' law and tracks particles as they gravitationally sink through the water column and interact with their biogeochemical environment. Particles are created in the surface ocean through a probabilistic model of phytoplankton functional type (PFT) seeding, coagulate with one another using Jackson (2001)'s collision model, and interact with zooplankton following Kiørboe (1997)'s encounters model. Consequently, large aggregates can be formed and there is biotic reprocessing into faecal pellets. Further to that, the biogenic silica (bSi) and calcium carbonate ($CaCO_3$) content of particles provided by calcifying and silicifying phyto- and zooplankton is dissolved at the saturation horizon, and the POC is remineralised through heterotrophic metabolism following Q_{10} kinetics and O_2 limitation. Particles disaggregate into daughter particles following an empirical approach to particle breakup, and loss of solubilised material happens during zooplankton sloppy feeding and bacterial growth. The key insight is to reinterpret the deterministic equations that govern phytoplankton growth, coagulation, grazing, etc., as probabilities that govern whether a particle is likely to be created as one of four PFTs (diatom, coccolithophore, dinoflagellate or picophytoplankton), whether it will coagulate with another particle or not, or whether it will be grazed upon or not. The fluxes of POC and accompanying biogenic ballasting materials, bSi and $CaCO_3$, emerge from the stochastic dynamics and composition of the ensemble of particles, instead of being prescribed using a decaying function. Hence, a distinct feature of SLAMS-2.0 is that Martin's b and T_{eff} are extracted using a first-principles, mechanistic approach from a surface-to-depth profile of sinking Lagrangian particles.

Ocean biogeochemical modelling is a discipline with standardised procedures, and the steps commonly followed to develop a model have been used to divide this thesis into three parts and seven chapters.

Part I reviews the research topic of the ocean's BCP, from the context of why it matters to the metrics used to study it. In Chapter 1, I introduce the role of the BCP in the coevolution of the Earth's ocean and climate system. I present the different components that make up the BCP (sub-pumps and food web components) and stress the relationship between those and the environmental-sensitive factors that control the strength and efficiency of the downward transfer of POC flux, and which are changing in the light of anthropogenic climate change. I review two main competing theories about our current understanding of the factors that control the global pattern of Martin's b, Henson et al. (2012) vs Marsay et al. (2015), and present the lack of consensus in published work as to how to quantify the BCP's magnitude and efficiency. I conclude explaining how a Lagrangian and mechanistic model of marine particles becomes a suitable tool to map particle characteristics and dynamics to the factors that control oceanic POC flux. The first part of Chapter 2 contains the mathematical realisation of SLAMS-2.0, that is, the mathematical characteristics of Lagrangian and stochastic particle modelling. The second half of the chapter reviews common biogeochemical models that use particle-based descriptions. Those models offer an illustrative example of what makes SLAMS-2.0 a computationally efficient simulation to model billions of marine particles.

Part II is dedicated to methodology. In Chapter 3, I present the model equations and I navigate through the intellectual challenge of model development, describing the biogeochemical theory and the assumptions made to implement those into a Lagrangian and stochastic framework. SLAMS-2.0 is a modular model, where food web processes and biogeochemical forcing can be easily switched on and off to explore their effect on particle fluxes. As such, I have made an effort to find the appropriate level of complexity and mechanisation that will allow the simulation to have the most skill in predicting POC, $CaCO_3$ and bSi fluxes while keeping time and computational costs low. Chapter 4 contains the oceanographic data compilation that I will use to constrain the response of SLAMS-2.0 in the following chapter. The data are POC, PIC (particulate inorganic carbon, or $CaCO_3$) and bSi fluxes from five distinct ocean locations with long particle flux time-series: EqPac, HAUSGARTEN, OSP, PAP-SO and BATS/OFP. I also discuss the challenges of aggregating flux data in time and depth to produce quasi climatologies of particle fluxes. Chapter 5 addresses model calibration, for which I have used a derivative-free optimisation algorithm called DFO-LS, a method that finds local minima solutions and that has been shown to perform more efficiently than other optimisation methods for global biogeochemical models. Seven parameters are chosen for calibration and the skill of the optimiser is assessed in terms of model overfitting and reproduction of locally observed metrics of the BCP at the five time-series sites (Martin's b, T_{eff},

remineralisation length scale, particle export efficiency, POC/PIC ratio, bSi/PIC ratio and the slope of the particle size distribution).

Part III aims to answer the research questions outlined at the end of Chapter 1 by running SLAMS-2.0 on a global grid with the locally calibrated parameter set. Chapter 6 presents the simulated global distributions of various metrics of the BCP, as well as particle fluxes, particle attributes averaged by depth, PFTs relative abundance, coagulation intensity and zooplankton grazing intensity. I then compare the outputs of SLAMS-2.0 against observations and modelling results from data-constrained (Weber et al., 2016; DeVries and Weber, 2017) and empirical (Henson et al., 2012; Marsay et al., 2015) models. SLAMS-2.0 shows a global annual mean pattern of Martin's b and T_{eff} that is qualitatively comparable to published studies and is able to reconcile, or find a common ground, for Henson et al. (2012) and Marsay et al. (2015) schools of thought, which had shown opposite global patterns for Martin's b. Through a sensitivity analysis, I investigate three biogeochemical variables (NPP, diatom relative abundance and seawater temperature) that I suspect have the most effect on Martin's b and T_{eff}. Regression analyses of various particle attributes and the inspection of how global patterns of fluxes and ecosystem structure overlap, allows me to trace back the ocean's surface and mesopelagic mechanisms by which modelled POC is allocated into different particle size classes, velocity classes and compositions. This way, I manage to separate out the effects of the different particle attributes on the pattern of Martin's b. Lastly, a practical goal was set to extract from the Lagrangian ensemble of particles a linear relationship between Martin's b and its most relevant biogeochemical controls. Such a functional form can be easily plugged in Earth system models, overcoming the impracticalities of incorporating an expensive simulation of mechanistic Lagrangian particles whilst providing a parameterisation for Martin's b driven by local forcing, instead of the canonical prescribed form. Chapter 7 rounds off by summarising the key points of the methodological approach presented in this thesis and its main findings. It discusses model limitations, future work and provides suggestions for improving measurements and the representation of the particle processes needed to improve our mechanistic understanding of the ocean's BCP.

Contents

List of Figures

List of Tables

List of symbols, abbreviations and acronyms

1-D, 2-D, 3-D . One-, two- or three-dimensional, referring to spatial dimensions

AVHRR . . . . . Advanced Very High Resolution Radiometer

b . . . . . Slope of the Martin curve, also referred to as Martin's b, the rate at which POC flux decreases with depth

BATS/OFP . . Bermuda Atlantic Time-Series/Oceanic Flux Program

BCP . . . . . Biological carbon pump

bSi . . . . . Biogenic silica, also known as opal, an hydrated form of silica ($SiO_2 \cdot 0.4\,H_2O$)

C . . . . . Carbon, the element

$\mathbf{C_{org}}$ . . . . . Organic carbon, the form of carbon bounded to hydrogen atoms

$\mathbf{Ca^{2+}}$ . . . . . Calcium ion

$\mathbf{CaCO_3}$ . . . . . Calcium carbonate, and its forms calcite and aragonite

CAFE . . . . . Carbon, Absorption, and Fluorescence Euphotic-resolving model

CbPM . . . . . Carbon-based Productivity Model

CDF . . . . . Cumulative distribution function

chl *a* . . . . . Chlorophyll a

$\mathbf{CO_2}$ . . . . . Carbon dioxide

$\mathbf{CO_3^{2-}}$ . . . . . Carbonate ion

COMICS . . . . Controls over Ocean Mesopelagic Interior Carbon Storage

DFO-LS . . . . . Derivative-Free Optimisation using Least Squares, an optimisation algorithm

DIC . . . . . Dissolved inorganic carbon

DOC . . . . . Dissolved organic carbon (organic carbon in the < 0.2 µm size fraction)

DVM . . . . . Diel (diurnal) vertical migration

ESD . . . . . Equivalent spherical diameter

ESM . . . . . Earth system model

F . . . . . Flux of POC

F_{dust} . . . . . Flux of aeolian dust

$F(z)$ . . . . . Flux of POC at depth z

$F(z_0)$ . . . . . Flux of POC at the base of the euphotic zone, also known as export production

GLODAP . . . . Global Data Analysis Project

Gt . . . . . Gigaton (10^{15} g = 1 Pg)

H . . . . . Depth of the water column

$\mathbf{H^+}$ . . . . . Proton

$\mathbf{H_2CO_3}$ . . . . . Carbonic acid

$\mathbf{HCO_3^-}$ . . . . . Bicarbonate ion

HNLC . . . . . High-nutrient, low-chlorophyll

IPCC . . . . . Intergovernmental Panel on Climate Change

JGOFS . . . . . Joint Global Ocean Flux Study

LTER . . . . . Long-Term Ecological Research

MCM . . . . . Monte Carlo method

MLD . . . . . Mixed layer depth

MODIS . . . . . Moderate Resolution Imaging Spectroradiometer

$\mathbf{NO_3^-}$ . . . . . Nitrate

NPP . . . . . Net primary production, daily integrated, expressed in units of rate ($\mathrm{mg\,C\,m^{-2}\,d^{-1}}$)

NPZD . . . . . Nutrient-phytoplankton-zooplankton-detritus, referred to a type of ocean biogeochemical model

NSP . . . . . Net secondary production, equivalent to zooplankton biomass ($\mathrm{mol\,C}$)

$\mathbf{O_2}$ . . . . . Oxygen

OBM . . . . . Ocean biogeochemical modelling

OGCM . . . . . Ocean general circulation model

OM . . . . . Organic matter

OIND . . . . . Optimisation experiment where the fit to data is done individually on a location basis

OSIM . . . . . Optimisation experiment where fit to data is done simultaneously for my five studied ocean locations

OSP . . . . . Ocean Station Papa

P–I . . . . . Photosynthesis–Irradiance

PAP-SO . . . . . Porcupine Abyssal Plain time-Series Observatory

PAR . . . . . Photosynthetic active radiation (PAR_0 is PAR at the surface ocean)

$\boldsymbol{pCO_2}$ . . . . . Partial pressure of CO_2

PDE . . . . . Partial derivative equation

PDF . . . . . Probability density function

$\mathbf{PE_{eff}}$ . . . . . Particle export efficiency, a.k.a. export ratio (*e*-ratio) and particle export ratio (*pe*-ratio)

PFT . . . . . Phytoplankton functional type

Pg . . . . . Petagram (10^{15} g)

PIC . . . . . Particulate inorganic carbon

POC . . . . . Particulate organic carbon

ppm . . . . . Parts per million

PSD . . . . . Particle size distribution

r . . . . . Radius

SDM . . . . . Super droplet method

SeaWiFS . . . . Sea-viewing Wide Field-of-view Sensor

Si . . . . . Silicon, the element

$\mathbf{SiOH_4}$ . . . . . Silicic acid, or silicate in dissolution ($SiO_2 + 2\,H_2O$)

SLAMS-2.0 . . . Stochastic, Lagrangian Aggregate Model of Sinking particles, v. 2.0

SST . . . . . Sea surface temperature

T . . . . . Temperature of seawater

t . . . . . Time

$\mathbf{T_{eff}}$ . . . . . transfer efficiency

TEC . . . . . Transparent exopolymer carbon

TEP . . . . . Transparent exopolymer particle

$\mathbf{^{234}Th}$ . . . . . Thorium-234, an isotope of thorium

TKE . . . . . Turbulent kinetic energy

UVP . . . . . Underwater vision profiler

VGPM . . . . . Vertically-Generalised Production Model

w . . . . . Sinking velocity, also referred to as settling speed

WOA . . . . . World Ocean Atlas

z . . . . . Depth

z_{bathy} . . . . . A depth in the bathypelagic zone

z_{eu} **or** z_0 . . . . . Base of the euphotic zone (typically, the depth of 100 m)

z_{meso} . . . . . Base of the mesopelagic zone (typically, the depth of 1,000 m)

z^* . . . . . Remineralisation length scale, or depth interval over which POC flux declines by 63%

ZFT . . . . . Zooplankton functional type

ρ . . . . . Density

ξ . . . . . Slope of the particle size distribution (PSD)

Synonyms for terminology used in this thesis

bathypelagic depth . . . . . any depth below 1,000 m

cluster . . . . . super particle, super droplet, computational particle, individual particle, aggregate class

coagulation . . . . . aggregation, flocculation

data . . . . . observations, real-world observations, empirical observations, training data set, assimilated data

deep ocean . . . . . water column section beneath 1,000 m

error . . . . . uncertainty, deviation

euphotic zone . . . . . epipelagic zone, surface ocean, 0–100 m

export depth . . . . . around 100 m, depth of the euphotic zone, z_{eu}, depth of 1% of incident photosynthetic active radiation

export flux . . . . . POC flux leaving the 100 m depth horizon, POC flux leaving the euphotic zone

mesopelagic zone . . . . . twilight zone, 100–1,000 m (belongs to the ocean interior)

model outputs . . . . . model estimates, model predictions, model solutions

ocean interior . . . . . water column section beneath 100 m

optimiser . . . . . optimisation algorithm

particle attributes . . . . . particle variables, particle internal coordinates, particle physicochemical characteristics, particle properties

particle size distribution . . . . particle density distribution, particle size spectrum, particle number spectrum

particles . . . . . marine snow, marine aggregates, aggregates

sequestration depth . . . . . around 1,000 m, depth of the mesopelagic zone, z_{meso}

sequestration flux . . . . . POC leaving the 1,000 m depth horizon, POC leaving the mesopelagic zone

stochastic . . . . . probabilistic

Part I

INTRODUCTION

1

Introduction to my research problem

Contents

1.1 Motivation

1.1.1 A fundamental reaction for life on Earth

Our planet Earth has an exceptional atmosphere that has allowed life to proliferate beneath. The amount of atmospheric heat-trapping gases, popularly known as "greenhouse" gases (carbon dioxide, water vapour, methane...), is neither elevated, which would create a

planet with a scorching surface, like that of Venus, nor too small, which would originate a frozen surface, like that of Mars. At some 0.04% by volume of our atmosphere, the most abundant greenhouse gas, carbon dioxide, represents only a tiny fraction of the total Earth's atmospheric composition, which is fundamentally nitrogen (78% by volume) and oxygen (21% by volume). But that partitioning between oxygen (O_2) and carbon dioxide (CO_2) has created the perfect balance for life to thrive. However, the composition of our atmosphere has not always been like that.

During the Precambrian period (comprising over 80% of the Earth's history), there were trace amounts of O_2 in the atmosphere and plenty of CO_2, the latter believed to have been up to 100–1000 times the present atmospheric CO_2 concentration (pCO_2, or partial pressure of CO_2) of 418 ppm[1] (Kasting, 1993, Fig. 2). Since the Sun was fainter back then, a CO_2-driven atmospheric greenhouse would have been the factor keeping the Earth's surface temperature above freezing and thus maintaining water in a liquid form, a prerequisite for life (Catling and Zahnle, 2020). A popular theory in science is that life might have originated at the bottom of the sea, around hydrothermal vents (Martin et al., 2008), where it was protected from uninhabitable surface conditions, with high volcanic activity, the constant bombardment of meteorites and deadly UV radiation (as there was not a protective ozone (O_3) layer yet). The first organisms, which appeared around 4000 Mya, had a chemosynthetic metabolism and would synthesise their proteins, sugars and fats from the CO_2 dissolved in the water using the chemical energy released by the inorganic substances that spew from the vents. The photosynthetic metabolism would have appeared later and closer to the surface ocean, using the radiant energy of the Sun to drive the synthesis of complex biomolecules. First photosynthetic organisms did not generate O_2 as a byproduct; photosynthesis was anoxygenic. Later on, around 3200–2700 Mya (Catling and Zahnle, 2020), a biotic innovation appeared that would change our planet forever: photosynthetic organisms learned to use an unlimited source they had around, water. By splitting it using sunlight, they would free hydrogen (needed to reduce CO_2 into sugars, $C_6H_{12}O_2$) and release O_2; oxygenic photosynthesis had appeared. Photosynthesis is the only reaction on Earth that produces all the available O_2. As the photosynthetic metabolism became more widespread, the gradual increase of O_2 in the atmosphere catalysed two transcendental events. One was the evolution of aerobic metabolic pathways, known as aerobic respiration, which used O_2 to synthesise biomolecules, a strategy that offered a higher energy yield than the original anaerobic metabolism. That released a constrain for the appearance of multicellular life. The second one was the formation of an ozone layer that provided a shield against UV

[1] Atmospheric pCO_2 measurement from July 2021. The units, ppm (parts per million), indicate that there are 418 parts of CO_2 in one million parts of air, or that a little bit more than 0.04% by volume of our atmosphere is made of CO_2.

radiation. That allowed organisms to colonise land. The overall reaction of (oxygenic) photosynthesis is the reverse of (aerobic) respiration:

$$6CO_2 + 6H_2O + \text{light} \underset{\text{respiration}}{\overset{\text{photosynthesis}}{\rightleftharpoons}} C_6H_{12}O_6 + 6O_2. \tag{1.1}$$

The metabolic success of O_2 is materialised in the burst of sea life known as the Cambrian explosion (541–530 Mya), with the rapid surge of hard-bodied, large animals and most of present-day lineages.

With the emergence of O_2, a coupled biological–geochemical mechanism that would regulate the planet's surface temperature had emerged: the carbon cycle. By moving carbon (C) between the atmosphere, photosynthetic organisms, soils, the ocean, sediments, and rocks, the carbon cycle recycles an atom that results in warmer temperatures when released into the atmosphere. The carbon cycle balances the inputs of CO_2 into the atmosphere, where volcanism has been the main contributor in the geological past, with the removal via (i) silicate rock weathering on the continents and bounding to calcium ions in the oceans, forming carbonate rocks[2], (ii) the absorption by photosynthetic organisms on land and the surface ocean, forming living biomass, (iii) the burial of organic carbon (dead biomass) in sediments on land and the seafloor, forming fossil fuels, and (iv) the dissolution in the oceans, making seawater more acidic. The last time our planet would see a high atmospheric pCO_2, of 2000 ppm, would be in the Triassic, around 200 Mya (Foster et al., 2017). Since then, a long-term decline in atmospheric CO_2 to < 300 ppm (preindustrial value) has been achieved through silicate weathering and organic carbon burial.

The past 800,000 years (late Quaternary) is the only time in the geological record for which scientists have robust evidence of Earth's past climate. During that period, the carbon cycle has gone through cyclical changes called glacial–interglacial cycles (Fig. 1.1). The frequency of the cycles can be compared to that of the variations in the Earth's orbit, the Milankovitch cycles, which altered the amount of solar energy received at the Earth's surface, leading to cycles of ice ages (or glacial periods, with decreasing atmospheric pCO_2) and warm epochs (or interglacial periods, times with increasing pCO_2). Glacial–interglacial cycles last around 100,000 years, where glacials develop when the Northern Hemisphere is less tilted towards the Sun, decreasing summer insolation and, consequently, ice sheets built on land. The mechanisms by which a decrease in insolation, and thus temperature, lowers atmospheric pCO_2 in the glacial periods is an active area of research involving the ocean system, the largest reservoir of carbon ready to be mobilised on glacial–interglacial timescales (10^3–10^5 years) (Broecker, 1982; Archer et al., 2000;

[2] Chemical weathering of silicate rock by acid rain releases calcium and bicarbonate ions ($CaSiO_3 + 2CO_2 + H_2O \rightarrow Ca^{2+} + 2HCO_3^- + SiO_2$) that are carried by rivers into the oceans, where precipitation of calcium carbonate happens ($2HCO_3^- + Ca^{2+} \rightarrow CaCO_3 + CO_2 + H_2O$) and carbonate rocks form over time as $CaCO_3$ is buried in the sediments and turns to rock (limestone).

Sigman and Boyle, 2000; Kohfeld and Ridgwell, 2009; Sigman et al., 2010). Briefly, those mechanisms include (i) ocean physics, like increased solubility of CO_2 in the colder surface ocean (Khatiwala et al., 2019) or ice sheets stopping the upwelling of deep CO_2-rich waters (Stephens and Keeling, 2000), (ii) ocean biology, like increased phytoplankton growth due to increased iron fertilisation favoured by winds from colder, drier continental areas (Martin, 1990), and (iii) ocean chemistry, like changes in the "rain rates" of organic carbon to calcium carbonate, which affect whole-ocean alkalinity[3] and, thus, the solubility of atmospheric CO_2 (Archer and Maier-Reimer, 1994). During the interglacials, warming must have liberated to the atmosphere the carbon stored in the ocean during the glacial periods. Since CO_2 is a greenhouse gas, its release and removal from the atmosphere has acted amplifying the changes in temperature triggered by changes in the Earth's orbit.

1.1.2 The anthropogenic perturbation

In the past three centuries, humans, with our industrial and technological revolutions supported by changes in agricultural land use (e.g., deforestation) and burning of fossil fuels (e.g., generation of electricity), are perturbing the carbon cycle. The Global Carbon Project (Friedlingstein et al., 2020) estimates the total discharge of CO_2 into the atmosphere since 1750 of ~2560 Gt CO_2 (equivalent to ~700 Gt C[4]), half of which was released only between 1990 and 2020. Atmospheric pCO_2 has risen from a preindustrial of ~280 ppm to

[3] Here are some concepts needed to understand the relationship between seawater alkalinity and atmospheric CO_2. *Alkalinity* (ALK) is the seawater's acid buffering capacity; otherwise said, the excess of bases (proton acceptors) over acids (proton donors), ALK = $[HCO_3^-] + 2[CO_3^{2-}] + [OH^-] - [H^+]$. The ocean receives ALK from riverine inputs after chemical weathering of emergent rocks on land. Alkalinity is a conservative property and thus riverine inputs must balance with marine precipitation and burial of $CaCO_3$ ($CaCO_3$ formation removes ALK). Thus, in the eventuality of a deficit or excess of ALK (i.e., after enhanced or decreased $CaCO_3$ burial), the ocean readjusts the ALK budget through $CaCO_3$ dissolution. The depth below which seawater is undersaturated with respect to $CaCO_3$, so dissolution becomes thermodynamically favourable, is called the *carbonate saturation horizon* (CSH), and depends on the concentration of carbonate ion (CO_3^{2-}). The CSH is a transition separating shallower seafloor sediments where $CaCO_3$ is preserved (as there is abundant CO_3^{2-}), from deeper seafloor sediments, where $CaCO_3$ starts to dissolve (as CO_3^{2-} become scarce). Variations of the CSH controls ocean's ALK and atmospheric CO_2 through the redistribution of $CaCO_3$ in the water, a mechanism called *carbonate compensation*. Here is how it works. In the situation of increased transfer of organic carbon to the ocean interior (for instance, after iron fertilisation), there will be an increase in the amount of organic carbon degraded in the deep ocean and, thus, in the amount of CO_2 regenerated at depth. Increased CO_2 lowers the concentration of CO_3^{2-} in the deep ocean as the CO_3^{2-} soaks up the excess protons (H^+) released by the dissolution of CO_2 in water, following Le Chatelier's principle ($\uparrow CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons HCO_3^- + \uparrow H^+ \rightleftharpoons \downarrow CO_3^{2-} + \uparrow H^+$). As CO_3^{2-} ions disappear (and thus ALK decreases), $CaCO_3$ must readily dissolve to replenish CO_3^{2-} and restore ALK. That pushes up (or shoals) the CSH, with the consequence that $CaCO_3$ starts to dissolve closer to the surface, decreasing the burial rate of $CaCO_3$. Since $CaCO_3$ burial is the main mechanism by which the ocean removes the ALK input from rivers, a decrease in $CaCO_3$ burial translates into an excess in the input of ALK. Higher whole-ocean ALK increases the ability of seawater to hold more atmospheric CO_2 (which adds protons to the water), i.e., a greater ALK increases CO_2 solubility. The increase in the stock of surface CO_3^{2-} (after $CaCO_3$ dissolution) propagates gradually through depth and, thus, the CSH bounces back down and $CaCO_3$ burial rate is reestablished. By this point, steady state between riverine inputs and sedimentary burial has been reached, carrying with it a decrease in atmospheric CO_2.

[4] 1 Gt C = 1 Pg C = 1 billion metric tons = 10^{15} g C = 3.67 Gt CO_2.

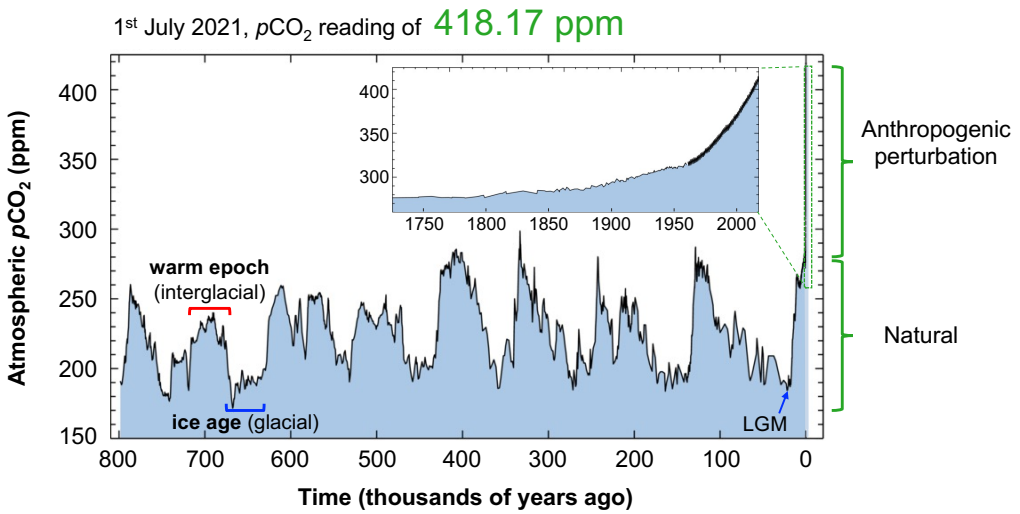


Figure 1.1. Time evolution of the atmospheric levels of CO_2 (pCO_2) in parts per million (ppm) over the past 800,000 years (late Quaternary). A cyclical pattern is present, with ice ages (or glacial periods, times with decreasing pCO_2) and warm epochs (or interglacial periods, times with increasing pCO_2) clearly distinct. Currently, we are in an interglacial period, and the last glacial maximum (LGM) occurred 18,000–22,000 years ago. Enlarged are the postindustrial years (the Industrial Revolution started in 1750), with the anthropogenic perturbation increasing atmospheric pCO_2 from a preindustrial of ~280 ppm (< 0.03% by volume) to the current 418 ppm (> 0.04% by volume). Since 1958, data are from the Mauna Loa (Hawaii) station, which has the longest-running record of direct measurements of atmospheric pCO_2, known as the Keeling curve. Before 1958 and all the way back to 800,000 years ago, data are from Antarctic ice cores, where ancient air was trapped in bubbles in polar ice, a proxy for past atmospheric pCO_2 conditions. ppm = ratio of the number of moles CO_2 in a given volume of dry air to the total number of moles of all air constituents in this volume. Source: author's elaboration with data from https://scripps.ucsd.edu/programs/keelingcurve/.

the current 418 ppm, a rate of increase unprecedented in the past 800,000 years (Fig. 1.1), pushing the Earth away from known past climate conditions. Despite Callendar first warning us in 1938 about the warming effect that increased amounts of the greenhouse gas CO_2 would have in the atmosphere, humanity turned a blind eye to it.

The oceans are estimated to have absorbed already a quarter of the total anthropogenic CO_2 emissions (155–170 Gt C, Khatiwala et al., 2013; Gruber et al., 2019; Friedlingstein et al., 2020) (Fig. 1.2). Had not been by the capacity of the ocean to sequester CO_2, today there would be at least ~200 ppm[5] more of atmospheric CO_2 (Parekh et al., 2006), making our planet warmer and thus an uninhabitable place for most forms of life –e.g., enhanced wildfire regimes and tropical storms, desertification and erosion with loss of soil nutrients, reduction of freshwater reserves, meltdown of polar ice and release of deposits of the greenhouse gas methane, sea-level raise, change in the intensity and direction of the global

[5]Value estimated by idealised box models of the ocean–atmosphere carbon cycle, like that of Sarmiento and Toggweiler (1984) and Toggweiler (1999), and which attribute the ocean's uptake of ~200 ppm to the "soft-tissue" biological carbon pump. That value of 200 ppm is in the low end of the full capacity of the ocean to reduce atmospheric pCO_2 as there are other mechanisms acting on top of the soft-tissue pump, like changes in the solubility of CO_2 mediated by temperature and alkalinity.

ocean conveyor belt, ocean deoxygenation and acidification, etc. (Cavallaro et al., 2018, SOCCR2). Life on Earth depends on the capacity of the sea to sequester CO_2.

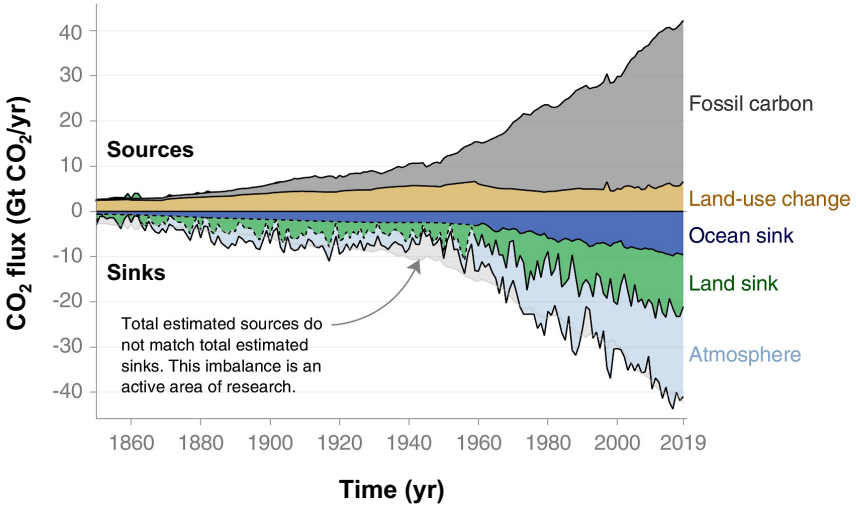


Figure 1.2. Balance of sources and sinks of CO_2 in gigatons of CO_2 released per year (Gt CO_2 yr^{-1}) over the period 1850–2020. At present, the annual increase of CO_2 in the atmosphere is less than half of the calculated emissions, with some quarter being absorbed by land (via photosynthesis) and another some quarter by the ocean (via dissolution in cold waters and its transport to depth). Source: Global Carbon Project, supplemental data of Global Carbon Budget 2020, v1.0, https://doi.org/10.18160/GCP-2020.

1.1.3 The ocean carbon pump

The ocean water column is the largest reservoir of carbon in our planet (38,703 Gt C) after sedimentary carbonate rocks (~100,000,000 Gt C). It contains 45 times more carbon than the atmosphere and 15 times more than land plants and soils (Fig. 1.3). Practically all of the carbon stored in the ocean is in its dissolved phase. The predominant form (38,000 Gt C) is the dissolved inorganic carbon (DIC), the sum of dissolved CO_2 and the carbon species it produces when it reacts with water, i.e., acid carbonic (H_2CO_3), bicarbonate ion (HCO_3^-) and carbonate ion (CO_3^{2-}). The second most abundant form (700 Gt C) is the dissolved organic carbon (DOC, operationally defined as organic particles < 0.2 μm in diameter), released by the marine plankton (phyto-, zoo- and bacterioplankton) as a result of their metabolic activity. DOC storage represents a small fraction of the total water-column carbon storage, however a quantity similar to the carbon stored in the atmosphere (860 Gt C). Notice how if only 2% of the amount of DOC stored in the ocean would oxidise, it would release to the atmosphere more CO_2 (2% × 700 Gt C × 3.67 = 51 Gt CO_2) than that already emitted by fossil fuels in a year (37 Gt CO_2 yr^{-1}, Fig. 1.2). Two other less abundant forms of carbon exist in the ocean and are in the transient, particulate phase. One is the particulate organic carbon (POC), which in Fig. 1.3 is in the form of marine biota (or all plankton, 3 Gt C), but also encompasses

non-living material (plankton waste products) that sinks in the water column in the form of marine particles. The other is the particulate inorganic carbon (PIC), the least abundant form of oceanic carbon (~0.03 Gt C, Brewin et al., 2021), encompassing all the calcium carbonate ($CaCO_3$), predominantly formed by calcifying organisms (mostly plankton, but also corals or molluscs) and, to a lesser degree, formed spontaneously in supersaturated waters. PIC and POC formed in the surface ocean rain to the seafloor sediments, where they accumulate and cement, forming over time deposits of carbonate rocks (limestone) and fossil fuels, respectively.

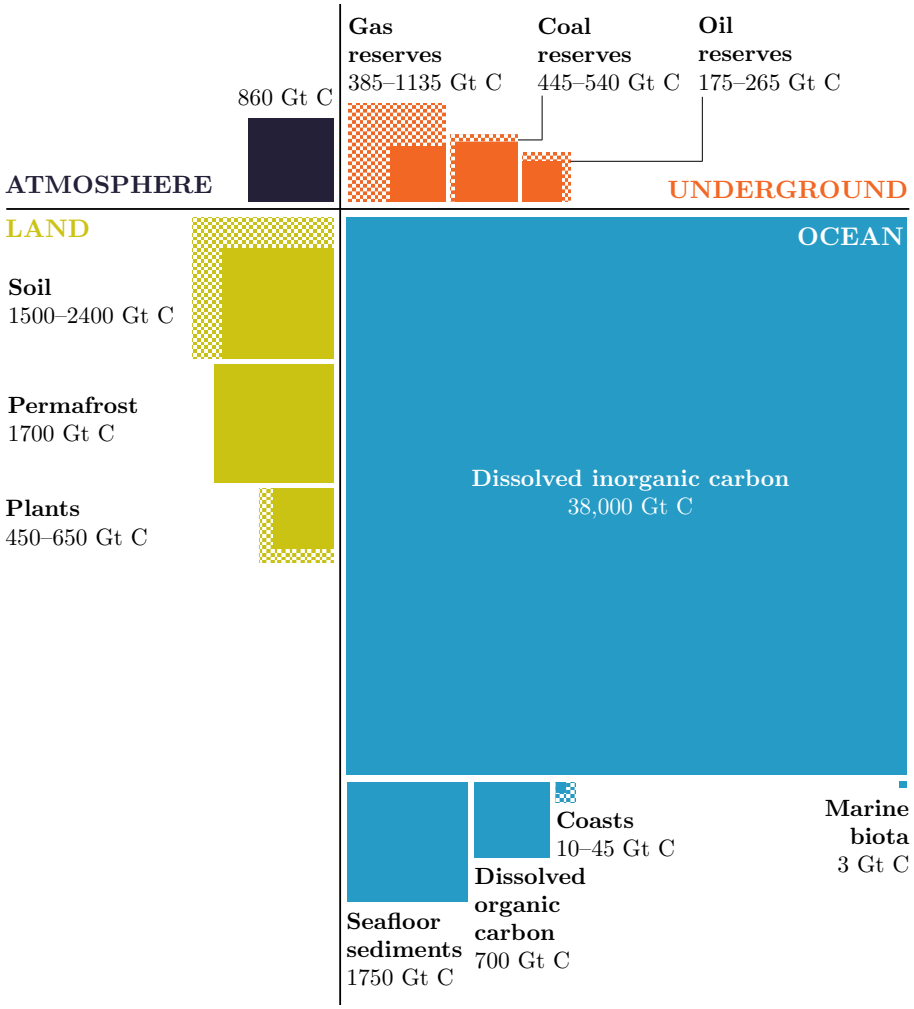


Figure 1.3. Global carbon storage in 2020 in gigatons of carbon (Gt C), with its four main short-term reservoirs: the atmosphere (860 Gt C), land (~3700–4700 Gt C), the ocean (~40,500 Gt C) and the underground on land and sea (~1000–2000 Gt C). The long-term reservoir of carbon, the sedimentary carbonate rocks (~100,000,000 Gt C), is not shown. Dissolved inorganic carbon (DIC) includes four carbon species: dissolved carbon dioxide (CO_2), carbonic acid (H_2CO_3), bicarbonate ions (HCO_3^-) and carbonate ions (CO_3^{2-}). Marine biota refers to plankton biomass, which is a form of particulate organic carbon (POC) and biogenic calcium carbonate ($CaCO_3$). Dissolved organic carbon (DOC) is a byproduct of the biological carbon pump generated from POC, a transient form of carbon. Seafloor sediments represent the amount of POC and $CaCO_3$ that reach the seafloor. Source: redrawn from Friedlingstein et al. (2020) after REUTERS.

The reason the ocean is able to hold much more carbon than the atmosphere is due to the particular chemistry of CO_2 in seawater. Due to the slightly basic pH of seawater, most of the atmospheric CO_2 gas that reacts with water gets converted into nongaseous DIC species (~90% of DIC is HCO_3^- and ~9% is CO_3^{2-})[6], and very little remains as CO_2 gas ($< 1\%$). That particular chemistry of CO_2 in seawater causes a reduction in the CO_2 gas pressure in the surface water that allows more diffusion from the atmosphere, following Henry's law. Once in the water, DIC is further pulled downwards, to the deep ocean, by a combination of ocean physical, chemical and biological mechanisms collectively known as the ocean carbon pump (after Volk and Hoffert, 1985). The ocean carbon pump takes up surface water CO_2 and transfers it to the deep ocean (> 1000 m), where it accumulates and becomes isolated from the atmosphere on a millennial timescale (the ocean overturning timescale) before it is upwelled back to the surface ocean and outgasses into the atmosphere. Thus, the ocean carbon pump creates a surface-to-depth gradient in the distribution of carbon in the ocean, with deep waters enriched in DIC with respect to the upper ocean, further lowering surface water pCO_2 and thus reinforcing the diffusion of more atmospheric CO_2 into the ocean. The CO_2 sequestered by the ocean carbon pump has tempered the Earth during glacial ages (Fig. 1.1).

Volk and Hoffert (1985) distinguished three ocean carbon sub-pumps, corresponding to different forms of carbon found in the marine environment (Fig. 1.4):

- the solubility carbon pump, a physicochemical pump controlled by seawater temperature and operating on DIC;
- the carbonate counter pump (or hard-tissue pump), a biochemical pump controlled by calcifying phyto- and zooplankton (coccolithophores and foraminifera) and alkalinity and operating on $CaCO_3$ (a form of PIC), and
- the soft-tissue pump, a biochemical pump that depends on the photosynthesis of phytoplankton and operating on biologically produced (biogenic) organic carbon (POC and DOC).

The carbonate counter pump and the soft-tissue pump collectively form the biological carbon pump (BCP). The term "biological" refers to the biologically produced material (POC, DOC and PIC) (Boyd et al., 2019)[7], and the term "pump" refers to the transport

[6] HCO_3^- and CO_3^{2-} are proton acceptors (i.e., bases) thus, dominant forms in a basic solution like seawater (see a Bjerrum diagram).

[7] The term "biological" in BCP originally meant that the transfer of biogenic carbon to the deep ocean was mediated by *biological processes*, which generated fast-sinking particles transferred to depth by gravity (Volk and Hoffert, 1985; Buesseler et al., 2007b). That idea is currently outdated, as evidenced by the presence of deep carbon in the form of neutrally-buoyant and slowly-sinking POC as well as DOC that would not reach the deep ocean if fast gravitational sinking (mediated by biological processes) was the sole mechanism supplying carbon to depth (Giering et al., 2014; Omand et al., 2015; Dall'Olmo et al., 2016; Villa-Alfageme et al., 2016; Baker et al., 2017; Bol et al., 2018). Thus, the definition of the BCP has been broadened (see the reviews of Boyd et al., 2019; Le Moigne, 2019) to also include *physical processes*, like ocean mixing and subduction, which mediate the supply of an important fraction of the deep biogenic carbon.

of that carbon against a concentration gradient, from the surface waters (low DIC) to the deep ocean (high DIC) (Volk and Hoffert, 1985). The BCP is thought to be responsible for two-thirds of the vertical gradient of DIC, with the solubility pump accounting for the other third (Sarmiento and Gruber, 2006; Marinov et al., 2008).

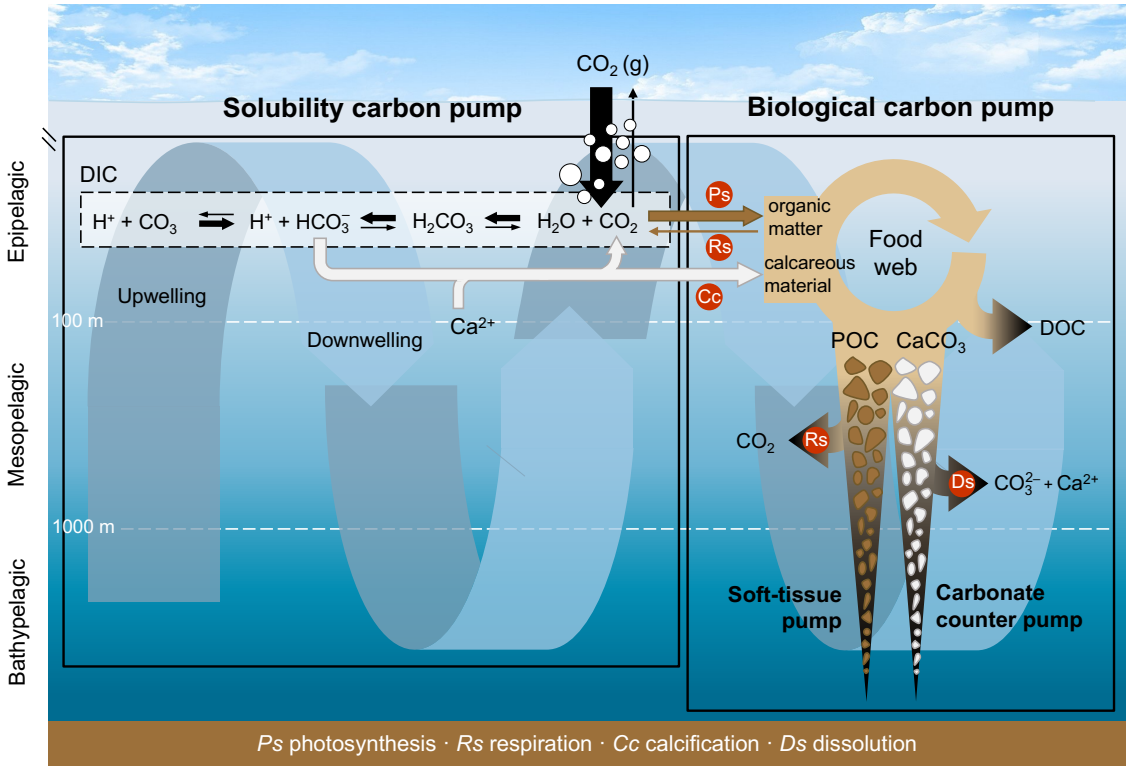


Figure 1.4. Schematic representation of the ocean carbon pump, with its three sub-pumps: solubility carbon pump, soft-tissue pump and carbonate counter pump (the latter two collectively known as the biological carbon pump, BCP). This artistic depiction of the BCP emphasises the role of the gravitational pump, the BCP mechanism contributing the most in the sequestration of biogenic carbon (food-web generated). Here is some information on magnitudes. POC export (POC flux below 100 m) represents ~10% of the surface ocean primary production, POC sequestration (POC flux below 1000 m) represents ~1%, and POC burial (at the seafloor) is < 0.1% (Passow and Carlson, 2012). The soft-tissue carbon pump exports ~10 times more carbon than the carbonate counter pump (Sarmiento et al., 2002). On a century timescale, the solubility carbon pump dominates the sink of carbon, but on a millennial timescale the BCP is thought to be controlling the glacial–interglacial oscillations (Parekh et al., 2006). In terms of carbon inventory size, from larger to smaller: DIC > DOC > POC > PIC (Brewin et al., 2021). DIC, dissolved inorganic carbon; CO_2, carbon dioxide; H_2CO_3, carbonic acid; HCO_3^-, bicarbonate ion; CO_3^{2-} carbonate ion; Ca^{2+}, calcium ion; DOC, dissolved organic carbon; POC, particulate organic carbon; $CaCO_3$, calcium carbonate (also termed particulate inorganic carbon, PIC).

1.1.3.1 The solubility carbon pump

The solubility carbon pump is propelled by low temperatures, which increase the solubility of gases in seawater. Thus, the solubility carbon pump is mainly active at high latitudes, where the cooling of surface waters favours the dissolution of atmospheric CO_2 as well as

increases water density. Those heavy, cold surface water masses rich in DIC sink to the ocean interior (downwelling), becoming part of the deep ocean circulation "conveyor belt" and isolating carbon from the atmosphere from centuries to millennia. Eventually, ocean mixing and upwelling currents bring these deep waters to the ocean surface and CO_2 outgasses into the atmosphere. CO_2 outgassing is especially strong in the tropics, where warm waters cannot retain as much CO_2, and the Southern Ocean, with a wind-driven upwelling.

1.1.3.2 The biological carbon pump (BCP)

The BCP is propelled by the photosynthesis of phytoplankton living in the sunlit surface ocean (the euphotic zone, or upper ~100 m), which uptake the CO_2 dissolved in seawater and fix it, producing POC and reducing atmospheric pCO_2. That process is what we know as marine primary production and accounts for about half of the global net primary production (NPP) (Field et al., 1998). On their part, calcifying phytoplankton (coccolithophores) take up HCO_3^- to form their carbonate shells (coccoliths), producing $CaCO_3$. That process, calcification, releases CO_2 to the surrounding shallow water, which escapes to the atmosphere, counteracting the soft-tissue pump.

Phytoplankton sustain the marine food web and, as they get eaten or die, primary production and calcareous material get transferred to "marine snow", a myriad of biogenic detrital particles left behind by the operations of the marine food web (see the reviews of Stemmann and Boss, 2012; Turner, 2015). Marine snow (> 0.5 mm in diameter) include things like clumps of aggregated phytoplankton (phytodetritus), faecal pellets produced by zooplankton, discarded organic and mineral components (e.g., phytoplankton exudates or zooplankton moults), zooplankton dead bodies, and combinations of those, which mostly aggregate by sorption onto larger particles. The association of POC with calcareous particles (PIC), denser than organic material, adds ballast, increasing particle sinking velocity and enabling a faster transfer to depth (e.g., Klaas and Archer, 2002). Those larger and denser particles sink away to the ocean interior, pulling carbon from the surface into the deep ocean and thus forming a rain, or flux, of POC and PIC.

The IPCC (Intergovernmental Panel on Climate Change) stipulates that an atom of carbon is considered *sequestered* in the ocean when the return to the atmosphere happens more than a hundred years since it was initially fixed by phytoplankton. That criterion is met when the atom of carbon is transported below the 1000 m depth horizon (the deep, or bathypelagic, ocean), where the ocean ventilation timescale is of 100–1000 years (Passow and Carlson, 2012). There are three main mechanisms by which the BCP sequesters biogenic carbon in the deep ocean (Boyd et al., 2019; Le Moigne, 2019).

- Gravitational sinking of particles due to their large size and/or high density. This mechanism is commonly known as the gravitational pump and it is the most studied BCP mechanism as, on average, has the highest contribution in the sequestration of biogenic carbon.
- Transport by vertically migrant zooplankton (both migrating diurnally and ontogenetically), which ingest particles in the surface ocean and migrate at depth (typically ~400 m), where they avoid predation and release the metabolised food as compact faecal pellets (egestion), DOC (excretion), DIC (respiration) and body carcasses (death).
- Transport by physical processes of downward water motion (subduction and displacement of the mixed layer depth), which carry surface particulate material that has been fragmented into suspended POC and solubilised into DOC to the ocean interior.

Whereas the particles carried by the migrant and the physical transport pumps bypass the upper mesopelagic ocean (100–500 m), where the decomposition (remineralisation) of the organic carbon begins, the particles carried by the gravitational pump do not short-circuit the upper mesopelagic. Sinking gravitational particles are thus fully exposed to processes that will transform them and attenuate the particulate flux. Thus, before being sequestered, the sinking POC experiences (Boyd and Trull, 2007; Collins et al., 2015): (i) remineralisation into CO_2 as the organic carbon is consumed and respired by zooplankton and heterotrophic bacteria; (ii) disaggregation into smaller particles either abiotically (by spontaneous breakup induced by water shear or by collision to other particles) or mediated by zooplankton as they mishandle food particles before ingestion (termed "sloppy feeding" after Jumars et al., 1989), and (iii) solubilisation into DOC and suspended POC by particle-attached bacteria (termed "microbial gardening" after Mayor et al., 2014). Respiration by bacteria is the main sink of POC in the mesopelagic ocean and thus the primary source of DIC and regenerated dissolved nutrients in the ocean interior (Buesseler and Boyd, 2009; Sanders et al., 2014). Depending on how close to the surface ocean POC is regenerated into DIC, oceanic carbon may be stored for as little as months to as long as millennia (DeVries et al., 2012).

The attenuation of POC flux in the mesopelagic presents a global distribution that is thought to be controlled either by factors that stimulate the sinking flux of POC (e.g., NPP or mineral composition of the phytoplankton community) or by factors that stimulate its mechanical attenuation (e.g., temperature or O_2 concentration, on which respiration depends) (Weber et al., 2016; DeVries and Weber, 2017; Cram et al., 2018). Those ecosystem and environmental factors have been studied using an amalgam of different *in situ* sampling, remote sensing and modelling approaches (see review in Brewin et al.,

2021) without a common framework on how to bring together contrasting spatiotemporal scales (e.g., Bisson et al., 2018) or apply reference metrics (e.g., Buesseler et al., 2020b). The variety of methods has generated a wide range of values around the magnitude of the downward flux of POC. Thus, the BCP is though to export to the ocean interior between 3.8 and 11 Gt C yr^{-1}, or 7–29% of the global NPP (50 Gt C yr^{-1}, Carr et al., 2006), and to sequester 0.12–0.66 Gt C yr^{-1}, or 0.24–1.3% of the global NPP (Table 1.1). That implies that ~83–99% of the exported carbon flux is lost as DIC into the mesopelagic, or that 1–17% of the exported carbon flux is sequestered in the mesopelagic. Only a tiny fraction of the NPP, $< 0.1\%$, is eventually buried into the sediments. Oil, gas and coal are examples of marine snow falling to the ocean floor and being buried and crushed under tons of sediments for millions of years.

Table 1.1. Global POC flux estimates at export depth (~100 m) and sequestration depth (~1000–2000 m) from a variety of measurement techniques, biogeographical scales and temporal scales.

Reference	Export flux (Gt C yr^{-1})	Sequestration flux (Gt C yr^{-1})	Export efficiency (%)	Sequestration efficiency (%)
Martin et al. (1987)		0.12		
Lampitt and Antia (1997)		0.34		
Laws et al. (2000)	11		21	
Gehlen et al. (2006)	5–10		14–29	
Dunne et al. (2007)	9.6		19*	
Lutz et al. (2007)	4.6	0.31	9.2*	6.7†
Honjo et al. (2008)	5.7	0.43	16	7.6
Henson et al. (2012)	4	0.66	10	19
Lima et al. (2014)	6.0	0.21	13	5.5
Siegel et al. (2014)	5.7		11	
Guidi et al. (2015)		0.33		
DeVries and Weber (2017)	6.7		13*	
Bisson et al. (2018)	3.8–5.5		7–10	
Stock et al. (2020)	6.4	0.47	13*	7.3†

* Export efficiency value not reported in the study but calculated afterwards assuming a global NPP of 50 Gt C yr^{-1} (Carr et al., 2006) as 100 × (POC export flux/NPP).
† Sequestration efficiency value not reported in the study but calculated afterwards as 100 × (POC sequestration flux/POC export flux).

1.1.4 Hypotheses on the effects of a higher pCO_2 world on the BCP

Scientists expect a gradual slowing of the ocean BCP-mediated uptake of CO_2 if fossil fuel emissions continue to accelerate (Boyd, 2015; Moore et al., 2018; Friedlingstein et al., 2020; Stock et al., 2020). Nevertheless, it is challenging to predict with certainty

the net effect of the anthropogenic perturbation on the BCP. Think of the cumulative effects of multiple climate-derived stressors (increased seawater temperature, sea-ice melt, water stratification) with a differing sign of action (positive vs negative feedbacks), an associated seasonality and operating at different spatial scales (local vs ubiquitous) (Boyd and Trull, 2007; Turner, 2015; Moore et al., 2018). It is hypothesised that, in the light of climate change, these four biogeochemical variables will have the highest impact on the BCP: seawater temperature, phytoplankton community composition, O_2 levels and light and nutrient availability for NPP (Passow and Carlson, 2012; Sanders et al., 2016). Here is a checklist of their proposed effects on the BCP, where negative feedbacks on the BCP (i.e., processes that mitigate the effect of a perturbation) are denoted by a "−" sign and positive feedbacks (i.e., processes that amplify the effect of a perturbation) are denoted by a "+".

- Seawater temperature, with direct controls on (i) vertical supply of deep nutrients to surface waters, (ii) heterotrophic and (iii) autotrophic metabolism and (iv) water thermodynamics.
  - (i) A warmer ocean will be more stratified and less well mixed, which will reduce the supply of deep macronutrients (nitrate, phosphate and silicic acid) to surface waters through convection. Two consequences can be drawn.
    - A more starved phytoplankton, which will mean a reduced POC export flux to the ocean interior (+).
    - But an increase in stratification will confine phytoplankton in the very surface layer, where there is more light. NPP could be boosted (and therefore increase POC export flux) if it is sustained by some other form of nutrient supply that is not vertical mixing, e.g., through exogenous supply (warming of the upper ocean will increase storm frequency, promoting pulsed delivery of nutrients into the euphotic zone), or through autochthonous supply (stratification could increase the retention and utilisation of regenerated nutrients) (−).
  - (ii) Heterotrophic respiration rates increase with temperature (Iversen and Ploug, 2013; Marsay et al., 2015; Laufkötter et al., 2017; Cram et al., 2018).
    - With higher respiration rates, POC in the ocean interior will be remineralised into DIC at a faster pace, meaning that DIC stock will increase in the deep layers (−). At the same time, surface, warmer layers will be outgassing more CO_2. Both effects could overall increase the vertical gradient of DIC, favouring the drawdown of atmospheric CO_2 by the ocean in the short term (−).

    - But higher respiration rates of surface POC will reduce the export of organic carbon to the deeper ocean, thus reducing, in the long term, DIC stock in the deep ocean (+).
    - Increased remineralisation of surface organic matter will also result in an increased nutrient supply (regenerated production) to the euphotic zone, boosting NPP in the short term (−).
  - (iii) Phytoplankton growth rates increase with temperature.
    - Higher growth rates increase NPP and, thus, POC export flux (−).
    - Higher growth rates lead to a higher intake of macronutrients, which get, in the long term, depleted, causing phytoplankton starvation (+).
  - (iv) A warmer ocean has a higher kinetic energy and water is less viscous.
    - In a higher kinetic energy environment and in less viscous water, particles will be colliding at higher rates, thus favouring the formation of larger particles, which sink faster to the ocean interior (−).
    - But higher collision rates could also work the other way round, favouring particle fragmentation after impact, thus, breaking up large particles into smaller, slow-sinking particles (+).
    - Warmer waters are less viscous, and particles in less viscous water settle faster (De La Rocha and Passow, 2007) (−).

- Phytoplankton starvation (induced after a reduced nutrient supply from deep waters) will prompt shifts in phytoplankton community composition (Lam et al., 2011; Henson et al., 2012; Boyd, 2015; Bach et al., 2019). Smaller taxa (picophytoplankton and coccolithophores), which can grow with less nutrients, will be favoured over larger taxa (diatoms). That affects (i) the magnitude of the POC flux exported from the surface layer, (ii) the mineral composition of the flux (ballast effect and calcification rates), (iii) the lability of the organic carbon and (iv) grazing relationships.
  - (i) If the flora composition shifts to one dominated by smaller phytoplankton, two things can happen in terms of the intensity of the flux of POC.
    - The average size of a particle will be reduced, thus reducing the sinking speed of aggregates (Stemmann and Boss, 2012; Bach et al., 2019) and, therefore, reducing POC export flux (+).
    - The release of transparent exopolymer particles (TEP), primarily produced by larger phytoplankton (Passow, 2002), will decrease. Since TEP create a sticky matrix that holds aggregates together, the reduction in the TEP stock will reduce the formation of large, fast-sinking aggregates, and

thus reduce POC flux (+). But it has been observed that a more acidic environment, alongside the higher metabolic rates in a warmer ocean, could also increase the release of TEP (Campoverde et al., 2018) (−).

  - (ii) Changes in the flora composition come with changes in the mineral composition of the flux and the amount of calcification.
    * If picophytoplankton (i.e., non-mineralising phytoplankton) increase their abundance over calcifying coccolithophores and silicifying diatoms, there will be a reduction in the ballast effect on POC flux (+).
    * If coccolithophores, denser than diatoms, increase their relative abundance, there will be an increase in the ballast effect on POC flux (Armstrong et al., 2002) (−).
    * Calcification will increase if coccolithophores become more abundant. Calcification is a process that increases CO_2 in the water column, thus offsetting the positive influence of the ballast effect of coccolithophores (+). Furthermore, increasing amounts of dissolved CO_2 in the surface ocean will create a more acidic environment, which in turn will reduce calcification, thus reducing mineral ballast (+).
  - (iii) If there is a reduced biomineral content in particles, the organic carbon embedded in them will become more labile, as it will be more accessible to particle-attached bacteria (reduced physical protection), thus affecting negatively particle resilience in the water column (Armstrong et al., 2002; Francois et al., 2002) (+).
  - (iv) Grazing relationships (i.e., coupling between a phytoplankton cell and the predator's mouth) will be affected. If the fora composition shifts to one dominated by smaller cells before the zooplankton's feeding structure can adapt to smaller prey cells, there will be a decoupling between the sizes of prey and predator. Consequently, there will be an accumulation of non-grazed biomass in the euphotic zone and a reduction in the amount of particles that could have potentially been exported to the mesopelagic via fast-sinking faecal pellets egested by zooplankton (Lam et al., 2011) (+).

- Reduced O_2 levels in the water column (triggered by an increase in stratification in a warmer ocean) will affect (i) respiration and (ii) zooplankton diel vertical migration (DVM) (DeVries and Weber, 2017; Laufkötter et al., 2017; Cram et al., 2018).
  - (i) Increased stratification, and thus reduced ventilation of the ocean interior, will increase deoxygenation and the expansion of hypoxic waters, poisonous for aerobic metabolism (+); but it could also have POC preserved by depressing heterotrophic respiration (−).

  - (ii) Zooplankton, limited by the availability of O_2, will get confined to better ventilated waters, which are shallower. Heterotrophic respiration will increase in surface layers (+) as well as the production of compact faecal pellets (−).

- Increased light availability and nutrient stocks (due to melting of sea ice in a warmer world) will affect NPP.
  - (i) A lesser extent of sea ice cover will increase light penetration, and thus light availability in the surface ocean. That will boost NPP (−) and the growth of larger phytoplankton species (diatoms), with higher light requirements than smaller species (Marañón, 2009) (−).
  - (ii) Melting of the ice will release nutrients into the water, boosting NPP in the short term (−). But, in the long term, nutrients stripped out from the surface ocean and not replenished will drive phytoplankton to starvation (+).
  - (iii) The input of freshwater (meltwater runoff), which is more buoyant than seawater, will weaken vertical mixing, causing shoaling of the mixed layer depth and confining phytoplankton even closer to their light source (the uppermost surface ocean) (−).

1.1.4.1 The scientific and the public response

Our inability to resolve the direction of change in the BCP-mediated sequestration of atmospheric CO_2 has made of this an active area of ocean research in the past 30 years (Boyd and Trull, 2007; Jin et al., 2020), with major ocean POC flux surveys starting in the 1980s (VERTEX[8]) and 1990s (JGOFS[9]). Since then, progress has been made in three key areas: (i) advancing our understanding of the climate-sensitive mechanisms that control the efficiency of the biological pump of POC to the deep ocean through mesocosm experiments, field campaigns and numerical modelling (Doney et al., 2002; Ito and Follows, 2005; Boyd and Trull, 2007; Buesseler et al., 2007b; Sanders et al., 2016; Siegel et al., 2016; Boyd et al., 2019; Henson et al., 2019; Le Moigne, 2019); (ii) improving the various technological equipment that sample the ocean to obtain a wider coverage and more reliable measurements that allow intercomparison between methodologies (Bisson et al., 2018; Giering et al., 2020), and (iii) improving the still crude representation of ocean biology in Earth system models (ESMs), including those used by the IPCC (Séférian et al., 2013, 2020; Moore et al., 2018; Cavicchioli et al., 2019; Stock et al., 2020). In that light, several interdisciplinary field programs, like

[8] Vertical Carbon and Exchange (Martin et al., 1987).
[9] Joint Global Ocean Flux Study (Doney et al., 2002).

the NASA's EXPORTS[10] (U.S.), COMICS[11] (U.K.), GOCART[12] (U.K.), CUSTARD[13] (U.K.), FLUXES (Spain), WHOI's Ocean Twilight Zone[14] (U.S.), and more, have taken off in the past five years to provide an increased number of observations and improved observational technologies, and platforms like JETZON[15] have been created to improve data sharing. This thesis is part of COMICS (Controls over Ocean Mesopelagic Interior Carbon Storage, NE/M020835/2), a large grant project funded in 2015 by the Natural Environmental Research Council (Sanders et al., 2016). It aims to achieve a mechanistic understanding of the key processes that control mesopelagic carbon remineralisation, both from an observational and modelling perspective.

The situation has also raised public awareness. Our incomplete understanding of ocean biological CO_2 sequestration poses a drawback for improving the estimates of the relative partitioning of carbon between the ocean and the atmosphere in future climatic scenarios and paleoenvironmental reconstructions (Heinze et al., 2015; Jin et al., 2020). That is a necessary step to inform the public of the appropriate socioeconomic strategies to be taken to reduce our footprint on the ocean system and mitigate the adverse effects of a weaker BCP. In 2009, the United Nations Environment Programme (UNEP) coined the term "blue carbon" to refer to the carbon sequestered by marine ecosystems via the BCP (Nellemann et al., 2009, UNEP). It brought to light that half of the carbon captured by photosynthesis globally is blue carbon, and most of it is sequestered around the coasts, with their enhanced primary production and advective waters. That highlighted to the world's governments the essential ecosystem service the oceans provide by sequestering atmospheric CO_2 through the BCP and the importance of adequately managing coastal ecosystems –amongst others– to slow down climate change.

1.1.5 Metrics to characterise the BCP

The gravitational pump is the most studied BCP mechanism to date (Boyd et al., 2019) and the one for which metrics have been defined to quantify the sequestration of biogenic carbon. Three main types of metrics exist.

- According to the "strength" of the BCP, which is set in the surface ocean, and corresponds to the amount of POC exported from the surface to the mesopelagic ocean. It is measured by the metric particle export efficiency (PE_{eff}, after Buesseler and Boyd, 2009).

[10] https://oceanexports.org
[11] https://www.comics.ac.uk
[12] https://projects.noc.ac.uk/gocart/
[13] https://roses.ac.uk/category/custard/
[14] https://twilightzone.whoi.edu
[15] www.jetzon.org

- According to the "efficiency" of the BCP, which is set in the subsurface ocean, and has four complementary views in the literature: (a) how much of the POC exported survives respiration in the mesopelagic, measured by the metric transfer efficiency (T_{eff}, after Buesseler and Boyd, 2009), (b) how deep POC penetrates the subsurface ocean before being respired, measured by the metric remineralisation length scale (z^*, after Lutz et al., 2002), (c) the rate of attenuation of POC flux in the subsurface ocean, measured by the metric Martin's b (after Martin et al., 1987), and (d) the association of POC with ballast minerals, measured by the metric POC/PIC ratio (after Armstrong et al., 2002) and bSi/PIC ratio (after Honjo et al., 2008).
- According to the "fate" of the exported carbon, which is determined by the particle size distribution (PSD) and is measured by the slope of that distribution (ξ, after Guidi et al., 2009).

A strong and efficient BCP sends more atmospheric CO_2 to the deep ocean via fast-sinking organic particles that escape remineralisation/disaggregation/solubilisation and thus reach deeper depths. The fate of the carbon is intimately associated to the size of the carrier particles, with different sizes having a different life history. In the next Secs. 1.1.5.1–1.1.5.6, I describe the metrics aforementioned (PE_{eff}, T_{eff}, Martin's b, z^*, POC/PIC, bSi/PIC and ξ), which will be used in this thesis to characterise the BCP.

This thesis uses specific terminology to refer to ocean depths, which I want to clarify before moving on. *Surface ocean* is used as a synonym of *euphotic zone* (depth of 1% of incident irradiance), *epipelagic ocean* or the first ~100 m of the water column. *Subsurface ocean* is the water column that lies below the surface ocean and I also refer to it as the *ocean interior*. *Mesopelagic ocean* is that section of the ocean interior between 100–1000 m and is also known as *twilight zone*. Lastly, *bathypelagic ocean* is the section of the ocean interior below the mesopelagic (< 1000 m) and I frequently refer to it as the *deep ocean*.

1.1.5.1 The particle export efficiency (PE_{eff})

Particle export efficiency (PE_{eff})[16], is the fraction of NPP that leaves the base of the euphotic zone (z_0) as POC flux (F),

$$PE_{eff} = \frac{F(z_0)}{NPP}, \tag{1.2}$$

where $F(z_0)$, the POC flux leaving the base of the euphotic zone, is often referred to as export production or POC export flux. Areas with a strong BCP are those with a high PE_{eff}, where most of the new particles produced by NPP events leave the surface ocean intact, as sinking POC. PE_{eff} ranges globally from 0.02 to 0.96 (depending on the reference

[16] PE_{eff} has various synonym terms in the literature: particle export ratio, *pe*-ratio, *p*-ratio or *e*-ratio.

surface ocean depth, Buesseler et al., 2020b), meaning that as little as 2% and up to 96% of the depth-integrated NPP leaves the surface ocean. PE_{eff} is typically low in subtropical and tropical areas because of the high temperatures facilitating the remineralisation of POC in the surface layer (Henson et al., 2012; DeVries and Weber, 2017). On the contrary, PE_{eff} seems to be high in the high-nutrient, low-chlorophyll (HNLC) regions (Southern Ocean, subarctic North Pacific and the eastern equatorial Pacific)[17], where cool advective waters might be tempering the surface ocean, reducing the decomposition of POC.

1.1.5.2 The transfer efficiency (T_{eff})

The transfer efficiency (T_{eff}) is the fraction of POC export flux ($F(z_0)$) that survives attenuation above a reference ocean interior depth, typically 1000 or 2000 m,

$$T_{eff}(z_0 \rightarrow 1000\,\mathrm{m}) = \frac{F(1000)}{F(z_0)}. \tag{1.3}$$

Ocean regions with an efficient BCP are those with a high T_{eff}, where most of the POC exported from the surface ocean reaches the deep ocean intact, having escaped mesopelagic remineralisation (conversion of POC to DIC) and solubilisation/disaggregation (conversion of particulate matter to dissolved matter) during downward transport. According to some modelling studies (Marsay et al., 2015; Weber et al., 2016; DeVries and Weber, 2017), global T_{eff} values seem to correlate positively with NPP and negatively with temperature, following this gradient (from high to low T_{eff}): polar regions > tropics > subtropical gyres. That distribution shows that carbon exported in polar and tropical regions is stored for longer in the deep ocean, constituting that a more efficient BCP.

1.1.5.3 The rate of attenuation of POC flux with depth (Martin's *b*)

An efficient BCP has a low vertical attenuation of POC flux, so that most of the POC export flux reaches the deep ocean. A marine POC flux profile has a decaying shape with depth and it typically shows two sections: in the upper 1000 m of the water column, an exponential decrease, and beneath 1000 m, the decrease becomes asymptotic (Fig. 1.5). That shape implies that most of the attenuation occurs < 1000 m, and what passes that

[17]High-nutrient, low-chlorophyll (HNLC) regions are repleted with surface macronutrients but, paradoxically, phytoplankton biomass (or chlorophyll) remains low all year round. HNLC areas cover ~30% of the world's oceans and consist of three upwelling regions: the open subarctic North Pacific, the eastern equatorial Pacific and the ice-free region of the Southern Ocean. Martin and Fitzwater (1988) suggested that the relatively low productivity in HNLC regions is due to the low availability of micronutrient iron. That happens because HNLC areas are remote offshore provinces far from continental aeolian iron inputs. Iron is an indispensable micronutrient for phytoplankton growth and it is found primarily in the catalytic centres of enzymes. Iron scarcity knocks out larger phytoplankton species, with higher nutrient, and thus iron, requirements. Alongside the low levels of iron, there are high levels of zooplankton grazing that crop small phytoplankton biomass (Frost, 1991). The dual control iron–zooplankton grazing (so-called "ecumenical iron hypothesis" after Morel et al., 1991) keeps phytoplankton, and thus chlorophyll levels, low, and leaves a large stock of surface macronutrients unused.

threshold depth mostly reaches the seafloor. A common practice in biogeochemical models to approach the attenuation of POC flux along depth is by fitting the power-law model proposed by Martin et al. (1987), known as the "Martin curve" (Eq. 1.4 and Fig. 1.5),

$$F(z) = F(z_0)\left(\frac{z}{z_0}\right)^{-b}, \tag{1.4}$$

where $F(z)$ ($\mathrm{mg\,C\,m^{-2}\,d^{-1}}$) is the POC flux at depth z (m), z_0 (m) is the base of the euphotic zone (usually taken to be the fixed depth of 100 m, Buesseler et al., 2020b) and b is the dimensionless exponent of the power-law and corresponds to the slope of the attenuation curve of the POC flux.

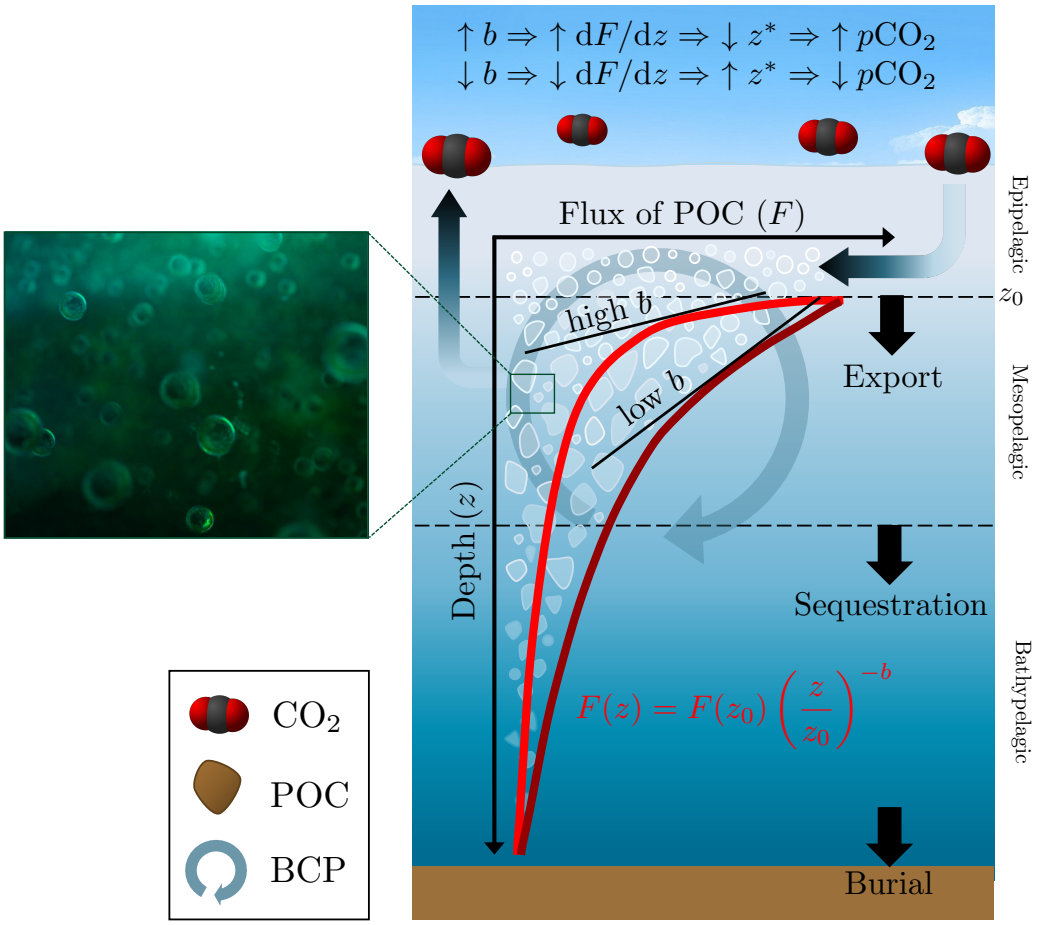


Figure 1.5. Schematic representation of a generic Martin curve and its effects on atmospheric $p\mathrm{CO_2}$ depending on the slope, b, of the curve. Zoomed in is an underwater image of marine particles (image credit: BBC Blue Planet II series). z_0 is the base of the epipelagic, or euphotic, zone; z^* is the remineralisation length scale; POC is particulate organic carbon; BCP is biological carbon pump. Export flux is the name given to the flux of POC below the epipelagic (~100 m, 1% incident irradiance), sequestration flux is the flux of POC below the mesopelagic (~1000 m, nitrate maximum and the greatest extent of vertically migrating zooplankton) and burial is the flux of POC to the sediments (~3500 m). Those three types of fluxes equilibrate at different timescales depending on the latitude: POC is exported from the epipelagic in a matter of weeks to decades, POC is sequestered in a matter of centuries and POC is buried in a millennium/geologic timescale.

The exponent b is known as the "Martin's b" (or just b) and describes how quickly the export flux is attenuated with depth, being essentially a metric for the attenuation rate of POC flux with depth, $\frac{\mathrm{d}F(z)}{\mathrm{d}z}$. Higher values of b mean that most of the remineralisation happens at shallower depths, and thus the attenuation rate of POC flux with depth is high, with potential impacts on atmospheric $p\mathrm{CO}_2$ (Kwon et al., 2009; Wilson et al., 2019). In many studies, b is assumed to have a global average value of 0.858 (after the dataset of Martin et al., 1987) and ranges regionally from 0.20 to 2.0 (Francois et al., 2002; Henson et al., 2012; Marsay et al., 2015). There are two main schools of thought regarding the global ocean spatial patterns of b and $\mathrm{T_{eff}}$ (summary in Table 1.2).

- School of though of Henson et al. (2012): high $\mathrm{T_{eff}}$, along with low b, are expected in the oligotrophic ocean (supported by Buesseler and Boyd, 2009; Guidi et al., 2009; Lam et al., 2011; Lima et al., 2014; Guidi et al., 2015). Henson et al. (2012) argue that higher $\mathrm{T_{eff}}$ along with lower b observed in low-latitude, oligotrophic systems are attributed to a higher recycling ecosystem in the upper ocean due to the tight coupling between phytoplankton and zooplankton growth. Thus, zooplankton quickly graze phytoplankton cells, which never enter a bloom phase (low $\mathrm{PE_{eff}}$). Zooplankton grazing packages phytoplankton cells into compact, fast-sinking faecal pellets (so-called "packaging effect"), which explains the high $\mathrm{T_{eff}}$ to the ocean interior. Phytoplankton communities in oligotrophic ecosystems are dominated by picophytoplankton and coccolithophores. Those cells are less labile than high-latitude dominating diatoms, thus more difficult to digest. Consequently, organic carbon decomposes more slowly and the attenuation of the flux is slower, which explains the low b associated with the high $\mathrm{T_{eff}}$. Opposite to the regime of oligotrophic systems is that of high-latitude, blooming systems (high $\mathrm{PE_{eff}}$). There, the phytoplankton community is dominated by the more labile diatoms (resulting in high b) and there is decoupling between the accelerated phytoplankton growth and zooplankton grazing, where blooming diatoms outpace zooplankton, and a lot of material that had potential to sink fast as faecal pellets is left ungrazed (low $\mathrm{T_{eff}}$). Overlapping with that division oligotrophic/blooming systems, the so-called "ballast effect" could also be at work by favouring higher transfer of carbonate-rich particles (those that contain coccolithophores) vs slower transfer of the less dense opal-rich particles (those that contain diatoms).
- School of thought of Marsay et al. (2015): low $\mathrm{T_{eff}}$, along with high b, are expected in the oligotrophic ocean (supported by Weber et al., 2016; DeVries and Weber, 2017; Cram et al., 2018). Marsay et al. (2015) and supporting studies argue that the lower b (or, conversely, higher $\mathrm{T_{eff}}$) observed in high-latitude and equatorial upwelling regions is due to the lower temperatures. In those colder regions, heterotrophic respiration

in the upper water column is lower, which attenuates POC more slowly (low b). At the same time, the larger size of the exported particles (large phytoplankton cells) contributes to a stronger transfer to deeper depths (high T_{eff}). In subtropical, oligotrophic systems, the reverse pattern is observed: higher respiration rates due to warmer temperatures result in higher b values, and the smaller size of the exported material results in a lower T_{eff}.

Table 1.2. Henson et al. (2012) vs Marsay et al. (2015) paradigms on the downward transfer of POC flux as characterised by b, z^*, PE_{eff} and T_{eff}. Oligotrophic refers to the nutrient-poor regions of the ocean, i.e., the subtropical and tropical open ocean, and eutrophic refers to the nutrient-rich regions, i.e., high-latitude regions and the upwelling equatorial Pacific. The size of picophytoplankton is 0.2–2 μm in diameter, nanophytoplankton (coccolitophores) are 2–20 μm and microphytoplankton (diatoms) are 20–200 μm.

	Henson et al. (2012)		Marsay et al. (2015)	
	Oligotrophic	Eutrophic	Oligotrophic	Eutrophic
Phytoplankton composition	Nano/pico- sized, $CaCO_3$-rich	Micro- sized, opal-rich	Nano/pico- sized, $CaCO_3$-rich	Micro- sized, opal-rich
b	↓	↑	↑	↓
z^*	↑	↓	↓	↑
PE_{eff}	↓	↑	↓	↑
T_{eff}	↑	↓	↓	↑

Two critiques can be made to the former two positions. Firstly, Henson et al. (2012) and Marsay et al. (2015) paradigms are based on curve-fitting procedures of average data. Therefore, they do not capture the seasonal differences of of the ecosystems, where intra-annual variability in a given location can be larger than regional differences (Buesseler and Boyd, 2009). Secondly, Henson et al. (2012) and Marsay et al. (2015) studies are held on the depth range over which POC flux observations have been made, which is the upper 2000 m for Henson et al. (2012) vs the upper 500 m for Marsay et al. (2015). Reliance on deep observations (2000 m) neglects the upper water column, where most of the respiration occurs, but reliance on shallower measurements (500 m) misses the attenuation mechanisms that occur deeper in the water column, like zooplankton-mediated disaggregation (Briggs et al., 2020). Therefore, one could argue that Henson et al. (2012) and Marsay et al. (2015) miss the mechanisms that control POC flux attenuation, and what they capture is just the plain attenuation of POC flux at two different water column sections.

Buesseler and Boyd (2009) developed another body of theory to highlight the role of the seasonality of the ecosystem in the transfer of POC to the deep ocean, as illustrated in Fig. 1.6. As can be seen, the strength and efficiency of the BCP change for a given location throughout the year (KIWI, OSP, K2). The seasonality (the composition of

phytoplankton species in terms of their size), the bloom stage, and the nature of particle degradation (solubilisation, respiration and fragmentation by zooplankton) determine the overall strength and efficiency of the BCP and its regional variability. Buesseler and Boyd (2009) also challenge the ballast hypothesis. They argue that the site with the strongest and most efficient BCP, the northern North Atlantic (represented by NABE in Fig. 1.6), has a pump driven by diatoms, and thus dominated by opal, instead of $CaCO_3$, as would be expected by the ballast hypothesis (Klaas and Archer, 2002).

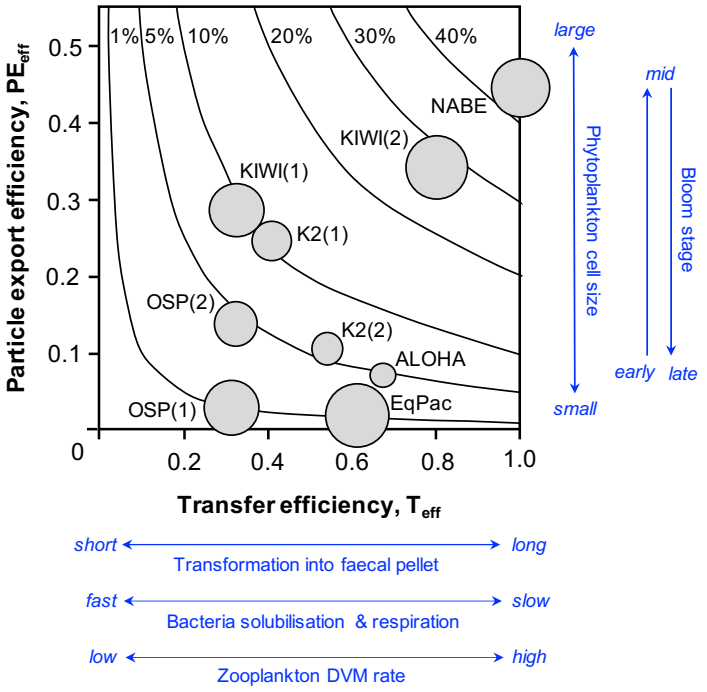


Figure 1.6. Buesseler and Boyd (2009)'s empirical relationship between transfer efficiency, T_{eff}, and particle export efficiency, PE_{eff}. Low T_{eff} are indicative of high attenuation of POC flux in the mesopelagic and low PE_{eff} are indicative of high attenuation of POC flux in the euphotic zone. The circle sizes are proportional to NPP in the ocean sites, which are located in the NE Atlantic (NABE), Antarctic (KIWI), NW Pacific (K2), subarctic NE Pacific (OSP), subtropical Pacific (ALOHA) and tropical Pacific (EqPac). A number is added in between parentheses for a site that has more than one sample represented. The contour lines from 1 to 40% show the fraction of NPP that reaches a depth 100 m below the euphotic zone. Source: redrawn from Buesseler and Boyd (2009).

The Martin curve (Eq. 1.4), despite originally being derived empirically to a local dataset of few[18] sediment traps in the oligotrophic NE Pacific, is of widespread use because it offers a simple relationship between POC flux and depth. Thus, it has been assumed to have global application and has been adopted as the gold standard equation in 3-D models of the marine carbon cycle (e.g., Gnanadesikan et al., 2004; Najjar et al., 2007; DeVries et al., 2012; Bernardello et al., 2014; Maerz et al., 2020; Omand et al., 2020) and observational studies (e.g., Lampitt and Antia, 1997; Francois et al., 2002; Honjo et al., 2008; Henson et al., 2012; Guidi et al., 2015; Marsay et al., 2015) to calculate subsurface POC fluxes and thus obtain distribution of particle fluxes, remineralised nutrients and respiration rates in the water column. Nevertheless, it is possible to find a mechanistic

[18] A total of 87 flux observations in the upper 2000 m of the water column.

(i.e., from first principles) derivation of the Martin curve using the advection-diffusion-reaction equation of a tracer (see Appendix A). It can be shown that b can be derived assuming steady state and that particle sinking velocity increases linearly with depth –a reasonable assumption considering that, as particles sink, they accumulate more mineral materials that get sorbed into them and add ballast. The faster the particle sinks, the more chances to escape remineralisation by bacteria or zooplankton. Thus, a low b (low attenuation of POC flux at the subsurface ocean) could be achieved either by a low remineralisation rate or by a high particle sinking velocity.

But the Martin curve has also had its detractors. Lutz et al. (2002) and Armstrong et al. (2002) were the first to disapprove of it because it had been derived from not too deep ($>$ 2000 m) sediment traps, thereby underestimating the flux in the deep ocean. Lutz et al. (2002) proposed an exponential fit for the attenuation of POC flux with depth (see Sec. 1.1.5.4 below) and Armstrong et al. (2002) proposed a "ballast model" where POC flux is modelled alongside ballast mineral flux. In any case, the interest of oceanographers in empirical fits to the relationship POC flux vs depth dates from the beginning of the 1980s, before the seminal work of Martin et al. (1987) and, as a result of that, today we have a range of functions modelling the attenuation of POC flux with depth: from power-law (e.g., Betzer et al., 1984; Martin et al., 1987; Pace et al., 1987; Antia et al., 2001) and exponential equations (e.g., Lutz et al., 2002; Boyd and Trull, 2007), to rational equations (e.g., Suess, 1980) and ballast equations (e.g., Armstrong et al., 2002; Francois et al., 2002).

All those models, although empirically derived, have been found to have distinct mechanistic interpretations (see synthesis of Cael and Bisson, 2018), be that sinking velocity is constant with depth (exponential model), that remineralisation is a second-order kinetic process (rational model) or that mineral components add more ballast and protection from bacteria (ballast model). But here is the problem: although all those equations are based on different mechanistic assumptions, they seem to fit equally well a given set of local observations (Gloege et al., 2017; Cael and Bisson, 2018). Consequently, it is not possible to identify characteristic sets of mechanisms controlling POC flux for given locations. More concerning is that when those equations (with an equally good fit locally, but based on different mechanisms) are used in ESMs to make future projections of atmospheric pCO_2, they yield different estimates of pCO_2 (Lauderdale and Cael, 2021). Antia et al. (2001) expressed that the quest for a universal fit of POC flux with depth would prove difficult. Such a general equation would assume some common, universal underlying mechanisms processing the material that rains from the euphotic zone, something that it is well known not to hold true; contrasting regional differences exist in the ocean, for instance in terms of mode of export (bloom-pulsed, eutrophic systems vs more stable, oligotrophic systems, e.g., Lampitt and Antia, 1997) and mineral composition of the particles (opal vs carbonate ocean, e.g., Honjo et al., 2008).

1.1.5.4 The remineralisation length scale (z^*)

The depth by which most of the POC flux (specifically, 63% of it) has already been remineralised in the water column is known as remineralisation length scale (z^*, an *e*-folding length scale). Lutz et al. (2002) derived z^* empirically from an exponential fit to POC flux data as

$$F(z) = F(z_0)\, e^{-(z-z_0)/z^*}. \tag{1.5}$$

Nevertheless, a mechanistic derivation for z^* can be found from the advection-diffusion-reaction equation after assuming constant sinking speed with depth (see Appendix A).

The POC flux parameters z^* and b offer complementary information. An efficient BCP is such with an associated b that is low or, complementarilly, with a z^* that is high. The lower the b, the lower the POC flux attenuation rate over depth implying that, at any depth horizon, less POC has already been remineralised, and that there is more POC intact. Consequently, relatively more material will be remineralised at deeper depths than at shallower depths, thus increasing z^*. A deeper z^* means that CO_2 is returned to poorly ventilated, deep waters, for which it takes more time to outgas to the atmosphere, and therefore increase atmospheric pCO_2. Corollary: a high z^* and a low b correlate with a low atmospheric pCO_2 and, thus, with an efficient biological pump of carbon to the deep ocean (Kwon et al., 2009; Marsay et al., 2015).

1.1.5.5 Ballast minerals and the POC/PIC and bSi/PIC ratios

Since POC is a labile material and is almost neutrally buoyant[19], it must be associated with heavier materials, called ballast, to settle gravitationally in the ocean water column. Two biogenic minerals serve as ballast: biogenic silica (bSi)[20] in diatom frustules (mainly) and radiolarian tests (secondarily), and $CaCO_3$ in coccoliths (mainly), foraminiferal tests and pteropod shells (secondarily). Three seminal papers published in 2002 around the ballast hypothesis (Armstrong et al., 2002; Francois et al., 2002; Klaas and Archer, 2002) advocated that since $CaCO_3$ is denser than bSi (2.9 vs 1.9[21] $\mathrm{g\,cm^{-3}}$), it must be globally the most important facilitator of POC transfer to the deep ocean.

PIC, or the carbon in $CaCO_3$, is primarily produced in the ocean by coccolithophores, which are thought to enjoy good growing conditions in warmer, less seasonal, low-latitude systems. In those systems, PIC might increase T_{eff} to the deep ocean compared to high-latitude systems, where coccolithophores are less abundant (Lam et al., 2011; Henson et al., 2012). But, despite calcification dominates biomineralisation in low latitudes, PIC

[19] POC density, 1.06 $\mathrm{g\,cm^{-3}}$, is only a little higher than that of seawater, 1.03 $\mathrm{g\,cm^{-3}}$.
[20] Also termed biogenic opal, or just opal.
[21] Density corresponding to hydrated amorphous silica, $SiO_2 \cdot 0.4\,H_2O$.

flux does not correlate with POC flux at export depth (Lam et al., 2015). Correlation is only found in sediment traps > 1000 m, and the correlation is higher for $CaCO_3$ than for bSi fluxes (Klaas and Archer, 2002). Although correlation does not mean causation, that observation is the one in which the ballast hypothesis is supported. According to Klaas and Archer (2002), bSi is produced by diatoms and, if it were the main ballast material, POC fluxes would present a pattern correlating with diatom abundance, which is not observed in sediment trap data nor modelling studies (Henson et al., 2012; Lam et al., 2015; Weber et al., 2016).

Be PIC or be bSi the main carrier of POC in the sinking flux, ballast minerals affect T_{eff}. Furthermore, their relative contribution to POC flux reveals information on the mechanisms that control the transfer of POC flux to the deep sea, either by increasing the density of particles and, therefore, their sinking velocity (Klaas and Archer, 2002), or from protecting soft organic matter (OM) from particle-attached bacteria (Francois et al., 2002). Honjo et al. (2008) used the POC/PIC ratio along with the bSi/PIC ratio at 2000 m to draw the line between two ocean BCP regimes, with different efficiency considering the main ballasting material: (a) the "silica ocean", where those ratios are ≥ 1 and covers 20% of the world pelagic ocean, including the North Pacific Boreal Gyres and the Antarctic Zone in the Southern Ocean, and (b) the "carbonate ocean", where molar POC/PIC and bSi/PIC flux ratios at 2000 m would be < 1, and applies to the 80% of the ocean. In the silica ocean, few carbonate particles are available to ballast POC and a massive export of opal assumes the ballasting role for $CaCO_3$. The BCP sequesters more carbon in the silica ocean than in the carbonate ocean because the BCP exports lots of POC with little loss of alkalinity from the surface (Honjo et al., 2008).

1.1.5.6 The slope of the particle size distribution (ξ)

In observational studies, particulate fluxes have been typically calculated by normalising the weight of the material collected by the sediment trap to the trap collection area and deployment time. However, larger and fluffy aggregates, which make up the largest fraction by volume of the sinking POC flux, are difficult to collect and study from the trap samples as their fragile nature leaves them reduced into mush (Guidi et al., 2008). Nevertheless, advances in *in situ* imaging techniques in the past 20 years (Guidi et al., 2008; Stemmann and Boss, 2012) have made possible studying aggregate sizes from images and producing particle size distributions (PSDs) that can be used to calculate carbon fluxes as the product of particle concentration per size class × mass × average sinking velocity[22](Guidi et al., 2008; Stemmann et al., 2008).

[22]Particle sinking velocities are difficult to estimate accurately (McDonnell and Buesseler, 2010). A common method to estimate particle sinking velocities is with the help of sediment traps moored in

Particle size data are necessary to better understand the fate of carbon fluxes in the ocean (Guidi et al., 2009; Stemmann and Boss, 2012; Sanders et al., 2014; DeVries and Weber, 2017; Bach et al., 2019) as it appears that the distribution of the mass flux into size classes provides information about the processes forming/consuming the aggregates and the transfer of material along with size classes. For instance, PSDs can reveal processes like faecal pellet production after zooplankton grazing events or TEP-mediated coagulation of particles, all of which increase the average size of a particle by transforming small particles into larger, rapidly sinking ones, thus increasing POC flux to the deep ocean. Conversely, microbial degradation or zooplankton fragmentation decrease the average size of an aggregate, thus reducing POC flux to the deep ocean.

Nowadays, marine particle size data can be readily accessed from the global database EcoTaxa and its Particle module[23]. It is a collaborative platform and repository where particle size measurements come from the UVP (Underwater Vision Profiler, sees in the size range 60–10,000 μm) and the LISST (Laser In Situ Scattering and Transmissometry, sees in the size range 1–500 μm) instruments (see reviews of Stemmann and Boss, 2012; Giering et al., 2020). From the data on particle sizes and abundances, PSDs can be built as

$$N(d) = N_0 \left(\frac{d}{d_0} \right)^{-\xi}, \tag{1.6}$$

where $N(d)$ is number of particles per volume of seawater normalised by the particle size range d (num. $\mathrm{L}^{-1}\,\mu\mathrm{m}^{-1}$), or number spectrum for short; $N(d_0)$ is particle concentration at reference diameter d_0 (commonly taken to be the smallest size category "seen" by the imaging system); ξ is the dimensionless slope of the distribution. The spectral slope has a negative value meaning that the particle count decreases with increasing size following a power-law. Indeed, large particles have a smaller contribution to POC fluxes in terms of total mass and particle count (Sheldon et al., 1972). ξ varies widely regionally, in time and with depth, ranging between 5 and 2 (Guidi et al., 2009), where 5 entails a steeper slope, and indicates a large proportion of smaller particles in the flux, whereas 2 indicates a less higher count of small particles in the flux (but still more abundant than larger ones).

The value of ξ can be typically calculated using UVP data (Guidi et al., 2008; Cram et al., 2018), but it has also been determined from satellite backscattering observations at multiple wavelengths (Kostadinov et al., 2009) and in models (DeVries et al., 2014). It has been found that ξ is steepest (higher) in the subtropical oligotrophic areas, implying a dominance of smaller particles (i.e., picophytoplankton and coccolithophores), whereas ξ is more even in higher latitudes, implying a relatively higher concentration of larger particles (i.e., diatoms) (Kostadinov et al., 2009; DeVries et al., 2014; Cram et al., 2018).

different water depths (Berelson, 2002). However, the method suffers from this (wrong, but necessary) assumption: that the ocean is in steady state and, thus, that flux that leaves some surface depth (where the upper trap is) sinks to some deeper depth (where the lower trap is) at a constant speed, vertically (as in a funnel) and not being affected by sinks (remineralisation losses, advection) nor sources (lateral influxes from exogenous material) (Berelson, 2002; Giering et al., 2017). Nothing far from reality.

[23] http://ecotaxa.obs-vlfr.fr/part/

1.2 Research problem

The observational community has not agreed yet on a common spatiotemporal framework for sampling POC flux. In the meantime, the modelling community is still coming to terms as to what is the best mathematical treatment for the rate of attenuation of POC flux with depth. Those two aspects combined have created sound incongruities in published results of the BCP's magnitude and efficiency. To cite some.

- Prescription of a spatially uniform b in many global biogeochemical models –e.g., b = 0.97 in Kwon and Primeau (2008), b = 0.86 in Bernardello et al. (2014) and the CMIP2 (Najjar et al., 2007), b = 0.82 in DeVries et al. (2012) or b = 0.40 in Aumont et al. (2015). In reality, it is well documented that b is highly variable in space (trophic region) and time (season), by up to an order of magnitude global wide (Henson et al., 2012; Guidi et al., 2015; Weber et al., 2016).
- Conflicting global patterns of b and T_{eff} (i.e., school of thought of Henson et al. (2012) vs school of thought of Marsay et al. (2015)), which arise due to the use of observational data sets of POC fluxes collected in different sections of the water column, different moments of the year and with different technologies[24]. The variability in calculated b introduced by differing sampling times and depths is called *extrinsic variability* and comes from: (i) sampling campaigns carried before or after the bloom period and thus collecting data with different degrees of representativity of the annual cycle (Buesseler and Boyd, 2009; Lam et al., 2011), (ii) the inconsistent definition of the deep ocean as the 500, 1000 or the 2000 m depth (Marsay et al., 2015), (iii) different treatment of the definition of the reference depth used to calculate export flux in observational studies (which use the euphotic zone depth) and modelling studies (which use the 100 m depth to facilitate the analyses) (Bisson et al., 2018; Buesseler et al., 2020b), or (iv) the placement of the deep trap with regard to the surface trap for a strategy of 1-D sampling of the water column, avoiding lateral additions of materials (Siegel and Deuser, 1997). On the other hand, *intrinsic variability* is inherent to the use of different observational technologies, from sediment traps and radionuclide methods to emerging autonomous optical techniques and satellite data products. Those differ in how well they capture different features of the ecosystem: (i) resolution of the smallest possible particle seen or filtered, (ii) distinction of living vs non-living material, (iii) maximum possible sampling (or seen) depth or (iv) preservation of the vertical signal of POC flux, avoiding additions from horizontal sources (not in equilibrium). Overall, the extrinsic and

[24] Examples of current *in situ* methods to measure POC flux: (i) gel-filled sediment traps, (ii) marine snow catchers, (iii) UVPs, (iv) underwater cameras like P-cam and LISST-Holo, (v) backscatter sensors on gliders and BGC-Argo floats, (vi) radioactive pairs disequilibrium (^{234}Th and ^{210}Po) and (vii) mesocosms.

intrinsic factors result in undersampling and oversampling biases and yield POC flux observations with a large error associated. Bisson et al. (2018) showed that, locally, observed inter-annual monthly variations could be as large as the long-term (climatological) monthly mean.

- Mismatching temporal scales between the data used to drive global models of the BCP (climatological scales) and the POC flux data sets available to force/validate those models (from punctual, short-term campaigns that miss most of the year's cycle) (e.g., Siegel et al., 2014; DeVries and Weber, 2017). There are only a few long-term sediment trap measurements made over the globe, like those of BATS and HOT. The use of short-term measurements to optimise global models of POC flux, which use climatological scales, further contributes to the huge spread in the estimates of modelled POC export fluxes, ranging 4–11 Gt C yr^{-1} (see Table 1.1).
- Uncertainty as to what should be the exact shape of the particle flux profile (e.g., power-law, exponential, rational). Different shapes are statistically indistinguishable from the power-law Martin curve but carry different implications for flux attenuation and the mechanisms controlling the BCP (Cael and Bisson, 2018).
- Imbalance between the metabolic demand supported by the exported POC flux and the carbon demand of heterotrophic metabolism, which is two to three times higher (Steinberg et al., 2008; Burd et al., 2010; Herndl and Reinthaler, 2013; Giering et al., 2014).

Despite more than 30 years of field and modelling work on the BCP (Doney et al., 2002), oceanographers still do not understand the exact operating mechanisms that relate biogeochemical forcing to the attenuation rate of POC flux with water depth (Boyd and Trull, 2007; Sanders et al., 2016; Siegel et al., 2016; Boyd et al., 2019). Several hypotheses have been already formulated to link biogeochemical forcing to key parts of the BCP –e.g., seawater temperature and O_2 availability to remineralisation rates of organic carbon, phytoplankton community composition to particle ballast, light and nutrient limitation to the amount of exported material, or euphotic zone food-web dynamics to the degree of packaging of the exported particles. Nevertheless, to build a mechanistic theory on the controls of the biological pump of organic carbon to the ocean interior, a new generation of modelling approaches and new observational technologies need to be developed (Bisson et al., 2018; Cael and Bisson, 2018; Brewin et al., 2021).

Although many factors are intertwined in defining the pattern for the attenuation rate of POC flux with depth (i.e., Martin's b), the question can be vastly simplified from the general agreement that b is dependent on the balance between particle sinking velocity and the rate of organic carbon remineralisation (as per derivation of the Martin curve from the advection-diffusion-reaction equation of a tracer, see Appendix A). Thus,

fast-sinking particles, which are likely to reach greater depths before most of their organic carbon is remineralised, and low levels of remineralisation favour the transfer of organic carbon to the deep ocean (Sanders et al., 2014).

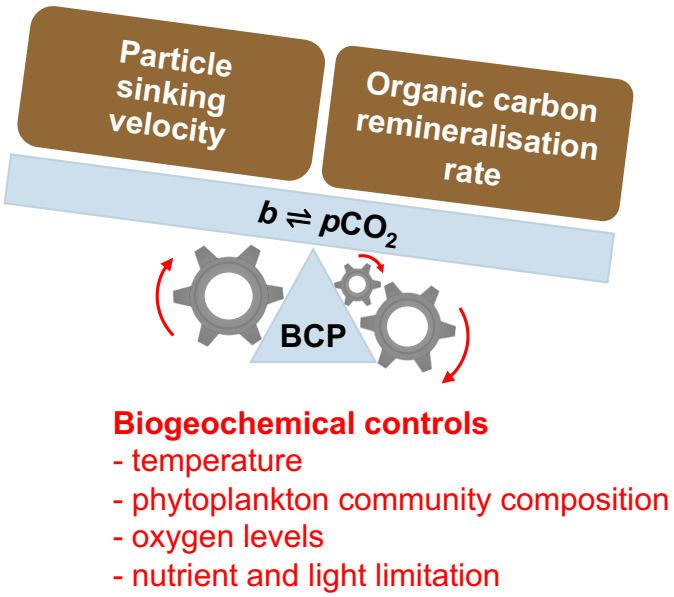


Figure 1.7. Schematic linking the concepts of biological carbon pump (BCP), proposed biogeochemical controls, Martin's b, atmospheric pCO_2, particle sinking velocity and remineralisation rate. Note that the ratio between the particle sinking velocity and the remineralisation rate is the metric remineralisation length scale, z^* (see Appendix A).

Current Eulerian biogeochemical models have proven successful in explaining the global distribution of dissolved nutrients in the ocean. However, they are unsuited for improving our understanding of the controls that affect the attenuation of particulated nutrients, e.g., POC and PIC, which are Lagrangian in nature. Furthermore, the biogeochemical modelling community has been using models based on an obsolete and parameterised view of the BCP (the Martin curve) that needs to be replaced by a mechanistic one that varies in space and time. The new mechanistic approach must put marine particles –instead of dissolved tracers– at the centrepiece of the model, and consider particle chemical composition, microstructure and its transformations (Burd et al., 2010; Sanders et al., 2014; Boyd et al., 2019). This thesis precisely addresses the development of a Lagrangian model of marine particles –the mechanistic units of the BCP– to provide a Lagrangian-based parameterisation for the Martin's b. It is expected that this method can contribute to build a body of theory on the biogeochemical factors that control particle sinking velocity and remineralisation rates, the two parameters that define b.

1.2.1 Marine particles, the mechanistic units of the BCP

To date, there are still no global biogeochemical models that approach the BCP as an ensemble of Lagrangian, discrete particles that are tracked along the water column, since they are created until they disappear. Current formulations are Eulerian and thus model particulate matter as average concentrations in water volumes (that is discussed

in Chapter 2). Consequently, the attenuation of POC flux with depth is simulated coarsely, not considering that the particles responsible for the bulk of POC range from a myriad of sizes (from < 1 μm to > 1 cm) and sinking velocities (< 1 m d^{-1} for small algae to > 100 m d^{-1} for large marine snow, Stemmann and Boss, 2012). In Eulerian models, POC fluxes are calculated from a few predefined particle sinking speeds and then propagated to the ocean interior using empirical POC flux vs depth relationships (e.g., Martin's b) that are spatially uniform and lack mechanistic foundations. That has lead to a considerable spread in the calculations of particle fluxes to the ocean interior (as shown in Table 1.1).

That scenario conflicts with the widespread availability of supercomputers as well as improved mathematical realisations to efficiently model millions of particles (Shima et al., 2009). Nowadays, it is already possible to develop models that consider the millions of particles that are in the ocean using Lagrangian descriptions and thus simulating the BCP from a bottom-up approach (e.g., Jokulsdottir and Archer, 2016). To advance our knowledge on the BCP-mediated sequestration of carbon in the ocean, marine particles not only need to be put at the core of the mechanistic descriptions of the BCP, but need to be related to the ecosystem that produces and transforms them correctly (Sanders et al., 2016). The details of that relationship are not understood yet because *in situ* data are difficult to collect, oceanographic fieldwork campaigns are expensive and the ocean is a vast and variable system, altogether making *in situ* data scarce and patchy (Boyd and Trull, 2007; Boyd, 2015).

Marine particles[25] are at the core of the BCP and constitute its mechanistic units. Particle attributes record the biogeochemical transformations that particles have undergone throughout their life history, since they were created at the surface ocean until they disappear as remineralisation, disaggregation and solubilisation attack the particulate carbon and return it back into its dissolved form. To which extent each of those mechanisms consume the particulate carbon and how their intensity is affected by the surrounding biogeochemical conditions of temperature, O_2 levels or phytoplankton community composition is a question that could be resolved by paying attention to how the particle attribute "size" changes as the particle sinks (Weber et al., 2016; DeVries and Weber, 2017). Size is a central property in ecology because it is a structuring dimension along which ecosystem dynamics of various types can be projected (Stemmann and Boss, 2012, and references therein): particle sinking velocity, amount of packaged nutrients, particle-to-particle interactions (particle coagulation rates), particle-to-zooplankton encounters (zooplankton encounter rates), extent of colonisation by bacteria, scavenging rates, etc. Thereby, a model capable of resolving the size of the individual marine particles could provide an improved perspective on the main mechanisms that control the temporal and spatial variations of the downward flux of POC (Weber et al., 2016; DeVries and Weber, 2017).

[25] By definition, a particle is a solid body with a diameter > 0.2 μm.

1.2.2 My approach: a mechanistic model of marine particles

Jokulsdottir and Archer (2016) built a 1-D model of Lagrangian marine particles called SLAMS (Stochastic, Lagrangian Aggregate Model of Sinking particles), which provided insights into how the physicochemical attributes of aggregates that coagulate and fragment in a streamlined BCP had dependencies on sea surface temperature and the magnitude and seasonality of primary production. SLAMS explicitly represented fluxes of particulate materials allowing direct comparison to a growing number of such observations (e.g., Lutz et al., 2002; Guidi et al., 2009). Mathematically, SLAMS is characterised by three aspects: (i) it uses a Lagrangian framework, which tracks individual, discrete particles as they move through the model space and calculates the physicochemical attributes of the particles (e.g., mass, size, porosity, sinking velocity) at each point in space; (ii) it is a stochastic model, which uses probabilities to make decisions about the outcomes of processes highly uncertain, and uses the Monte Carlo method to sample the probability distributions of such processes; (iii) the model is solved algebraically to focus on the value of a model with a good compromise between time and memory costs.

The model presented in this thesis, SLAMS-2.0, expands the work of Jokulsdottir and Archer (2016) by (i) improving the representation of phyto- and zooplankton types and food web processes, (ii) adding more spatially varying forcing variables that control the BCP (temperature, nutrients, oxygen, irradiance, mixed layer depth) and (iii) extending the 1-D model grid to a global array.

1.3 Aim, objectives and research questions

The overall goal of this thesis is to improve our understanding of the mechanisms that control the global pattern of Martin's b, which characterises the attenuation of the sinking flux of POC.

The specific objectives that will facilitate the achievement of that goal are:

- to develop a model of Lagrangian marine particles in the context of the BCP (SLAMS-2.0) with simple dynamical rules and that is modular in design, creating a virtual environment where particle processes and biogeochemical forcing can be easily switched on and off to explore their effect on the pool of particles;
- to calibrate the model using available flux observations of POC, PIC and bSi for contrasting ocean biogeochemical sites; and
- to apply the model to explore the biogeochemical (top-down) controls and particle (bottom-up) controls on Martin's b and T_{eff} and help reconcile the two canonical views on their global patterns, i.e., Henson et al. (2012) vs Marsay et al. (2015).

This thesis is part of the COMICS project (NE/M020835/2) and contributes to accomplishing the COMICS modelling objective: to extract from the ensemble particle dynamics of SLAMS-2.0 a first-order relationship of the most relevant controls that characterise the attenuation of POC flux with depth. Given the computational expense of incorporating Lagrangian descriptions of particle dynamics in current ESMs, and the need to overcome the use of prescribed attenuation rates (the Martin's b), such a simple functional form could be easily implemented in ESMs to govern the vertical distribution of organic carbon in the ocean as a function of local forcing and improve current estimates of the biological sequestration of POC.

SLAMS-2.0 is used to answer the following research questions.

- How relevant is the presumed dominant control of temperature on b through its effects on the dynamics of marine particles?
- How relevant is the control of diatoms vs coccolithophores abundance (ballast hypothesis) on the transfer of POC to the deep ocean?
- What are the physicochemical attributes that characterise the particles that reach the deep ocean?
- What is the role of coagulation and zooplankton repackaging in the attenuation of POC flux?
- How relevant are surface water conditions (primary production and phytoplankton community composition) with respect to mesopelagic conditions (particle transformation processes) in characterising particle fluxes?

2

Introduction to my methodological approach

Contents

2.1 A mechanistic model

I inherited the first version of my simulation, SLAMS-1.0, from Jokulsdottir and Archer (2016). They built a particle-based simulation that routinely tracks biogenic marine aggregates along the 1-D vertical scale of a motionless water column as they sink and engage in particle dynamics, chemical and biological processes. The representation of the biology and the physicochemical environment in SLAMS-1.0 was still rudimentary and did not offer the whole picture of the proposed mechanisms controlling the BCP (see Fig. 1.7). The first task of my thesis project was to continue developing SLAMS-1.0 into a more sophisticated, or mechanised, simulation to answer the research questions laid out in Sec. 1.3.

I have developed SLAMS-2.0 to understand the particle attributes, the particle dynamics and the biogeochemical factors that control the transfer of POC flux from the

surface to the deep ocean. It is thus an *explanatory model* of the causal relationship between particle and biogeochemical factors and their effects on the sinking flux of POC. An explanatory model is a bridge built between the hypotheses that the researcher wants to test and the observable measurements that the researcher has collected from the literature. The model is the construct (the model equations, parameters and the model forcing) that reflects the hypotheses to be tested. If it successfully represents the observations, it is often assumed to inherently possess predictive power and, therefore, be used to predict new, future or past observations, becoming a *predictive model*.

Models can be classified as *empirical* or *mechanistic* depending on how much we know about the laws governing the causal factors. A mechanistic model of marine particles would use equations for primary production, particle dynamics or zooplankton grazing based on biogeochemical principles (laws). In the face of incomplete mechanistic knowledge, one would build an empirical model, where equations are parameterised and bear no relationship to underlying mechanisms. Published models of ocean biogeochemistry often claim to be mechanistic when they are semi-empirical, or somewhere between empirical and mechanistic. Whereas physical models of the ocean are derived from first principles, to date there is still no holistic mechanistic view of the BCP, with all its components fully mechanised, but only some, like the aggregation of sinking particles (Armstrong et al., 2002; DeVries et al., 2014; Jokulsdottir and Archer, 2016; Maerz et al., 2020; Omand et al., 2020). SLAMS-2.0 is a mechanistic description of marine particle dynamics in the context of the BCP, where planktonic processes and particle interactions have been mechanised. The degree of mechanisation stops at the particle level (it does not enter the enzymatic level), and I do not have a food web structure self-regulated by evolutionary mechanisms of competition and exclusion, like the trait-based model of emergent phytoplankton communities of Follows et al. (2007) (the Darwin model).

The degree of mechanisation of a model is linked to the degree of complexity needed in a model. How much complexity is needed for the model to have explanatory and descriptive power is a topic that has been hotly contested in the literature (e.g., Flynn, 2003; Anderson, 2005; Friedrichs et al., 2007; Kriest et al., 2010; Ward et al., 2013, 2014; Schartau et al., 2017). The answer is an open research question. It depends on various aspects, namely: (i) the purpose of the model, that is, the hypothesis that needs to be tested; (ii) how much resolution the research question needs, that is, the degree of biogeochemical detail; (iii) how much data we have available to justify the choices made in the model, so that the choices are not unfounded; and (iv) the portability of the model to broader geographic scales, which requires more flexibility in model's formulations.

2.2 SLAMS's approach to simulating marine particles

2.2.1 Lagrangian modelling

Lagrangian, and its analogous Eulerian, are two terms that come from the realm of fluid dynamics and refer to two ways to sample the water. Whereas Eulerian measurements are made over time from a fixed point in space, like bottom-moored sediment traps, Lagrangian measurements are made by following a parcel of water as it moves over time and space, using instruments like freely-drifting neutrally buoyant sediment traps. Likewise, biogeochemical ocean models can be classified as Eulerian or Lagrangian according to how the flux of particles is measured (Fig. 2.1).

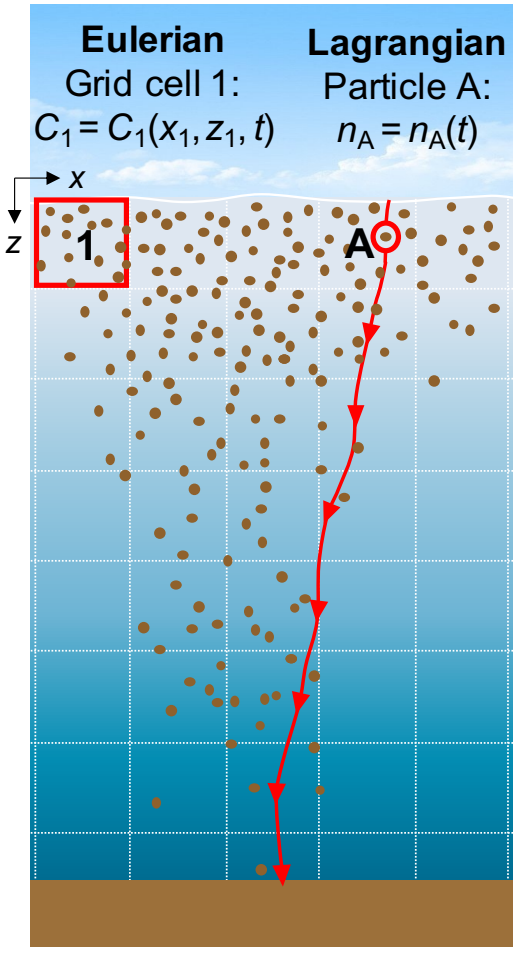


Figure 2.1. Schematic of Lagrangian vs Eulerian modelling perspectives of marine particles. Each squared cell represents a model grid cell. Lots of Lagrangian particles move through a fixed Eulerian spatial grid. In the Eulerian scheme, the concentration of the particulate tracer, C, is measured at each grid cell, which implicitly assumes that the tracer is homogeneously distributed across the cell and each particle contributing to the concentration has the same physicochemical characteristics. In the Lagrangian scheme, each particle in a cell is individually tracked. The concentration of a specific material in a cell results from the ensemble of particles, each carrying n moles of the tracer.

Eulerian biogeochemical models simulate the concentration of particulate tracers over time and space at each model grid box (or cell) according to local environmental conditions. Those tracers, like POC flux or phytoplankton biomass, are assumed to be evenly distributed in the grid box and have the same physicochemical properties throughout, which is the same as assuming that particles are all identical to each other in the modelled water volume. Therefore, particles are modelled for each grid box as a

continuum property of seawater by calculating the sources-minus-sinks (SMS) term of the advection-diffusion-reaction equation of a tracer (Eq. 2.1).

$$\frac{\partial \mathrm{C}}{\partial t} = \nabla \cdot (\boldsymbol{u}\mathrm{C}) + \nabla \cdot (\kappa \nabla \mathrm{C}) + \mathrm{SMS}(\mathrm{C}), \tag{2.1}$$

where C is a tracer, or substance; $\boldsymbol{u}$ is 3-D velocity field; κ is the turbulent diffusion coefficient, and SMS is the reaction sources-minus-sinks term.

In reality, there are not two particles alike in the ocean. Each particle has undergone a unique life history that has left a characteristic imprint of properties such as material composition, density, size and porosity. Thus, the Eulerian way of averaging entire populations of dissimilar particles using a mean value and a unique variable (e.g., particle size) leads to loss of descriptive power of the particle population that makes up the flux of POC. The Lagrangian approach, instead, considers the entire population of individual particles as well as all particle variables (or attributes) of interest, like size, porosity, density, stickiness or sinking velocity. A Lagrangian model tracks each individual, discrete particle as it moves throughout the model grid space, since it is created (e.g., as a phytoplankton cell after primary production events) until it disappears (e.g., by remineralisation). Consequently, the Lagrangian framework of SLAMS-2.0 simulates the chemical and microstructural attributes of the individual particles at each point in the grid space as they gravitationally sink and experience changes in their material composition by biogeochemically interacting with their ecosystem (e.g., via coagulation, grazing, bacteria attachment, dissolution below the lysocline, etc.). Each particle in the model space has its associated set of equations describing the particle state. There are as many equations per particle as attributes need to be modelled, where particle attributes are calculated from first principles from the amount of material packed in the particle and a particle shape descriptor. It can be easily seen that the Lagrangian approach becomes computationally expensive when modelling thousands of particle histories. That is the main reason why, to date, there are still no Lagrangian models of the planktonic system coupled to ocean general circulation models (OGCMs). With billions of particles in the ocean, a Lagrangian treatment of them needs to use computational tricks to achieve computational economy. In the case of SLAMS-2.0, that trick is the super droplet method of Shima et al. (2009), where clone particles are grouped into clusters whose number keeps unchanged during coagulation (see Sec. 2.3.1.2 below).

The Lagrangian framework is regarded because particle macro properties such as average sinking velocity, POC export flux or the rate of attenuation of POC flux with depth (Martin's b) emerge from the dynamics of the ensemble. Contrarily, in the Eulerian framework those are parameterised. Usually, there are no more than a few prescribed particle sinking velocities (e.g., Yool et al., 2013), or those can be altered with depth using

depth-dependent functions, such as the Martin curve (Gloege et al., 2017). As a result, either the diversity of marine particles is reduced to a few parameterised particle classes, or macro properties like Martin's b need to be imposed to calculate derived variables and cannot be studied causatively to understand what mechanisms control them.

2.2.2 Governing equation in a particle-oriented model

The governing equation of a Lagrangian, particle-oriented model is a population balance equation called the Smoluchowski coagulation equation (Smoluchowski, 1916) (Eq. 2.2). A population balance equation describes how changes in the particle state (size, velocity, chemical composition) propagate through the population of particles in time and space as particles collide and coagulate, fragment, are digested, etc. It essentially describes the spatiotemporal evolution of the population of particles. The rationale for Eq. 2.2 is as follows. Consider a dispersion of initially identical marine particles (i.e., primary particles) where, after some time of coagulation, appear aggregates of various sizes and with different particle concentrations associated to these sizes (e.g., N_i aggregates of size r_i, N_j aggregates of size r_j...),

$$\begin{aligned}\frac{\partial N(r_k,t)}{\partial t} = &\frac{\alpha}{2}\sum_{i+j=k}\beta(r_i,r_j)N(r_i,t)N(r_j,t)\\ &-\alpha N(r_k,t)\sum_{i=1}^{\infty}\beta(r_k,r_i)N(r_i,t)\\ &-\frac{w(r_k)}{z}N(r_k,t)\\ &+\mathrm{SMS}(r_k,t).\end{aligned} \tag{2.2}$$

where: $N(r_k,t)$ = particle number concentration (num. m^{-3}) of an aggregate of radius r_k at time t, such that $r_k = r_i + r_j$

β = coagulation kernel, i.e., number of particle pair collisions occurring per unit volume and unit time ($\mathrm{m}^3\,\mathrm{s}^{-1}$)

α = stickiness coefficient (unit-less, between 0 and 1)

$w(r_k)$ = sinking velocity of a particle of size r_k

z = thickness of the water layer (or grid box)

SMS = biological sources-minus-sinks term

$N(r_k,t)$ is the time evolution of the average number concentration of particles of size r_k at a given time t and gets plotted as a particle size distribution (PSD). The first term of Eq. 2.2 is the birth of particles of size r_k as a result of coagulation of particles of sizes r_i and r_j and the second term, preceded by a negative sign, is the death of particles of size r_k when coagulating with any other particles. Physical coagulation is the primary source transferring particles from smaller to larger size classes in the surface ocean (Jackson

and Burd, 2015). It is a two-step process involving (1) collision of particles brought together by transport mechanisms (the coagulation kernel β denotes the intensity of collision) and (2) coalescence of the particles if they remain stuck to each other after collision (the stickiness coefficient, α, denotes the intensity of coalescence). The third term in Eq. 2.2 is the gravitational sinking of particles, a loss term that happens when big, fast-sinking particles (created from smaller ones) escape the depth interface, leaving behind the smaller ones. The fourth term, the SMS, represents the addition or loss of particles due to food web processes, like the production of faecal pellets (source), which competes with coagulation in the formation of large aggregates (Jackson, 2001; Stemmann et al., 2004), or remineralisation (sink), which degrades particles. Thus, there are three processes in Eq. 2.2[1] competing in the particle distribution: coagulation, gravitational sinking and biological sources and sinks of particles. The Smoluchowski equation shows that coagulation is a density-dependent process (notice how the coagulation rates are multiplied by $N(r_k, t)$), which implies that it unfolds quickly, for instance during times of high phytoplankton abundance.

The coagulation submodel of SLAMS-2.0 (after Jackson, 2001) is based on the Smoluchowski equation, which is no more than a particle size distribution. But instead of being solved numerically or analytically, the solution of the distribution is an emergent property of the simulated Lagrangian particle dynamics.

2.2.3 Stochastic modelling and the Monte Carlo method

In stochastic simulations, the outcomes of processes that are highly variable are treated probabilistically. For instance, in SLAMS-2.0, one of several processes that are modelled stochastically is the seeding of phytoplankton functional types (PFTs), which has four possible outcomes: (a) coccolithophores, (b) diatoms, (c) picophytoplankton and (d) dinoflagellates (Fig. 2.2). If PFT seeding was treated deterministically, the outcome (*a*, *b*, *c* or *d*) would be decided upon solving a phytoplankton growth-limited equation, so that the PFT displaying the highest growth rate at that specific time step would be chosen. In a stochastic, or probabilistic, treatment, the four possible solutions of the growth-limited equation are first converted into relative frequencies[2] and next sampled

[1] Eq. 2.2 is a modified version of the original Smoluchowski equation (1916), which has been expanded to incorporate particle biogeochemical processes (i.e., SMS) after Burd and Jackson (1997); Jackson (2001); Burd and Jackson (2009).

[2] The terms *relative frequency*, *frequency* and *probability* mean different things but sometimes are used interchangeably. Frequency is how many times an outcome DO has occurred. It is a term reserved for observed data, e.g., tossing a coin ten times and getting six heads gives a frequency of six (an absolute number). Relative frequency is how often an outcome DO occurs. It is the absolute frequency divided by the number of trials, e.g., 6/10 = 0.6 in the previous example. Probability is how many times we THINK an event will occur. Probability is a term used for expected data, e.g., the probability of getting a head from each coin flip is the number of outcomes we are interested divided by the number of possible outcomes, 1/2 = 0.5. The probability of an outcome A in an experiment is the limit of the relative frequency of the outcome when the experiment is repeated an infinite number of times.

lots of times based on computer-generated random numbers. Outcomes that are more probable (i.e., have a higher relative frequency) will be sampled more times. This random sampling of the PFT frequency distribution is repeated many times over for each model time step, enough times to get a decent sampling of the PFT seeding response. By doing that, random noise is injected into the model solution and the variability of the system can be captured. Thus, a given set of initial conditions will never yield the exact same results, but an average outcome with an associated error. This way, the outcome of a process modelled stochastically becomes statistical rather than exact, a convenient way to capture the fluctuations that occur in natural systems. Conventional ocean biogeochemical models are deterministic, thus cannot simulate the variability in natural systems. Stochastic models do recognise it and therefore are likely to be a more accurate representation of nature (Gillespie, 1975).

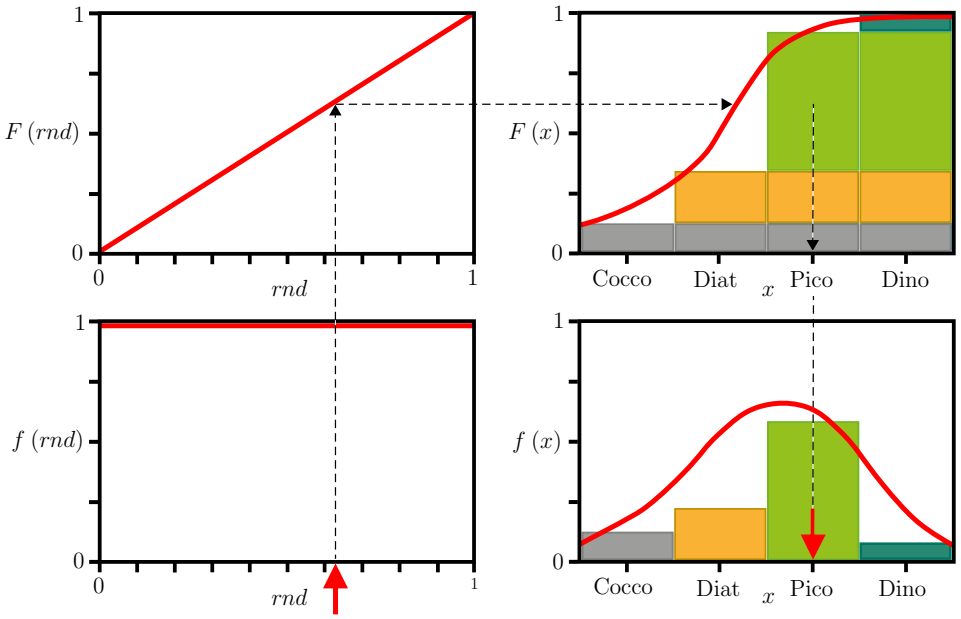


Figure 2.2. The Monte Carlo method applied to sampling phytoplankton functional types (PFTs) in SLAMS-2.0. *rnd* is a random number uniformly distributed between 0 and 1; x is the random variable "phytoplankton functional type", with its four possible outcomes as Cocco (coccolithophores), Diat (diatoms), Pico (picophytoplankton) and Dino (dinoflagellates); f is the relative frequency distribution function; F is the cumulative distribution function. Consider a *rnd* drawn by the computer from a uniform distribution between 0 and 1 (red arrow pointing to *rnd*). This random number is used to select a particular PFT, x_i, from the cumulative distribution function $F(x)$. Each $F(x_i)$ are the relative frequencies assigned to each PFT according to the biogeochemical laws that control phytoplankton growth (based on limitation by light, temperature and nutrient), and therefore $F(x_i)$ are values known by the researcher. By randomly sampling $F(x)$ many times, we get the $f(x)$ of the PFTs.

The Monte Carlo method (MCM)[3] is the name given to the random sampling of the probabilistically weighted response of a process. Besides PFT seeding, there are

[3] After the Monte Carlo Casino in Monaco.

other processes in SLAMS-2.0 that, due to their intrinsic large stochastic component, also receive a stochastic treatment via the MCM: selection of particle pairs for collision, execution of coagulation or particle encounter by zooplankton. Realise that in all processes modelled using the MCM, the computer-generated random number operates on the possible outcomes of a variable that has been generated based on biogeochemical laws. For instance, in the collision of particles, the MCM operates on the mechanistic collision-coalescence formulation of Jackson (2001).

2.3 A review of particle-oriented biogeochemical models

2.3.1 Smoluchowski-centered particle models

Particle coagulation is a phenomenon not unique to marine particles. It has been widely studied in the fields of atmospheric physics (formation of snow flakes from ice crystals, formation of clouds from water droplets, coagulation of pollutants), geology (sedimentology, formation of basins, CO_2 storage by soil particles), planetary sciences (protoplanet formation) or chemical engineering (compound formation in bioreactors, polymerisation). The subject is also recently attracting lots of attention in the study of marine microplastics to clean seawater (e.g., Van Sebille et al., 2020). Smoluchowski (1916) was the first to put together the processes of birth of particles of different size classes due to coagulation. His equation was later modified to be applied in marine systems with the addition of gravitational sinking and the biological input of new sources of particles (Eq. 2.2).

Ocean biogeochemical models attempting to resolve the distribution of marine particulate material into various particle size classes need to solve the Smoluchowski equation or parameterise it. Fig. 2.3 provides an overview of the types of methods that common published models in the ocean biogeochemical modelling (OBM) literature use to solve the Smoluchowski equation, with branches into deterministic (inherently Eulerian) and stochastic (inherently Lagrangian) modelling methods. The most widespread approach (Gehlen et al., 2006; Karakaş et al., 2009; DeVries et al., 2014; Cram et al., 2018; Omand et al., 2020) is the deterministic, Eulerian framework, which becomes computationally expensive when more than 2–4 particle attributes are being modelled. Per contra, stochastic, Lagrangian frameworks do not have a problem with the number of attributes being modelled. However, they are still an emergent technique in OBM as those have a more complex implementation. To the author's knowledge, to date SLAMS-1.0 (Jokulsdottir and Archer, 2016) is the only published model of Lagrangian marine particles that coagulate, sink and interact in a simplified ocean ecosystem without those particle processes being parameterised. The Lagrangian implementation of particle aggregation and biogeochemical interactions with their ecosystem is still a challenge due to the lack of mechanistic understanding of those processes.

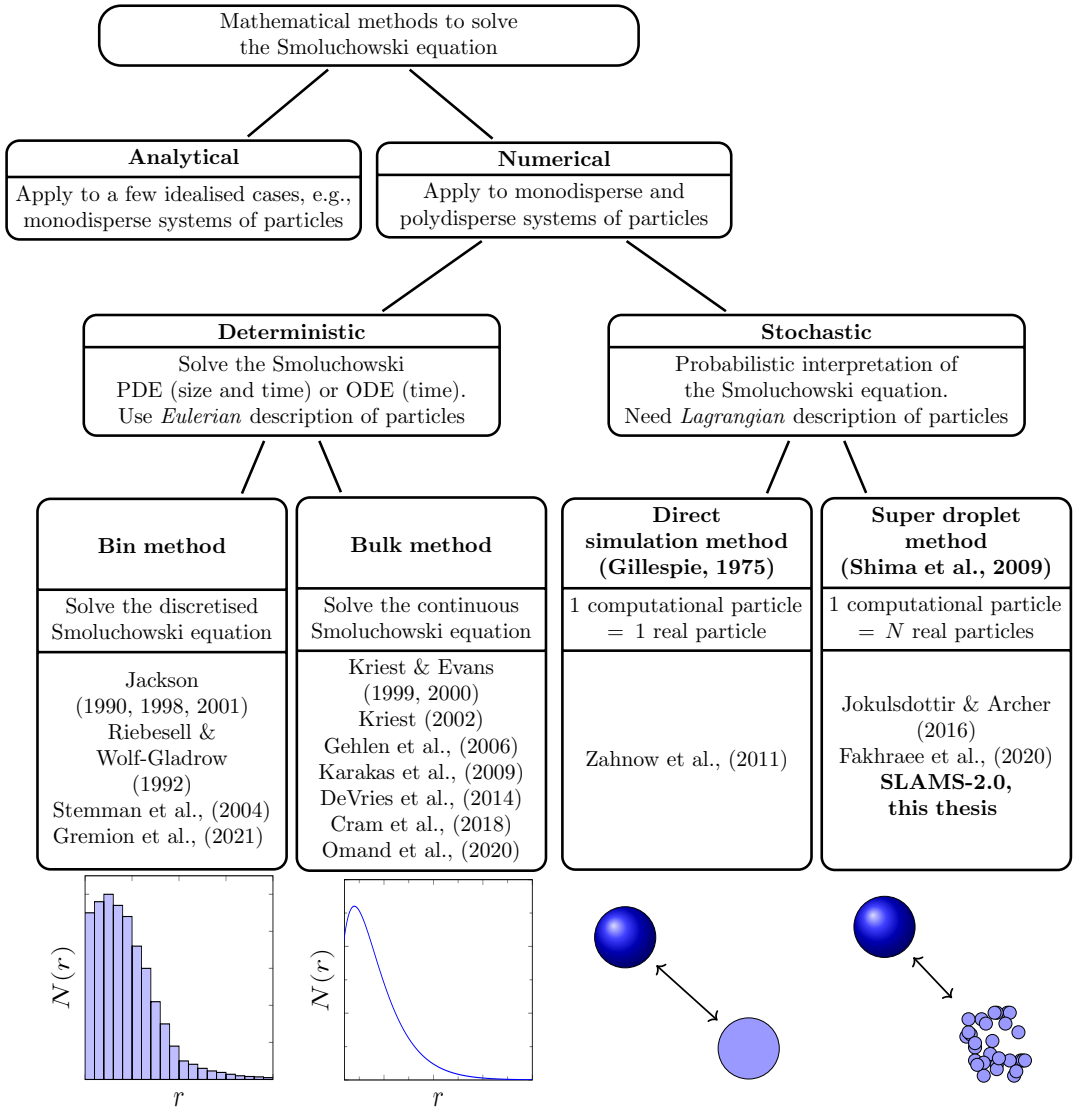


Figure 2.3. Tree diagram with common approaches in the ocean biogeochemical modelling literature to solve the Smoluchowski equation. The Smoluchowski equation models how rapidly the diversity of particles present at some moment in the system is changed. PDE stands for partial differential equation and ODE stands for ordinary differential equation. Stochastic methods are also known as particle-based methods.

2.3.1.1 Eulerian methods

Numerical deterministic (Eulerian) methods have two sub-branches (Fig. 2.3).

- Bin methods, which solve the discretised version of the Smoluchowski equation, which is pieced into particle sizes (bins, or sections). Those are illustrated with the pioneering work of Jackson (1990, 1998, 2001).
- Bulk methods, which solve the continuous version of the Smoluchowski equation. Those are illustrated with the pioneering work of Kriest and Evans (1999, 2000).

Those two methods can easily solve the PSD if the number of particle attributes (e.g., mass, radius, particle number, etc.) is kept low (typically 1–2, Grabowski et al., 2019).

Jackson (1990)'s (J90 hereafter) model has multiple particle size classes (bins, or sections), which are explicitly represented as model variables. These size classes (r) are co-dominant via fractal scaling with mass classes (m), as $m(r) = cr^D$ (where c is a factor of proportionality and D is the fractal power), thus forming a 2-D lattice of variables (i.e., the modelled distribution). The model uses standard coagulation laws (as those described in Burd and Jackson, 2009) to form new particles after the coagulation of two smaller ones. There is a substantial computational cost in resolving the discretised version of the Smoluchowski distribution as there are many variables (as many as particles) for which the aggregation kernels have to be computed. The problem of solving for the continuous evolution of the distribution (a PDE) is transformed into one of solving for all of the bin concentrations using a set of coupled ODEs for each $N(r_k, t)$. A bin method will therefore describe the PSD in great detail (multiple size classes) at the expense of becoming computationally expensive when used in a global biogeochemical model with high spatial resolution, with many grid points for which to solve ODEs.

Kriest and Evans (1999) (KE99 hereafter) describe the PSD using a continuous Smoluchowski distribution. To describe particle state, KE99 use three particle attributes: mass (m) and particle number (N), which are related to the radius of the particles (r) via a power-law relationship, i.e., $n(r) = \frac{\mathrm{d}N}{\mathrm{d}r} = c_N r^{-\epsilon}$ and $M(r) = \frac{\mathrm{d}m}{\mathrm{d}r} = c_m r^{-\eta}$. Particle mass and number allow simulating the concentration of phytoplankton as both a nitrogen concentration and a number of plankters after the aggregates are integrated over the model space. The continuous Smoluchowski equation is solved for the size distribution (not for the individual size classes $N(r_k, t)$, as there are not any) using: (i) standard coagulation laws, which transfer particles from one size class to another and (ii) a sinking velocity (w) computed as a function of the particle size using another power-law relationship $w(r) = c_w r^{\zeta}$. The main assumptions are that the particles are well-mixed and that there is a sufficiently high number of particles, allowing to treat the populations of particles (pools) as concentrations varying as continuous, Eulerian concentrations.

The rest of Eulerian biogeochemical models presented in Fig. 2.3 (Stemmann et al., 2004; Gehlen et al., 2006; Karakaş et al., 2009; DeVries et al., 2014; Cram et al., 2018; Omand et al., 2020) are an application of the slab (0-D) PSD formulations of J90 and KE99 to larger model spaces, essentially 1-D and 3-D, and with more complex ocean biogeochemical settings, which include disaggregation, zooplankton grazing and microbial activity, amongst others. Stemmann et al. (2004) implemented J90's bin approach to model particle dynamics in a 1-D Nutrient-Phytoplankton-Zooplankton-Detritus (NPZD) biogeochemical model to simulate PSD in the subsurface ocean (100–1000 m) of the French

JGOFS site in the NW Mediterranean Sea. Gehlen et al. (2006) implement KE99 bulk mathematical description of the PSD into the 3-D global biogeochemical circulation model PISCES. Karakaş et al. (2009) apply the continuous PSD of KE99 into an ocean coupled ecosystem model to simulate the 1-D, water-column marine snow aggregation dynamics in the NW African upwelling zone. DeVries et al. (2014) implement KE99's description of the PSD in a particle aggregation/disaggregation model that they call PRiSM, which they couple to the 3-D ocean biogeochemical model ROMS, and use it to simulate the water-column PSD in the NW African upwelling region. Cram et al. (2018) mechanise PRiSM by including the effects of particle mineral composition (ballast effect), oxygen and temperature. They use the new mechanistic model to globally evaluate the effect of such factors on the transfer efficiency of POC between the euphotic zone and 1000 m (T_{eff}). Omand et al. (2020) modify PRiSM and derive the total flux of particles from the depth and volume of each particle size class (instead of mass and particle number).

In bulk method models, the calculation of particle macro properties like the total number of particles and the mass flux typically becomes something like

$$T(z) = \int_r^\infty \frac{\mathrm{d}N}{\mathrm{d}r} = c_N \int_r^\infty r(z)^{-\epsilon}\,\mathrm{d}r, \tag{2.3}$$

$$F(z) = \int_r^\infty n(z,r)\,M(z,r)\,w(r)\,\mathrm{d}r. \tag{2.4}$$

The above Eulerian descriptions contrast with SLAMS-2.0, where there are no equations such as Eqs. 2.3 and 2.4 to calculate particle numbers and fluxes of materials. In a Lagrangian model, those emerge from the integration of the ensemble of sinking individual particles –see the corresponding Eqs. 5.4 and 5.5 in Chapter 5.

2.3.1.2 Lagrangian methods

The models of J90 and KE99 are based on a *deterministic* interpretation of the processes that modify the distribution of particles in size classes. That means that once the initial conditions of the system are known, the same starting conditions will always yield the same result. However, real particle aggregation events are *stochastic* as not every collision between two colliding particles leads to an aggregation event. Therefore, aggregation events are better represented by a probabilistic function that estimates the outcome of aggregation events, rather than a deterministic function for the expected values. Numerical stochastic methods, or simply *particle-based methods*, offer the solution. Interactions between particles occur in a probabilistic manner, where probabilities are based on standard coagulation laws (e.g., Burd and Jackson, 2009) and particles are modelled as Lagrangian, discrete entities. Particle-based methods have two sub-branches (Fig. 2.3).

- The direct simulation method, developed by Gillespie (1975), and which follows each individual particle in the system.
- The super droplet method, developed by Shima et al. (2009), and which clusters particles that are identical to each other (in terms of average depth and physicochemical attributes) into a single computational entity called super droplet (or super particle). Each computational super droplet represents a given number of real particles.

Gillespie (1975), for his direct simulation method, developed a Monte Carlo algorithm for droplet populations that is now called the stochastic simulation algorithm (SSA). Gillespie (1975) uses the Monte Carlo method to repeatedly draw a random waiting time for which the next pair of particles in the grid cell will coalesce. Gillespie (1975)'s SSA, though, becomes expensive for a large population of particles. Note that when two particles collide, they will create a third one; thus, particle collision will lead to rapid growth in the number of particles –scaled by a power of two in binary collisions–, making the coagulation problem unmanageable (Grabowski et al., 2019). Based on the SSA, Shima et al. (2009) developed a new Monte Carlo algorithm for their super droplet method (SDM) that avoids the squared growth of the number of particles after coagulation. In the SDM, particle collision scales linearly with the number of super particles; otherwise said, the number of super particles keeps unchanged as coagulation proceeds. Shima et al. (2009)'s Monte Carlo algorithm proceeds in four steps: (1) all the N super droplets in the grid cell are listed, (2) from that list, a sublist of $N/2$ randomly generated, non-overlapping pairs is selected for coagulation, (3) the probability that two super droplets will coagulate is calculated based on coagulation laws and a computer generated random number, (4) coagulation is executed (or not) for each of the $N/2$ pairs.

Gillespie's direct simulation method has been more recently used by Zahnow et al. (2011). The authors came up with a fractal-like structure representation of marine particles in a 3-D synthetic turbulent flow. They used Stokes' equation for the vertical motion of particles and applied the standard rules of particle coagulation and fragmentation, using random numbers to help make decisions about particle splitting and time before reaching a steady state. The PSD evolves from the dynamics of the ensemble. However, the model does not contain ocean biogeochemical processes of sources and sinks of particles. It is purely a model of dynamic particles for coastal, turbid areas.

Jokulsdottir (2011) showed that the SDM implemented into SLAMS-1.0 replicates both the numerical solution of the discretised version of the Smoluchowski equation (after Burd and Jackson, 2009) as well as the analytical solution (after Wetherill, 1990), proving that the SDM is a good proxy to replicate coagulation kinetics similarly to coagulation theory. Other applications of the SDM have also proven successful in approximating the behaviour

of real particle coagulation (see a review in Grabowski et al., 2019). More recently, SLAMS-1.0 has been applied by Fakhraee et al. (2020) to explore the effects of changes in the atmospheric concentration of O_2 in the geological past as well as changes in the dominant plankton type, dust input and SST in the T_{eff} of the BCP through Earth's history.

The accuracy of the solution yielded by the SDM is controlled by the size of the population of particles that is sampled and the number of repetitions of the sampling process in one time step (Boje et al., 2019). Since the MCM converges to the exact solution as the size of the sample increases, the memory costs to store particle information, and the runtime costs to run a model with lots of particles, can be very large if the random sampling is set to occur many times per time step. Thus, the modeller must make a compromise between accuracy of the solution (i.e., number of sampling events) and computational economy. In SLAMS-2.0, thanks to an optimised[4] model design, run time to reach equilibrium at a single location is short: ~20 min (equivalent to 10 model run years) for a single water column of 4000 m depth in a Mac machine with four processor cores.

Shima's SDM not only is an excellent option to simulate particle populations at a low computational cost, but it is also the only viable method to resolve the internal state of a particle with many attributes associated (Shima et al., 2009; Jokulsdottir, 2011; Boje et al., 2019; Grabowski et al., 2019). Although adding more particle attributes has a cost in that more super particles are needed to cover the increased variety of particles produced without losing accuracy, model run time is independent of the number of internal coordinates a particle has. In other words, in the SDM, model run time does not scale with the number of particle attributes but with the number of computational particles. The SDM thus becomes a computational trick to reduce model runtime and memory costs in SLAMS-2.0 due to three main reasons.

- It represents a large number of real particles that share the same properties (i.e., are clones) within a single computational Lagrangian entity. It is called *super droplet* in Shima et al. (2009) and *cluster* in SLAMS-2.0.
- The number of clusters does not increase with stochastic binary coagulation –which would grow the number of clusters to the squared power– but with the addition of particles from biological sources –which grow the number of clusters linearly.
- It can hold a large number of particle attributes without slowing down the model.

[4]SLAMS-2.0 uses a structure of arrays of "derived type" (Fortran90). The structure is the particle and the arrays are the members of the structure (the particle attributes). The arrays have a common length corresponding to the maximum number of particles that can be simulated. When SLAMS-2.0 modifies particle attributes after biogeochemical interactions, it accesses structure members. Derived types allows holding various data types (real, integer, character, logical) in the structure (e.g., particle radius is a real number, but particle initial type is an integer number).

2.3.2 Non-Smoluchowski particle models

There is a second class of particle-based models that model marine particles in a Lagrangian, discrete way but do not recreate PSDs because coagulation is not represented in the model formulations. Those models are what I call non-Smoluchowski particle models. I review two cases: Lagrangian models of marine particles that are living cells (instead of marine snow) (Sec. 2.3.2.1) and hybrid Lagrangian-Eulerian planktonic models (Sec. 2.3.2.2). With this, I will finish the review of particle-oriented models in the OBM literature.

2.3.2.1 Lagrangian modelling of plankton

In the field of oceanography, Lagrangian modelling of marine particles is not unique to marine snow. There are other entities treated as Lagrangian marine particles, like phytoplankton (Woods and Onken, 1982; Woods, 2005), planktonic foraminifera (Van Sebille et al., 2015), microbes (Hellweger et al., 2014), fish eggs/larvae (Paris et al., 2013; Ådlandsvik et al., 2009) or squid (Sinerchia et al., 2012). Quite remarkable is the Lagrangian plankton model of Woods and Onken (1982). They broke with the classical Eulerian-continuum method used in NPZD modelling, where plankters are treated as a continuum property of water (e.g., Fasham et al., 1990), and developed the Lagrangian ensemble (LE) method, which follows an ensemble of individual plankters (phyto- and zooplankton) along their life history, simulating at a biochemical level the cellular processes of growth, reproduction and death (Sinerchia et al., 2008). This way, processes that are highly non-linear, and thus problematic to parameterise in Eulerian modelling, like the development of phytoplankton blooms as a consequence of cell division, do not proceed at an average rate, where no account is taken of the variability of individual plankters. Instead, LE modelling considers all the individual cellular growth rates and computes the average from the ensemble of plankters. That is a more mechanistic approach to understanding the interactions between plankters and their ecosystem. The LE model has been exclusively used by the research group lead by John Woods at Imperial College. Their publications, from 1982 to 2012 (last year of published records), are listed here: http://www.virtualecology.org/docs/index.cgi.

In LE modelling, as in the SDM, each computational plankter represents a subpopulation of identical plankton particles based on the properties of one plankter. The LE model, like SLAMS, does not use partial differential equations. However there are other fundamental differences between the LE model and SLAMS. (a) The LE model does not consider particle-to-particle interaction (collision, aggregation, disaggregation, prey detection by zooplankton), as SLAMS does; it only allows particle-to-concentration interaction. (b) The LE models plankters (phyto- and zooplankton) whereas SLAMS models the wider realm of marine particles, which includes plankters and also spans

to faecal pellets, TEPs, dead plankters and aggregates of those. (c) The LE model is focused on the dynamics and energetics of plankters, whereas SLAMS is focused on the characterisation of the sinking flux of biogenic carbon in the BCP.

2.3.2.2 Hybrid Lagrangian–Eulerian planktonic models

Here are two cases of biogeochemical models that apply a hybrid Lagrangian–Eulerian framework to simulate non-coagulating Lagrangian particles in their aquatic environment.

Xue et al. (2018) developed a property-carrying particle model where Lagrangian, discrete estuarine particles carry with them many internal coordinates that correspond to the state variables of an NPZD Eulerian model. NPZD time-derivatives operate on each particle to determine the concentration of a generic particulate tracer remaining after all the ecosystem model processes have operated at the end of the time step. Note that those particles do not carry physicochemical attributes (as in SLAMS-2.0), but ecosystem time-derivates and a 3-D location property. Thus, the calculation of the trajectory of a particle is not based on its physicochemical attributes but pre-computed based on the output of an off-line hydrodynamic model. Xue et al. (2018) applied their model to a freshwater estuary. For each grid cell of the model space, an initial particle number is prescribed proportional to the water volume input from the river that discharges into the great lake. Particle number evolves according to the NPZD model acting on each particle, and particles move according to the hydrodynamic model.

Maerz et al. (2020) developed a fractal-like structure representation of marine particles and implemented it into the global biogeochemical model HAMOCC (HAMburg Ocean Carbon Cycle Model). Maerz and colleagues developed a detailed representation of marine particles that considers their size, sinking velocity, volume, stickiness, number of primary particles or amount of packed material. Particles are created from the HAMOCC available tracer concentration at a given grid cell of POC, CaCO3, opal and dust. Those materials are encapsulated into primary particles according to reference sizes, volumes and densities of diatoms, coccolithophores, detritus and dust particles. Those particles undergo remineralisation of organic carbon and dissolution of opal and calcium carbonate. However, the Lagrangian representation of particulate material finishes at the level of particle aspect description and sinking. Aggregation and fragmentation are implicitly represented by imposing a PSD slope that assumes homogeneous composition of aggregates. Thus, Maerz et al. (2020) model is not a particle-tracking simulation that can resolve the Smoluchowski equation as an emergent property, as that is imposed. But interestingly, the Martin curve slope does emerge from the fluxes of POC.

Part II

METHODS

3 Model description

Contents

3.1 Overview of SLAMS-2.0

SLAMS-2.0 is a stochastic model of Lagrangian marine particles written in Fortran90. It simulates particle physicochemical attributes as they biogeochemically interact with their ecosystem and gravitationally sink through a 2-D water column. I have developed SLAMS-2.0 to mechanistically link environmental forcing to critical living components of the BCP and thus gain a quantitative understanding of the environmental controls that affect the attenuation of POC flux in the ocean. The model has two modules (Fig. 3.1).

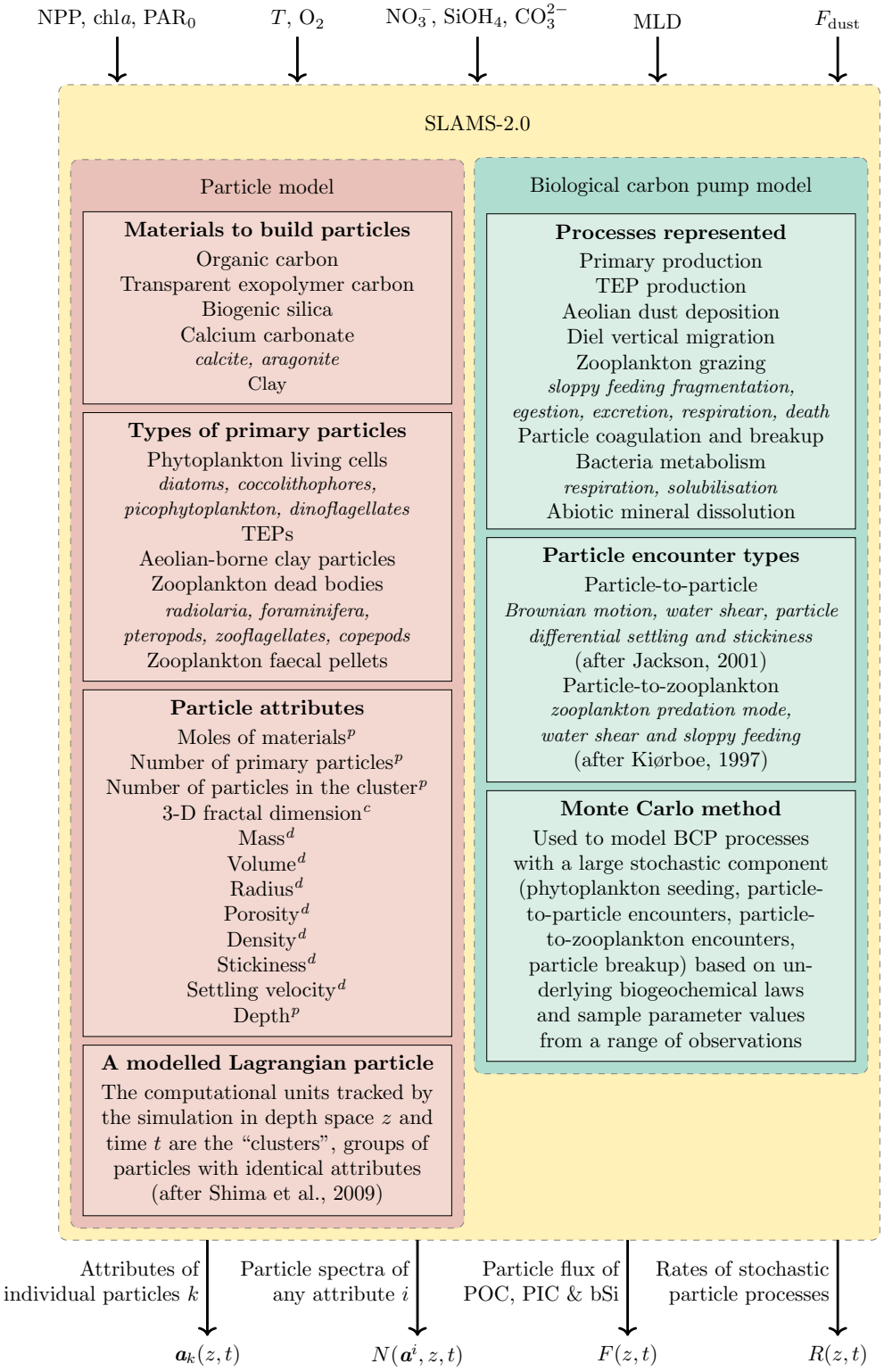


Figure 3.1. Schematic of the characteristics of the components of SLAMS-2.0 (Stochastic, Lagrangian Aggregate Model of Sinking particles, version 2.0), model forcing (input arrows) and model outputs (output arrows). Particle attributes with superindex d are diagnostic variables, those with superindex p are prognostic variables, and those with superindex c are held constant.

- A particle module that resolves particle physicochemical attributes and tracks their depth position as they interact with each other and their biogeochemical environment (Sec. 3.3). A modelled *particle* consists of an amount of material distributed amongst many *primary particles* organised within the structure according to a fractal dimension value. Identical particles are grouped into *clusters*, the computational units tracked by SLAMS-2.0 (Fig. 3.2). Particle transformations are simulated after particle-particle or particle-environment interactions by updating the physicochemical attributes of mass, volume, radius, porosity, density, stickiness and sinking velocity according to fractal scaling law and Stokes' law. There are 18 particle attributes (16 physicochemical attributes, a depth position in the water column and a particle multiplicity factor), 12 sources of primary particles (diatoms, coccolithophores, dinoflagellates, picophytoplankton, TEPs, clays, zooplankton faecal pellets, copepods, radiolarians, foraminifera, pteropods and zooflagellates) and six sources of materials to form primary particles (organic carbon, transparent exopolymer carbon, biogenic silica, calcite, aragonite and clay) (Fig. 3.2).
- A biological carbon pump (BCP) module that contains the biogeochemical processes that create, transform and degrade particles (Sec. 3.4). The BCP module has access only to two particle attributes: the number of primary particles and the moles of material embedded in the particle –the rest of attributes are calculated from those two using fractal scaling and Stokes' law. The BCP is made of these components: two macronutrients (NO_3^- and $SiOH_4$), CO_3^{2-}, O_2, aeolian clay, PAR, temperature, MLD, phytoplankton cells (silicifying diatoms, calcifying coccolithophores, dinoflagellates and picophytoplankton), zooplankton bodies (silicifying radiolaria, calcifying foraminifera and pteropods, zooflagellates and copepods) and implicit heterotrophic bacteria. The BCP processes represented in SLAMS-2.0 are: primary production, which follows standard $P-I$ kinetics (Geider et al., 1997) and is limited by temperature and macronutrient availability; TEP production, which is dependant on large phytoplankton cells; particle coagulation, which follows Jackson (2001)'s collision kinetics and uses TEP-mediated sticking; particle-zooplankton encounters, which depend on the standing stock of POC, DVM rules of feeding vs mortality risk and feeding behaviour strategy following Kiørboe (1997)'s formulations; zooplankton grazing, which follows a size-dependent relationship as well as standard $CaCO_3$ dissolution kinetics (Jansen and Wolf-Gladrow, 2001) and Q_{10} and O_2 limitation kinetics (Cram et al., 2018); bacteria respiration and solubilisation, which follows the same Q_{10} and O_2 respiration kinetics as zooplankton; abiotic mineral dissolution of $CaCO_3$ and bSi, which follows standard dissolution kinetics based on saturation coefficients and dissolution rates; and particle physical breakup, which uses an empirical model of particle fragmentation (Goldthwait et al., 2004).

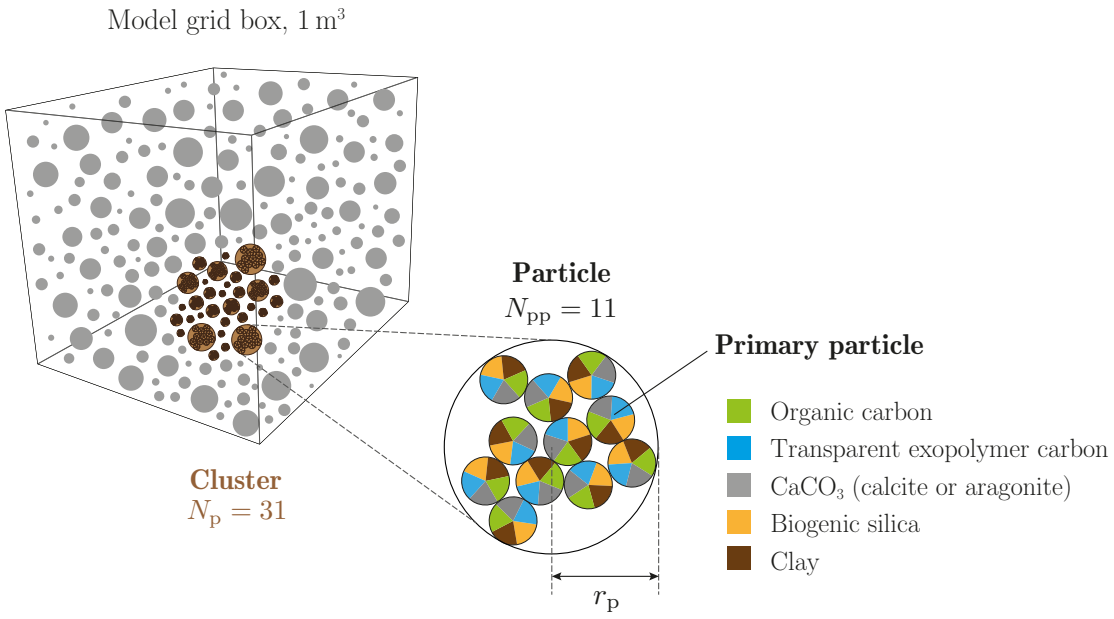


Figure 3.2. Schematic of a Lagrangian particle in SLAMS-2.0. A model grid box with sinking computational particles (clusters) is depicted. I have highlighted in brown a cluster that represents 31 identical particles. From those 31 identical particles, one has been enlarged. The enlarged particle is an aggregate made of 11 identical primary particles formed by a combination of the six model materials. N_p is number of particles that make up the cluster, N_{pp} is number of primary particles in the aggregate structure, and r_p is the radius of the particle.

SLAMS-2.0 improves the mechanisation of the BCP offered by its predecessor version, SLAMS-1.0 (Jokulsdottir and Archer, 2016), where (i) biological processes were imposed regardless of the local environment conditions (e.g., there is no macronutrient, nor temperature (T) nor photosynthetic active radiation (PAR) control on phytoplankton growth), (ii) there are not different zooplankton types, and the processes of encountering food particles and grazing are all parameterised, (iii) remineralisation is not coupled to O_2, (iv) mineral dissolution is not coupled to depth, nor CO_3^{2-}, and (v) biogenic particle creation does not conserve biomass. The main changes with respect to SLAMS-1.0 are the development of two plankton probabilistic models: one for seeding phytoplankton cells based on environmental forcing and another for particle encountering and grazing by zooplankton. The code has also been modified to be implemented on a global grid of environmental forcing and uses efficient structures to store particle information and optimised loops[1], which have reduced model runtime.

At the core of SLAMS are two realisations that make it different from other models in the OBM literature: that POC fluxes are made of discrete marine particles and that BCP events are hugely stochastic. Thus SLAMS becomes a stochastic particle-based

[1]For instance, by vectorising loops (or getting rid of *for* loops whenever possible), placing independent code outside loops or having the inner loop operating on the most left-hand side dimension of an array (in Fortran90), preallocating space or explicitly declaring the size of all variables.

simulation, where the mathematical facets of *Lagrangian* and *stochastic* modelling become computational tricks to improve computational economy.

- In Lagrangian modelling, particles are modelled discretely, rather than as a continuous Eulerian mass. That feature is made computationally feasible by grouping lots of identical particles into computational particles (the clusters).
- In stochastic modelling, (i) BCP processes are approached using probabilities that are randomly sampled with the Monte Carlo method (MCM) and (ii) unknown parameters are approached with a MCM extraction from a range of possible values. From all simulated BCP processes in SLAMS-2.0, particle coagulation has a distinct Monte Carlo algorithm, which was developed by Shima et al. (2009) and called the super droplet method (SDM). The SDM helps to keep SLAMS-2.0 computational expense low by keeping the number of computational particles unchanged as binary coagulation proceeds. That process would make particle numbers grow exponentially in standard binary coagulation procedures (as explained in Sec. 2.3.1.2 in Chapter 2).

Contrary to most biogeochemical particle-oriented models (see Fig. 2.3), SLAMS-2.0 does not use time derivatives to model the impact of the BCP on the particle attributes but algebraic equations. BCP rate processes are modelled using the analytical solution of the the time derivatives. Those analytically-derived solutions operate on the amount of particle material and number of primary particles. From the new particle state of moles and primary particles, the rest of physicochemical attributes are calculated diagnostically (via fractal-scaling law and Stokes' law) and the depth position is calculated prognostically. As a consequence of resolving thousands of particle cluster states, a rain of particles is generated. Particle fluxes of POC, PIC and bSi emerge from the statistical ensemble of particles, particle size spectra can be built and average BCP rates of stochastic processes are calculated for every depth layer and on monthly basis.

Next, I describe the model equations and the assumptions used. All cited parameters are in Table C.1 in Appendix C, with a detailed description of how values have been assigned to them. The model code is available at https://doi.org/10.5281/zenodo.5570253, and a brief user guide is provided in Appendix B, outlining how to set up a model experiment and a summary of model output files.

3.2 Model temporal and spatial framework

SLAMS-2.0 uses a large time step size (Δt = 8 h) and a small grid box volume (ΔV = 1 m^3) when compared to conventional biogeochemical models (Δt < 1 h and ΔV > 1000 m^3). Those two parameters control the number of clusters circulating when a repeating annual cycle of POC flux is reached (i.e., model equilibrium state). Equilibrium is desired with

an accompanying sufficiently large number of clusters to ensure that the population of particles is as diverse as possible, but not too abundant to reduce the computational load particles and avoid a memory-demanding simulation. The choice of Δt and ΔV is inherited from SLAMS-1.0, where the temporal and spatial framework of the model was chosen accordingly to keep computational costs low while producing enough clusters in each grid box for model results to be comparable to observations (Jokulsdottir, 2011).

With model time step Δt = 8 h, POC flux equilibrium is reached in within 10 model run years, which take a convenient runtime of ~20 min on a Mac machine with four computer cores for a water column of ~4000 m. Since SLAMS-2.0 does not use differential equations but algebraic equations, the long Δt (longer than some ecosystem's rate parameters) does not cause numerical instability. Furthermore, an adaptive time stepping is used for particle coagulation, where Δt is recomputed internally by the SDM algorithm for collisions that occur in a narrower time frame.

SLAMS-2.0 represents the cycle of day and night, as required by some BCP events that can only happen during the daylight period, like primary production, or during the nighttime, like the vertical migration of zooplankton to the surface ocean. A modelled day (24 h) contains three time steps where, for convenience, daylight lasts for two time steps (16 h) and nighttime one time step (8 h). That partitioning of daylight/nighttime duration is typical of mid-latitude regions in the summer and has been applied homogeneously throughout the global grid of SLAMS-2.0, regardless of the latitude and time of the year. The use of that fixed partitioning is justified based solely on debugging reasons and needs to be made spatiotemporal dependent in the next application of SLAMS-2.0. However unnatural the diurnal time stepping is for some locations and moments of the year, the PAR_0 ($W\,m^{-2} = J\,m^{-2}\,s^{-1}$) and the NPP ($mol\,C\,m^{-2}\,s^{-1}$) forced into the model have been normalised to the number of daylight, sunny seconds of a day. Thus, the intensity of radiation and the amount of organic carbon injected into SLAMS-2.0 are not affected by the choice of diurnal time stepping. What is affected is the frequency at which the model is pulsed with primary production events.

The model physical space is a water column closed by a sea surface and a seafloor. A water column has a height H given by the local seafloor depth and it is sliced into depth layers of a thickness Δz = 10 m each. A depth layer in the water column has a volume ΔV = 1 m^3 and thus a theoretical cross-sectional area of ($\Delta V/\Delta z$) 0.1 m^2. There is no advection nor diffusion moving the particles in the 3-D water field as SLAMS-2.0 has not been coupled to an OGCM. Thus, water motion dynamics are parameterised via a turbulent kinetic energy dissipation rate and a temperature-dependent water dynamic viscosity. See Appendix C.1 for the equations of parameterised water physics.

Each water column is forced at the surface with data of local PAR_0, NPP, chl a and F_{dust}. It has a depth-layer distribution of T and nutrients (NO_3^-, $SiOH_4$, CO_3^{2-}, O_2) along H, and it has an imposed mixed layer depth (MLD). Unlike other popular biogeochemical simulations (e.g., MEDUSA, COBALT, HAMOCC), variables like dissolved nutrients or NPP are not updated time step after time step from the model pool but are constantly forced into the simulation. That is because the BCP module of SLAMS-2.0 does not resolve well the flows of such variables given the large time step Δt of the simulation. See Sec. 4.2 in Chapter 4 for the reference sources of the forcing variables.

SLAMS-2.0 can be run at many water columns simultaneously, each with their independent, enclosed particle and BCP system, i.e., there is no inter-column exchange of materials. A global ocean representation can be achieved by creating a global grid of water columns. In that case, the computational cross-sectional area of the water column still remains 0.1 m^2, although the actual area is given by the resolution of the global latitude/longitude grid chosen. Thus, an assumption is made that the processes happening in the 0.1 m^2 cross-section are representative of the actual and much larger cross-sectional area. Currently, SLAMS-2.0 uses a coarse latitude/longitude grid to speed up computations, which is of 2.8° resolution (~280 km), equivalent to 64 × 128 grid points on the horizontal space.

3.3 The particle module

3.3.1 Description of particle state

In SLAMS-2.0, modelled particle attributes (Table 3.1) are local to the Lagrangian particle. Particle aggregates are built up from a heterodisperse suspension of primary particles coagulating or being packaged by zooplankton into faecal pellets. The concept of *primary particle* in SLAMS-2.0 refers to any non-porous, compact particle that cannot be disaggregated into sub-entities, and there are six types (Fig. 3.3), all with distinct origins, material compositions and sizes. A particle is called *aggregate* when it is made up of more than one primary particle aggregated together. Groups of identical particles, be aggregates or primary particles, are tracked together by the Lagrangian framework in computational units called *clusters*. A cluster thus represents a large number of clone particles (i.e., with shared average water depth and physicochemical attributes). A certain number of clusters can be created before reaching low computational speed –subject to machine specifications where the simulation is running, but generally ~20,000 clusters.

All particles are first created from a carbon and a mineral quota specific for a type of primary particle. The quota is randomly selected from a range of possible values according to prescribed boundary quotas, e.g., $[\theta_C^{min}, \theta_C^{max}]$. The amount of material is then shaped

Table 3.1. Variables that define particle state in SLAMS-2.0. Those can be grouped into three types: 16 physicochemical attributes, depth position and particle multiplicity in the cluster.

Attribute description	Symbol	Units
amount of organic carbon (C_{org})	n_{Corg}	mol
amount of transparent exopolymer carbon (TEC) from TEPs	n_{TEC}	mol
amount of bSi, the hydrated form of silica, $SiO_2 \cdot 0.4\,H_2O$	n_{bSi}	mol
amount of biogenic calcite, a form of $CaCO_3$	n_{calc}	mol
amount of biogenic aragonite, a form of $CaCO_3$	n_{arag}	mol
amount of clay in the form of kaolinite, $Al_2O_3 \cdot 2\,SiO_2 \cdot 2\,H_2O$	n_{clay}	mol
number of primary particles in the aggregate particle	N_{pp}	
3-D fractal dimension	D_3	
dry mass	m_p	g
volume	V_p	µm^3
radius of each primary particle	r_{pp}	µm
radius of the particle	r_p	µm
porosity	P_p	0–1
density	ρ_p	g cm^{-3}
stickiness	α_p	0–1
sinking velocity	w_p	m d^{-1}
depth position in the water column	z_p	m
number of particles in the cluster (multiplicity factor)	N_p	

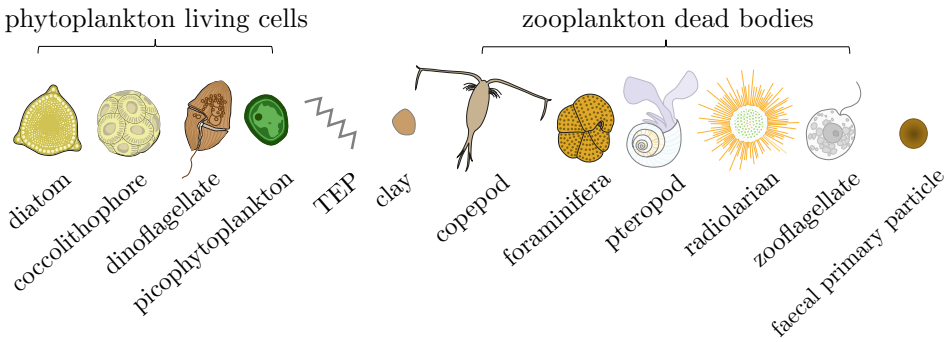


Figure 3.3. Types of primary particles represented in SLAMS-2.0. The diatom represented is a triangular diatom of the genus *Triceratium* sp.; the coccolithophore is the species *Emiliania huxleyi*; the dinoflagellate is of the genus *Gonyaulax* sp.; the picophytoplankton is a *Prochlorococcus* sp.; the copepod is from the Calanoida order; the foraminifera belongs to the genus *Globigerina* sp.; the pteropod belongs to the Thecosomata clade; the radiolarian is from the genus *Actinomma* sp., and the zooflagellate is a generic heterotrophic nanoflagellate. *Living* zooplankton are not modelled as Lagrangian particles but as a concentration that interacts with the particles.

into a particle according to fractal scaling and Stokes' law for the calculation of sinking velocity. The 3-D fractal dimension, D_3, is the only particle attribute held constant. Four particle attributes are calculated prognostically: the number of moles of the different materials i (n_i), the number of primary particles in the aggregate (N_{pp}), the multiplicity factor (N_p) and the depth position (z_p). For n_i, N_{pp} and N_p the increment or decrement is given by the changing surrounding BCP (by using the analytical expression of the

time-derivative process that affects the particle) and for z_p the new value is calculated from the new sinking velocity (using Euler time stepping, as $z_p(t) = z_p(t-1) + w_p(t)\Delta t$). The rest of the physicochemical attributes are calculated diagnostically every time step. Their equations are summarised in Fig. 3.4 and defined in Appendix C.2. The calculations of those attributes are the same as in SLAMS-1.0 (Jokulsdottir and Archer, 2016) except for the porosity and sinking velocity, which have slight changes. For a comparison with observed particles in seawater, Table 3.2 shows the modelled attributes of density, volume, size and sinking velocity of the six primary particle types that are seeded into the model in the surface ocean.

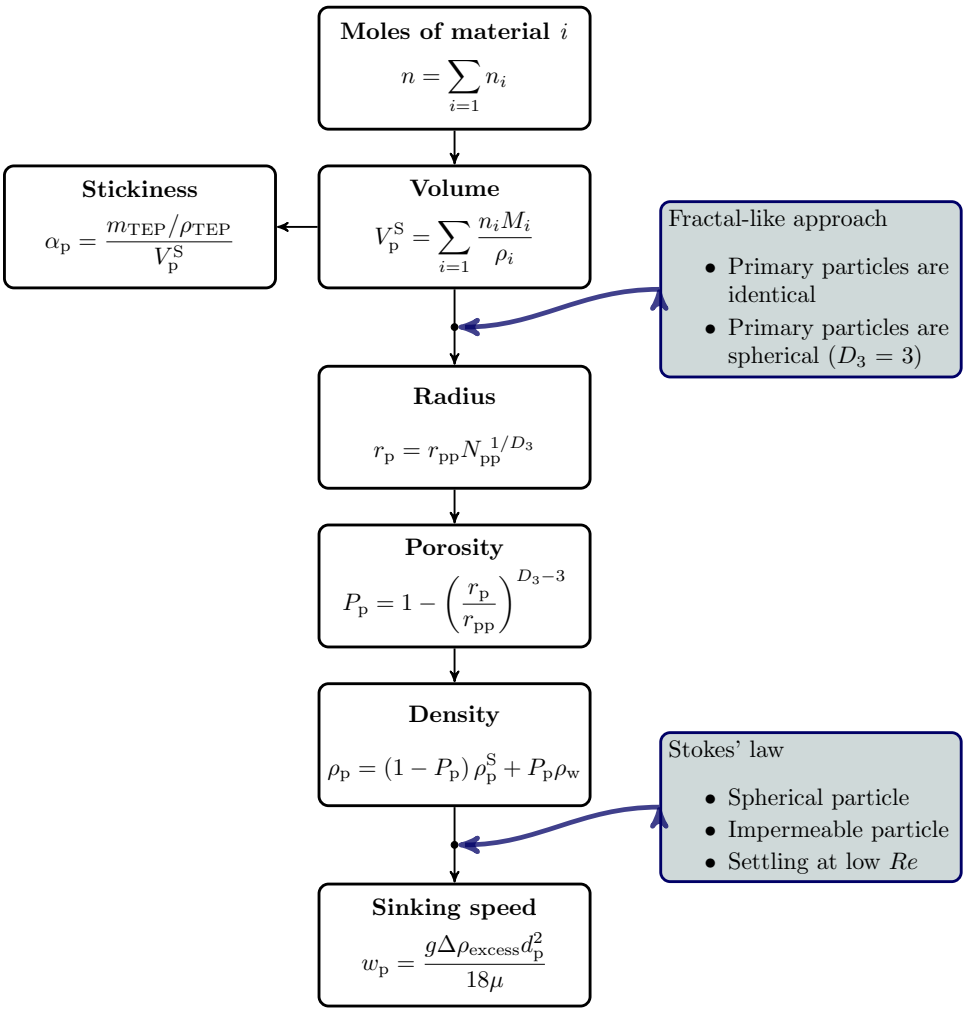


Figure 3.4. Schematic of SLAMS-2.0's diagnostic particle attributes and their equations. Discussed in Appendix C.2.

3.4 The biological carbon pump module

The material that fuels the BCP is the phytoplankton cells that are seeded in the ocean's sunlit layer according to temperature, light and nutrient availability (Sec. 3.4.1) (Fig. 3.5).

Table 3.2. Comparison of modelled vs observed particle attributes of density, volume, diameter and sinking velocity in six types of primary particle. Values are provided as ranges. Observations are from a compilation of data, where a total of 32 references were screened for values and which is available at https://doi.org/10.5281/zenodo.5233652. Observed diameters, when not available, have been computed from the observed volume considering sphericity.

Primary particle	Density ($g\,cm^{-3}$)	Volume (μm^3)		Diameter (μm)		Sinking velocity ($m\,d^{-1}$)	
		Mod.	Obs.	Mod.	Obs.	Mod.	Obs.
Diatom	1.13	$184 - 23,000$	$100 - 10^7$	$7.1 - 35$	$4 - 200$	$0.18 - 4.6$	$0.30 - 35$[a]
Cocco.	1.90	$79 - 314$	$25 - 10^4$	$5.3 - 8.4$	$3 - 14$	$0.87 - 2.2$	$0.15 - 10$[b]
Dino.	1.06	$8,400 - 210,000$	$200 - 10^7$	$25 - 74$	$10 - 200$	$0.74 - 6.3$	
Pico.	1.06	$0.42 - 21$	$0.10 - 25$	$0.93 - 3.4$	$0.6 - 4$	$0.00 - 0.01$	
TEP	0.800	$26,000 - 30,000$		$37 - 39$	$1 - 100$[d]	$-(11 - 12)$	
Clay	2.70	8.38		2.52	$0.10 - 20$[c]	0.37	

[a] Miklasz and Denny (2010); [b] Jansen et al. (2002); [c] Schulz et al. (2012); [d] Passow (2002).

In regions of the ocean exposed to terrigenous dust, there is aeolian deposition of clay particles (Sec. 3.4.3). During their lifetime, large phytoplankton cells (diatoms and dinoflagellates) that are under stress exude TEPs that glue marine particles together (Sec. 3.4.2). After that, zooplankton consume living phytoplankton and detrital particles. The amount of zooplankton found in the water column depends on the moment of the day and depth in the water column following DVM rules (Sec. 3.4.5.1), and particles will be encountered by zooplankton depending on the turbulence of the surrounding environment and zooplankton feeding behaviour (Sec. 3.4.5.2). Ingestion of food particles and assimilation of the material leave behind new forms of particles: zooplankton faecal pellets and dead bodies. Zooplankton feeding also includes the fragmentation of food particles as the plankter mishandle them (sloppy feeding), dissolution of the $CaCO_3$ present in the particle in the zooplankton digestive system, excretion and respiration (Sec. 3.4.5.3). Aggregates are formed with the various types of primary particles available after the described food web processes. Aggregation consists in physical collision followed by coalescence aided by TEP (Sec. 3.4.4). But aggregates are also degraded and disintegrated due to: bacterial metabolism (Sec. 3.4.7), dissolution of the bSi and $CaCO_3$ content below the lysocline (Sec. 3.4.8), physical breakup of unstable particles and disintegration by photolysis of buoyant particles (Sec. 3.4.6).

All those BCP processes are evaluated sequentially at every time step. An auxiliary function is then run to clean up the model space from single primary particles that are poorly sticky and, thus, will never aggregate, by packing them into clusters.

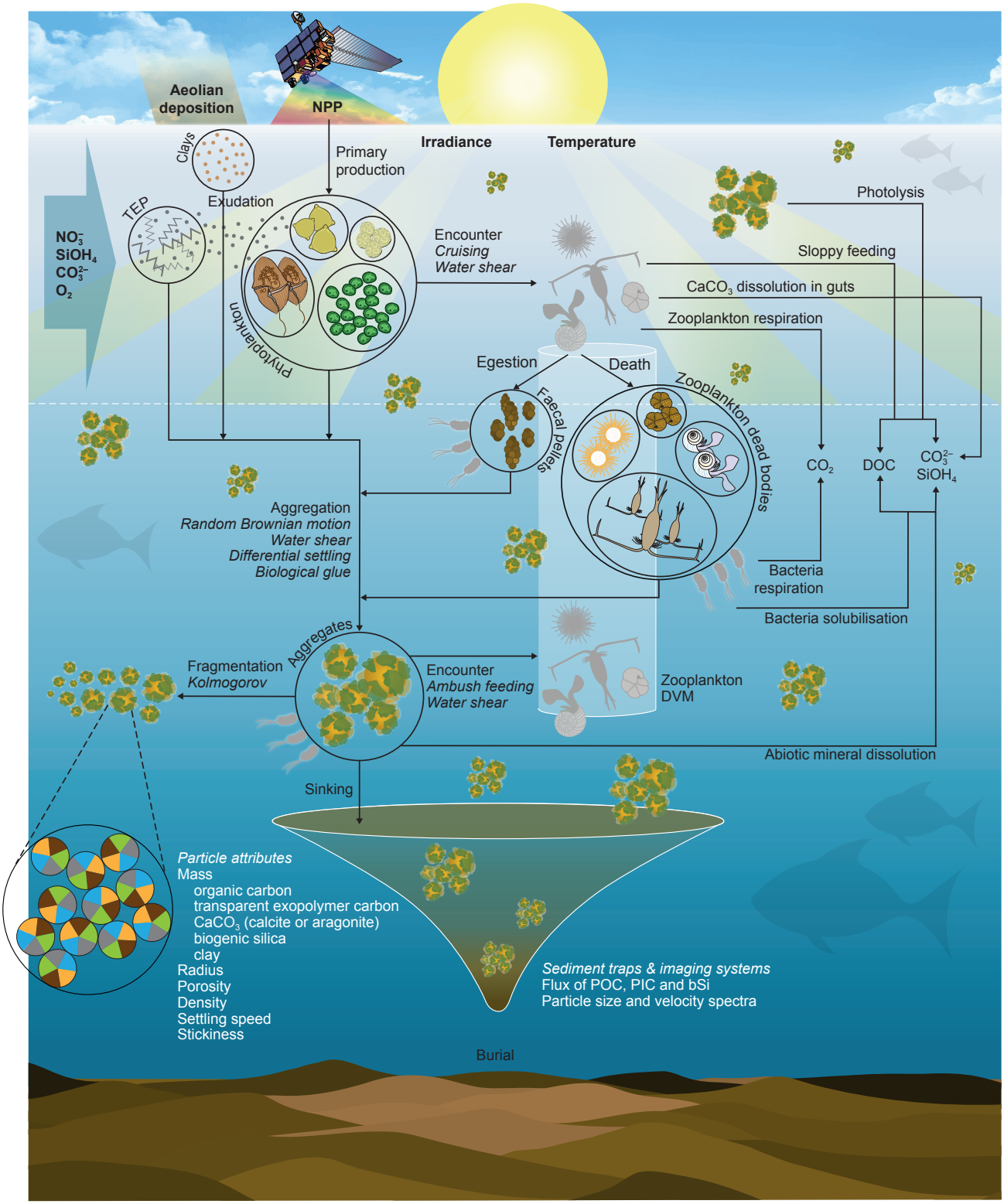


Figure 3.5. Schematic of the mechanistic model of Lagrangian marine particles in the context of the biological carbon pump (BCP) simulated in SLAMS-2.0.

3.4.1 Production of phytoplankton cells

SLAMS-2.0 represents four phytoplankton functional types (PFTs) as primary particles, in correspondence with the four major ecological groups in the ocean (Le Quéré et al., 2005; Sarmiento and Gruber, 2006, Panel 4.2.2): silicifying diatoms, calcifying coccolithophores, large motile dinoflagellates and biomass-dominant picophytoplankton. Those PFTs have distinct biogeochemical roles, respond differently to environmental conditions and are of quantitative importance in at least some region of the ocean (Margalef, 1978; Li, 2002; Zwirglmaier et al., 2008).

The mineral parts of the modelled phytoplankton are made directly from the available dissolved $SiOH_4$ (for diatoms) and CO_3^{2-} (for coccolithophores) forced into the model, and the organic component is made based on the locally forced NPP. At every time step, the NPP is transformed into 20 clusters[2] of phytoplankton and then each cluster is assigned a PFT and a number of identical phytoplankton cells (primary particles). To seed those clusters, there are three decisions to be made: (i) how many of each PFTs will be seeded, (ii) how many primary particles should each cluster contain and (iii) at which depth of the euphotic zone the PFT cluster should be allocated. I have developed a probabilistic model to execute those three steps (illustrated in Fig. 3.6), where a stock of phytoplankton primary particles is seeded from NPP and modulated according to local temperature, macronutrient availability (NO_3^- and $SiOH_4$) and PAR_0.

Mathematically, the probabilistic model to seed PFTs relies on the calculation of phytoplankton growth rates, which follow a standard growth limitation equation,

$$\mu^i = \mu_0^{\max,i} \times L_{\mathrm{T}}^i \times L_{\mathrm{N}}^i \times L_{\mathrm{PAR}}^i, \quad i = 1 \ldots 4 \text{ PFTs}, \tag{3.1}$$

where μ^i (divisions d^{-1}) is the PFT specific growth rate and it is scaled down by three dimensionless factors: L_{T}^i (Eq. 3.2a), a temperature factor that follows Q_{10} kinetics and limits $\mu_0^{\max,i}$ (divisions d^{-1}), the maximum growth rate at 0°C; L_{N}^i (Eq. 3.2b), a nutrient factor that considers limitation by NO_3^- and $SiOH_4$ and follows Michaelis–Menten/Monod kinetics; and L_{PAR}^i (Eq. 3.2c), a PAR_0 limitation factor that uses the Smith (1936) function. The dependency of μ^i on the limitation factors is written using the multiplicative law, which assumes that the three factors act simultaneously, as the product.

$$L_{\mathrm{T}}^i = Q_{10}^{i\ \frac{T}{10}}, \tag{3.2a}$$

$$L_{\mathrm{N}}^i = \frac{[NO_3^-]}{[NO_3^-] + k_{\mathrm{N}}^i} \times \frac{[SiOH_4]}{[SiOH_4] + k_{\mathrm{Si}}^{\mathrm{diat}}}, \tag{3.2b}$$

$$L_{\mathrm{PAR}}^i = \frac{\mathrm{PAR}_z}{\sqrt{{\mathrm{PAR}_z}^2 + \left(\frac{\mu_{\mathrm{T}}^i}{\alpha_{\mathrm{chl}}^i \, \theta_{\mathrm{chl:C}}^i}\right)^2}}, \tag{3.2c}$$

[2] As explained in Jokulsdottir and Archer (2016), 20 is a number that was estimated from a trade-off between computational economy and particle diversity.

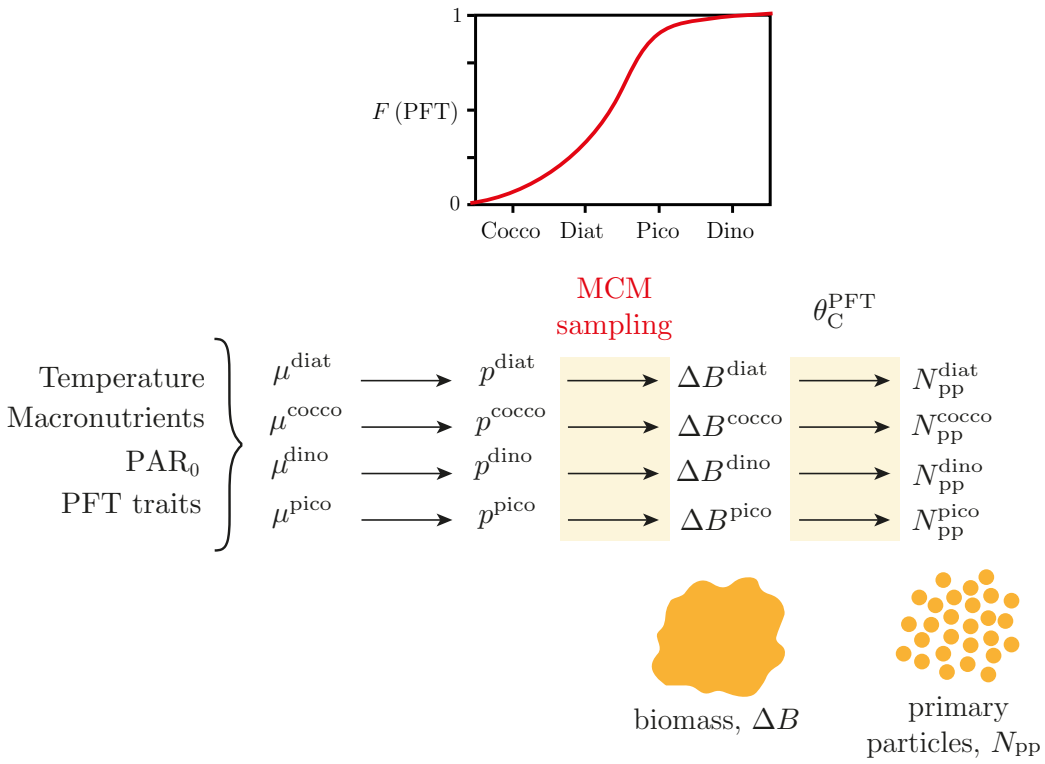


Figure 3.6. Schematic of the probabilistic model used in SLAMS-2.0 to seed phytoplankton primary particles based on environmental forcing (temperature, macronutrients and PAR_0) and phytoplankton physiological traits (the photosynthetic parameters). Four phytoplankton functional types (PFTs) are considered: diatoms, coccolithophores, dinoflagellates and picophytoplankton. Photosynthetic parameters and forcing variables are combined using the laws of phytoplankton growth (see Eq. 3.1) to calculate the growth rates of the PFTs at the different depth layers of the euphotic zone, $\mu^{PFT}(z)$. Next, growth rates are transformed into probabilities p^{PFT} ($p^{PFT} = \mu^{PFT} / \sum \mu^{PFT}$), obtaining (i) the depth distribution of those PFTs probabilities, $p^{PFT}(z)$ and (ii) the overall p^{PFT} at the euphotic zone. I then sample 20 times those two probability distribution functions (F(PFT)) with the Monte Carlo method (MCM). By sampling p^{PFT}, I get the most likely type of PFT that can grow in the overall euphotic zone, and by sampling $p^{PFT}(z)$, I get the most likely depth for allocating it. I thus make 20 clusters of the most likely PFTs and seed them into their likely depths. The number of primary particles (cells) in each of the 20 clusters, N_{pp}^{PFT}, is decided by dividing the local NPP forced into the model by 20, thus allocating the same biomass ΔB into each cluster. Next, ΔB^{PFT} is discretised into N_{pp}^{PFT} cells by dividing it by a C_{org} cell quota (θ_C^{PFT}) of the corresponding PFT randomly sampled from the possible range of values [$\theta_C^{min,PFT}$, $\theta_C^{max,PFT}$].

where Q_{10}^i is a temperature dependence constant that states that for every 10°C increase of T, the physiological rate (μ) changes by a factor of Q_{10}, which for most biological systems doubles or triples the base rate value; k_N^i (mmol N m^{-3}) and k_{Si}^{diat} (mmol Si m^{-3}) are the half-saturation constants for NO_3^- and $SiOH_4$ uptake, two frequently limiting macronutrients in the ocean; PAR_z is the PAR at depth z (Eq. 3.3); $\mu_T^i = \mu_0^{max,i} \times L_T^i$ (first two terms in Eq. 3.1); α_{chl}^i (g C (g chl)$^{-1}$ m^2 (μmol photons)$^{-1}$) is the chl-specific initial slope of the P–I curve; and $\theta_{chl:C}^i$ (g chl g C^{-1}) is the chl:C ratio (Eq. 3.4).

PAR_z is calculated from PAR_0 (W m^{-2}) according to the Beer–Lambert law with

a PAR attenuation coefficient k_{PAR}:

$$\mathrm{PAR}_z = \mathrm{PAR}_0\, e^{-k_{\mathrm{PAR}} z_{\mathrm{eu}}}, \tag{3.3}$$

where k_{PAR} (m^{-1}) is a single euphotic zone z_{eu} (m) value and is the sum of the attenuation properties of seawater, k_{w} (m^{-1}), and phytoplankton shading, k_{chl} ($\mathrm{m}^2\,(\mathrm{mg\,chl})^{-1}$): $k_{\mathrm{PAR}} = k_{\mathrm{w}} + k_{\mathrm{chl}}\,[\mathrm{chl}\,a]$. The parameter z_{eu} is defined as the upper layer of the ocean where $\mathrm{PAR}_z > 1\%$ of PAR_0. Therefore, the phytoplankton growth rate $\mu^i(z)$ is evaluated at all the depth layers where the previous condition is true.

$\theta^i_{\mathrm{chl:C}}$ is based on Geider et al. (1997)'s model describing phytoplankton chl:C under balanced growth,

$$\theta^i_{\mathrm{chl:C}} = \frac{\theta_{\mathrm{chl:C}}}{1 + \dfrac{\alpha^i_{\mathrm{chl}}\,\theta_{\mathrm{chl:C}}\,\mathrm{PAR}_z}{2\,\mu^i_{\mathrm{T}}}}, \tag{3.4}$$

where $\theta_{\mathrm{chl:C}}$ ($\mathrm{g\,chl\,(g\,C)^{-1}}$) is an average observed chl:C ratio in phytoplankton

Once the four μ^i are calculated, the actual growth rates that are passed into the probabilistic model are the MIN value between μ^i and a hypothetical maximum threshold value for phytoplankton growth, $\mu^{\max}$. The optimal $\mu^i(z)$ lies beneath the uppermost surface, where a better compromise between surface light and bottom-upwelled nutrients exists.

3.4.2 Production of TEPs

Transparent exopolymer particles (TEPs) are an excretion of mucilaginous, sticky polysaccharides released by large phytoplankton in stress (e.g., when they are in bloom conditions). TEPs act as a sticky matrix for aggregation, boosting particle coagulation (Passow, 2002; Azetsu-Scott and Passow, 2004) and, since are less dense than water (density of 0.70–0.84 $\mathrm{g\,cm^{-3}}$), TEPs make aggregates more buoyant.

In SLAMS-2.0, TEPs are produced by the large PFT groups, diatoms and dinoflagellates. Out of the 20 PFT clusters seeded, I select those made of diatoms and dinoflagellates and extract from them 30% (f_{TEP}) of the total cluster organic carbon ($\mathrm{C_{org}}$). I recalculate the properties of the phytoplankton cells after they have produced the exudates, so that those shrink after releasing $\mathrm{C_{org}}$. The exuded sticky biomass is encased into two TEP clusters, or a 10% of the amount of PFT clusters (20), coinciding with the fact that TEP production accounts for some 5–10% of the NPP (Mari et al., 2017). The ratio between the number of TEP clusters and the number of PFT clusters is a balance that matters for coagulation to develop correctly. The number of primary particles in each TEP cluster, N_{pp}, is just the cluster biomass divided by the carbon content of a TEP, which is a random value from $[\theta_{\mathrm{C}}^{\max,\,\mathrm{TEP}},\ \theta_{\mathrm{C}}^{\min,\,\mathrm{TEP}}]$. There are no reported values in the literature for the carbon content of a TEP, so my choice for the values of $\theta_{\mathrm{C}}^{\max,\,\mathrm{TEP}}$ and

$\theta_{\mathrm{C}}^{\mathrm{min,\,TEP}}$ are such that allow modelled TEPs to stay in Mari (1999)'s carbon content range ($0.25 \times 10^{-6} r_{\mathrm{p}}^{2.55}$ μg C (TEP)$^{-1}$) and Passow (2002)'s size range (1–100 μm). The TEPs are exuded at the depths where the selected PFT clusters are or, if there are less than two PFTs exuding, the depth is probabilistically chosen.

3.4.3 Dust deposition of clay particles

The main origin of marine aeolian dust inputs are the deserts, being North Africa, the Middle East and Southeast America the most active sources (Schulz et al., 2012). Dust is composed of a mixture of minerals such as quartz, carbonates, soluble salts and aluminosilicates (feldspars and clays). The relative composition is variable, depending on the source region, but clay, and in particular kaolinite, usually dominates the bulk of the total aerosol mass (Journet et al., 2008).

In SLAMS-2.0, all the aeolian-borne dust particles are clays of type kaolinite ($Al_2O_3 \cdot 2\,SiO_2 \cdot 2\,H_2O$). These terrigenous particles are inert as they do not interact with the BCP module (e.g., there is no dissolution of dust and it does not provide micronutrients, like iron for phytoplankton growth). Their only function is to provide ballast for the sinking rain of particles. To minimise the influence of clusters of dust type in the model, dust particles are seeded every 10 days all in one cluster. For that, I accumulate the aeolian dust flux deposition, F_{dust} (g dust m^{-2} s^{-1}), that has been forced into the model in that ten-day period into the 10th day. The amount of primary particles in the dust cluster, N_{pp}, is the dust biomass divided by the clay particle quota, θ_{clay}, with a value of 0.10 pmol clay (particle)$^{-1}$ to keep dust particle diameter at ~2.5 μm, following reported equivalent spherical diameter (ESD) values of 0.1–20 μm (Schulz et al., 2012).

3.4.4 Particle coagulation

Small primary particles coagulate to form larger ones following the rules of coagulation theory. Three physical transport mechanisms (Brownian motion, water shear and differential settling; Fig 3.7) bring particles together, making *collision* possible, and a TEP-mediated mechanism sticks particles together, making *coalescence* possible.

The development of coagulation theory started with McCave (1984), who built a marine particle collision model. Later on, Jackson (1990) consolidated the main transport mechanisms by which marine particles collide and combined his collision model with phytoplankton growth kinetics to develop a two-state model. There, coagulation processes were unimportant at low phytoplankton concentrations but dominated at high concentrations, with a transition phase characterised by decreasing fluid shear and increasing phytoplankton size and stickiness. The mechanisms by which particles stuck to each other were still unknown but Hill (1992) realised that the inclusion of

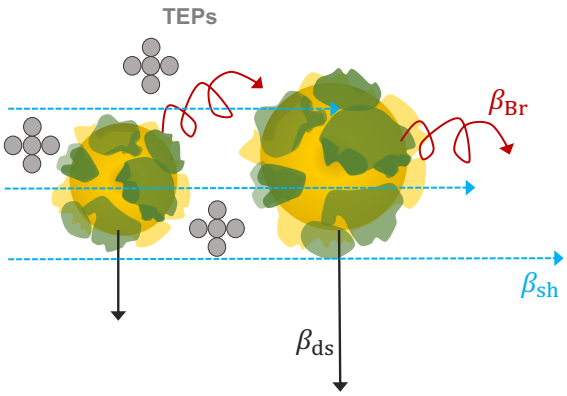


Figure 3.7. Schematic of the coagulation model used in SLAMS-2.0, with three particle collision mechanisms and a TEP-mediated coalescence. The collision mechanisms are: random thermal Brownian motion (β_{Br}), velocity gradients in the fluid flow (shear stress, β_{sh}), and differences in the settling velocities of individual particles (differential settling, β_{ds}). A single collision mechanism can end up controlling collision rates, for instance, if particles are too small, $< 1\ \mu$m (β_{Br} dominates) or if particles are too big (β_{ds} dominates).

abundant non-phytoplankton particles was needed to produce model results comparable to field observations of fluxes. Soon after, laboratory and fieldwork by Alldredge et al. (1993) proved the existence of sticky, non-phytoplankton particles responsible for particle aggregation, the TEPs. To date, marine coagulation theory is still developing, with three areas of critical importance: the incorporation of the fractal nature of aggregates in coagulation kernels, the development of a mechanistic model for particle disaggregation and the development of a mechanistic model for particle stickiness (Burd and Jackson, 2009).

SLAMS-2.0 uses a modified version of the particle collision algorithm of SLAMS-1.0, which was based on the formulation presented in Burd and Jackson (2009). The collision kernels used in standard coagulation theory (i.e., Burd and Jackson, 2009) have been theoretically derived for impermeable and non-fractal particles and proven to do not capture observed kernels in experiments where porous and fractal latex spheres are involved (Li and Logan, 1997a,b). Thus, in SLAMS-2.0 I have applied, instead, Jackson (2001)'s empirically derived rectilinear collision kernels from the experimental data of porous, fractal spheres of Li and Logan (1997a,b),

$$\beta_{\mathrm{Br}} = \frac{2}{3}\frac{k_{\mathrm{B}}T}{\mu}\frac{(r_{\mathrm{p}_i} + r_{\mathrm{p}_j})^2}{r_{\mathrm{p}_i}\, r_{\mathrm{p}_j}}, \tag{3.5a}$$

$$\beta_{\mathrm{sh}}^{\mathrm{lam}} = \frac{4}{3}\gamma(r_{\mathrm{p}_i} + r_{\mathrm{p}_j})^3 \quad (\text{laminar shear}, \mathrm{MAX}(r_{\mathrm{p}_i}, r_{\mathrm{p}_j}) < \eta), \tag{3.5b}$$

$$\beta_{\mathrm{sh}}^{\mathrm{turb}} = 1.3\gamma(r_{\mathrm{p}_i} + r_{\mathrm{p}_j})^3 q^{0.88} \quad (\text{turbulent shear}, \mathrm{MAX}(r_{\mathrm{p}_i}, r_{\mathrm{p}_j}) > \eta), \tag{3.5c}$$

$$\beta_{\mathrm{ds}} = \pi(r_{\mathrm{p}_i} + r_{\mathrm{p}_j})^2 \left|w_{\mathrm{p}_i} - w_{\mathrm{p}_j}\right| q^{0.984}, \tag{3.5d}$$

where the collision kernels β ($\mathrm{m^{-3}\,s^{-1}}$) represent the volume swept by a particle in a unit of time on its way to colliding into another particle; k_{B} ($\mathrm{J\,K^{-1}}$) is the Boltzmann constant; T (K) is water temperature; μ ($\mathrm{kg\,m^{-1}\,s^{-1}}$) is the dynamic viscosity of water; γ ($\mathrm{s^{-1}}$) is water shear rate; q is $\mathrm{MIN}(r_{\mathrm{p}_i}, r_{\mathrm{p}_j})/\mathrm{MAX}(r_{\mathrm{p}_i}, r_{\mathrm{p}_j})$; r_{p_i} and r_{p_j} (m) are the radii of

particles i and j; w_{p_i} and w_{p_j} ($\mathrm{m\,s^{-1}}$) are their corresponding sinking velocities; and η (m) is the Kolmogorov length scale. η is used to separate between turbulent shear (Eq. 3.5b) and laminar shear (Eq. 3.5c), and thus apply one of the two water shear collision kernels.

Particles will coalesce after collision only if, at least, one of the two particles involved in the collision has some degree of stickiness (i.e., $\alpha_{\mathrm{p}} > 0$), following SLAMS-1.0's formulation. With the three collision kernels added together and the collision efficiency given by the stickiest particle, the probability that the particles of one cluster will coagulate with the particles of the other cluster is

$$p_{\mathrm{coagu}} = (\beta_{\mathrm{Br}} + \beta_{\mathrm{sh}} + \beta_{\mathrm{ds}}) \times \mathrm{MAX}(\alpha_{\mathrm{p}_i}, \alpha_{\mathrm{p}_j}) \times \frac{\Delta t}{\Delta V}. \tag{3.6}$$

p_{coagu} is used to build the probabilistic model for cluster coagulation of Shima et al. (2009), the SDM, which proved of successful implementation in SLAMS-1.0. The SDM represents collision events between two clusters of many particles, which significantly speeds up the computation. It uses the MCM to randomly sample non-overlapping pairs of clusters that are evaluated for collision, and rescales Δt if necessary for particles to meet in the appropriate window time frame. Jokulsdottir and Archer (2016) developed a formulation based on the rules of conditional probability (their Eq. 17) to calculate the number of particles in a cluster involved in the collision with the particles of the other. That calculation operates inside the coagulation algorithm.

The probability of collision between two particles is thus a function of the concentration of particles (the more particles, the more clusters are sampled), their sizes, the difference between the sinking velocities of two particles, seawater temperature and the intensity of turbulent shear. The main advantage of the SDM, and what keeps it computationally stable, is that the number of cluster particles does not change as particles coagulate. After coagulation, the number of particles always decreases, but the number of clusters is unchanged thanks to the SDM, which keeps the simulation computationally stable. Mass and the number of primary particles are conserved in the process. It is all about the redistribution of primary particles.

3.4.5 Zooplankton

In SLAMS-2.0, I have separated the zooplankton community into micro- and mesosized zooplankton (i.e., 20–200 and 200–20,000 μm). Modelled mesozooplankton are representative of Calanoid copepods (metazoans), which are present in all marine environments and usually make up 70% by mass of all net-collected zooplankton (Lalli and Parsons, 1997). Modelled microzooplankton consist of protists (unicellular) and I have partitioned the assemblage into four major groups (Biard et al., 2016; Steinberg and Landry, 2017): radiolarians (silicified shelled amoebas), foraminifera (calcitic shelled amoebas), pteropods

(aragonitic snails) and zooflagellates (heterotrophic dinoflagellates and ciliates). Thus, SLAMS-2.0 models a total of five zooplankton functional types (ZFTs). Zooplankton, while in their living state, are not modelled as Lagrangian particles but as a *biomass*; when they die, the living zooplankton biomass is encased into Lagrangian particles. In Sec. 3.4.5.1 I describe the diel vertical migration (DVM) approach that I have adopted to distribute zooplankton biomass spatiotemporally in the water column. In Sec. 3.4.5.2 I describe how the zooplankton biomass encounters the raining flux of particles (prey items) according to two main feeding strategies, cruisers and ambushers. In Sec. 3.4.5.3 I describe how zooplankton ingest and digest the POC biomass encountered.

3.4.5.1 Zooplankton biomass model

The first step to model zooplankton biomass is to link it to the standing stock of POC biomass –the potential food items. Nevertheless, an empirical relationship between zooplankton biomass and POC biomass that holds globally has not been found yet. Thus, in SLAMS-2.0, I have adopted an approach that uses the rules of DVM to spatiotemporally distribute the biomass of zooplankton in the water column. Those rules are based on the trade-off between the gain from feeding and the risk of mortality following day/night time stepping and depth in the water column.

Backscattering observations (Bianchi et al., 2013; Behrenfeld et al., 2019) have shown that diel vertical migrants (mostly mesozooplankton) feed in the euphotic waters during night, when their fish predators are resting. After consuming phytoplankton and other non-migrating zooplankton, they migrate down into the twilight zone, just before daylight arrives, to escape predation. They will then metabolise the ingested surface POC at depth through respiration, egestion and excretion (Steinberg and Landry, 2017). In SLAMS-2.0, I have implemented the empirical formulation of Bianchi et al. (2013) to estimate the depth where zooplankton descend at daylight to escape predation and digest the food,

$$z_{\mathrm{DVM}}^{\mathrm{day}} = 398 - 0.56\,\Delta \mathrm{O}_2 - 115\,\log_{10}(\mathrm{chl}\,a) + 0.36\,\mathrm{MLD} - 2.4\,\Delta T, \tag{3.7}$$

where $z_{\mathrm{DVM}}^{\mathrm{day}}$ is in metres, $\Delta \mathrm{O}_2$ ($\mathrm{mmol\,m^{-3}}$) and ΔT (°C) are the O_2 and T gradients between the euphotic zone (0–25 m) and the twilight zone (150–500 m), and $\mathrm{chl}\,a$ ($\mathrm{mg\,m^{-3}}$) is the chlorophyll forced into the model.

In SLAMS-2.0, the DVM depth is not a discrete depth but a range of depths Δz_{DVM} (or migrating layer with upper and lower boundaries) where zooplankton are more abundant than anywhere else in the water column. Zooplankton abundance is modelled using a characteristic ratio to POC biomass, as explained in the next paragraph. Fig. 3.8 illustrates SLAMS-2.0's DVM algorithm, with four types of DVM depth boundaries encasing two Δz_{DVM}: the upper and lower boundaries of the night

DVM depth ($z_{\mathrm{DVM}}^{\mathrm{night,\ upp}}$ and $z_{\mathrm{DVM}}^{\mathrm{night,\ low}}$), and the upper and lower boundaries of the day DVM depth ($z_{\mathrm{DVM}}^{\mathrm{day,\ upp}}$ and $z_{\mathrm{DVM}}^{\mathrm{day,\ low}}$). I assume a thickness for Δz_{DVM} randomly chosen between 10 and 100 m (after Bianchi et al., 2013) and calculate the upper and lower boundaries around that thickness. The day DVM depth is $z_{\mathrm{DVM}}^{\mathrm{day}}$ (Eq. 3.7) whereas for the night DVM depth I consider the depth where there is maximum concentration of phytoplankton cells, z_{bounty},

$$z_{\mathrm{DVM}}^{\mathrm{night,\ upp}} = \mathrm{MAX}\left(z_{\mathrm{bounty}} - \frac{1}{2}\Delta z_{\mathrm{DVM}}^{\mathrm{night}},\, 0\right) \quad (\Delta z_{\mathrm{DVM}}^{\mathrm{night}} = rand[10, 100]), \tag{3.8a}$$

$$z_{\mathrm{DVM}}^{\mathrm{night,\ low}} = \mathrm{MIN}\left(z_{\mathrm{bounty}} + \frac{1}{2}\Delta z_{\mathrm{DVM}}^{\mathrm{night}},\, H\right), \tag{3.8b}$$

$$z_{\mathrm{DVM}}^{\mathrm{day,\ upp}} = \mathrm{MAX}\left(z_{\mathrm{DVM}}^{\mathrm{day}} - \frac{1}{2}\Delta z_{\mathrm{DVM}}^{\mathrm{day}},\, 0\right) \quad (\Delta z_{\mathrm{DVM}}^{\mathrm{day}} = rand[10, 100]), \tag{3.8c}$$

$$z_{\mathrm{DVM}}^{\mathrm{day,\ low}} = \mathrm{MIN}\left(z_{\mathrm{DVM}}^{\mathrm{day}} + \frac{1}{2}\Delta z_{\mathrm{DVM}}^{\mathrm{day}}.\, H\right), \tag{3.8d}$$

Each pair of DVM boundaries differentiates three sections in the water column (Fig. 3.8), each of them with distinct POC to zooplankton ratios: $\theta_{\mathrm{POC:Z}}^{\mathrm{DVM\,night}} = 0.50$ (↑Z:POC), $\theta_{\mathrm{POC:Z}}^{\mathrm{DVM\,day}} = 1.5$ (≃Z:POC) and $\theta_{\mathrm{POC:Z}} = 3.0$ elsewhere (↓Z:POC). Those ratios are multiplied by the standing stock of POC biomass at a specific depth to transform it into zooplankton biomass, B_{zoo} (mol C). Notice that there can only be zooplankton (i.e., $B_{\mathrm{zoo}} > 0$) when POC "food" is present. The value of the POC:zooplankton ratio is not a trivial one. Buitenhuis et al. (2013) show that observed global autotrophic (picophytoplankton, diazotrophs, coccolithophores, diatoms, phaeocystis) to heterotrophic (bacteria, micro-, meso-, and macrozooplankton, pteropods and foraminifera) ratio in the ocean is ~1:1, and the depth profiles of the biomass of phytoplankton (diatoms+picophytoplankton+coccolithophores) would be ~3 times those of the biomass of zooplankton (microzooplankton+mesozooplankton). In SLAMS-2.0, I use that fraction, 3:1, as the bottom line for the POC:zooplankton ratio (i.e., $\theta_{\mathrm{POC:Z}}$; from that ratio, I derive the other two indicated above). However, considering that POC matter is more than just phytoplankton matter, the actual ratio could be higher.

Each time step, all the estimated B_{zoo} at a particular depth is assigned to one of the five ZFTs probabilistically, according to the proportion of large vs small phytoplankton. Thus, if $(p^{\mathrm{diat}} + p^{\mathrm{dino}}) \geq 50\%$, B_{zoo} has a higher chance of being of type mesozooplankton than of microzooplankton. However, if microzooplankton type is chosen, a further decision has to be made favouring one of four microzooplankton types (radiolaria, foraminifera, pteropods or zooflagellates). Each type has an assigned *a priori* probability (p^{radiol}, p^{foram}, p^{pterop} and p^{flagel}) which will be *a posteriori* manipulated according to the MCM and the availability of $SiOH_4$ and CO_3^{2-}, the nutrients needed to form the mineral parts of silicifying and calcifying microzooplankton. If the concentrations of such nutrients at $z \leq 1000$ m are zero, the probabilities of microzooplankton with mineral parts are assigned a 0.

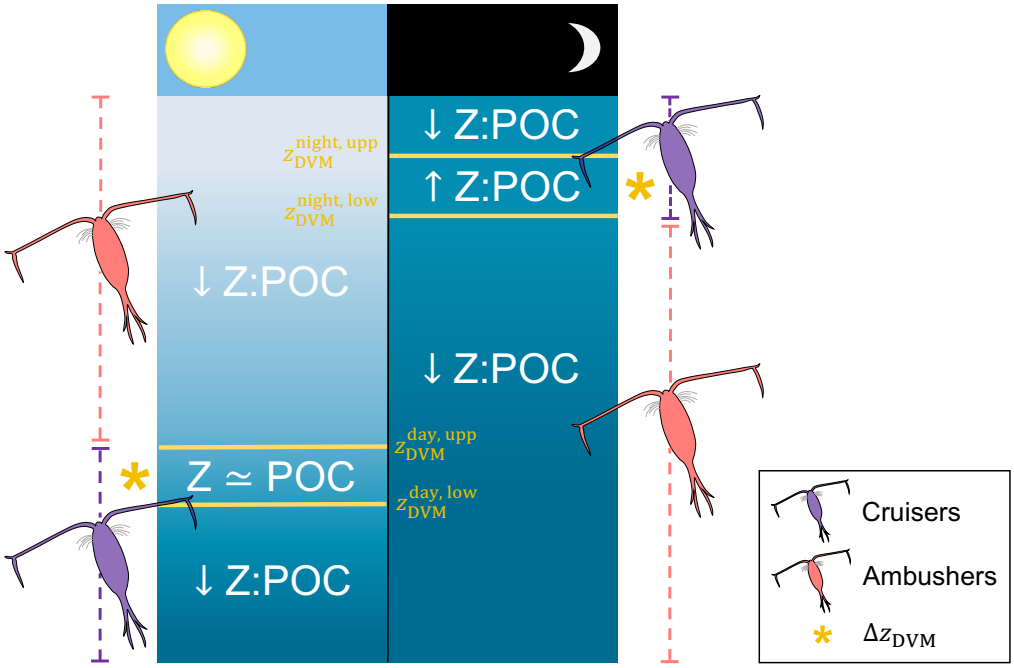


Figure 3.8. Schematic of the zooplankton diel vertical migration (DVM) model used in SLAMS-2.0 to compute zooplankton biomass and zooplankton-particle encounter kernels. Four boundaries are represented: the upper and lower boundaries of the night DVM depth ($z_{\mathrm{DVM}}^{\mathrm{night,\ upp}}$ and $z_{\mathrm{DVM}}^{\mathrm{night,\ low}}$), and the upper and lower boundaries of the day DVM depth ($z_{\mathrm{DVM}}^{\mathrm{day,\ upp}}$ and $z_{\mathrm{DVM}}^{\mathrm{day,\ low}}$). Those differentiate three regions in the water column at day and night: the DVM depth layer (Δz_{DVM}), the region above Δz_{DVM} and the region below Δz_{DVM}, with a zooplankton biomass (Z) that is characteristically smaller ($\downarrow$), higher ($\uparrow$) or slightly equal ($\simeq$) to particulate organic carbon (POC) biomass, following the DVM rules based on the trade-off between gain from feeding and risk of mortality. Two zooplankton feeding strategies that show vertical structure are represented to calculate zooplankton-particle encounter kernels: cruisers, or active zooplankton that search prey while swimming, and ambushers, or passive zooplankton that remain motionless waiting for their prey. During the day, cruisers hide at depth to reduce their risk of mortality –cruising is a behaviour that increases conspicuousness to fish predators. During the night, cruisers are back in the surface waters to feed while their predators are less able to detect them in the darkness. Ambushers do not display migrating behaviour; they are spread throughout the water column and are second to abundance when they share a water region with cruisers.

3.4.5.2 Zooplankton encounter model

In the same way that coagulation occurs when two particles encounter each other, zooplankton grazing is only possible if there is an encounter between a zooplankton individual (predator) and a particle (the prey). The calculation of the encounter probability between a prey particle and a zooplankton predator requires B_{zoo} to be encased into a number of zooplankton organisms, N_{zoo}, that have a defined radius, r_{zoo}, and a swimming speed, v_{zoo}. The equations to characterise those zooplankton variables are in Appendix C.3 and are based on the empirical formulations of Kiørboe (2013), Brun et al. (2019) and Pinti et al. (2019), respectively.

Encounters in SLAMS-2.0 are viewed from the perspective of the particles (the prey). Thus, the encounter probability (p_{enc}) that needs to be calculated is that of a particle

being encountered by a zooplankton individual, and not the other way around.There are two main encounter mechanisms between a particle and a zooplankton individual in SLAMS-2.0 (after Kiørboe, 1997): behavioural locomotion, β_{be}, and the physical transport of water shear, β_{sh}. Mirroring the formulation of p_{coagu} (Eq. 3.6),

$$p_{\mathrm{enc}} = (\beta_{\mathrm{be}} + \beta_{\mathrm{sh}}) \times N_{\mathrm{zoo}} \times \frac{\Delta t}{\Delta V}, \tag{3.9}$$

where $(\beta_{\mathrm{be}} + \beta_{\mathrm{sh}})$ $(\mathrm{m^3\,s^{-1}})$ represents the volume of water ΔV that a particle navigates in a time step Δt on its way to encountering zooplankton, and N_{zoo} is the number of zooplankton individuals in the grid box (calculated after encasing B_{zoo} into a number of zooplankters according to a zooplankton carbon quota, calculated in Appendix C.3).

The calculation of β_{be} depends on the zooplankton locomotion (or feeding) behaviour, with two main strategies distinguished: cruisers vs ambushers (Kiørboe, 2011), which switch according to depth in the water column and day/night time stepping (Fig. 3.8). Cruisers are those predators that search prey while actively swimming (e.g., copepods and zooflagellates). Since a cruising behaviour increases conspicuousness to predators (Ohman and Romagnan, 2016; Prowe et al., 2019), to reduce the risk of detection and capture by visually hunting predators, like planktivorous fish, cruisers migrate at depth at dawn. Thus, during the day they are mostly found at depth, in dark waters below $z_{\mathrm{DVM}}^{\mathrm{day,\ upp}}$. At night, cruisers migrate upwards to surface waters to feed and are mostly found at depths above $z_{\mathrm{DVM}}^{\mathrm{night,\ low}}$. Ambushers are motionless zooplankton that wait to intercept a prey (i.e., foraminifera, radiolarians and pteropods). They do not display migrating behaviour as cruisers do, and thus constitute a resident community. Their abundance is second to that of cruisers when both groups are sharing a region in the water column. The encounter kernels presented in Eq. 3.9 are (after Kiørboe, 1997):

$$\beta_{\mathrm{be}}^{\mathrm{crui}} = \pi\left(r_{\mathrm{detect}}^{i} + r_{\mathrm{p}}\right)^2 v_{\mathrm{zoo}}^{i} \quad \begin{array}{l}\text{(cruisers, if } \mathrm{PAR_0} > 0 \text{ and } z > z_{\mathrm{DVM}}^{\mathrm{day,\ upp}}\text{, or}\\ \qquad\qquad \text{if } \mathrm{PAR_0} = 0 \text{ and } z \le z_{\mathrm{DVM}}^{\mathrm{night,\ low}}\text{)}\end{array} \tag{3.10a}$$

$$\beta_{\mathrm{be}}^{\mathrm{amb}} = \pi\left(r_{\mathrm{detect}}^{i} + r_{\mathrm{p}}\right)^2 w_{\mathrm{p}} \quad \text{(ambushers, otherwise)}, \tag{3.10b}$$

$$\beta_{\mathrm{sh}}^{\mathrm{lam}} = \frac{4}{3}\gamma\left(r_{\mathrm{detect}}^{i} + r_{\mathrm{p}}\right)^3 \quad \text{(laminar shear, } \left(r_{\mathrm{detect}}^{i} + r_{\mathrm{p}}\right) < \eta\text{)}, \tag{3.10c}$$

$$\beta_{\mathrm{sh}}^{\mathrm{turb}} = 1.3\gamma\left(r_{\mathrm{detect}}^{i} + r_{\mathrm{p}}\right)^3 \quad \text{(turbulent shear, } \left(r_{\mathrm{detect}}^{i} + r_{\mathrm{p}}\right) > \eta\text{)}, \tag{3.10d}$$

where $(r_{\mathrm{detect}}^{i} + r_{\mathrm{p}})$ (m) is called the reaction distance and is the sum of the detection radius of the animal (i = Mz or μz) and the radius of the particle; $r_{\mathrm{detect}}^{i} = R_i\, r_{\mathrm{i}}$, where R_i is a detection radius factor that accounts for zooplankton bodily extensions or chemical cues (in the parameter list) and r_{i} is from Eqs. C.19b and C.20b; v_{zoo}^{i} $(\mathrm{m\,s^{-1}})$ is the zooplankton swimming speed (Eq. C.22). Notice that $\beta_{\mathrm{be}}^{\mathrm{crui}} > \beta_{\mathrm{be}}^{\mathrm{amb}}$ because the zooplankton swimming speed is higher than the sinking velocity of a particle, which is neglected in Eq. 3.10a.

In summary, SLAMS-2.0 uses the DVM rules that reduce the risk of mortality to allocate zooplankton biomass in the water column. The migrating vs non-migrating behaviour in SLAMS-2.0 is given by two locomotion (or feeding) strategies, cruisers (migrants) and ambushers (non-migrants).

3.4.5.3 Zooplankton grazing

All the particle clusters in the grid box are evaluated individually for encountering zooplankton, where the calculated p_{enc} is compared to a random number (*rand*) from a homogeneous distribution [0, 1]. If $p_{\text{enc}} > rand$, the particle is encountered and evaluated for ingestion depending on the particle's OM content. To prevent that unpalatable food particles, with a low OM content, are ingested, only particle clusters with a fraction of OM $> f_{\text{OM}}^{\text{ing, surf}}$ at depths < 1000 m and OM $< f_{\text{OM}}^{\text{ing, deep}}$ at depths > 1000 m are ingested.

In the absence of a mechanistic model to treat fragmentation induced by the action of zooplankton swimming-induced currents, SLAMS-1.0 used an empirical formulation based on the observations of Goldthwait et al. (2004). In that formulation, the number of daughter particles a parent particle fragments into is based on the radius of the particle. In SLAMS-2.0, though, I have decided to treat the process of fragmentation by zooplankton parametrically to test the effect of a single controlling factor (a parameter, instead of a range of particle sizes) on the model output. A fraction, f_{sloppy}, of the total mass of the intercepted particle cluster is fragmented and released as dissolved material (DOC + DIC + $SiOH_4$) as the zooplankton handle it in a sloppy manner with their appendages, rather than fragmenting it through swimming-induced currents into daughter particles.

The amount of particle cluster ingested will be metabolised by zooplankton in five steps (Fig. 3.9): (1) dissolution in the zooplankton guts of the $CaCO_3$ present in the ingested food, releasing DIC; (2) egestion of a fraction of the ingested material that is not absorbed, packaging it into faecal pellet clusters; (3) excretion of a fraction of the absorbed material that is not assimilated, releasing DOC, DIC and $SiOH_4$; (4) respiration of the assimilated organic material that is not transformed into net secondary production (NSP), remineralising POC into CO_2, and (5) death, where dead biomass is encased into ZFTs clusters.

The ingested cluster is not treated as discrete particles anymore but as food biomass. Since digestion is primarily an acidic process, there is dissolution of the ingested $CaCO_3$ food biomass (B_{CaCO_3}) in the digestive system of a zooplankton organism, which is released as DIC (B_{DIC}). For the dissolution of $CaCO_3$ in the zooplankton digestive system, I use the model of Jansen and Wolf-Gladrow (2001) in which dissolution limitation factors for the two forms of $CaCO_3$, L_{calc} and L_{arag}, are calculated. Those two depend on the

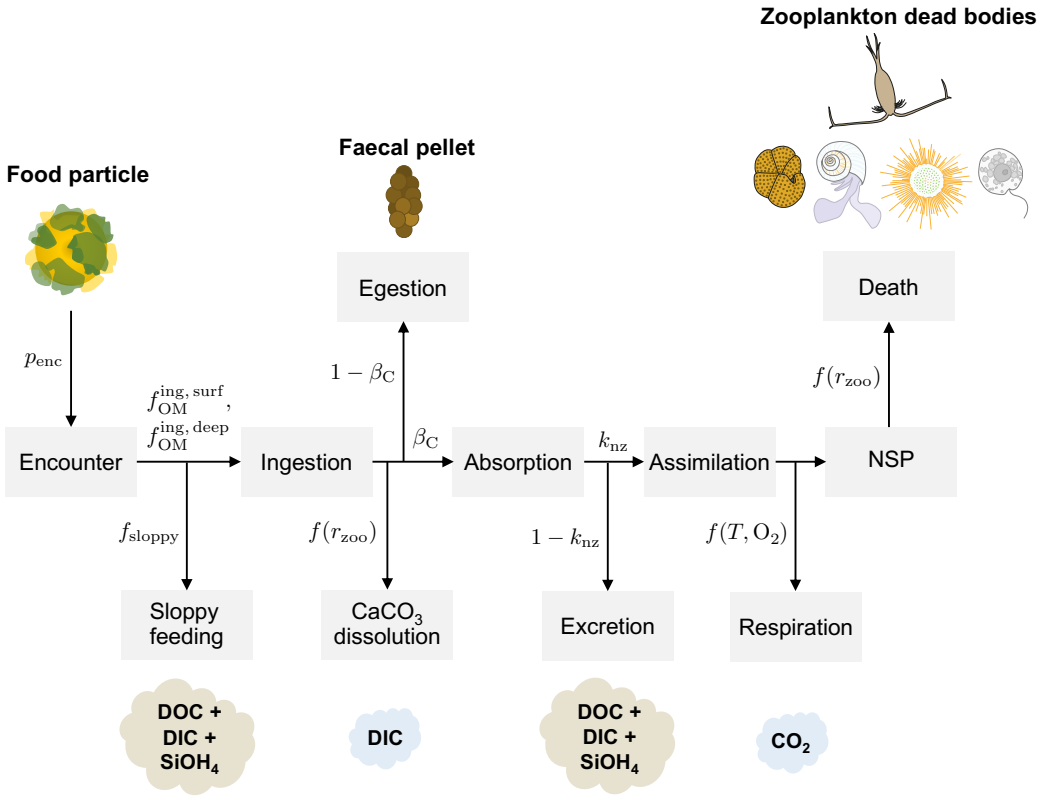


Figure 3.9. Flow chart of the zooplankton grazing model used in SLAMS-2.0. p_{enc} is the probability of particle-zooplankton encounter; $f_{OM}^{ing,surf}$ and $f_{OM}^{ing,deep}$ are the fractions of organic matter (OM) in a surface ocean and deep ocean food particle, respectively; f_{sloppy} is the fraction of intercepted material fragmented due to sloppy handling of the food particle; r_{zoo} is the radius of a zooplankton organism; β_C is the absorption efficiency of carbon (C); k_{nz} is the C-specific net growth efficiency; T is seawater temperature; O_2 is oxygen concentration. Ingestion is the process of intake of food into the feeding apparatus. Absorption is the process of diffusion of food from the gut/vacuole to the circulatory system. Egestion refers to the production of faecal pellets. Excretion refers to the production of urine products. Assimilation is the incorporation of matter into tissues as gross secondary production. After respiration takes place, the assimilated biomass is considered net secondary production (NSP). Encounters depend on the zooplankter size (meso- vs microzooplankton), so it is the production of faecal pellets. At each time step and each depth layer, zooplankton dead bodies can be one of five main zooplankton functional types (ZFTs): copepods (mesozooplankton), foraminifera, pteropods, radiolarians or zooflagellates (the latter four, microzooplankton). ZFT is decided probabilistically upon the abundance of large vs small phytoplankton functional types.

magnitude of $CaCO_3$ undersaturation in the zooplankton digestive systems (guts for mesozooplankton, food vacuoles for microzooplankton):

$$L_{calc} = \begin{cases} \left(1 - \frac{[CO_3^{2-}]_{zoo}}{[CO_3^{2-}]_{sat}^{calc}}\right)^{\eta_{calc}} & \text{if } [CO_3^{2-}]_{zoo} < [CO_3^{2-}]_{sat}^{calc}, \\ 0 & \text{if } [CO_3^{2-}]_{zoo} \geq [CO_3^{2-}]_{sat}^{calc}, \end{cases} \tag{3.11a}$$

$$L_{arag} = \begin{cases} \left(1 - \frac{[CO_3^{2-}]_{zoo}}{[CO_3^{2-}]_{sat}^{arag}}\right)^{\eta_{arag}} & \text{if } [CO_3^{2-}]_{zoo} < [CO_3^{2-}]_{sat}^{arag}, \\ 0 & \text{if } [CO_3^{2-}]_{zoo} \geq [CO_3^{2-}]_{sat}^{arag}, \end{cases} \tag{3.11b}$$

where L_{calc} and L_{arag} are the fractions of carbonate dissolving per minute in the

zooplankton digestive systems; $[CO_3^{2-}]_{sat}^{calc}$ and $[CO_3^{2-}]_{sat}^{arag}$ ($mmol\,m^{-3}$) are the calcite and aragonite saturation concentrations, for which a surface ocean (100 m) value has been assumed (42.7 $mmol\,m^{-3}$ after Jansen, 2001); $[CO_3^{2-}]_{zoo}$ is the concentration of CO_3^{2-} in the zooplankton digestive system, calculated as the quotient between the amount of CO_3^{2-} in the ingested food biomass and the volume of the gut or vacuole ($V_{gut} = \frac{4}{3}\pi r_{Mz}^3$ and $V_{vacu} = \frac{4}{3}\pi r_{\mu z}^3$, where digestive system volumes are assumed to equal the volume of the zooplankton organism); and η_{calc} and η_{arag} are the calcite and aragonite dissolution order reactions. When the digestive system is supersaturated with respect to $CaCO_3$ ($[CO_3^{2-}]_{zoo} \geq [CO_3^{2-}]_{sat}^{calc}$), no dissolution occurs, and when the digestive system is undersaturated with respect to $CaCO_3$ ($[CO_3^{2-}]_{zoo} < [CO_3^{2-}]_{sat}^{calc}$), dissolution is possible.

Finally, the released DIC biomass (B_{DIC}) is the analytical solution of the differential equation $dB_{DIC}/dt = (\kappa_{CaCO_3} L_{CaCO_3} + 1/\phi^i) \times B_{CaCO_3}$ (Jansen and Wolf-Gladrow, 2001, Eq. 1), which is an exponential equation,

$$\begin{aligned} B_{DIC} = {} & B_{calc} - B_{calc} \exp\left(\left(-\kappa_{calc} L_{calc} + \frac{1}{\phi^i}\right)\Delta t\right) \\ & + B_{arag} - B_{arag} \exp\left(\left(-\kappa_{arag} L_{arag} + \frac{1}{\phi^i}\right)\Delta t\right), \end{aligned} \tag{3.12}$$

where κ_{calc} and κ_{arag} (s) are the calcite and aragonite dissolution timescales and ϕ^i (s) is the gut passage time (i = Mz or μz).

The fraction of $CaCO_3$ not dissolved in the zooplankton digestive system will be allocated into faecal pellets as mineral components cannot be absorbed, only organic material. The fraction of matter doomed for egestion, $B_{egest} = (B_{ing} - B_{DIC}) \times (1 - \beta_C)$, is packaged into faecal pellet particles. A faecal (pellet) particle in SLAMS-2.0 is usually an aggregate of N_{fpp} faecal primary particles with a porosity between 40–70% ($[P_{fp}^{min}, P_{fp}^{max}]$, in the parameter list) (Ploug et al., 2008). A faecal cluster is created per food cluster ingested, with a number of faecal particles N_{fp}.

To decide the number N_{fp} into which the egested biomass B_{egest} will be split into, I have designed an algorithm that keeps the size of faecal particles proportional to the size of the food particles ingested. The method uses the bulk volume of the faecal particle, V_{fp}^B, to scale the material content of the ingested particle –I use the solid volume and not, for instance, the radius because the radius implies already a porosity, which might be very different in the ingested vs faecal particles. I calculate V_{fp}^B (μm^3) as a function of the prosome length, PL (mm), of the zooplankton individual, following Stamieszkin et al. (2015) empirical formulation,

$$V_{fp}^B = 10^{(5.4 + 2.58 \times \log_{10}(PL))}, \tag{3.13}$$

where PL is calculated in Appendix C.3 using a temperature–PL empirical relationship. I use $V_{\mathrm{fp}}^{\mathrm{B}}$ to calculate the ratio between the bulk volume of the faecal pellet and the solid volume of the particle ingested, as $q_{\mathrm{vol}} = V_{\mathrm{fp}}^{\mathrm{B}}/V_{\mathrm{p}}^{\mathrm{S}}$. That value is used to scale up or down the ingested particle's molar content (as $q_{\mathrm{vol}} \sum_i n_{\mathrm{p}}^i$), mass (as $q_{\mathrm{vol}}\, m_{\mathrm{p}}$) and solid volume (as $q_{\mathrm{vol}}\, V_{\mathrm{p}}^{\mathrm{S}}$). Those scaled quantities become the potential attributes of the new faecal particle (m_{fp}, $\sum_i n_{\mathrm{fp}}^i$ and $V_{\mathrm{fp}}^{\mathrm{S}}$). Next I calculate the number of faecal particles in the faecal cluster as $N_{\mathrm{fp}} = (N_{\mathrm{p}}\, m_{\mathrm{p}})/m_{\mathrm{fp}}$, where the numerator is the total mass of the ingested cluster of particles and the denominator is the mass of one faecal particle (notice: particle, not cluster). The last attribute needed to constitute the faecal particle is N_{fpp}, which through fractal scaling law can compute all the other faecal attributes left to be known. Two cases arise.

- Case 1: $N_{\mathrm{fp}} > 1$, the total mass of the ingested cluster is larger than the mass of a single faecal particle. In that case, I split the ingested cluster mass into faecal particles according to the calculated $V_{\mathrm{fp}}^{\mathrm{B}}$, $V_{\mathrm{fp}}^{\mathrm{S}}$ and porosity chosen from a range $[P_{\mathrm{fp}}^{\mathrm{min}}, P_{\mathrm{fp}}^{\mathrm{max}}]$ modified according to mineral content (higher mineral content favours more porous structures). Those three attributes can compute N_{fpp} using fractal scaling.
- Case 2: $N_{\mathrm{fp}} \leq 1$, the mass of a single faecal particle is larger than the total mass of the ingested cluster. In that case, it is necessary to scale down the original target volume of the faecal particle, $V_{\mathrm{fp}}^{\mathrm{B}}$, until $N_{\mathrm{p}}\, m_{\mathrm{p}} = N_{\mathrm{fp}}\, m_{\mathrm{fp}}$. For that, I scale down $V_{\mathrm{fp}}^{\mathrm{S}}$ by factors of 10, and when the mass balance is achieved, I am ready to compute the faecal particle attributes following Case 1.

After putting aside a fraction $(1 - \beta_{\mathrm{C}})$ for egestion and a fraction $(1 - k_{\mathrm{nz}})$ for excretion (Fig. 3.9), I calculate the amount of assimilated POC food biomass (B_{POC}) that will be respired and remineralised into CO_2 (B_{CO_2}). For that, I calculate first the zooplankton respiration rate, λ, which follows Q_{10} kinetics for temperature limitation and Michaelis–Menten/Monod kinetics for O_2 limitation (after Cram et al., 2018),

$$\lambda^i = \mathrm{MIN}\left(\lambda_0^{\max,i}\, Q_{10}^{i\,\frac{T}{10}},\, \lambda^{\max,i}\right) \times \frac{[O_2]}{[O_2] + k_{O_2}}, \tag{3.14}$$

where $\lambda_0^{\max,i}$ (d^{-1}) is the maximum zooplankton C-specific respiration rate at 0°C, $\lambda^{\max,i}$ (d^{-1}) is a threshold zooplankton C-specific respiration rate, and k_{O_2} ($\mathrm{mL\,L^{-1}}$) is the half-saturation constant for O_2 uptake. B_{POC} is the analytical solution of the differential equation $\mathrm{d}B_{CO_2}/\mathrm{d}t = \lambda\, B_{\mathrm{POC}}$, which is an exponential equation,

$$B_{CO_2} = B_{\mathrm{POC}} - B_{\mathrm{POC}}\, \exp\left(-\lambda^i\, \Delta t\right). \tag{3.15}$$

In SLAMS-2.0, zooplankton die due to natural causes only and are encased into a new type of primary particles: the dead zooplankton bodies (Fig. 3.3). The calculation of the dying zooplankton biomass, $B_{\text{zoo}}^{\text{dead}}$, happens when all particles in the depth layer have been evaluated for ingestion and only if a fraction of the ingested biomass has been assimilated into NSP,

$$B_{\text{zoo}}^{\text{dead}} = B_{\text{zoo}} - B_{\text{zoo}} \exp\left(-m^i \, \Delta t\right). \tag{3.16}$$

$B_{\text{zoo}}^{\text{dead}}$ is encased into N_{pp} primary particle dead bodies according to the zooplankton carbon body quota, i.e., $N_{\text{pp}} = B_{\text{zoo}}^{\text{dead}} / \theta_{\text{C}}^{i}$ (θ_{C}^{i} is the zooplankton body C quota, computed in Appendix C.3). The ZFT of the N_{pp} formed had already been assigned at the moment of forming B_{zoo}.

3.4.6 Particle physical breakup and photolysis

Particle breakup happens in SLAMS-2.0 if an aggregate meets one of two criteria: (a) it has an unstable structure because it is poorly sticky (mathematically executed as $\alpha_{\text{p}} < 10^{-20}$ and $rand > 0.99$) and/or (b) it is sinking through a turbulent environment ($r_{\text{p}} > \eta$). There is still no mechanistic formulation to model particle physical disaggregation. However, it is understood (Goldthwait et al., 2004) that it depends on (i) the size of the particle, where the number of fragments increases exponentially with increasing aggregate size, and (ii) particle lability, where less labile (or more fragile) aggregates tend to fragment more readily and into more fragments than more labile ones. Following those two premises, I have developed a mechanistic formulation for aggregate breakup in SLAMS-2.0. The number of fragments, or daughter particles, in which an aggregate will break up, N_{fr}, is a function of r_{p} and the mineral fraction in the particle with respect to (C_{org} + TEC), the proxy to particle fragility (i.e., higher mineral content destabilises particles as mineral spicules and shells make the structure more open). Following Goldthwait et al., 2004, Fig. 5, the function can be modelled with a power-law shape equivalent to

$$N_{\text{fr}} = 2\, r_{\text{p}}^{\left(\frac{\sum_i n_i}{n_{\text{Corg}} + n_{\text{TEC}} + \sum_i n_i}\right)}, \quad i = \text{calc, arag, bSi \& clay}, \tag{3.17}$$

and the particle is fragmented if the product $N_{\text{fr}} \times rand > 2$, a threshold value to avoid excessive particle breakup. It follows that the parent aggregate must have an amount of primary particles $N_{\text{pp}} \geq N_{\text{fr}}$ to be able to disintegrate.

Photolysis is applied to particles that meet the following three criteria: (a) are in the upper 1 m of the water column, and (b) are big enough to have a weak structure ($r_{\text{p}} > 10$ mm), and (c) are positively buoyant ($\rho_{\text{p}} < \rho_{\text{w}}$). Photolysis breaks those particles into their dissolved material components. Photolysis rarely happens in SLAMS-2.0 as not

many particles meet the criterion of being bigger than 10 mm (quite a large number for marine aggregates). The proposed mathematical realisation for photolysis, inherited from SLAMS-1.0, is not a mechanisation of photolysis, but more of a fudge function of what is expected to happen to particles that will never sink and that are in direct contact with surface ocean radiation. The bar on particle size for photolysis has been set high to do not immediately photolyse all particles in the surface ocean that become relatively large, so to give them some more lifetime to interact with the surrounding particles, offering them an opportunity to sink if their density becomes larger than that of seawater. Thus, the large particle size is a proxy for photolytic lifetime, retarding degradation by light.

3.4.7 Bacteria metabolism

Around 50% of the POC produced by primary production is converted into DOC by processes such as zooplankton excretion, zooplankton fragmentation during sloppy feeding, viral lysis of phytoplankton cells and microbial solubilisation (Anderson and Tang, 2010). Many marine heterotrophic bacteria can produce sticky polysaccharides (Herndl and Reinthaler, 2013), which help attach them to the aggregates they feed on by producing hydrolytic enzymes to solubilise particulate material and use the generated DOC as a resource for growth and maintenance. In SLAMS-2.0, the metabolism of bacteria is represented with two processes: (i) respiration, which follows the formulation kinetics based on T and O_2 used in zooplankton respiration (Eqs. 3.15 and 3.15) and releases CO_2, and (ii) solubilisation of small particles ($r_p < 0.2$ μm, particles operationally defined as DOC), which releases DOC, DIC and $SiOH_4$. Those small particles will never collide with another particle as they became too small after remineralisation, disaggregation and dissolution processes. In SLAMS-2.0, I collect all the solubilised material belonging to small particles (in that depth layer, in that time step) and accumulate the organic carbon until I have put together enough to form a cluster of several sticky TEP particles.

3.4.8 Abiotic mineral dissolution

Abiotic mineral dissolution operates on both $CaCO_3$ and bSi (there is no dissolution of clay). The abiotic dissolution of $CaCO_3$ is homologous to that of zooplankton digestive system dissolution of $CaCO_3$ (biotic) using distinct dissolution limitation factors for calcite and aragonite, L_{calc} and L_{arag} (Eqs. 3.11a and 3.11b). In the case of abiotic dissolution, those depend on ambient saturation concentrations, which are functions of depth (or pressure). Following the empirical formulation in Jansen et al. (2002),

$$[CO_3^{2-}]_{sat}^{calc} = 88.7 \exp(0.189\,(z - 3.82)), \tag{3.18a}$$

$$[CO_3^{2-}]_{sat}^{arag} = 117.5 \exp(0.176\,(z - 3.06)), \tag{3.18b}$$

where z is depth in km and 3.82 km and 3.06 km are the saturation horizon depths for calcite and aragonite, respectively. Waters where $[CO_3^{2-}] > [CO_3^{2-}]_{sat}$ are supersaturated with respect to mineral $CaCO_3$ and no dissolution occurs; it is the case of surface waters. If $[CO_3^{2-}] < [CO_3^{2-}]_{sat}$, the water is undersaturated with respect to mineral $CaCO_3$ and dissolution occurs; it is the case of deep waters, which are corrosive to calcareous shells and skeletons.

The released DIC biomass (B_{DIC}) is also the analytical solution to the differential equation of dissolution kinetics,

$$\begin{aligned} B_{DIC} = {} & B_{calc} - B_{calc} \exp\left(-\kappa_{calc} L_{calc} \Delta t\right) \\ & + B_{arag} - B_{arag} \exp\left(-\kappa_{arag} L_{arag} \Delta t\right). \end{aligned} \tag{3.19}$$

The dissolution kinetics of bSi mirrors that of $CaCO_3$. Since most of the water column is so highly undersaturated in $SiOH_4$, i.e., $[SiOH_4] < [SiOH_4]_{sat}$ (Sarmiento and Gruber, 2006), bSi dissolution readily occurs and one can express the dissolution of opal directly without a limitation factor,

$$B_{SiOH_4} = B_{bSi} - B_{bSi} \exp\left(-\kappa_{opal} \Delta t\right), \tag{3.20}$$

where κ_{opal} (s) is the bSi dissolution timescale.

4 Data

Contents

4.1 Introduction

Climatological environmental data are needed to drive SLAMS-2.0's ecosystem functioning and to calibrate the model output. The *forcing data sets* used to run SLAMS-2.0 (NPP, chl a, PAR_0, NO_3^-, $SiOH_4$, CO_3^{2-}, O_2, T, MLD, F_{dust}) are of more conventional acquisition compared to the *calibration data sets* used to constrain its response (POC, PIC and bSi fluxes as well as particle number concentrations). Forcing data sets are already found in a globally gridded format and statistically processed. On the contrary, particle flux data are available for point locations in the ocean and are heterogeneous in time and space, making it difficult for building global climatologies. Furthermore, the collection

of particle flux data is not synchronised with closely related variables, like *in situ* NPP data, needed to force SLAMS-2.0 and calculate flux-related metrics.

A decision was made to calibrate SLAMS-2.0 only with particle flux data from long-term, fixed-point time series, which provide quasi climatological data on surface-to-depth profiles. That criterion has restricted my compilation of calibration data sets to five locations only: EqPac, in the upwelling equatorial Pacific (0°N, 140°W), HAUSGARTEN, in the Arctic Atlantic (79.0°N, 4.3°E), Ocean Station Papa, in the subarctic eastern Pacific (50°N, 145°W), Porcupine Abyssal Plain time-Series Observatory, in the blooming NE Atlantic (49°N, 16.5°W) and Bermuda Atlantic Time-Series/Oceanic Flux Program joint site, in the subtropical NW Atlantic (31.6°N, 64.2°W). This chapter provides the reference sources of the calibration and forcing data sets and addresses the problem of data choice and data aggregation (i.e., statistical treatment) to produce local quasi climatologies of particle fluxes, particle number concentrations and NPP.

4.2 Model forcing data sets

SLAMS-2.0 is driven by ten local forcing variables, introduced into the model every time step: PAR_0, the solar radiation in the range of 400–700 nm incident on the surface ocean; [chl *a*], the concentration of photosynthetic pigment chlorophyll *a* integrated for the euphotic layer; NPP, the production of phytoplankton biomass integrated for the euphotic layer; F_{dust}, the wind-borne, aeolian flux of dust deposited into the surface ocean; $[O_2]$, the seawater concentration of dissolved oxygen along a depth profile; T, the seawater temperature along a depth profile; $[NO_3^-]$ the seawater concentration of macronutrient nitrate along a depth profile; $[SiOH_4]$, the seawater concentration of macronutrient silicic acid along a depth profile; $[CO_3^{2-}]$, the seawater concentration of carbonate ion along a depth profile, and MLD, the mixed layer depth. These ten biogeochemical variables are based on oceanographic and satellite observations. Next I detail the sources and processing of the model forcing data sets and Fig. 4.1 shows the global distribution of their annual climatological means.

4.2.1 Light

Data sets of PAR_0 exist from satellite products, but those produce incomplete pictures above 60°N and below 60°S during the cold months of the year as the satellites cannot see through the cloud thickness and the winters have days without sunlight. Thus, I have created my own PAR_0 ($W\,m^{-2}$) product from the solar constant using standard trigonometric/astronomical equations. I have calculated PAR_0 with dependencies on: (i) latitude, (ii) day of the year (Julian day), (iii) the atmospheric properties of absorption

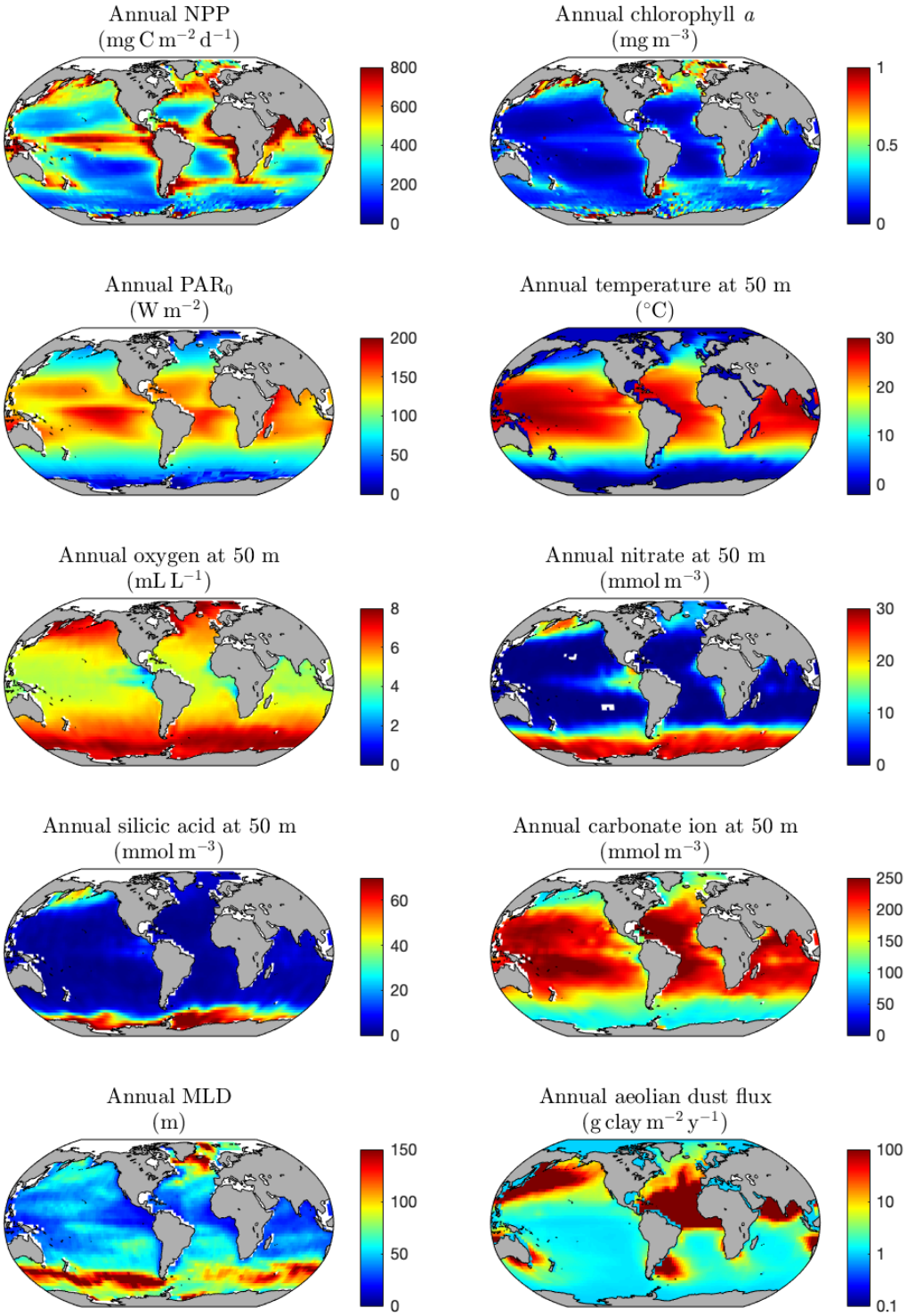


Figure 4.1. Global ocean distribution of the climatological annual mean of SLAMS-2.0's forcing variables at a spatial resolution of 1°×1°. NPP is a satellite estimate from Carr (2002), chlorophyll a is a MODIS-Aqua product, PAR$_0$ is from trigonometric/astronomical equations combined with input data of the fraction of clouds and ice covering the sea (obtained from CMIP model estimates), temperature, oxygen and macronutrients (nitrate, silicic acid) are from the World Ocean Atlas 2013 (WOA13), carbonate ion concentration is from the Global Ocean Data Analysis Project for Carbon (GLODAPv2), MLD is from the climatology of De Boyer Montégut et al. (2004) that uses the variable density criterion, and aeolian dust flux is a composite model estimate.

and scattering, (iv) the ocean surface reflectivity (albedo), (v) the amount of cloud and ice covering the sea (monthly climatologies from https://psl.noaa.gov/data/gridded/cmip3-clouds-rad-precip/ and composite model estimate borrowed from my supervisor, respectively), and (vi) the time of the day. For any modelled location in SLAMS-2.0, the daily cycle of day and night alternates with two time steps dedicated to the daytime period and one time step to the night period (16 h in daylight and 8 h in the darkness). PAR_0 (see calculation in Appendix D) is the average PAR_0 received per unit area during the daylight period of a particular Julian day.

4.2.2 Primary production

Traditionally, observations of NPP have been obtained *in situ* in shipboard incubation experiments of phytoplankton where cells would uptake radioactive inorganic ^{14}C through photosynthesis. But the elevated cost of oceanographic campaigns has produced a reduced set of *in situ* NPP observations, with poor spatial and temporal coverage, making them not the first choice for validating biogeochemical models (Carr et al., 2006). Luckily, alongside oceanographic campaigns, satellites have been recording worldwide, all-year-round ocean colour. The optical signal detected by ocean colour radiometers is influenced by the concentration of seawater components that interact with light, such as chl *a* in phytoplankton cells. That chl-based signal is transformed into NPP. Satellite-based estimates of NPP outnumber *in situ* measurements as they are available at synoptic scales. Today, there exist more than 20 satellite-based models that estimate NPP from ocean optics in combination with other remotely sensed properties, such as SST and PAR_0, and assumptions about the shape of the $P-I$ curve, the depth of NPP integration and the value of the photosynthetic parameters (see review of Carr et al., 2006).

I have compiled NPP data from four satellite-based models commonly used in the OBM literature: the Vertically Generalized Productivity Model (VGPM, Behrenfeld and Falkowski, 1997), the Carbon-based Productivity Model-2 (CbPM, Westberry et al., 2008), the Carbon, Absorption, and Fluorescence Euphotic-resolving model (CAFE model, Silsbe et al., 2016) and the Carr (2002) model. Those four models produce the spatial maps of annual averages of NPP shown in Fig. 4.2 after applying strategic interpolation to fill in those areas of the ocean not seen by the satellite when covered by clouds. The interpolation algorithm that I have designed operates over locations where there have been ≥ 2 months in the year when an NPP data value was available and interpolates linearly in time using the Matlab's function `interp1`. VGPM, CbPM and CAFE are driven with data products from the MODIS-Aqua satellite sensor (chl *a*, PAR_0 and SST) and are already available as finished data sets at the Ocean Productivity site (http://sites.science.oregonstate.edu/ocean.productivity/). Carr (2002)'s model is driven

with data products from a different satellite sensor, the SeaWiFS (chl a and PAR_0, but SST from AVHRR satellite sensors). I have recalculated the Carr (2002) model output using a carbon to chl a ratio of 50 g C $(g\,chl a)^{-1}$ (from Fasham et al., 1990, after value not reported in the Carr model), and input data from extended periods compared to the originally chosen by the author. Thus, chl a and PAR_0 are from SeaWiFs 9 km monthly climatology for the whole satellite lifespan, 1997–2010 (hhttps://oceandata.sci.gsfc.nasa.gov/SeaWiFS), SST is from AVHRR Pathfinder v5.0 global 4 km monthly climatology for 1985–2001 (https://www.ncei.noaa.gov/products/avhrr-pathfinder-sst) and seawater temperature to calculate the maximum photosynthetic rate at the base of the mixed layer is from WOA13 (https://www.nodc.noaa.gov/OC5/woa13/woa13data.html).

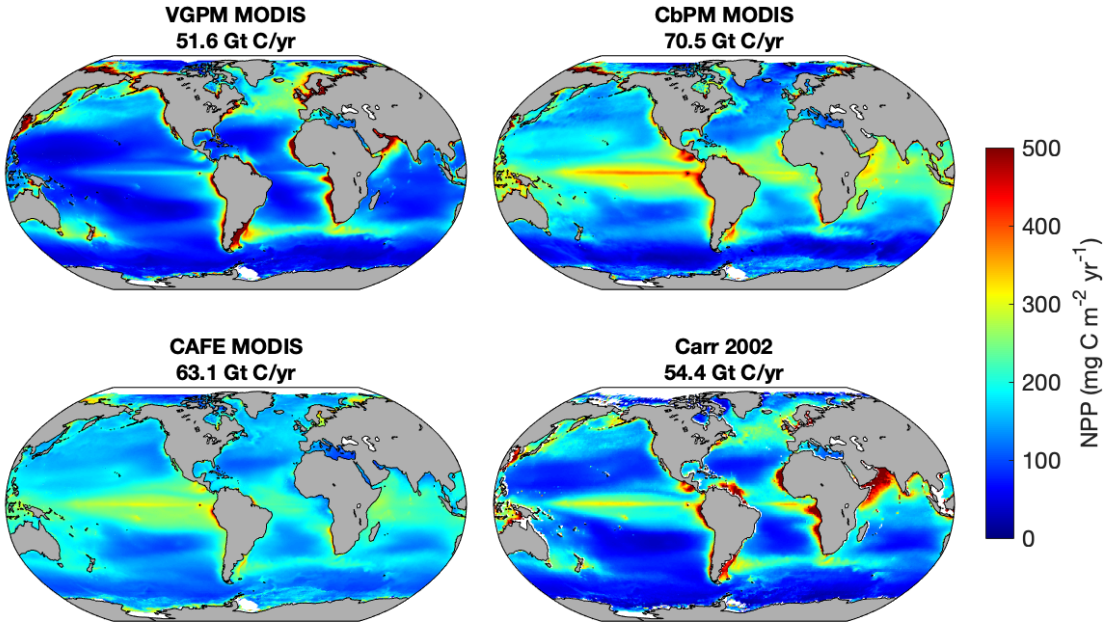


Figure 4.2. Global map of net primary production (NPP) estimated from four algorithms: (top left) the VGPM (Vertically Generalized Productivity Model) based on MODIS-Aqua products, (top right) the CbPM (Carbon-based Production Model-2) based on MODIS-Aqua products, (bottom left) the CAFE model based on MODIS-Aqua products and (bottom right) the Carr (2002) model based on SeaWiFs and AVHRR products. Printed on top of each map is the globally integrated average values of NPP for each model.

The VGPM and Carr (2002) model predict higher NPP in the northern North Atlantic and lower NPP in oligotrophic areas than the CbPM and CAFE models (Fig. 4.2). It has been already noted that the VGPM model has a large uncertainty in the northern North Atlantic when compared to *in situ* observations (Tilstone et al., 2009, 2015; Milutinovic and Bertino, 2011) due to errors in photosynthetic rate functions and irradiance-depth dependent functions. Per contra, the CAFE and CbPM model predict higher NPP estimations in the Equatorial upwelling region, tropical and subtropical oceans. Those four models differ in how NPP is globally distributed, and their globally integrated average

values (VGPM MODIS, 51.6 Gt C yr^{-1} < Carr (2002), 54.4 Gt C yr^{-1} < CAFE MODIS, 63.1 Gt C yr^{-1} < CbPM MODIS, 70.5 Gt C yr^{-1}) are not in all cases close to the global value of 50 Gt C yr^{-1} based on *in situ* data (Carr, 2002; Carr et al., 2006).

The NPP from Carr (2002) has been chosen to run SLAMS-2.0 due to closer resemblance with *in situ* NPP estimates at five ocean locations selected for model calibration (see Sec. 4.3.2 below). Considering that not all three time steps in a modelled day receive PAR_0, the NPP (mol C m^{-2} d−1) is distributed over the sunny seconds of a day, and thus inputted into SLAMS-2.0 in units of mol C m^{-2} s−1 for the time steps exposed to daylight. Another primary productivity variable used by SLAMS-2.0 is [chl*a*] (mg m^{-3}), which is from a monthly climatology from MODIS Aqua product 9 km available at NASA's Ocean Color Web (https://oceancolor.gsfc.nasa.gov/atbd/chlor_a/). I have averaged monthly data of chl *a* for the period 2002–2019 and have applied the same extrapolation criterion as per the NPP data set to the regions covered by clouds.

4.2.3 Nutrients

[NO_3^-] (mmol m^{-3}), [$SiOH_4$] (mmol m^{-3}) and [O_2] (mL L^{-1}) are from the World Ocean Atlas 2013 (WOA13), where macronutrients are from García et al. (2013b) and oxygen is from García et al. (2013a), all available at https://www.nodc.noaa.gov/OC5/woa13/woa13data.html. Since monthly climatologies are only available for upper water column depths (monthly O_2 data go down to 1500 m and monthly macronutrients only go down to 500 m), I have created a water depth climatology using monthly climatologies for upper water column values and annual climatologies for the lower water column values. Monthly climatological values are interpolated to the days of the month (as well as to the depth grid and latitude/longitude grid chosen) using Matlab's function `griddedInterpolant`, and then allocated equally for the three time steps of the day.

[CO_3^{2-}] (mmol m^{-3}) is an annual, depth-distributed global climatology calculated using a program developed for CO_2 system calculations, CO2SYS (Lewis and Wallace, 1998; Van Heuven et al., 2011), using default dissociation constants and input data (alkalinity, CO_2, SST, salinity, silicate and phosphate) as compiled by the Global Ocean Data Analysis Project version 2 (GLODAPv2.2016b), available at https://www.nodc.noaa.gov/archive/arc0107/0162565/2.2/data/0-data/mapped/. For SLAMS-2.0, the annual climatological value is homogeneously spread over the three time steps of each modelled day. Thus, all year-round modelled time steps have the same [CO_3^{2-}] forcing value.

4.2.4 Physical forcing

T (°C) is also from WOA13 (Locarnini et al., 2013). Like for the case of O_2, monthly climatologies of T go down to depth of 1500 m so I have constructed a complete water depth climatology using monthly climatologies for upper water column values and the annual climatology for the lower water column values using `griddedInterpolant`.

MLD (m) is from the monthly climatology of De Boyer Montégut et al. (2004) that uses the variable density criterion, and is accessible at IFREMER/LOS Mixed Layer Depth Climatology website (http://www.ifremer.fr/cerweb/deboyer/mld).

F_{dust} ($\text{g m}^{-2}\,\text{s}^{-1}$) is an annual climatology from a composite model estimate (borrowed from my supervisor's data repository, but unknown source). The climatological annual value, in units of $\text{g m}^{-2}\,\text{y}^{-1}$, is converted into the units of $\text{g m}^{-2}\,\text{s}^{-1}$ and homogeneously spread over each modelled time step throughout the year. Thus, all year-round modelled time steps have the same dust forcing value.

4.3 Model calibration data sets

4.3.1 Particle flux data

4.3.1.1 Sediment traps, radionuclides and optical systems

Oceanographers have been collecting flux data since the 1970s using these two methods: the deployment of free-drifting and stationary sediment traps, which provide a direct estimation of the particle flux (Buesseler et al., 2000), or the measurement of the activity of particle-reactive radionuclides, such as the natural radionuclide pairs $^{238}\text{U}/^{234}\text{Th}$ and $^{210}\text{Po}/^{210}\text{Pb}$[1] (Buesseler et al., 2001), which constitute an indirect estimation of the particle flux. Traditionally, fixed-point sediment traps operating over decades have provided the backbone of particle flux observations (e.g., Honjo et al., 2008). However, they have many trapping issues, especially in the upper 1000 m. Three common factors impact sediment trap results: (i) hydrodynamics of the flows surrounding the traps, which subject the traps to motions that are faster than those of the sinking speed of particles, leaving the slow sinking pool of particles uncollected, (ii) the catch of motile zooplankton ("swimmers") that degrade the organic material in the trap cup and are preserved along with the true settling particles, and (iii) the dissolution of inorganic material after collection (Buesseler et al., 2007a; Lampitt et al., 2008). Those make upper ocean sediment traps prone to undersampling by as much as 90% of the flux (Buesseler et al., 2007a). Consequently, sediment traps became less of a popular method at the turn of the 21st century, being superseded by radionuclides, which are more flexible on the period they collect data over

[1] These radionuclide pairs differ in their spatial and timescale of use: regional and weeks for $^{238}\text{U}/^{234}\text{Th}$ and basin and months for $^{210}\text{Po}/^{210}\text{Pb}$.

(Th estimates typically integrate over ~3 weeks) and are also employed to assess the collection efficiency of traps and calibrate them (Buesseler et al., 2007a; Ceballos-Romero et al., 2016; Roca-Martí et al., 2016). Despite the caveats of sediment traps, they are still a recurrent choice for measuring POC fluxes as they are the only method that can provide direct estimations of flux into the deep ocean[2] and the seafloor as well as information on the chemical composition of the particles.

More recently, improved autonomous optical technologies, like profiling floats, gliders and sediment traps equipped with bio-optical sensors, have allowed estimating fluxes indirectly from the observation of particle sizes and assumptions about their relationship with mass and velocity (Guidi et al., 2009; McDonnell and Buesseler, 2012; Estapa et al., 2017, 2019; Iversen and Lampitt, 2020). *In situ* imaging systems such as the bio-optical profiling float UVP (Underwater Vision Profiler) can avoid the biases attributed to sediment traps and can sweep large areas taking snapshots of marine aggregates at a rate difficult to beat by other methods.

4.3.1.2 Time-series sites

Compilations of particle flux data spanning various oceanic regions are known to be dysfunctional (Buesseler et al., 2007a; Friedrichs et al., 2006, 2007) due to: (i) sparsity of the observations in the horizontal scale (oversampling of coastal areas and the Northern Hemisphere vs undersampling of polar oceans, open oceans and the Southern Hemisphere) as well as the vertical scale (most of the fluxes have been collected at > 500 m while the euphotic zone, where particles are created, and the upper part of the mesopelagic, where the vertical attenuation of POC flux is most prominent, remain undersampled); (ii) collection is often limited to the warmer seasons in the colder latitudes, with the winters undersampled; (iii) collection happens in time periods that remain unsynchronised between locations, some locations choosing to sample before the bloom peak, others after, and (iv) differing methodologies across sampling projects/campaigns, even when using the same technology (e.g., choice of the export depth to deploy sediment traps, Buesseler et al., 2020b). With data collection biased towards certain areas and periods, it is difficult to choose the right spatiotemporal interpolation method to create a global climatology. Thus, the field of oceanography has not seen the emergence of a global climatology of particle flux data yet, as there are already for dissolved tracers (e.g., World Ocean Atlas) or primary production (e.g., satellite-based maps of primary productivity).

[2]Deep ocean fluxes captured by sediment traps still need careful interpretation as the expected downward flux of particles is corrupted by ocean physics in the submesoscale (i.e., vertical scales of ~100 m, horizontal scales of ~1–10 km and temporal scales of ~1–30 days Lévy et al., 2012) that drift particles away from the original site and add new material from horizontal sources. That causes the vertical decoupling between the source of particles in the sunlit surface layer during NPP events and the transferred material to depth.

Particle flux data needed to drive a global model optimisation analysis must be selected and merged with care given the aforementioned dysfunctionalities. The ideal data sources are time-series sites, which have a wealth of particle flux data collected over a long period of time (from years to decades), usually complemented by other biogeochemical data such as *in situ* estimates of NPP, abundance on the different PFTs and ZFTs, or respiration rates. Time-series sites, although limited globally to some 20 projects (between still operational and defunct, Fig. 4.3), are the most reliable data sources. They use consistent methodologies through time, are robust (many replicates, quality controlled and calibration of *in situ* data with remote sensors) and provide true climatologies of the annual cycle after decades of operation.

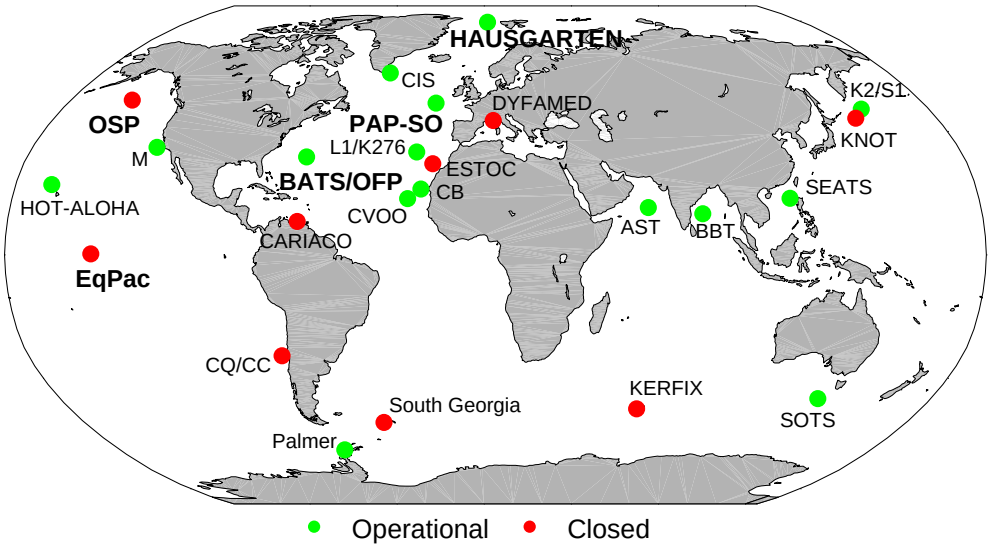


Figure 4.3. Map of global open-ocean (> 1000 m) flux sediment trap programs with fixed-point moorings. Locations include time-series, where moorings have been deployed repeatedly over many years, and temporary moorings, with a shorter time span, but no less than one year. Operational stations (green) are those currently active as of 2021. In bold text are the five locations that I have selected to calibrate SLAMS-2.0: EqPac (Equatorial Pacific Process Study), HAUSGARTEN, OSP (Ocean Station Papa), PAP-SO (Porcupine Abyssal Plain time-Series Observatory) and BATS/OFP (Bermuda Atlantic Time-Series/Oceanic Flux Program joint site). I have built this map based on the following global networks of time-series observatories: OceanSITES (http://www.oceansites.org/index.html), FixO3 for the Atlantic Ocean (http://www.fixo3.eu/observatory/), the International Group for Marine Ecological Time Series study (https://igmets.net/metabase) and WHOI's exhaustive data compilation on sediment traps.

As illustrated in Fig. 4.3, the lack of a global grid completely filled with long-term flux observations imposes a constrain on the calibration and validation of a global model of Lagrangian marine particles like SLAMS-2.0. On top of that, time-series data are not always made readily accessible to the potential users as they still need to be quality controlled or have restricted access due to publication privileges. For this thesis, I have only had ready access to flux data of POC, PIC and bSi ($\mathrm{mmol\,m^{-2}\,d^{-1}}$) from

five long-term studies carried out in fixed-point time-series stations distributed across distinct biogeochemical provinces.

- The U.S. JGOFS Equatorial Pacific Process Study experimental site (EqPac), in the upwelling region of the equatorial Pacific. It consists of an array of seven moored, time-series sediment traps deployed along the 140°W meridional from 9°N–12°S in 1992. I have selected data from the most equatorial stations, at 3°N–1°S, in the HNLC region of the equatorial Pacific, which is inside the Pacific Equatorial Divergence (PEQD) province of Longhurst (1995). Hydrographic details in Honjo et al. (1995).
- The Long-Term Ecological Research (LTER) observatory HAUSGARTEN, in the eastern Fram Strait, off Svalbard, at the Atlantic-Arctic boundary. It consists of an array of moored, time-series sediment traps between 78.3–80.0°N, 3.0–7.0°E. It was established in 1999 by the Alfred Wegener Institute. I have selected the central station, at around 79.0°N, 4.3°E, which falls within the Atlantic Subarctic (SARC) province of Longhurst (1995). Hydrographic details in Soltwedel et al. (2005).
- Ocean Station Papa (OSP), in the southern portion of the Alaska Gyre in the NE subarctic Pacific. Co-located with station P at 50°N, 145°W, it has had a sediment trap moored between 1982–2006. It is in one of three major HNLC regions, it is also part of the "silica ocean" of Honjo et al. (2008) and falls within the Pacific Subarctic Gyre-East (PSAG-E) province of Longhurst (1995). Although the bottom-anchored mooring was dismantled in 2006, the site is used for field campaigns, like the first EXPORTS expedition in 2018. Hydrographic details in Timothy et al. (2013).
- The Porcupine Abyssal Plain time-Series Observatory (PAP-SO), in the NE Atlantic off the SW of the U.K., at 49.0°N, 16.5°W, characterised by a strong spring NPP bloom, where a fixed-point mooring has been in placed since 2002. PAP-SO falls within the North Atlantic Drift (NADR) province of Longhurst (1995). Hydrographic details in Hartman et al. (2012).
- The Bermuda Atlantic Time-Series/Oceanic Flux Program joint site (BATS/OFP), in the northern Sargasso Sea, in the subtropical NW Atlantic. In 1978, the Woods Hole Oceanographic Institution started the OFP mooring, a sediment trap at 31.6°N, 64.2°W, making it the longest operational time series of POC, PIC and bSi flux data. The site is co-located with BATS, established during the U.S. JGOFS in 1988, collecting fluxes of POC alongside hydrographic, biological and chemical data in the upper Sargasso Sea since then. The study site falls within the North Atlantic Subtropical Gyre-West (NAST-W) province of Longhurst (1995). Hydrographic details in Conte and Weber (2014).

Those five sites are all open-ocean (> 1000 m), so that particle fluxes are dominated by local biogeochemical factors rather than land inputs. Furthermore, fluxes are well covered in the vertical scale, especially in the euphotic layer and the bottom of the mesopelagic, as well as in the temporal scale, with data points in most months of the year. Three out of five are still operational time-series stations (HAUSGARTEN, PAP-SO and BATS/OFP). Bottom-moored, sediment trap time-series data were easily available through publications and online platforms, and I have supplemented them with radioactive tracer data and surface-tethered trap data from shorter-term studies happening in those time-series sites. The source data can either be processed, with already applied statistical analysis to calculate means and standard deviations (e.g., OSP data from Roca-Martí et al., 2021), or it can come as the underlying "raw" data, as collected from the instruments without statistical processing (e.g., OSP data from Murray et al., 1996). I have focused data search around three key water depths, coinciding with the depths most frequently sampled by traps: the bottom of the euphotic layer (z_{eu}) at ~100 m, the bottom of the mesopelagic (z_{meso}) at ~1000 m and a bathypelagic depth (z_{bathy}) at ~3000 m.

4.3.1.3 Data compilation

Limiting the observations used to calibrate SLAMS-2.0 to five well-sampled locations makes it possible to construct annual averages that resemble valid climatological values. Thus, I expect to compensate the little worldwide spatial coverage that five locations can offer with values that can be meaningfully compared with the climatological output of SLAMS-2.0. In Table 4.1 I have outlined the three depth horizons at which observations of POC, PIC and bSi fluxes from sediment traps and radionuclides have been extracted to construct my local annual averages, as well as the abundance of data points in each of those depths and the data sources. The data set that I have put together has a total of 7013 data points, where BATS/OFP accumulates 50% of all of them, followed by OSP (28%), HAUSGARTEN (10%), PAP-SO (10%) and EqPac (2%). The bulk of the collected data is in z_{meso} (Fig. 4.4a) and POC flux has a higher count than PIC and bSi. The origin of the ~98% of the data is sediment traps, the remaining ~2% (157 data points) are surface fluxes derived from radionuclides. I have made use of trap and radionuclide data indistinctively. In some instances, radionuclides are in close agreement with sediment traps (e.g., OSP data from the first EXPORTS campaign, Estapa et al., 2021), in others radionuclides are the most reliable source of surface data (e.g., EqPac, Murray et al., 1996) and in some others are the major source of data in the surface ocean (e.g., PAP-SO, Le Moigne et al., 2013). All the data that I have compiled for each location are available on a table at https://doi.org/10.5281/zenodo.5233654 and plotted in Figs. 4.5–4.9 for visual inspection.

Table 4.1. Five ocean time-series where POC, PIC and bSi flux data have been collected for model optimisation, their location and bathymetry, the three ocean depth horizons searched for data (with the data depth ranges and the corresponding average depths used to aggregate the data), the method of data estimation (sediment traps, ST, and radionuclides, RD), the data count (N), the timespan covered and the data references. Data in https://doi.org/10.5281/zenodo.5233654.

Time-series	Basin	Lat (°N)	Lon (°E)	Bathy. (m)	Depths (m)		Method (N)		Timespan	References and flux provided (* = ST, † = RD)[b]
					Searched	Average[a]	ST	RD		
Equatorial Pacific Process Study (EqPac)	Central Equatorial Pacific	3 to -1	-140	4500	75–125 880 3618	100 900 3600	7 63 63	24 0 0	1992, 1994	Honjo et al. (1995): (POC, PIC, bSi)* Le Borgne and Gesbert. (1995): (POC, PIC, bSi)* Murray et al. (1996): POC†,[c]
LTER observatory HAUSGARTEN	Arctic Atlantic	79 to 79.7	4.3 to 5.2	2550	80–100 1225–1250 2495–2560	80 1230 2525	57 288 354	1 0 0	2005–2012	Bauerfeind et al. (2015): (POC, PIC, bSi)*
Ocean Station Papa (OSP)	NE subarctic Pacific	50	-145	4220	40–100 1000–1010 3800	80 1000 3800	85 908 969	22 0 0	1982–2006	Charette et al. (1999): POC† Wong et al. (2002): (POC, PIC, bSi)*,[d] Kawakami et al. (2010): POC† Timothy et al. (2013): (POC, PIC, bSi)*,[e] Mackinson et al. (2015): POC† Roca-Martí et al. (2021): (POC, PIC, bSi)†,[f]
Porcupine Abyssal Plain time-Series Observatory (PAP-SO)	NE Atlantic	49	-16.5	4800	48–68 1000 3000	50 1000 3000	12 92 481	83 0 0	1989–2009	Lampitt et al. (2001, 2010): POC*,[g] Marsay (2012): (POC, PIC, bSi)* Le Moigne et al. (2013): (POC, PIC, bSi)†
Bermuda Atlantic Time-Series/Oceanic Flux Program joint site (BATS/OFP)	Subtropical NW Atlantic	31.6	-64.2	4500	100–150 1500 3200	100 1500 3200	411 1354 1712	27 0 0	1978–2018	http://bats.bios.edu/bats-data/: POC* Kim and Church (2001): POC†,[h] Buesseler et al. (2008): (POC, bSi)*,† Maiti et al. (2009): (POC, bSi)† Stewart et al. (2011): POC*,† Owens et al. (2013): (POC, PIC, bSi)*,[h] Hayes et al. (2018): POC† M. Conte, pers. comm.: (POC, PIC, bSi)*,[i]

[a] The first of the three depths is approx. the euphotic layer depth (z_{eu}). It has been chosen as the surface depth where most of the ST and RD data have been measured, a depth choice made by the field data scientists and which approximates the true z_{eu} in some locations (EqPac, OSP, PAP-SO) but not in others (HAUSGARTEN, BATS/OFP).

[b] When necessary to convert between materials, I have used these molar ratios: POC/total carbon=0.82 (Mouw et al., 2016a); $CaCO_3$/PIC=8.33; bSi/Si=2.40, where bSi is the hydrated form of silica, $SiO_2 \cdot 0.4\,H_2O$, and Si is silicon, the element.

[c] Murray et al. (1996) reports POC flux data both estimated directly from traps and indirectly from radionuclides. I have only taken the radionuclide data since traps overestimate POC flux. Following Murray et al. (1996, Figs. 1–2,), I have eliminated data points of POC flux > 20 mmol C m^{-2} d^{-1} at 0–100 m and the associated PIC and bSi data points.

[d] I have eliminated POC and PIC flux data points (and associated bSi) where the molar PIC/POC ratio was < 0.20 or > 0.90 (Wong et al., 2002, Fig. 7).

[e] These data have been eliminated (Timothy et al., 2013): data from four extreme events (POC flux in the 2nd half of Aug 1983 at 3,800 m, all fluxes in May 1990, all fluxes in Feb 1997 and POC flux in the 1st half of Aug 2003); any POC flux ≥ 15, then any ≥ 4.0 at ≥ 200 m, ≥ 1.5 at ≥ 1000 m, ≥ 0.75 at ≥ 3800 m (units mmol C m^{-2} d^{-1}); any PIC flux ≥ 6.5, then any ≥ 2.0 at ≥ 200 m, ≥ 2.0 at ≥ 1000 m, ≥ 1.0 at ≥ 3800 m (units mmol C m^{-2} d^{-1}); bSi flux ≥ 3.0 at 200 m, then any ≥ 4.0 at 1000 m and ≥ 3.0 at 3800 m (units mmol Si m^{-2} d^{-1}).

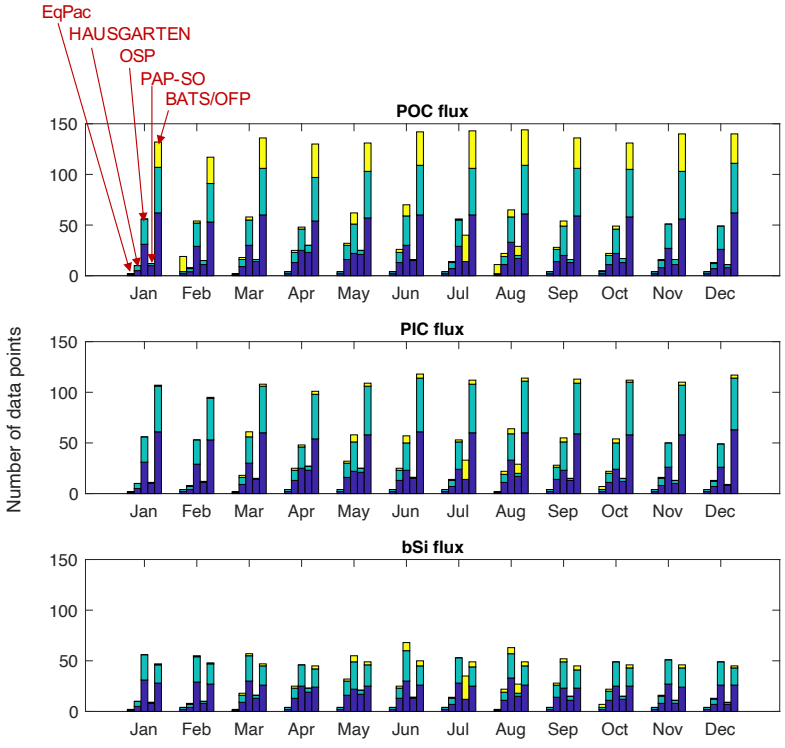
[f] EXPORTS data from sediment traps are reported in Buesseler et al. (2020a), but I have discarded them as they was noted to overestimate POC fluxes.

[g] Data in the year 2001 were eliminated as they were unusually high, probably affected by an aerosol deposition (Lampitt et al., 2010).

[h] Data were not made available on tables so I digitised the plots to extract flux values using the open-source software WebPlotDigitizer (https://automeris.io/WebPlotDigitizer).

[i] Data collected after hurricanes Fabian (Sep 2003) and Igor (Nov 2010) were eliminated, as well as some samples contaminated by fish in Sep 2002 (M. Conte, pers. comm.).

(a)



(b)

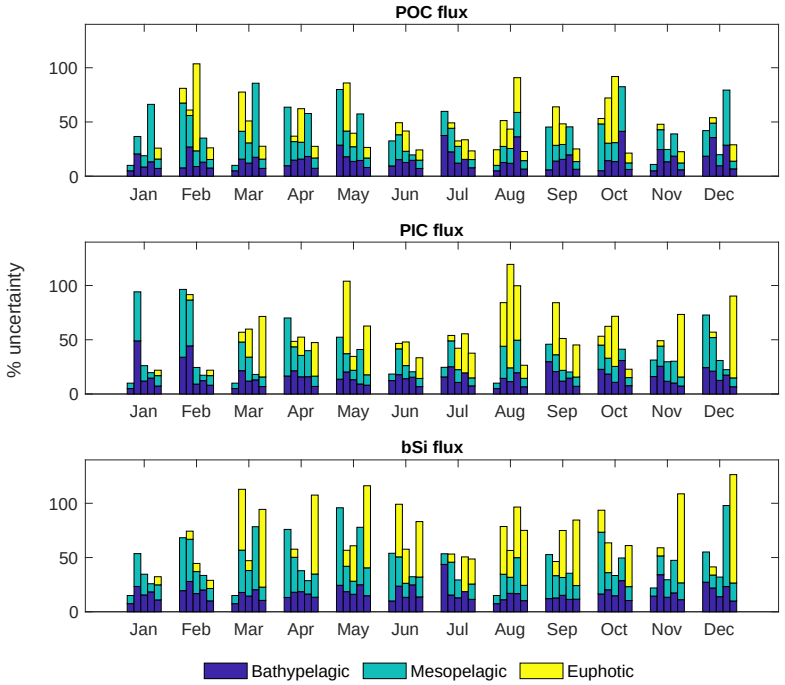


Figure 4.4. (a) Number of particle flux data points from sediment traps and radionuclides and (b) their associated average relative uncertainty compared on a monthly basis (*groups*) for each of the five locations used to calibrate SLAMS-2.0 (*stacks*: EqPac, HAUSGARTEN, OSP, PAP-SO and BATS/OFP) and three key ocean depths (*stack elements*: euphotic, mesopelagic and bathypelagic). The total number of data points is 7013, where BATS/OFP accumulates 50% of all of them, followed by OSP (28%), HAUSGARTEN (10%), PAP-SO (10%) and EqPac (2%). The % uncertainty is the sum of systematic and random errors divided by the monthly average value.

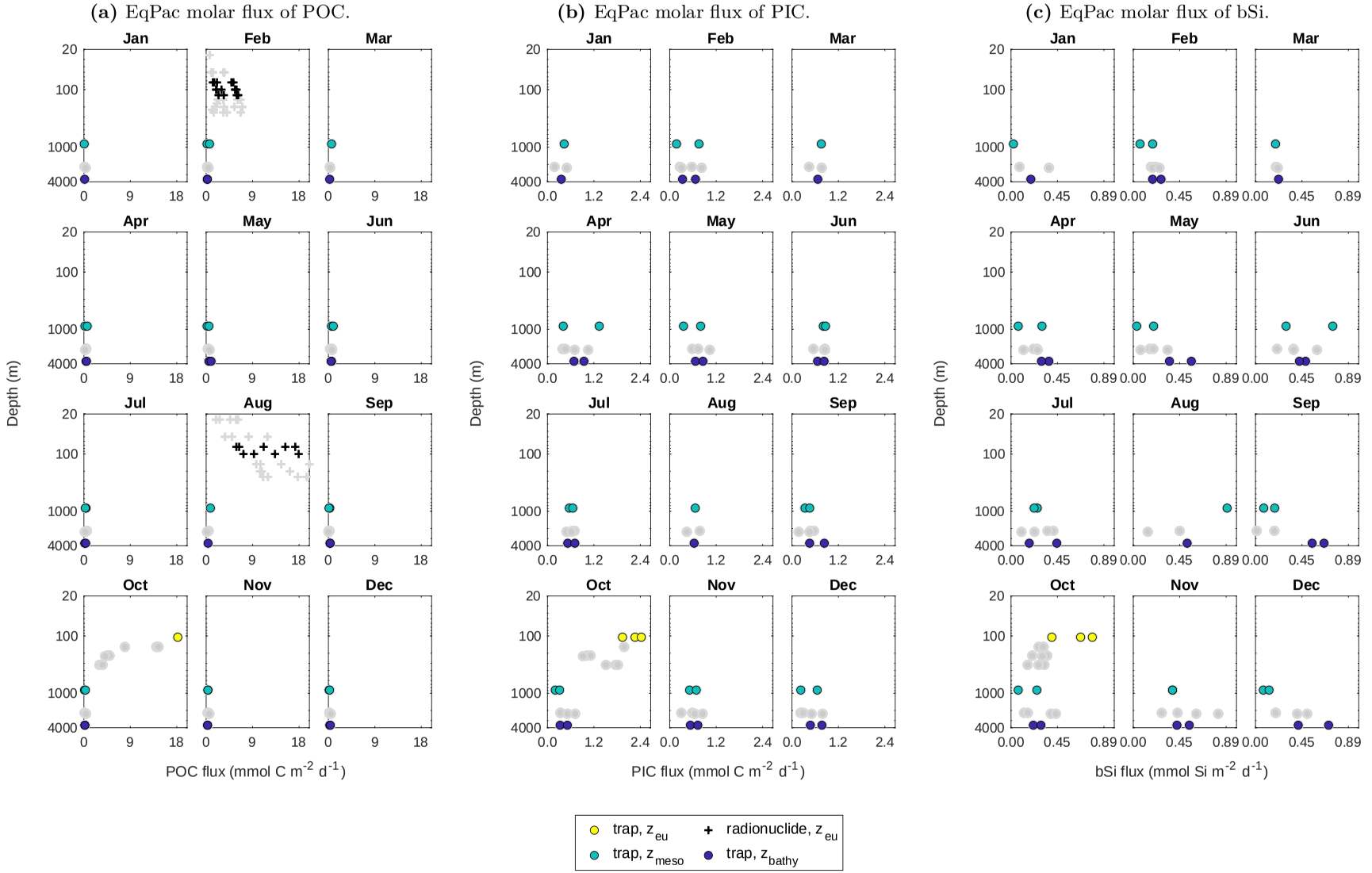


Figure 4.5. EqPac's monthly observations of fluxes of (a) POC ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$), (b) PIC ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$) and (c) bSi ($\mathrm{mmol\,Si\,m^{-2}\,d^{-1}}$). Coloured on top of grey symbols are those data points at the depths of interest (Table 4.1), with a colour scheme according to depth horizon and method of observation: yellow circle for trap data at the bottom of the euphotic layer, z_{eu}; turquoise circle for trap data at the bottom of the mesopelagic, z_{meso}; indigo circle for trap data at a bathypelagic depth, z_{bathy}; black cross for radionuclide data at z_{eu}. Notice that the y axis is in $\log_{10}$ and that the x is bounded to the year's maximum POC, PIC or bSi flux value. Data are from: the U.S. JGOFS time-series published work in Honjo et al. (1995) (meso- and bathypelagic traps) and Murray et al. (1996) (surface radionuclides), and the FLUPAC field work published in Le Borgne and Gesbert. (1995) (surface traps).

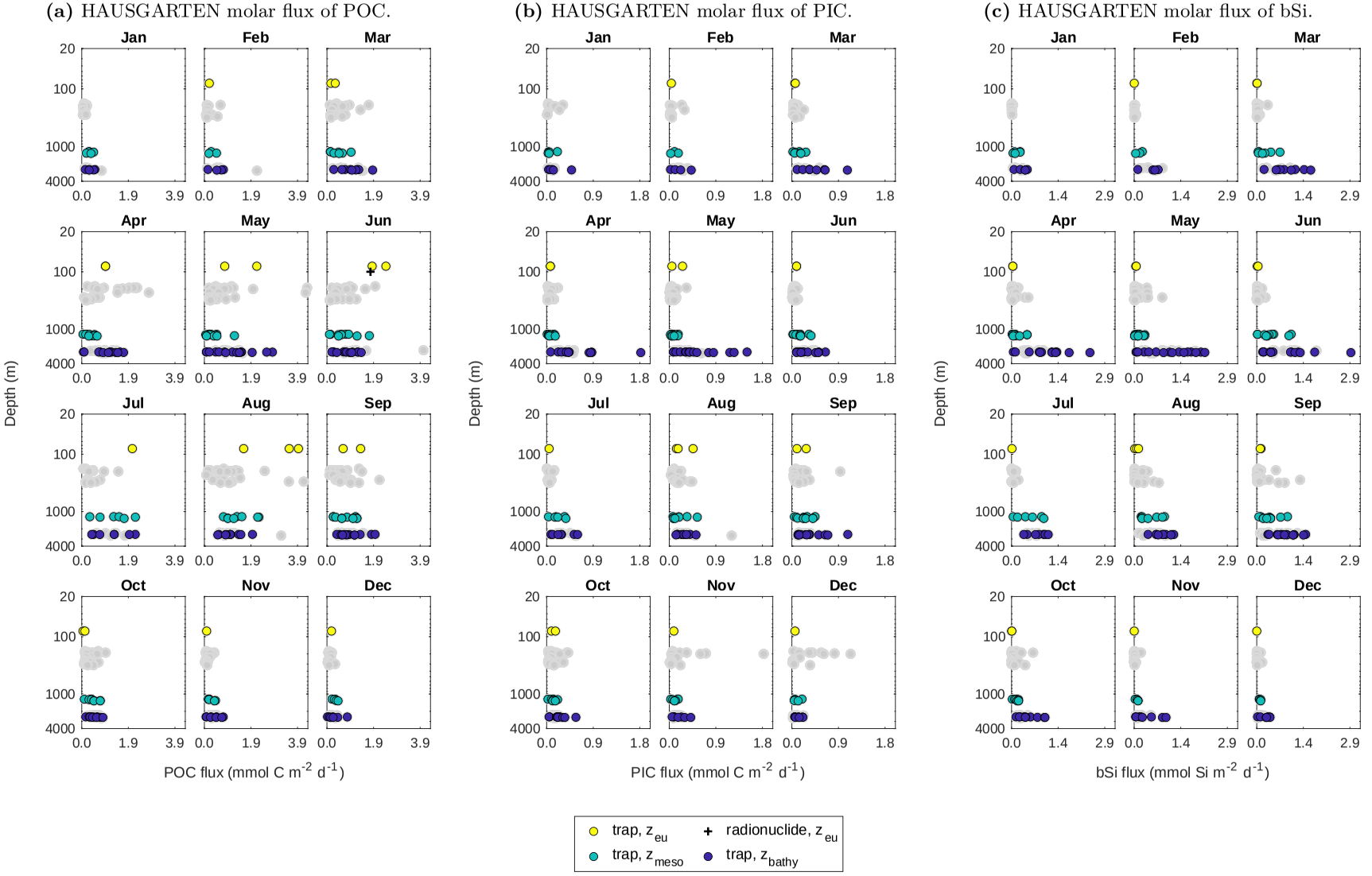


Figure 4.6. HAUSGARTEN's monthly observations of fluxes of (a) POC ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$), (b) PIC ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$) and (c) bSi ($\mathrm{mmol\,Si\,m^{-2}\,d^{-1}}$). Symbols and axes are as explained in Fig. 4.5. All data are from Bauerfeind et al. (2015) (traps).

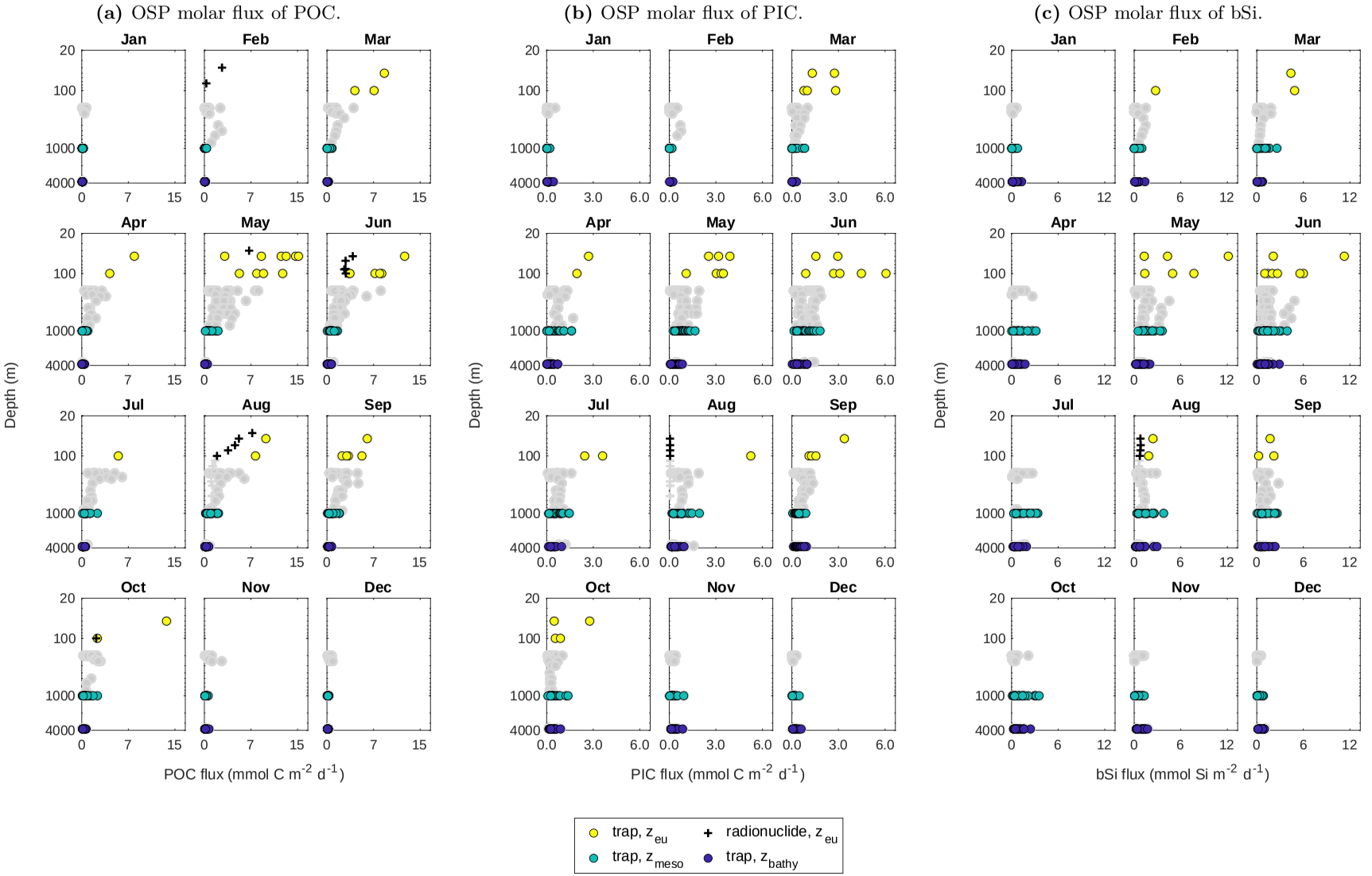


Figure 4.7. OSP's monthly observations of fluxes of (a) POC ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$), (b) PIC ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$) and (c) bSi ($\mathrm{mmol\,Si\,m^{-2}\,d^{-1}}$). Symbols and axes are as explained in Fig. 4.5. Data are from: the OSP time-series synthesis of Timothy et al. (2013) (surface, meso- and bathypelagic bottom-moored traps), the EXPORTS field work published in Roca-Martí et al. (2021) (surface and mesopelagic radionuclides), Charette et al. (1999) (surface radionuclides), Wong et al. (2002) (surface and mesopelagic free-drifting traps), Kawakami et al. (2010) (surface radionuclides) and Mackinson et al. (2015) (surface radionuclides).

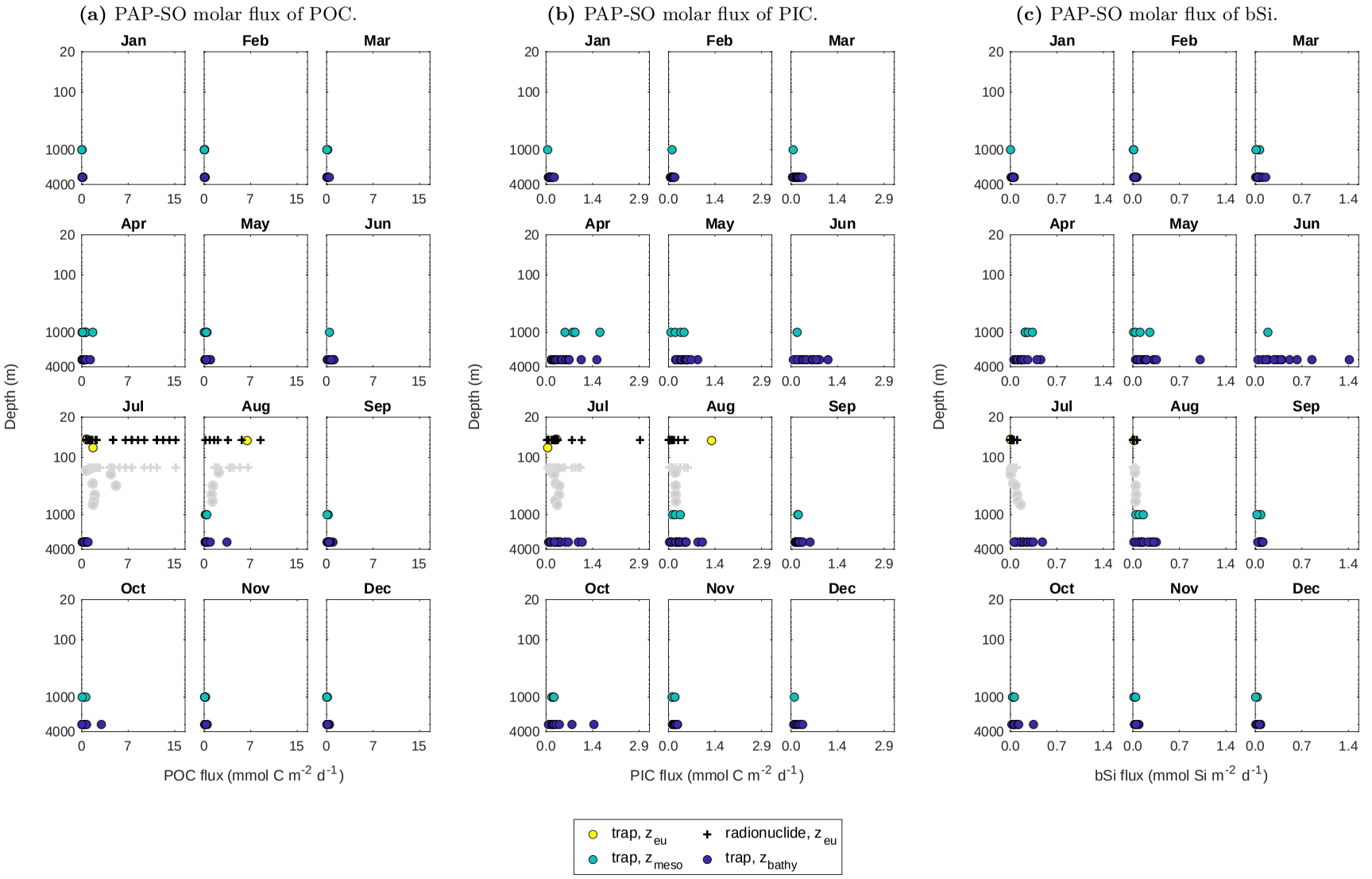


Figure 4.8. PAP-SO's monthly observations of fluxes of (a) POC ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$), (b) PIC ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$) and (c) bSi ($\mathrm{mmol\,Si\,m^{-2}\,d^{-1}}$). Symbols and axes are as explained in Fig. 4.5. Data are from: the PAP-SO time-series published in Lampitt et al. (2001, 2010) (meso- and bathypelagic traps) and field work data at the PAP site published in Marsay (2012) (surface and mesopelagic traps) and Le Moigne et al. (2013) (surface radionuclides).

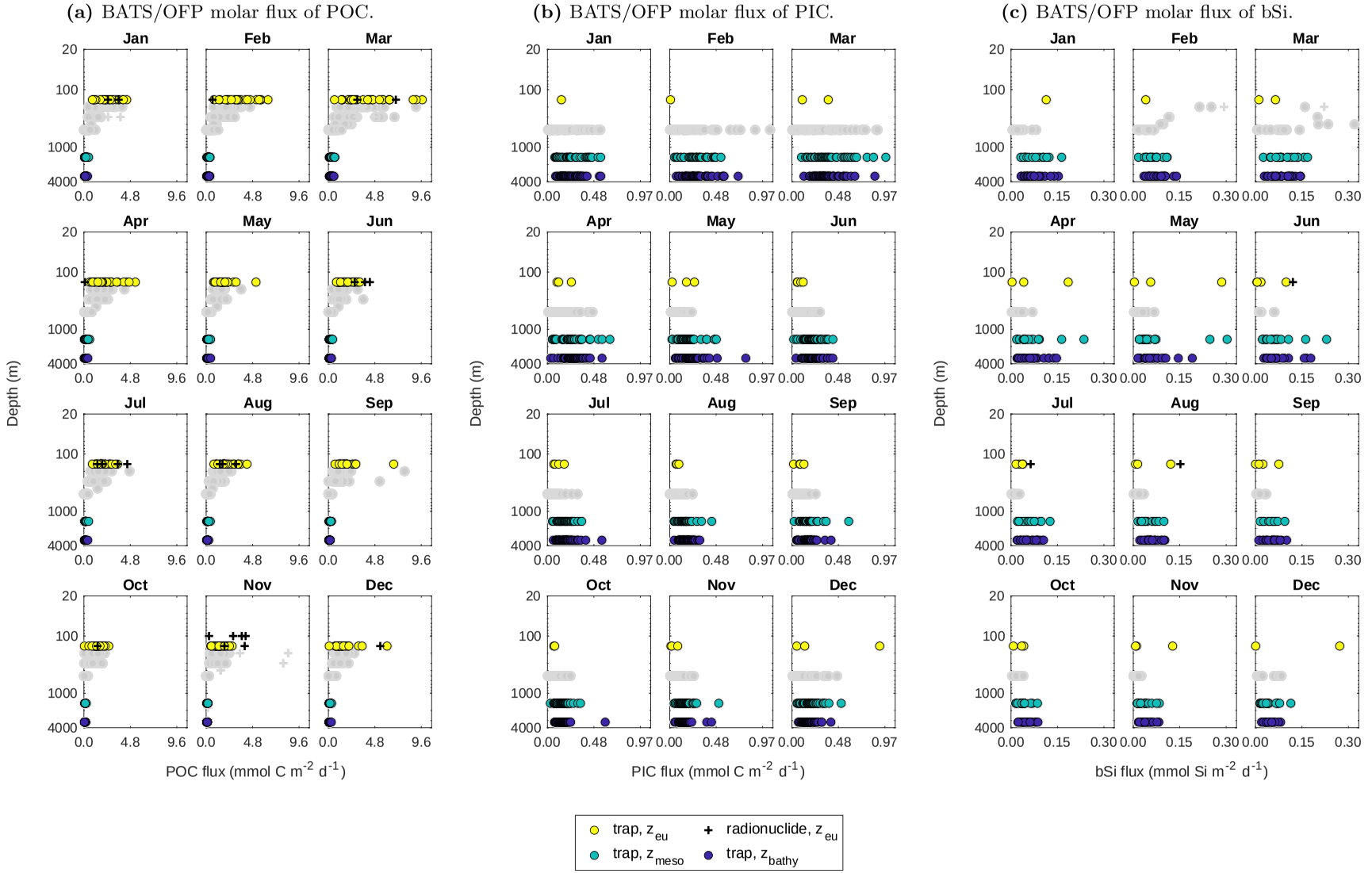


Figure 4.9. BATS/OFP's monthly observations of fluxes of (a) POC ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$), (b) PIC ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$) and (c) bSi ($\mathrm{mmol\,Si\,m^{-2}\,d^{-1}}$). Symbols and axes are as explained in Fig. 4.5. Data are from: the BATS time-series in http://bats.bios.edu/bats-data/ (surface and mesopelagic traps), the OFP time-series from M. Conte, pers. comm. (meso- and bathypelagic traps), and various field work carried out at the BATS site published in Kim and Church (2001) (surface radionuclides), Maiti et al. (2009) (surface radionuclides), Stewart et al. (2011) (surface and mesopelagic radionuclides and traps), Owens et al. (2013) (surface traps), as well as field work from the U.S. GEOTRACES campaign published in Hayes et al. (2018) (surface, meso- and bathypelagic radionuclides) and the EDDIES project published in Buesseler et al. (2008) (surface radionuclides).

A thorough analysis of the data plotted reveals the temporal heterogeneity of the data points: some months are better sampled than others (e.g., Feb and Aug in EqPac, the warmer months for OSP and PAP-SO) and PIC, and specially bSi, rely on fewer monthly measurements, and in some extreme cases all measurements are reduced to one month in the whole year (e.g., Oct in EqPac for surface PIC and bSi). There is also spatial heterogeneity, with the euphotic layer depth being missed in HAUSGARTEN and BATS/OFP and a general better coverage of the mesopelagic and bathypelagic ocean compared to the surface. Notice as well the decaying shape in the POC flux plots, and the lack of such a trend in PIC and bSi fluxes, with a messier vertical structure that, in most locations, does not even appear to decrease with depth, but the opposite (e.g., BATS/OFP and the rest, but mind OSP). And a final remark. When plotting flux data on a depth axis, one is assuming that those observations have been taken from a water column that is in steady state (i.e., between an observation made at an upper depth and an observation made at a lower depth, there has been time for the pulse at the upper depth to arrive at the lower depth). However, observations could not reflect that vertical funnel if not taken with care. Sometimes, there is an addition of foreign material from horizontal sources due to unaccounted lateral advection. That often explains the unexpected vertical trends in data (e.g., increase of the flux of material with depth, like PIC flux, when a decrease was expected).

4.3.1.4 Data aggregation

To transform the compiled flux data of POC, PIC, and bSi into values comparable to SLAMS-2.0 flux outputs, I have to aggregate the data into monthly means and then calculate the annual averages of those fluxes. Dealing with data estimated using different methods (sediment traps vs radionuclides) and with differing statistical treatments applied to them (raw data vs data that have already been averaged from raw data points), is a challenge that welcomes the consideration of the simplest method to aggregate the data points: aggregating the multiple observations made in each month into a single value (the monthly average) –instead of separating by seasons or bloom phases–, and propagating the uncertainties associated with the data using the standard method of adding in quadrature, as explained further below.

Particle flux observations have associated uncertainties due to a combination of bad measurement practices (e.g., a miscalibrated trap or a technique that undersamples/oversamples) and unknown and unpredictable changes in the experimental environment (i.e., electronic noise in the circuit or currents deviating the trap). I call the uncertainty due to challenges with the instrumentation *systematic error* and it can be corrected if experiments are designed carefully. The uncertainty caused by unknown

and unpredictable variations in the environment is the *random error*, it has a Gaussian distribution and can be dealt with if enough replicates are taken. Field data uncertainty is presented in a variety of ways that make no easy treatment of the uncertainty. The most common case is not to find a reported measure of uncertainty associated with the data. It is the case of the POC, PIC and bSi fluxes of the OFP time-series, probably because the systematic error is too low and replicates for the same day are not taken. However, when observations do come with an associated error, it can be the systematic error (e.g., the ^{234}Th-derived fluxes of POC, PIC and bSi in Le Moigne et al. (2013) for PAP-SO) or, if replicates have been taken, the standard error of the mean, which is a random error (e.g., the trap-derived POC, PIC and bSi fluxes in Marsay (2012) for PAP-SO). Other times, the nature of the reported uncertainty is not specified, like BATS fluxes of POC in Stewart et al. (2011). To overcome that heterogeneity in reported uncertainties, I have assigned a common systematic error to all observations of POC, PIC and bSi flux following the average systematic errors found in sediment traps after the U.S. JGOFS. Honjo et al. (2008) report those errors as being 5% for POC and PIC and <10% for bSi, which I have approximated to 7.5%.

Next, I outline how I have aggregated the data points, propagated their error and calculated the BCP flux-related metrics of PE_{eff}, T_{eff}, molar ratio of POC/PIC and bSi/PIC.

1. I have binned monthly the flux data of each of the three materials i to calculate monthly flux averages, $\overline{x_i}$.
2. I have calculated monthly values of BCP flux-related metrics. The molar ratios of POC/PIC and bSi/PIC and the T_{eff} are computed with the corresponding monthly flux averages. To compute PE_{eff} (POC flux at z_{eu}/NPP), I have used NPP estimates from the Carr (2002) model (see Sec. 4.3.2 below for a justification of the choice of this model). To facilitate the analysis, I have assumed that NPP estimates have no uncertainties associated, i.e., that they are the true values.
3. I have calculated the associated error of the monthly binned values, a combination of systematic and random errors. Since all individual data points for a given i have the same assigned systematic error, the monthly systematic error of flux is simply the product of the fractional systematic error times the monthly flux average i.e., $\Delta\overline{x}_{i,\,\mathrm{sys}} = \frac{\Delta x_i}{x_i} \cdot \overline{x}_i$. The systematic error of BCP metrics, which are quotients, is calculated by propagating the errors of the divided quantities. For instance, the systematic error associated with POC/PIC is the sum in quadrature of the individual systematic errors, $\Delta\overline{x}_{\mathrm{POC/PIC,\,sys}} = \overline{x}_{\mathrm{POC/PIC}} \times \sqrt{\left(\frac{\Delta x_{\mathrm{POC}}}{x_{\mathrm{POC}}}\right)^2 + \left(\frac{\Delta x_{\mathrm{PIC}}}{x_{\mathrm{PIC}}}\right)^2}$. The random error of monthly averaged flux data corresponds to the standard error of the monthly mean, $SE_{i,m}$, and it is calculated as the standard deviation of each monthly

variable, $\sigma_{i,m}$, divided by the square root of the number of monthly observations N_i, i.e., $\mathrm{SE}_{i,m} = \sigma_{i,m}/\sqrt{N_i}$. The random error of BCP metrics is calculated by applying a sum in the quadrature of the standard errors, $\Delta\overline{x}_{\mathrm{POC/PIC,rand}} = \overline{x}_{\mathrm{POC/PIC}} \times \sqrt{\left(\frac{\mathrm{SE_{POC}}}{\overline{x}_{\mathrm{POC}}}\right)^2 + \left(\frac{\mathrm{SE_{PIC}}}{\overline{x}_{\mathrm{PIC}}}\right)^2}$.

4. I have calculated the annual averages of fluxes and BCP metrics as weighted arithmetic means, dividing the sum of monthly means pondered by their number of associated observations by the total number of observations.
5. I have calculated the corresponding associated uncertainty by propagating the monthly random and systematic errors using the error propagation formula for weighted averages, which divides the uncertainty summed in quadrature by N. The N chosen for the annual average of POC/PIC and bSi/PIC is the associated to the POC, PIC or bSi measurement in the quotient that has the minimum number of entries; for $\mathrm{PE_{eff}}$, N is the associated to POC at z_{eu}, and, for $\mathrm{T_{eff}}$, N is the associated to the POC measurement in the quotient that has the minimum number of entries. I have computed the net annual error by summing the annual systematic and random errors in quadrature, i.e., $\mathrm{net} = \sqrt{\mathrm{sys}^2 + \mathrm{rand}^2}$.

The above procedure generates the annual means and associated uncertainties of POC, PIC and bSi fluxes summarised in Table 4.2. Of the five locations selected, BATS/OFP has the minimum amount of uncertainty associated with observations across all depth intervals. Of all three depth horizons, z_{eu} has the largest uncertainties due to large inter-monthly variations in export fluxes –notice in Figs. 4.5–4.9 how spread the data are for any given month, which can equal the total spread in one year.

4.3.2 The Carr (2002) model for NPP

NPP observations are required to calculate $\mathrm{PE_{eff}}$. To evaluate which satellite-based model (VGPM, CbPM, CAFE or Carr (2002)) produces NPP estimates that best resemble *in situ* data at my five ocean locations, I have carried out a correlation analysis. After constructing the global compilations of satellite estimates of NPP shown in Fig. 4.2, I have extracted the NPP values at my five locations and have compared them with *in situ* NPP observations collected from the literature (Table 4.3).

As noted in the recent publication of Marra et al. (2020), *in situ* estimates of NPP are scarce worldwide, which makes finding and putting together a data set of *in situ* NPP observations a challenging task not unique to this thesis. Marra and colleagues claim to present a data set of *in situ* NPP with "wide geographic coverage"; however, they could only put together data for nine locations in the entire ocean, which are not even covered for all months of the year. At my five locations, the lack of NPP measurements is

Table 4.2. Annual means of particle flux data of POC, PIC and bSi ($\mathrm{mmol\,m^{-2}\,d^{-1}}$) and their associated absolute uncertainty (combination of systematic and random errors, summed in quadrature and expressed as absolute error) for the five studied ocean locations (EqPac, HAUSGARTEN, OSP, PAP-SO, BATS/OFP) at three relevant ocean depths (z_{eu}, z_{meso} and z_{bathy}). The sources of the data are annotated in Table 4.1.

Depth	EqPac	HAUSGARTEN	OSP	PAP-SO	BATS/OFP
POC flux, $\mathrm{mmol\,m^{-2}\,d^{-1}}$					
z_{eu}	7.34 ± 0.69	1.27 ± 0.14	6.51 ± 0.52	5.08 ± 0.81	2.21 ± 0.07
z_{meso}	0.39 ± 0.04	0.61 ± 0.04	0.51 ± 0.02	0.25 ± 0.05	0.19 ± 0.01
z_{bathy}	0.36 ± 0.02	0.85 ± 0.05	0.19 ± 0.01	0.36 ± 0.03	0.15 ± 0.00
PIC flux, $\mathrm{mmol\,m^{-2}\,d^{-1}}$					
z_{eu}	2.23 ± 0.18	0.13 ± 0.02	2.21 ± 0.24	0.38 ± 0.11	0.13 ± 0.03
z_{meso}	0.57 ± 0.06	0.14 ± 0.01	0.44 ± 0.02	0.33 ± 0.04	0.22 ± 0.01
z_{bathy}	0.63 ± 0.04	0.36 ± 0.03	0.29 ± 0.01	0.34 ± 0.02	0.22 ± 0.00
bSi flux, $\mathrm{mmol\,m^{-2}\,d^{-1}}$					
z_{eu}	0.61 ± 0.12	0.04 ± 0.01	3.36 ± 0.57	0.01 ± 0.00	0.06 ± 0.01
z_{meso}	0.24 ± 0.03	0.26 ± 0.02	0.93 ± 0.05	0.08 ± 0.01	0.06 ± 0.00
z_{bathy}	0.39 ± 0.03	0.80 ± 0.05	0.64 ± 0.03	0.15 ± 0.01	0.06 ± 0.00

also noted, with yearly-complete NPP *in situ* observations only at BATS/OFP, almost complete at OSP, sporadic observations at EqPac and HAUSGARTEN, and none at all at PAP-SO. Thus, only *in situ* NPP estimates for BATS/OFP, OSP and EqPac are used in the correlation exercise between the satellite models and *in situ* NPP (Table 4.4), which shows that Carr (2002) is the model that most closely fits local *in situ* NPP.

4.3.3 Particle number concentration data

I downloaded particle number concentration ($\mathrm{num.\,L^{-1}}$) data from the particle imaging online repository EcoTaxa (http://ecotaxa.obs-vlfr.fr/part/), where data are stored by sampling events in 15 size classes geometrically spaced in a sequence spanning 1, 2, 4, 8, ..., 8190, 16400 μm. I could only find particle number data for all my calibration sites except the EqPac site. To aggregate the data, at each of the four locations I have searched for nearby depth profiles containing particle number. Next, I have summed over particle size classes to get the total particle number, as shown on Table 4.5. I have calculated the associated uncertainty in two steps. First, I have calculated the standard deviation of the particle number for each size class amongst the different depth profiles found nearby the location. Second, when summing over particle size classes, I have propagated the size class standard deviations by adding them in quadrature. I have selected the calculated averages for three depths: 50, 200 and 500 m. Particle number data carry most of the uncertainty at the uppermost depth, which is remarkably high for OSP.

Table 4.3. Monthly means of net primary production (NPP) depth-integrated ($mg\,C\,m^{-2}\,d^{-1}$) for five ocean locations (EqPac, HAUSGARTEN, OSP, PAP-SO, BATS/OFP) with a comparison of satellite estimates from four ocean-colour-based models (VGPM, CbPM, CAFE and Carr 2002) and *in situ* estimates using the ^{14}C uptake technique. The different models use one of these two distinct chl *a* and PAR_0 forcings: from MODIS-Aqua products (M.) or from SeaWiFS products (S.). The estimates from the VGPM (Behrenfeld and Falkowski, 1997), CbPM (Westberry et al., 2008) and CAFE (Silsbe et al., 2016) models are from the Ocean Productivity site (http://sites.science.oregonstate.edu/ocean.productivity/). The Carr 2002 values are estimated from Carr (2002, Eqs. 1–4) using a C to chla *a* ratio of 50 $g\,C\,(g\,Chl\,a)^{-1}$ (from Fasham et al., 1990, after value not reported in the Carr 2002 model). The input data to drive the Carr (2002) algorithm are from slightly different periods as originally chosen by the author after some data sets have been extended with new data over the past two decades. Thus, chl *a* and PAR_0 are from SeaWiFS 9 km monthly climatology for the whole satellite lifespan, 1997–2010 (hhttps://oceandata.sci.gsfc.nasa.gov/SeaWiFS), SST is from AVHRR Pathfinder v5.0 global 4 km monthly climatology for 1985–2001 (https://www.ncei.noaa.gov/products/avhrr-pathfinder-sst) and seawater temperature to calculate maximum photosynthetic rate at the base of the mixed layer is from WOA13 (https://www.nodc.noaa.gov/OC5/woa13/woa13data.html). The *in situ* estimates of NPP based on the ^{14}C uptake technique are: for the EqPac, from Murray et al. (1996, Table 5, 3°N–1°S, Feb, Aug, Sep) and Le Borgne and Gesbert. (1995, Ch. 11, avg. st. 94 and 100, Oct); for HAUSGARTEN, data are from Le Moigne et al. (2015, Table 2, avg. st. 11 and 14); for OSP, data are as reported in Timothy et al. (Fig 2(i), period 1987–1996 2013, data extracted with the open-source software WebPlotDigitizer, available at https://automeris.io/WebPlotDigitizer); for the PAP-SO, there are no primary production in situ measurements (R. Lampitt, pers. comm.); for BATS/OFP data are monthly averages for the BATS site from 1990–2018 put together by M. Conte (pers. comm., not published yet).

	EqPac (0°N, 140°W)						HAUSGARTEN (79°N, 4.3°E)						OSP (50°N, 145°W)						PAP-SO (49°N, 16.5°W)						BATS/OFP (31.6°N, 64.2°W)					
	VG'	Cb'	CAFE		Carr	^{14}C	VG'	Cb'	CAFE		Carr	^{14}C	VG'	Cb'	CAFE		Carr	^{14}C	VG'	Cb'	CAFE		Carr	^{14}C	VG'	Cb'	CAFE		Carr	^{14}C
	M.	M.	M.	S.	S.		M.	M.	M.	S.	S.		M.	M.	M.	S.	S.		M.	M.	M.	S.	S.		M.	M.	M.	S.	S.	
Jan	470	1034	793	738	766	–	567	131	361	341	0	–	165	130	232	186	164	–	187	2	188	154	159	–	372	302	316	277	305	394
Feb	479	1065	845	810	805	718	530	100	346	344	0	–	219	204	271	214	279	267	265	2	252	207	253	–	427	300	359	353	405	613
Mar	454	1046	879	878	828	–	494	69	331	348	444	–	299	253	346	300	398	–	407	3	351	331	461	–	497	343	402	423	494	576
Apr	411	1035	855	859	786	–	458	38	316	351	847	–	362	360	372	333	471	557	720	83	539	511	833	–	400	428	393	433	440	547
May	443	1098	846	821	819	–	765	125	637	630	872	–	461	548	509	453	517	527	989	563	710	710	1407	–	296	360	371	405	323	466
Jun	461	1112	825	800	905	–	1432	378	872	810	1281	477	579	698	641	599	469	823	1361	1066	894	847	1200	–	245	372	370	397	276	449
Jul	502	1127	827	814	1070	–	1155	759	745	625	952	–	701	748	662	604	546	377	1271	1022	771	778	934	–	184	361	354	376	250	443
Aug	493	1072	861	856	1021	1289	921	466	511	425	796	–	778	738	606	559	574	619	1120	922	642	615	705	–	153	361	342	355	217	403
Sep	480	1054	900	871	890	1011	712	255	420	327	514	–	695	718	516	476	572	833	887	688	500	456	655	–	159	378	337	356	229	395
Oct	477	1026	867	841	880	1373	675	224	405	330	87	–	472	508	384	322	437	362	619	461	397	345	574	–	190	391	332	323	213	355
Nov	461	992	811	771	767	–	639	193	390	334	0	–	251	252	289	236	244	230	313	129	288	245	242	–	265	419	317	288	251	349
Dec	464	989	765	734	796	–	603	162	376	337	0	–	208	191	261	211	152	–	248	66	233	194	148	248	321	389	293	249	268	303

Table 4.4. Squared correlation coefficients (R^2) between five satellite-based estimates of NPP (VGPM MODIS, CbPM MODIS, CAFE MODIS, CAFE SeaWiFS, Carr 2002) and their *in situ* observations counterparts at EqPac, OSP and BATS/OFP. The highest R^2 are highlighted for each location.

VGPM MODIS	CbPM MODIS	CAFE MODIS	CAFE SeaWiFS	Carr 2002
EqPac (0°N, 140°W)				
0.29	-0.47	0.20	0.53	0.68
OSP (50°N, 145°W)				
0.63	0.68	0.62	0.66	0.70
BATS/OFP (31.6°N, 64.2°W)				
0.65	-0.37	0.85	0.74	0.84

Table 4.5. Total particle number observations (num. L^{-1}) from EcoTaxa and their associated error (random uncertainty propagated from particle size classes and expressed as absolute error) for four ocean locations where data was available (HAUSGARTEN, OSP, PAP-SO, BATS/OFP) at three relevant ocean depths (50, 200 and 500 m).

Depth	HAUSGARTEN	OSP	PAP-SO	BATS/OFP
50 m	402 ± 353	519 ± 1311	59.6 ± 31.4	186 ± 88.4
200 m	32.2 ± 14.5	340 ± 884	14.8 ± 6.00	74.7 ± 11.1
500 m	24.9 ± 1.40	101 ± 98.8	14.9 ± 5.16	132 ± 31.5

5

Model calibration and validation

Contents

5.1 Introduction

Any simulation, like SLAMS-2.0, models the relationship between x inputs (ocean forcing) and y outputs (particle fluxes and BCP metrics). That relationship is built using parameters, the gears of a model, and SLAMS-2.0 counts on 84 (Table C.1). After a model is setup, most often, it will not represent reality accurately as there is a mismatch, or misfit, between model outputs (predictions) and empirical observations (actual target

values). Thus, assuming that the mismatch is attributable to the parameters and not a poor model construct, model parameters need to be tuned, or calibrated, for the model to match observations before being used in an application.

This chapter aims to constrain the particle response of SLAMS-2.0 through its sensitive (adjustable) parameters by assimilating *in situ* flux observations of POC, PIC and bSi ($\mathrm{mmol\,m^{-2}\,d^{-1}}$), as compiled in Chapter 4. Although I did also compile observations of particle number concentration ($\mathrm{num.\,L^{-1}}$), I will not use those to calibrate the model for three reasons: (i) those are not long-term observations, (ii) they have an unspecified uncertainty (not provided by EcoTaxa, but could be relatively large) and (iii) SLAMS-2.0's coagulation module, based on Jackson (2001)'s formulation, is not fit for reproducing the high observed particle abundances of porous and fractal aggregates, but impermeable spheres, or beads. What SLAMS-2.0 can correctly simulate in terms of particles are relative abundances of small vs large particles, and thus the modelled slope of the PSD should be comparable to that of observations.

Calibration has two steps: identifying the most sensitive parameters that will be subject to adjustment (*parameter sensitivity analysis*) and determining their optimal value, that which minimises the mismatch between model predictions and real-world observations (*parameter optimisation analysis*)[1]. Model output and observations are used to build a statistical model (misfit function) of the mismatch between those. An optimiser minimises the misfit function by computing new parameter values and reassessing the misfit that those generate iteratively until a minimum is found. The parameter values become optimal when generating model results that have a minimum data–model misfit. Amongst several optimisation algorithms available, I have chosen a derivative-free, least squares, local optimiser, DFO-LS (Derivative-Free Optimisation using Least Squares) recently developed by Cartis et al. (2019). DFO-LS is suited for (i) inverse problems (where there are more unknown parameters than data available to constrain them), (ii) model outputs that have noise due to the use of stochasticity, (iii) it is a bounded algorithm, limiting the search to parameter range values chosen by the modeller, and (iv) it is quick to run (Roberts, 2018). This optimisation exercise is not intended to get a perfect fit of the model to data, which is problematic given the insufficient number of particle flux observations available and the

[1] A parameter optimisation analysis is also referred in the oceanography literature as *inverse problem* (e.g., Fennel et al., 2001) and *data assimilation process* (e.g., Schartau et al., 2017). Indeed, whereas model development is a forward problem, where parameters are mapped to the observations, parameter optimisation is an inverse problem, where parameters are estimated from observations (i.e., the inverse of a forward problem). Data assimilation is a particular class of inverse problems in which the unknown parameters are treated as the initial conditions of the modelled system. They take over the role of the state variables and are estimated by combining the model with available observations of the actual system.

high non-linearity (i.e., flexibility) of SLAMS-2.0, with a propensity to overfit[2] data due to non-mechanistic reasons. Instead, the aim is to improve the match to data of a default model run for simultaneously five ocean sites placed in distinct ocean biogeochemical provinces while using a parameter set kept in between reasonable physiological boundaries.

Multiple factors can affect the success of an optimisation analysis (Evans, 2003; Friedrichs et al., 2006, 2007; Ward et al., 2010).

(a) The model construct –is it lacking processes, or the design is overly complicated? can the processes represented in the model explain the observed data?
(b) The parameter set used to calibrate the model –how many parameters does the configuration set contain? how sensitive is the model output to the parameter set selected for calibration? what is our knowledge of the range of possible values that a parameter can take?
(c) The calibration training data set –how many observations of the type we are interested in are available? how robust is the averaged data set (random errors)? are there large errors associated to their collection (systematic errors)?
(d) The model forcing data –is the forcing data not synchronised with the timing of the observation data set used to train the optimisation algorithm? are there large errors associated with those data? do the data available offer strong enough constraints on the output fields I am most interested in?
(e) The misfit function design, which considers the shape used to define the misfit (e.g.,absolute difference vs squared difference), the weighting factors (mean, variation) and how different data types have been combined.
(f) The type of optimisation algorithm and the algorithm's settings used to avoid getting trapped in local optima.

This is how SLAMS-2.0 treats some of the factors mentioned above. (a) The model has a complex construct and a relatively large parameter set. (b) Ecosystem parameters are difficult to quantify empirically as they are highly variable in time, space and across taxa. (c) Unlike data on dissolved inorganic tracers, the flux data that I am using to calibrate SLAMS-2.0 (POC, PIC and bSi fluxes), are still limited in time and space, as explained in Chapter 4. Data sparsity in time and space also make data aggregation and statistical treatment noisy, increasing the errors associated with data. SLAMS-2.0 finds itself in a situation where there are more parameters than observations available to constrain them.

[2]Overfitting is an issue that can occur when the number of parameters in a model exceeds the number of observations available to constrain them. It is in essence a memorisation of the training data (and any associated random noise). To illustrate it, think of a polynomial equation, where each degree we add (i.e., each one more parameter we add to our set) is one more flexor point that helps to better fit the data. The higher the degree of the polynomial equation, or the higher the number of parameters, the better fit to data.

Thus, more data than the currently available would be needed to determine whether SLAMS-2.0 reproduces observations for the right mechanistic reasons and not due to over parameterisation, which overfits the data. However, that problem cannot be dealt with in this thesis. (d) The model forcing variable NPP and the observations of POC, PIC and bSi fluxes are disjointed as they are not measured simultaneously. It is not ideal constraining a model output using forcing variables that are not connected to the model output, but I have no choice. (e and f) There exist more suited optimisers than DFO-LS to optimise SLAMS-2.0, like Bayesian approaches (e.g., Bisson et al., 2018), as well as misfit function designs with weighting factors that would factor in the different nature of the data that I am using; those could (or not) provide a more optimum calibrated parameter set. Nevertheless, I have opted for the simplest method to optimise SLAMS-2.0 as the first step towards more advanced optimisation methods that I will use in the future.

5.2 Parameter optimisation scheme

5.2.1 Definition variables and training dataset

To improve the performance of the particle cycling of SLAMS-2.0, the definition variables that I have chosen are annual means of POC flux ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$), PIC flux ($\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$) and bSi flux ($\mathrm{mmol\,Si\,m^{-2}\,d^{-1}}$). I have not included in the pool of definition variables BCP metrics, like Martin's b or T_{eff}, because those are post-processed fluxes that would confuse the optimiser's response. The corresponding observations for those three types of flux materials, which constitute the training data for the optimiser, were obtained for five locations with abundant flux data and that are representative of distinct oceanographic regimes (see Chapter 4): EqPac experiment site in the equatorial Pacific, Long-Term Ecological Research observatory HAUSGARTEN in the arctic Atlantic, Ocean Station Papa (OSP) in the subarctic Pacific, Porcupine Abyssal Plain time-Series Observatory (PAP-SO) in the North Atlantic and the Bermuda Atlantic Time-Series/Oceanic Flux Program (BATS/OFP) site in the subtropical North Atlantic.

Whereas SLAMS-2.0 can reproduce the surface-to-depth profile of those particulate flux tracers, available observations are usually restricted to a few key depths (export depth, sequestration depth and bathypelagic depth). Thus, I have chosen those three depth horizons to draw my model outputs and observations from. I have not interpolated the data between depths to avoid introducing bias to the observational dataset. Thus, I have a total of 45 (= 3 flux material types × 3 depths × 5 locations) data points to optimise SLAMS-2.0. The number of observations available to optimise SLAMS-2.0 gets surpassed by the number of model parameters (84) in which the model output pivots. Nevertheless, it is beyond the scope of this thesis to build a globally gridded dataset of particulate flux tracers.

5.2.2 The misfit function design

A misfit function quantifies the model–data mismatch. I use a two-step calculation to construct my misfit function: (1) calculation of the differences between the modelled flux outputs and the corresponding observations at each of the five ocean locations and (2) combination of those five local misfits into an overall misfit.

The function for the local misfits f_j at each ocean location j is a square root of a sum of weighted squared residuals:

$$f_j = \sqrt{\sum_{i=1}^{9} \left(p_{i,j} \frac{m_{i,j} - o_{i,j}}{\langle o_i \rangle} \right)^2}, \tag{5.1}$$

where $i = 1 \ldots 9$ are the nine flux variables (3 flux materials × 3 depths) available at each location j; $\langle o_i \rangle$ is a mean of all available observations of type i; the squared residual, $(m_{i,j} - o_{i,j})^2$, is the difference between $m_{i,j}$ model outputs of i variables at j locations and $o_{i,j}$ observations of the same i, j; $p_{i,j}$ is a coefficient that equals 1 if the output variable is present in the observation set or 0 otherwise. Eq. 5.1 does not have weighting factors, which appear in other misfit function designs (e.g., Schartau and Oschlies, 2003; Friedrichs et al., 2007; Ward et al., 2010), and are implemented based on subjective preferences of the authors; thus, I give all my observations equal weight, regardless of their associated error. That is the most straightforward approach to reducing the model-data misfit and avoiding introducing biases that negatively affect calibration results (Hemmings and Challenor, 2012). I calculate individual misfits f_j on a per-location j basis rather than a variable i basis. I have summed over output variables i rather than locations j because the correlation of variables at each location is higher than that of a specific variable across different locations. Said otherwise, the variation that a variable offers says more about the characteristics of the different locations than the characteristics of the flux materials. Note that Eq. 5.1 will emphasise shallow fluxes more because they have a bigger absolute value compared to deeper fluxes.

The shape of Eq. 5.1 is that of the square root of the sum of squared errors (SSE) normalised beforehand. It follows the design of Schartau and Oschlies (2003) and Kriest et al. (2017). Squaring the pairwise errors (or residuals) prevents positive and negative residuals from cancelling each other out in the summation, which could lead to underestimating the model–data misfit. It also amplifies larger residuals due to the quadratic growth of the squared function, which dominate the misfit value and therefore are penalised more strongly.

Next, f_j is passed onto the optimiser, which internally calculates the overall misfit, f, the sum of the squared f_j over the five studied ocean locations j,

$$f = \sum_{j=1}^{5} {f_j}^2. \tag{5.2}$$

An f close to 0 ($f^{\min}$) represents a good fit between model and observations and is the optimiser's target value. DFO-LS will search for $f^{\min}$ following a least-squares procedure. Based on previous misfit history from previous parameter configurations, DFO-LS tries to solve a curve fitting (or regression) problem, where the target curve is that with the least sum of all pairwise squared residuals, mathematically expressed as least $\sum_{i=1}^{N} (o_i - m_i(\boldsymbol{p}))^2$ (a mathematical construct, DFO-LS knows nothing about o nor m), where $\boldsymbol{p} = p_1, \ldots p_n$ is the parameter space created by the n parameters that are being tuned. DFO-LS will search the parameter space $\boldsymbol{p}$ as it reevaluates the least-squares model internally built. In the course of several iterations, the misfit f will converge to a minimum, $f \to f^{\min}$. Due to the highly non-linear nature of SLAMS-2.0, DFO-LS is likely to get trapped in one of many local minima (Fig. 5.1). Therefore, it is not possible to know if the optimised parameters $\boldsymbol{p}$ correspond to the true global minimum nor how good they are in relation to that global minimum. That is a limitation that local optimisers, like DFO-LS, face. However, there are mechanisms to avoid the optimiser getting stuck in a local minimum, like restarting the parameter space searches.

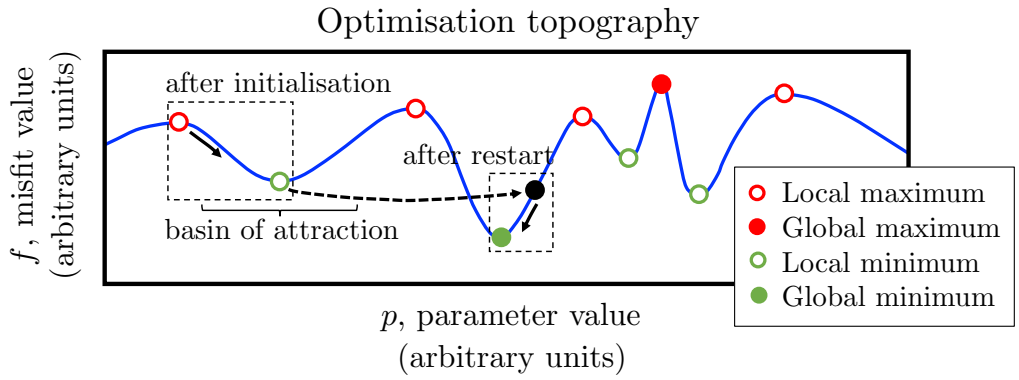


Figure 5.1. Schematic representation of a 1-D optimisation topography, where the misfit function is mapped to one parameter. Squares indicate hypothetic windows of action and arrows the direction of local search paths. A global minimum in the parameter search space is a point where the misfit value is smaller than any other point in the search space and a local minimum is a point where the misfit value is smaller than nearby points. Each misfit minimum has an associated basin of attraction. When the optimiser is initialised, the search starts anywhere in the optimisation topography and will only descend. A random restart is triggered after the optimiser gets stuck in a local minimum (black point). Source: redrawn and adapted from Hengenius et al. (2014).

5.2.3 Selection of the parameters to calibrate

In an ecosystem model, a parameter summarises a physiological or a biogeochemical process that has not been mechanistically described yet by a functional relationship. Parameters are often measured in idealised empirical settings, where plankton cultures are grown under controlled environmental conditions. That empirically estimated value must represent a broad spectrum of environmental conditions (from temperate to high

latitude systems, from the surface to the deep ocean) as well as a heterogeneous mixture of species within a plankton group. Parameters are thus the Achilles heel of models: black boxes of physiological/biogeochemical processes with a questionable representation of the actual process. Modellers often turn to model optimisation techniques to constrain highly uncertain parameters.

As our knowledge of ocean biogeochemistry increases, models incorporate more parameterised processes. A general trend has emerged: the number of unconstrained parameters exceeds the amount of data available to validate those. As the number of parameters increases, so does the mathematical flexibility (degrees of freedom) to better fit a model to a given data set, like a polynomial equation of a higher degree. Models with more parameters are more capable of reproducing specific data sets. However, they do a poor job reproducing new data from other locations where they have not been calibrated for. Thus, overparameterised models are expected to suffer from little portability and less explanatory power (Friedrichs et al., 2006).

The large number of ecosystem parameters in SLAMS-2.0 (Table C.1) is a common feature with most biogeochemical models (Table 5.1). Optimisation is usually applied on a subset of model parameters, those that are more fundamental to the model output, and which we expect the data available to constrain best. To get a ballpark estimate of how sensitive the model output is to a subset of 43 tuneable parameters (from the list of 84, as presented in Appendix C), I have calculated the misfit change of the overall misfit function for each tuneable parameter p_i, as:

$$\Delta f(p_i) = \left| f(p_i^{\text{upp}}) - f(p_i^{\text{low}}) \right|, \tag{5.3}$$

where f is the overall model-data misfit value at the upper p_i^{upp} and lower p_i^{low} boundary of a parameter p_i. Only the parameter of interest is changed to the target boundary; the rest of the parameters are kept fixed in their default value. The boundaries are estimates obtained from a literature research, with a posterior guess of their reasonable physiological limits. More robust methods exist for choosing the subset of parameters for optimisation (e.g., that presented in Friedrichs et al., 2006, 2007). However, the more first-hand approach of Eq. 5.3 is enough for a more ballpark assessment.

I have plotted the $\Delta f(p_i)$ results in Fig. 5.2, where I have cut off parameters that had a negligible effect on $\Delta f(p_i)$. In red are the seven parameters that I have selected for optimisation and have ranked higher on the $\Delta f(p_i)$ calculation. I have put the cut in seven after running twin experiments[3] varying the number of parameters between 6–11 to assess how parameter number could affect the outcome of the optimisation experiments.

[3] A twin experiment is a calibration experiment where real observations are replaced by synthetic observations.

Table 5.1. Number of ecosystem parameters in other common marine biogeochemical models. The number of parameters reported on this table comes from a recount after screening tables in the main text and/or appendixes.

Acronym	Full model's name	No. params.	Reference
BEC	Biogeochemical Elemental Cycling	175	Moore et al. (2001)
BLING	Biogeochemistry with Light, Iron, Nutrients and Gases	40	Galbraith et al. (2010)
COBALT2	Carbon, Ocean Biogeochemistry and Lower Trophics	80	Stock et al. (2020)
Diat-HadOCC	Hadley Centre Ocean Carbon Cycle	60	Totterdell (2019)
HAMOCC5.1	Hamburg Oceanic Carbon Cycle	90	Maier-Reimer et al. (2005)
M^4AGO	Microstructure, Multiscale, Mechanistic, Marine Aggregates in the Global Ocean	30	Maerz et al. (2020)
MEDUSA-2.0	Model of Ecosystem Dynamics, nutrient Utilisation, Sequestration and Acidification	75	Yool et al. (2013)
PlankTOM10	–	160	Le Quéré et al. (2016)
PISCES-v2	Pelagic Interactions Scheme for Carbon and Ecosystem Studies, volume 2	90	Aumont et al. (2015)
TOPAZ	Tracers of Ocean Phytoplankton with Allometric Zooplankton	80	Dunne et al. (2010)

Those twin experiments showed that it is not the number of parameters that affect the calibration performance of SLAMS-2.0, but their value at the start of the optimisation analysis. Seven is high enough to incorporate sufficient BCP processes and low enough to avoid model overfitting (Friedrichs et al., 2007).

The seven selected parameters are listed in Table 5.2. Two phytoplankton cell mineral quotas, $\theta^{\text{diat}}_{\text{Si:C}}$ and $\theta^{\text{cocco}}_{\text{calc:C}}$, with controls on the chemical composition, size and sinking velocity of the phytoplankton cells seeded. Two zooplankton grazing parameters, $f^{\text{ing, surf}}_{\text{OM}}$ and f_{sloppy}, the first with controls on the chance of a particle being ingested by zooplankton and therefore processed into faecal pellets and dead bodies, and the second with controls on the amount of food particles left behind as dissolved material during an ingestion event. Two zooplankton DVM parameters, $\theta_{\text{POC:Z}}$ and $\theta^{\text{DVM night}}_{\text{POC:Z}}$, with controls on the abundance of zooplankton in different depths in the water column. One bSi cycle-related parameter, κ_{opal}, with controls on the amount of bSi that returns to the dissolved pool.

5.2.4 The optimiser

DFO-LS (Cartis et al., 2019) is a derivative-free optimisation method that does not require the user to provide information on the derivatives (gradients) of the misfit function, nor

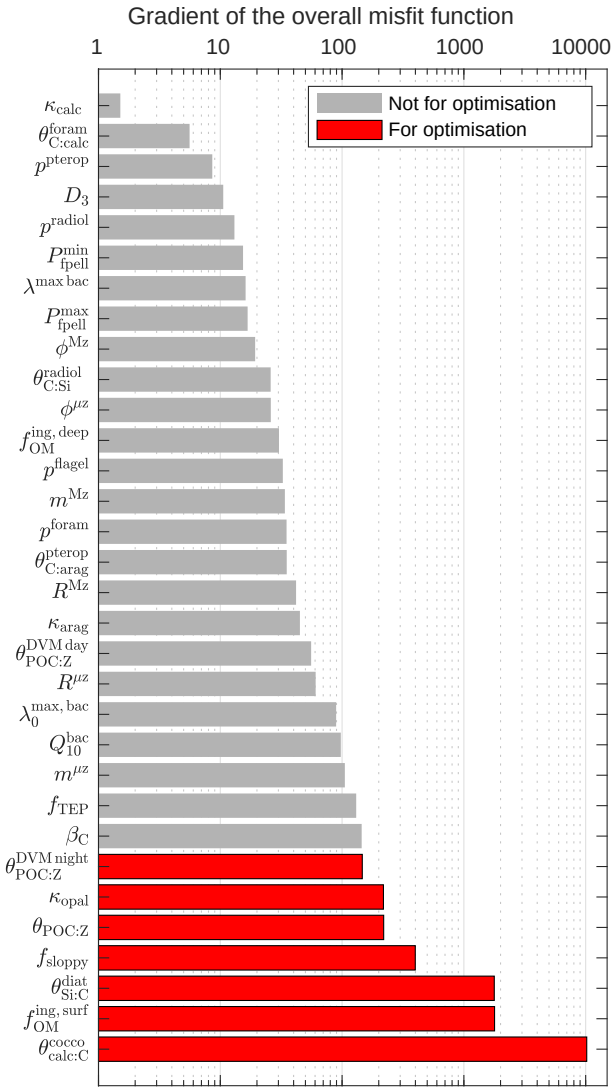


Figure 5.2. Parameter misfit sensitivity analysis showing the parameters ranked according to which ones the optimiser (DFO-LS) is more sensitive to. Highlighted in red are the parameters that will be subject to optimisation. See Table C.1 for the meaning of each parameter. Note the $\log_{10}$ axis.

Table 5.2. List of SLAMS-2.0 parameters subject to optimisation with their rank in the parameter sensitivity analysis, their default value and the boundary values.

Symbol	Meaning	Rank	Value	Boundaries	Units
$\theta^{\mathrm{diat}}_{\mathrm{Si:C}}$	Diatom cell Si:C ratio	3	0.30	0.03 – 0.65	$\mathrm{mol\,Si\,(mol\,C)^{-1}}$
$\theta^{\mathrm{cocco}}_{\mathrm{calc:C}}$	Coccolithophore cell PIC:POC ratio	1*	0.90	0.50 – 2.50	$\mathrm{mol\,PIC\,(mol\,POC)^{-1}}$
$\theta_{\mathrm{POC:Z}}$	Standard POC:zoo. biomass	5	3.0	1.5 – 10	–
$\theta^{\mathrm{DVM\,night}}_{\mathrm{POC:Z}}$	POC:zoo. biomass at the night DVM depth	7	0.50	0.20 – 2.0	–
f_{sloppy}	Sloppy feeding loss fraction	4	0.20	0.05 – 0.40	–
$f^{\mathrm{ing,\,surf}}_{\mathrm{OM}}$	Fraction of OM in an appetitive surface particle	2	0.30	0.10 – 0.80	–
κ_{opal}	Opal dissolution timescale	6	0.008	0.005 – 0.050	$\mathrm{d^{-1}}$

* Parameter to which the model flux output has more sensitivity to, as shown in Fig. 5.2.

does it attempt to estimate them internally (by using finite differencing, for instance). Instead, it navigates the multi-dimensional parameter space using least squares. The DFO-LS way of navigating the parameter space is prescribed explicitly for noisy model evaluations (for which the accurate evaluation of derivatives is impossible) or expensive function evaluations (so finite differencing is too costly). SLAMS-2.0 outputs are neither noisy (to optimise it, I am running it using a sequence of pseudo-random numbers, so repeated calls to the model produce the same outputs) nor function evaluations are costly (on average, a function evaluation with the optimiser running at five locations is completed within 1 h). Instead, DFO-LS has become convenient for SLAMS-2.0 for other reasons, mainly: (i) a least-square misfit problem is inherently faster than calculating the derivatives of my misfit function, (ii) DFO-LS allows imposing bound constraints on parameter values or outputs to guide the solution within theoretically justifiable limits, and (iii) DFO-LS has been shown to perform more efficiently than other derivative-free methods for global biogeochemical models (Oliver et al., 2021)[4]. DFO-LS is a local optimisation method, in that it finds local minima, and it is paired with a globalisation method to aide the optimiser to approach the global minimum. Typically this is done by restarting the optimiser from different initial parameter values and testing whether the optimiser converges to the same solution (Fig. 5.1).

I have obtained DFO-LS from the software package OptClimSO (Oliver, 2019), a Python interface that regulates the sequential workflow between the output files of a given climate model, the associated model parameter configuration and an optimiser (the user can choose between four: DFO-LS, Py-BOBYQA, Gauss-Newton, CMA-ES). OptClimSO and its documentation can be downloaded from https://github.com/SophyOliver/OptClim. See Appendix E for details on how I have implemented OptClimSO in SLAMS-2.0.

5.3 Results and discussion of the optimisation

The calibration of SLAMS-2.0 has involved, in first place, running so-called twin experiments, which aim to assess how suitable the optimiser is for calibrating SLAMS-2.0 and, in second place, optimisation experiments, where the parameters of SLAMS-2.0 are tuned in accordance to the model misfit to data. Each twin and optimisation experiment has 10 replicates that differ in the initial parameter configuration values (Table F.1), selected randomly using Matlab's `lhsdesign` function. The definition variables are particle fluxes of POC, PIC and bSi × 3 depths × 5 locations and the number of parameters to optimise is seven and listed on Table 5.2.

[4] Oliver et al. (2021) found that DFO-LS requires less function evaluations than CMA-ES to recover parameters from a model calibrated with synthetic data.

5.3.1 Twin experiments

A twin experiment is almost identical to an optimisation experiment. The difference is that in a twin experiment, real observations of the modelled system are replaced with synthetic data that a previous model run have produced. Therefore, in a twin experiment, the actual system is known (the synthetic observations), and since it comes from model-generated data, it should not be excessively complicated for the model to eventually reproduce it in the course of some optimiser generations. A twin experiment is expected to easily produce a misfit of 0, or very close to it, and therefore retrieve the exact reference parameter set or be very close to it.

Fig. 5.3 shows the ability of DFO-LS to retrieve a synthetic parameter configuration of a SLAMS-2.0 run in the form of parameter trajectory plots. Those show the trajectory that the values of the seven parameters chosen for model tuning take in the course of up to 255 generations. All the possible values taken by the parameters are realistic as parameters are bounded to physiological/biogeochemical known limits. The trajectories are displayed along with the target value, the one given to the parameter to produce the synthetic data set, and which a twin experiment aims to recover. The filled-in circles shown mark the generation number when the lowest misfit was achieved for a twin configuration. Not all parameter trajectories make it until the last 255th iteration as an experiment will stop when reaching the maximum number of unsuccessful restarts. When a low misfit occurs early, that means that the optimiser is getting trapped in local minimum basins (Fig. 5.1). All parameters except $\theta^{\text{diat}}_{\text{Si:C}}$ show a dispersed trajectory amongst the replicates, meaning that not all of them point towards the target value. For instance, $\theta^{\text{diat}}_{\text{Si:C}}$, the third-most sensitive parameter, has the 10 replicates approaching the synthetic target value after ~100 generations, and thus all hit the target regardless of the initial parameter configuration. For less sensitive parameters ($\theta_{\text{POC:Z}}$, $\theta^{\text{DVM night}}_{\text{POC:Z}}$, f_{sloppy} and κ_{opal}), the evolution of the trajectory does get influenced by the initial parameter value and are more difficult to displace from the initial trajectory towards the target.

$\theta^{\text{diat}}_{\text{Si:C}}$ presents the ideal *steep* misfit function shape, with a large decrease in either side of the target value, meaning that $\theta^{\text{diat}}_{\text{Si:C}}$ carries a lower uncertainty in its value or, otherwise said, a small deviation from its optimal value increases the model-data misfit. DFO-LS starts effectively tuning parameters after the first $N_{\text{parameters}} + 1$ generations so that after eight generations one would expect finding the steep phase of the misfit curve, and that is the case for the three most sensitive parameters ($\theta^{\text{cocco}}_{\text{calc:C}}$, $f^{\text{ing, surf}}_{\text{OM}}$, $\theta^{\text{diat}}_{\text{Si:C}}$). The least sensitive parameters present a *flat* misfit function shape, meaning that those parameters carry a higher uncertainty, otherwise said, a small deviation from their optimal value has little effect on the model–data misfit. Parameters with a flat misfit function shape can take any value in their bounded parameter space while providing a

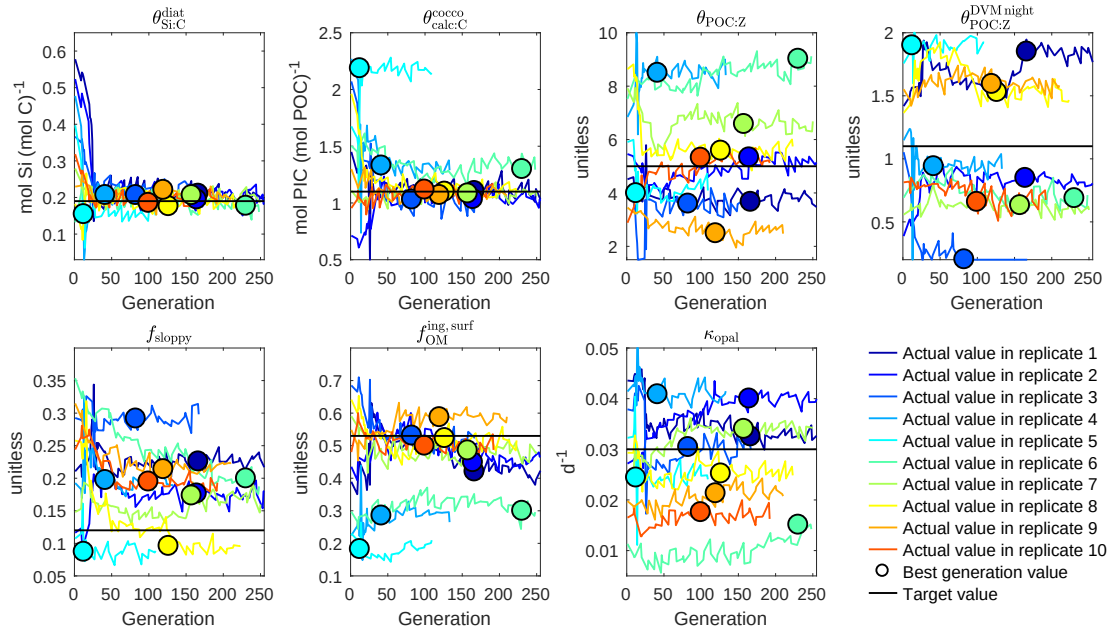


Figure 5.3. Parameter trajectory results of of the twin experiment, with the 10 experimental replicates. The evolution of the value of the seven parameters chosen for calibration is shown for up to 255 generations, or optimiser iterations. The y axis is limited to the boundaries of each parameter. The black horizontal line is the target value (i.e., the parameter's value in the synthetic observations set) and the filled-in circles mark the parameter's value at the best generation (i.e., the optimiser iteration with the lowest misfit). Parameter symbols are: $\theta^{diat}_{Si:C}$, diatom cell Si:C ratio; $\theta^{cocco}_{calc:C}$, coccolithophore cell PIC:POC ratio; $\theta_{POC:Z}$, standard POC:zooplankton biomass; $\theta^{DVM\,night}_{POC:Z}$, POC:zooplankton biomass at the night DVM depth; f_{sloppy}, sloppy feeding loss fraction; $f^{ing,\,surf}_{OM}$, fraction of OM in an appetitive surface particle; κ_{opal}, opal dissolution timescale.

similarly low model–data misfit (Ward et al., 2010). The rest of results produced by the twin experiment (overall misfit function reduction and fraction of the misfit attributed to each ocean site) are in Figs. F.1 and F.2. Replicate 7 achieves the lowest overall misfit, $f^{min} = 0.134$, a value that represents a percentage decline with respect to the first generation's misfit of 98.1% and it is achieved in the generation number 158.

The general outcome of the twin experiment is that DFO-LS is capable of dealing with the rough optimisation topography of SLAMS-2.0 when given enough initial parameter options and some 150 generations to iterate. Due to the rough optimisation topography of SLAMS-2.0, the optimiser gets trapped in several local minima. Thus, model calibration ends up having multiple solutions (best generations) as I have several sets of best parameters that produce local minima. That is the scenario that the calibration of SLAMS-2.0 poses.

5.3.2 Optimisation experiments

I have carried out optimisation experiments for two cases: (a) fitting model to data *simultaneously* for the five ocean locations (OSIM experiment) and (b) fitting model to

data *individually* on a location basis (OIND experiment). OIND experiments aim to inspect whether optimisation on a location basis produces BCP metrics more adjusted for each location than the simultaneous fitting in OSIM.

5.3.2.1 Misfit function minimisation for OSIM

Fig. 5.4 shows the overall misfit reduction in the OSIM experiment in each of the 10 replicates. Given the complexity of simultaneously adjusting model output to three types of flux data (POC, PIC and bSi), DFO-LS seems to do quite well reducing the overall model-data misfit. The misfit function shows higher initial misfits when compared to the mirror twin plot (Fig. F.1). Notice that not all replicates make it till the last generation (255) as the optimiser stops when it reaches a maximum number of unsuccessful restarts. Restarts are decided internally by the algorithm when it gets stuck in a local minimum and cannot reduce the misfit any further. The highest misfit reduction, $f^{\min} = 40.8$, occurs for replicate 4 at iteration 52, drawing a baseline that can be taken as the minimum misfit that DFO-LS is capable of achieving. That baseline is more than 10 times higher than that of the twin experiment ($f^{\min} = 0.134$) but more than 10 times lower than that of the default, non-optimised run ($f^{\min} = 776$, not shown here).

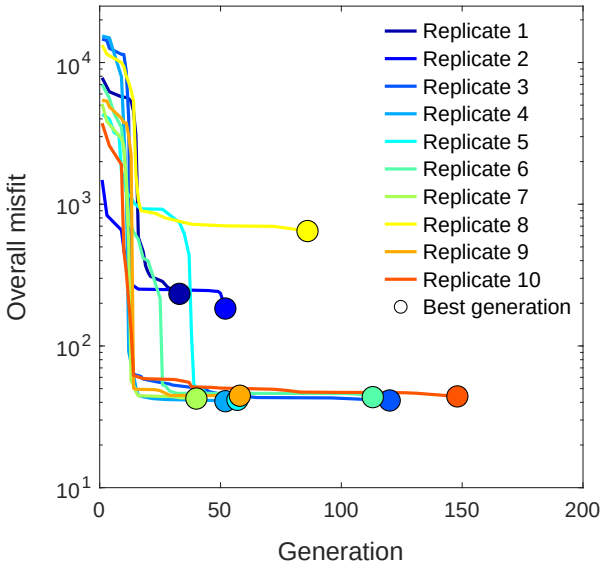
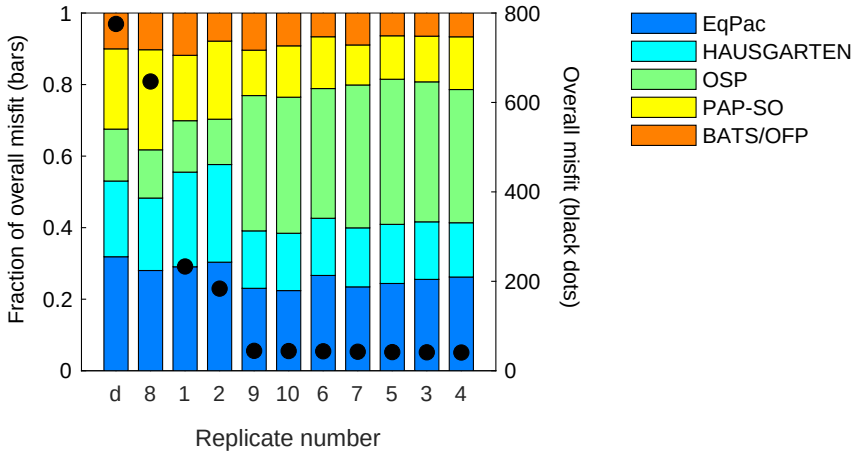


Figure 5.4. Overall misfit trajectory of the optimisation experiment, with the 10 experimental replicates. The evolution of the overall misfit is shown for up to 255 generations, or optimiser iterations. The filled-in circles mark the value of the best generation, or optimiser iteration with the lowest misfit, and the shown replicate trajectories stop there.

Fig. 5.5 shows the contribution of each location (Fig. 5.5a) and each pair material-depth of collection (Fig. 5.5b) to the overall misfit in the best iteration of each experimental replicate of OSIM. Black dots show the $f^{\min}$ reached by each of the optimisation replicates, which are more than 10 times higher than those in the mirror twin runs (Fig. F.2). In the best replicates, the pair OSP-bSi at export depth carry most of the misfit, whereas it is EqPac-PIC and EqPac-bSi at sequestration depth the variables carrying most of the misfit in the worst replicates.

(a) Fraction of the overall misfit explained by each location variable.



(b) Fraction of the overall misfit explained by each material-depth variable.

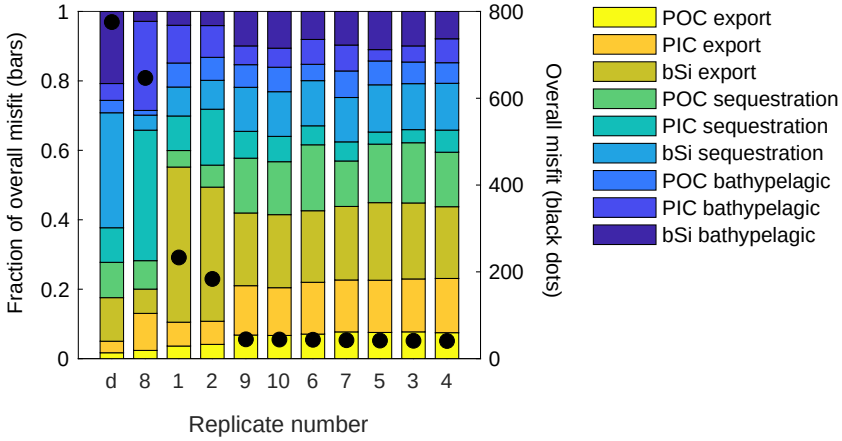


Figure 5.5. (a) Optimisation's best generation misfit and fraction of the overall misfit accounted by each location and (b) material-depth of settling pairs in each of the 10 experimental replicates. Replicates are ordered decreasingly by their overall misfit value. The default model run (d) is the first one being represented.

5.3.2.2 Misfit function minimisation for OIND

For the OIND runs, the misfit reduction trajectory is in Fig. F.3. OIND runs have only six replicates, which use the initial parameter configurations of the OSIM replicates 5–10. Given that, in an OSIM experiment, five local misfits are added together (compared to just one misfit term in an OIND experiment), the misfit reduction achieved by individually fitting each location to data will be greater than for simultaneous fitting over all sites. Thus, in an OSIM experiment, $f^{\min} = 40.8$, whereas for the OIND experiments: EqPac has $f^{\min} = 4.00$, HAUSGARTEN has $f^{\min} = 37.3$, OSP has $f^{\min} = 1.90$ and BATS/OFP has $f^{\min} = 0.82$. PAP-SO is the exception, with a large $f^{\min} = 198$ (larger than OSIM's $f^{\min}$), proving the hardest location to optimise individually (see Table F.2) and, thus, the one benefitting the most from the simultaneous fitting as its large OIND misfit gets offset

by the misfit of the locations with lower misfits, like BATS/OFP and OSP. The better model fit to BATS/OFP data results from that location providing the highest percentage of data points for the calibration dataset (50%, as shown in Fig. 4.4 in Chapter 4).

5.3.2.3 Optimised modelled fluxes vs observations

In Fig. 5.6a, the optimised parameter set produced by the OSIM replicate with the lowest misfit (replicate 4, $f^{\min}$ = 40.8) is used to visualise how the optimised model response fits to observations of annual means of POC, PIC and bSi flux. That response is compared with the fit achieved by individually fitting locations to data in each OIND experiment (each row of subplots in Fig. 5.6b).

A first glance at Fig 5.6a reveals that the model fit provided by the OSIM experiment (blue line) is similar to that provided by the OIND experiments. The more locations (data) an optimisation experiment incorporates, if it were overfitting, we would expect a worse fit, which is not observed here (there is equally good performance in OSIM compared to OIND). That suggests that the model is not overfitting data for individual stations in the OSIM experiment and that DFO-LS can be applied in SLAMS-2.0 to simultaneously fit data across a variety of locations with distinct biogeochemistry. Fig 5.6a further illustrates that it is not possible to get a *perfect* fit to all three types of fluxes for all five locations simultaneously. The most considerable improvement in the OSIM model response occurs for bSi flux, for which all locations displayed a default value a lot higher than the observed. Regarding the POC flux response, all locations have got a better fit to data at intermediate and deeper depths, but not so much at the ocean's surface. The optimised PIC flux response is a lot more smooth than the default one. Observed surface PIC flux at EqPac and OSP are markedly higher than at deeper depths, a feature that the optimised model response does not capture. The OIND fits in Fig. 5.6b are not markedly better for POC fluxes than the OSIM fit, but an improvement is seen at OSP and BATS/OFP. PIC flux fits for OIND are also slightly more in line with observations than for OSIM, but cannot reproduce the surface peaks at EqPac and OSP. The same can be noted for bSi fluxes, with OIND fitting observations better.

None of the OSIM and OIND runs can reproduce surface POC fluxes as high as those observed (except HAUSGARTEN), which suggests that low modelled surface POC flux could result from an NPP forcing that is too low to reproduce the observed POC fluxes. To test whether the choice of the NPP data set used to run SLAMS-2.0 affects its ability to simulate high surface POC fluxes, I have forced SLAMS-2.0 with four other satellite-based NPP models (VGPM, CbPM, CAFE and SeaWiFS) as well as *in situ* NPP data for BATS/OFP and OSP (data as presented in Table 4.3) using the calibrated parameter set produced by the best OSIM experiment. The plotted results (Fig. 5.7)

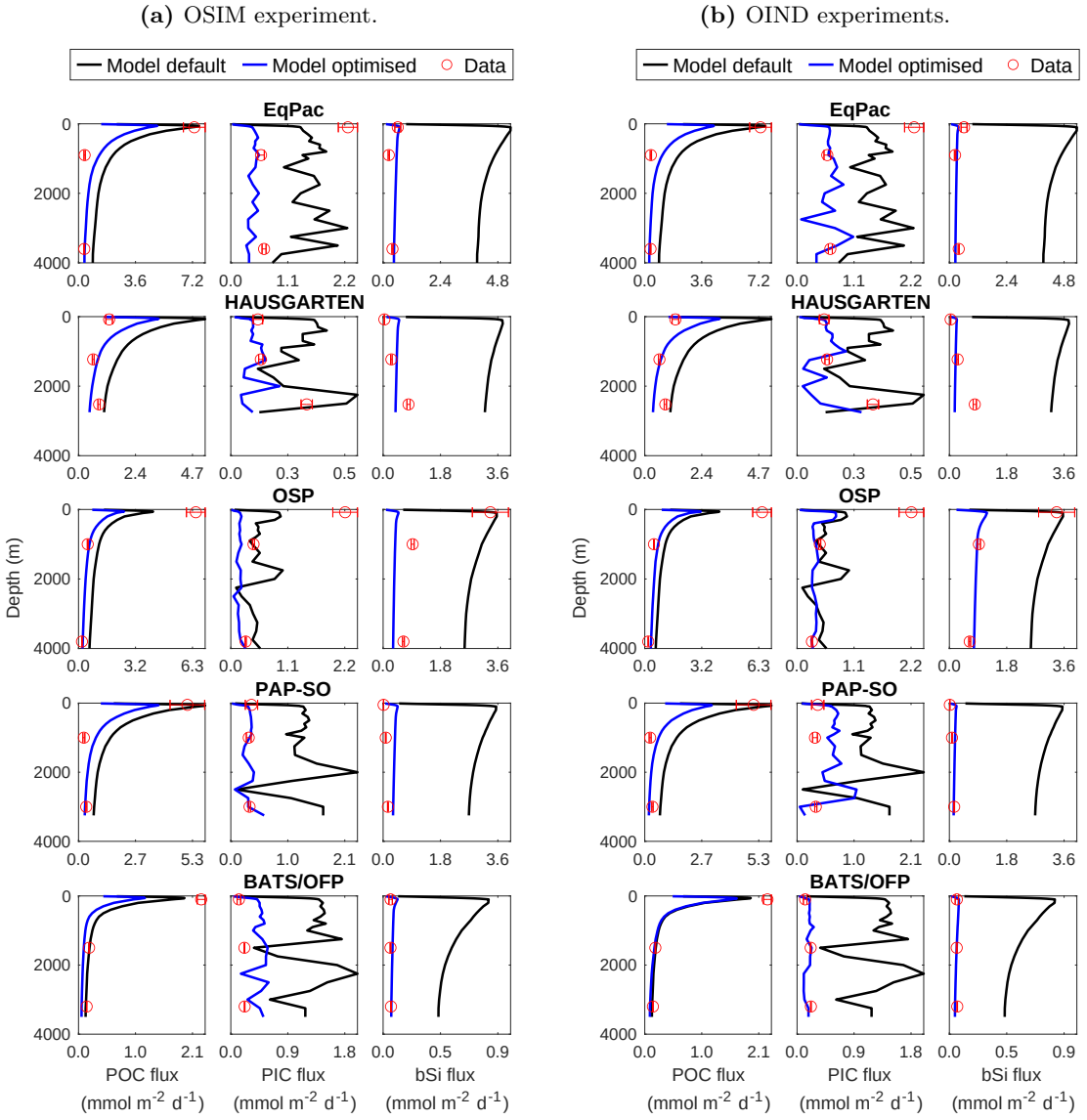


Figure 5.6. Modelled vs observed annual means of POC, PIC and bSi fluxes for (a) the simultaneous fit to the five ocean locations (OSIM) and (b) the individual fit to each ocean location (OIND). Two model instances are shown: the default response (black line) and the optimised (blue line). Data (red circles) display the standard deviation. The default run is based on the best guess of the parameter set before optimisation. The misfits associated are: OSIM, f = 40.8; default model run, f = 776; EqPac OIND, f = 4.00; HAUSGARTEN OIND, f = 27.3; OSP OIND, f = 1.90; PAP-SO OIND, f = 198; BATS/OFP OIND, f = 0.82.

suggest that *in situ* estimates of NPP improve the model fit to observed POC fluxes. POC flux sourced from satellite estimates of NPP lags slightly behind the fuelled by *in situ* NPP (for the two time-series sites where that comparison can be made) and presents misfit variability depending on the choice of satellite-based model. That test shows that model performance is sensitive to choice of the NPP data set used to force the model as primary production is the source of POC. Unfortunately, *in situ* NPP forcing, which naturally keeps a close matchup with *in situ* POC observations, is not available for all

locations nor on a global gridded format, as NPP satellite-based estimates are. Thus, it is not possible to use it to scale up model fit to data.

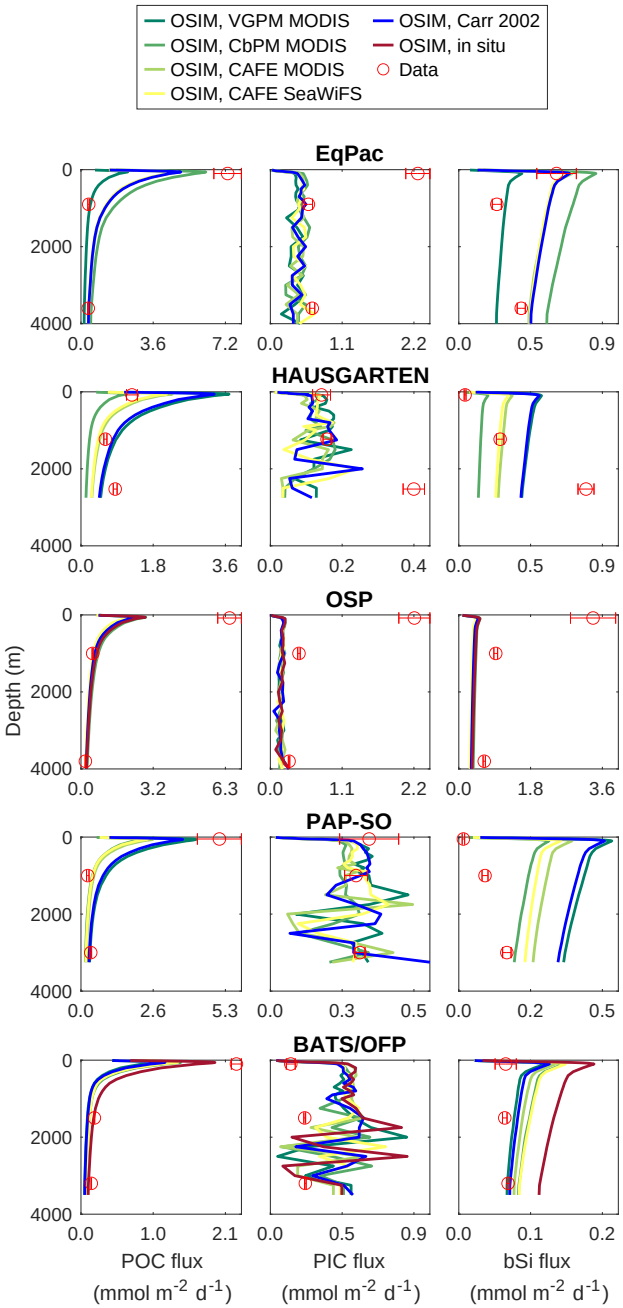


Figure 5.7. Effect of different NPP forcings (five satellite-based estimates and *in situ* estimates) on the modelled POC, PIC and bSi fluxes. Runs use the optimised parameter set produced by the best OSIM experiment (simultaneous fit to data from the five locations using the NPP forcing of the Carr (2002) model). The satellite-based estimates of NPP are from: the VGPM (Vertically Generalized Productivity Model) and the CbPM (Carbon-based Production Model-2), both based on MODIS-Aqua products; the CAFE model, based on MODIS-Aqua and SeaWiFS products, and the Carr (2002) model, based on SeaWiFS and AVHRR products. *In situ* NPP estimates with a complete annual cycle were only available for OSP and BATS/OFP.

5.3.2.4 Optimised parameter set

An illustration of the spread presented by the values of the most optimum parameters that DFO-LS has found is shown in Fig. 5.8. It collects the parameter values and misfits associated to the best OSIM and OIND experiments shown in Fig. 5.6 (exact parameter and misfit values listed in Table F.2) as well as the next six best OSIM experiments, which had a similar minimum misfit of around 43 (see the seven runs hitting the baseline in Fig. 5.4).

Fig. 5.8 displays how the parameters to which DFO-LS manifested highest sensitivity ($\theta^{\text{cocco}}_{\text{calc:C}}$, $f^{\text{ing, surf}}_{\text{OM}}$, $\theta^{\text{diat}}_{\text{Si:C}}$ and f_{sloppy}, as per Fig. 5.2) have a rather consistent low/high value across runs, whereas the remnant less sensitive parameters ($\theta_{\text{POC:Z}}$, κ_{opal}, $\theta^{\text{DVM night}}_{\text{POC:Z}}$) have values inconsistently spread along the parameter range. That indicates that DFO-LS has a more steady response (i.e., it is more predictable where it is going to) regarding parameters that show a higher sensitivity to the optimiser. For those more sensitive parameters, DFO-LS finds a more consistent pattern across runs (e.g., all $\theta^{\text{diat}}_{\text{Si:C}}$ values are in the lower boundary) than when optimising less sensitive parameters, for which DFO-LS gets inconsistently stuck in the upper or lower boundary of the parameter range (e.g., κ_{opal}'s experiment 4 vs experiment 5). The misfits associated to the OSIM experiments (in red, varying between 40.8–44.5) cannot be compared with the misfits of the OIND experiments (in maroon, varying between 0.82–198) because the OIND experiments have less terms in the misfit function construct (45 for OSIM vs 9 for OIND), which will result in a relatively smaller misfit.

Overall, DFO-LS finds better fit model–data for these parameter conditions (after the prescribed boundaries):

- less silicified diatoms (with $\theta^{\text{diat}}_{\text{Si:C}} < 0.15$ mol Si (mol C)$^{-1}$), in agreement with the values of 0.11 mol Si (mol C)$^{-1}$ (Sarthou et al., 2005) and 0.13 mol Si (mol C)$^{-1}$ (Brzezinski, 1985) reported for general diatoms;
- less calcified coccolithophores (with $\theta^{\text{cocco}}_{\text{calc:C}} < 1.2$ mol PIC (mol POC)$^{-1}$), also in agreement with the values reported by Daniels et al. (2014) for the two most ubiquitous coccolithophore species in the ocean, *Emiliania huxleyi* ($\theta^{\text{cocco}}_{\text{calc:C}} = 0.70$ mol PIC (mol POC)$^{-1}$) and *Coccolithus* sp. ($\theta^{\text{cocco}}_{\text{calc:C}} = 1.20$–$1.55$ mol PIC (mol POC)$^{-1}$);
- higher amount of POC lost due to sloppy feeding fragmentation ($f_{\text{sloppy}} > 25\%$);
- surface ocean particles do not require a high C_{org} content to be deemed appetitive by hungry zooplankton ($f^{\text{ing, surf}}_{\text{OM}} = 0.10$);
- zooplankton biomass is, on average, 1/7 of POC biomass ($\theta_{\text{POC:Z}}$ ~7);
- at the night DVM depth, zooplankton biomass constitutes, on average, 77% of POC biomass ($\theta^{\text{DVM night}}_{\text{POC:Z}}$ ~1.3).

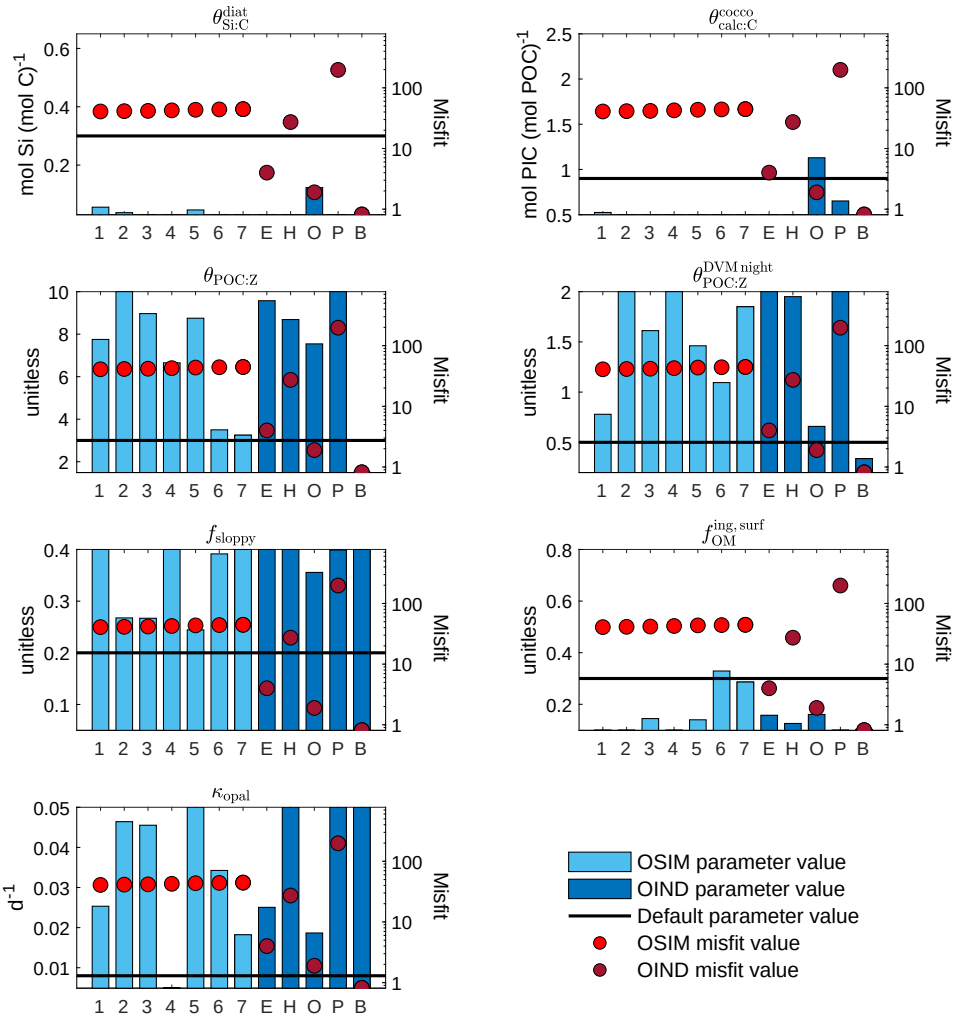


Figure 5.8. Comparison of the parameter values and associated model–data misfits (f) given by the seven SLAMS-2.0's parameters optimised. Shown are the seven best seven experimental replicates of OSIM (model fit to all available data simultaneously) and each best replicate of OIND (model fit to local data individually). For reference, the default parameter value, i.e., my best guess for the non-optimised run, is shown on top (black horizontal line). The misfits associated to the OSIM runs (red dots) and OIND runs (maroon dots) are the same for each parameter subplot. The three OSIM light blue bars (1–7) correspond to the experimental replicates 4 (f = 40.8), 3 (f = 41.2), 5 (f = 41.7), 7 (f = 42.5), 6 (f = 43.5), 10 (f = 44.1), 9 (f = 44.5). The dark blue OIND bars correspond to EqPac's (E) replicate 9 (4.0), HAUSGARTEN's (H) replicate 5 (f = 27.3), OSP's (O) replicate 6 (f = 1.9), PAP-SO's (P) replicate 7 (f = 198) and BATS/OFP's (B) replicate 8 (f = 0.82). The left-hand side y axis is limited to the boundaries of each parameter and the right-hand side, in $\log_{10}$, is bounded between 0.82–776 (the minimum f found in the optimised runs and the f of the default, non-optimised, run). Parameter symbols are: $\theta^{\text{diat}}_{\text{Si:C}}$, diatom cell Si:C ratio; $\theta^{\text{cocco}}_{\text{calc:C}}$, coccolithophore cell PIC:POC ratio; $\theta_{\text{POC:Z}}$, standard POC:zooplankton biomass; $\theta^{\text{DVM night}}_{\text{POC:Z}}$, POC:zooplankton biomass at the night DVM depth; f_{sloppy}, sloppy feeding loss fraction; $f^{\text{ing, surf}}_{\text{OM}}$, fraction of OM in an appetitive surface particle; κ_{opal}, opal dissolution timescale.

From that analysis, I can conclude that it is not possible to identify a unique parameter set solution after assimilating particle flux data for multiple locations with SLAMS-2.0.

That agrees with the conclusions reached by Gehlen et al. (2006) and Kriest and Oschlies (2013) when calibrating global models against individual observations of particle flux. However, the current calibration analysis cannot distinguish whether that is due to (i) the few number of observations with which SLAMS-2.0 has been calibrated, (ii) the use of data that, once averaged to smooth out bloom periods (i.e., observed surface peaks of POC, PIC and bSi fluxes), are still not representative of a climatological average, the period of time in which the model forcing operates, (iii) the fact that SLAMS-2.0 might be missing or incorrectly representing some significant BCP processes, or (iv) the naturally non-linear response of a complex model like SLAMS-2.0, with the optimiser getting trapped in local optima.

5.4 Results and discussion of the validation calculations

Validation experiments are commonly carried out to test whether an optimised parameter set can fit well a new set of data withhold from the optimisation runs. However, due to the scarcity of new POC, PIC and bSi flux data available to carry out a validation analysis and the still not perfect fit of the optimised model to data (Fig. 5.6), I will validate the optimised parameter set calculating BCP metrics that are of general use in the literature. Thus, for the five ocean locations SLAMS-2.0 has been calibrated with, I have calculated various relevant metrics of the BCP, as defined in Sec. 1.1.5: b, z^*, $\mathrm{PE}_{\mathrm{eff}}$, $\mathrm{T}_{\mathrm{eff}}$, molar ratio of POC/PIC flux, molar ratio of bSi/PIC flux and ξ.

5.4.1 Calculation of SLAMS-2.0's emergent properties

5.4.1.1 Particle fluxes

The calculation of BCP-related metrics in SLAMS-2.0 needs, first of all, calculating the flux associated with the rain of sinking Lagrangian particles. Surface-to-depth flux profiles of POC ($\mathrm{C}_{\mathrm{org}}$+TEC), PIC (calcite+aragonite) and bSi are calculated prognostically at the end of each month from the ensemble of Lagrangian particles crossing virtual sediment trap deployment depths, spaced out every 10 m for the top 100 m of the water column, then every 100 m for the mesopelagic (100–1000 m) and next every 250 m for the bathypelagic (1000–4000 m). The flux of the material is calculated as it would be from a real oceanic trap:

$$\begin{aligned}&\text{flux of material } (\mathrm{mg\,m^{-2}\,d^{-1}})\\ &\quad = \text{material in trap } (\mathrm{mg}) \cdot \frac{1}{\text{catchment area}}(\mathrm{m^{-2}}) \cdot \frac{1}{\text{deployment time}}(\mathrm{d^{-1}}),\end{aligned} \tag{5.4}$$

where the trap's catchment area in SLAMS-2.0 corresponds to the surface area of a modelled water column and the deployment time is 30 days. Fluxes are thus normalised

monthly. SLAMS-2.0 does not have 3-D movement of particles, which could actually disperse the depth position of particles in a horizontal area of ~100–500 km (Siegel and Deuser, 1997). But since the modelled water columns were particles are gravitationally sinking are wide enough (~280 km), it can be assumed that the catchment area of my virtual sediment traps are large enough to encompass particles that could have been transported horizontally, and thus we are correct assuming that horizontal influxes of particles are negligible compared to their vertical gravitational sinking. Once the monthly fluxes are computed, I compute from those some BCP metrics (POC/PIC, bSi/PIC, T_{eff} and PE_{eff}). I then calculate the annually averaged fluxes. The annually averaged BCP metrics are of two types: those computed from the monthly BCP metrics (POC/PIC, bSi/PIC, T_{eff} and PE_{eff}) and those computed from the annually averaged fluxes (b, z^* and ξ).

5.4.1.2 Particle spectra

The calculation of particle abundances are also prognostic at the end of each month from the ensemble of particles crossing virtual imaging systems deployment depths with the same spacing as the sediment traps,

$$\begin{aligned}&\text{number concentration of particles in bin } d \text{ (num.}\,\mathrm{L}^{-1}\text{)}\\ &= \text{no. particles imaged for } d \text{ (num.)} \cdot \frac{1}{\text{searched volume}}(\mathrm{L}^{-1}) \cdot \frac{1}{\text{no. picture frames}},\end{aligned} \tag{5.5}$$

where the so-called bins are the diameter categories d in which particles are binned (16 bins between 0.2 and $>$ 32,768 μm, following EcoTaxa's classification), and the volume searched by the imaging system is the total sampled height around the deployment depth times the area of a modelled water column. The imaging system samples the ±5 m depth boundaries around the deployment depths of the top 100 m of the water column, then the ±50 m depth boundaries for the mesopelagic deployment depths and next the ±125 m depth boundaries for the bathypelagic deployment depths. The number of picture frames is the number of times the imaging system has "photographed" particles and equals 90 (the number of time steps in a day times 30 days).

To build discrete particle size number distributions (PSD), I have followed the power-law model of Eq. 1.6, where $N(d)$ (num. $\mathrm{L}^{-1}\,\mu\mathrm{m}^{-1}$) is calculated by dividing the number concentration of particles in a bin by the bin width d, and the equation's reference diameter d_0 corresponds to that bin size with the largest particle count. For those bins with particle counts ≤ 1, particles have been added into one last bin. The actual range of size bins used in SLAMS-2.0 oscillates between 0.2 and 4100 μm, thus extending to the smallest possible particle size and limited in the upper boundary by the maximum size permitted after applying the collision kernels of Jackson (2001).

5.4.2 Comparison of modelled metrics of the BCP against observations

In Table 5.3, I show how the metrics of the BCP compare for different cases: (i) SLAMS-2.0 before and after optimisation, (ii) the observations data set, (iii) four empirical models (Guidi et al., 2015; Henson et al., 2012; Marsay et al., 2015; Dunne et al., 2005) and (iv) published values from fieldwork. In order to compare SLAMS-2.0's modelled metrics with the outputs of published empirical models, I have recalculated the metrics of Henson et al. (2012), Marsay et al. (2015) and Dunne et al. (2005) using the forcing climatologies I had available, which can differ in period of time from the ones used originally by the authors as some climatologies have been superseded from online repositories (see explanation below). In the case of SLAMS-2.0 and the empirical models, the BCP metrics are annual averages, whereas observed metrics (Other) can be anything between a one-time measurement, a monthly average, a seasonal average or an annual average.

Guidi et al. (2015) estimated b from a global collated sample of carbon flux profiles reconstructed from a median of (i) UVP-observed PSDs (0–1000 m) and (ii) measurements of $^{238}U/^{234}Th$ disequilibrium combined with deep sediment traps. The local median flux values were then regionalised according to Longhurst (1995)' 56 ocean biogeographic provinces. Henson et al. (2012) estimates of b, PE_{eff} and T_{eff} are based on the combination of three algorithms (Carr (2002) to estimate NPP, Henson et al. (2011)'s statistical fit to SST to compute PE_{eff} from observations of ^{234}Th-derived POC fluxes, and Lutz et al. (2007)'s fit to NPP to compute POC flux at 2000 m from deep sediment trap flux observations) and a global database of POC fluxes (^{234}Th-derived POC export fluxes from various authors and sediment trap-derived POC fluxes at 2000 m from Honjo et al. (2008)). I have recalculated Henson et al. (2012)'s metrics using climatological data from their same sources but slightly different periods: SST from AVHRR Pathfinder version 5.0 global monthly climatology for 1985–2001 (https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.nodc:AVHRR_Pathfinder-NODC-v5.0-climatologies) and chl a and PAR_0 from SeaWiFS monthly climatology for the whole satellite lifespan, 1997–2010 (hhttps://oceandata.sci.gsfc.nasa.gov/SeaWiFS). Dunne et al. (2005)'s PE_{eff} is based on a statistical fit to T and chl a after a global compilation of export measurements from a combination of shallow sediment traps and ^{234}Th. I have recalculated it using the equation reported in the supplementary material of Henson et al. (2012), which slightly differs from Dunne et al. (2005)'s original formulation, and have used the same SST and chl a data sets as per Henson et al. (2012). Marsay et al. (2015)'s global b and z^* are calculated from a statistical fit to average upper water (500 m) T that was constructed using sediment trap POC flux data from four North Atlantic and four North Pacific locations. I have recalculated Marsay et al. (2015)'s metrics using seawater T from WOA18 (https://accession.nodc.noaa.gov/NCEI-WOA18), which I have averaged for the upper 500 m of the water column.

Table 5.3. Comparison of relevant metrics of the BCP (b, z^*, PE_{eff}, T_{eff}, molar ratio of POC/PIC flux, molar ratio of bSi/PIC flux and ξ) at five ocean locations (EqPac, HAUSGARTEN, OSP, PAP-SO and BATS/OFP) in different cases: (i) SLAMS-2.0 before (DEF) and after (OSIM, OIND) optimisation, (ii) the observations data set (particulate flux and particle number concentration, as per Chapter 4), (iii) four published empirical models (Guidi et al., 2015; Henson et al., 2012; Marsay et al., 2015; Dunne et al., 2005) and (iv) available on-site fieldwork values (Other). Values are expressed as mean annual values (± error) [R^2-adj]. b is calculated from the power-law fit in Eq. 1.4 and z^* is calculated from the exponential fit in Eq. 1.5. Both metrics use the reference depth z_0 = 100 m in SLAMS-2.0 and the location's z_{eu} for the observations. PE_{eff} uses NPP estimates from the Carr (2002) model. In SLAMS-2.0, POC is C_{org} + TEC and PIC is calcite + aragonite. The depths z_{eu}, z_{meso} and z_{bathy} represent are in Table 4.1. ξ is calculated using the Junge-type distribution in Eq. 1.6. Henson et al. (2012) values are split into two: as per Guidi et al. (2015) recalculation (first) and as per my recalculation (second).

BCP metrics	SLAMS-2.0 results			Observations	Published empirical models				Other
	OSIM	OIND	DEF		Guidi	Henson	Marsay	Dunne	
				EqPac (0°N, 140°W)					
b	0.60 [0.95]	0.58 [0.95]	0.53 [0.98]	1.25 [1.0]	0.69	0.48/0.43	1.50		0.59–0.81, 1.03, 1.35–1.69[c,f,i]
z^*	650 [0.92]	612 [0.94]	681 [0.82]	272 [1.0]			116		217, 307–495[f,l]
PE_{eff}	0.06 (<0.01)	0.05 (<0.01)	0.10 (<0.01)	0.095 (<0.01)		0.03		0.09	0.020, 0.140[f,e]
T_{eff} ($z_{eu} \rightarrow z_{meso}$)	0.30 (0.02)	0.28 (0.02)	0.31 (0.02)	0.074 (0.033)					0.35, 0.61[i,f]
T_{eff} ($z_{meso} \rightarrow z_{bathy}$)	0.09 (<0.01)	0.08 (0.01)	0.13 (0.01)	0.041 (<0.01)		0.28			0.077–0.125[c]
POC/PIC (z_{eu})	11.9 (5.1)	6.6 (1.3)	5.5 (1.3)	9.5 (0.7)					
POC/PIC (z_{meso})	7.0 (10.8)	9.0 (13.5)	9.0 (9.9)	0.67 (0.01)					
POC/PIC (z_{bathy})	7.7 (4.5)	10.3 (12.1)	8.5 (8.2)	0.57 (<0.01)					0.60[m]
bSi/PIC (z_{eu})	1.9 (0.7)	0.63 (0.10)	4.1 (0.8)	0.28 (0.02)					
bSi/PIC (z_{meso})	3.0 (4.5)	2.7 (4.1)	19.3 (21.5)	0.41 (0.01)					
bSi/PIC (z_{bathy})	9.0 (5.2)	8.9 (10.8)	34.7 (33.3)	0.63 (0.01)					0.70[e]
ξ (50 m)	2.6 [1.0]	2.3 [1.0]	2.6 [1.0]	–					3.6–4.1[g]
ξ (100 m)	2.5 [1.0]	2.6 [1.0]	2.0 [1.0]	–					
ξ (500 m)	13.9 [1.0]	12.4 [1.0]	8.6 [1.0]	–					

Table 5.3 (Continued) Comparison of metrics of the BCP.

BCP metrics	SLAMS-2.0 results			Observations	Published empirical models				Other
	OSIM	OIND	DEF		Guidi	Henson	Marsay	Dunne	
HAUSGARTEN (79°N, 4.3°E)									
b	0.49 [0.96]	0.52 [0.94]	0.42 [0.98]	0.27 [1.0]		0.89/0.98	0.38		1.61[c]
z^*	745 [0.89]	679 [0.93]	938 [0.81]	1561 [1.0]			460		
PE_{eff}	0.05 (0.03)	0.05 (0.03)	0.08 (0.05)	>1 (0.95)		0.22		0.48	0.09–1.3[n]
T_{eff} ($z_{eu} \rightarrow z_{meso}$)	0.18 (0.11)	0.16 (0.09)	0.23 (0.13)	>1 (0.24)					
T_{eff} ($z_{meso} \rightarrow z_{bathy}$)	0.11 (0.07)	0.09 (0.05)	0.16 (0.09)	>1 (0.30)					
POC/PIC (z_{eu})	42.0 (50.3)	19.8 (16.2)	15.5 (16.9)	11.3 (0.25)					6.7–6.9[m]
POC/PIC (z_{meso})	13.9 (16.3)	139 (457)	14.2 (15.7)	5.0 (0.04)					2.8[e]
POC/PIC (z_{bathy})	9.5 (8.7)	7.6 (7.3)	11.5 (11.1)	3.1 (0.03)					
bSi/PIC (z_{eu})	7.0 (9.1)	1.7 (1.3)	12.5 (15.6)	0.32 (<0.01)					
bSi/PIC (z_{meso})	9.7 (9.8)	44.4 (145)	166 (447)	1.9 (0.02)					1.2[e]
bSi/PIC (z_{bathy})	10.4 (8.3)	3.8 (3.6)	338 (1059)	2.7 (0.03)					
ξ (50 m)	5.9 [1.0]	0.94 [0.79]	4.8 [1.0]	1.7 [1.0]					
ξ (100 m)	1.7 [0.99]	1.4 [0.98]	1.1 [0.93]	2.0 [1.0]					
ξ (500 m)	4.2 [1.0]	–	8.5 [1.0]	2.2 [1.0]					
OSP (50°N, 145°W)									
b	0.53 [0.98]	0.51 [0.99]	0.49 [1.0]	0.99 [1.0]		0.87/0.83	0.67		0.80, 1.05–1.16, 1.17–1.98[c,f,i]
z^*	696 [0.82]	714 [0.70]	778 [0.53]	362 [0.99]			371		77–78, 81[f,l]
PE_{eff}	0.06 (<0.01)	0.08 (<0.01)	0.11 (<0.01)	0.16 (0.01)		0.11		0.27	0.03, 0.09, 0.13, 0.03–0.14[k,a,l,f]
T_{eff} ($z_{eu} \rightarrow z_{meso}$)	0.26 (0.01)	0.26 (0.02)	0.29 (0.02)	0.12 (0.01)					0.10, 0.24, 0.16–0.32, 0.56[i,l,f,k]
T_{eff} ($z_{meso} \rightarrow z_{bathy}$)	0.11 (0.01)	0.12 (<0.01)	0.16 (0.01)	0.04 (<0.01)		0.08			0.06[k]

Table 5.3 (Continued) Comparison of metrics of the BCP.

BCP metrics	SLAMS-2.0 results			Observations	Published empirical models				Other
	OSIM	OIND	DEF		Guidi	Henson	Marsay	Dunne	
POC/PIC (z_{eu})	15.5 (5.6)	4.1 (1.2)	4.5 (1.1)	13.2 (0.41)					
POC/PIC (z_{meso})	5.2 (5.4)	2.4 (1.7)	5.5 (9.1)	1.7 (<0.01)					
POC/PIC (z_{bathy})	9.8 (7.3)	12.8 (19.4)	17.7 (27.0)	0.75 (<0.01)					2.0[e]
bSi/PIC (z_{eu})	3.3 (1.1)	1.7 (0.6)	3.98 (0.93)	4.7 (0.16)					
bSi/PIC (z_{meso})	3.2 (3.2)	2.7 (1.8)	14.6 (23.7)	2.7 (0.02)					
bSi/PIC (z_{bathy})	13.8 (10.4)	29.0 (41.4)	74.5 (115.5)	2.5 (0.01)					8.8[e]
ξ (50 m)	0.90 [0.95]	0.85 [0.93]	3.2 [1.0]	2.1 [1.0]					3.8–4.0[g]
ξ (100 m)	1.0 [0.98]	0.61 [0.80]	0.69 [0.83]	1.4 [0.99]					
ξ (500 m)	4.2 [1.0]	3.5 [1.0]	2.1 []0.99	2.6 [1.0]					
PAP-SO (49°N, 16.5°W)									
b	0.58 [0.95]	0.60 [0.94]	0.54 [0.97]	1.01 [1.0]	0.92	0.69/0.67	1.05		1.7[d]
z^*	587 [0.93]	560 [0.96]	646 [0.87]	315 [1.0]			255		154–291[l]
PE_{eff}	0.06 (<0.01)	0.05 (<0.01)	0.10 (0.01)	0.07 (0.01)		0.08		0.24	0.06–0.12[h]
T_{eff} ($z_{eu} \rightarrow z_{meso}$)	0.21 (0.04)	0.18 (0.04)	0.24 (0.06)	0.09 (0.04)			0.41		
T_{eff} ($z_{meso} \rightarrow z_{bathy}$)	0.09 (0.02)	0.06 (0.01)	0.13 (0.02)	0.16 (0.08)		0.14			
POC/PIC (z_{eu})	25.8 (17.3)	7.3 (5.0)	6.6 (4.1)	28.2 (0.54)					
POC/PIC (z_{meso})	5.6 (6.7)	4.3 (6.3)	7.7 (9.1)	1.2 (0.02)					
POC/PIC (z_{bathy})	4.6 (4.6)	5.0 (2.8)	8.9 (7.7)	1.2 (<0.01)					3.2[e]
bSi/PIC (z_{eu})	2.7 (1.6)	0.42 (0.24)	3.2 (1.9)	0.07 (<0.01)					
bSi/PIC (z_{meso})	3.0 (4.2)	1.5 (2.0)	19.1 (25.6)	0.33 (<0.01)					
bSi/PIC (z_{bathy})	5.1 (4.8)	4.0 (2.3)	34.6 (29.1)	0.40 (<0.01)					0.9[e]

Table 5.3 (Continued) Comparison of metrics of the BCP.

BCP metrics	SLAMS-2.0 results			Observations	Published empirical models				Other
	OSIM	OIND	DEF		Guidi	Henson	Marsay	Dunne	
ξ (50 m)	2.3 [1.0]	2.5 [1.0]	2.1 [1.0]	1.6 [0.99]					3.6–4.0[g]
ξ (100 m)	2.1 [0.99]	2.1 [1.0]	2.4 [0.83]	1.8 [1.0]					
ξ (500 m)	2.7 [1.0]	2.1 [1.0]	1.7 [0.98]	–					
BATS/OFP (31.6°N, 64.2°W)									
b	0.84 [0.99]	0.78 [0.99]	0.78 [0.99]	0.86 [1.0]	0.68	0.49/0.46	1.54		0.97[b]
z^*	300 [0.93]	335 [0.92]	327 [0.89]	591 [0.99]			104		158[l]
PE_{eff}	0.04 (<0.01)	0.06 (<0.01)	0.06 (<0.01)	0.09 (<0.01)		0.04		0.06	0.06, 0.09[j,l]
T_{eff} ($z_{eu} \rightarrow z_{meso}$)	0.10 (0.01)	0.11 (0.01)	0.12 (0.01)	0.09 (<0.01)					0.56[l]
T_{eff} ($z_{meso} \rightarrow z_{bathy}$)	0.06 (<0.01)	0.06 (<0.01)	0.08 (<0.01)	0.07 (<0.01)		0.26			0.04[c]
POC/PIC (z_{eu})	2.6 (1.7)	8.4 (3.7)	1.3 (0.6)	22.3 (0.45)					
POC/PIC (z_{meso})	2.0 (1.1)	1.3 (0.5)	1.6 (1.1)	0.91 (<0.01)					
POC/PIC (z_{bathy})	0.99 (0.70)	0.78 (0.4)	1.3 (0.8)	0.68 (<0.01)					1.0[e]
bSi/PIC (z_{eu})	0.30 (0.18)	0.41 (0.14)	0.63 (0.26)	0.82 (0.03)					
bSi/PIC (z_{meso})	1.6 (0.9)	0.43 (0.17)	4.34 (3.10)	0.24 (<0.01)					
bSi/PIC (z_{bathy})	1.2 (0.9)	0.37 (0.16)	4.61 (2.89)	0.28 (<0.01)					0.40[e]
ξ (50 m)	2.7 [1.0]	2.4 [1.0]	2.3 [1.0]	1.5 [1.0]					
ξ (100 m)	2.5 [1.0]	2.7 [1.0]	2.5 [1.0]	1.6 [1.0]					
ξ (500 m)	2.1 [0.99]	–	1.2 [0.91]	–					

[a] Charette et al. (1999). [b] Conte et al. (2001). [c] Francois et al. (2002). [d] Lampitt et al. (2008). [e] Honjo et al. (2008). [f] Buesseler and Boyd (2009).
[g] Kostadinov et al. (2009). [h] Lampitt et al. (2010). [i] Lam et al. (2011). [j] Lomas et al. (2013); Owens et al. (2013). [k] Timothy et al. (2013).
[l] Mouw et al. (2016b). [m] Bauerfeind et al. (2009). [n] Le Moigne et al. (2015).

Some general inferences can be made from the above Table: (i) that SLAMS-2.0's optimised simulation offers an improved performance regarding b, z^* and PE_{eff}, (ii) that the mineral ratios (POC/PIC and bSi/PIC) are the metrics more notoriously away from observed values, (iii) that there is a lot of dispersion in the data (metrics from compiled observations, metrics from published empirical models and metrics from published results), and (iv) that there are not enough published values of ξ to assess how well SLAMS-2.0 reproduces that metric. Overall, the most important fact is that model performance has improved not only in terms of a better fit to depth profile flux observations (as shown in Fig. 5.6), but also in that the BCP metrics that can be calculated from those flux profiles are also closer to observed values.

To help with the analysis of Table 5.3, I have plotted in Fig. 5.9 the BCP metrics calculated from SLAMS-2.0's output (grey bars and black lines) as well as the metrics calculated from the fluxes compiled in Chapter 4 (black crosses) and obtained from published results, which are a mix of empirical fits and on-site fieldwork observations (coloured stars and reddish lines). Attempting a case by case analysis is too exhaustive, so I will comment on general trends that emerge from Fig. 5.9. Most importantly, it can be seen that there is no common pattern between SLAMS-2.0's outputs and observations (i.e., data and published values). Instead, it can be said that modelled outputs are within the ranges drawn by the observations. The large dispersion shown by observations is a limiting factor to constrain the output of SLAMS-2.0 as there is no such a thing as a *correct* value, but a range of values that must be obtained through the exhaustive collection of published values.

I have identified four reasons to explain the dispersion in the observations shown in Table 5.3 and Fig. 5.9.

- Different authors use different statistical methods to compute b, i.e., different reference depths, different collection depths and different integration times. Fig. 5.10 illustrates the problem of the statistical treatment of b in different data availability scenarios. b is a metric particularly sensitive to the amount of data in the depth profile from which it is computed. I show that for the same *in situ* POC export flux, the b calculated in scenarios 1, 2 or 3 will differ depending on how many POC flux data points the researcher has access to.
- Different observational technologies (e.g., sediment traps vs radionuclides) use different time frames of material collection, up to a year for a sediment trap and around 20 days for Th radioactive pairs.
- Different authors collect observations at different times of the year (e.g., pre-bloom, bloom or post-bloom period), which can be more or less representative of the annual mean.

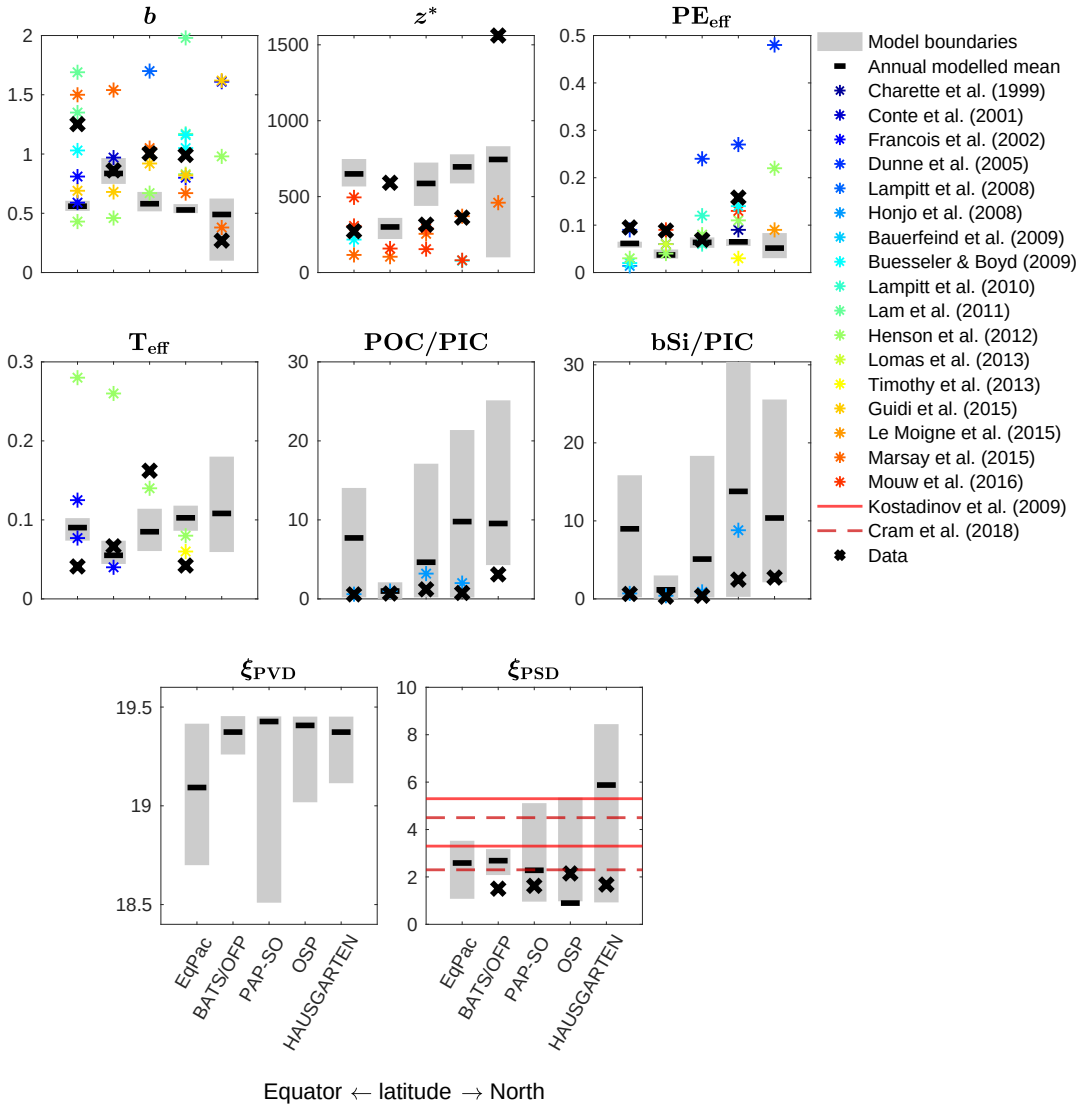
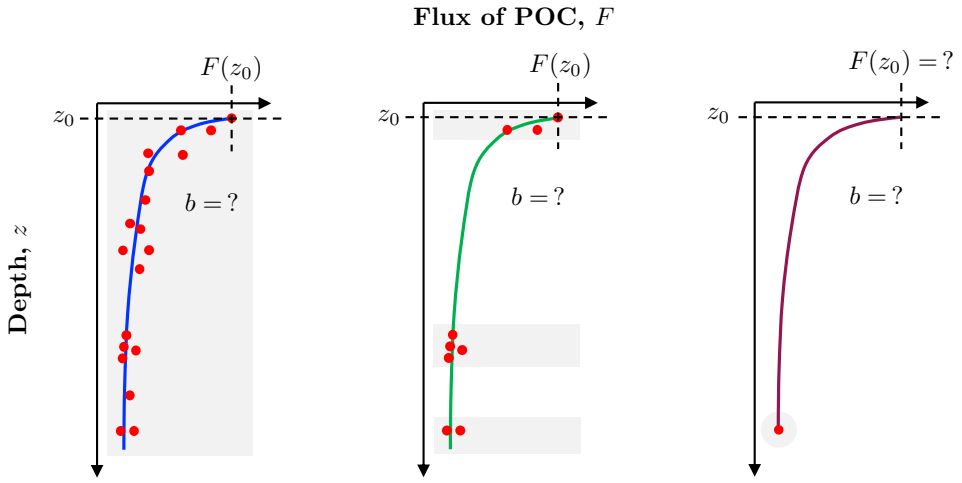


Figure 5.9. Comparison of SLAMS-2.0's modelled metrics of the BCP at five ocean time-series sites against local observations from my compiled data set (Data), empirical modelling studies and on-site fieldwork studies (coloured stars and reddish lines). The grey bars show the amplitude in the range of monthly estimates generated by SLAMS-2.0 and the black line shows the annual mean. The BCP metrics are: the rate of attenuation of POC flux with depth (b, after Martin et al., 1987), the depth where 63% of the POC flux has already been remineralised (remineralisation length scale, z^*, after Lutz et al., 2002), the fraction of the NPP that is exported out of the ocean's surface layer (particle export efficiency, PE_{eff}, after Buesseler and Boyd, 2009), the fraction of exported POC flux that reaches the 2000 m depth horizon (transfer efficiency to the deep ocean, T_{eff} 100 → 2000, after Buesseler and Boyd, 2009), the molar ratio of the POC to PIC flux in the deep ocean (~3000 m) (POC/PIC, after Honjo et al., 2008), the molar ratio of the bSi to PIC flux in the deep ocean (bSi/PIC, after Honjo et al., 2008), the slope of the particle velocity distribution at 50 m (ξ_{PVD}) and the slope of the particle size distribution at 50 m (ξ_{PSD}). The locations are: EqPac, in the upwelling equatorial Pacific (0°N, 140°W); HAUSGARTEN, in the Arctic Atlantic (79.0°N, 4.3°E); OSP (Ocean Station Papa), in the subarctic eastern Pacific (50°N, 145°W); PAP-SO (Porcupine Abyssal Plain time-Series Observatory), in the blooming NE Atlantic (49°N, 16.5°W); BATS/OFP (Bermuda Atlantic Time-Series/Oceanic Flux Program) joint site, in the subtropical NW Atlantic (31.6°N, 64.2°W).

- If published BCP metrics are calculated from a one-time flux measurement, that value will show strong dependency on the moment of the year the flux value was collected, and thus the sort of productive mode the ocean was in. Therefore, their degree of representativeness of the actual annual mean value can be questioned (e.g., Charette et al., 1999; Lampitt et al., 2008; Buesseler and Boyd, 2009; Marsay et al., 2015).



Mode 1: complete profile
e.g., Martin et al., (1987)

Mode 2: key depths profile
e.g., Marsay et al., (2015)

Mode 3: pivot profile
e.g., Henson et al., (2012)

(*i*)	$b = \frac{\log(F(z)) - \log(F(z_0))}{\log(z) - \log(z_0)}$	$F(z_0) = \frac{F(z)}{\left(\frac{z}{z_0}\right)^{-0.858}}$
(*ii*)		$b = \frac{\log(F(z)) - \log(F(z_0))}{\log(z) - \log(z_0)}$

Figure 5.10. Modes in which Martin's b is calculated in the literature according to the amount of data points in the depth profile and their distribution in key depths. In modes 1 and 2, b can be directly calculated from the Martin curve by isolating it, which has the solution shown. The value of b is sensitive to the amount of F data points, which differs in modes 1 and 2. In mode 3, b needs to be calculated in two steps. First of all, an export flux value $F(z_0)$ needs to be extrapolated from the only deep flux value $F(z)$ available by assuming $b = -0.858$ (canonical b). Once $F(z_0)$ is known, b can be recalculated.

5.5 Conclusions

It can be concluded that SLAMS-2.0's optimised model response is better than that of the default model, especially for midwater depths and the bSi and PIC fluxes. The good performance of the optimisation is also supported by a comparison of modelled BCP metrics against observations, where SLAMS-2.0 can reproduce ballpark figures of b, z^*, $\mathrm{PE}_{\mathrm{eff}}$ and

T_{eff}. Due to (i) the non-linearity of SLAMS-2.0 response, (ii) the possibility that SLAMS-2.0 may be lacking or inaccurately representing some significant processes, and (iii) the few and noisy data available to constrain it, DFO-LS struggles to navigate the topography accurately and, thus, to recreate surface POC peak fluxes in the observations data set.

I had more POC, PIC and bSi flux observations available for DFO-LS to assimilate, and with more robust annual averages, the performance of the optimiser would have probably improved, drawing a less noisy landscape. Optimising a model with particulate flux data is a challenge given the limited availability of data and the different methods of obtention and statistical processing, with no unique way of doing it. It is also worth emphasising the uncertainty introduced by the choice of of the NPP data set used to force the model, with different satellite and *in situ* estimates having an impact on the magnitude of the POC flux profiles. I cannot help but stress how important the collaboration between modellers and observationalists is to elaborate particle flux and NPP data sets that can be meaningfully used in models of marine particles by identifying the data points that are less biased and help to differentiate amongst various optimal solutions in terms of supporting parameters and their physiological meaning. The relevance of such collaborations cannot be underestimated, as it currently is.

The optimisation exercise presented in this thesis could have benefited from a misfit function that considers the uncertainty in the data (with various sources of error) and the parameters (with values that can vary across a wide range without a significant effect on the model-data misfit). Those uncertainties draw regions of variation around point solutions (local minima) and produce parameter sets that, although different, do not produce significantly different model outputs. Looking at *regions of solution* instead of *point solutions* could be the next step to improve the performance of DFO-LS. A further step taken towards improving the optimisation of SLAMS-2.0 could be using a Bayesian approach (e.g., Bisson et al., 2018) instead of a data assimilation approach, as DFO-LS is. Bayesian approaches are probabilistic and can efficiently infer solutions from noisy observations without using expensive integration methods. Optimised model outputs are viewed as realisations of a stochastic process, with a distribution that has a mean and a region of variation.

Part III

RESULTS AND DISCUSSION

6

A global application of SLAMS-2.0

Contents

6.1 Introduction

In Chapter 5, I have calibrated SLAMS-2.0 with the best available set of POC, PIC and bSi flux observations at five well-sampled ocean locations (EqPac, HAUSGARTEN, OSP, PAP-SO and BATS/OFP). I have demonstrated the complexities of calibrating SLAMS-2.0 with POC, PIC and bSi flux data that are sparse in time and depth, as well as the challenge of dealing with an optimiser that gets trapped in local minima. In this chapter, the locally calibrated SLAMS-2.0 is run on a global grid with local biogeochemical forcing. I then apply the model to investigate the controls that affect the transfer of POC flux to the ocean interior and their impact on the global patterns of various BCP metrics

(b, PE_{eff}, T_{eff}, z^*, POC/PIC, bSi/PIC) derived from the Lagrangian ensemble of particles. A specific objective of this chapter is to reconcile Henson et al. (2012) and Marsay et al. (2015) school of thoughts, with opposite global patterns for the distribution of Martin's b.

To include the uncertainty offered by the optimiser for different local minima, this chapter offers global scale results for two optimised parameter sets from the simultaneous fit to data (OSIM) experiments: (i) the parameter set with the lowest misfit (OSIM1), i.e., the best parameter set found by the optimiser, and (ii) the parameter set that is most different to the first best (OSIM2), inferred from a root-mean squared analysis (Table 6.1). By running the model globally with two parameter sets different enough, it is possible to attribute the differences in the modelled patterns of the BCP to the underlying parameters that have been tuned to produce them.

Table 6.1. Statistical comparison of the seven best OSIM optimised parameter sets with the first best OSIM parameter set (OSIM1) based on a root-mean squared (RMS) difference normalised.

	OSIM1	OSIM2	OSIM3	OSIM4	OSIM5	OSIM6	OSIM7
OSIM1	0	1.41	1.18	1.34	1.31	0.78	1.07

6.2 Description of the results

6.2.1 Global patterns of modelled efficiencies and biogenic fluxes

Fig. 6.1 shows global maps of four commonly used BCP metrics annually averaged with the OSIM1 (Fig. 6.1a) and (Fig. 6.1b) OSIM2 runs: Martin's b, PE_{eff} and T_{eff} from 100 to 1000 m and T_{eff} from 1000 to 2000 m. Both OSIM1 and OSIM2 runs show the same patterns in the four metrics plotted, with just slight differences in the magnitudes of those.

The global maps of Martin's b (top left panels) show high values of b (i.e., high attenuation rate of POC flux with depth) in the subtropics (0.70–1.0), and lower values in the polar oceans and coastal and equatorial upwelling regions (0.35–0.50). Martin et al. (1987)'s b of 0.86 is reproduced at the VERTEX study site, a group of several locations in the Pacific off the western coast of the U.S. and Mexico (14–35°N, 108–139°W). The simulated global b ranges from 0.21 to 0.94 in OSIM1, with a mean of 0.63, and it ranges from 0.15 to 0.99 in OSIM2, with a mean of 0.64.

The global maps of PE_{eff} (top right panels of Figs. 6.1a and 6.1b) are anticorrelated with the global maps of b; higher PE_{eff}, or fraction of NPP that is exported as POC, are found in regions where b is low. Higher values of PE_{eff} occur in polar systems and upwelling equatorial and coastal regions (~0.07), where 7% of the NPP is exported out of the surface ocean . The oligotrophic subtropical regions have a lower PE_{eff} (~0.03) and, thus, a BCP that injects less POC to the mesopelagic ocean compared to the polar and

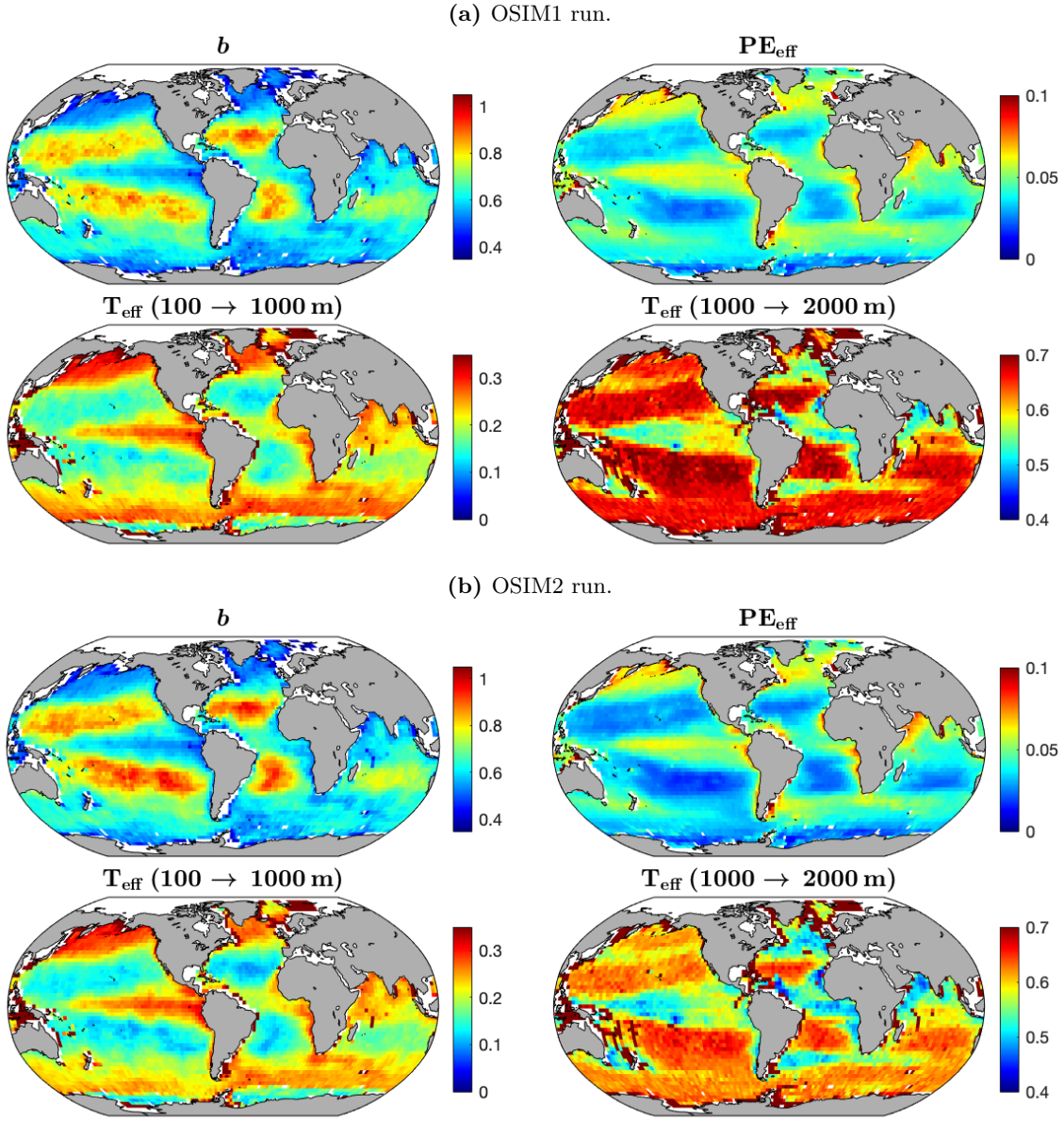


Figure 6.1. Global maps of SLAMS-2.0's modelled annual averages of (top left) Martin's b, (top right) particle export efficiency and (bottom left) transfer efficiency from 100 to 1000 m (surface to mesopelagic), and (bottom right) transfer efficiency from 1000 to 2000 m (mesopelagic to deep ocean) produced with (a) the first best parameter set found by the optimiser (OSIM1 run) and (b) the optimised parameter set that is most different to the first best (OSIM2 run).

upwelling regions. The simulated global PE_{eff} ranges from 1.5% to 10% in OSIM1, with a mean of 4.4%, and it ranges from 1.1% to 12% in OSIM2, with a mean of 4.1%.

The global maps of T_{eff} 100 → 1000 m, or surface to mesopelagic ocean (bottom left panels of Figs. 6.1a and 6.1b), mirror the pattern of PE_{eff}, with high T_{eff}, or transfer of export POC flux, in the polar oceans and coastal and equatorial upwelling regions. The simulated global T_{eff} 100 → 1000 m has a lowest value of 11% in the subtropics in OSIM1, with a mean of 24%, and spanning a wide range between 30–100% in the high latitudes.

In OSIM2, the lowest value found is of 9.2%, with a mean of 23%. If T_{eff} is looked at the transition between the mesopelagic and the deep ocean (1000 → 2000 m), as shown in the bottom right panels of Figs. 6.1a and 6.1b, an inversion can be seen with respect to the pattern displayed by the 100 → 1000 m. T_{eff} 1000 → 2000 m is highest in the subtropics and lowest in the high latitude North Atlantic (but not in the North Pacific nor the Southern Ocean) and coastal and equatorial upwelling regions. It is also worth noting that T_{eff} is typically higher in the range 1000–2000 m (deep ocean) than in the range 100–1000 m (mesopelagic ocean). The simulated global T_{eff} 1000 → 2000 m has a lowest value of 44% in OSIM1 and 43% in OSIM2, with a mean of 66% in OSIM1 and 63% in OSIM2.

Those four BCP metrics described are supported by the flux landscapes depicted in Fig. 6.2, with the polar oceans and equatorial and costal upwelling regions exhibiting (top left) high POC export flux (~30–80 mg C m^{-2} d^{-1}), (top right) high POC sequestration flux (~7.5–20 mg C m^{-2} d^{-1}) and (bottom right) high opal export flux (~30–80 mg bSi m^{-2} d^{-1}). The globally integrated total POC export is 2.7 Gt C yr^{-1} in OSIM1 and 2.6 Gt C yr^{-1} in OSIM2. The (bottom left) PIC export flux pattern is more monotonous than those of POC and opal, where it is not possible to ascertain a latitudinal PIC export pattern. It must be noted the big difference in the pattern of PIC export flux between OSIM1 and OSIM2 runs. The integrated PIC export flux is the same for both the Pacific and the Atlantic basins, 0.13 Gt C yr^{-1} (value for OSIM1). Per contra, the Pacific ocean presents a smaller delivery of export bSi flux to the deeper ocean (0.26 Gt Si yr^{-1} in OSIM1) than the Atlantic (0.30 Gt Si yr^{-1} in OSIM1).

POC export flux has a smaller range of variation in OSIM1 (2.0–335 mg C m^{-2} d^{-1}) than in OSIM2 (1.5–369 mg C m^{-2} d^{-1}), with a correspondingly larger average value in OSIM1 (23.0 mg C m^{-2} d^{-1}) than in OSIM2 (22.3 mg C m^{-2} d^{-1}). POC sequestration flux has a smaller average, coinciding in 7.0 mg C m^{-2} d^{-1} for both OSIM1 and OSIM2. PIC export flux has a smaller range of variation in OSIM1 (10.8–66.2 mg $CaCO_3$ m^{-2} d^{-1}) than in OSIM2 (11.8–83.9 mg $CaCO_3$ m^{-2} d^{-1}), with a correspondingly smaller average value in OSIM1 (33.6 mg $CaCO_3$ m^{-2} d^{-1}) than in OSIM2 (49.1 mg $CaCO_3$ m^{-2} d^{-1}). Opal export flux, per contra, has a larger range of variation in OSIM1 (1.8–195 mg bSi m^{-2} d^{-1}) than in OSIM2 (0.98–157 mg bSi m^{-2} d^{-1}), with a correspondingly higher average value in OSIM1 (22.5 mg bSi m^{-2} d^{-1}) than in OSIM2 (15.7 mg bSi m^{-2} d^{-1}).

To further investigate the relationship between POC flux and the two major ballast materials ($CaCO_3$ and opal), Fig. 6.3 displays the relative contribution of the ballast minerals in two sets of molar ratios: (left panels) organic carbon vs inorganic carbon in $CaCO_3$ (POC/PIC) and (right panels) Si in biogenic silica vs inorganic carbon in $CaCO_3$ (bSi/PIC). These are mole weight ratios obtained from monthly metric values annually averaged. The distribution of POC/PIC is generally similar to POC export

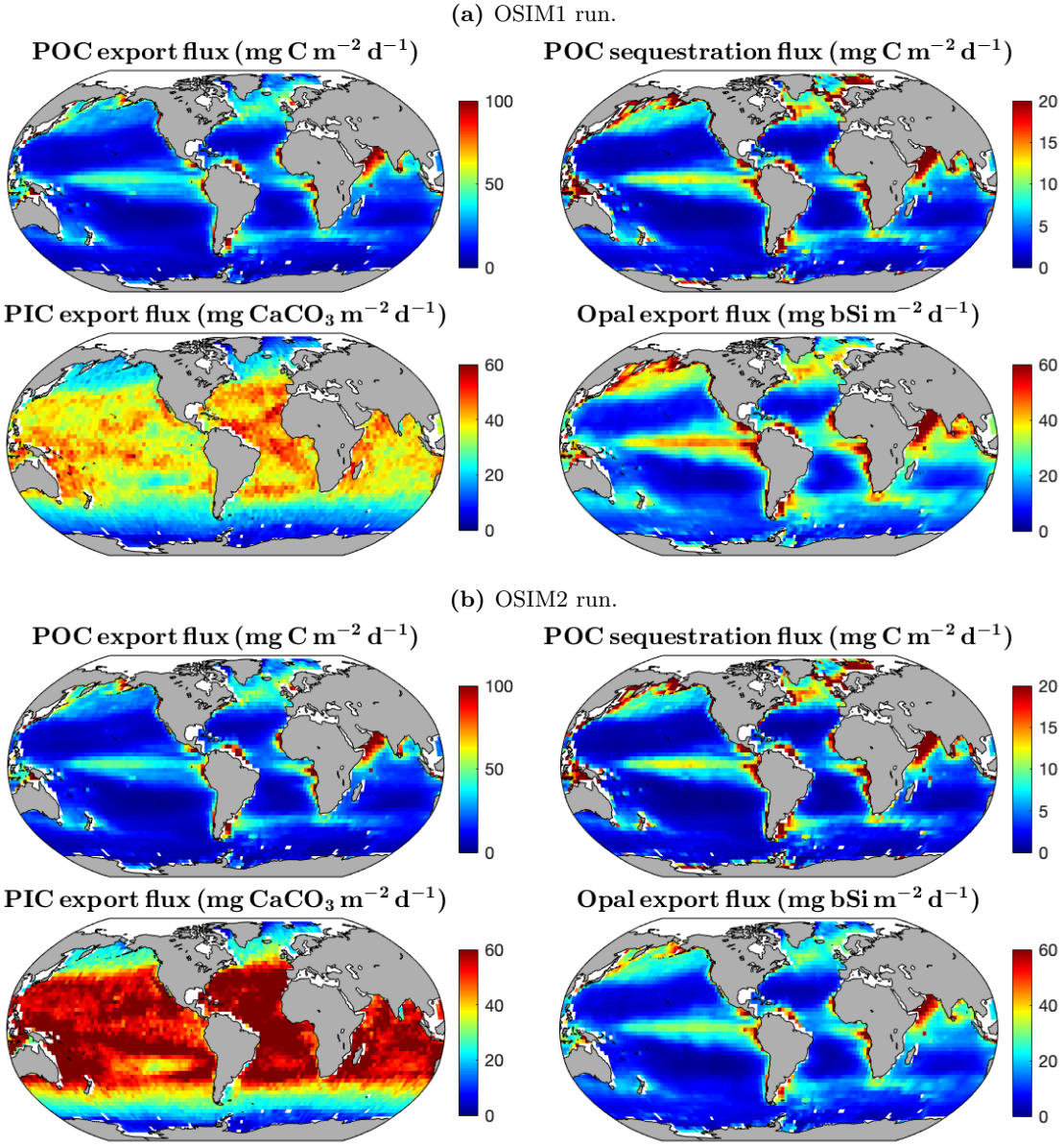


Figure 6.2. Global maps of SLAMS-2.0's modelled annual averages of (top left) POC flux at export depth (100 m), (top right) POC flux at sequestration depth (1000 m), (bottom left) PIC flux at export depth and (bottom right) opal flux at export depth produced with (a) the first best parameter set found by the optimiser (OSIM1 run) and (b) the optimised parameter set that is most different to the first best (OSIM2 run).

flux, with higher values in the northern polar oceans, equatorial pacific upwelling region and Atlantic coastal upwelling regions. POC/PIC is a metric that indicates the amount of $CaCO_3$ required in particles to deliver POC to the ocean interior. Thus, in regions with high POC/PIC, there is relatively low $CaCO_3$ available to ballast particles, and it is opal the mineral that assumes the ballasting role. Thus, in regions with high POC/PIC, there is high bSi/PIC (Fig. 6.3, right panels) and, correspondingly, high opal export (Fig. 6.2, bottom left panels).

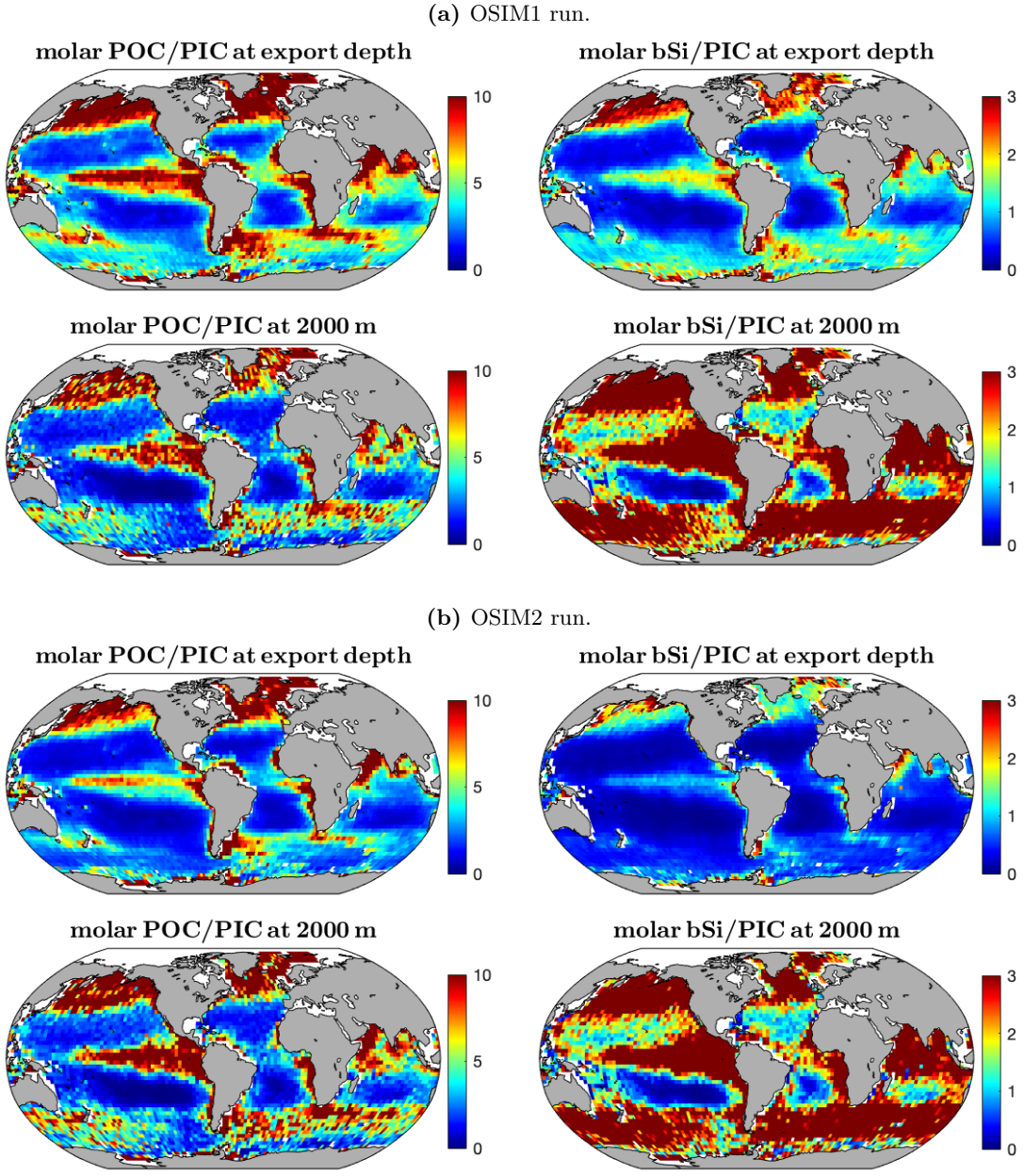


Figure 6.3. Global maps of SLAMS-2.0's modelled annual averages of molar (top left) POC/PIC ratio at export depth (100 m), (bottom left) POC/PIC ratio at 2000 m, (top right) bSi/PIC at export depth and (bottom right) bSi/PIC at 2000 m produced with (a) the first best parameter set found by the optimiser (OSIM1 run) and (b) the optimised parameter set that is most different to the first best (OSIM2 run).

From Figs. 6.1–6.3 it can be concluded that both optimised solutions (OSIM1 and OSIM2) offer the same global patterns for the transfer of fluxes of biogenic material to the ocean interior, just with slight differences in the magnitudes of those (Table 6.2, upper row). Since the similar dynamics of OSIM1 and OSIM2 are supported by slightly different parameter sets (Table 6.2, lower row), it can be inferred that differences in the parameters

do not alter the model dynamics significantly. OSIM1 shows a lower b, or smaller rate of change of POC flux with depth, which translates into a higher fraction of NPP being exported out of the euphotic zone (higher PE_{eff}) and a higher transfer efficiency of the exported material to deeper depths (higher T_{eff}). Those relationships between b, PE_{eff} and T_{eff} are supported by a higher POC export flux (F_{POC}) that is positively correlated with a higher opal export flux (F_{bSi}) but negatively correlated with PIC export flux (F_{PIC}). The parameter values (Fig. 5.8 in Chapter 5) reveal that those flux relationships are based on (i) higher silicified diatoms (higher $\theta^{diat}_{Si:C}$) and (ii) smaller opal dissolution timescale (lower κ_{opal}), (iii) a coccolithophore community of higher calcified cells (higher $\theta^{cocco}_{calc:C}$), (iv) a zooplankton community with higher biomass abundance in relation to POC food biomass (lower $\theta_{POC:Z}$), (v) specially at the night DVM depth (lower $\theta^{DVM\,night}_{POC:Z}$), (vi) and which fragments a higher amount of POC due to sloppy feeding (higher f_{sloppy}).

Table 6.2. Comparison of modelled average metrics of the BCP (b, PE_{eff} and T_{eff}) and biogenic fluxes at 100 m (F_{POC}, F_{PIC} and F_{bSi}, in $mg\,m^{-2}\,d^{-1}$) obtained from the two best SLAMS-2.0's optimised runs (OSIM1 and OSIM2) and the corresponding optimised parameter sets associated to those two runs. Highlighted in red is the variable or parameter displaying the highest value in one of the two optimised runs, and yellow for the other one. Parameter symbols are: $\theta^{diat}_{Si:C}$, diatom cell Si:C ratio $(mol\,Si\,(mol\,C)^{-1})$; $\theta^{cocco}_{calc:C}$, coccolithophore cell PIC:POC ratio $(mol\,PIC\,(mol\,POC)^{-1})$; $\theta_{POC:Z}$, standard POC:zooplankton biomass; $\theta^{DVM\,night}_{POC:Z}$, POC:zooplankton biomass at the night DVM depth; f_{sloppy}, sloppy feeding loss fraction; $f^{ing,\,surf}_{OM}$, fraction of OM in an appetitive surface particle; κ_{opal}, opal dissolution timescale (d^{-1}).

BCP metrics and fluxes							
	b	PE_{eff}	T_{eff} (100 → 1000 m)	T_{eff} (1000 → 2000 m)	F_{POC}	F_{PIC}	F_{bSi}
OSIM1	0.63	4.4%	24%	66%	23.0	33.6	22.5
OSIM2	0.64	4.1%	23%	63%	22.3	49.1	15.2
Parameters							
	$\theta^{diat}_{Si:C}$	$\theta^{cocco}_{calc:C}$	$\theta_{POC:Z}$	$\theta^{DVM\,night}_{POC:Z}$	f_{sloppy}	$f^{ing,\,surf}_{OM}$	κ_{opal}
OSIM1	0.06	0.52	7.8	0.78	0.40	0.10	0.03
OSIM2	0.04	0.50	10	2.0	0.27	0.10	0.05

Collectively, those results present a global spatial pattern of particle flux attenuation through the water column that is coherent: the subtropical, oligotrophic oceans deliver less surface POC to the deep ocean and present higher attenuation rates and smaller transfer efficiencies, whereas the high latitudes and upwelling regions deliver the largest amount of surface POC and present smaller attenuation rates and larger transfer efficiencies. The modelled mineral ballast components accompanying the flux of POC to the deeper ocean show that opal, and not calcium carbonate, presents a synergistic flux pattern to that of POC. Whereas the delivery of PIC to the deeper ocean is similar for both the Atlantic and the Pacific basins, the Atlantic presents a higher delivery of bSi than the Pacific.

6.2.2 Global patterns of ecosystem structure

The patterns of the fluxes of the ballast minerals presented in Fig. 6.2 fundamentally arise from the global distribution of relative biomass stock of diatoms (the primary source of bSi) and coccolithophores (the primary source of PIC) prescribed in SLAMS-2.0. In the model, PFTs are seeded in the water column following an *a priori* distribution of PFT probabilities calculated according to phytoplankton rules of growth that are randomly sampled using the Monte Carlo method. The seeding of those PFTs produces the global pattern of the relative biomass (in %) distribution of diatoms, dinoflagellates, coccolithophores and picophytoplankton shown in Fig. 6.4. That pattern of relative biomass gives rise to the global pattern of PFT biomass stock (in $\mathrm{mg\,C\,m^{-2}}$) shown in Fig. G.1. According to the phytoplankton growth rules used for seeding PFTs, diatoms, which are better at taking up nutrients in high-nutrient, low-light (or chlorophyll) environments, dominate in the seasonal high latitudes, the upwelling and convective regions of the equatorial Pacific and off the west coast of continents. Phytoplankton rules of growth dictate that picophytoplankton have significant abundances in the subtropics and tropic. A combination of coccolithophores and dinoflagellates make up the remaining background concentration.

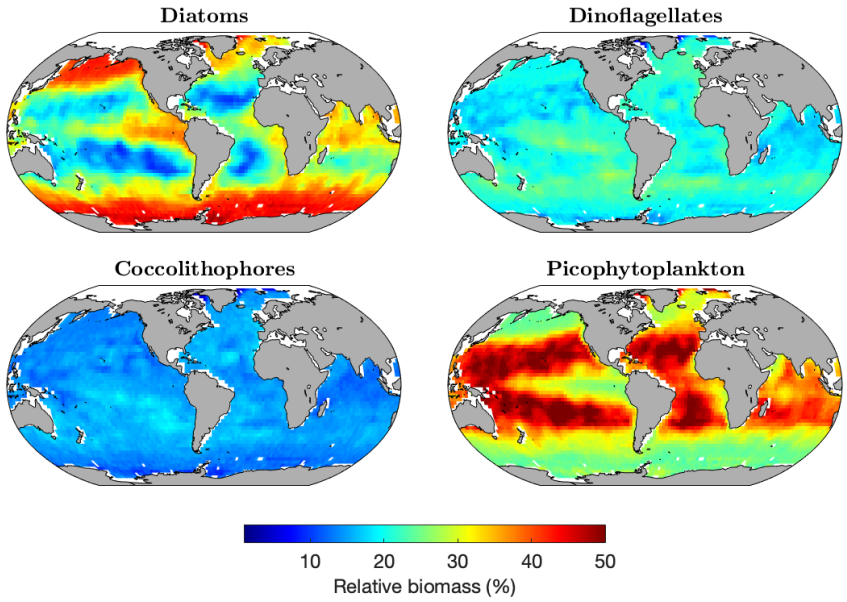


Figure 6.4. SLAMS-2.0's prescribed global distribution of the relative biomass stock of the four modelled phytoplankton functional types (PFTs): (top left) diatoms, (top right) dinoflagellates, (bottom left) coccolithophores and (bottom right) picophytoplankton. Relative biomass stock (in percentage) is defined as the fractional contribution of each PFT biomass to the total phytoplankton biomass, and it is an equivalent metric to the probability of seeding PFTs in SLAMS-2.0.

Particle size and particle sinking velocity are two particle attributes often used to link POC transfer efficiency patterns to the BCP mechanisms associated with POC attenuation. Particle size and sinking velocity emerge from the chemical composition

of the particles and their fluffiness, both of which are controlled by the phytoplankton community composition. In Figs. 6.5–6.6, I have mapped over depth the evolution of the average particle size and sinking velocity for two main categories of particles (large vs small), and in Appendix G (Figs. G.2–G.9) I provide the rest of the particle attributes simulated by SLAMS-2.0 (density, porosity, radius of the primary particles, stickiness, POC content, PIC content, bSi content and clay content). Based on the mapped particle attribute trends, the size of an average small particle increases with depth (Fig. 6.5a). However, there is an increase in size for a large particle up to the depth of 500 m (1000 m in the Equatorial Pacific) and below that depth the particle decreases in size (Fig. 6.5b). The sinking velocity of both an average small and large particle (Fig. 6.6) increases with depth. Those changes in size and sinking velocity are supported by an overall increase in average density for both average small and large particles (Fig. G.2), an overall decrease in stickiness (Fig. G.5), an overall decrease in POC content (Fig. G.6) and an overall increase in PIC content (Fig. G.7) and opal content (Fig. G.8). Thus, larger densities at deeper depths arise from relatively larger contributions of PIC and opal to the total weight of the particle. The trend for porosity (Fig. G.3) is not so clear for small particles, as they are already quite compact, but for large particles an increase is seen, especially off the west coast of North Africa, a region known for the supply of aeolian dust to the surface ocean. That signal off the west coast of North Africa is also seen for the other attributes in more or less degree, especially at deeper depths, $>$ 500 m. Larger particles have an average porosity (given the same depth) higher than small particles, a feature that emerges due to the fractal nature of aggregates (Alldredge and Gotschalk, 1988).

In order to infer to which particle attributes size and sinking velocity are more sensitive to, I have correlated those two attributes to the rest of particle attributes simulated by SLAMS-2.0 for the 500 m depth horizon for an average small and large particle (Figs. G.10–G.11). For large particles, size is mostly correlated with particle volume (R^2 = 0.45) and POC content (R^2 = 0.39), and sinking velocity is mostly correlated with density (R^2 = 0.53) and bSi content (R^2 = 0.23). For small particles, the correlations are stronger and size is mostly correlated with clay content (R^2 = 1.0), density (R^2 = 0.98) and sinking velocity (R^2 = 0.94), whereas sinking velocity is mostly correlated with clay content (R^2 = 0.94), size (R^2 = 0.94) and density (R^2 = 0.90).

The seeded primary production in SLAMS-2.0 is processed by a variety of modelled processes, such as aggregation and zooplankton grazing, which result in the production of fast-sinking and large particles (as shown in Figs. 6.5–6.6) regardless of whether the ecosystems is dominated by large (diatoms and dinoflagellates) or small PFTs (coccolithophores and picophytoplankton). The global distribution of the intensity of particle-particle interactions (collision and coagulation) and particle-zooplankton

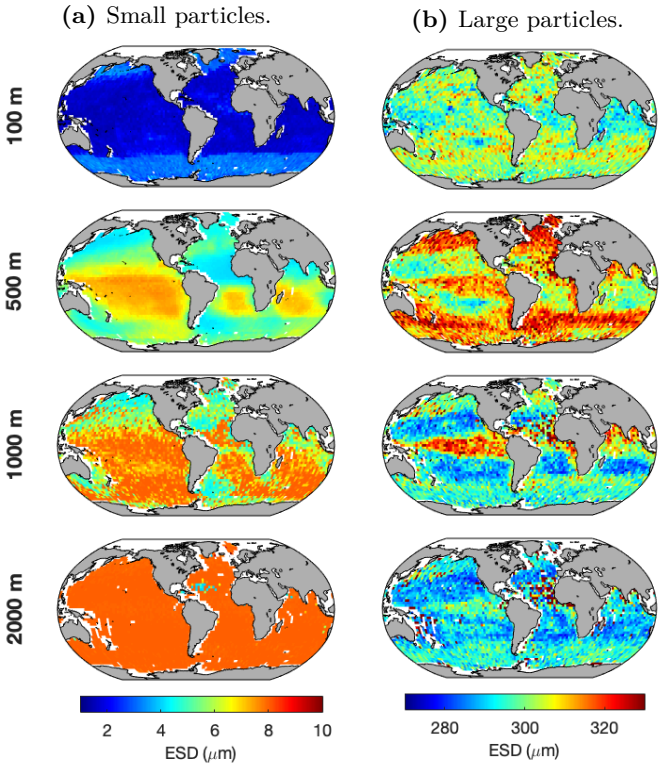


Figure 6.5. Size (ESD) of an average modelled particle at four main ocean depths according to two main size categories, (a) small (≤ 150 μm) and (b) large (> 150 μm).

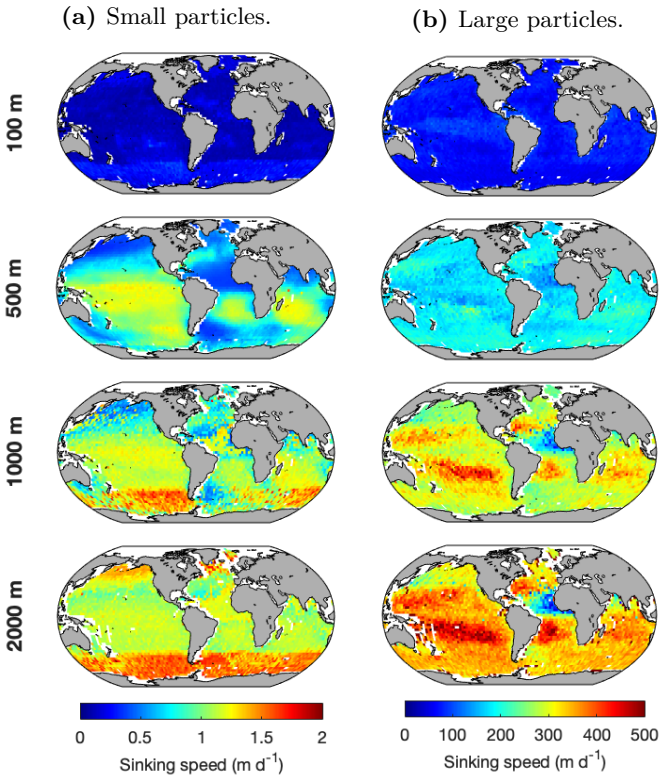


Figure 6.6. Sinking velocity of an average average modelled particle at four main ocean depths according to two main size categories, (a) small (≤ 150 μm) and (b) large (> 150 μm).

interactions is plotted in Fig. 6.7 for four depth horizons (100, 500, 1000 and 2000 m). The spatial pattern shows that particle collision rates (Fig. 6.7a) are higher in northern high latitudes, the upwelling equatorial Pacific, the upwelling regions off the west coast of continents and the subtropical front bordering the southernmost continental perimeter of the southern hemisphere. Particle collisions end up in successful coagulation in regions where collision rates are high, as expected (Fig. 6.7b). A prominent feature in the pattern of successful coagulation is the high coagulation in the equatorial Atlantic offshore the west coast of North Africa, following the pattern of Saharan aeolian dust deposition, the major desert dust source region in the world (Jickells and Moore, 2015). The spatial pattern of particle-zooplankton encounter rates (Fig. 6.7c) is similar to the pattern drawn by particle-particle interactions, with the encounters being less abundant in the subtropics and more abundant in the high-latitude systems and the equatorial Pacific.

To gain a deeper insight into the activity of zooplankton, I have plotted in Fig. G.12 the global distribution of the intensity of zooplankton grazing with its different facets of sloppy feeding, ingestion, egestion (i.e., production of faecal pellets) and respiration. The pattern reveals that modelled grazing pressure is mostly concentrated in northern high latitudes, the upwelling equatorial Pacific and coastal regions, resembling the distribution of the higher biomass stocks of phytoplankton (Fig. G.1), and where the export of NPP to the deep ocean in form of POC is more efficient (higher PE_{eff}, Fig. 6.1). It can be noted that higher respiration flows (Fig. G.12d) do not overlap with higher attenuation rates of POC flux (e.g., Fig. 6.1a) since regions where zooplankton heterotrophic metabolism is high are also regions with high zooplankton egestion fluxes. Thus, the function of repackaging that zooplankton grazing does on ingested particles, mixing lighter OM with ballast minerals, generates ballast fluxes that transfer POC efficiently to the deeper ocean, compensating for the loss of POC due to respiration.

6.3 Comparison of SLAMS-2.0's modelled fluxes to other modelling studies

Martin's b is the most widely used of the BCP metrics and the most litigious one. Due to the nature of its mathematical fit (a power-law instead of, for instance, a fraction, like T_{eff}), b is highly sensitive to the availability of sufficiently depth-resolved data below the reference depth, as well as the choice of that reference depth (Buesseler et al., 2020b), which makes it highly variable amongst studies. The pattern of b predicted by SLAMS-2.0 can be compared against that of Henson et al. (2012) and Marsay et al. (2015), the most cited studies of the global distribution of b (Fig. 6.8). SLAMS-2.0 predicts a b that follows the same qualitative pattern as that of Marsay et al. (2015), high in the

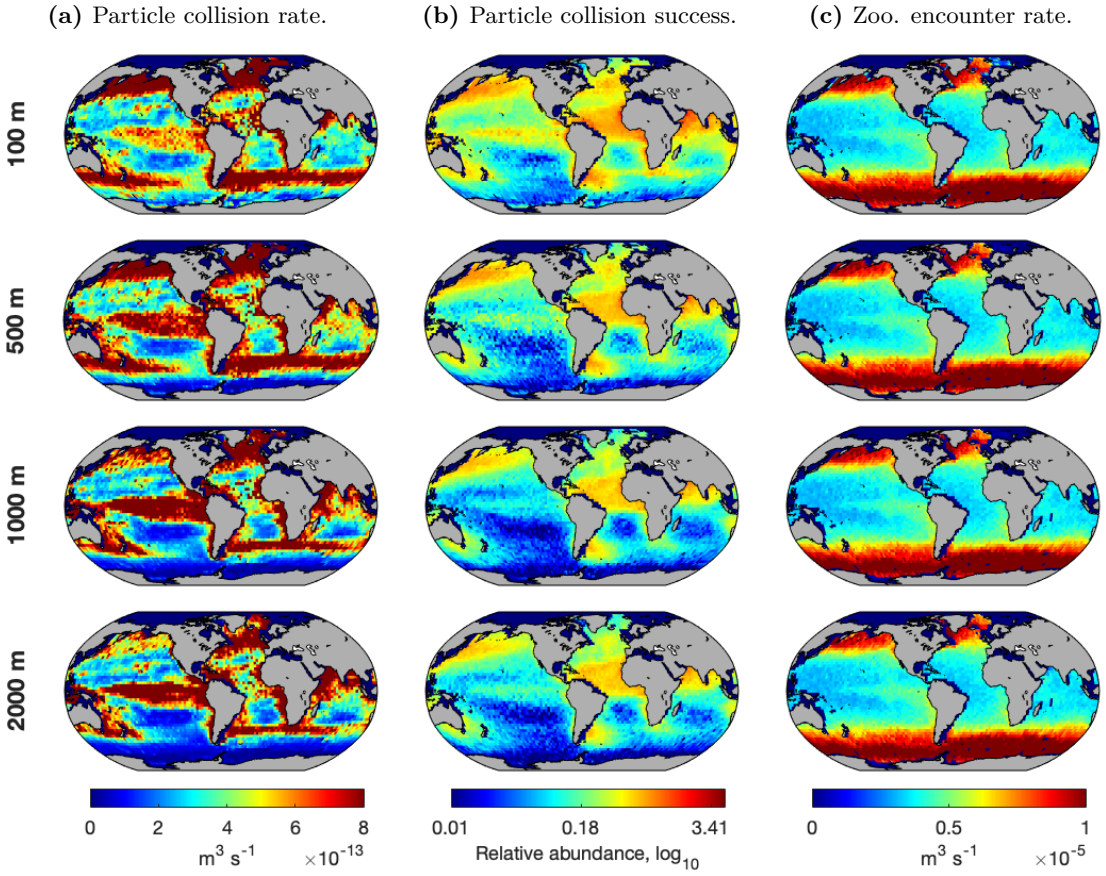


Figure 6.7. SLAMS-2.0's modelled global distributions of (a) particle-particle collision rates, (b) particle coagulation success and (c) particle-zooplankton encounter rates by depth horizon 100, 500, 1000 and 2000 m.

subtropics and low in the colder and upwelling equatorial and coastal regions. That pattern is opposite to that presented in Henson et al. (2012), where b is lower in the subtropics and the equatorial Pacific. Nevertheless, SLAMS-2.0's b range of variation (0.21–0.94) is more similar to that of Henson et al. (2012) (0.22–1.2) than that of Marsay et al. (2015) (0.19–1.8), and so SLAMS-2.0 predicts substantially lower attenuation rates in the subtropics than Marsay et al. (2015).

The pattern of PE_{eff} simulated by SLAMS-2.0 can also be compared with previous global studies (Fig. 6.9a)[1]. Compared to SLAMS-2.0's estimates, the studies of Henson et al. (2012) and DeVries and Weber (2017) predict higher maximum PE_{eff} and display higher PE_{eff} values in all regions. DeVries and Weber (2017) use a satellite-driven ecosystem model to estimate fluxes of POC and predict a PE_{eff} higher than that of Henson et al. (2012), and with more zonal variability towards the equator, displaying maximum values in the poles and modestly elevated along the equatorial upwelling region,

[1]SLAMS-2.0's panel for PE_{eff} in Fig. 6.9a is the same as that displayed in Fig. 6.2a (top right panel); notice how the colour bar has been tuned to macth that of the studies of Henson et al. (2012) and DeVries and Weber (2017).

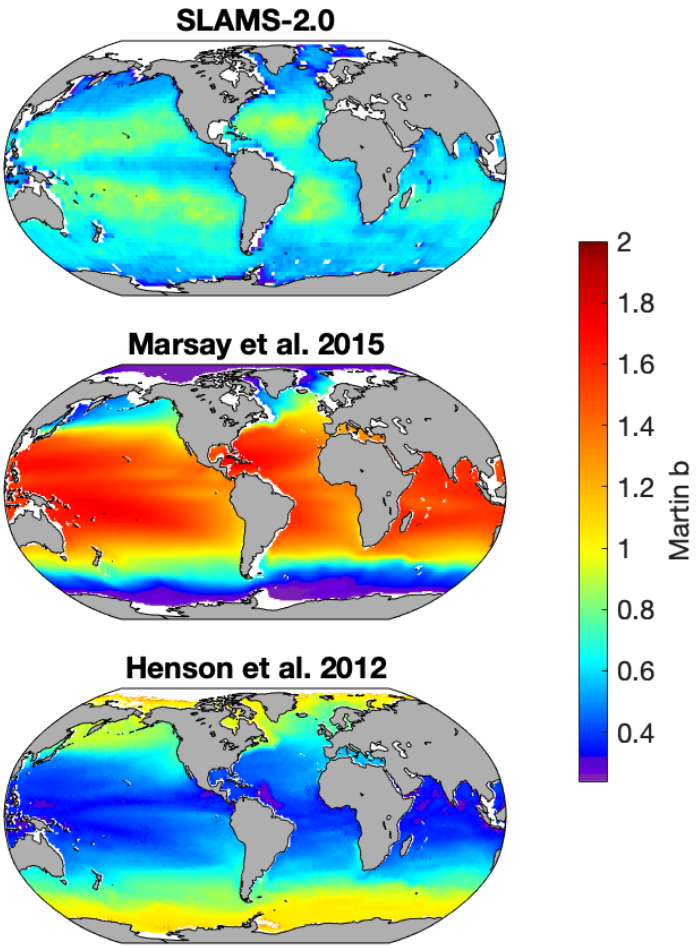


Figure 6.8. Global maps of Martin's b for SLAMS-2.0 compared to the empirical models of Marsay et al. (2015) and Henson et al. (2012). Marsay et al. (2015) model is based on a regression of upper ocean (0–500 m) temperature against b whereas Henson et al. (2012) is based on an application of algorithms to satellite data of SST and NPP based on a compilation of Th-derived POC export flux and sediment trap POC flux at 2000 m. Notice that the colour bar has been customised to reproduce that of Marsay et al. (2015)'s Fig. 3.

in agreement with the studies of Siegel et al. (2014) and Dunne et al. (2005) (not shown, see Fig. 3 in DeVries and Weber (2017)).

The pattern of POC flux at export depth (100 m) can also be compared with that of the studies of Henson et al. (2012) and DeVries and Weber (2017) (Fig. 6.9b). Modelled POC export flux presents lower values than those in the literature but has similar spatial patterns, high in the northern high-latitudes and upwelling equatorial and coastal regions, areas characterised by abundant large phytoplankton and high NPP. SLAMS-2.0's integrated POC flux at 100 m of 2.7 Gt C yr^{-1} is at the lower end of published values, which span from 2.8 Gt C yr^{-1} at 150 m (Buesseler et al., 2020b), 4.0 Gt C yr^{-1} at z_{eu} (Henson et al., 2012), 5.7 Gt C yr^{-1} at z_{eu} (Buesseler et al., 2020b), 6.3 Gt C yr^{-1} at 100 m (Stock et al., 2020), 6.6 Gt C yr^{-1} at z_{eu} (Siegel et al., 2014), 6.7 Gt C yr^{-1} at

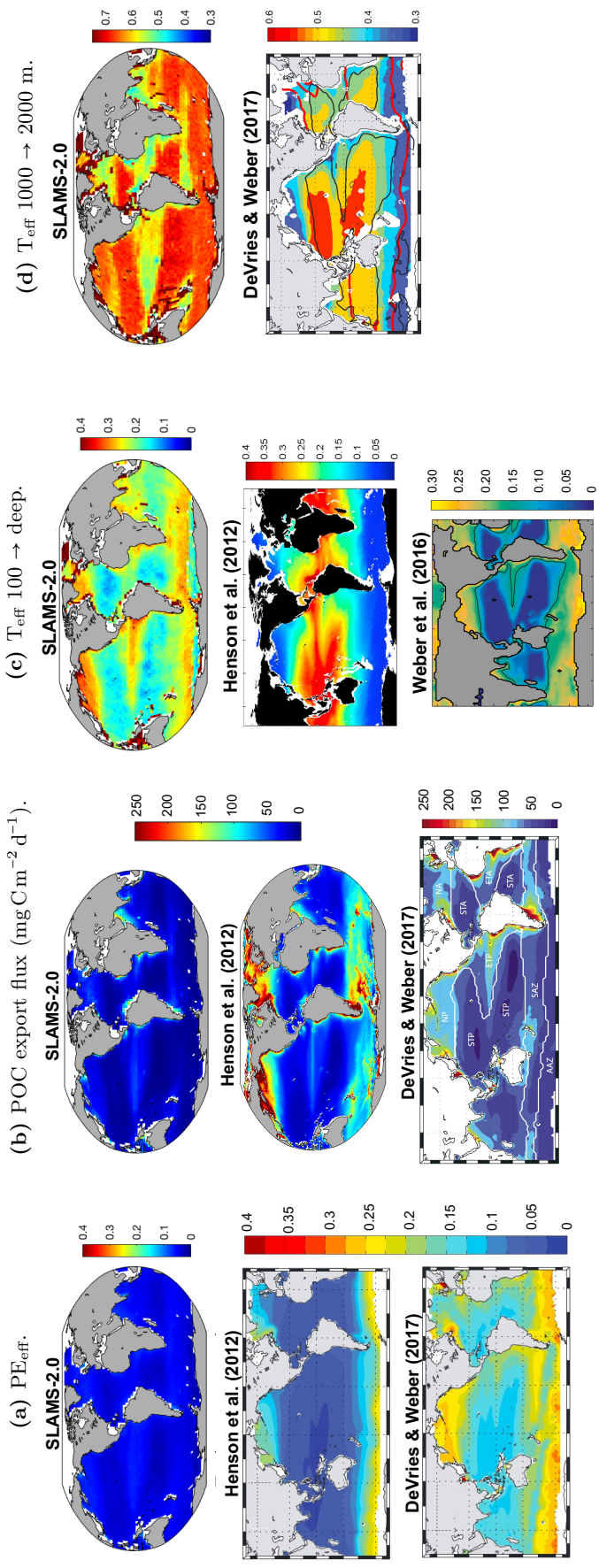


Figure 6.9. Comparison of SLAMS-2.0's modelled (a) PE_{eff}, (b) POC export flux, (c) T_{eff} surface → deep ocean and (d) T_{eff} 1000 → 2000 m (mesopelagic to deep ocean) against available published global assessments. Notice that the colour bar of SLAMS-2.0 output (from the OSIM1 run) has been changed to match that of published results, but that the SLAMS-2.0's results are the same as those shown in Figs. 6.1a–6.2a. Reprinted from Henson et al. (2012), DeVries and Weber (2017) and Weber et al. (2016) (no permission requested).

100 m (DeVries and Weber, 2017) to 9.1 Gt C yr^{-1} at z_{eu} (DeVries and Weber, 2017). Notice how the choice of the depth of integration matters as it shifts the value of carbon export up or down, with same author's reported values of POC flux lower for deeper depths. Since most of the attenuation of POC happens just below z_{eu}, a flux measured at 100 m (which is well below z_{eu} in polar areas) underestimates the *in situ* POC export flux. The integrated POC flux at 2000 m simulated by SLAMS-2.0 of 0.62 Gt C yr^{-1} is comparable to that of 0.43 Gt C yr^{-1} estimated by Honjo et al. (2008) and Stock et al. (2020) and 0.66 Gt C yr^{-1} estimated by Henson et al. (2012).

Compared to PE_{eff} and POC flux at export, the pattern of T_{eff} from the surface to the deep ocean (Fig. 6.9c) keeps closer resemblance in magnitude to that determined by other authors (Henson et al., 2012; Weber et al., 2016). It is worth emphasising that whereas the displayed T_{eff} for Weber et al. (2016) is calculated from 100 → 1000 m, that of Henson et al. (2012) is from 100 → 2000 m, thus running deeper and thus, reflective of a flux that has undergone more attenuation processes . SLAMS-2.0 has the very same regional pattern as that presented by the data-driven ocean circulation model of phosphate of Weber et al. (2016), with T_{eff} relatively low in the subtropics compared to the rest of the ocean. Contrarily, Henson et al. (2012) show relatively low T_{eff} in the poles and high values in subtropical and tropical regions. An inversion in the pattern of T_{eff} is observed for both SLAMS-2.0 and DeVries and Weber (2017) from the mesopelagic to the deeper ocean (Fig. 6.9d), where the subtropics have a higher T_{eff} (instead of the upwelling regions, as it is for the T_{eff} 100 → 1000 m) and the metric is less spatially variable (Fig. 6.9c).

The pattern of PIC and bSi export (Fig. 6.2a, bottom row) can be compared against the global observations of PIC and bSi fluxes of Honjo et al. (2008) for the 2000 m depth horizon (the mesopelagic/bathypelagic boundary) (Fig. 6.10, upper row). SLAMS-2.0's modelled PIC and bSi fluxes at export depth (100 m) are connected with those at 2000 m depth, and thus comparable in relative magnitude to those of Fig. 6.10. Honjo et al. (2008) show that observed PIC flux varies less geographically than bSi flux, which also occurs in SLAMS-2.0. The Atlantic presents a higher maximum PIC flux than the Pacific, and bSi flux is higher in the Pacific than in the Atlantic, which is at odds with the pattern modelled by SLAMS-2.0. The global-basin integrated PIC flux at export depth calculated in SLAMS-2.0 (0.48 Gt C yr^{-1}) is on the low end of prior estimates at export depth (0.38–1.8 Gt C yr^{-1}, as summarised in Dunne et al., 2007) and closer to the recent estimate made by Stock et al. (2020) of 0.40 Gt C yr^{-1}. At 2000 m, the global-basin integrated PIC flux in SLAMS-2.0 (0.47 Gt C yr^{-1}) is similar to that reported in Honjo et al. (2008) (0.41 Gt C yr^{-1}). When it comes to the global-basin integrated flux of bSi at export depth, the value estimated by SLAMS-2.0 (1.10 Gt Si yr^{-1}) is below the range estimated by Dunne et al. (2007) (1.9–5.0 Gt Si yr^{-1}) and still under the recent estimate

made by Stock et al. (2020) of 2.4 Gt Si yr^{-1}. Contrarily, SLAMS-2.0's modelled value at 2000 m (0.79 Gt Si yr^{-1}) is comparable to that of Honjo et al. (2008) (0.68 Gt Si yr^{-1}).

Lastly, the global pattern of SLAMS-2.0's modelled molar ratios of POC/PIC and bSi/PIC, can also be compared with the observations of deep-sea (2000 m) sediment trap fluxes of Honjo et al. (2008) (Fig. 6.10, bottom row, c.f. Fig. 6.3, bottom row) and the modelling study at export depth of Jin et al. (2006), where POC export is driven by diatoms and coccolithophores (see their Fig. 3, c.f. Fig. 6.3, upper row). In SLAMS-2.0, calcification by coccolithophores and, thus, biogenic production of PIC, does not account for the saturation state of the ambient water with respect to calcite (e.g., Gehlen et al., 2007) (see discussion in Sec. 6.5.5). That makes PIC fluxes at export depth higher than would be expected if a limitation factor for calcification had been accounted for, and thus POC/PIC tend to be lower than those estimated by the data-driven model of Jin et al. (2006). In Jin et al. (2006), POC/PIC export is maximum at the poles, followed by the subtropics and the Equator, with the Atlantic basin displaying the highest POC/PIC export ratio. In SLAMS-2.0, the poles, followed by the Equator, display maximum (i.e., > 1) POC/PIC, with the highest value also occurring in the Atlantic. Honjo et al. (2008) also display maximum POC/PIC at the poles, whereas the maximum bSi/PIC occurs in the North Pacific boreal gyres and the Antarctic Zone. SLAMS-2.0 also presents maximum bSi/PIC in the North Pacific, followed by the subpolar and polar North Atlantic and the upwelling coastal regions. According to Honjo et al. (2008), locations where POC/PIC and bSi/PIC are ≥ 1 constitute the "silica ocean", a diatom-dominated area where molar export of Si $>$ PIC, and molar export of POC $>$ PIC. Contrarily, regions where those rations are < 1 constitute the "carbonate ocean". In silica ocean mode, CO_2 is removed from the upper ocean more efficiently than in carbonate ocean mode as there are lots of export of POC flux with little loss of alkalinity from the surface waters. The notion that little loss of alkalinity from the surface ocean constitutes an efficient removal of surface POC was also expressed by Kwon et al. (2009), who demonstrated that an increase of alkalinity makes the ocean pull more atmospheric CO_2. Following that ocean divide, in SLAMS-2.0 the silica ocean comprises the high latitudes, the equatorial Pacific and the coastal upwelling regions, areas which, presumably, have a more efficient BCP.

6.4 Sensitivity experiments

I have investigated the sensitivity of SLAMS-2.0 to three biogeochemical forcings hypothesised to control the efficiency of the transfer of POC flux to the deep ocean through changes in particle characteristics: phytoplankton community composition (Guidi et al., 2009; Lam et al., 2011; Bach et al., 2016), NPP (Boyd and Trull, 2007; Jokulsdottir and

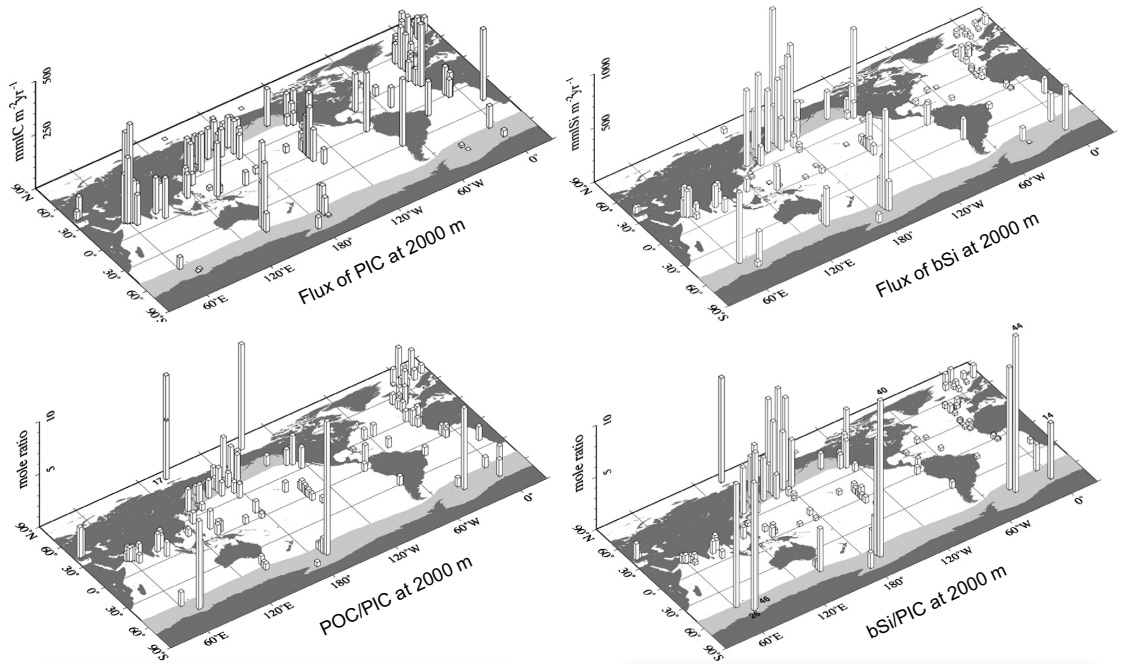


Figure 6.10. Honjo et al. (2008)'s compilation of observed molar annual fluxes of (top left) PIC, (top right) bSi, (bottom left) POC/PIC and (bottom right) bSi/PIC at 2000 m. The silica ocean, with bSi/PIC and POC/PIC ≥ 1, is depicted in grey, covering the North Pacific boreal gyres and the Antarctic Zone. The rest of the ocean is the carbonate ocean, with bSi/PIC and POC/PIC < 1. Reprinted from Honjo et al. (2008) (no permission requested).

Archer, 2016; Henson et al., 2019) and temperature (Weber et al., 2016; Cram et al., 2018; DeVries and Weber, 2017). For that, I have carried out a sensitivity analysis at two ocean locations that are distinct enough biogeochemically: Ocean Station Papa (OSP), in the northern North Pacific, and the Bermuda Atlantic Time-Series/Oceanic Flux Program (BATS/OFP) joint site, in the subtropical North Atlantic.

The sensitivity analysis has consisted in assessing how sensitive the pattern of POC flux with depth is to (i) varying surface NPP, (ii) varying temperature profile and (iii) varying relative biomass abundance of each of the four modelled PFTs. I have varied the forcing variables within wide ranges according to fractional increments while keeping the rest of the forcing variables fixed. Of the three variables, temperature perturbations present the most complex implementation as a change in temperature simultaneously affects many processes in SLAMS-2.0 that can offset each other and, thus, mask the actual effect of temperature in the system: temperature-controlled phytoplankton growth, the viscosity of the water (and thus, particle coagulation rates, particle-zooplankton encounter rates and particle sinking velocity), respiration rates of zooplankton and bacteria and zooplankton biomass through the calculation of the DVM depth at night. It can be seen that the effects of temperature can be split according to a physical/biological divide. Temperature perturbations have been implemented by keeping unchanged alterations to phytoplankton

growth (I fix the PFTs probabilities to their annual mean) while affecting the rest of processes. The results in Fig. 6.11 demonstrate that the pattern of POC flux with depth is essentially insensitive to temperature. In contrast, the greatest changes in POC flux occur when altering NPP and diatom relative abundance. As the forcing NPP and the diatom relative abundance increase, so does the amount of POC exported from the surface ocean. Interestingly, an inverse relation with the amount of POC flux exported is seen for the other PFTs (dinoflagellates, coccolithophores and picophytoplankton), where an increase in the relative abundance generates a decrease in the amount of exported POC flux.

To assess whether NPP and phytoplankton community composition have a significant impact on the transfer characteristics of POC to the deep ocean, I have calculated the values of Martin b, z^*, PE_{eff} and T_{eff} 100 → 1000 m that emerge from the perturbed POC flux profiles (Fig. 6.12). Those show that b (first column across all six rows) is negatively correlated with the NPP and the relative abundance of the large and silicified diatoms, meaning that higher attenuation of POC flux is expected at lower NPP levels and lower relative abundance of diatoms. Conversely, b is positively correlated with the relative abundance of dinoflagellates, coccolithophores and picophytoplankton, expecting higher attenuation rates of POC when those smaller PFTs are more abundant. b does not respond to varying temperature, as expected from the perturbed POC flux profiles (Fig. 6.11). For all the six forcings to which the sensitivity of the model output has been tested, OSP, in a subpolar region, presents lower b values compared to BATS/OFP, a subtropical location. The pattern for z^* (second column across all six rows) has the opposite sign to that of b, as expected from the definition of the metrics. Thus, changes in the NPP and PFT forcings that cause b to decrease will make z^* increase. Notice how z^* went off the scale for higher relative contributions of diatoms to the phytoplankton biomass. Temperature has, again, no effect on z^*. The change in the magnitude of PE_{eff} (third column across all six rows) presents a less neat pattern. A higher PE_{eff} is expected for higher NPP values, higher diatom abundance and lower dinoflagellate abundance, but the trend is less clear for coccolithophores and picophytoplankton. PE_{eff} seems to respond to changes in temperature at OSP: as temperature increases, PE_{eff} decreases. Since PE_{eff} is proportional to POC flux at export (by definition of the metric), and POC flux at export presents a pattern where it is higher in upwelling and convective systems, which are colder (Fig. 6.2, top left), PE_{eff} will then reflect this temperature effect through the amount of POC export. Lastly, T_{eff} 100 → 1000 m (fourth column across all six rows), like z^*, is a metric negatively correlated with b and thus should display an opposite behaviour to b. Consequently, higher transfer of surface POC to depth is observed at higher NPP (at OSP, but the trend is less clear for BATS/OFP)

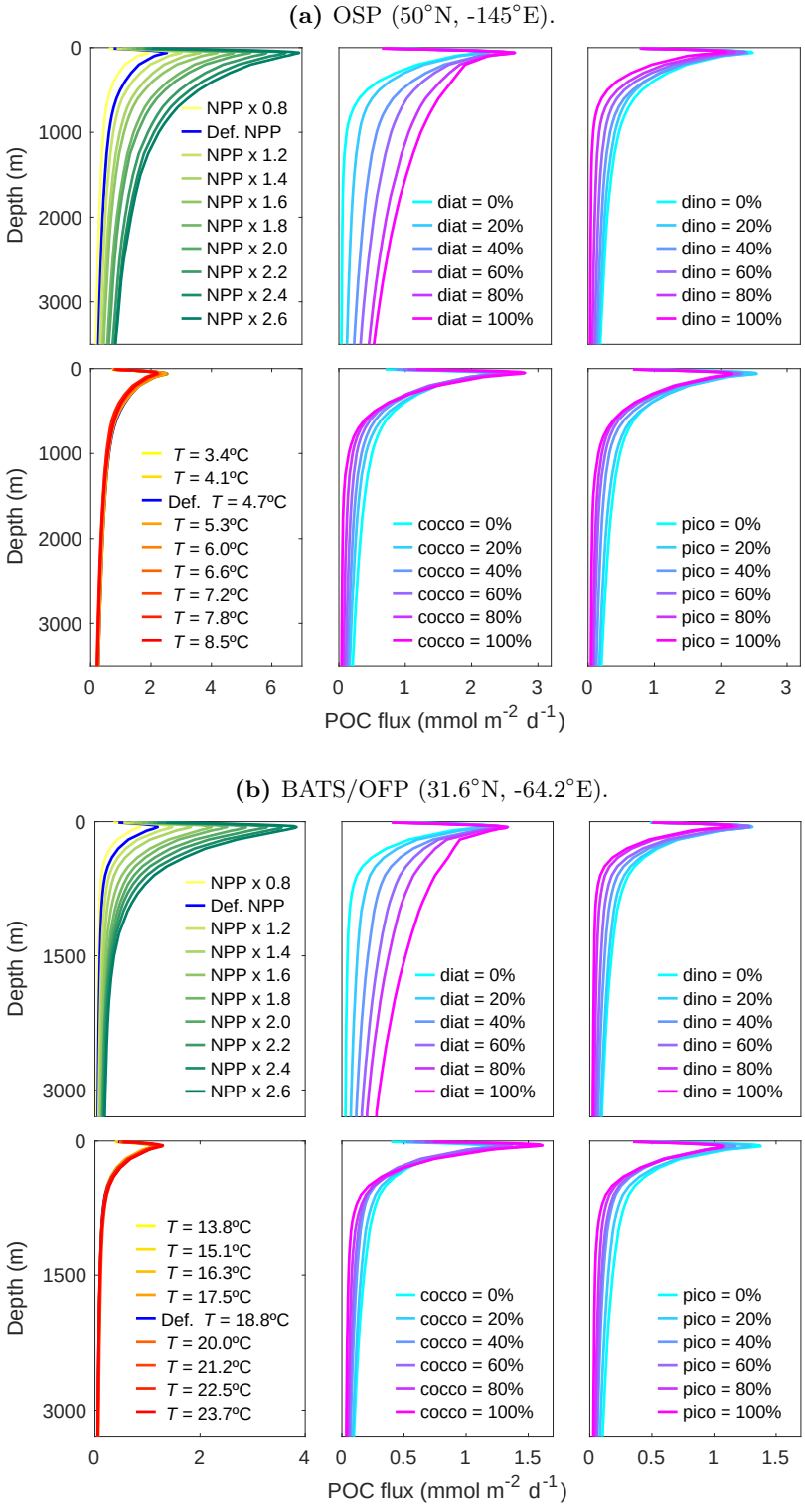


Figure 6.11. Sensitivity of modelled POC flux at location (a) OSP and (b) BATS/OFP to varying (top left) net primary production (NPP), (bottom left) temperature (T) profile, shown as the average upper (500 m) water column temperature, and (centre and right) phytoplankton relative biomass abundance (diatoms, dinoflagellates, coccolithophores and picophytoplankton). Default (Def.) experiments show the default forcing used to run SLAMS-2.0. NPP and phytoplankton seeding probabilities are imposed as one unique value in the model, whereas T is imposed as a depth profile for which I have altered the annual mean in the upper 700 m perturbed by a factor (the T mean shown is for the first 500 m).

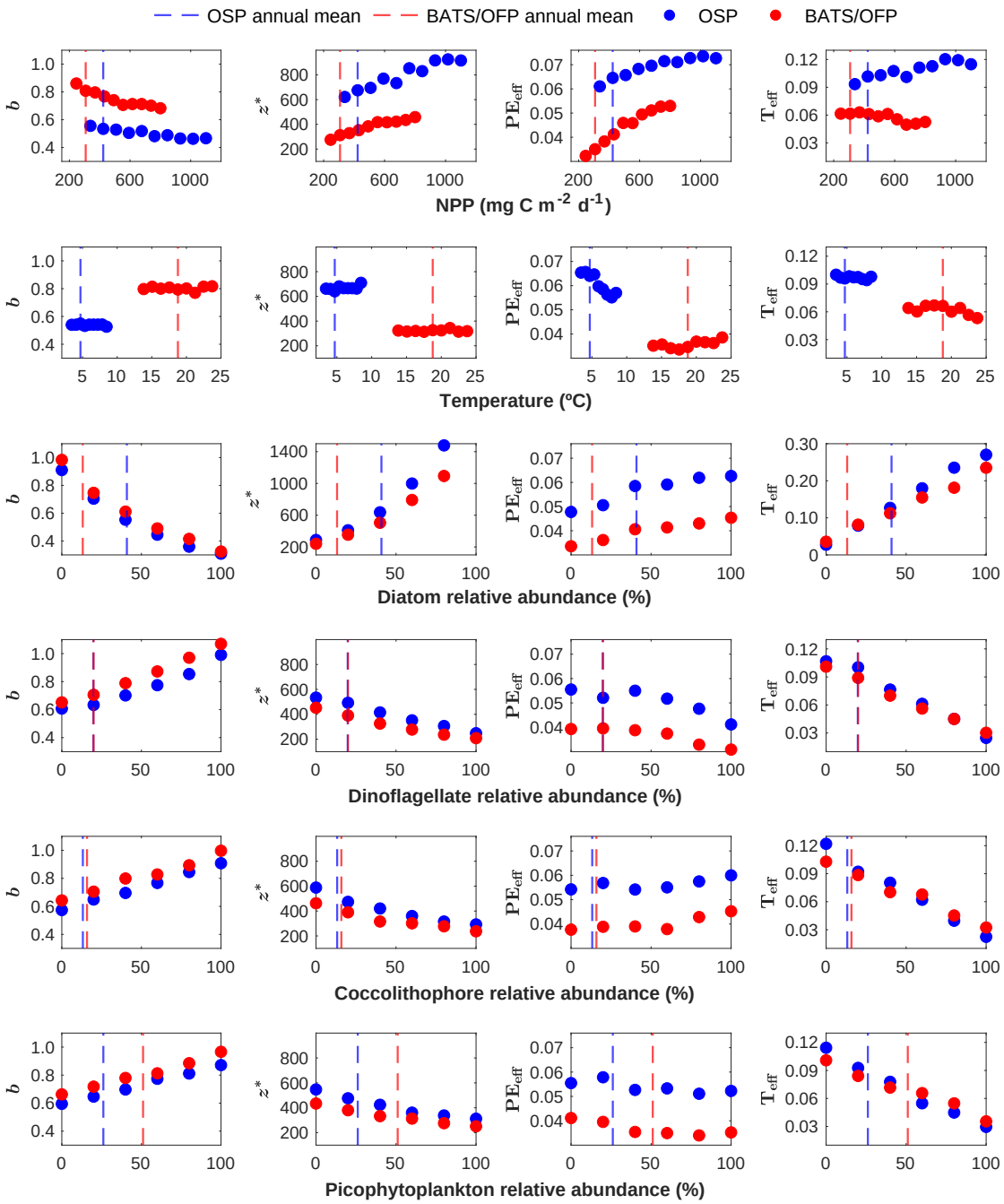


Figure 6.12. Relationship between six forcing variables (NPP, temperature and relative biomass abundance of diatoms, dinoflagellates, coccolithophores and picophytoplankton) and four metrics of the BCP (Martin b, z^*, PE_{eff} and T_{eff} 100 → 1000 m) at two different ocean locations (OSP, in blue, and BATS/OFP, in red). The default annual means of each of the six forcing variables used to run SLAMS-2.0 are shown for reference.

and higher diatom dominance. On the other hand, lower transfer efficiency is observed for increasing levels of the smaller phytoplankton groups.

Overall, Fig. 6.12 suggests that lower flux attenuation (b) or, conversely, higher transfer of surface POC to the deep ocean (T_{eff}), is expected for higher levels of NPP and diatom

dominance. There is no temperature signal in such metrics as derived from a model of Lagrangian marine particles like SLAMS-2.0. Nevertheless, temperature does have an effect on those metrics when derived empirically (e.g., Henson et al., 2012; Marsay et al., 2015) or using Eulerian models of the BCP (e.g., Weber et al., 2016; DeVries and Weber, 2017). Temperature is a variable whose various effects on the ecosystem (i.e., temperature-controlled biological effects vs temperature-controlled physical effects) are challenging to disentangle. For instance, in the sensitivity analysis carried out with SLAMS-2.0, those biological vs physical effects have not been fully separated yet on the particle dynamics. They are probably still offsetting each other, explaining the lack of correlation between temperature and the BCP metrics as modelled by SLAMS-2.0. Nevertheless, when it comes to empirical studies and data-driven models of the BCP, which are run with biogeochemical forcing that results from temperature (e.g., NPP is the result of temperature-controlled phytoplankton growth), it becomes easier to find a correlation between BCP metrics derived from such biogeochemical forcing and temperature. Almost all ecosystem variables in the ocean are correlated with temperature in one way or another. However, when it comes to particle dynamics, the mechanistic correlation with temperature becomes less clear and masked by the combination of several biological and physical processes.

6.5 Discussion

Mapping out the metrics that quantify the transfer of the flux of C_{org} to the ocean interior, like Martin's b or T_{eff}, to the underlying marine particle attributes that control the magnitude and efficiency of that transfer has become a research priority (Buesseler and Boyd, 2009; Sanders et al., 2014; Siegel et al., 2016; Sanders et al., 2016). I have applied a Lagrangian model of marine particles globally to illustrate which environmental variables and particle attributes explain the global pattern of b and T_{eff}. Modelled results show that a low b (or, conversely, a high T_{eff}) occurs in high latitudes and the upwelling regions of the equatorial Pacific and the west coast off the continents; b is highest (and T_{eff} is lowest) in the subtropics. That pattern responds to NPP and percentage abundance of diatoms, and overlaps with regions of differential zooplankton activity and coagulation. That latitudinal pattern found by SLAMS-2.0 agrees with the findings of the empirical study of Marsay et al. (2015) and the modelling studies of Weber et al. (2016) and DeVries and Weber (2017), and it is at odds with the findings of the empirical study of Henson et al. (2012) (further supported by Lam et al. (2011) or Guidi et al. (2015)). My modelled b and T_{eff} emerge from tracking an ensemble of discrete particles built from structural and chemical characteristics (fractal scaling law and Stokes' velocity). That aspect makes this study the first to address the global pattern of b and T_{eff} based on detailed information of *particle attributes* of composition and structure and relates them to the *biogeochemical variables* and *particle dynamics* that control them.

6.5.1 Biogeochemical controls

A range of BCP metrics exist to understand the efficiency and magnitude of the transfer of POC from the surface to the deep ocean (b, z^*, T_{eff}, PE_{eff}, β). Out of those, b, or the attenuation rate of POC flux with depth, and T_{eff} (100 → 1000 m), or the fraction of surface POC flux transferred to the mesopelagic ocean, have become the standards in the biogeochemical literature to inform of the characteristics of the flux of POC to the ocean interior. Thus, the following discussion uses one or another indistinctively to evaluate the functioning of the BCP.

Marsay et al. (2015) found a strong correlation between b and temperature (R^2 = 0.82), where the two variables are positively correlated (i.e., the attenuation rate of POC increases with increasing temperature). Henson et al. (2012)'s study, although it does not explicitly mention the relationship that their estimated b bears with temperature, a b–T relationship can be calculated from their model. A contour plot between Henson et al. (2012)'s algorithm input variables (NPP and SST) and b reveals that temperature strongly controls Henson et al. (2012)'s b (Fig. 6.13a, with R^2 = 0.98), and that the two variables are negatively correlated (i.e., the attenuation rate of POC increases with decreasing temperature). That is at odds with the findings of Marsay et al. (2015), where the attenuation rate of POC increases with increasing temperature. When it comes to NPP, Fig. 6.13a's contours are essentially vertical for that variable, indicating that Henson et al. (2012)'s estimated b is quite insensitive to NPP. To meaningfully compare the distribution of b values, it is important to bear in mind how differently Henson et al. (2012) and Marsay et al. (2015)'s b are calculated (summarised for reference in Table G.1). Whereas the Henson et al. (2012) algorithm heavily relies on an empirically-derived PE_{eff} using ^{234}Th-derived export measurements regressed against SST and traps at 2000 m to derive POC flux in the deep ocean, Marsay et al. (2015)'s equation relies on sediment trap fluxes of POC in the upper water column (50–800 m), which are used to estimate a b that is regressed against upper-water column (0–500 m) temperature.

I have quantified the sensitivity of SLAMS-2.0's Lagrangian-derived b and T_{eff} to local changes in NPP, temperature and phytoplankton community composition and have shown that T_{eff} and b are most sensitive to changes in NPP and diatom relative abundance, and insensitive to temperature (effects summarised in Fig. 6.13b for comparison with Henson et al. (2012)'s findings). That demonstrates that, in SLAMS-2.0, the transfer of surface POC to the deep ocean is strongly dependent on how much organic carbon is produced in the surface layer (i.e., NPP) by silicified, sinking phytoplankton cells (diatoms). Thus, ballast makes a significant contribution to the pattern of T_{eff} and b, as hypothesised by Francois et al. (2002), Klaas and Archer (2002) and Armstrong et al. (2002) (the so-called "ballast effect").

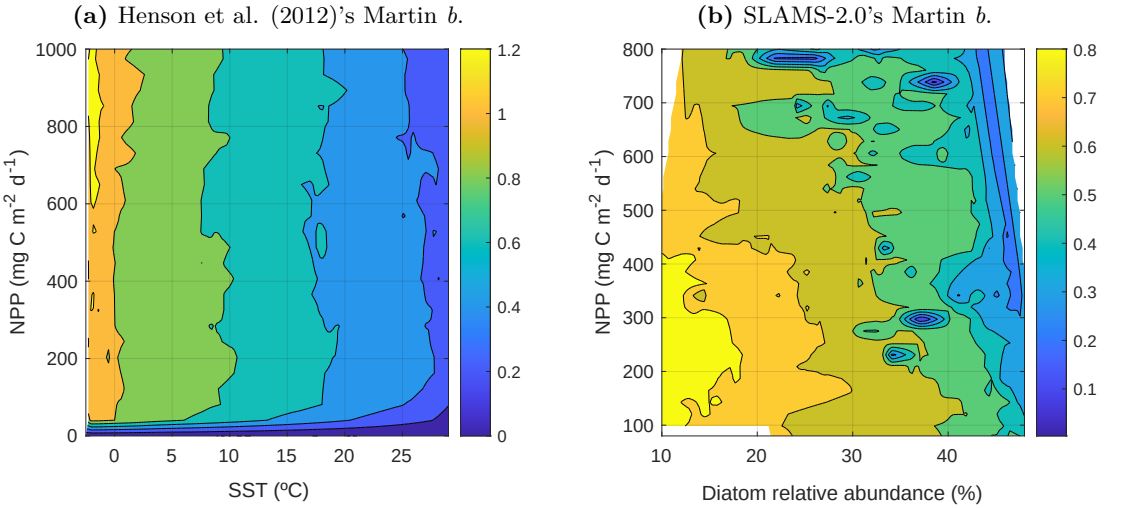


Figure 6.13. Contour plot of Martin's b calculated from (a) the empirical model of Henson et al. (2012), and its dependency on net primary production (NPP) and sea surface temperature (SST) and (b) the Lagrangian model SLAMS-2.0, and its dependency on NPP and diatom relative biomass abundance.

Weber et al. (2016) used an ocean circulation inverse model driven with phosphate data to reconstruct particle flux profiles below 300 m (see details in Table G.1). Similarly to SLAMS-2.0, Weber and colleagues found that T_{eff} is highest in the equatorial Pacific and high-latitude regions and lowest in the subtropics. They supported their findings based on a strong correlation of T_{eff} with picophytoplankton dominance and calcite+silicate export to POC export (ballast ratio). DeVries and Weber (2017) used the same ocean circulation inverse model as that of Weber et al. (2016) but had a biological model embedded in it (see details in Table G.1) and found a similar pattern for T_{eff} in the upper water column (100 → 1000 m) which, in turn, was controlled by temperature and O_2 levels. From those two modelling studies and supporting findings in the literature (Guidi et al., 2009; Lam et al., 2011; Cram et al., 2018), it can be concluded that phytoplankton community structure, ballast minerals, temperature and O_2 are popular factors hypothesised to control the attenuation of POC flux with depth.

The negligible influence of temperature on the sinking flux of Lagrangian particles in SLAMS-2.0 does not deny the existence of an effect of temperature on b or T_{eff} in nature, but that its effect on b or T_{eff} through changes in Lagrangian particle dynamics (particle sinking velocity, particle-particle coagulation, particle-zooplankton encounters, remineralisation rates of organic carbon, depth of zooplankton DVM) is not strong enough. Temperature is a proxy for emergent ecological variables, like NPP, which is imposed in SLAMS-2.0 as model forcing. Thus top-down controls of temperature in the model through the amount of phytoplankton produced vanish. It can also be hypothesised that a b–T correlation, as that inferred by Marsay et al. (2015) ($b = 0.062T + 0.303$,

R^2 = 0.82), does not prove direct connection between POC flux attenuation rates and temperature, but that both variables can be mechanistically linked through various other underlying processes (e.g., water viscosity, stratification, subsurface chlorophyll maxima, etc.). Correlation does not always imply direct causation.

In SLAMS-2.0, the high-latitude and upwelling regions displaying high T_{eff} overlap with regions displaying low b (Fig. 6.1a), which in turn overlap with regions of high opal export (Fig. 6.2a) that exhibit high (i.e., > 1) molar POC/PIC and bSi/PIC ratios (Fig. 6.3a). In those opal-rich regions, fewer carbonate particles are available to ballast POC, and thus bSi assumes the ballast role for $CaCO_3$, following the silica ocean regime of Honjo et al. (2008). Since that way of BCP's functioning comes with minor loss of surface alkalinity , the removal of surface POC is more efficient in high-latitude and upwelling regions than for the rest of the ocean (Honjo et al., 2008; Kwon et al., 2009).

I have extracted a first-order relationship to link SLAMS-2.0's modelled b to its most relevant biogeochemical controls (as per Fig. 6.12): F_{diat} (%), the fraction of total phytoplankton biomass contributed by diatoms, and NPP ($g\,C\,m^{-2}\,d^{-1}$), the net primary production (R^2 = 0.63, p-value < 0.05),

$$b = 1.0 - 0.01\,F_{diat} - 0.20\,\mathrm{NPP}. \tag{6.1}$$

Eq. 6.1 offers a parameterisation for Martin's b that leaves aside assumptions about the shape of the decaying POC flux (i.e., power-law or exponential decay), separates its most relevant biogeochemical controls (NPP and opal ballast effects) and provides an equation that is easier to implement in an ESM than a Lagrangian model of marine particles. Such an equation is just a starting point towards a mechanistic parameterisation for the attenuation of POC flux with depth and will be tested in further applications of SLAMS-2.0 to understand its implications.

6.5.2 Particle attribute controls

To date, there are no descriptions of global patterns of particle attributes as presented in this thesis, only local descriptions made in mesocosms, like those shown in Ploug et al. (2008), Iversen and Ploug (2010), Laurenceau-Cornec et al. (2015) or Iversen and Lampitt (2020). The present thesis is thus the first study to present modelled particle attributes on a global scale in the ocean. The Lagrangian nature of SLAMS-2.0 makes it possible to get a "photographic" impression of the average particle in the ocean as it goes through depth. The model predicts that, at deeper depths, particles are, on average, faster and larger (Figs. 6.5 and 6.6) as they have had more time to undergo coagulations (which add primary particles to the aggregate structure) and being added more ballast materials. Indeed, deeper particles are impoverished in POC with respect to surface particles (Fig. G.6) and

enriched in opal (Fig. G.8), making them denser as the balance of material composition shifts from lighter POC to heavier bSi (Fig. G.2) and less sticky as the amount of TEC decreases (Fig. G.5). Deeper particles are also slightly more porous as the accumulation of more primary particles in the aggregate structure expands the particle in size following fractal scaling law, as dictated by the particle structure model used in SLAMS-2.0. The higher porosities observed in the diatom-dominated area of the equatorial Pacific and clay-dominated area off the west coast of North Africa (Fig. G.3) can be explained by the effect that minerals have on the porosities of modelled faecal particles: diatom spicules (equatorial Pacific) and dust-borne clays (North African coast) favour more open particle structures, increasing the porosity of faecal particles rather than compressing them.

Modelled sinking velocities (Fig. 6.6b) compare well with observations, where aggregates created in mesocosm experiments are seen sinking at speeds 100–500 $\mathrm{m\,d^{-1}}$ (Iversen et al., 2010). Regarding modelled particle sizes (Fig. 6.5), modelled large particles are under the observed ranges for large marine snow (300–3000 μm, Iversen et al., 2010; McDonnell and Buesseler, 2010). That is because the coagulation model used in SLAMS-2.0 (Jackson, 2001) has limitations to creating large particles if it is not using an unrealistically high shear stress regime.

Whereas the average sinking velocity of a particle (small or large) keeps increasing with depth, the size of larger aggregates does not. Below some threshold depth (~1000 m), the average size of a large aggregate decreases with depth, thus showing signs of compression (or compactness) as the increase in its density due to the accumulation of ballast materials overcomes the increase in porosity (or fluffiness) due to the addition of more primary particles, following fractal scaling. That explains why particles that sink faster are not necessarily, on average, bigger, as shown by the disassociation of correlation between larger particle sizes and sinking velocity (Figs. G.10 and G.11).

The lack of correlation between the size of an average modelled large aggregate and its sinking velocity (R^2 = 0.05, but not for small particles, R^2 = 0.94, Figs. G.10–G.11) challenges the paradigm that fast-sinking aggregates must be bigger, as generally found for laboratory aggregates of similar composition (e.g., Alldredge and Gotschalk, 1988; Iversen and Ploug, 2010) as well as commonly assumed in many biogeochemical models. However, *in situ* aggregate sinking velocities have also been seen offering a poor correlation with aggregate size (Laurenceau-Cornec et al., 2015, 2019; Iversen and Lampitt, 2020), and even an overall decrease of sinking velocity with size (McDonnell and Buesseler, 2010), in agreement with the findings of SLAMS-2.0. The variety of relationships size-sinking velocity found in the literature demonstrates that the various factors that determine that relationship, such as ballast minerals (Klaas and Archer, 2002), sticky TEPs (Engel and Schartau, 1999), porosity or degree of packaging (Passow et al., 2019) or ambient

temperature (Iversen and Ploug, 2013), when combined together, obscure the overall size-dependency of sinking velocities. In SLAMS-2.0, I have found that modelled sinking velocity of similar-sized aggregates is largely controlled by density, which for large particles is mostly controlled by the bSi content and for small particles is mostly controlled by the $CaCO_3$ and clay content (Figs. G.10b and G.11b). Regarding modelled sizes, results show that larger particles consist of voluminous aggregates rich in C_{org}, whereas smaller particles consist of densely packed material with a high content of clay and $CaCO_3$, in a similar fashion to the findings of Ploug et al. (2008). Thus, SLAMS-2.0 finds that the ballast effect of opal and calcite in particles is divided according to particle size class: bSi is the ballast material of large particles (R^2 = 0.23) whereas $CaCO_3$ is the ballast mineral of small particles (R^2 = 0.86), supporting the established paradigm that ballast by both $CaCO_3$ and bSi may enhance T_{eff} of POC (Francois et al., 2002; Klaas and Archer, 2002; Armstrong et al., 2002). Porosity, though, does not seem to be a trait that characterises a specific cohort of particles, as hypothesised by Iversen and Ploug (2010) and Bach et al. (2016), as it does not show a strong correlation with sinking speed nor radius ($R^2 < 0.20$). Taken altogether, the results of the particle attribute relationships suggest that chemical composition and density predict the largest part in the variation of aggregate sinking velocities, overshadowing the effect of size.

For the ballast hypothesis to be correct, minerals must either increase the sinking velocities of large aggregates or protect them from microbial degradation (Francois et al., 2002; Klaas and Archer, 2002; Armstrong et al., 2002). Whereas the protection effect cannot be proved with SLAMS-2.0 (it has not been explicitly modelled using a mechanistic parameterisation), an increase in sinking velocity can be seen due to the increasing content of both bSi and $CaCO_3$ for both small and large particles (Figs. G.10b and G.11b). That proves the existence of a ballast effect in SLAMS-2.0.

To assess which particle characteristics have the most influence on the attenuation rate of POC flux with depth, I have plotted b as a function of various modelled particle attributes for an average fast-sinking particle at 500 m depth, the upper mesopelagic (Fig. 6.14). I have chosen to highlight fast-sinking particles because of their overall higher correlation with other attributes when compared to slow-sinking particles. From the generally low correlation coefficients, it can be speculated that it is not possible to infer a universal scaling between b and particle attributes easily as, at any depth in the ocean, there is a highly heterogeneous pool of particles. POC, followed by bSi, are the two material components that are most correlated with b. For higher b, or higher POC attenuation rates with depth, a reduction in the volume, and thus radius of the fast-sinking aggregates, is observed, following the reduction in POC content. Especially striking is the lack of any correlation between the average sinking velocity of a large

particle with b. That means that it is not possible to justify lower attenuation rates of POC flux (or lower b) based on the presence of fast-sinking particles that escape remineralisation, which is at odds with current understanding of the relationship between b and particle sinking velocity (e.g., Herndl and Reinthaler, 2013; Cavan et al., 2015; Cram et al., 2018). The mathematical construct of b from the advection-diffusion-reaction equation of a tracer (see Appendix A) predicts that higher sinking velocities will lower b. However, that relationship is not observed in Fig. 6.14; instead, for decreasing values of b, the sinking velocity is not markedly higher. That can be explained based on the simultaneous increasing porosities (i.e., decreasing aggregate compactness, or density), which compensate the increase in ballast by bSi, and so sinking velocities are unchanged.

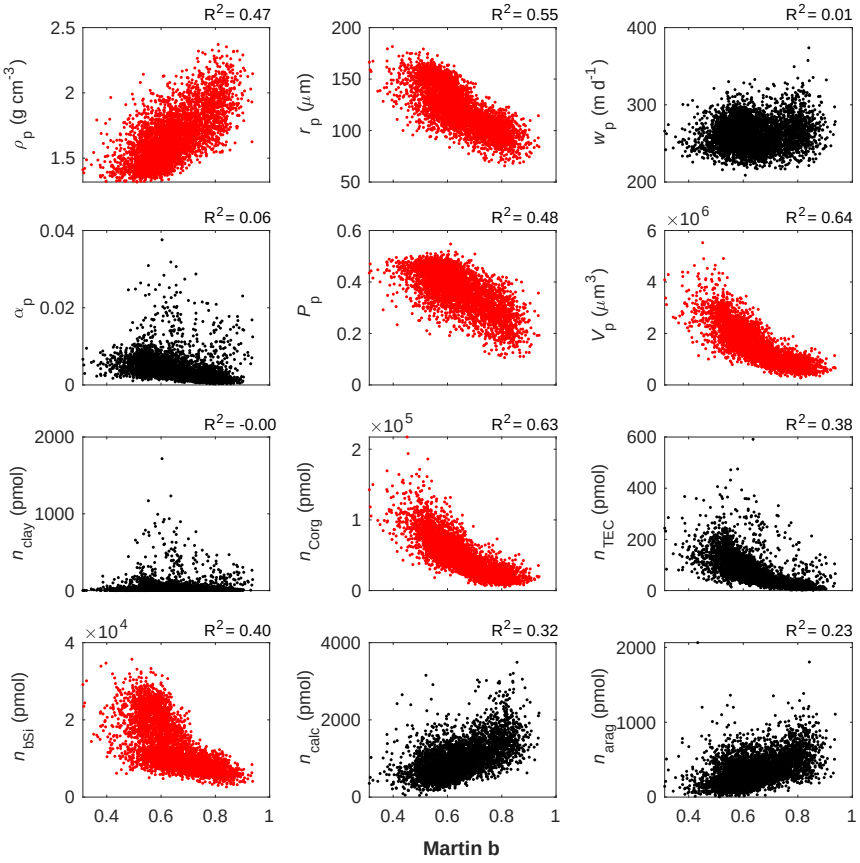
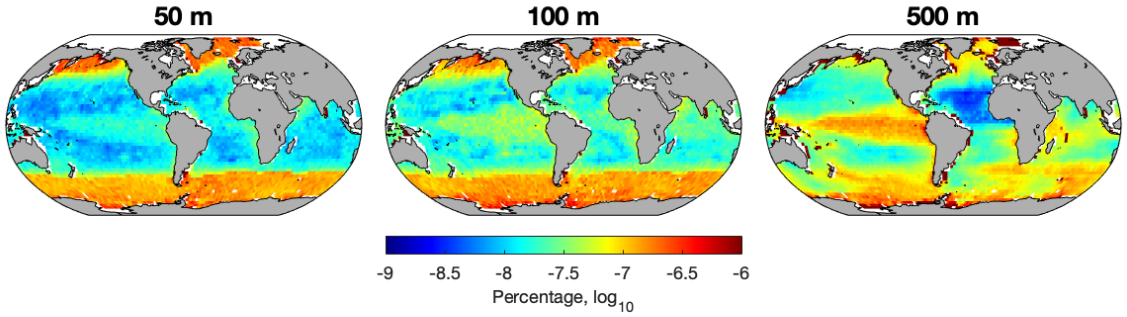


Figure 6.14. Relationship between Martin's b and the modelled particle attributes at the 500 m depth horizon for fast-sinking (> 100 m d^{-1}) particles. Highlighted in red are those relationships with higher correlation coefficient (R^2). Data plotted are from the OSIM1 experiment. Particle attribute symbols are: ρ_p, density; r_p, radius; w_p, sinking speed; α_p, stickiness; P_p, porosity; V_p, solid volume; n_{clay}, clay content; n_{Corg}, C_{org} content; n_{TEC}, TEC content; n_{bSi}, bSi content; n_{calc}, calcite content; n_{arag}, aragonite content.

6.5.3 Particle dynamics controls

The global distributions of average particle attributes presented in Figs. 6.5–6.6 and Figs. G.2–G.9 do not show a strong latitudinal pattern. Nevertheless, if each modelled small (≤ 150 μm) and large (> 150 μm) particle contributing to the average global is counted (Fig. 6.15), it can be seen that areas of high T_{eff}, like the equatorial Pacific, have characteristically a higher fraction of large and fast-sinking particles (below 100 m). That pattern is expected from the more porous, sticky and enriched in POC and bSi particulate environment material. Those higher numbers of large and fast-sinking particles are generated by relatively higher collision and coagulation rates (Figs. 6.7a and 6.7c) and a relatively higher zooplankton activity (Fig. G.12).

(a) Fraction of large particles (> 150 μm).



(b) Fraction of fast-sinking particles (> 100 m d^{-1}).

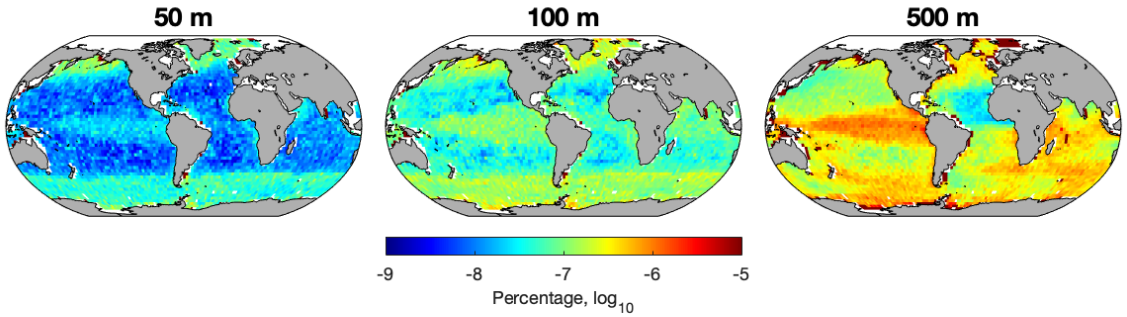


Figure 6.15. Global maps of SLAMS-2.0's modelled particle distributions in terms of (a) fraction of large particles and (b) fraction of fast-sinking particles. Notice that the axes are in $\log_{10}$.

To what extent the modelled zooplankton activities of sloppy feeding (attenuating POC flux), faecal pellet production (increasing POC flux) and respiration (attenuating POC flux) contribute to the particle attributes shown and the patterns of T_{eff} is beyond the scope of the present thesis. Nevertheless, a qualitative inspection of Figs. 6.7c and G.12 suggest that the low T_{eff} regime of the subtropics could be a consequence of lower zooplankton-particle encounter rates and, therefore, lower grazing activities. I propose that, through grazing, zooplankton recycle the rain of particulate material by reprocessing it in their guts and producing a new ensemble of particles, the faecal pellets, with a

more well-mixed material composition that includes all the organic and mineral material ingested. It could well be that if that gut mixing mechanism is not so intense, as it happens in the subtropics, with a lower zooplankton activity, the rain of particles remains more "untouched". Those particles do not get the opportunity to well-mix with others, remaining smaller and slower-sinking. Otherwise said, a less rich variety of particles is sinking in the subtropics. Subtropical particles are, in turn, less porous (or more compact), less sticky and with higher PIC content as this has not had the chance to undergo dissolution in zooplankton guts. Conversely, particles more intensely processed by zooplankton (equatorial Pacific regime and northern high latitudes) would offer the opposite picture: those are more diverse in material composition, stickiness, size and sinking speed, and those attributes accelerate their aggregation and thus the transfer signal. A higher abundance of fast-sinking particles in the equatorial Pacific regime and the high latitudes was also found by the modelling study of Niemeyer et al. (2019).

Phytoplankton abundance results (Fig. 6.4) showed that diatom-dominated communities overlap with areas where zooplankton metabolic activity is high (Fig. G.12). Those areas (the northern high-latitudes and the upwelling regions) are high regenerative systems where opal-rich biogenic particles are potentially subjected to a higher degree of biotic reprocessing (mixing of the material in zooplankton guts), distributing the bSi initially present in diatom cells amongst the broader community of particles. The export of OM that has been more extensively processed by the food web accelerates the transfer of surface material to the deeper ocean and leads to the high T_{eff}, or low b values, observed in those ecosystems. That pattern results not only from high NPP and relatively high diatom abundances (as shown by the sensitivity analyses), but also from food-web particle dynamics that revolve around the primary producers.

6.5.4 Reconciling Henson et al. (2012) and Marsay et al. (2015) schools of thought

The particle density-based control on POC transfer efficiency advocated in this thesis is generated by an active zooplankton community that mix well the ballast opal generated by diatoms amongst the broader pool of marine particles. The importance of the role of zooplankton repackaging in the transfer efficiency of POC is supported by an increasing number of studies (Guidi et al., 2009; Lam et al., 2011; Henson et al., 2012; Villa-Alfageme et al., 2016; Henson et al., 2019), following Henson et al. (2012)'s school of thought. Nevertheless, the mechanisms invoked by those observational studies are different from the ones put forward by SLAMS-2.0.

In Henson et al. (2012)'s school of thought, high T_{eff}, or low b, are seen in low-latitude (subtropical and tropical) systems, which correlate with higher temperature. Those

warmer regions are typically non-bloom systems, dominated by small phytoplankton cells (picophytoplankton and coccolithophores), where there is a paucity between phytoplankton growth and zooplankton grazing, which crop down the phytoplankton biomass stock. Henson et al. (2012) argue that zooplankton repackage the phytoplankton cells into tightly packed, fast-sinking aggregates that lead to a high T_{eff} (100 → 2000 m) signal seen in the deep-sea sediment trap observations. In productive, high-latitude systems, NPP outpaces zooplankton grazing. In those regions, large and fluffy, diatom-rich aggregates are not packed into smaller aggregates by the zooplankton community, which cannot catch up with phytoplankton growth, leading to the reverse pattern in T_{eff} (low) and b (high). Note that in SLAMS-2.0 zooplankton grazing is a mechanism used not to "package", but to spread opal amongst the community of particles, increasing their average density. Furthermore, modelled grazing is more intense in high-latitude and the equatorial Pacific, not overlapping so much with a temperature spatial pattern but with a diatom-abundance pattern.

The results of SLAMS-2.0 support the school of thought of Marsay et al. (2015), where high T_{eff}, or low b, are seen in upwelling ecosystems (the high-latitudes followed by the equatorial Pacific), which in turn are colder than neighbouring regions. b is thus positively correlated with temperature (low b at low temperatures), whereas for Henson et al. (2012), b was negatively correlated with temperature. Marsay et al. (2015) hypothesise that, in warmer regions, due to the respiration metabolism accelerated by temperature, the T_{eff} (100 → 1000 m) signal is weak (thus, b is high) as most of the material is remineralised in the upper water column. Furthermore, in those regions, with a phytoplankton community dominated by less labile cells compared to diatoms, the fewer fraction of exported material to the deep ocean remains more intact as it reaches deeper depths (e.g., 2000 m). That explains the inversion in the pattern seen between low T_{eff} (100 → 1000 m) but highest deep ocean T_{eff} (1000 → 2000 m) in the subtropics . The opposite pattern between surface and deep transfer efficiency is justified based on different mechanisms controlling T_{eff} in the two zones: a temperature-dependence of respiration in the upper water column and a lability-dependence in the lower water column, with less labile particles (more resistant to remineralisation) producing higher T_{eff}. The inversion in the T_{eff} pattern observed in SLAMS-2.0 for oligotrophic regions can be attributed to the increasing load of $CaCO_3$ in the sinking flux of particles. As POC is remineralised through depth, the relative weight of $CaCO_3$ increases through depth, adding ballast to the particle, which increases the transfer signal. Since $CaCO_3$ is denser than opal, the ballast in oligotrophic systems at deeper depths might be higher than in opal-dominated areas.

Henson et al. (2012) suggested that T_{eff} to 2000 m is highest in low-latitude systems, where zooplankton are repackaging material more intensively. That coincided with the also higher lability-dependant deep T_{eff} of Marsay et al. (2015) for those warmer regions (see their Fig. 4). SLAMS-2.0 just reveals another dependance for T_{eff}, that which comes with NPP and phytoplankton community composition.

6.5.5 Unexpected modelling results

The BCP metrics discussed above are supported by POC, PIC and bSi fluxes at export depth that, in SLAMS-2.0, have globally integrated values lower than those found elsewhere in the literature (F_{POC} = 2.7 Gt C yr^{-1}, F_{PIC} = 0.48 Gt C yr^{-1} and F_{Si} = 1.1 Gt Si yr^{-1}, c.f. ranges in Dunne et al. (2007), F_{POC} = 4.0–11 Gt C yr^{-1}, F_{PIC} = 0.38–1.8 Gt C yr^{-1} and F_{Si} = 1.9–5.0 Gt Si yr^{-1}). I do not expect SLAMS-2.0 to reproduce published magnitudes of export fluxes of POC, PIC and bSi because, in the current version of the model, particles are only transported vertically due to gravity; SLAMS-2.0 does not have ocean circulation embedded yet. Coupling SLAMS-2.0 to an OGCM, which would move particles in the 3-D space, is a necessary step to be able to claim that the simulated fluxes of materials are more or less correct. Notice that those low fluxes of POC at export depth modelled by SLAMS-2.0 are responsible for the low values of b and $\mathrm{T_{eff}}$ when compared to other studies (Fig. 6.9).

Despite the lower modelled globally integrated POC flux, its latitudinal pattern is consistent with that in the literature, being highest in the northern high-latitudes and upwelling regions, where there is abundant large phytoplankton (and high NPP) (Henson et al., 2012; DeVries and Weber, 2017; Buesseler et al., 2020b). Three reasons might explain why SLAMS-2.0 predicts a POC export flux lower than in the rest of the literature. First, is the choice of the reference depth used to integrate the flux globally (Buesseler et al., 2020b). For instance, in the polar oceans, where z_{eu} is shallower than the reference depth (100 m), the modelled POC export will underestimate *in situ* POC export, contributing to the lower globally integrated POC export fluxes. Second, is the choice of the NPP forcing used to run SLAMS-2.0, the Carr (2002) model. It is the same choice as that used by the study of Henson et al. (2012), with the difference that they use z_{eu} to integrate their export flux, obtaining a higher annually integrated flux of 4.0 Gt C yr^{-1}, and which is lower than that of DeVries and Weber (2017) (9.1 Gt C yr^{-1} at z_{eu}), probably due to the different choice of NPP forcing (VGPM and CbPM in the latter). Third, SLAMS-2.0 has been run on a coarse grid for reasons of computational economy; as a result, the model cannot resolve the high pulses of NPP observed in coastal upwelling regions, which lowers the globally integrated POC flux.

SLAMS-2.0 predicts a latitudinal pattern for bSi flux at export that completely misses the high input of bSi flux into the Southern Ocean (an area with abundant $SiOH_4$ ready to be uptaken by diatoms) and predicts quite a strong injection of bSi into the equatorial Pacific (an area more limited by $SiOH_4$ availability). The lower export fluxes of bSi in the Southern Ocean are due to the low diatom biomass modelled in that region (Fig. G.1), despite being the PFT exhibiting the largest relative abundance (Fig. 6.4). That disconnection between relative abundance and actual biomass is probably caused by

the use of an NPP forcing in SLAMS-2.0 (Carr (2002)'s model) that is poorly resolved for the Southern Ocean, not sourcing enough phytoplankton biomass in that area. Carr et al. (2006) noted that ocean-colour-based models struggle to recreate NPP in HNLC regions, like the Southern Ocean and the equatorial Pacific, as most satellite models were not developed with data subject to micronutrient limitation, like iron.

The modelled PIC flux at export depth is the most controversial of all three modelled material fluxes. Honjo et al. (2008) predict a basin-integrated PIC export flux higher for the Atlantic than the Pacific basin, whereas the opposite happens in SLAMS-2.0. On top of that disagreement, Jin et al. (2006) predict a zonal pattern for POC/PIC in the Pacific basin as: northern high-latitude > subtropic > equator (c.f. SLAMS-2.0, equator > northern high-latitude > subtropics). For the Atlantic basin, the zonal pattern predicted by Jin et al. (2006) is: poles > subtropics > equator (c.f. SLAMS-2.0, northern high-latitudes > equator > subtropics). The modelled PIC export pattern in SLAMS-2.0 overlaps with that of CO_3^{2-} ion availability forced into the model. Most probably, the way in which $CaCO_3$ (PIC) is produced in the model by coccolithophores taking up CO_3^{2-} could be curved by introducing a CO_3^{2-} uptake limitation factor (e.g., Gehlen et al., 2007, Eq. 3). Currently, the uptake of CO_3^{2-} by calcifying organisms is not limited by the saturation state of ambient water, in the same way other growth factors are (nitrate and silicic acid uptake). This way, the growth of coccolithophores could hopefully be restricted into the equatorial and subtropical ocean, following Jin et al. (2006).

6.6 Conclusions

I find that the global pattern for the attenuation rate of POC flux with depth (Martin's b) is high in the oligotrophic ocean and low in the productive ecosystems. The latitudinal variability in the Lagrangian-derived b manifests an explicit coupling with NPP and diatom relative abundance and decoupling with temperature. In productive systems, there is a diatom-driven export production that penetrates through the mesopelagic ocean thanks to the characteristically higher density of particles ballasted by opal. Indeed, opal-rich diatoms have associated higher transfer efficiencies (T_{eff}) or, correspondingly, low b. Mesopelagic processes that modify the exported flux in the mesopelagic ocean, like coagulation and grazing by zooplankton, spread the opal embedded in diatom frustules across the broader pool of particles. That fact supports the idea that surface water conditions (NPP and phytoplankton community composition), as well as upper mesopelagic ecosystem transformations, affect the BCP-mediated sequestration of POC in the deep ocean.

7 Summary

Contents

7.1 A new methodology to understand the transfer of POC flux in the ocean

Linking marine particle attributes of composition and structure to the global patterns of the transfer of POC flux from the surface to the deep ocean is a research priority in Ocean Biogeochemistry. Current understanding of the mechanisms that control global patterns of the transfer of POC flux through the mesopelagic are mainly based on (a) the observation of particle attributes in deep-sea sediment traps, which are scarce for the surface and mesopelagic ocean; (b) the global application of a fixed parameterisation for the the attenuation rate of POC flux with depth (Martin's b), when it is known to be dependent on local forcing, and (c) the combination of flux estimates from various methods (radionuclides, remineralised nutrient budgets, remote sensing, sediment traps) to build empirical relationships connecting surface ocean properties and deep sea fluxes, without direct observations of flux in between these two water sections (the mesopelagic). The mesopelagic ocean remains undersampled due to the technical challenges of maintaining fixed-point sediment traps in those waters, the high economic cost of research expeditions and the still emergent technology of free-floating

sediment traps. Thus, to date, oceanographers are still speculating about how the surface ocean system and the deep sea are connected in terms of transport of organic carbon in sinking particles. The large unknowns have created significant uncertainties as to which biogeochemical controls mostly affect the spatiotemporal variation of POC flux with depth (e.g., temperature, oxygen, phytoplankton community composition, NPP zooplankton repackaging). A turning point in this research field has been the conflicting patterns for the latitudinal variation of Martin's b found by the empirical studies of Henson et al. (2012) and Marsay et al. (2015). It is speculated that those two patterns are influenced by the spatial undersampling of the ocean and the methodology used to estimate b, using data from two different sections of the water column (deep vs surface ocean, respectively). Alongside, various data assimilation biogeochemical models have been run to obtain a global view of POC flux and understand the mechanisms that control its global transfer pattern through the mesopelagic.

This thesis develops a probabilistic model of Lagrangian marine particles that biogeochemically interact with their ecosystem and gravitationally sink through a 2-D water column. The model is called the model SLAMS-2.0 and the particle processes represented in it are: particle creation through a probabilistic model of PFT seeding, particle-to-particle interactions using Jackson (2001)'s model, particle-to-zooplankton interactions using Kiørboe (1997)'s model, a mechanistic model for the egestion of faecal pellets, bSi and $CaCO_3$ dissolution kinetics, heterotrophic metabolism following Q_{10} kinetics and O_2 limitation, and an empirical approach for particle disaggregation. The BCP has been streamlined and it has dependencies on common forcing variables (PAR_0, NPP, chl a, T, O_2, NO_3^-, CO_3^{2-}, $SiOH_4$, F_{dust}, MLD) and has an explicit representation of phytoplankton and zooplankton functional types. Equations are solved algebraically and hugely stochastic processes have a probabilistic interpretation. What makes SLAMS-2.0 distinct from other modelling approaches (mostly Eulerian) to study the transfer of POC flux to the deep ocean is that the metrics of the BCP (e.g., b and T_{eff}) emerge from an ensemble of Lagrangian marine particles, rather than averaged concentrations of particulate tracers. In SLAMS-2.0, fractal particles that follow Stokes' law interact, and their material is mixed according to the laws of coagulation and zooplankton encounters. Their composition and structure are transformed from first principles rather than from imposing a parameterisation on the fraction of mineral material they should contain or the size category they should belong to.

SLAMS-2.0 has been locally calibrated to improve its representation of local fluxes of POC, PIC and bSi. The five ocean sites chosen for calibration (EqPac, HAUSGARTEN, OSP, PAP-SO and BATS/OFP) were selected based on the availability of flux data for the surface and mesopelagic ocean and relatively good coverage of the annual cycle. Flux

observations are a limiting factor to constraining the output of a Lagrangian model of marine particles. A model like SLAMS-2.0 would greatly benefit from a larger number of long-term time-series sites, with especial emphasis on covering the bottom of the euphotic layer and the mesopelagic, two depth horizons largely undersampled. Of particular importance would be pairing POC flux observations (needed for a bottom-up calibration of SLAMS-2.0) with *in situ* NPP measurements (needed for a top-down calibration). It is also worth highlighting that the choice of the optimiser affects the success of the calibration. In SLAMS-2.0, I have adopted a data assimilation approach, where a derivative-free optimiser that uses least-squares (DFO-LS) gets trapped in local minima, thus offering multiple solutions for the calibration of SLAMS-2.0. I propose replacing DFO-LS by a Bayesian approach in future calibration exercises of SLAMS-2.0. Bayesian approaches are probabilistic and can efficiently infer solutions from noisy observations without using expensive integration methods.

7.2 Findings of the research

After calibration, SLAMS-2.0 has been applied globally to understand the mechanisms that control the global pattern of Martin's b, which characterises the attenuation rate of the sinking flux of POC in the mesopelagic. Empirical algorithms driven by satellite ocean colour and particle flux data suggest that a low attenuation rate of POC flux with depth (low b or, conversely, high T_{eff}) is expected in low and mid latitudes (Buesseler and Boyd, 2009; Guidi et al., 2009; Lam et al., 2011; Henson et al., 2012; Lima et al., 2014; Guidi et al., 2015). Nevertheless, SLAMS-2.0 predicts an opposite pattern, with low b or, conversely, high T_{eff}, in high latitudes and upwelling systems, a pattern that follows other data-driven models (Weber et al., 2016; DeVries and Weber, 2017; Cram et al., 2018) as well as empirical algorithms based on shallow sediment trap data (Marsay et al., 2015).

In SLAMS-2.0, there is a particle density and ballast control through opal on the transfer of POC to the deep ocean. High-latitude and upwelling regions are productive, or high NPP, ecosystems where the phytoplankton community is naturally dominated by diatoms. A variety of dynamic BCP processes, such as particle coagulation and zooplankton grazing, which are more intense in those productive ecosystems as a consequence of the higher abundance of particles, mix well the assemblage of phytoplankton cells. That spreads the opal initially associated with diatoms to the broader community of particles, resulting in the formation of opal-rich, dense particles. Those opal-rich particles are ballasted to the deeper ocean, generating a signal of high T_{eff} or, conversely, low b. In ocean regions where the phytoplankton assemblage is dominated by small picophytoplankton (oligotrophic regions), opal production is minimal, and thus T_{eff} is low, and b is high. Thus, the relative

importance of opal controlling the mesopelagic transfer of POC flux needs to be framed in the context of its local production (accompanied by high inputs of POC) and mesopelagic processes that spread it (coagulation and zooplankton grazing).

The modelled results convey the following key messages.

1. High T_{eff} or, conversely, low b, are predicted for areas with higher NPP and diatom abundance in the euphotic zone.
2. Those diatom-rich phytoplankton communities are associated with opal ballast properties.
3. Particles loaded with opal are denser and sink faster, but are not necessarily bigger as the increase in their density due to the accumulation of ballast materials overcomes the increase in porosity (or fluffiness) that comes with the addition of more primary particles following fractal scaling.
4. The downward flux of particles is affected by the particle dynamics of coagulation and zooplankton grazing, which spread opal components amongst the broader pool of sinking particles.
5. The intensity of particle dynamics are linked to the structure of the phytoplankton community in the euphotic zone, with more grazing and coagulation occurring in areas where more particles are produced (i.e., high NPP areas).

Thus, phytoplankton community structure influences the global patterns of b and T_{eff} through (i) a mechanistic link to the biogeochemical control of phytoplankton biomass, or NPP, with more productive areas producing a higher load of export particles, (ii) a mechanical control on particle density and ballast exerted by the opal that is sourced by diatoms, and (iii) its coupling to ecosystem functioning processes of grazing and coagulation, which spread the opal amongst the broader pool interacting particles. These results highlight phytoplankton community structure and particle dynamics of coagulation and zooplankton grazing as proxies to understand the transfer of POC from the surface to the deep ocean and deviate the attention from the overly cited temperature. SLAMS-2.0 does not respond to temperature changes in the ecosystem, which is the most distinctive finding of this study. The temperature control found in e.g., Henson et al. (2012), Marsay et al. (2015), Weber et al. (2016) or DeVries and Weber (2017) is probably an artefact of temperature being a proxy for Eulerian variables used to force those models (NPP, O_2, PAR_0). SLAMS-2.0 gives us a hint that Lagrangian-derived b or T_{eff} do not respond directly to changes in temperature if those do not affect first the composition of the phyto- and zooplankton system that generates and reprocesses the particles. Further tests are needed to understand this in more detail, with a sensitivity analysis that separates the effects of temperature on (a) physical properties (coagulation rates and viscosity-mediated

sinking velocities) and (b) biological properties (phytoplankton growth, heterotrophic metabolism and zooplankton DVM). The absence of a b-temperature correlation in SLAMS-2.0 highlights the message of the important role of complex ecological relationships over that of simple correlations with ecosystem forcing (Henson et al., 2012; Marsay et al., 2015) on our way to understanding the controls that affect particle flux attenuation.

Simulating billions of Lagrangian particles is an expensive task to be performed in Earth system models (ESMs). Thus, a key aim of a Lagrangian model of marine particles should be deriving a mechanistic parameterisation to inform of the relationship between ecological controls and particulate fluxes that could be easily implemented in ESMs to constrain particle fluxes and improve the budgets of the BCP-mediated sequestration of carbon in the deep ocean. Thus, in this thesis, I have come up with a parameterisation (Eq. 6.1) that summarises the ecological effects of NPP and diatom relative abundance on Martin's b, both of which enhance most notoriously the transfer efficiency of POC. The equation still needs to be tried and tested. However, it opens the door to Lagrangian simulations of marine particles becoming testbed models for the next generation of equations on the attenuation rate of POC flux with depth, leaving behind the overarching paradigm of the power-law Martin curve.

The present study is the first to model from first principles (fractal scaling law and Stokes' velocity) individual marine particle attributes of structure, chemical composition and sinking velocity on a global scale across all depths, with the accompanying particle dynamics of particle-particle interaction and zooplankton-particle encounters. That allows SLAMS-2.0 to unravel the particle attribute and dynamic mechanisms that control the global patterns of the BCP-mediated sequestration of carbon in the deep ocean. The results help connecting surface-water particle sources to mesopelagic fluxes.

7.3 Outlook

As demonstrated, a probabilistic model of Lagrangian marine particles like SLAMS-2.0 has shown great promise in reproducing the global patterns of b and T_{eff} and providing a mechanistic explanation for the factors controlling those patterns from a bottom-up approach. On top of that skill, SLAMS-2.0 has low computational costs due to an efficient design and relatively low complexity of the ecosystem model. It also offers a modular design, where it is easy to implement new processes and study their implications on the flux of particles in an isolated manner. Nonetheless, SLAMS-2.0 has limitations: (i) ocean circulation has not been included yet, which conditions the magnitude and spatial variability of particle fluxes due to the exclusion of physical effects like advection and diffusion, lateral transport, stratification and mixed layer depth variations, some

of which are currently parameterised; (ii) the particle coagulation algorithm of Jackson (2001) was developed for non-porous spheres (beads) and, thus, struggles to reproduce the high coagulation rates expected for porous marine particles; (iii) the approach to simulate particle disaggregation has not been tested in other models, and it is an empirical approximation to particle fragility using the fraction of mineral content in a particle; (iv) biogenic calcification has not been made dependent on saturation state of ambient water to curve the growth of coccolithophores and thus constrain PIC fluxes in the model, which at present have little latitudinal variation; (v) there is neither a self-regulated, numerical phytoplankton model embedded in SLAMS-2.0 nor a zooplankton numerical model –those two components are currently prescribed through stochastic sampling, numerically modelling them will require the current model forcing to become state variables of the model first.

However, introducing that additional complexity is beyond the scope of this thesis. In future applications of SLAMS-2.0, all of the processes mentioned above will be considered, in conjunction with a model calibration that will include observed particle size distributions (PSDs, obtained from the EcoTaxa repository) and will expand the data set of POC, PIC and bSi fluxes to other locations. A calibration analysis that considers PSDs may help to better constrain the contributions of small vs large particles in the transfer of POC to the ocean interior and, thus, assist in validating simulated particle dynamics. I also believe that a regionalisation of the patterns of b and T_{eff} according to biogeochemical provinces (instead of considering them on a grid point basis) might offer neater global distributions. I expect to run SLAMS-2.0 on more fine-scale global grids to reproduce the high NPP pulses in coastal areas, which are currently underrepresented in SLAMS-2.0 due to the more coarse grid used. A final consideration would be integrating fluxes and obtaining the metrics of the BCP on a seasonal basis instead of averaging those for the annual cycle, where the particle pulses of the bloom periods disappear.

In the results presented in this thesis, due to the need to prioritise the most relevant findings, I have not explored the relative contributions of faecal pellets, marine snow and phytodetritus to total particulate flux, nor the average characteristics of such particles. However, all of that information is currently saved in the model output files, along with particle size distributions and their slopes and the particle number distribution for any of the particle attributes modelled. A more in-depth analysis is needed for the zooplankton grazing flow rates presented in the results, which will benefit from examining the characteristics of the faecal pellets produced. Nevertheless, zooplankton flow rates face the problem of insufficiently depth-resolved observations of zooplankton activity (e.g., MAREDAT).

The findings of SLAMS-2.0 also have implications for the future of the oceanic organic carbon sequestration in the light of a warming climate. Because of climate change, predicted ocean warming, increase in stratification, and decrease in nutrient supply are expected to favour picophytoplankton growth. Thus, oligotrophic biomes could be in expansion and diatom-rich biomes in contraction. Oligotrophic regimes have characteristically less efficient particle export out of the surface ocean and lower transfer efficiencies to the deep ocean. Consequently, if the BCP system behaves in a linear way, a warming climate could see a reduction in the oceanic sequestration of organic carbon. That, in turn, could have positive feedbacks on climate change, amplifying atmospheric warming. Feedbacks between increased ocean temperatures, reduced global organic carbon sequestration and hence higher atmospheric CO_2 levels leading to higher atmospheric temperatures could further reduce future ocean CO_2 uptake. Predicting and quantifying climate feedbacks will require improving the mechanisation of Lagrangian models of marine particles, which in turn will require:

- improving the formulation of coagulation kernels for porous particles since particle coagulation has extensive implications in the sinking flux of particles;
- improving existing algorithms of satellite estimates of depth-integrated NPP, which are used to force the model and to which POC fluxes are sensitive to, making the choice of the NPP algorithm a source of uncertainty in the model output;
- promoting the collection of particle flux data (POC, PIC and bSi fluxes) over sufficiently long timescales, at relevant depths in the water column (export depth and sequestration depth most importantly), and at a high enough spatiotemporal resolution, a task needed to calibrate particle models;
- encouraging the collection of *in situ* data of zooplankton activity and NPP alongside observations of particle flux data; and
- urging the ocean biogeochemical community to come to terms with the statistical methods to compute the attenuation rate of POC flux with depth (Martin's b), agreeing at least in reference sampling depths or, alternatively, adopt a new metric used as the standard, like T_{eff}, which is less sensitive to noisy POC flux depth profiles.

Appendices

A

Derivation of the Martin model

The Martin model was first derived in 1987 by John Martin and colleagues as an empirical fit of POC flux sediment trap data to depth (Martin et al., 1987). However, it is also possible to mathematically derive the Martin model from the advection-reaction-diffusion equation using some assumptions. That process gives us a mechanistic interpretation for Martin's b.

A.1 The advection-diffusion-reaction model

The advection-diffusion-reaction equation of a tracer, also known as the tracer conservation equation, has this generic shape:

$$\frac{\partial \mathrm{C}}{\partial t} = \left.\frac{\partial \mathrm{C}}{\partial t}\right|_{\text{advection}} + \left.\frac{\partial \mathrm{C}}{\partial t}\right|_{\text{diffusion}} + \mathrm{SMS(C)}, \tag{A.1}$$

where the time evolution of the concentration of a tracer, C, is regulated by the balance between the transport of the tracer due to the physical process of water motion (advection and diffusion) and the biogeochemical processes that become internal sources of C and sinks that degrade C (sources-minus-sinks, SMS).

Suppose that the tracer in question is POC in a motionless water column, with no mixing nor horizontal transport. In that case, the transport terms are equal to a sinking term (i.e., the downwards differential of POC keeping the time constant is $\frac{\mathrm{D\,POC}}{\mathrm{D}\,t} = \cancel{\frac{\partial \mathrm{POC}}{\partial t}} + w\frac{\partial \mathrm{POC}}{\partial z}$) and the SMS term is equal to the biogeochemical process of remineralisation, like so:

$$\frac{\partial \mathrm{POC}}{\partial t} = -w\frac{\partial \mathrm{POC}}{\partial z} - \lambda\ \mathrm{POC}, \tag{A.2}$$

where w is the sinking velocity of marine particles, λ is the remineralisation rate of POC and POC is a concentration (e.g., mmol POC m^{-3}). The two most common assumptions used in the literature to solve Eq. A.2 are:

- w is constant throughout depth, and it has as a solution an exponential model, and
- w increases with depth, and it has as a solution a power-law model (the Martin model).

The exponential model makes sense when one considers non-aggregated, primary particles. Field observations of non-aggregated particles show constant particle sinking velocities, regardless of the depth. Nevertheless, when one considers aggregated particles and their diversity in properties, such as size or density, an increase of sinking velocity with depth is observed. Next, I will derive both the exponential and the power-law models for POC flux attenuation with depth. It will become evident that the mathematical assumptions used to derive those models are their mechanistic interpretations.

A.2 The exponential model

The exponential model assumes constant POC sinking velocity throughout depth (w = cnt). The starting point is Eq. A.2 and at equilibrium, or steady state conditions, $\frac{\partial \mathrm{POC}}{\partial t} = 0$, and so Eq. A.2 can be written as

$$w\frac{\partial \mathrm{POC}}{\partial z} = -\lambda\ \mathrm{POC}. \tag{A.3}$$

Reorganising and integrating Eq. A.3 to get rid of the differentials,

$$\begin{aligned}
&\frac{\partial \mathrm{POC}}{\mathrm{POC}} = -\frac{\lambda}{w}\partial z \\
&\rightarrow \int_{z_0}^{z_1} \frac{1}{\mathrm{POC}}\partial \mathrm{POC} = -\frac{\lambda}{w}\int_{z_0}^{z_1} \partial z \\
&\rightarrow \ln(\mathrm{POC}) + C_1 = -\frac{\lambda}{w}\ z\Big|_{z_0}^{z_1} + C_2 \\
&\rightarrow \ln(\mathrm{POC}) = -\frac{\lambda}{w}z + (C_2 - C_1) = -\frac{\lambda}{w}z + C_3 \\
&\rightarrow e^{\ln(\mathrm{POC})} = e^{-\frac{\lambda}{w}z + C_3} = e^{-\frac{\lambda}{w}z}\ e^{C_3} \\
&\rightarrow \mathrm{POC} = e^{-\frac{\lambda}{w}z}\ e^{C_3};\ \text{at } z = 0,\ \mathrm{POC} = e^0\ e^{C_3};\ \mathrm{POC}(z_0) = e^{C_3} \\
&\rightarrow \mathrm{POC}(z) = e^{-\frac{\lambda}{w}(z - z_0)}\ \mathrm{POC}(z_0).
\end{aligned} \tag{A.4}$$

where the depth integration limits are the base of the euphotic layer (z_0, usually ~100 m) and any depth that is below z_0 (z_1). Likewise, POC flux, F (mmol POC $m^{-2}\,d^{-1}$), is given by

$$F(z) = F(z_0)\ e^{-\frac{\lambda}{w}(z - z_0)} = F(z_0)\ e^{-\frac{(z - z_0)}{z^*}}, \tag{A.5}$$

Eq. A.5 is the exponential model, where z^* is the remineralisation length scale, or w/λ.

A.3 The power-law model

The Martin, power-law model assumes that the particle sinking velocity increases linearly with depth, i.e., $w(z) = mz + n$ ($n = 0$). Solving Eq. A.2 with that assumption,

$$\begin{aligned}
\frac{\partial^2 \mathrm{POC}}{\partial t^2} &= \frac{\partial (w(z)\ \mathrm{POC})}{\partial z} - \lambda\ \mathrm{POC} \\
&\overset{(1)}{=} -\left(w(z) \frac{\partial \mathrm{POC}}{\partial z} + \mathrm{POC} \frac{\partial w(z)}{\partial z} \right) - \lambda\ \mathrm{POC} \\
&= -\left(mz \frac{\partial \mathrm{POC}}{\partial z} + \mathrm{POC} \frac{\partial (mz)}{\partial z} \right) - \lambda\ \mathrm{POC} \\
&\overset{(2)}{=} -\left(mz \frac{\partial \mathrm{POC}}{\partial z} + m\ \mathrm{POC} \right) - \lambda\ \mathrm{POC} \\
&= -mz \frac{\partial \mathrm{POC}}{\partial z} - (m + \lambda)\ \mathrm{POC}.
\end{aligned} \tag{A.6}$$

(1) chain rule

(2) since m does not depend on z, $\frac{\cancel{\partial (mz)}}{\cancel{\partial z}}$

At equilibrium, $\frac{\partial \mathrm{POC}}{\partial t} = 0$, and so reorganising and integrating Eq. A.6,

$$\begin{aligned}
& mz\ \frac{\partial \mathrm{POC}}{\partial z} = -(m + \lambda)\ \mathrm{POC} \\
& \rightarrow \frac{\partial \mathrm{POC}}{\mathrm{POC}} = -\frac{(m + \lambda)}{mz} \partial z \\
& \rightarrow \int_{z_0}^{z_1} \frac{1}{\mathrm{POC}} \partial \mathrm{POC} = -\frac{(m + \lambda)}{m} \int_{z_0}^{z_1} \frac{1}{z} \partial z \\
& \rightarrow \ln(\mathrm{POC}) + C_1 = -\left(1 + \frac{\lambda}{m}\right) \ln z \Big|_{z_0}^{z_1} + C_2 \\
& \rightarrow \ln(\mathrm{POC}) = -\left(1 + \frac{\lambda}{m}\right) (\ln z_1 - \ln z_0) + (C_2 - C_1) \\
& \rightarrow \ln(\mathrm{POC}) = C_3 - \left(1 + \frac{\lambda}{m}\right) \ln\left(\frac{z_1}{z_0}\right).
\end{aligned} \tag{A.7}$$

Eq. A.7 has the familiar shape $\ln y = \ln m + n \ln x \leftrightarrow y = mx^n$. Furthermore, after solving Eq. A.4 we know that $e^{C_3} = \mathrm{POC}(z_0)$. Thus,

$$\mathrm{POC}(z) = \mathrm{POC}(z_0)\ \left(\frac{z}{z_0}\right)^{-\left(1 + \frac{\lambda}{m}\right)}. \tag{A.8}$$

Likewise, POC flux, F (mmol POC $\mathrm{m}^{-2}\,\mathrm{d}^{-1}$), is given by

$$F(z) = F(z_0)\ \left(\frac{z}{z_0}\right)^{-\left(1 + \frac{\lambda}{a}\right)}. \tag{A.9}$$

Eq. A.9 is the Martin model, also known as the Martin curve. In the literature, the exponent $1 + (\lambda/m)$ is referred to as Martin's b, where m is the slope of the function that predicts particle sinking velocity. Note that m, and thus w, is anticorrelated with b.

B
Code availability

SLAMS-2.0 code, available at https://doi.org/10.5281/zenodo.5570253, is written in Fortran90, a compiled programming language. The code is fully documented with in-code comments and for a detailed description of the algorithms, refer to Chapter 3 of this dissertation.

The code is modular and it is made up of 24 Fortran90 source files (ended with the extension `.F90`) and two header files (ended with the extension `.h`). It comes with two supporting files necessary to write information to output binary (`.bin`) files and read from those: `io.F` and `FNread_binary.m`. Binary, unformatted files take less disk space than ASCII (formatted) files. The code is compiled and run in any machine that has installed gfortran or ifort compilers using the `Makefile` provided alongside. Post-processing scripts (`readSlamsOutputBCP.m` and accompanying functions) and scripts used to create model forcing files (`createSlamsInputDataAndDirectories.m` and accompanying functions) are available in the same URL above.

B.1 Experiment configuration

Each SLAMS-2.0 experiment must be run in a separate directory with (i) its own `namelist` file, which includes various model parameters and configuration options, (ii) 10 forcing `.bin` files, which contain the variables that support ecosystem functioning at each time step (PAR_0, NPP, [chl a], F_{dust}, $[NO_3^-]$, $[SiOH_4]$, $[CO_3^{2-}]$, $[O_2]$, T and MLD), and (iii) soft links to the code files. I use Matlab to create the `.bin` forcing files from `.mat` data arrays. SLAMS-2.0 can be run in one of two modes: (a) on a single water column or (b) on a global grid (an ensemble of water columns). Forcing files must be created

accordingly using the location coordinates (latitude and longitude) and seafloor depth in which the user wants to run the model.

Amongst the source files, `main.F90` is the master file. It contains the program and calls to all functions and subroutines that make up SLAMS-2.0. As a user, you should open main.F90 and read the comments at the preamble (preceded by "!"), which contain information on the structure of output files as well as input files needed to run the model. The header file `blockdefinitions.h` controls big chunks of code by activating them or deactivating them. If you want to run SLAMS2.0 in a lighter mode, keep undefined `BLOCK_PRINT_INFO_TO_THE_SCREEN` and `BLOCK_PRINT_INFO_TO_THE_SCREEN_EXTENDED`, and defined `SAVE_DISK_SPACE`.

B.2 Model output files

SLAMS-2.0 produces `.bin` model output files every modelled month for every depth layer and are listed in Table B.1. Raw model output is post-processed to calculate particle and BCP derived properties, classified into the following six types.

- Individual particle attributes, $\mathbf{a}_k$ (k is for the k$^{\text{th}}$ cluster). This type of information is produced as a model final snapshot of all the particles that are in the simulation, which are collected into an array that presents their 18 attributes (16 physicochemical, 1 depth position and 1 particle multiplicity factor in the cluster).
- Average particle attributes by depth z and month t, $\mathbf{a}_k(z,t)$, which produce a picture of a typical particle in terms of its physicochemical attributes. Those average attributes can also be optionally classified according to two main particle size categories (small and large, with 150 μm marking the threshold, Giering et al. (2020), two main particle velocity categories (slow and fast, with 100 m d^{-1} marking the threshold, McDonnell and Buesseler (2010)) and eight particle typologies (living phytoplankton cell, dead phytoplankton cell, TEP, clay particle, dead zooplankton body, faecal particle, non-faecal aggregate and non-faecal single particle).
- Integrated particle attributes by z and t: particle number distribution, $N(\mathbf{a}^i,z,t)$, expressed as a function i of particle size classes and particle velocity classes; the slope of the particle number distribution, ξ; fluxes of POC, PIC and bSi, $F(z,t)$; the slope of the vertical attenuation of POC flux with depth, Martin's b; particle export efficiency, PE_{eff}; transfer efficiency, T_{eff}; the ratio of organic to inorganic carbon, POC/PIC; and the ratio of biogenic silica to inorganic carbon, bSi/PIC.
- Volumetric fluxes of material (C, Si, CO_3^{2-}) from one ecosystem component to another, i.e.,the sources-minus-sinks terms.

- Biomass of phytoplankton and zooplankton, abundance of zooplankton individuals and average carbon cell quota of a phytoplankton cell.
- Average rates of remineralisation, particle collision and zooplankton-particle encounter.

Table B.1. Description of SLAMS-2.0 output files. All files are binary (.bin). Abbreviations: *descr*, description; *z/t*, space/time step of recording; *dim*, array dimensions.

File name	Characteristics
`out_cluster_i`	**descr** collects in an array the particle attributes calculated as integer values of all modelled particles **z/t** recorded three times in the final run year **dim** `maxNumClusters` × 6 × 3
`out_cluster_r`	**descr** collects in an array the particle attributes calculated as real values of all modelled particles **z/t** recorded three times in the final run year **dim** `maxNumClusters` × 20 × 3
`out_stats_nParticles`	**descr** records the number of extant particles **z/t** for every month of every modelled year (`nTimesStats`) **dim** `nTimesStats` × 1
`out_stats_nClusters`	**descr** records the number of extant particle clusters **z/t** for every month of every modelled year **dim** `nTimesStats` × 1
`out_control`	**descr** records simulation state information: (1) no. Δt past, (2) last position occupied in the particle array, (3) no. times monthly data have been recorded, `nTimesFlux`, (4) no. sediment trap deployment depths, `nSedTrapDeployDepths`, (5) no. imaging system deployment depths, `nImagingDeployDepths`, (6) no. times average particle data have been recorded, `nTimesAvgAtt`, (7) no. times stats. data have been recorded, `nTimesStats`, (8) no. times aux. and SMS data have been recorded, `nTimesAux` **z/t** at the end of the simulation **dim** 8 × 1
`out_flux`	**descr** for the water column, records the fluxes of six particulate materials (POC, TEC, bSi, calcite, aragonite, clay) in units of flow rate per unit area ($\mathrm{mg\,m^{-2}\,d^{-1}}$) **z/t** for all particles crossing every depth where the virtual sediment trap is deployed, for every month of every modelled year **dim** 6 × `nSedTrapDeployDepths` × `nTimesFlux`
`out_flux_seafloor`	**descr** for the seafloor, *does as before* **z/t** for all particles hitting the seafloor, for every month of every modelled year **dim** 6 × 1 × `nTimesFlux`

Table B.1 (Continued) Description of SLAMS-2.0 output files.

File name	Characteristics
`out_avgparticleatt_sizeclass`	**descr** for the water column particles, and according to 16 particle size classes (`nSizeClasses`), averages 15 particle attributes: (1) no. particles, (2) density ($g\,cm^{-3}$), (3) sinking velocity ($m\,d^{-1}$), (4) stickiness, (5) porosity, (6) solid volume (μm^3), (7) fractal dimension, (8) radius primary particle, µm, (9) depth (m), (10) mol C_{org}, (11) mol TEC, (12) mol bSi, (13) mol calcite, (14) mol aragonite, (15) mol clay **z/t** for all particles crossing every depth where the virtual imaging system is deployed, for every month of every modelled year **dim** `nSizeClasses` × `nImagingDeployDepths` × 15 × `nTimesAvgAtt`
`out_avgparticleatt_sizeclass_seafloor`	**descr** for all particles hitting the seafloor, *does as before* **z/t** for all particles hitting the seafloor, for every month of every modelled year **dim** `nSizeClasses` × 1 × 15 × `nTimesAvgAtt`
`out_avgparticleatt_veloclass`	**descr** for the water column particles, and according to 11 particle velocity classes (`nVeloClasses`), averages 15 particle attributes **z/t** for all particles crossing every depth where the virtual imaging system is deployed, for every month of every modelled year **dim** `nVeloClasses` × `nImagingDeployDepths` × 15 × `nTimesAvgAtt`
`out_avgparticleatt_veloclass_seafloor`	**descr** for all particles hitting the seafloor, *does as before* **z/t** for all particles hitting the seafloor, for every month of every modelled year **dim** `nVeloClasses` × 1 × 15 × `nTimesAvgAtt`
`out_avgparticleatt_maintype`	**descr** for the water column particles, calculates four average attributes (radius, density, settling velocity and porosity) of a typical particle belonging to one of the following eight types (`nMainParticleTypes`): (1) faecal particle, (2) non-faecal aggregate, (3) living phytoplankton cell, (4) dead phytoplankton cell, (5) TEP, (6) clay particle, (7) dead zooplankton body, (8) non-faecal single particle **z/t** for all particles crossing every depth where the virtual imaging system is deployed, for every month of every modelled year **dim** `nMainParticleTypes` × `nImagingDeployDepths` × 4 × `nTimesAvgAtt`
`out_avgparticleatt_maintype_seafloor`	**descr** for all particles hitting the seafloor, *does as before* **z/t** for all particles hitting the seafloor, for every month of every modelled year **dim** `nMainParticleTypes` × 1 × 4 × `nTimesAvgAtt`

Table B.1 (Continued) Description of SLAMS-2.0 output files.

File name	Characteristics
`out_sms`	**descr** records the model sources-minus-sinks (SMS) terms (`maxNumSmsTerms`) in units of flow rate per unit volume ($\mathrm{mol\,m^{-3}\,d^{-1}}$): (1) C taken up by phytoplankton, (2) CO_3^{2-} taken up by coccolithophores, (3) $SiOH_4$ taken up by diatoms, (4) aeolian dust deposition, (5) TEC released by phytoplankton, (6) TEC released by bacteria, (7) DOC released due to photolysis, (8) dissolved TEC released due to photolysis, (9) CO_3^{2-} released due to photolysis, (10) $SiOH_4$ released due to photolysis, (11) dissolved clay released due to photolysis, (12) DOC released due to zooplankton messy feeding, (13) dissolved TEC released due to zooplankton messy feeding, (14) CO_3^{2-} released due to zooplankton messy feeding, (15) $SiOH_4$ released due to zooplankton messy feeding, (16) dissolved clay released due to zooplankton messy feeding, (17) POC ingested by zooplankton, (18) TEC ingested by zooplankton, (19) CO_2 released due to zooplankton respiration of POC, (20) CO_2 released due to zooplankton respiration of TEC, (21) POC egested by zooplankton, (22) TEC egested by zooplankton, (23) CO_3^{2-} released due to biotic dissolution of $CaCO_3$, (24) DOC released due to zooplankton excretion, (25) dissolved TEC released due to zooplankton excretion, (26) CO_2 released by mesozooplankton respiration, (27) CO_2 released by microzooplankton respiration, (28) POC egested by mesozooplankton, (29) POC egested by microzooplankton, (30) POC released as mesozooplankton dead bodies, (31) POC released as microzooplankton dead bodies, (32) C associated to secondary production of foraminifera, (33) C associated to secondary production of pteropods, (34) C associated to secondary production of radiolaria, (35) C associated to secondary production of microzooplankton, (36) C associated to secondary production of mesozooplankton (37) POC released as zooplankton dead bodies, (38) $CaCO_3$ released as zooplankton dead bodies, (39) bSi released as zooplankton dead bodies, (40) CO_2 released due to bacteria respiration of POC, (41) CO_2 released due to bacteria respiration of TEC, (42) DOC released due to bacteria solubilisation, (43) dissolved TEC released due to bacteria solubilisation, (44) CO_3^{2-} released due to abiotic solubilisation, (45) $SiOH_4$ released due to abiotic solubilisation, (46) dissolved clay released due to abiotic solubilisation, (47) CO_3^{2-} released due to abiotic dissolution of $CaCO_3$, (48) $SiOH_4$ released due to abiotic dissolution of bSi **z/t** for every modelled depth layer (`nDepthLayers`), for every month of every modelled year **dim** `maxNumSmsTerms` × `nDepthLayers` × `nTimesAux`
`out_sms_annual`	**descr** record a water column average of the sources-minus-sinks (SMS) terms (`maxNumSmsTerms`) in units of flow rate per unit area, or flux ($\mathrm{mol\,m^{-2}\,yr^{-1}}$) **z/t** at the end of every modelled year (`nYears`) **dim** `maxNumSmsTerms` × `nYears`

Table B.1 (Continued) Description of SLAMS-2.0 output files.

File name	Characteristics
`out_aux`	**descr** records the model auxiliary terms (`maxNumAuxTerms`): (1) mesozooplankton biomass ($\mathrm{mol\,C\,m^{-2}}$), (2) microzooplankton biomass ($\mathrm{mol\,C\,m^{-2}}$), (3) no. mesozooplankton individuals ($\mathrm{num\,m^{-3}}$), (4) no. microzooplankton individuals ($\mathrm{num\,m^{-3}}$), (5) dead mesozooplankton biomass ($\mathrm{mol\,C\,m^{-3}}$), (6) dead microzooplankton biomass ($\mathrm{mol\,C\,m^{-3}}$), (7) diatom biomass ($\mathrm{mol\,C\,m^{-3}}$), (8) dinoflagellate biomass ($\mathrm{mol\,C\,m^{-3}}$), (9) coccolithophore biomass ($\mathrm{mol\,C\,m^{-3}}$), (10) picophytoplankton biomass ($\mathrm{mol\,C\,m^{-3}}$), (11) C cell quota of an average diatom ($\mathrm{mol\,C\,cell^{-1}}$), (12) C cell quota of an average dinoflagellate ($\mathrm{mol\,C\,cell^{-1}}$), (13) C cell quota of an average coccolithophore ($\mathrm{mol\,C\,cell^{-1}}$), (14) C cell quota of an average picophytoplankton ($\mathrm{mol\,C\,cell^{-1}}$), (15) average probability of having mesozooplankton, (16) average probability of having microzooplankton, (17) average zooplankton specific respiration rate ($\mathrm{s^{-1}}$), (18) average bacteria specific respiration rate ($\mathrm{s^{-1}}$), (19) average upper boundary of the DVM depth of zooplankton at night (m), (20) average lower boundary of the DVM depth of zooplankton at night (m), (21) average upper boundary of the DVM depth of zooplankton during the day (m), (22) average lower boundary of the DVM depth of zooplankton during the day (m), (23) average euphotic layer depth (m), (24) average particle collision kernel, ($\mathrm{m^3\,s^{-1}}$), (25) average zooplankton-particle encounter kernel ($\mathrm{m^3\,s^{-1}}$), (26) average particle coagulation success **z/t** for every modelled depth layer, for every month of every modelled year **dim** `maxNumAuxTerms` × `nDepthLayers` × `nTimesAux`

C

Additional tables and equations for Chapter 3

Table C.1. SLAMS-2.0 parameters. The values provided are mostly based on observations but also on values used in other models. *Primary production parameters* have been tuned to replicate the world distribution of phytoplankton of the Darwin Project (https://svs.gsfc.nasa.gov/30669). Phytoplankton cell quotas are based on a compilation of 32 references available at the url link provided. The following parameters have been optimised in a calibration analysis (see Chapter 5): $\theta^{\text{diat}}_{\text{Si:C}}$, $\theta^{\text{cocco}}_{\text{calc:C}}$, $f^{\text{ing, surf}}_{\text{OM}}$, $\theta_{\text{POC:Z}}$, $\theta^{\text{DVM night}}_{\text{POC:Z}}$, f_{sloppy} and κ_{opal}.

Symbol	Meaning	Value	Units	Literature values
Particle parameters				
D_3	3-D fractal dimension	2.0	–	1.3–2.3 (Burd and Jackson, 2009)
$P^{\text{max}}_{\text{fp}}$	Max. porosity of a faecal pellet	0.70	–	0.43–0.65 (Ploug et al., 2008)
$P^{\text{min}}_{\text{fp}}$	Min. porosity of a faecal pellet	0.40	–	*as above*
$\theta^{\text{max, diat}}_{\text{C}}$	Max. diatom cell C quota	1,000	pmol C cell^{-1}	0.49–6,494 (https://doi.org/10.5281/zenodo.5233652)
$\theta^{\text{min, diat}}_{\text{C}}$	Min. diatom cell C quota	8.0	pmol C cell^{-1}	*as above*
$\theta^{\text{max, dino}}_{\text{C}}$	Max. dinoflagellate cell C quota	10,000	pmol C cell^{-1}	3.2–1,900 (https://doi.org/10.5281/zenodo.5233652)
$\theta^{\text{min, dino}}_{\text{C}}$	Min. dinoflagellate cell C quota	400	pmol C cell^{-1}	*as above*
$\theta^{\text{max, cocco}}_{\text{C}}$	Max. coccolithophore cell C quota	8.0	pmol C cell^{-1}	0.35–35 (https://doi.org/10.5281/zenodo.5233652)
$\theta^{\text{min, cocco}}_{\text{C}}$	Min. coccolithophore cell C quota	2.0	pmol C cell^{-1}	*as above*
$\theta^{\text{max, pico}}_{\text{C}}$	Max. picophytoplankton cell C quota	1.0	pmol C cell^{-1}	0.0023–0.30 (https://doi.org/10.5281/zenodo.5233652)
$\theta^{\text{min, pico}}_{\text{C}}$	Min. picophytoplankton cell C quota	0.020	pmol C cell^{-1}	*as above*
$\theta^{\text{diat}}_{\text{Si:C}}$	Diatom cell Si:C ratio	0.30	mol Si (mol C)$^{-1}$	0.04–0.96 (https://doi.org/10.5281/zenodo.5233652)
$\theta^{\text{cocco}}_{\text{calc:C}}$	Coccolithophore cell PIC:POC ratio	0.90	mol PIC (mol POC)$^{-1}$	0.70–1.55 (https://doi.org/10.5281/zenodo.5233652)
$\theta^{\text{radiol}}_{\text{C:Si}}$	Radiolarian cell C:Si ratio	2.0	mol C (mol Si)$^{-1}$	*as per diatoms*
$\theta^{\text{foram}}_{\text{C:calc}}$	Foraminifera cell POC:PIC ratio	1.0	mol POC (mol PIC)$^{-1}$	*as per coccolithophores*
$\theta^{\text{pterop}}_{\text{C:arag}}$	Pteropod body POC:PIC ratio	1.0	mol POC (mol PIC)$^{-1}$	*as per coccolithophores*
$\theta^{\text{max, TEP}}_{\text{C}}$	Max. TEC quota	150	pmol C (mol TEP)$^{-1}$	Estimated combining Mari (1999)'s TEC ($0.25 \times 10^{-6} r^{2.55}$ μg C (TEP)$^{-1}$) and Passow (2002)'s TEP size (1–100 μm)
$\theta^{\text{min, TEP}}_{\text{C}}$	Min. transparent TEC quota	10	pmol C (mol TEP)$^{-1}$	*as above*

Table C.1 (Continued) SLAMS-2.0 parameters.

Symbol	Meaning	Value	Units	Literature values
θ_{clay}	Clay particle quota	0.10	pmol clay (particle)$^{-1}$	Value that makes possible a size for clay particles of 0.1–20 μm in diameter (Schulz et al., 2012)
$f_{\text{OM}}^{\text{ing, surf}}$	Min. fraction of OM in a surface particle ($\leq$ 1000 m) to be deemed appetitive	0.30	–	0.25 and 0.15 (Yool et al., 2013), 0.33 (Anderson et al., 2015), both detritus preferences
$f_{\text{OM}}^{\text{ing, deep}}$	Min. fraction of OM in a deep particle ($>$ 1000 m) to be deemed appetitive	0.30	–	*as above*
$\theta_{\text{chl:C}}$	chl:C ratio	0.025	g chl (g C)$^{-1}$	$<$ 0.01 to 0.050 (Arteaga et al., 2016), $<$ 0.01 to $>$ 0.1 (Geider et al., 1997)
$\theta_{\text{C:N}}$	C:N ratio	6.625	mol C (mol N)$^{-1}$	Based on Redfield ratio (106:16)
$\theta_{\text{C:O}}$	C:O_2 ratio	0.7681	mol C (mol O_2)$^{-1}$	Based on Redfield ratio (106:138)
$\theta_{\text{C:P}}$	C:P ratio	106	mol C (mol P)$^{-1}$	Based on Redfield ratio (106:1)
$f_{\text{C:OM}}$	Fraction of C_{org} in OM	0.54	–	Calculated from values of the elemental composition of CHONP in standard carbohydrates, lipids, proteins and nucleic acids in OM (Anderson, 1995)
$f_{\text{C:TEP}}$	Fraction of TEC in a (dry) TEP	0.75	–	0.51–0.88 (Passow, 2002)
$f_{\text{d:w}}^{\text{TEP}}$	Fraction of dry matter in a wet TEP (the rest is water)	0.10	–	0.99 (Mari, 1999; Azetsu-Scott and Passow, 2004)
Primary production parameters				
$\mu_0^{\text{max, diat}}$	Max. diatom-specific growth rate at 0°C	0.70	divisions d^{-1}	0.60 (Le Quéré et al., 2005), 0.33 (Dunne et al., 2013), 0.44 (Le Quéré et al., 2016)
$\mu_0^{\text{max, dino}}$	Max. dinoflagellate-specific growth rate at 0°C	0.53	divisions d^{-1}	0.35 (Le Quéré et al., 2016)
$\mu_0^{\text{max, cocco}}$	Max. coccolithophore-specific growth rate at 0°C	0.38	divisions d^{-1}	0.20 (Le Quéré et al., 2005), 0.23 (Dunne et al., 2013), 0.70 (Le Quéré et al., 2016)
$\mu_0^{\text{max, pico}}$	Max. picophytoplankton-specific growth rate at 0°C	0.35	divisions d^{-1}	0.60 (Le Quéré et al., 2005), 0.26 (Le Quéré et al., 2016)

Table C.1 (Continued) SLAMS-2.0 parameters.

Symbol	Meaning	Value	Units	Literature values
$\mu^{\max}$	Max. phytoplankton-specific growth rate	2.0	divisions d^{-1}	1.34–2.0 (Gregg et al., 2003, 30°C), 0.4–3.3 (Sarthou et al., 2005, diatoms), 1.3–2.9 (Oguz and Merico, 2006), 1.0–2.92 (Irwin et al., 2006), 0.52–1.47 (Litchman et al., 2006), 0.1–2.1 (Edwards et al., 2015), 1.3 (Flynn and Raven, 2017)
$\alpha_{\text{chl}}^{\text{diat}}$	Diatom chl-specific initial slope of the P–I curve	3.7×10^{-5}	g C (g chl)$^{-1}$ m^2 (μmol photons)$^{-1}$	0.35–2.42×10^{-5} (Geider et al., 1997)
$\alpha_{\text{chl}}^{\text{dino}}$	Dinoflagellate chl-specific initial slope of the P–I curve	4.8×10^{-5}	g C (g chl)$^{-1}$ m^2 (μmol photons)$^{-1}$	0.14–1.49×10^{-5} (Geider et al., 1997)
$\alpha_{\text{chl}}^{\text{cocco}}$	Coccolithophore chl-specific initial slope of the P–I curve	3.0×10^{-5}	g C (g chl)$^{-1}$ m^2 (μmol photons)$^{-1}$	0.86–1.50×10^{-5} (Geider et al., 1997)
$\alpha_{\text{chl}}^{\text{pico}}$	Picophytoplankton chl-specific initial slope of the P–I curve	2.5×10^{-5}	g C (g chl)$^{-1}$ m^2 (μmol photons)$^{-1}$	0.95–2.40×10^{-5} (Geider et al., 1997)
$k_{\text{N}}^{\text{diat}}$	Diatom half-saturation cnt. for NO_3^- uptake	0.65	mmol N m^{-3}	0.2–3.6 (Lehman et al., 1975), 0.1–5.1 (Smayda, 1997), 1.0 (Gregg et al., 2003), 1.6 ± 1.9 (Sarthou et al., 2005), 0.93–1.58 (Litchman et al., 2006), 0.50 (Oguz and Merico, 2006), 1.0–2.0 (Litchman et al., 2007)
$k_{\text{N}}^{\text{dino}}$	Dinoflagellate half-saturation cnt. for NO_3^- uptake	0.50	mmol N m^{-3}	3.8–9.5 (Lehman et al., 1975), 1.0–29 (Smayda, 1997), 0.23 (Oguz and Merico, 2006), 2.5–6.3 (Litchman et al., 2006), 2.5–6.0 (Litchman et al., 2007)
$k_{\text{N}}^{\text{cocco}}$	Coccolithophore half-saturation cnt. for NO_3^- uptake	0.40	mmol N m^{-3}	0.10 (Lehman et al., 1975), 0.50 (Gregg et al., 2003), 0.38 (Oguz and Merico, 2006), 0.15–0.21 (Litchman et al., 2006), 0.20 (Litchman et al., 2007)
$k_{\text{N}}^{\text{pico}}$	Picophytoplankton half-saturation cnt. for NO_3^- uptake	0.23	mmol N m^{-3}	0.50 (Gregg et al., 2003), 0.30 (Oguz and Merico, 2006)
$k_{\text{Si}}^{\text{diat}}$	Diatom half-saturation cnt. for $SiOH_4$ uptake	2.0	mmol Si m^{-3}	4.0 (Le Quéré et al., 2005), 3.9 ± 5.0 (Sarthou et al., 2005), 1.0 (Litchman et al., 2006)

Table C.1 (Continued) SLAMS-2.0 parameters.

Symbol	Meaning	Value	Units	Literature values
f_{TEP}	Fraction of a phytoplankton cell exuded as transparent exopolymers	0.30	–	0.02–0.50 (Thornton, 2014)
Zooplankton parameters				
$\rho_{\mathrm{C}}^{\mathrm{zoo}}$	Zooplankton C density	0.12	$\mathrm{g\,cm^{-3}}$	Kiørboe (2011)
$p^{\mathrm{Mz,\,high}}$	Mesozooplankton high probability for growth	0.80	–	My guess
$p^{\mathrm{Mz,\,low}}$	Mesozooplankton low probability for growth	0.65	–	My guess
p^{radiol}	Radiolarian probability for growth	0.10	–	Based on biomass profiles (MAREDAT, Buitenhuis et al., 2013) and the value of 5.2% for the total oceanic biota C reservoir for Rhizarians (Biard et al., 2016)
p^{flagel}	Zooflagellate probability for growth	0.60	–	*as above*
p^{foram}	Foraminifera probability for growth	0.20	–	*as above*
p^{pterop}	Pteropod probability for growth	0.10	–	*as above*
$\theta_{\mathrm{POC:Z}}$	Standard POC:zooplankton biomass	3.0	–	3:1 phytoplankton:zooplankton biomass (Buitenhuis et al., 2013, depth profiles), 2:1 phytoplankton:zooplankton biomass (Laufkötter et al., 2013, global value)
$\theta_{\mathrm{POC:Z}}^{\mathrm{DVM\,day}}$	POC:zooplankton biomass at the day DVM depth	1.5	–	My guess, *based on above*
$\theta_{\mathrm{POC:Z}}^{\mathrm{DVM\,night}}$	POC:zooplankton biomass at the night DVM depth	0.50	–	My guess, *based on above*
R_{Mz}	Mesozooplankton detection radius factor	3.0	–	Based on Visser (2007)
$R_{\mu\mathrm{z}}$	Microzooplankton detection radius factor	2.0	–	Based on Visser (2007)
f_{sloppy}	Sloppy feeding loss fraction	0.20	–	0.20 (Yool et al., 2013)
β_{C}	Zooplankton absorption efficiency of C	0.70	–	0.69 (Anderson et al., 2015)

Table C.1 (Continued) SLAMS-2.0 parameters.

Symbol	Meaning	Value	Units	Literature values
k_{nz}	Zooplankton C-specific net growth efficiency	0.75	–	0.75 (Anderson et al., 2015)
ϕ^{Mz}	Mesozooplankton gut passage time	4,800	s	1200 (Irigoien, 1998), 600–18,000 (Besiktepe and Dam, 2002), 1500–7800 (Hansen and Visser, 2016), 800–1200 (Visser and Fiksen, 2013)
$\phi^{\mu\mathrm{z}}$	Microzooplankton gut passage time	1,200	s	*as above*
m^{Mz}	Mesozooplankton natural mortality rate	0.050	d^{-1}	0.10 (Visser, 2007, copepods)
$m^{\mu\mathrm{z}}$	Microzooplankton natural mortality rate	0.080	d^{-1}	*as above*
Respiration parameters				
Q_{10}^{diat}	Diatom Q_{10}	1.65	–	1.14–1.93 (Le Quéré et al., 2016), 1.5 (Oguz and Merico, 2006), 1.51–1.65 (Sherman et al., 2016), 1.88 (Eppley, 1972), 1.4–2.8 (Raven and Geider, 1988)
Q_{10}^{dino}	Dinoflagellate Q_{10}	1.65	–	*as above*
Q_{10}^{cocco}	Coccolithophore Q_{10}	1.65	–	*as above*
Q_{10}^{pico}	Picophytoplankton Q_{10}	1.65	–	*as above*
Q_{10}^{Mz}	Mesozooplankton Q_{10}	1.80	–	1.8–1.9 (Ikeda, 2014), 2.0 (Oguz and Merico, 2006)
$Q_{10}^{\mu\mathrm{z}}$	Microzooplankton Q_{10}	1.90	–	2.2–3.0 (Nielsen and Thomas, 1994), 2.0 (Oguz and Merico, 2006)
Q_{10}^{bac}	Bacteria Q_{10}	2.10	–	2.0 (Kirchman et al., 2009), 1.5–2.01 (Laufkötter et al., 2017), 1.5–2.7 (DeVries and Weber, 2017)
$\lambda_0^{\mathrm{max,\,Mz}}$	Max. mesozooplankton C-specific respiration rate at 0°C	0.080	d^{-1}	0.010–0.10 (Hernández-León and Ikeda, 2005)
$\lambda_0^{\mathrm{max},\,\mu\mathrm{z}}$	Max. microzooplankton C-specific respiration rate at 0°C	0.080	d^{-1}	*as above*

Table C.1 (Continued) SLAMS-2.0 parameters.

Symbol	Meaning	Value	Units	Literature values
$\lambda_0^{\max,\,\mathrm{bac}}$	Max. bacteria C-specific respiration rate at 0°C	0.085	d^{-1}	0.12 at 15°C and 0.03 at 4°C (Iversen and Ploug, 2013)
$\lambda^{\max,\,\mathrm{Mz}}$	Max. mesozooplankton C-specific respiration rate	0.25	d^{-1}	0.010–0.10 (Hernández-León and Ikeda, 2005, mesozoo.)
$\lambda^{\max,\,\mu\mathrm{z}}$	Max. microzooplankton C-specific respiration rate	0.30	d^{-1}	0.13 (Cavan et al., 2017), 0.01–0.5 (references in Cavan et al., 2017), 0.12 at 15°C, 0.03 at 4°C (Iversen and Ploug, 2013)
$\lambda^{\max,\,\mathrm{bac}}$	Max. bacteria C-specific respiration rate	0.30	d^{-1}	*as above*
k_{O_2}	O_2 half-saturation cnt. for respiration	0.020	$\mathrm{mL\,L^{-1}}$	0.34–1.3 (García-Robledo et al., 2016, Costa Rica), 0.0022–0.067 (references in Ploug, 2001), 0.22–0.67 (DeVries and Weber, 2017, model), 0.090–0.27 (Laufkötter et al., 2017)
f_{solub}	Bacteria solubility loss fraction	0.10	–	0.5 (Anderson and Tang, 2010), 0.4–0.7 (Stemmann et al., 2004, use avg. bacteria respiration rate of 0.10 d^{-1})
Mineral dissolution parameters				
κ_{calc}	Calcite dissolution timescale	3.5	d^{-1}	1–7 (Jansen et al., 2002, calcite is more stable than aragonite so it has a lower κ)
κ_{arag}	Aragonite dissolution timescale	4.5	d^{-1}	*as above*
κ_{opal}	Opal dissolution timescale	8.0×10^{-3}	d^{-1}	3.0×10^{-4} to 0.30 (Fig. 4, Van Cappellen et al., 2002, surface ocean), 8.0×10^{-3} to 0.18 (Bidle and Azam, 1999, colonised by bacteria)
η_{calc}	Calcite dissolution order reaction	2.0	–	1–5 (Jansen et al., 2002)
η_{arag}	Aragonite dissolution order reaction	2.0	–	*as above*
$[\mathrm{CO}_3^{2-}]_{\mathrm{sat}}^{\mathrm{calc}}$	Calcite saturation concentration in the surface ocean	42.7	$\mathrm{mmol\,m^{-3}}$	Jansen (2001) (Table 3.1), after Millero (1995)
$[\mathrm{CO}_3^{2-}]_{\mathrm{sat}}^{\mathrm{arag}}$	Aragonite saturation concentration in the surface ocean	42.7	$\mathrm{mmol\,m^{-3}}$	*as above*

Table C.1 (Continued) SLAMS-2.0 parameters.

Symbol	Meaning	Value	Units	Literature values
Physical and chemical constants.				
k_w	Absorption coefficient of water	0.040	m^{-1}	
k_{chl}	Absorption coefficient of chl *a*	0.015	$m^2\,(mg\,chl)^{-1}$	
E	Energy of a photon	3.90×10^{-19}	J	
k_B	Boltzmann cnt.	1.38×10^{-23}	$J\,K^{-1}$	
N_A	Avogadro cnt.	6.02×10^{23}	mol^{-1}	
g	Gravity cnt.	9.81	$m\,s^{-2}$	
M_C	Molar mass of C	12.011	$g\,mol^{-1}$	
M_{CaCO_3}	Molar mass of $CaCO_3$	100.1	$g\,mol^{-1}$	
M_{opal}	Molar mass of opal (hydrated amorphous silica, $SiO_2 \cdot 0.4\,H_2O$)	67.3	$g\,mol^{-1}$	
M_{clay}	Molar mass of clay (kaolinite, $Al_2O_3 \cdot 2\,SiO_2 \cdot 2\,H_2O$)	226.2	$g\,mol^{-1}$	
ρ_{OM}	Density of OM	1.06	$g\,cm^{-3}$	
ρ_{TEP}	Density of a TEP	0.800	$g\,cm^{-3}$	Mari et al. (2017); Azetsu-Scott and Passow (2004)
ρ_{calc}	Density of calcite	2.87	$g\,cm^{-3}$	
ρ_{arag}	Density of aragonite	2.90	$g\,cm^{-3}$	
ρ_{opal}	Density of opal	1.90	$g\,cm^{-3}$	
ρ_{clay}	Density of clay	2.70	$g\,cm^{-3}$	
ρ_w	Density of water	1.027	$g\,cm^{-3}$	
μ_{w_0}	Water viscosity parameter	6.829×10^{-6}	$g\,cm^{-1}\,s^{-1}$	Based on an empirical relationship dynamic viscosity-temperature from tabulated values
b_μ	Exponent of the water viscosity curve	2160	–	*as above*
$V_m(O_2)$	Molar volume of O_2 at STP	22.4	$L\,mol^{-1}$	

C.1 Water physics

Water physics are parameterised through two foundational hydrodynamic properties: the dynamic viscosity of seawater (μ) and the turbulent kinetic energy (TKE) dissipation rate (ϵ). The variable μ ($\mathrm{g\,cm^{-1}\,s^{-1}}$) tells how much internal resistance the fluid offers to flow. It is a function of pressure, salinity and T, amongst which T is by far the most important factor in a marine system. In SLAMS-2.0, I calculate μ as a function of T as

$$\mu = \mu_{\mathrm{w}_0} e^{(b_\mu / T)}, \tag{C.1}$$

where $\mu_{\mathrm{w}_0} = 6.829 \cdot 10^{-6}$ $\mathrm{g\,cm^{-1}\,s^{-1}}$ is the water viscosity parameter, $b_\mu = 2,160$ is the exponent of the curve and T is in degrees Kelvin. The parameters μ_{w_0} and b_μ can be obtained from tabulated values of seawater μ and T. A parameter closely related to water dynamic viscosity is water kinematic viscosity, ν ($\mathrm{cm^2\,s^{-1}}$), which is derived as the quotient

$$\nu = \frac{\mu}{\rho_w}, \tag{C.2}$$

where $\rho_w = 1.027$ $\mathrm{g\,cm^{-1}}$ is the density of seawater. The TKE dissipation rate, ϵ ($\mathrm{m^2\,s^{-3}}$), is the dissipation of the kinetic energy that stirs the ocean, which originates from the winds and tides. That energy is dissipated by water viscosity at a micro-scale and it generates turbulent mixing. The variable ϵ varies in horizontal space by up to several orders of magnitude, ranging from 10^{-7}–10^{-4} $\mathrm{m^2\,s^{-3}}$ in shelf seas and the coastal ocean up to 10^{-10}–10^{-6} $\mathrm{m^2\,s^{-3}}$ in the open ocean in response to physical forcing (Saiz and Kiørboe, 1995). In SLAMS-2.0, ϵ is approximated by a log-normal distribution curve constructed with the values indicated in the previous ranges (after Pearson and Fox-Kemper, 2018),

$$f(\epsilon) = \frac{1}{\epsilon\sigma\sqrt{2\pi}}\, e^{-\frac{(\ln\epsilon - \mu)^2}{2\sigma^2}}, \tag{C.3}$$

where $\epsilon = \texttt{logspace}(-7, -4, 3)$ if the water column depth $H < 200$ m (shallow, coastal waters) and $\epsilon = \texttt{logspace}(-10, -6, 4)$ if $H \geq 200$ m (open ocean waters); μ is the mean of the distribution and its value is the logarithm of the mean value in the ϵ range, i.e., $\mu = \ln[\sum(\epsilon)/n]$; σ is the variance of the distribution and equals 1.2. The distribution $f(\epsilon)$ is randomly sampled using the MCM. It is known that ϵ not only varies in the horizontal space but also with depth, although to a lesser extent (Pearson and Fox-Kemper, 2018). Thus, in SLAMS-2.0 I have assumed that the vertical, depth distribution of ϵ is constant.

The combination of ϵ (Eq. C.3) and ν (Eq. C.2) allows the derivation of two other hydrodynamic properties used in SLAMS-2.0. One is the water shear rate, γ ($\mathrm{s^{-1}}$), which measures the water velocity gradients derived from turbulent mixing,

$$\gamma = \sqrt{\frac{\epsilon}{\nu}}. \tag{C.4}$$

The second one is the Kolmogorov length scale, η (m), which is the size of the smallest possible eddy in the water below which fluid viscosity dominates and all turbulent kinetic energy is dissipated (as heat) instead of creating turbulent motion,

$$\eta = \left(\frac{\nu^3}{\epsilon}\right)^{1/4}. \tag{C.5}$$

C.2 The diagnostic particle attributes

C.2.1 Fractal law-derived attributes

Real marine primary particles range in shapes. Nevertheless, modelled primary particles are assumed to be spherical, to behave hydrodynamically as rounded 3-D bodies. Yet, modelled aggregates, built up from spherical primary particles, are not rounded but fractal (after Logan and Wilkinson, 1990). The concept of fractality in marine particles is used as a measure of the internal chaotic organisation of an aggregate (Kilps, 1993). Thus, when the primary particles coalesce, they do it disorderly, leaving holes in between and making the (fractal) aggregates have a rough perimeter instead of a perfectly rounded one.

The fractality of modelled aggregates is represented by a single parameter, D_3, the three-dimensional fractal dimension. In mathematical abstraction, D_3 has a value between 3 (for volumetric Euclidean, 3-D structures) and 1 (for lines, Euclidean 1-D structures). The reported values of D_3 in the marine literature range between 1.26–2.49 (Logan and Wilkinson, 1990; Li and Logan, 1995; Jackson et al., 1997) and are estimates obtained by measuring on images of particles properties like radius, area, porosity or counts of primary particles and finding statistical correlations between them –D_3 is the power of those fits. Smaller values of D_3 indicate more porous aggregates, and therefore larger in size. Higher values of D_3 are associated with less porous and smaller aggregates (Maggi, 2007). Thus aggregates have a variable D_3 depending on their phase of growth. In SLAMS-2.0, a crucial assumption is made by keeping the fractal dimension invariant across all aggregates, with a fixed value of $D_3 = 2.0$ chosen after the synthesis of Logan and Wilkinson (1990). Keeping D_3 fixed is necessary because, unlike observational studies, it is not a derived parameter from other properties, but the parameter that is used in the simulation precisely to derive those other particle properties like size, porosity or density. If we could visualise the aggregates in SLAMS-2.0, with a $D_3 = 2.0$, those would look like somewhat stringy bodies where newly added primary particles do not hide behind others but lie instead on the side, trying to spread the body to other directions. Next, I detail how fractal law scaling is applied to derive the modelled particle physicochemical attributes.

The first attribute of a particle that is computed after it interacts with its environment and the moles of the different constituent materials are changed is dry mass, m_p (g),

$$m_\mathrm{p} = \sum_{i=1} n_i \; M_i = \sum_{i=1} m_i, \qquad i = \mathrm{C_{org}},\ \mathrm{TEC},\ \mathrm{bSi},\ \mathrm{calc},\ \mathrm{arag},\ \mathrm{clay} \tag{C.6}$$

where n_i is the moles of material i and M_i is the molar mass of the material i (see Table C.1). Next, I calculate the volume occupied by the solid material (i.e., excluding the pore spaces), V_p^S ($\mu\mathrm{m}^3$),

$$V_\mathrm{p}^\mathrm{S} = \sum_{i=1} \frac{m_i}{\rho_i}. \tag{C.7}$$

The radius of the particle, r_p, needs three attribute inputs: the number of constituent primary particles (N_pp), their radius (r_pp) and the fractal dimension of the particle (D_3) (Meakin, 1998). Fractal law requires making two assumptions about the constituent primary particles: that they are all identical and, therefore, of the same size, and that they are spherical. Equally sized primary particles are achieved as $V_\mathrm{pp} = V_\mathrm{p}^\mathrm{S}/N_\mathrm{pp}$. Under the assumption of sphericity, $r_\mathrm{pp} = [3V_\mathrm{pp}/(4\pi)]^{1/3}$. Thus,

$$r_\mathrm{p} = r_\mathrm{pp} \; N_\mathrm{pp}^{1/D_3}, \tag{C.8}$$

Eq. C.8 shows that the radius of the aggregate, r_p (μm), increases as the fractal dimension decreases. That is because, with decreasing fractal dimension, the more open the aggregate structure is, and so the radius increases.

The calculation of the porosity, P_p, of the particle is the next stage. Porosity refers to the fraction of a particle occupied by interstitial water, i.e., the fraction occupied by pores filled with water. It can be calculated as 1 minus the fraction of solid material. The fraction of solid material is given by the quotient of the solid volume of the particle to its bulk volume, $V_\mathrm{p}^\mathrm{S}/V_\mathrm{p}^\mathrm{B}$, where V_p^B is still unknown at this stage. But I can derive a known expression for the fraction of solid material using the attributes that I already know,

$$\begin{aligned} P_\mathrm{p} &= 1 - \frac{V_\mathrm{p}^\mathrm{S}}{V_\mathrm{p}^\mathrm{B}} = 1 - \frac{N_\mathrm{pp} V_\mathrm{pp}}{V_\mathrm{p}^\mathrm{B}} = 1 - \frac{N_\mathrm{pp} \; \frac{4}{3}\pi r_\mathrm{pp}^3}{\frac{4}{3}\pi r_\mathrm{p}^{D_3}} \\ &= 1 - \left[\left(\frac{r_\mathrm{p}}{r_\mathrm{pp}} \right)^{D_3} \left(\frac{r_\mathrm{p}}{r_\mathrm{pp}} \right)^{-3} \right] = 1 - \left(\frac{r_\mathrm{p}}{r_\mathrm{pp}} \right)^{D_3 - 3}, \qquad r_\mathrm{p} > r_\mathrm{pp}. \end{aligned} \tag{C.9}$$

Notice how non-fractal aggregates ($D_3 = 3$) are impermeable spheres, with no porosity ($P_\mathrm{p} = 0$). Fractal theory (Eq. C.9) implies that aggregates have a size-dependent porosity, where porosity increases with increasing size. The fractal nature of marine aggregates implies that small aggregates are expected to have primary particles separated by a smaller distance; therefore, small aggregates will be more compact overall, or less porous. In larger aggregates, the distance between primary particles increases as they have a

higher fraction of fluid medium incorporated within the particle structure and therefore are more porous. The mechanistic derivation of P is in agreement with observations made, for instance, by Engel (2000) and Laurenceau-Cornec et al. (2019).

The next stop is the calculation of the density of the bulk particle, ρ_p ($g\,cm^{-3}$), which results from the addition of the densities of the solid fraction and the aqueous (porous) fraction,

$$\rho_p = (1 - P_p)\,\rho_p^S + P_p\ \rho_w. \tag{C.10}$$

where ρ_w is the density of the water and ρ_p^S is the density of the solid particle, which excludes the water in the pores ($\rho_p^S = m_p/V_p^S$). According to fractal law, ρ_p depends on the aggregate size in that, as particle size increases, aggregates incorporate more void spaces into their internal structure, so their P_p increases and ρ_p decreases. That fractal decrease in particle density as its size increases is a fundamental property of marine aggregates, also corroborated by observations (Hill et al., 1998; Khelifa and Hill, 2006; Maggi, 2007, 2013). If aggregates were non-fractal objects, i.e., regular structures of primary particles separated by a fixed distance, their P_p would be constant with increasing size, and so ρ_p would stabilise at larger sizes (Bowers et al., 2017).

C.2.2 Surface reactivity, or stickiness

In inorganic substances, stickiness is a property that manifests as electrostatic forces (e.g., electrostatic attraction, hydrogen bonding, van der Waals forces), where surface charges create electrostatic bridges between particles, getting adsorbed into each other. However, in the presence of OM, stickiness is a force that also arises from an extra element: organic gel-like sticky surfaces that behave as biological glues. It is the case of a special case of carbohydrates, TEPs (Alldredge et al., 1993), hydrogels where the constituent polysaccharides are embedded within a mucilaginous gel-like matrix stabilised by divalent cation and half-ester sulfate bridging between the individual polysaccharides (Meng and Liu, 2016; Passow, 2002). TEPs are the only stickiness factor considered in SLAMS-2.0.

In empirical studies of marine particles, stickiness is measured as the probability that two colliding particles have of sticking to each other due to the presence of TEPs in the surrounding environment (Kiørboe and Hansen, 1993; Dam and Drapeau, 1995; Engel, 2000). Although there is still no set formula to compute the stickiness of a particle as a function of the presence of TEP, it is possible to intuitively derive a mechanistic formulation based on the findings of empirical studies, as did Jokulsdottir and Archer (2016). Thus, TEPs have maximum stickiness, $\alpha_p = 1$, and as they get embedded into aggregates with other particles, the stickiness of the aggregate is calculated as the fraction of particle volume occupied by TEP volume, as

$$\alpha_p = \frac{m_{TEP}/\rho_{TEP}}{V_p^S} = \frac{V_{TEP}}{V_p^S}. \tag{C.11}$$

C.2.3 Stokes' sinking velocity

Shape affects the particle's motion as frictional forces arise from the surface of a sinking particle. Shape should be accounted in calculations of the sinking velocity of a particle, w_p, but the impact of complex particle shapes on particle motion is a contested subject where, to date, only empirical equations can be derived for specific problem sets (e.g., Dörgens et al., 2015). In SLAMS-2.0, the calculation of w_p does not explicitly include in the equation the only considered particle shape-like factor, D_3. Thus, in the calculation of w_p, the particle is assumed to be a perfect sphere with a smooth surface, simplifying the treatment of the hydrodynamic effects of particle shape on w_p. That characteristic does not necessarily imply that modelled aggregates must appear round-like at the surface but, instead, that can be described by quantities equivalent to those of spheres.

The settling velocity results from a balance of forces. A particle falling in viscous fluid experiences three (e.g., Maggi, 2013): a downwards gravitational force (F_g), an upwards buoyancy force that emerges from the weight of the fluid displaced by the particle (F_b), and a frictional force that pulls the particle upwards, called drag force (F_d).

$$F_g = m_p g = \rho_p V_p^B g, \tag{C.12a}$$

$$F_b = m_w g = \rho_w V_p^B g, \tag{C.12b}$$

$$F_d = \frac{1}{2}\rho_w w_p^2 C_D A, \tag{C.12c}$$

where g is the gravitational constant, C_D is the drag coefficient (calculated next), and A is the projected area of the particle (πr_p^2). When the two retarding forces balance the gravitational force,

$$F_{net} = F_g - (F_b + F_d) = 0, \tag{C.13}$$

the net acceleration of the particle goes to zero, and the particle keeps falling at a constant velocity, its terminal velocity w_p, which can be computed from F_d in Eq. C.12c. For that, one must derive first an expression for the drag coefficient C_D, which is a function of the fluid type around the particle.

The water flowing around a sinking particle at terminal velocity can either be streamlined (laminar fluid flow) or can be forming vortices and turbulence wakes at the back (turbulent fluid flow). The Reynolds number, Re, makes that distinction. Re is a dimensionless number that comes from the ratio between the inertial forces (forces of the motion of the liquid) and the viscous forces (frictional force that arises from the resistance of the liquid to flow),

$$Re = \frac{\text{inertial force}}{\text{viscous force}} = \frac{2 r_p \rho_w w_w}{\mu}, \tag{C.14}$$

where w_w is the velocity of the fluid. At low values of Re, viscous effects dominate over inertial effects, and flows are orderly and laminar. When the fluid velocity increases, or so does the size of the particle, Re increases leading to vortices and, eventually, to turbulence. A mechanistic model for the relationship between Re and C_D does not exist in the literature yet, C_D has been derived experimentally only for impermeable spheres (Maggi, 2013), with the White (1974) relationship commonly being used,

$$C_D = \frac{24}{Re}, \quad Re < 1 \text{ (laminar flow)}, \tag{C.15a}$$

$$C_D = \frac{24}{Re} + \frac{6}{1+\sqrt{Re}} + 0.4, \quad Re >> 1 \text{ (turbulent flow)}. \tag{C.15b}$$

C_D derived here for impermeable spheres is assumed to be applicable to fractal aggregates (Maggi, 2013). Thus, there exist two expressions for w_p: one derived for laminar flow and another for turbulent flow.

To derive w_p for laminar flow, C_D of Eq. C.15a is substituted into Eq. C.12c and A is replaced by πr_p^2, leaving

$$F_d = 6\pi r_p \mu w_p, \tag{C.16}$$

an expression known as the Stokes' law, a formula describing the upwards force acting on a sphere of radius r_p moving with a steady velocity w_p held in a laminar fluid of viscosity μ. By substituting Eqs. C.12a, C.12b and C.16 into the force balance in Eq. C.13, and replacing V_p^B by $(4/3)\pi r_p^3$ (after having assumed sphericity), the Stokes' terminal settling velocity for laminar flow is isolated as

$$w_p = \frac{g(\rho_p - \rho_w){d_p}^2}{18\mu}, \tag{C.17}$$

where $d_p = 2r_p$ is the diameter of the particle (also known as equivalent spherical diameter, ESD), and $(\rho_p - \rho_w)$ is the particle excess density. According to Eq. C.17, w_p increases linearly with density and to the squared power of the radius, with larger particles experiencing a parabolic increase in velocity. Eq. C.17 is commonly used to compute particle Stokes' velocity (Burd and Jackson, 1997; Ploug et al., 2008; Bach et al., 2016; Laurenceau-Cornec et al., 2019; Maerz et al., 2020) and comes with three important assumptions: that the flow is laminar ($Re < 1$) and that the particle is spherical and impermeable (no fractality nor porosity). To derive w_p for turbulent flow, C_D of Eq. C.15b is substituted into Eq. C.12c, leaving

$$w_p = \sqrt{\frac{4g(\rho_p - \rho_w)d_p}{3\rho_w C_D}}, \tag{C.18}$$

an expression that is solved iteratively using Newton's method as the Re has to be adjusted for the new turbulent w_p (mind that Re is calculated with the velocity of the particle, instead of that of the fluid, for which I don't have a value in SLAMS-2.0).

C.3 Zooplankton number, carbon quota, size and swimming speed

Following the method of discretisation of NPP into phytoplankton cells, the discretisation of the zooplankton biomass, B_{zoo}, present at a depth into a number of zooplankton organisms, N_{zoo}, is achieved by dividing B_{zoo} by the average amount of organic carbon that makes a zooplanktoner. The calculation of the zooplankton organic carbon body quota (θ^{i}_{C}) follows the empirical expressions of Kiørboe (2013, Table 2), with a distinct radius for mesozooplankton and for microzooplankton. For mesozooplankton (Mz),

$$\theta^{\text{Mz}}_{\text{C}} = 10^{-0.93 + 0.95 \log_{10}\left(\frac{4}{3}\pi r^3_{\text{Mz}} \times \rho_{\text{zoo}}\right)}, \tag{C.19a}$$

$$\text{where:} \quad r_{\text{Mz}} = \frac{1}{2}\left(0.38^{2/3} \times \text{PL}^{Cala}\right), \tag{C.19b}$$

and for microzooplankton (μz),

$$\theta^{\mu\text{z}}_{\text{C}} = 10^{-1.37 + 0.88 \log_{10}\left(\frac{4}{3}\pi r^3_{\mu\text{z}} \times \rho_{\text{zoo}}\right)}, \tag{C.20a}$$

$$\text{where:} \quad r_{\mu\text{z}} = \frac{1}{2}\left(\frac{\text{PL}^{Cala}}{10}\right), \tag{C.20b}$$

where $\theta^{\text{i}}_{\text{C}}$ (i = Mz or μz) is the carbon mass (mg C (body)$^{-1}$); the expression inside the $\log_{10}$ is the wet mass; ρ_{zoo} is the zooplankton carbon density (120 mg C cm^{-3}, Kiørboe, 2011, Fig. 1); and r_i (cm) is the radius of a zooplankter and it is estimated from the prosome length of a *Calanus* individual (PL^{Cala}) (see Eq. C.21 below). Eq. C.19b assumes that mesozooplankton shape is that of a prolate sphere (Kiørboe, 2011), where 0.38 is the aspect ratio and the radius is 1/2 of the ESD of the prolate sphere. Eq. C.20b corrects PL^{Cala} for microzooplankton individual by reducing it by a factor of 10, and assumes sphericity of the microzooplankton by implicitly equating the corrected PL to ESD.

Since copepod PL is strongly related to the temperatures, T, at which growth occurs (Wilson et al., 2015), I have calculated PL based on a T–PL empirical relationship that I have built using a data set of PL from non-polar *Calanus* species. The data set is an extraction from the bigger copepod trait data set of Brun et al. (2016), from which I have taken 21 references, and it is available at https://doi.org/10.5281/zenodo.5247486. It gives this T–PL^{Cala} relationship (n = 77, R^2 = 0.48):

$$\text{PL}^{Cala} = -0.0319\,T + 2.8574, \tag{C.21}$$

where T (°C) is SST (the average temperature in the first 20 m in SLAMS-2.0) and PL^{Cala} is between 2–3 mm. To take into account the high uncertainty associated to PL^{Cala}, its

value in Eqs. C.19b and C.20b is chosen randomly between [(PL^{Cala}/10), ($\mathrm{PL}^{Cala} \times 10$)] (boundaries chosen considering that mesozooplankton are 0.2–20 mm in size).

The zooplankton swimming speed is calculated from Pinti et al. (2019)'s empirical formulation (Eq. in their associated code),

$$v^i_{\mathrm{zoo}} = 0.0204\,(\mathrm{PL}^i)^{0.825} \quad \begin{array}{l}(\text{if } i = \mathrm{Mz},\ \mathrm{PL}^i = \mathrm{PL}^{Cala},\ \text{or} \\ \text{if } i = \mu z,\ \mathrm{PL}^i = \mathrm{PL}^{Cala}/10).\end{array} \tag{C.22}$$

where v^i_{zoo} is in $\mathrm{m\,s^{-1}}$ and PL^i is in cm.

D

Calculation of incident PAR at the surface ocean

The calculation of photosynthetic active radiation at the surface ocean (PAR_0) from extraterrestrial solar radiation uses standard trigonometric/astronomical equations and input observations of cloud cover and ice cover. The process can be split into seven steps.

The first step is the calculation of the solar irradiance at the top of the atmosphere, I_{top}, which varies according to latitude (ϕ) and Julian day (t). For that, I have adopted the classic model of Brock (1981, Eqs. 1–8),

$$I_{\text{top}}(\phi, t) = \frac{I_{\text{SC}} \cos Z(\phi, t)}{R_{\text{v}}(t)^2}, \tag{D.1}$$

where $I_{\text{top}}(\phi, t)$ is in W m^{-2}; ϕ is in radians; t is the day number in the year (e.g., 32 for the 1st of February); I_{SC} is the solar constant, or average intensity per square meter at the surface of the Earth, and has a value of 1368 W m^{-2} (equivalent to units of $\text{J s}^{-1}\,\text{m}^{-2}$); $Z(\phi, t)$ is the sun zenith angle in radians according to latitude ϕ and Julian day t; $R_{\text{v}}(t)$ is the Earth's radius vector for Julian day t. The equation for $Z(\phi, t)$ is

$$Z(\phi, t) = \arccos[\sin(\delta(t)) \sin(\phi)] + \cos(\delta(t)) \cos(\phi), \tag{D.2}$$

where $\delta(t)$ is the sun declination angle in radians, or the angle between the equatorial plane of the Earth and the axis joining the centres of the Earth and the Sun, and is calculated as

$$\delta(t) = 23.45 \sin\left(\frac{2\pi(284 + t)}{365}\right) \frac{2\pi}{360}, \tag{D.3}$$

where 23.45 is the Earth's tilt from the orbit plane in degrees, and the last term, $2\pi/360$, is a conversion factor from degrees to radians. Lastly, the expression for $R_{\text{v}}(t)$ in Eq. D.1 is

$$R_{\text{v}}(t) = \frac{1}{\sqrt{1 + 0.033 \cos\left(\frac{2\pi t}{365}\right)}}. \tag{D.4}$$

The second step towards the calculation of PAR_0 is the calculation of the radiation that reaches the Earth's surface at solar noon under clear skies, $I_{0,\mathrm{clear}}$, for which I have adopted the model of Rosati and Miyakoda (1988, Eqs. 3.3–3.7):

$$I_{\mathrm{beam}}(\phi,t) = I_{\mathrm{top}}(\phi,t)\ T_a^{[1/\cos(Z(\phi,t))]}, \tag{D.5a}$$

$$I_{\mathrm{diffuse}}(\phi,t) = [(1-A_a)I_{\mathrm{top}}(\phi,t) - I_{\mathrm{beam}}(\phi,t)]\frac{1}{2}, \tag{D.5b}$$

$$I_{0,\mathrm{clear}} = I_{\mathrm{beam}}(\phi,t) + I_{\mathrm{diffuse}}(\phi,t). \tag{D.5c}$$

$I_{0,\mathrm{clear}}$ is calculated from the sum of a direct component of solar radiation (I_{beam}) and a diffuse sky radiation component under cloudless conditions (I_{diffuse}); T_a is the atmospheric transmission coefficient and has a typical value of 0.7; A_a is the absorption by vapour and ozone and has a typical value of 0.09.

Thirdly, one has to account for the attenuation by clouds on $I_{0,\mathrm{clear}}$, for which I have used the empirical formula derived by Reed (1977),

$$I_0(\phi,\lambda,t) = I_{0,\mathrm{clear}}(1 - 0.62\ f_C(\phi,\lambda,t) + 0.0019\ \beta(\phi,t)). \tag{D.6}$$

$I_0(\phi,\lambda,t)$ ($\mathrm{W\,m^{-2}}$) is the solar radiation that reaches the Earth's surface at solar noon with clouds on the skies at latitude ϕ, longitude λ and Julian day t. The fractional cloud cover, f_C (0–1), is calculated as $C(\phi,\lambda,t)/8$, where C is cloud cover data in oktas (units: 1–8) obtained from a monthly climatology of satellite observations and CMIP3 models from Pincus et al. (2008) available at the NOAA Physical Sciences Laboratory site (https://psl.noaa.gov/data/gridded/cmip3-clouds-rad-precip/) and interpolated to daily values. Lastly, β is the Sun elevation angle in degrees, and is calculated as

$$\begin{aligned}\beta(\phi,t) = \arcsin\Bigg\{&\sin(\phi)\sin\left[23.45\frac{2\pi}{360}\sin\left((t-82)\frac{2\pi}{360}\right)\right]\\ &+\cos(\phi)\cos\left[23.45\frac{2\pi}{360}\sin\left((t-82)\frac{2\pi}{360}\right)\right]\Bigg\}\frac{360}{2\pi},\end{aligned} \tag{D.7}$$

where $360/(2\pi)$ is the conversion factor from radians to degrees.

The fourth step is discounting from $I_0(\phi,t)$ the radiation that is reflected to the atmosphere by the surface ocean, termed ocean albedo (α), as well as by the ice covering the ocean (f_{ice}). Added to that, since only the solar radiation in the spectral range of 400–700 nm (visible light, or shortwave radiation) can stimulate the chlorophyll in the photosystems, i.e., the photosynthetic active radiation (PAR), I_0 will be converted into PAR_0 by applying a conversion factor, f_{PAR},

$$\mathrm{PAR}_{0,\mathrm{noon}}(\phi,\lambda,t) = I_0(\phi,\lambda,t)\ (1-\alpha)\ (1 - f_{\mathrm{ice}}(\phi,\lambda))\ f_{\mathrm{PAR}}. \tag{D.8}$$

$\mathrm{PAR}_{0,\mathrm{noon}}(\phi,\lambda,t)$ ($\mathrm{W\,m^{-2}}$) is the photosynthetic active radiation reaching the surface ocean at solar noon. The ocean albedo, α, has a typical value of 0.06, meaning that

the ocean reflects 6% of solar radiation. The fraction of ocean covered by ice, f_{ice}, is an annual climatology from a MIT OGCM output interpolated to daily values. The PAR conversion factor, f_{PAR}, has a typical value of 0.43, meaning that PAR comprises 43% of solar radiation.

Fifth is the calculation of the variation of PAR_0 with time of day h, $\text{PAR}_0(\phi, \lambda, t, h)$. In Eq. D.8 I have calculated the variation of PAR_0 over the calendar year, $\text{PAR}_{0,\text{noon}}(\phi, \lambda, t)$, which has a sinusoidal pattern over the 1–365 day cycle, peaking at the summer solstice. The variation of PAR_0 with the time of day, $\text{PAR}_0(\phi, \lambda, t, h)$, also follows a sinusoidal pattern, from sunset to sunrise and peaking at noon, after Platt et al. (1990),

$$\text{PAR}_0(\phi, \lambda, t, h) = \text{PAR}_{0,\text{noon}}(\phi, \lambda, t, h) \sin\left(\frac{\pi(h - h_{\text{sunrise}}(\phi, t))}{n_h(\phi, t)}\right), \tag{D.9}$$

where h is the time of the day in hours; $h_{\text{sunrise}}(\phi, t)$ is the sunrise time in hours from midnight (midnight = 00:00), and $n_h(\phi, t)$ is the number of daylight hours in a day. The calculation of $n_h(\phi, t)$ is

$$n_h(\phi, t) = \frac{2}{15} \arccos\left[-\tan(\phi)\tan(\delta(t))\right] \frac{360}{2\pi}, \tag{D.10}$$

and $h_{\text{sunrise}}(\phi, t)$ is calculated as noontime minus half the hours of daylight,

$$h_{\text{sunrise}}(\phi, t) = 12 - n_h(\phi, t)\frac{1}{2}, \tag{D.11a}$$

$$h_{\text{sunset}}(\phi, t) = 12 + n_h(\phi, t)\frac{1}{2}. \tag{D.11b}$$

Sixth, the daily curve of $\text{PAR}_0(\phi, \lambda, t, h)$ (W m^{-2}) is integrated over the discrete number of daylight hours (from sunrise to sunset time) to get the total amount of radiation received per unit area in a day at that location,

$$T_{\text{PAR}_0}(\phi, \lambda, t_p) = \sum_{h_{\text{sunrise}}(\phi,t)}^{h_{\text{sunset}}(\phi,t)} \text{PAR}_0(\phi, \lambda, t, h)\ 3600. \tag{D.12}$$

where 3600 is the number of seconds in an hour. Thus, $T_{\text{PAR}_0}(\phi, \lambda, t_p)$ is the total amount of radiation received per unit area in the daylight hours, or photoperiod (t_p), and has units of J m^{-2}, which come from the integral in time of W m^{-2} (=$\text{J s}^{-1}\,\text{m}^{-2}$).

Seventh, and lastly, $T_{\text{PAR}_0}(\phi, \lambda, t_p)$ is homogeneously distributed over the number of time steps of a day that are exposed to daylight, $n_{\Delta t}$. Each day in SLAMS-2.0 contains three time steps, thus of a duration of eight hours each (Δt = 8 h); consequently, $n_{\Delta t}$ can take the value of 0, 1, 2 or 3 depending on the location and time of the year. But a decision was made to fix $n_{\Delta t}$ to 2 time steps, regardless of the location and time, distributing light everywhere on the globe with a $n_{\Delta t}$ that is typical of mid-latitude regions in the summer. The distribution of $T_{\text{PAR}_0}(\phi, \lambda, t_p)$ (J m^{-2}) into time steps exposed to daylight gives $\text{PAR}_0(\phi, \lambda, \Delta t)$ (W m^{-2}), the average PAR_0 received per unit area during the time steps exposed to daylight,

$$\text{PAR}_0(\phi, \lambda, \Delta t) = \frac{T_{\text{PAR}_0}(\phi, \lambda, t_p)}{n_{\Delta t}\ \Delta t\ 3600}. \tag{D.13}$$

E

Implementation of DFO-LS in SLAMS-2.0

OptClimSO needs a specific folder and file architecture to work. As a matter of example, Fig. E.1 shows the directory structure I have used with SLAMS-2.0. There are two files in OptClimSO that will need some changes before running the optimiser: `runOptimise.py` and `dfols_optim.json`. I have implemented DFO-LS with settings for noiseless (or smooth) runs by setting trust-region parameters to $tr\_radius.gamma\_dec$ = 0.5, $tr\_radius.alpha1$ = 0.1, $tr\_radius.alpha2$ = 0.5 in the `.json` file and specifying $objfun\_has\_noise = False$ in the `.py` file, which also changes some trust-region-related parameters. Also, in the `.json` file, I have increased the value of the lower bound of the trust region radius, *rhoend*, from the default 10^{-7} to 0.05 so that the parameter restart functionality is triggered earlier when DFO-LS gets stuck in a local optimum. When a local optimiser deals with complex optimisation topographies, it can get trapped in local minimum points that are not the actual absolute lowest (i.e.,the global minimum). To circumvent that, the user can either restart DFO-LS when the misfit decreases towards the minimum very slowly –by using the *rhoend* functionality–, and/or reinitialise DFO-LS with a different initial parameter set –by setting up different runs– to ensure more of the parameter space is searched. Thus, to globalise DFO-LS and search more of the parameter space, I have chosen 10 different initial parameter sets distributed in a parameter space randomly using Matlab's function `lhsdesign`. That improves the coverage of the parameter space, which is especially needed with SLAMS-2.0 as twin experiments have shown that the ability of DFO-LS to find the target parameter value is dependent on the initial parameter set. I have imposed boundaries on each parameter subject to calibration, with an upper and lower boundary specified in the `.json` file. Finally, I have imposed a limit of 255 function evaluations ($maxfun$ = 255 in the `.json` file). For the rest of the options, I have used default presets. The output from DFO-LS is read from a final `.json` file that it creates inside the loop folder, with the new misfits and parameter vector.

Figure E.1. Tree view of the directory structure that I have used to implement OptClimSO in SLAMS-2.0.

F

Additional tables and figures for Chapter 5

This Appendix contains ancillary tables and figures referenced throughout the text in Chapter 5. Table F.1 shows the initial parameter configuration used in the 10 twin and optimisation replicates. Table F.2 displays the optimised parameter sets obtained in the best experimental replicates of the optimisation experiments OSIM and OIND. Figs. F.1 and F.2 show twin experiment results. Fig. F.3 shows results for the optimisation experiment where locations have been fit individually to data (OIND).

Table F.1. Parameter values of the 10 replicates used in the twin and optimisation experiments. The parameter values have been obtained with Matlab's Latin hypercube design function and imposing a limit to 10% of the parameter boundaries, a limit imposed by DFO-LS search algorithm. Parameter symbols are: $\theta^{\text{diat}}_{\text{Si:C}}$, diatom cell Si:C ratio; $\theta^{\text{cocco}}_{\text{calc:C}}$, coccolithophore cell PIC:POC ratio; $\theta_{\text{POC:Z}}$, standard POC to zooplankton biomass; $\theta^{\text{DVM night}}_{\text{POC:Z}}$, POC to zooplankton biomass at the night DVM depth; f_{sloppy}, sloppy feeding loss fraction; $f^{\text{ing, surf}}_{\text{OM}}$, fraction of OM in an appetitive surface particle; κ_{opal}, opal dissolution timescale.

Symbol	Boundaries	Replicate initial value									
		1	2	3	4	5	6	7	8	9	10
$\theta^{\text{diat}}_{\text{Si:C}}$	0.03 – 0.65	0.58	0.52	0.37	0.48	0.40	0.13	0.28	0.18	0.23	0.32
$\theta^{\text{cocco}}_{\text{calc:C}}$	0.50 – 2.50	1.45	0.71	2.07	2.27	1.68	1.50	1.11	1.96	1.30	0.95
$\theta_{\text{POC:Z}}$	1.5 – 10	4.51	5.46	3.79	6.84	5.91	7.52	7.92	8.64	3.45	2.88
$\theta^{\text{DVM night}}_{\text{POC:Z}}$	0.20 – 2.0	1.42	0.39	1.05	1.15	1.76	0.69	0.56	1.36	1.55	0.81
f_{sloppy}	0.05 – 0.40	0.21	0.12	0.29	0.19	0.11	0.35	0.15	0.26	0.31	0.24
$f^{\text{ing, surf}}_{\text{OM}}$	0.10 – 0.80	0.39	0.44	0.69	0.22	0.23	0.30	0.59	0.64	0.55	0.48
κ_{opal}	0.005 – 0.050	0.044	0.027	0.038	0.041	0.032	0.011	0.023	0.029	0.020	0.017

Table F.2. Optimised parameter set for the five locations optimised simultaneously (OSIM) and on an individual basis (OIND) plus accompanying misfit values. The locations are: EqPac, in the upwelling equatorial Pacific (0°N, 140°W); HAUSGARTEN, in the Arctic Atlantic (79.0°N, 4.3°E); OSP (Ocean Station Papa), in the subarctic eastern Pacific (50°N, 145°W); PAP-SO (Porcupine Abyssal Plain time-Series Observatory), in the blooming NE Atlantic (49°N, 16.5°W); BATS/OFP (Bermuda Atlantic Time-Series/Oceanic Flux Program) joint site, in the subtropical NW Atlantic (31.6°N, 64.2°W). Parameter symbols are: $\theta^{\text{diat}}_{\text{Si:C}}$, diatom cell Si:C ratio; $\theta^{\text{cocco}}_{\text{calc:C}}$, coccolithophore cell PIC:POC ratio; $\theta_{\text{POC:Z}}$, standard POC:zooplankton biomass; $\theta^{\text{DVM night}}_{\text{POC:Z}}$, POC:zooplankton biomass at the night DVM depth; f_{sloppy}, sloppy feeding loss fraction; $f^{\text{ing, surf}}_{\text{OM}}$, fraction of OM in an appetitive surface particle; κ_{opal}, opal dissolution timescale.

Symbol	Boundaries	OSIM	OIND				
			EqPac	HAUSG.	OSP	PAP-SO	BATS
$\theta^{\text{diat}}_{\text{Si:C}}$	0.03 – 0.65	0.06	0.03	0.03	0.12	0.03	0.03
$\theta^{\text{cocco}}_{\text{calc:C}}$	0.50 – 2.50	0.52	0.50	0.50	1.13	0.65	0.50
$\theta_{\text{POC:Z}}$	1.5 – 10	7.8	9.6	8.7	7.5	10	1.5
$\theta^{\text{DVM night}}_{\text{POC:Z}}$	0.20 – 2.0	0.78	2.0	2.0	0.66	2.0	0.34
f_{sloppy}	0.05 – 0.40	0.40	0.40	0.40	0.36	0.40	0.40
$f^{\text{ing, surf}}_{\text{OM}}$	0.10 – 0.80	0.10	0.16	0.13	0.16	0.10	0.10
κ_{opal}	0.005 – 0.050	0.027	0.025	0.050	0.019	0.050	0.050
f initial*		15,374	822	18,520	59.6	46,367	5,517
f^{min}		40.8	4.00	27.3	1.90	198	0.82
% reduction in f		99.7	99.5	99.9	96.8	99.6	100.0
Generations		52	108	153	134	41	284
Replicate		4	9	5	6	7	8

* f initial is the misfit associated to the initial parameter configuration, that selected randomly using a latin hypercube design. f = 776 for the default parameter configuration set.

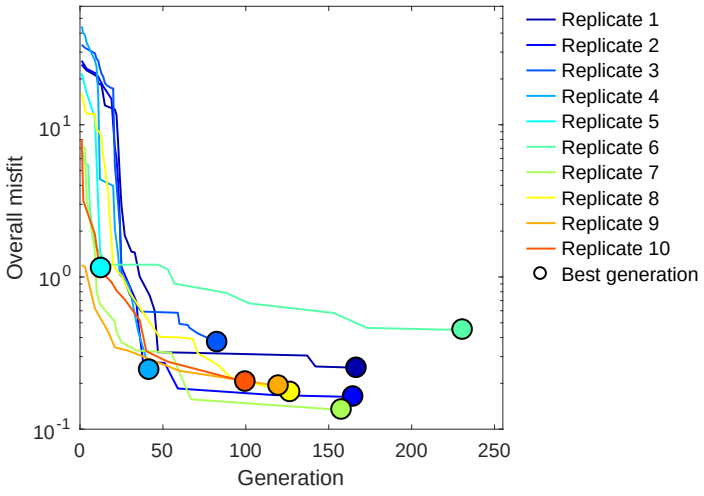


Figure F.1. Overall misfit trajectory of the twin experiment, with the 10 experimental replicates. The evolution of the overall misfit is shown for up to 255 generations, or optimiser iterations. The filled-in circles mark the value of the best generation, or optimiser iteration with the lowest misfit, and the shown replicate trajectories stop there. In a twin experiment we look for the best misfit reduction, which typically requires 100–150 generations in those replicates.

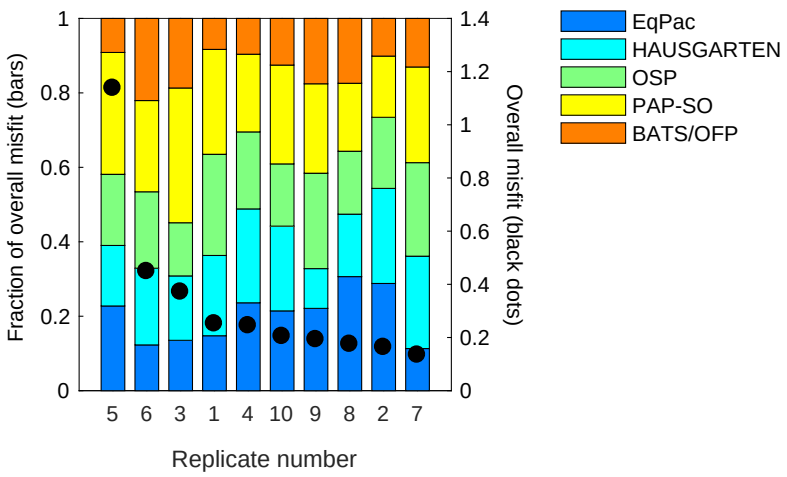


Figure F.2. Twin's best generation misfit and fraction of the overall misfit accounted by each location in each of the 10 experimental replicates. Replicates are ordered decreasingly by the overall misfit. All locations contribute similarly, ~20%, to the overall misfit.

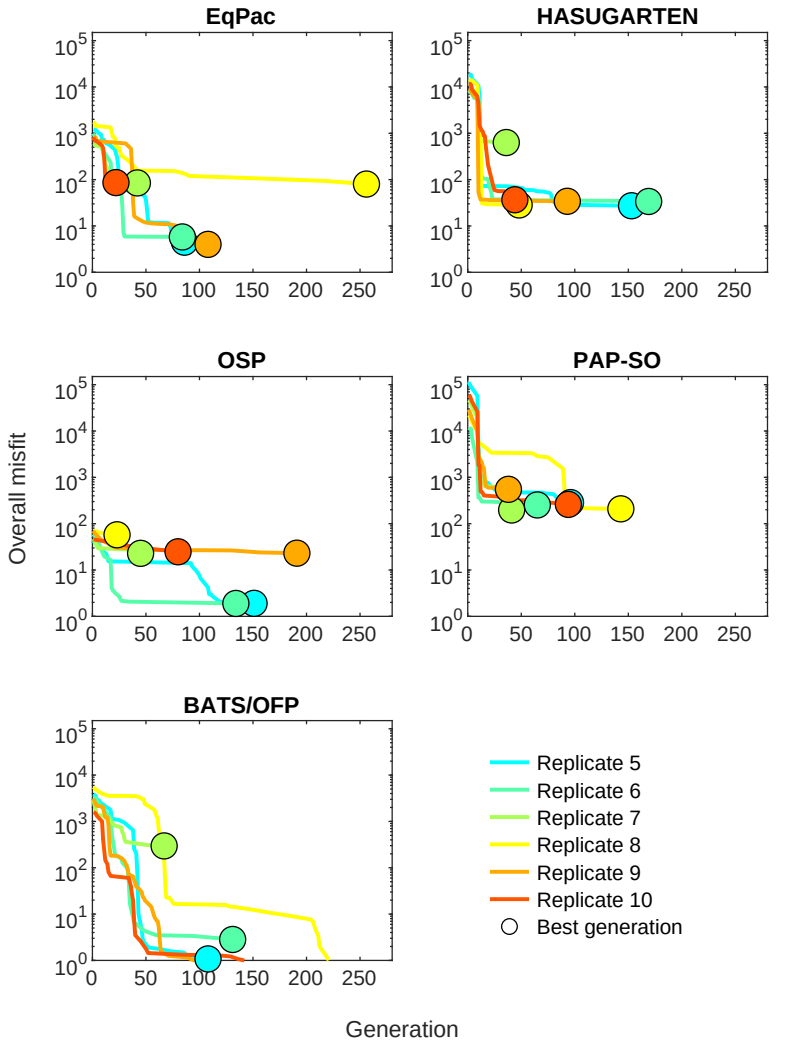


Figure F.3. Overall misfit trajectory of the OIND experiment, with the six experimental replicates. The evolution of the overall misfit is shown for up to 255 generations, or optimiser iterations. The filled-in circles mark the value of the best generation, or optimiser iteration with the lowest misfit, and the shown replicate trajectories stop there.

G

Additional tables and figures for Chapter 6

This Appendix contains ancillary tables and figures referenced throughout the text in Chapter 6. Fig. G.1 shows the global distribution of the four phytoplankton functional types modelled by SLAMS-2.0. Figs. G.2–G.9 show the global distribution of modelled particle attributes. Figs. G.11 and G.10 show the relationship between pairs of modelled particle attributes. Fig. G.12 shows modelled zooplankton flow rates. Table G.1 displays key aspects of the main published studies the outputs of SLAMS-2.0 are compared with.

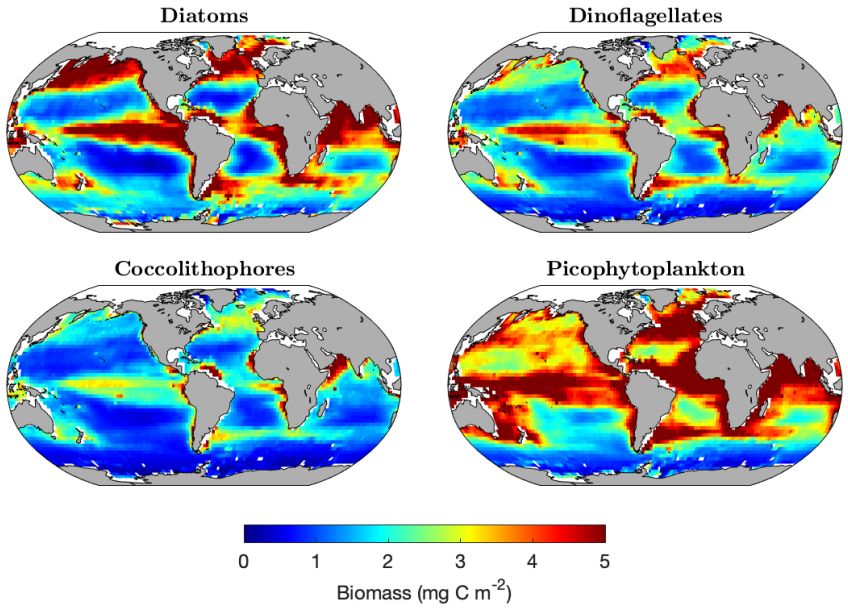


Figure G.1. SLAMS-2.0's simulated global distribution of the biomass of four phytoplankton functional types integrated over the euphotic layer and expressed in $\mathrm{mg\,C\,m^{-2}}$: (top left) diatoms, (top right) dinoflagellates, (bottom left) coccolithophores and (bottom right) picophytoplankton.

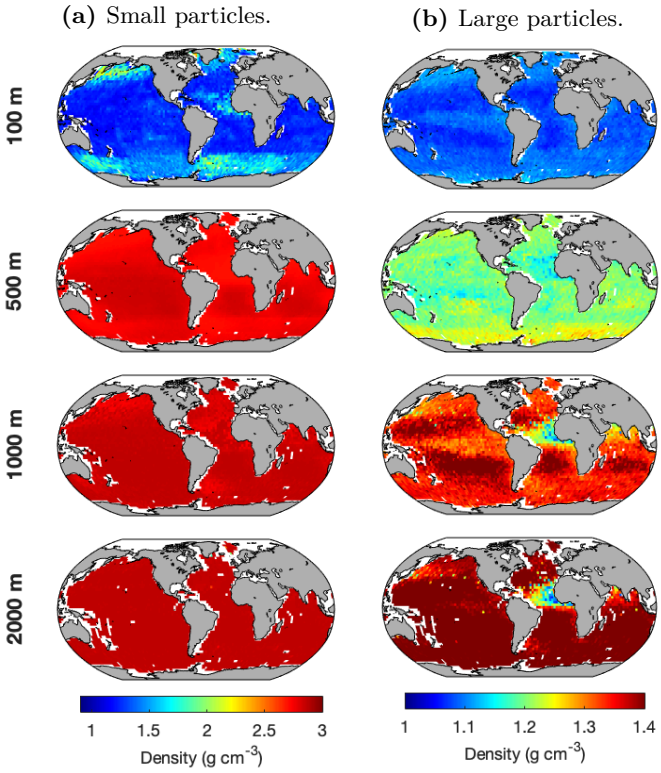


Figure G.2. Density of an average modelled particle at four main ocean depths according to two main size categories, (a) small (≤ 150 μm) and (b) large (>150 μm).

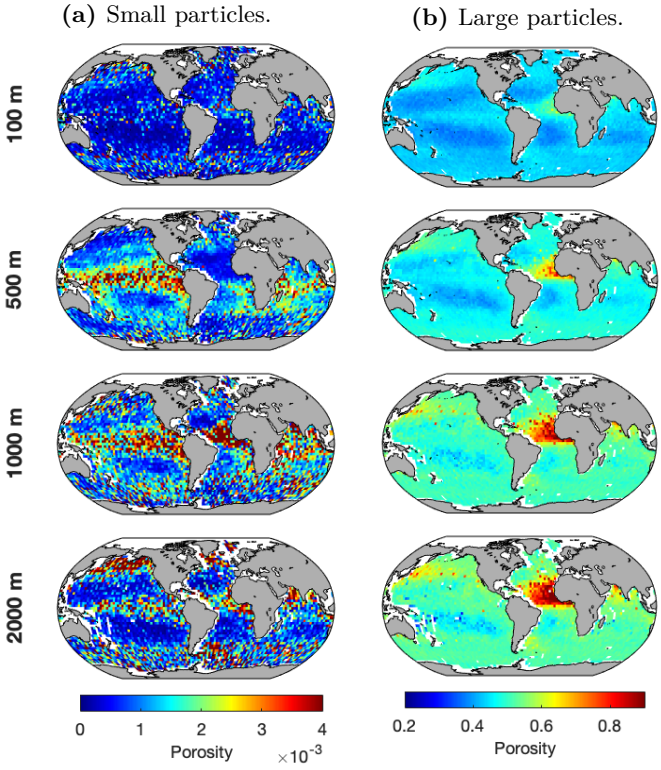


Figure G.3. Porosity of an average modelled particle at four main ocean depths according to two main size categories, (a) small (≤ 150 μm) and (b) large (>150 μm).

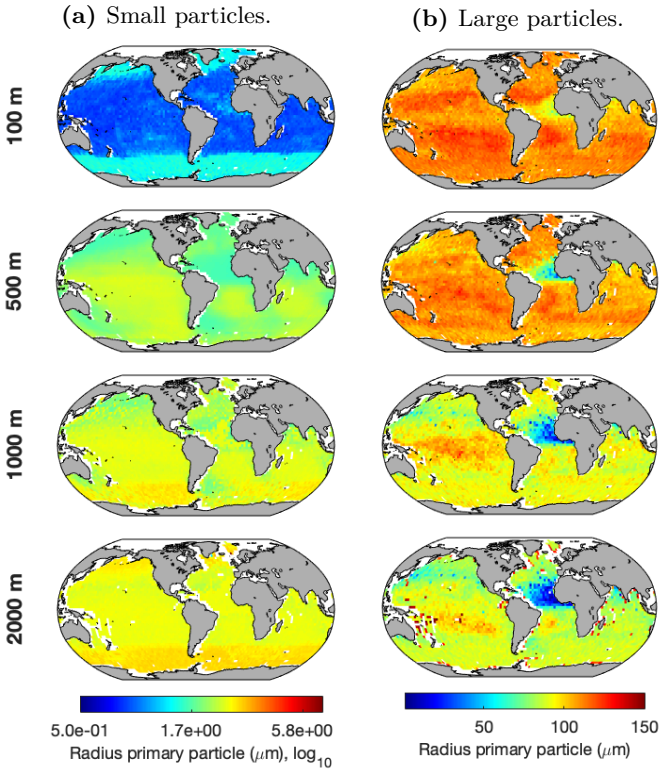


Figure G.4. Radius of the primary particles of an average modelled particle at four main ocean depths according to two main size categories, (a) small (≤ 150 μm) and (b) large (>150 μm).

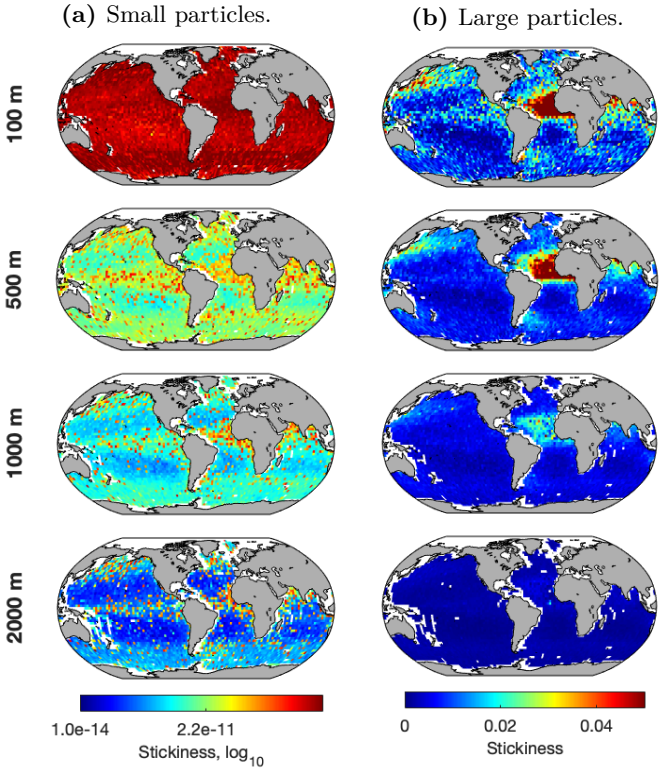


Figure G.5. Stickiness of an average modelled particle at four main ocean depths according to two main size categories, (a) small (≤ 150 μm) and (b) large (>150 μm).

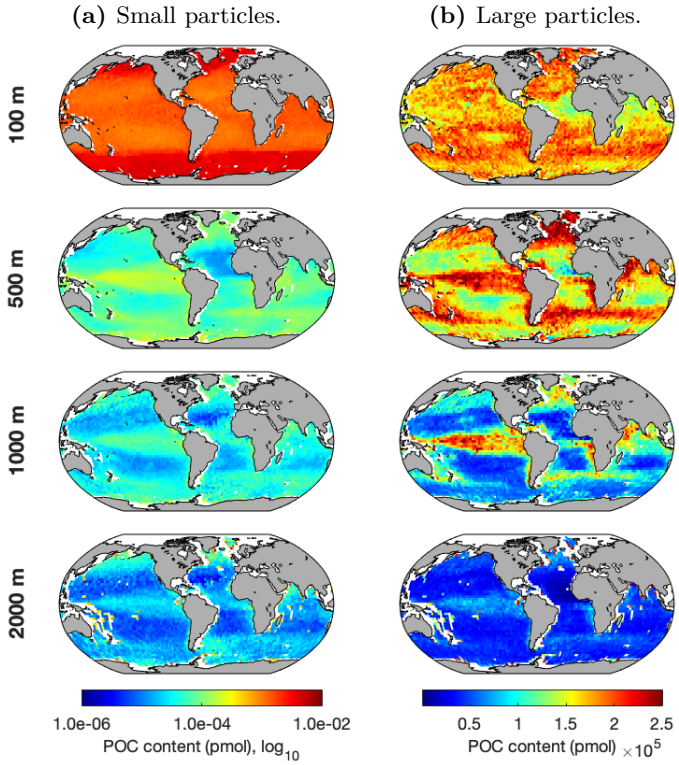


Figure G.6. POC content of an average modelled particle at four main ocean depths according to two main size categories, (a) small (≤ 150 μm) and (b) large (>150 μm).

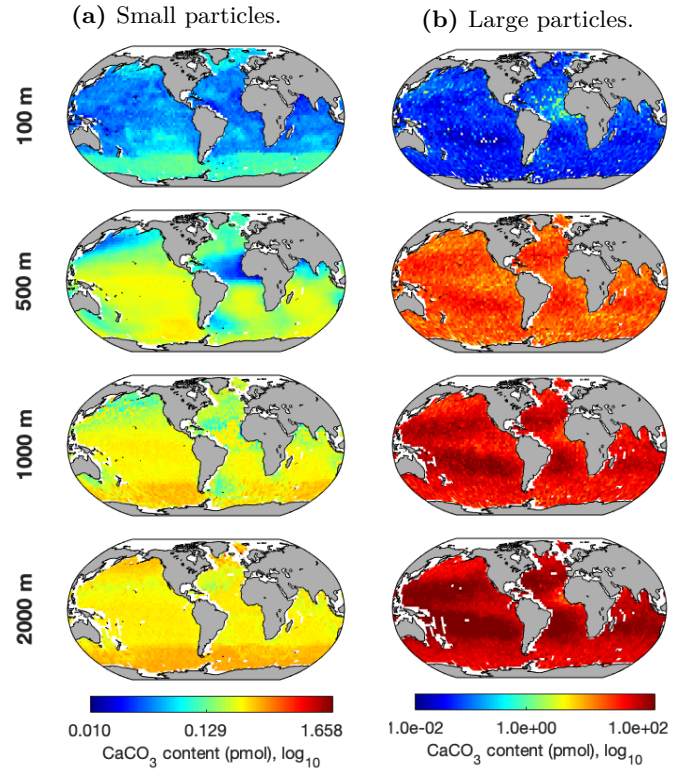


Figure G.7. PIC content of an average modelled particle at four main ocean depths according to two main size categories, (a) small (≤ 150 μm) and (b) large (>150 μm).

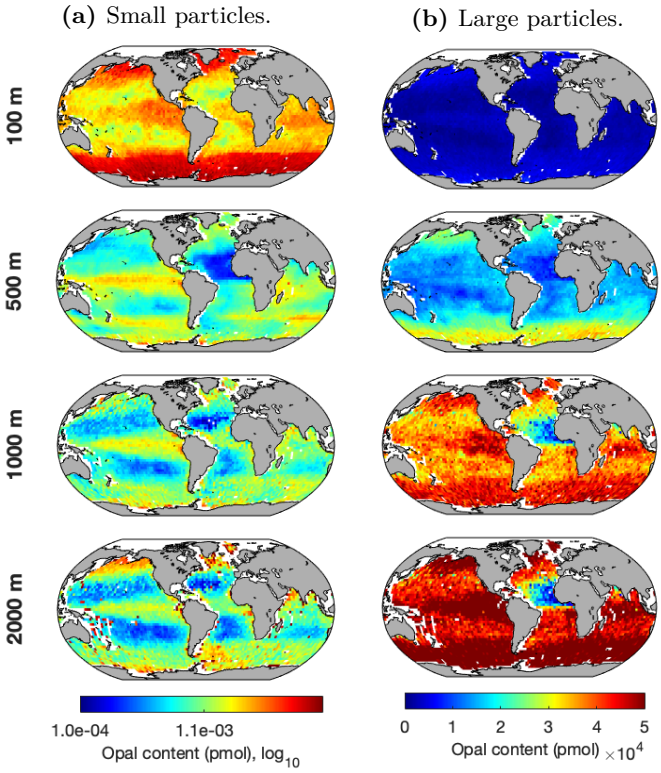


Figure G.8. Opal content of an average modelled particle at four main ocean depths according to two main size categories, (a) small (≤ 150 μm) and (b) large (>150 μm).

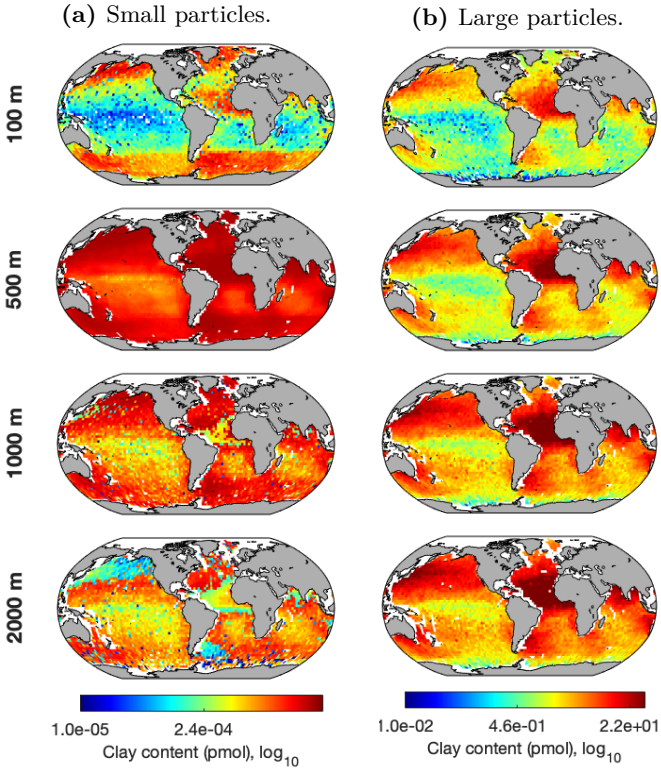


Figure G.9. Clay content of an average modelled particle at four main ocean depths according to two main size categories, (a) small (≤ 150 μm) and (b) large (>150 μm).

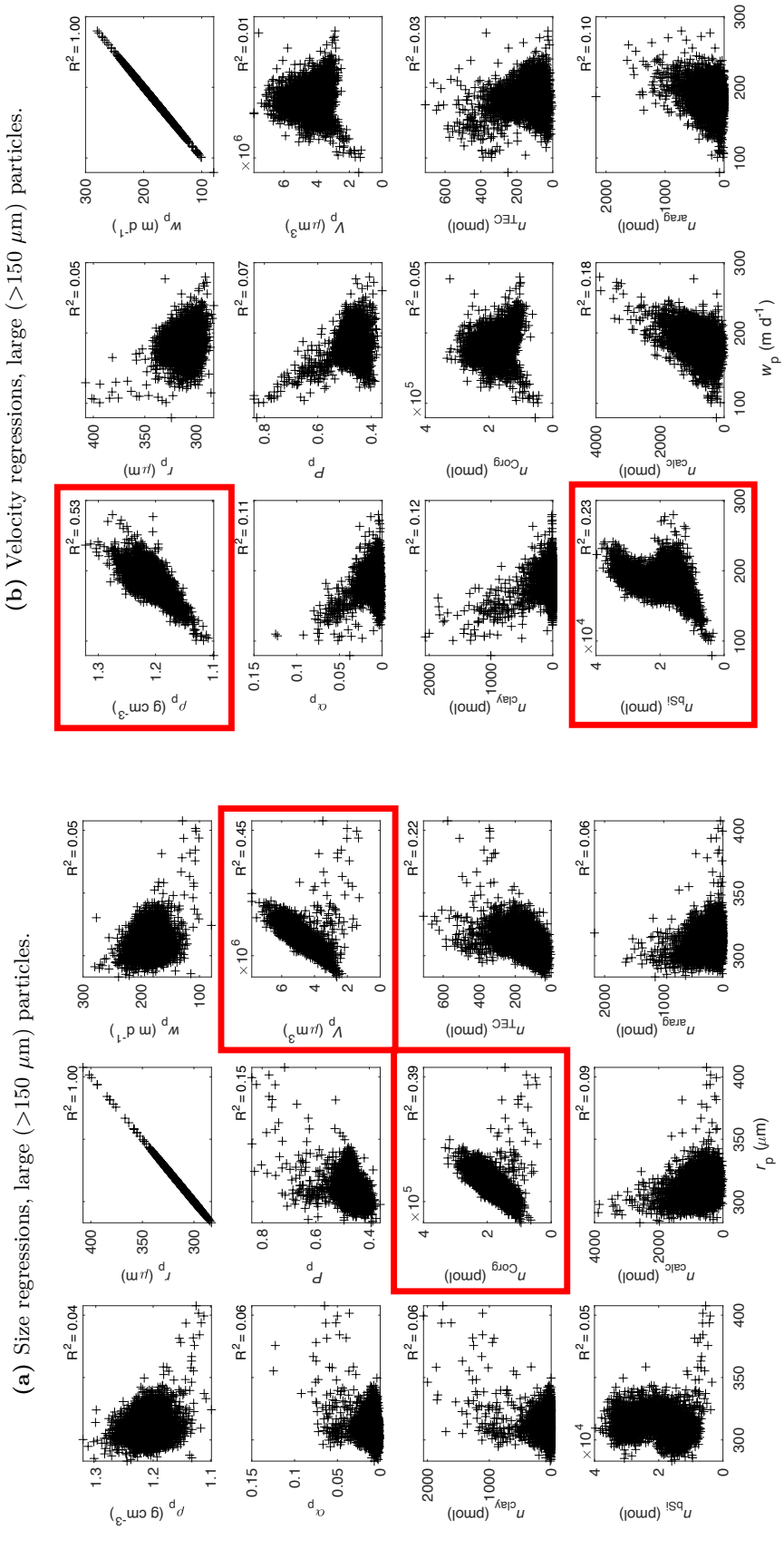


Figure G.10. Relationship between (a) particle size, r_p, and (b) particle sinking velocity, w_p, and the rest of modelled particle attributes at the 500 m depth horizon for large (>150 μm) particles. Squared in red are those regressions with a higher correlation coefficient (R^2). Data plotted are from the OSIM1 experiment. Particle attribute symbols are: ρ_p, density; r_p, radius; w_p, sinking velocity; α_p, stickiness; P_p, porosity; V_p, solid volume; n_{clay}, clay content; n_{Corg}, C_{org} content; n_{TEC}, TEC content; n_{bSi}, bSi content; n_{calc}, calcite content; n_{arag}, aragonite content.

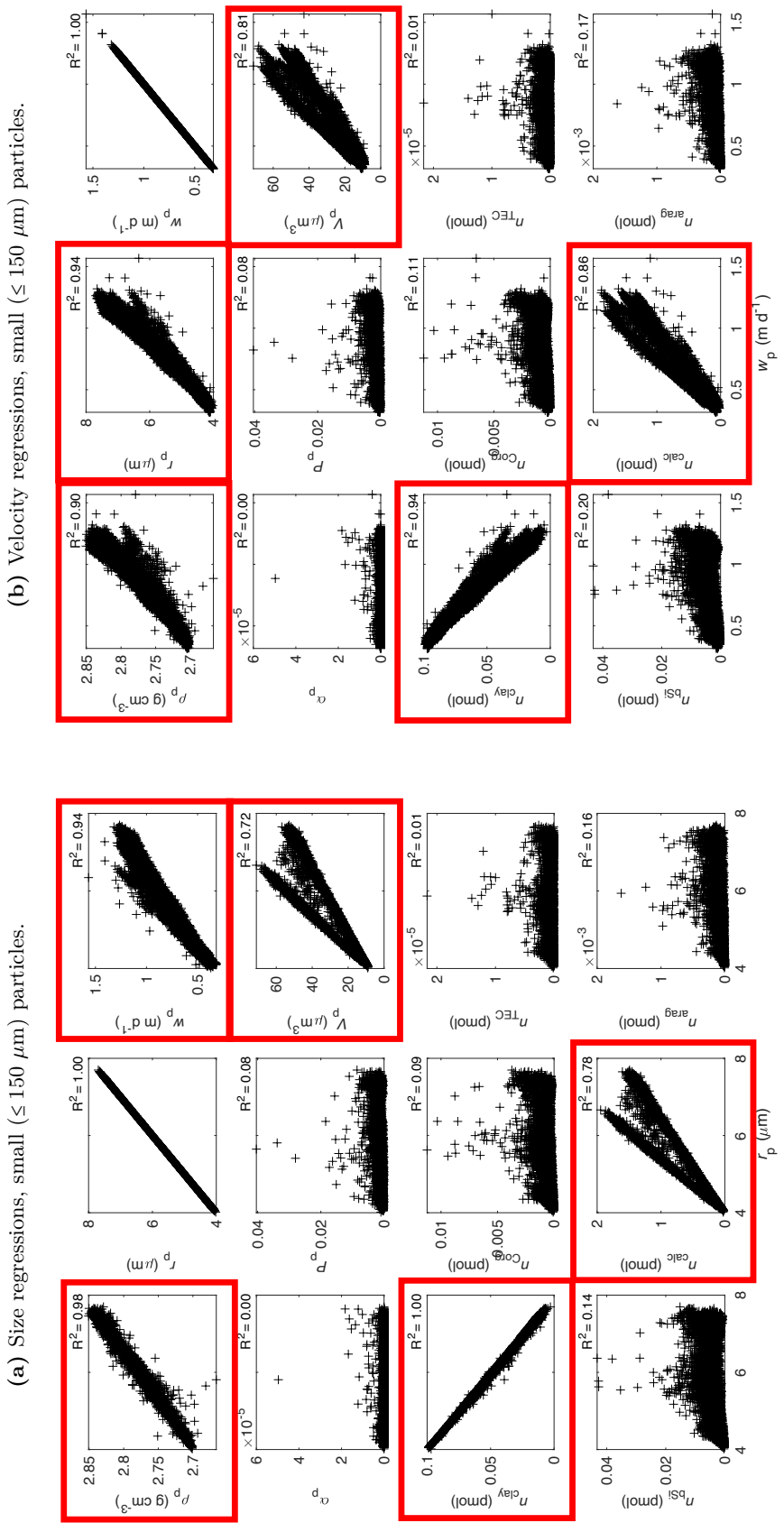


Figure G.11. Relationship between (a) particle size, r_p, and (b) particle sinking velocity, w_p, and the rest of modelled particle attributes at the 500 m depth horizon for small (≤ 150 μm) particles. Squared in red are those regressions with a higher correlation coefficient (R^2). Data plotted are from the OSIM1 experiment. Particle attribute symbols as in Fig. G.10.

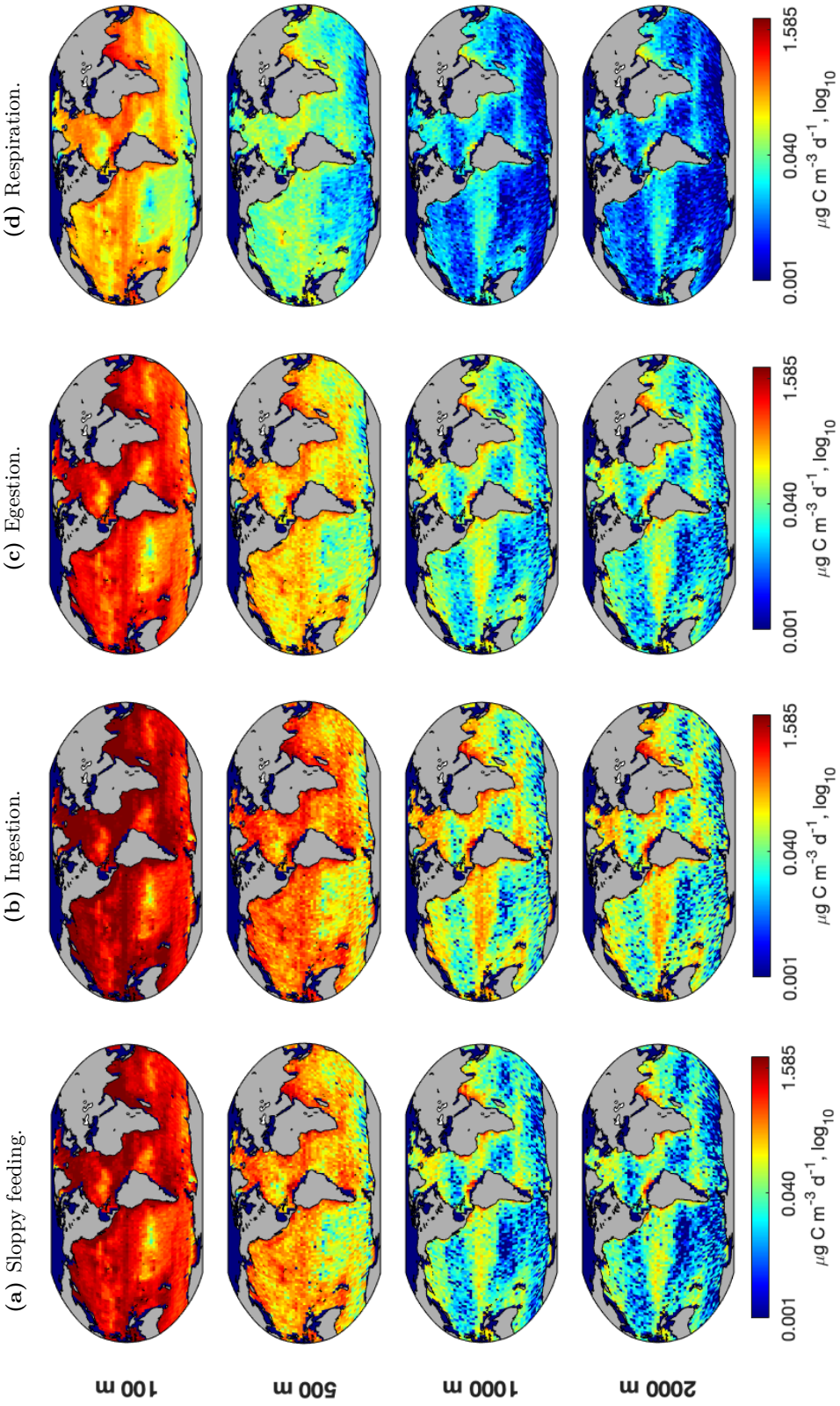


Figure G.12. SLAMS-2.0's simulated global distribution of zooplankton (microzooplankton + mesozooplankton) flow rates in $\mu g\,C\,m^{-3}\,d^{-1}$ by depth horizon: (a) sloppy feeding, (b) ingestion, (c) egestion and (d) respiration. Notice the $\log_{10}$ axis.

Table G.1. Comparison of key features of two empirical (Henson et al., 2012; Marsay et al., 2015) and two satellite-driven (Weber et al., 2016; DeVries and Weber, 2017) models commonly used in the literature to present regional patterns of Martin's b and T_{eff}.

Reference	Type of calculation	Data needed	Sources of the data
Henson et al. (2012)	b is calculated using an empirical algorithm that incorporates (i) Carr (2002) model to estimate NPP, (ii) Henson et al. (2011) equation for PE_{eff} and (iii) Lutz et al. (2007) equation for POC flux at 2000 m using a seasonal variation index of NPP.	NPP T	To *force* the algorithm: (i) NPP from the Carr (2002) model, which uses T at export depth (WOA), chl a and PAR_0 (SeaWiFs), and (ii) PE_{eff} from the Henson et al. (2011) equation, which uses SST (AVHRR). To *validate* the algorithm: (i) satellite-derived NPP and Th-derived estimates of POC flux at export depth (100 m) and (ii) sediment trap estimates of deep-sea POC flux (2000 m). POC flux are data from Honjo et al. (2008) spanning all ocean basins.
Marsay et al. (2015)	b is calculated using an empirical relationship with T.	T POC flux	T in the upper water column (0–500 m) from WOA and b calculated from published POC fluxes (50–800 m) at four North Atlantic locations (PAP-SO, Iceland Basin, Irminger Basin and NASG) and two locations in the North Pacific (ALOHA and K2).
Weber et al. (2016)	T_{eff} is the output of a data-driven ocean circulation model that uses phosphate ($[PO_4^{3-}]$) data. OM sinks as particulate organic phosphate (POP). POP export flux (at 300 m) is reconstructed by damping surface $[PO_4^{3-}]$ towards their observed values. POP flux in the surface ocean (0–300 m) is reconstructed using satellite-derived and CMIP5 estimates of POC flux, transformed into POC using a P:C ratio.	$[PO_4^{3-}]$ NPP T POC flux	POP > 300 m is a global distribution of $[PO_4^{3-}]$ from WOA. POP < 300 m are estimates calculated from global POC fluxes using a spatially varying P:C POC fluxes come from (i) satellite-based algorithms (an ensemble created from a permutation of estimates of NPP from the VGPM, VGPM-Eppley and CbPM and estimates of PE_{eff} using Laws et al. (2000), Dunne et al. (2005) and Laws et al. (2011) algorithms, each of which takes NPP and T as inputs) and (ii) CMIP5 model outputs, which explicitly model the BCP. Global observations are regionalised into eight major ocean biogeochemical regions between the Atlantic, Pacific and Antarctic.
DeVries and Weber (2017)	T_{eff} is the output of a data-driven model of the BCP embedded in an ocean circulation model (same as Weber et al. (2016)) and called SIMPLE-TRIM. The model is forced with NPP data that is converted into POC using the phytoplankton size fractionation provided by the satellite-based PSD of Kostadinov et al. (2010). The POC pool is further partitioned into slow vs fast-sinking and the DOC (produced from POC) has various levels of lability and driven to observed values. Attenuation of POC flux is parameterised using Q_{10} kinetics and dependency on O_2.	$[PO_4^{3-}]$ PSD NPP POC flux [DOC] $[O_2]$	POC is initially phytoplankton, which are fractionated into size classes using the PSD parameters (slope and abundance at reference diameter) retrieved from ocean colour data, following Kostadinov et al. (2010) (the procedure is based on the assumption that regions with higher chl a have larger particles, as well as higher satellite backscattering). NPP are estimates from the VGPM and the CbPM. For the cost function, the authors use: POC fluxes from the compilation of Lutz et al. (2007), [DOC] from a published compilation and $[O_2]$ from WOA.

Bibliography

Ådlandsvik, B., Bartsch, J., Brickman, D., Browman, H. I., Edwards, K., Fiksen, Ø., Hermann, A. J., Hinckley, S., Houde, E., Huret, M., Lacroix, G., Leis, J. M., McCloghrie, P., Megrey, B. A., Miller, T., Van der Molen, J., Mullon, C., Parada, C., Paris, C. B., Pepin, P., Rose, K., Thygesen, U. H., and Werner, C.: Manual of recommended practices for modelling physical: Biological interactions during fish early life, Tech. rep., International Council for the Exploration of the Sea (ICES), Copenhagen, Denmark, https://doi.org/10.25607/OBP-183, 2009.

Alldredge, A. L. and Gotschalk, C.: In situ settling behavior of marine snow, Limnology and Oceanography, 33, 339–351, https://doi.org/10.4319/lo.1988.33.3.0339, 1988.

Alldredge, A. L., Passow, U., and Logan, B. E.: The abundance and significance of a class of large, transparent organic particles in the ocean, https://doi.org/10.1016/0967-0637(93)90129-Q, 1993.

Anderson, L. A.: On the hydrogen and oxygen content of marine phytoplankton, Deep-Sea Research I, 42, 1675–1680, https://doi.org/10.1016/0967-0637(95)00072-E, 1995.

Anderson, T. R.: Plankton functional type modelling: Running before we can walk?, Journal of Plankton Research, 27, 1073–1081, https://doi.org/10.1093/plankt/fbi076, 2005.

Anderson, T. R. and Tang, K. W.: Carbon cycling and POC turnover in the mesopelagic zone of the ocean: Insights from a simple model, Deep-Sea Research II, 57, 1581–1592, https://doi.org/10.1016/j.dsr2.2010.02.024, 2010.

Anderson, T. R., Gentleman, W. C., and Yool, A.: EMPOWER-1.0: An Efficient Model of Planktonic ecOsystems WrittEn in R, Geoscientific Model Development, 8, 2231–2262, https://doi.org/10.5194/gmd-8-2231-2015, 2015.

Antia, A. N., Koeve, W., Fischer, G., Blanz, T., Schulz-Bull, D., Scholten, J., Neuer, S., Kremling, K., Kuss, J., Peinert, R., Hebbeln, D., Bathmann, U., Conte, M. H., Fehner, U., and Zeitzschel, B.: Basin-wide particulate carbon flux in the Atlantic Ocean: Regional export patterns and potential for atmospheric CO2 sequestration, Global Biogeochemical Cycles, 15, 845–862, https://doi.org/10.1029/2000GB001376, 2001.

Archer, D. and Maier-Reimer, E.: Effect of deep-sea sedimentary calcite preservation on atmospheric CO2 concentration, Nature, 367, 260–263, https://doi.org/10.1038/367260a0, 1994.

Archer, D. E., Eschel, G., Winguth, A., Broecker, W. S., Pierrehumbert, R., Tobis, M., and Jacob, R.: Atmospheric pCO2 sensitivity to the biological pump in the ocean, Global Biogeochemical Cycles, 14, 1219–1230, https://doi.org/10.1029/1999GB001216, 2000.

Armstrong, R. A., Lee, C., Hedges, J. I., Honjo, S., and Wakeham, S. G.: A new, mechanistic model for organic carbon fluxes in the ocean based on the quantitative association of POC with ballast minerals, Deep-Sea Research II, 49, 219–236, https://doi.org/10.1016/S0967-0645(01)00101-1, 2002.

Arteaga, L., Pahlow, M., and Oschlies, A.: Modeled Chl:C ratio and derived estimates of phytoplankton carbon biomass and its contribution to total particulate organic carbon in the global surface ocean, Global Biogeochemical Cycles, 30, 1791–1810, https://doi.org/10.1002/2016GB005458, 2016.

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., and Gehlen, M.: PISCES-v2: An ocean biogeochemical model for carbon and ecosystem studies, Geoscientific Model Development, 8, 2465–2513, https://doi.org/10.5194/gmd-8-2465-2015, 2015.

Azetsu-Scott, K. and Passow, U.: Ascending marine particles: Significance of transparent exopolymer particles (TEP) in the upper ocean, Limnology and Oceanography, 49, 741–748, https://doi.org/10.4319/lo.2004.49.3.0741, 2004.

Bach, L. T., Boxhammer, T., Larsen, A., Hildebrandt, N., Schulz, K. G., and Riebesell, U.: Influence of plankton community structure on the sinking velocity of marine aggregates, Global Biogeochemical Cycles, 30, 1145–1165, https://doi.org/10.1111/1462-2920.13280, 2016.

Bach, L. T., Stange, P., Taucher, J., Achterberg, E. P., Algueró-Muñiz, M., Horn, H., Esposito, M., and Riebesell, U.: The influcence of plankton community structure on sinking velocity and remineralization rate of marine aggregates, Global Biogeochemical Cycles, 33, 971–994, https://doi.org/10.1029/2019GB006256, 2019.

Baker, C. A., Henson, S. A., Cavan, E. L., Giering, S. L. C., Yool, A., Gehlen, M., Belcher, A., Riley, J. S., Smith, H. E. K., and Sanders, R. J.: Slow-sinking particulate organic carbon in the Atlantic Ocean: Magnitude, flux, and potential controls, Global Biogeochemical Cycles, 31, 1051–1065, https://doi.org/10.1002/2017GB005638, 2017.

Bauerfeind, E., Nöthig, E.-M., Beszczynska, A., Fahl, K., Kaleschke, L., Kreker, K., Klages, M., Soltwedel, T., Lorenzen, C., and Wegner, J.: Particle sedimentation patterns in the eastern Fram Strait during 2000–2005: Results from the Arctic long-term observatory HAUSGARTEN, Deep-Sea Research I, 56, 1471–1487, https://doi.org/10.1016/j.dsr.2009.04.011, 2009.

Bauerfeind, E., Nöthig, E.-M., Lorenzen, C., Knüppel, N., and Soltwedel, T.: Biogenic particle flux at AWI HAUSGARTEN central station HGIV from mooring FEVI13, https://doi.org/10.1594/PANGAEA.85546, 2015.

Behrenfeld, M. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, Limnology and Oceanography, 42, 1–20, https://doi.org/10.4319/lo.1997.42.1.0001, 1997.

Behrenfeld, M., Gaube, P., Della Penna, A., O'Malley, R. T., Burt, W. J., Hu, Y., Bontempi, P. S., Steinberg, D. K., Boss, E. S., Siegel, D. A., Hostetler, C. A., Tortell, P. D., and Doney, S. C.: Global satellite-observed daily vertical migrations of ocean animals, Nature, 576, 257–261, https://doi.org/10.1038/s41586-019-1796-9, 2019.

Berelson, W. M.: Particle settling rates increase with depth in the ocean, Deep-Sea Research II, 49, 237–251, https://doi.org/10.1016/S0967-0645(01)00102-3, 2002.

Bernardello, R., Marinov, I., Palter, J. B., Sarmiento, J. L., Galbraith, E. D., and Slater, R. D.: Response of the ocean natural carbon storage to projected twenty-first-century climate change, Journal of Climate, 27, 2033–2053, https://doi.org/10.1175/JCLI-D-13-00343.1, 2014.

Besiktepe, S. and Dam, H. G.: Coupling of ingestion and defecation as a function of diet in the calanoid copepod Acartia tonsa, Marine Ecology Progress Series, 229, 151–164, https://doi.org/10.3354/meps229151, 2002.

Betzer, P. R., Showers, W. J., Laws, E. A., Winn, C. D., DiTullio, G. R., and Kroopnick, P. M.: Primary productivity and particle fluxes on a transect of the equator at 153°W in the Pacific Ocean, Deep-Sea Research, 31, 1–11, https://doi.org/10.1016/0198-0149(84)90068-2, 1984.

Bianchi, D., Stock, C. A., Galbraith, E. D., and Sarmiento, J. L.: Diel vertical migration: Ecological controls and impacts on the biological pump in a one-dimensional ocean model, Global Biogeochemical Cycles, 27, 478–491, https://doi.org/10.1002/gbc.20031, 2013.

Biard, T., Stemmann, L., Picheral, M., Mayot, N., Vandromme, P., Hauss, H., Gorsky, G., Guidi, L., Kiko, R., and Not, F.: In situ imaging reveals the biomass of giant protists in the global ocean, Nature, 532, 504–507, https://doi.org/10.1038/nature17652, 2016.

Bidle, K. D. and Azam, F.: Accelerated dissolution of diatom silica by marine bacterial assemblages, Nature, 397, 508–512, https://doi.org/10.1038/17351, 1999.

Bisson, K. M., Siegel, D. A., DeVries, T., Cael, B. B., and Buesseler, K. O.: How data set characteristics influence ocean carbon export models, Global Biogeochemical Cycles, 32, 1312–1328, https://doi.org/10.1029/2018GB005934, 2018.

Boje, A., Akroyd, J., and Kraft, M.: A hybrid particle-number and particle model for efficient solution of population balance equations, Journal of Computational Physics, 389, 189–218, https://doi.org/10.1016/j.jcp.2019.03.033, 2019.

Bol, R., Henson, S. A., Rumyantseva, A., and Briggs, N.: High-frequency variability of small-particle carbon export flux in the Northeast Atlantic, Global Biogeochemical Cycles, 32, 1803–1814, https://doi.org/10.1029/2018GB005963, 2018.

Bowers, D. G., McKee, D., Jago, C. F., and Nimmo-Smith, W. A. M.: The area-to-mass ratio and fractal dimension of marine flocs, Estuarine, Coastal and Shelf Science, 189, 224–234, https://doi.org/10.1016/j.ecss.2017.03.026, 2017.

Boyd, P. W.: Toward quantifying the response of the oceans' biological pump to climate change, Frontiers in Marine Science, 2, 77, https://doi.org/10.3389/fmars.2015.00077, 2015.

Boyd, P. W. and Trull, T. W.: Understanding the export of biogenic particles in oceanic waters: Is there consensus?, Progress in Oceanography, 72, 276–312, https://doi.org/10.1016/j.pocean.2006.10.007, 2007.

Boyd, P. W., Claustre, H., Levy, M., Siegel, D. A., and Weber, T.: Multi-faceted particle pumps drive carbon sequestration in the ocean, Nature, 568, 327–335, https://doi.org/10.1038/s41586-019-1098-2, 2019.

Brewin, R. J. W., Sathyendranath, S., Platt, T., Bouman, H., Ciavatta, S., Dall'Olmo, G., Dingle, J., Groom, S. B., Bror, J., Kostadinov, T. S., Kulk, G., Laine, M., Martínez-Vicente, V., Psarra, S., Raitsos, D. E., Richardson, K., Salisbury, J., Shutler, J. D., and Walker, P.: Sensing the ocean biological carbon pump from space: A review of capabilities, concepts, research gaps and future developments, Earth-Science Reviews, 217, 103 604, https://doi.org/10.1016/j.earscirev.2021.103604, 2021.

Briggs, N., Dall'Olmo, G., and Claustre, H.: Major role of particle fragmentation in regulating biological sequestration of CO2 by the oceans, Science, 367, 791–793, https://doi.org/10.1126/science.aay1790, 2020.

Brock, T. D.: Calculating solar radiation for ecological studies, Ecological Modelling, 14, 1–19, https://doi.org/10.1016/0304-3800(81)90011-9, 1981.

Broecker, W. S.: Ocean chemistry during glacial time, Geochimica et Cosmochimica Acta, 46, 1689–1705, https://doi.org/10.1016/0016-7037(82)90110-7, 1982.

Brun, P., Payne, M. R., and Kiørboe, T.: Trait biogeography of marine copepods: An analysis across scales, Ecology Letters, 19, 1403–1413, https://doi.org/10.1111/ele.12688, 2016.

Brun, P., Stamieszkin, K., Visser, A. W., Licandro, P., Payne, M. R., and Kiørboe, T.: Climate change has altered zooplankton-fuelled carbon export in the North Atlantic, Nature Ecology and Evolution, 3, 416–423, https://doi.org/10.1038/s41559-018-0780-3, 2019.

Brzezinski, M. A.: The Si:C:N ratio of marine diatoms: interspecific variability and the effect of some environmental variables, Journal of Phycology, 21, 347–357, https://doi.org/10.1111/j.0022-3646.1985.00347.x, 1985.

Buesseler, K. O. and Boyd, P. W.: Shedding light on processes that control particle export and flux attenuation in the twilight zone of the open ocean, Limnology and Oceanography, 54, 1210–1232, https://doi.org/10.4319/lo.2009.54.4.1210, 2009.

Buesseler, K. O., Steinberg, D. K., Michaels, A. F., Johnson, R. J., Andrews, J. E., Valdes, J. R., and Price, J. F.: A comparison of the quantity and composition of material caught in a neutrally buoyant versus surface-tethered sediment trap, Deep-Sea Research I, 47, 277–294, https://doi.org/10.1016/S0967-0637(99)00056-4, 2000.

Buesseler, K. O., Benitez-Nelson, C. R., Van der Loeff, M. R., Andrews, J. E., Ball, L., Crossin, G., and Charette, M. A.: An intercomparison of small- and large-volume techniques for thorium-234 in seawater, Marine Chemistry, 74, 15–28, https://doi.org/10.1016/S0304-4203(00)00092-X, 2001.

Buesseler, K. O., Antia, A. N., Chen, M., Fowler, S. W., Gardner, W. D., Gustafsson, O., Harada, K., Michaels, A. F., Van der Loeff, M. R., Sarin, M., Steinberg, D. K., and Trull, T. W.: An assessment of the use of sediment traps for estimating upper ocean particle fluxes, Journal of Marine Research, 65, 345–416, https://doi.org/10.1357/002224007781567621, 2007a.

Buesseler, K. O., Lamborg, C. H., Boyd, P. W., Lam, P. J., Trull, T. W., Bidigare, R. R., Bishop, J. K. B., Casciotti, K. L., Dehairs, F., Elskens, M., Honda, M. C., Karl, D. M., Siegel, D. A., Silver, M. W., Steinberg, D. K., Valdes, J., Van Mooy, B. A. S., and Wilson, S.: Revisiting carbon flux through the ocean's twilight zone, Science, 316, 567–570, https://doi.org/10.1126/science.1137959, 2007b.

Buesseler, K. O., Lamborg, C. H., Cai, P., Escoube, R., Johnson, R., Pike, S., Masqué, P., McGillicuddy Jr., D. J., and Verdeny, E.: Particle fluxes associated with mesoscale eddies in the

Sargasso Sea, Deep-Sea Research II, 55, 1426–1444, https://doi.org/10.1016/j.dsr2.2008.02.007, 2008.

Buesseler, K. O., Benitez-Nelson, C. R., Roca-Martí, M., Wyatt, A. M., Resplandy, L., Clevenger, S. J., Drysdale, J. A., Estapa, M. L., Pike, S., and Umhau, B. P.: High-resolution spatial and temporal measurements of particulate organic carbon flux using thorium-234 in the northeast Pacific Ocean during the EXport Processes in the Ocean from RemoTe Sensing field campaign, Elementa: Science of the Anthropocene, 8, 1, https://doi.org/10.1525/elementa.030, 2020a.

Buesseler, K. O., Boyd, P. W., Black, E. E., and Siegel, D. A.: Metrics that matter for assessing the ocean biological carbon pump, Proceedings of the National Academy of Sciences, 117, 9679–9687, https://doi.org/10.1073/pnas.1918114117, 2020b.

Buitenhuis, E. T., Vogt, M., Moriarty, R., Bednaršek, N., Doney, S. C., Leblanc, K., Le Quéré, C., Luo, Y. W., O'Brien, C., O'Brien, T. D., Peloquin, J., Schiebel, R., and Swan, C.: MAREDAT: Towards a world atlas of MARine Ecosystem DATa, Earth System Science Data, 5, 227–239, https://doi.org/10.5194/essd-5-227-2013, 2013.

Burd, A. B. and Jackson, G. A.: Predicting particle coagulation and sedimentation rates for a pulsed input, Journal of Geophysical Research, 102, 10 545–10 561, https://doi.org/10.1029/96JC03592, 1997.

Burd, A. B. and Jackson, G. A.: Particle aggregation, Annual Review of Marine Science, 1, 65–90, https://doi.org/10.1146/annurev.marine.010908.163904, 2009.

Burd, A. B., Hansell, D. A., Steinberg, D. K., Anderson, T. R., Arístegui, J., Baltar, F., Beaupré, S. R., Buesseler, K. O., DeHairs, F., Jackson, G. A., Kadko, D. C., Koppelmann, R., Lampitt, R. S., Nagata, T., Reinthaler, T., Robinson, C., Robison, B. H., Tamburini, C., and Tanaka, T.: Assessing the apparent imbalance between geochemical and biochemical indicators of meso- and bathypelagic biological activity: What the @$#! is wrong with present calculations of carbon budgets?, Deep-Sea Research II, 57, 1557–1571, https://doi.org/10.1016/j.dsr2.2010.02.022, 2010.

Cael, B. B. and Bisson, K. M.: Particle flux parameterizations: Quantitative and mechanistic similarities and differences, Frontiers in Marine Science, 5, 395, https://doi.org/10.3389/fmars.2018.00395, 2018.

Callendar, G.: The artificial production of carbon dioxide and its influence on temperature, Quarterly Journal of the Royal Meteorological Society, pp. 223–240, 1938.

Campoverde, N. C. G., Hassenrück, C., Buttigieg, P. L., and Gärdes, A.: Characterization of bacterioplankton communities and quantification of organic carbon pools off the Galapagos Archipelago under contrasting environmental conditions, PeerJ, 6, e5984, https://doi.org/10.7717/peerj.5984, 2018.

Carr, M.-E.: Estimation of potential productivity in Eastern Boundary Currents using remote sensing, Deep-Sea Research II, 49, 59–80, https://doi.org/10.1016/S0967-0645(01)00094-7, 2002.

Carr, M.-E., Friedrichs, M. A. M., Schmeltz, M., Noguchi-Aita, M., Antoine, D., Arrigo, K. R., Asanuma, I., Aumont, O., Barber, R. T., Behrenfeld, M., Bidigare, R. R., Buitenhuis, E. T., Campbell, J., Ciotti, A., Dierssen, H., Dowell, M., Dunne, J. P., Esaias, W., Gentili, B., Gregg, W., Groom, S. B., Hoepffner, N., Ishizaka, J., Kameda, T., Le Quéré, C., Lohrenz, S., Marra, J. F., Mélin, F., Moore, J. K., Morel, A., Reddy, T. E., Ryan, J., Scardi, M., Smyth, T., Turpie, K., Tilstone, G., Waters, K., and Yamanaka, Y.: A comparison of global estimates of marine primary production from ocean color, Deep-Sea Research II, 53, 741–770, https://doi.org/10.1016/j.dsr2.2006.01.028, 2006.

Cartis, C., Fiala, J., Marteau, B., and Roberts, L.: Improving the flexibility and robustness of model-based derivative-free optimization solvers, ACM Transactions on Mathematical Software, 45, 32, https://doi.org/10.1145/3338517, 2019.

Catling, D. C. and Zahnle, K. J.: The Archean atmosphere, Science Advances, 6, https://doi.org/10.1126/sciadv.aax1420, 2020.

Cavallaro, N., Shrestha, G., Birdsey, R., Mayes, M. A., Najjar, R. G., Reed, S. C., Romero-Lankao, P., and Zhu, Z.: Second State of the Carbon Cycle Report (SOCCR2): A sustained assessment report, Tech. rep., Washington, DC, USA, https://doi.org/10.7930/Soccr2.2018, 2018.

Cavan, E. L., Le Moigne, F. A. C., Poulton, A. J., Tarling, G. A., Ward, P., Daniels, C. J., Fragoso, G. M., and Sanders, R. J.: Attenuation of particulate organic carbon flux in the Scotia Sea,

Southern Ocean, is controlled by zooplankton fecal pellets, Geophysical Research Letters, 42, 821–830, https://doi.org/10.1002/2014GL062744, 2015.

Cavan, E. L., Trimmer, M., Shelley, F., and Sanders, R. J.: Remineralization of particulate organic carbon in an ocean oxygen minimum zone, Nature Communications, 8, 14 847, https://doi.org/10.1038/ncomms14847, 2017.

Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., Behrenfeld, M., Boetius, A., Boyd, P. W., Classen, A. T., Crowther, T. W., Danovaro, R., Foreman, C. M., Huisman, J., Hutchins, D. A., Jansson, J. K., Karl, D. M., Koskella, B., Welch, D. B. M., Martiny, J. B., Moran, M. A., Orphan, V. J., Reay, D. S., Remais, J. V., Rich, V. I., Singh, B. K., Stein, L. Y., Stewart, F. J., Sullivan, M. B., Van Oppen, M. J. H., Weaver, S. C., Webb, E. A., and Webster, N. S.: Scientists' warning to humanity: Microorganisms and climate change, Nature Reviews Microbiology, 17, 569–586, https://doi.org/10.1038/s41579-019-0222-5, 2019.

Ceballos-Romero, E., Le Moigne, F. A. C., Henson, S. A., Marsay, C. M., Sanders, R. J., García-Tenorio, R., and Villa-Alfageme, M.: Influence of bloom dynamics on particle export efficiency in the North Atlantic: A comparative study of radioanalytical techniques and sediment traps, Marine Chemistry, 186, 198–210, https://doi.org/10.1016/j.marchem.2016.10.001, 2016.

Charette, M. A., Bradley Moran, S., and Bishop, J. K. B.: 234Th as a tracer of particulate organic carbon export in the subarctic northeast Pacific Ocean, Deep-Sea Research II, 46, 2833–2861, https://doi.org/10.1016/S0967-0645(99)00085-5, 1999.

Collins, J. R., Edwards, B. R., Thamatrakoln, K., Ossolinski, J. E., DiTullio, G. R., Bidle, K. D., Doney, S. C., and Van Mooy, B. A. S.: The multiple fates of sinking particles in the North Atlantic Ocean, Global Biogeochemical Cycles, 29, 1471–1494, https://doi.org/10.1111/1462-2920.13280, 2015.

Conte, M. H. and Weber, J. C.: Particle flux in the deep Sargasso Sea: The 35-year Oceanic Flux Program time series, Oceanography, 27, 142–147, https://doi.org/10.5670/oceanog.2014.17., 2014.

Conte, M. H., Ralph, N., and Ross, E. H.: Seasonal and interannual variability in deep ocean particle fluxes at the Oceanic Flux Program (OFP)/Bermuda Atlantic Time Series (BATS) site in the western Sargasso Sea near Bermuda, Deep-Sea Research II, 48, 1471–1505, https://doi.org/10.1016/S0967-0645(00)00150-8, 2001.

Cram, J. A., Weber, T., Leung, S. W., McDonnell, A. M. P., Liang, J. H., and Deutsch, C.: The role of particle size, ballast, temperature, and oxygen in the sinking flux to the deep sea, Global Biogeochemical Cycles, 32, 858–876, https://doi.org/10.1029/2017GB005710, 2018.

Dall'Olmo, G., Dingle, J., Polimene, L., Brewin, R. J., and Claustre, H.: Substantial energy input to the mesopelagic ecosystem from the seasonal mixed-layer pump, Nature Geoscience, 9, 820–823, https://doi.org/10.1038/ngeo2818, 2016.

Dam, H. G. and Drapeau, D. T.: Coagulation efficiency, organic-matter glues and the dynamics of particles during a phytoplankton bloom in a mesocosm study, Deep-Sea Research II, 42, 111–123, https://doi.org/10.1016/0967-0645(95)00007-D, 1995.

Daniels, C. J., Sheward, R. M., and Poulton, A. J.: Biogeochemical implications of comparative growth rates of Emiliania huxleyi and Coccolithus species, Biogeosciences, 11, 6915–6925, https://doi.org/10.5194/bg-11-6915-2014, 2014.

De Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology, Journal of Geophysical Research, 109, C12 003, https://doi.org/10.1029/2004JC002378, 2004.

De La Rocha, C. L. and Passow, U.: Factors influencing the sinking of POC and the efficiency of the biological carbon pump, Deep-Sea Research II, 54, 639–658, https://doi.org/10.1016/j.dsr2.2007.01.004, 2007.

DeVries, T. and Weber, T.: The export and fate of organic matter in the ocean: New constraints from combining satellite and oceanographic tracer observations, Global Biogeochemical Cycles, 31, 535–555, https://doi.org/10.1002/2016GB005551, 2017.

DeVries, T., Primeau, F. W., and Deutsch, C.: The sequestration efficiency of the biological pump, Geophysical Research Letters, 39, L13 601, https://doi.org/10.1029/2012GL051963, 2012.

DeVries, T., Liang, J. H., and Deutsch, C.: A mechanistic particle flux model applied to the oceanic

phosphorus cycle, Biogeosciences, 11, 5381–5398, https://doi.org/10.5194/bg-11-5381-2014, 2014.

Doney, S. C., Kleypas, J. A., Sarmiento, J. L., and Falkowski, P. G.: The U.S. JGOFS Synthesis and Modeling Project: An introduction, Deep-Sea Research II, 49, 1–20, https://doi.org/10.1016/S0967-0645(01)00092-3, 2002.

Dörgens, A. L., Ahmerkamp, S., Müssig, J., Stocker, R., Kuypers, M. M., Khalili, A., and Kindler, K.: A laboratory model of marine snow: Preparation and characterization of porous fiber particles, Limnology and Oceanography: Methods, 13, 664–671, https://doi.org/10.1002/lom3.10056, 2015.

Dunne, J. P., Armstrong, R. A., Gnanadesikan, A., and Sarmiento, J. L.: Empirical and mechanistic models for the particle export ratio, Global Biogeochemical Cycles, 19, GB4026, https://doi.org/10.1029/2004GB002390, 2005.

Dunne, J. P., Sarmiento, J. L., and Gnanadesikan, A.: A synthesis of global particle export from the surface ocean and cycling through the ocean interior and on the seafloor, Global Biogeochemical Cycles, 21, GB4006, https://doi.org/10.1029/2006GB002907, 2007.

Dunne, J. P., Gnanadesikan, A., Sarmiento, J. l., Deutsch, C., and Slater, R. D.: Technical description of the prototype version (v0) of Tracers Of Phytoplankton with Allometric Zooplankton (TOPAZ) ocean biogeochemical model as used in the Princeton IFMIP model, Biogeosciences, 7, 2010.

Dunne, J. P., John, J. G., Shevliakova, S., Stouffer, R. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Sentman, L. T., Adcroft, A. J., Cooke, W., Dunne, K. A., Griffies, S. M., Hallberg, R. W., Harrison, M. J., Levy, H., Wittenberg, A. T., Phillips, P. J., and Zadeh, N.: GFDL's ESM2 global coupled climate-carbon earth system models. Part II: Carbon system formulation and baseline simulation characteristics, Journal of Climate, 26, 2247–2267, https://doi.org/10.1175/JCLI-D-12-00150.1, 2013.

Edwards, K. F., Thomas, M. K., Klausmeier, C. A., and Litchman, E.: Light and growth in marine phytoplankton: Allometric, taxonomic, and environmental variation, Limnology and Oceanography, 60, 540–552, https://doi.org/10.1002/lno.10033, 2015.

Engel, A.: The role of transparent exopolymer particles (TEP) in the increase in apparent particle stickiness (alpha) during the decline of a diatom bloom, Journal of Plankton Research, 22, 485–497, https://doi.org/10.1093/plankt/22.3.485, 2000.

Engel, A. and Schartau, M.: Influence of transparent exopolymer particles (TEP) on sinking velocity of Nitzschia closterium aggregates, Marine Ecology Progress Series, 182, 69–76, https://doi.org/10.3354/MEPS182069, 1999.

Eppley, R. W.: Temperature and phytoplankton growth in the sea, Fishery Bulletin, 70, 1063–1085, 1972.

Estapa, M., Durkin, C. A., Buesseler, K. O., Johnson, R., and Feen, M. L.: Carbon flux from bio-optical profiling floats: Calibrating transmissometers for use as optical sediment traps, Deep-Sea Research I, 120, 100–111, https://doi.org/10.1016/j.dsr.2016.12.003, 2017.

Estapa, M. L., Feen, M. L., and Breves, E.: Direct observations of biological carbon export from profiling floats in the subtropical North Atlantic, Global Biogeochemical Cycles, 33, 282–300, https://doi.org/10.1029/2018GB006098, 2019.

Estapa, M. L., Buesseler, K. O., Durkin, C. A., Omand, M. M., Breves, E., Kelly, R. P., Benitez-Nelson, C. R., and Pike, S.: Biogenic sinking particle fluxes and sediment trap collection efficiency at Ocean Station Papa, Elementa: Science of the Anthropocene, 9, 0012, https://doi.org/10.1525/elementa.2020.00122, 2021.

Evans, G. T.: Defining misfit between biogeochemical models and data sets, Journal of Marine Systems, 40-41, 49–54, https://doi.org/10.1016/S0924-7963(03)00012-5, 2003.

Fakhraee, M., Planavsky, N. J., and Reinhard, C. T.: The role of environmental factors in the long-term evolution of the marine biological pump, Nature Geoscience, 13, 812–816, https://doi.org/10.1038/s41561-020-00660-6, 2020.

Fasham, M. J., Ducklow, H. W., and McKelvie, S. M.: A nitrogen-based model of plankton dynamics in the oceanic mixed layer, Journal of Marine Research, 48, 591–639, https://doi.org/10.1357/002224090784984678, 1990.

Fennel, K., Losch, M., Schroter, J., and Wenzel, M.: Testing a marine ecosystem model: Sensitivity analysis and parameter optimization, Journal of Marine Systems, 28, 45–63, https://doi.org/10.1016/S0924-7963(00)00083-X, 2001.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P.: Primary production of the biosphere: Integrating terrestrial and oceanic components, Science, 281, 237–240, https://doi.org/10.1126/science.281.5374.237, 1998.

Flynn, K. J.: Do we need complex mechanistic photoacclimation models for phytoplankton?, Limnology and Oceanography, 48, 2243–2249, https://doi.org/10.4319/lo.2003.48.6.2243, 2003.

Flynn, K. J. and Raven, J. A.: What is the limit for photoautotrophic plankton growth rates?, Journal of Plankton Research, 39, 13–22, https://doi.org/10.1093/plankt/fbw067, 2017.

Follows, M. J., Dutkiewicz, S., Grant, S., and Chisholm, S. W.: Emergent biogeography of microbial communities in a model ocean, Science, 315, 1843–1846, https://doi.org/10.1126/science.1138544, 2007.

Foster, G. L., Royer, D. L., and Lunt, D. J.: Future climate forcing potentially without precedent in the last 420 million years, Nature Communications, 8, https://doi.org/10.1038/ncomms14845, 2017.

Francois, R., Honjo, S., Krishfield, R. A., and Manganini, S. J.: Factors controlling the flux of organic carbon to the bathypelagic zone of the ocean, Global Biogeochemical Cycles, 16, 1087, https://doi.org/10.1029/2001gb001722, 2002.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R., Alin, S., Aragão, L. E. O. C., Arneth, A., Arora, V., Bates, N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., Van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, X., and Zaehle, S.: Global Carbon Budget 2020, Earth System Science Data, 12, 3269–3340, https://doi.org/10.5194/essd-12-3269-2020, 2020.

Friedrichs, M. A. M., Hood, R. R., and Wiggert, J. D.: Ecosystem model complexity versus physical forcing: Quantification of their relative impact with assimilated Arabian Sea data, Deep-Sea Research II, 53, 576–600, https://doi.org/10.1016/j.dsr2.2006.01.026, 2006.

Friedrichs, M. A. M., Dusenberry, J. A., Anderson, L. A., Armstrong, R. A., Chai, F., Christian, J. R., Doney, S. C., Dunne, J. P., Fujii, M., Hood, R. R., McGillicuddy Jr., D. J., Moore, J. K., Schartau, M., Spitz, Y. H., and Wiggert, J. D.: Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups, Journal of Geophysical Research, 112, C08 001, https://doi.org/10.1029/2006JC003852, 2007.

Frost, B. W.: The role of grazing in nutrient-rich areas of the open sea, Limnology and Oceanography, 36, 1616–1630, https://doi.org/10.4319/lo.1991.36.8.1616, 1991.

Galbraith, E. D., Gnanadesikan, A., Dunne, J. P., and Hiscock, M. R.: Regional impacts of iron-light colimitation in a global biogeochemical model, Biogeosciences, 7, 1043–1064, https://doi.org/10.5194/bg-7-1043-2010, 2010.

García, H. E., Boyer, T. P., Locarnini, R. A., Antonov, J. I., Mishonov, A. V., Baranova, O. K., Zweng, M. M., Reagan, J. R., and Johnson, D. R.: World Ocean Atlas 2013. Volume 3: Dissolved oxygen, apparent oxygen utilization, and oxygen saturation, Tech. rep., National Oceanic and Atmospheric Administration, 2013a.

García, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O. K., Zweng, M. M., Reagan, J. R., and Johnson, D. R.: World Ocean Atlas 2013. Volume 4: Dissolved inorganic nutrients (phosphate, nitrate, silicate), Tech. rep., National Oceanic and Atmospheric Administration, 2013b.

García-Robledo, E., Borisov, S., Klimant, I., and Revsbech, N. P.: Determination of respiration rates in water with sub-micromolar oxygen concentrations, Frontiers in Marine Science, 3, 244, https://doi.org/10.3389/fmars.2016.00244, 2016.

Gehlen, M., Bopp, L., Emprin, N., Aumont, O., Heinze, C., and Ragueneau, O.: Reconciling surface ocean productivity, export fluxes and sediment composition in a global biogeochemical ocean model, Biogeosciences, 3, 521–537, https://doi.org/10.5194/bg-3-521-2006, 2006.

Gehlen, M., Gangstø, R., Schneider, B., Bopp, L., Aumont, O., and Ethé, C.: The fate of pelagic CaCO3 production in a high CO2 ocean: A model study, Biogeosciences, 4, 505–519, https://doi.org/10.5194/bg-4-505-2007, 2007.

Geider, R. J., MacIntyre, H. L., and Kana, T. M.: Dynamic model of phytoplankton growth and acclimation: Responses of the balanced growth rate and the chlorophyll a:carbon ratio to light, nutrient-limitation and temperature, Marine Ecology Progress Series, 148, 187–200, https://doi.org/10.3354/meps148187, 1997.

Giering, S. L. C., Sanders, R. J., Lampitt, R. S., Anderson, T. R., Tamburini, C., Boutrif, M., Zubkov, M. V., Marsay, C. M., Henson, S. A., Saw, K., Cook, K., and Mayor, D. J.: Reconciliation of the carbon budget in the ocean's twilight zone, Nature, 507, 480–483, https://doi.org/10.1038/nature13123, 2014.

Giering, S. L. C., Sanders, R. J., Martin, A. P., Riley, J. S., Marsay, C. M., and Johns, D. G.: Particle flux in the oceans: Challenging the steady state assumption, Global Biogeochemical Cycles, 31, 159–171, https://doi.org/10.1002/2016GB005424, 2017.

Giering, S. L. C., Cavan, E. L., Basedow, S. L., Briggs, N., Burd, A. B., Darroch, L. J., Guidi, L., Irisson, J. O., Iversen, M. H., Kiko, R., Lindsay, D., Marcolin, C. R., McDonnell, A. M. P., Möller, K. O., Passow, U., Thomalla, S. J., Trull, T. W., and Waite, A. M.: Sinking organic particles in the ocean: Flux estimates from in situ optical devices, Frontiers in Marine Science, 6, 834, https://doi.org/10.3389/fmars.2019.00834, 2020.

Gillespie, D. T.: An exact method for numerically simulating the stochastic coalescence process in a cloud, Journal of the Atmospheric Sciences, 32, 1977–1989, 1975.

Gloege, L., McKinley, G. A., Mouw, C. B., and Ciochetto, A. B.: Global evaluation of particulate organic carbon flux parameterizations and implications for atmospheric pCO2, Global Biogeochemical Cycles, 31, 1192–1215, https://doi.org/10.1002/2016GB005535, 2017.

Gnanadesikan, A., Dunne, J. P., Key, R. M., Matsumoto, K., Sarmiento, J. L., Slater, R. D., and Swathi, P. S.: Oceanic ventilation and biogeochemical cycling: Understanding the physical mechanisms that produce realistic distributions of tracers and productivity, Global Biogeochemical Cycles, 18, GB4010, https://doi.org/10.1029/2003GB002097, 2004.

Goldthwait, S. A., Yen, J., Brown, J., and Alldredge, A. L.: Quantification of marine snow fragmentation by swimming euphausiids, Limnology and Oceanography, 49, 940–952, https://doi.org/10.4319/lo.2004.49.4.0940, 2004.

Grabowski, W. W., Morrison, H., Shima, S. I., Abade, G. C., Dziekan, P., and Pawlowska, H.: Modeling of cloud microphysics: Can we do better?, Bulletin of the American Meteorological Society, 100, 655–672, https://doi.org/10.1175/BAMS-D-18-0005.1, 2019.

Gregg, W., Ginoux, P., Schopf, P. S., and Casey, N. W.: Phytoplankton and iron: Validation of a global three-dimensional ocean biogeochemical model, Deep-Sea Research II, 50, 3143–3169, https://doi.org/10.1016/j.dsr2.2003.07.013, 2003.

Gremion, G., Nadeau, L.-P., Dufresne, C., Schloss, I. R., Archambault, P., and Dumont, D.: A discrete interaction numerical model for coagulation and fragmentation of marine detritic particulate matter (Coagfrag v.1), Geoscientific Model Development, 14, 4535–4554, https://doi.org/10.5194/gmd-14-4535-2021, 2021.

Gruber, N., Clement, D., Carter, B. R., Feely, R. A., Van Heuven, S., Hoppema, M., Ishii, M., and Key, R. M.: The oceanic sink for anthropogenic CO2 from 1994 to 2007, Science, 363, 1193–1199, https://doi.org/10.1126/science.aau5153, 2019.

Guidi, L., Jackson, G. A., Stemmann, L., Miquel, J. C., Picheral, M., and Gorsky, G.: Relationship between particle size distribution and flux in the mesopelagic zone, Deep-Sea Research I, 55, 1364–1374, https://doi.org/10.1016/j.dsr.2008.05.014, 2008.

Guidi, L., Stemmann, L., Jackson, G. A., Ibañez, F., Claustre, H., Legendre, L., Picheral, M., and Gorsky, G.: Effects of phytoplankton community on production, size and export

of large aggregates: A world-ocean analysis, Limnology and Oceanography, 54, 1951–1963, https://doi.org/10.4319/lo.2009.54.6.1951, 2009.

Guidi, L., Legendre, L., Reygondeau, G., Uitz, J., Stemmann, L., and Henson, S. A.: A new look at ocean carbon remineralization for estimating deepwater sequestration, Global Biogeochemical Cycles, 29, 1044–1059, https://doi.org/10.1002/2014GB005063, 2015.

Hansen, A. N. and Visser, A. W.: Carbon export by vertically migrating zooplankton: An optimal behavior model, Limnology and Oceanography, 61, 701–710, https://doi.org/10.1002/lno.10249, 2016.

Hartman, S. E., Lampitt, R. S., Larkin, K. E., Pagnani, M., Campbell, J., Gkritzalis, T., Jiang, Z.-P., Pebody, C. A., Ruhl, H. A., Gooday, A. J., Bett, B. J., Billett, D. S. M., Provost, P., McLachlan, R., Turton, J. D., and Lankester, S.: The Porcupine Abyssal Plain fixed-point sustained observatory (PAP-SO): Variations and trends from the Northeast Atlantic fixed-point time-series, ICES Journal of Marine Science, 69, 776–783, https://doi.org/10.1093/icesjms/fss077, 2012.

Hayes, C. T., Black, E. E., Anderson, R. F., Baskaran, M., Buesseler, K. O., Charette, M. A., Cheng, H., Cochran, J. K., Edwards, R. L., Fitzgerald, P., Lam, P. J., Lu, Y., Morris, S. O., Ohnemus, D. C., Pavia, F. J., Stewart, G., and Tang, Y.: Flux of particulate elements in the North Atlantic ocean constrained by multiple radionuclides, Global Biogeochemical Cycles, 32, 1738–1758, https://doi.org/10.1029/2018GB005994, 2018.

Heinze, C., Meyer, S., Goris, N., Anderson, L., Steinfeldt, R., Chang, N., Le Quéré, C., and Bakker, D. C. E.: The ocean carbon sink: Impacts, vulnerabilities and challenges, Earth System Dynamics, 6, 327–358, https://doi.org/10.5194/esd-6-327-2015, 2015.

Hellweger, F. L., Van Sebille, E., and Fredrick, N. D.: Biogeographic patterns in ocean microbes emerge in a neutral agent-based model, Science, 345, 1346–1349, https://doi.org/10.1126/science.1254421, 2014.

Hemmings, J. C. P. and Challenor, P. G.: 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, 471–498, https://doi.org/10.5194/gmd-5-471-2012, 2012.

Hengenius, J., Gribskov, M., Rundell, A., and Umulis, D.: Making models match measurements: Model optimization for morphogen patterning networks, Seminars in Cell and Developmental Biology, 35, 109–123, https://doi.org/10.1016/j.semcdb.2014.06.017.Making, 2014.

Henson, S. A., Sanders, R. J., Madsen, E., Morris, P. J., Le Moigne, F. A. C., and Quartly, G. D.: A reduced estimate of the strength of the ocean's biological carbon pump, Geophysical Research Letters, 38, L04 606, https://doi.org/10.1029/2011GL046735, 2011.

Henson, S. A., Sanders, R. J., and Madsen, E.: Global patterns in efficiency of particulate organic carbon export and transfer to the deep ocean, Global Biogeochemical Cycles, 26, GB1028, https://doi.org/10.1029/2011GB004099, 2012.

Henson, S. A., Le Moigne, F. A. C., and Giering, S. L. C.: Drivers of carbon export efficiency in the global ocean, Global Biogeochemical Cycles, 33, 891–903, https://doi.org/10.1029/2018GB006158, 2019.

Hernández-León, S. and Ikeda, T.: A global assessment of mesozooplankton respiration in the ocean, Journal of Plankton Research, 27, 153–158, https://doi.org/10.1093/plankt/fbh166, 2005.

Herndl, G. J. and Reinthaler, T.: Microbial control of the dark end of the biological pump, Nature Geoscience, 6, 718–724, https://doi.org/10.1038/ngeo1921, 2013.

Hill, P. S.: Reconciling aggregation theory with observed vertical fluxes following phytoplankton blooms, Journal of Geophysical Research, 97, 2295–2308, https://doi.org/10.1029/91JC02808, 1992.

Hill, P. S., Syvitski, J. P., Cowan, E. A., and Powell, R. D.: In situ observations of floe settling velocities in Glacier Bay, Alaska, Marine Geology, 145, 85–94, https://doi.org/10.1016/S0025-3227(97)00109-6, 1998.

Honjo, S., Dymond, J., Collier, R., and Manganini, S. J.: Export production of particles to the interior of the equatorial Pacific Ocean during the 1992 EqPac experiment, Deep-Sea Research II, 42, 831–870, https://doi.org/10.1016/0967-0645(95)00034-N, 1995.

Honjo, S., Manganini, S. J., Krishfield, R. A., and Francois, R.: Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: a synthesis of global sediment trap programs since 1983, Progress in Oceanography, 76, 217–285, https://doi.org/10.1016/j.pocean.2007.11.003, 2008.

Ikeda, T.: Respiration and ammonia excretion by marine metazooplankton taxa: synthesis toward a global-bathymetric model, Marine Biology, 161, 2753–2766, https://doi.org/10.1007/s00227-014-2540-5, 2014.

Irigoien, X.: Gut clearance rate constant, temperature and initial gut contents: A review, Journal of Plankton Research, 20, 997–1003, https://doi.org/10.1093/plankt/20.5.997, 1998.

Irwin, A. J., Finkel, Z. V., Schofield, O. M., and Falkowski, P. G.: Scaling-up from nutrient physiology to the size-structure of phytoplankton communities, Journal of Plankton Research, 28, 459–471, https://doi.org/10.1093/plankt/fbi148, 2006.

Ito, T. and Follows, M. J.: Preformed phosphate, soft tissue pump and atmospheric CO2, Journal of Marine Research, 63, 813–839, https://doi.org/10.1357/0022240054663231, 2005.

Iversen, M. H. and Lampitt, R. S.: Size does not matter after all: No evidence for a size-sinking relationship for marine snow, Progress in Oceanography, 189, 102 445, https://doi.org/10.1016/j.pocean.2020.102445, 2020.

Iversen, M. H. and Ploug, H.: Ballast minerals and the sinking carbon flux in the ocean: Carbon-specific respiration rates and sinking velocity of marine snow aggregates, Biogeosciences, 7, 2613–2624, https://doi.org/10.5194/bg-7-2613-2010, 2010.

Iversen, M. H. and Ploug, H.: Temperature effects on carbon-specific respiration rate and sinking velocity of diatom aggregates: Potential implications for deep ocean export processes, Biogeosciences, 10, 4073–4085, https://doi.org/10.5194/bg-10-4073-2013, 2013.

Iversen, M. H., Nowald, N., Ploug, H., Jackson, G. A., and Fischer, G.: High resolution profiles of vertical particulate organic matter export off Cape Blanc, Mauritania: Degradation processes and ballasting effects, Deep-Sea Research I, 57, 771–784, https://doi.org/10.1016/j.dsr.2010.03.007, 2010.

Jackson, G. A.: A model of the formation of marine algal flocs by physical coagulation processes, Deep-Sea Research, 37, 1197–1211, https://doi.org/10.1016/0198-0149(90)90038-W, 1990.

Jackson, G. A.: Using fractal scaling and two-dimensional particle size spectra to calculate coagulation rates for heterogeneous systems, Journal of Colloid and Interface Science, 202, 20–29, https://doi.org/10.1006/jcis.1998.5435, 1998.

Jackson, G. A.: Effect of coagulation on a model planktonic food web, Deep-Sea Research I, 48, 95–123, https://doi.org/10.1016/S0967-0637(00)00040-6, 2001.

Jackson, G. A. and Burd, A. B.: Simulating aggregate dynamics in ocean biogeochemical models, Progress in Oceanography, 133, 55–65, https://doi.org/10.1016/j.pocean.2014.08.014, 2015.

Jackson, G. A., Maffione, R., Costello, D. K., Alldredge, A. L., Logan, B. E., and Dam, H. G.: Particle size spectra between 1 μm and 1 cm at Monterey Bay determined using multiple instruments, Deep-Sea Research I, 44, 1739–1767, https://doi.org/10.1016/S0967-0637(97)00029-0, 1997.

Jansen, H.: Modelling the marine carbonate pump and its implications on the atmospheric CO2 concentration, Doctoral thesis, Universität Bremen, https://doi.org/10.23689/fidgeo-218, 2001.

Jansen, H. and Wolf-Gladrow, D. A.: Carbonate dissolution in copepod guts: A numerical model, Marine Ecology Progress Series, 221, 199–207, https://doi.org/10.3354/meps221199, 2001.

Jansen, H., Zeebe, R. E., and Wolf-Gladrow, D. A.: Modeling the dissolution of settling CaCO3 in the ocean, Global Biogeochemical Cycles, 16, 1027, https://doi.org/10.1029/2000gb001279, 2002.

Jickells, T. D. and Moore, C. M.: The importance of atmospheric deposition for ocean productivity, Annual Review of Ecology, Evolution, and Systematics, 46, 481–501, https://doi.org/10.1146/annurev-ecolsys-112414-054118, 2015.

Jin, D., Hoagland, P., and Buesseler, K. O.: The value of scientific research on the ocean's biological carbon pump, Science of the Total Environment, 749, 141 357, https://doi.org/10.1016/j.scitotenv.2020.141357, 2020.

Jin, X., Gruber, N., Dunne, J. P., Sarmiento, J. L., and Armstrong, R. A.: Diagnosing the contributions of phytoplankton functional groups to the production and export of particulate

organic carbon, CaCO3, and opal from global nutrient and alkalinity distributions, Global Biogeochemical Cycles, 20, GB2015, https://doi.org/10.1029/2005GB002532, 2006.

Jokulsdottir, T.: Sinking biological aggregates in the ocean: A modeling study, Doctor of philosophy, The University of Chicago, 2011.

Jokulsdottir, T. and Archer, D. E.: A stochastic, Lagrangian model of sinking biogenic aggregates in the ocean (SLAMS 1.0): Model formulation, validation and sensitivity, Geoscientific Model Development, 9, 1455–1476, https://doi.org/10.5194/gmd-9-1455-2016, 2016.

Journet, E., Desboeufs, K. V., Caquineau, S., and Colin, J. L.: Mineralogy as a critical factor of dust iron solubility, Geophysical Research Letters, 35, L07 805, https://doi.org/10.1029/2007GL031589, 2008.

Jumars, P. A., Penry, D. H., Baross, J. A., Perry, M. J., and Frost, B. W.: Closing the microbial loop: dissolved carbon pathway to heterotrophic bacteria from incomplete ingestion, digestion and absorption in animals, Deep-Sea Research, 36, 483–495, https://doi.org/10.1016/0198-0149(89)90001-0, 1989.

Karakaş, G., Nowald, N., Schäfer-Neth, C., Iversen, M. H., Barkmann, W., Fischer, G., Marchesiello, P., and Schlitzer, R.: Impact of particle aggregation on vertical fluxes of organic matter, Progress in Oceanography, 83, 331–341, https://doi.org/10.1016/j.pocean.2009.07.047, 2009.

Kasting, J. F.: Earth's early atmosphere, Science, 259, 920–926, https://doi.org/10.1126/science.11536547, 1993.

Kawakami, H., Honda, M. C., Matsumoto, K., Fujiki, T., and Watanabe. Shuichi: East-West distribution of POC fluxes estimated from 234Th in the northern North Pacific in autumn, Journal of Oceanography, 66, 71–83, https://doi.org/10.1007/s10872-010-0006-z, 2010.

Khatiwala, S., Tanhua, T., Mikaloff Fletcher, S., Gerber, M., Doney, S. C., Graven, H. D., Gruber, N., McKinley, G. A., Murata, A., Ríos, A. F., and Sabine, C. L.: Global ocean storage of anthropogenic carbon, Biogeosciences, 10, 2169–2191, https://doi.org/10.5194/bg-10-2169-2013, 2013.

Khatiwala, S., Schmittner, A., and Muglia, J.: Air-sea disequilibrium enhances ocean carbon storage during glacial periods, Science Advances, 5, https://doi.org/10.1126/sciadv.aaw4981, 2019.

Khelifa, A. and Hill, P. S.: Models for effective density and settling velocity of flocs, Journal of Hydraulic Research, 44, 390–401, https://doi.org/10.1080/00221686.2006.9521690, 2006.

Kilps, J. R.: Fractal dimensions of aggregates formed under natural and engineered fluid environments, Master of science, University of Arizona, 1993.

Kim, G. and Church, T. M.: Seasonal biogeochemical fluxes of 234Th and 210Po in the upper Sargasso Sea: Influence from atmosphere iron deposition, Global Biogeochemical Cycles, 15, 651–661, https://doi.org/10.1029/2000GB001313, 2001.

Kiørboe, T.: Small-scale turbulence, marine snow formation, and planktivorous feeding, Scientia Marina, 61, 141–158, 1997.

Kiørboe, T.: How zooplankton feed: Mechanisms, traits and trade-offs, Biological Reviews, 86, 311–339, https://doi.org/10.1111/j.1469-185X.2010.00148.x, 2011.

Kiørboe, T.: Zooplankton body composition, Limnology and Oceanography, 58, 1843–1850, https://doi.org/10.4319/lo.2013.58.5.1843, 2013.

Kiørboe, T. and Hansen, J. L. S.: Phytoplankton aggregate formation: Observations of patterns and mechanisms of cell sticking and the significance of exopolymeric material, Journal of Plankton Research, 15, 993–1018, https://doi.org/10.1093/plankt/15.9.993, 1993.

Kirchman, D. L., Morán, X. A. G., and Ducklow, H. W.: Microbial growth in the polar oceans - Role of temperature and potential impact of climate change, Nature Reviews Microbiology, 7, 451–459, https://doi.org/10.1038/nrmicro2115, 2009.

Klaas, C. M. and Archer, D. E.: Association of sinking organic matter with various types of mineral ballast in the deep sea: Implications for the rain ratio, Global Biogeochemical Cycles, 16, 1116, https://doi.org/10.1029/2001gb001765, 2002.

Kohfeld, K. E. and Ridgwell, A.: Glacial-interglacial variability in atmospheric CO2, in: Surface ocean: Lower atmosphere processes, edited by Le Quéré, C. and Saltzman, E. S., pp. 251–286,

Geophysical Monograph Series, AGU, Washington, DC, https://doi.org/10.1029/2008GM000845, 2009.

Kostadinov, T. S., Siegel, D. A., and Maritorena, S.: Retrieval of the particle size distribution from satellite ocean color observations, Journal of Geophysical Research, 114, C09 015, https://doi.org/10.1029/2009JC005303, 2009.

Kostadinov, T. S., Siegel, D. A., and Maritorena, S.: Global variability of phytoplankton functional types from space: assessment via the particle size distribution, Biogeosciences, 7, 3239–3257, https://doi.org/10.5194/bg-7-3239-2010, 2010.

Kriest, I.: Different parameterizations of marine snow in a 1D-model and their influence on representation of marine snow, nitrogen budget and sedimentation, Deep-Sea Research I, 49, 2133–2162, https://doi.org/10.1016/S0967-0637(02)00127-9, 2002.

Kriest, I. and Evans, G. T.: Representing phytoplankton aggregates in biogeochemical models, Deep-Sea Research I, 46, 1841–1859, https://doi.org/10.1016/S0967-0637(99)00032-1, 1999.

Kriest, I. and Evans, G. T.: A vertically resolved model for phytoplankton aggregation, Journal of Earth System Science, 109, 453–469, https://doi.org/10.1007/bf02708333, 2000.

Kriest, I. and Oschlies, A.: Swept under the carpet: Organic matter burial decreases global ocean biogeochemical model sensitivity to remineralization length scale, Biogeosciences, 10, 8401–8422, https://doi.org/10.5194/bg-10-8401-2013, 2013.

Kriest, I., Khatiwala, S., and Oschlies, A.: Towards an assessment of simple global marine biogeochemical models of different complexity, Progress in Oceanography, 86, 337–360, https://doi.org/10.1016/j.pocean.2010.05.002, 2010.

Kriest, I., Sauerland, V., Khatiwala, S., Srivastav, A., and Oschlies, A.: Calibrating a global three-dimensional biogeochemical ocean model (MOPS-1.0), Geoscientific Model Development, 10, 127–154, https://doi.org/10.5194/gmd-10-127-2017, 2017.

Kwon, E. Y. and Primeau, F. W.: Optimization and sensitivity of a global biogeochemistry ocean model using combined in situ DIC, alkalinity, and phosphate data, Journal of Geophysical Research, 113, C08 011, https://doi.org/10.1029/2007JC004520, 2008.

Kwon, E. Y., Primeau, F. W., and Sarmiento, J. L.: The impact of remineralization depth on the air-sea carbon balance, Nature Geoscience, 2, 630–635, https://doi.org/10.1038/ngeo612, 2009.

Lalli, C. M. and Parsons, T. R.: Biological oceanography: An introduction, Elsevier Butterworth-Heinemann, Vancouver, Canada, 2nd edn., https://doi.org/10.1016/s0022-0981(97)00038-5, 1997.

Lam, P. J., Doney, S. C., and Bishop, J. K. B.: The dynamic ocean biological pump: Insights from a global compilation of particulate organic carbon, CaCO3 and opal concentration profiles from the mesopelagic, Global Biogeochemical Cycles, 25, GB3009, https://doi.org/10.1029/2010GB003868, 2011.

Lam, P. J., Ohnemus, D. C., and Auro, M. E.: Size-fractionated major particle composition and concentrations from the U.S. GEOTRACES North Atlantic Zonal Transect, Deep-Sea Research II, 116, 303–320, https://doi.org/10.1016/j.dsr2.2014.11.020, 2015.

Lampitt, R. S. and Antia, A. N.: Particle flux in deep seas: Regional characteristics and temporal variability, Deep-Sea Research I, 44, 1377–1403, https://doi.org/10.1016/S0967-0637(97)00020-4, 1997.

Lampitt, R. S., Bett, B. J., Kiriakoulakis, K., Popova, E. E., Ragueneau, O., Vangriesheim, A., and Wolff, G. A.: Material supply to the abyssal seafloor in the northeast Atlantic, Progress in Oceanography, 50, 27–63, https://doi.org/10.1016/S0079-6611(01)00047-7, 2001.

Lampitt, R. S., Boorman, B., Brown, L., Lucas, M. I., Salter, I., Sanders, R. J., Saw, K., Seeyave, S., Thomalla, S. J., and Turnewitsch, R.: Particle export from the euphotic zone: Estimates using a novel drifting sediment trap, 234Th and new production, Deep-Sea Research I, 55, 1484–1502, https://doi.org/10.1016/j.dsr.2008.07.002, 2008.

Lampitt, R. S., Salter, I., De Cuevas, B. A., Hartman, S. E., Larkin, K. E., and Pebody, C. A.: Long-term variability of downward particle flux in the deep northeast Atlantic: Causes and trends, Deep-Sea Research II, 57, 1346–1361, https://doi.org/10.1016/j.dsr2.2010.01.011, 2010.

Lauderdale, J. M. and Cael, B. B.: Impact of remineralization profile shape on the air-sea carbon balance, Geophysical Research Letters, 48, e2020GL091 746, https://doi.org/10.1029/2020GL091746, 2021.

Laufkötter, C., Vogt, M., and Gruber, N.: Long-term trends in ocean plankton production and particle export between 1960-2006, Biogeosciences, 10, 7373–7393, https://doi.org/10.5194/bg-10-7373-2013, 2013.

Laufkötter, C., John, J. G., Stock, C. A., and Dunne, J. P.: Temperature and oxygen dependence of the remineralization of organic matter, Global Biogeochemical Cycles, 31, 1038–1050, https://doi.org/10.1002/2017GB005643, 2017.

Laurenceau-Cornec, E. C., Trull, T. W., Davies, D. M., De La Rocha, C. L., and Blain, S.: Phytoplankton morphology controls on marine snow sinking velocity, Marine Ecology Progress Series, 520, 35–56, https://doi.org/10.3354/meps11116, 2015.

Laurenceau-Cornec, E. C., Le Moigne, F. A. C., Gallinari, M., Moriceau, B., Toullec, J., Iversen, M. H., Engel, A., and De La Rocha, C. L.: New guidelines for the application of Stokes' models to the sinking velocity of marine aggregates, Limnology and Oceanography, 65, 1264–1285, https://doi.org/10.1002/lno.11388, 2019.

Laws, E. A., Falkowski, P. G., Smith, W. O., Ducklow, H. W., and McCarthy, J. J.: Temperature effects on export production in the open ocean, Global Biogeochemical Cycles, 14, 1231–1246, https://doi.org/10.1029/1999GB001229, 2000.

Laws, E. A., D'Sa, E., and Naik, P.: Simple equations to estimate ratios of new or export production to total production from satellite-derived estimates of sea surface temperature and primary production, Limnology and Oceanography: Methods, 9, 593–601, https://doi.org/10.4319/lom.2011.9.593, 2011.

Le Borgne, R. and Gesbert., H.: Campagne océanographique FLUPAC à bord du N. O. l'ATALANTE, 23 septembre au 29 octobre 1994. Recueil des données. Tome 2: Optique marine, matière organique dissoute, pigments photosynthétiques, observations microscopiques, production primaire, "broutage",, Tech. rep., Archives Sci. Mer Océanogr., 1995.

Le Moigne, F. A. C.: Pathways of organic carbon downward transport by the oceanic biological carbon pump, Frontiers in Marine Science, 6, 634, https://doi.org/10.3389/fmars.2019.00634, 2019.

Le Moigne, F. A. C., Villa-Alfageme, M., Sanders, R. J., Marsay, C. M., Henson, S. A., and García-Tenorio, R.: Export of organic carbon and biominerals derived from 234Th and 210Po at the Porcupine Abyssal Plain, Deep-Sea Research I, 72, 88–101, https://doi.org/10.1016/j.dsr.2012.10.010, 2013.

Le Moigne, F. A. C., Poulton, A. J., Henson, S. A., Daniels, C. J., Fragoso, G. M., Mitchell, E., Richier, S., Russell, B. C., Smith, H. E. K., Tarling, G. A., Young, J. R., and Zubkov, M.: Carbon export efficiency and phytoplankton community composition in the Atlantic sector of the Arctic Ocean, Journal of Geophysical Research: Oceans, 120, 3896–3912, https://doi.org/10.1002/2015JC010700, 2015.

Le Quéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Da Cunha, L. C., Geider, R. J., Giraud, X., Klaas, C. M., Kohfeld, K. E., Legendre, L., Manizza, M., Platt, T., Rivkin, R. B., Sathyendranath, S., Uitz, J., Watson, A. J., and Wolf-Gladrow, D. A.: Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models, Global Change Biology, 11, 2016–2040, https://doi.org/10.1111/j.1365-2486.2005.1004.x, 2005.

Le Quéré, C., Buitenhuis, E. T., Moriarty, R., Alvain, S., Aumont, O., Bopp, L., Chollet, S., Enright, C., Franklin, D. J., Geider, R. J., Harrison, S. P., Hirst, A. G., Larsen, S., Legendre, L., Platt, T., Prentice, I. C., Rivkin, R. B., Sailley, S., Sathyendranath, S., Stephens, N., Vogt, M., and Vallina, S. M.: Role of zooplankton dynamics for Southern Ocean phytoplankton biomass and global biogeochemical cycles, Biogeosciences, 13, 4111–4133, https://doi.org/10.5194/bg-13-4111-2016, 2016.

Lehman, J. T., Botkin, D. B., and Likens, G. E.: The assumptions and rationales of a computer model of phytoplankton population dynamics, Limnology and Oceanography, 20, 343–364, https://doi.org/10.4319/lo.1975.20.3.0343, 1975.

Lévy, M., Ferrari, R., Franks, P. J. S., Martin, A. P., and Rivière, P.: Bringing physics to life at the submesoscale, Geophysical Research Letters, 39, L14 602, https://doi.org/10.1029/2012GL052756, 2012.

Lewis, E. and Wallace, D. W. R.: Program developed for CO2 system calculations, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, TN, ornl/cdiac edn., URL http://cdiac.ornl.gov/oceans/co2rprt.html, 1998.

Li, W. K.: Macroecological patterns of phytoplankton in the northwestern North Atlantic Ocean, Nature, 419, 154–157, https://doi.org/10.1038/nature00994, 2002.

Li, X. and Logan, B. E.: Size distributions and fractal properties of particles during a simulated phytoplankton bloom in a mesocosm, Deep-Sea Research II, 42, 125–138, https://doi.org/10.1016/0967-0645(95)00008-E, 1995.

Li, X. and Logan, B. E.: Collision frequencies of fractal aggregates with small particles by differential sedimentation, Environmental Science and Technology, 31, 1229–1236, https://doi.org/10.1021/es960771w, 1997a.

Li, X. and Logan, B. E.: Collision frequencies between fractal aggregates and small particles in a turbulently sheared fluid, Environmental Science and Technology, 31, 1237–1242, https://doi.org/10.1021/es960772o, 1997b.

Lima, I. D., Lam, P. J., and Doney, S. C.: Dynamics of particulate organic carbon flux in a global ocean model, Biogeosciences, 11, 1177–1198, https://doi.org/10.5194/bg-11-1177-2014, 2014.

Litchman, E., Klausmeier, C. A., Miller, J. R., Schofield, O. M. E., and Falkowski, P. G.: Multi-nutrient, multi-group model of present and future oceanic phytoplankton communities, Biogeosciences, 3, 585–606, https://doi.org/10.5194/bg-3-585-2006, 2006.

Litchman, E., Klausmeier, C. A., Schofield, O. M. E., and Falkowski, P. G.: The role of functional traits and trade-offs in structuring phytoplankton communities: Scaling from cellular to ecosystem level, Ecology Letters, 10, 1170–1181, https://doi.org/10.1111/j.1461-0248.2007.01117.x, 2007.

Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., García, H. E., Baranova, O. K., Zweng, M. M., Paver, C. R., Reagan, J. R., Johnson, D. R., Hamilton, M., and Seidov, D.: World Ocean Atlas 2013. Volume 1: Temperature, Tech. rep., National Oceanic and Atmospheric Administration, 2013.

Logan, B. E. and Wilkinson, D. B.: Fractal geometry of marine snow and other biological aggregates, Limnology and Oceanography, 35, 130–136, https://doi.org/10.4319/lo.1990.35.1.0130, 1990.

Lomas, M. W., Bates, N. R., Johnson, R. J., Knap, A. H., Steinberg, D. K., and Carlson, C. A.: Two decades and counting: 24-years of sustained open ocean biogeochemical measurements in the Sargasso Sea, Deep-Sea Research II, 93, 16–32, https://doi.org/10.1016/j.dsr2.2013.01.008, 2013.

Longhurst, A.: Seasonal cycles of pelagic production and consumption, Progress in Oceanography, 36, 77–167, https://doi.org/10.1016/0079-6611(95)00015-1, 1995.

Lutz, M. J., Dunbar, R. B., and Caldeira, K.: Regional variability in the vertical flux of particulate organic carbon in the ocean interior, Global Biogeochemical Cycles, 16, 1037, https://doi.org/10.1029/2000gb001383, 2002.

Lutz, M. J., Caldeira, K., Dunbar, R. B., and Behrenfeld, M.: Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean, Journal of Geophysical Research, 112, C10 011, https://doi.org/10.1029/2006JC003706, 2007.

Mackinson, B. L., Moran, S. B., Lomas, M. W., Stewart, G. M., and Kelly, R. P.: Estimates of micro- , nano- , and picoplankton contributions to particle export in the northeast Pacific, Biogeosciences, 12, 3429–3446, https://doi.org/10.5194/bg-12-3429-2015, 2015.

Maerz, J., Six, K. D., Stemmler, I., Ahmerkamp, S., and Ilyina, T.: Microstructure and composition of marine aggregates as co-determinants for vertical particulate organic carbon transfer in the global ocean, Biogeosciences, 17, 1765–1803, https://doi.org/10.5194/bg-17-1765-2020, 2020.

Maggi, F.: Variable fractal dimension: A major control for floc structure and flocculation kinematics of suspended cohesive sediment, Journal of Geophysical Research, 112, C07 012, https://doi.org/10.1029/2006JC003951, 2007.

Maggi, F.: The settling velocity of mineral, biomineral, and biological particles and aggregates in water, Journal of Geophysical Research: Oceans, 118, 2118–2132, https://doi.org/10.1002/jgrc.20086, 2013.

Maier-Reimer, E., Kriest, I., Segschneider, J., and Wetzel, P.: The HAMburg Ocean Carbon Cycle Model HAMOCC 5.1. Technical Description Release 1.1, Tech. rep., Max Planck Institute for Meteorology, Hamburg, 2005.

Maiti, K., Benitez-Nelson, C. R., Lomas, M. W., and Krause, J. W.: Biogeochemical responses to late-winter storms in the Sargasso Sea. Part III-Estimates of export production using 234Th:238U disequilibria and sediment traps, Deep-Sea Research I, 56, 875–891, https://doi.org/10.1016/j.dsr.2009.01.008, 2009.

Marañón, E.: Phytoplankton size structure, Encyclopedia of Ocean Sciences, pp. 445–452, https://doi.org/10.1016/B978-012374473-9.00661-5, 2009.

Margalef, R.: Life-forms of phytoplankton as survival alternatives in an unstable environment, Oceanologica Acta, 1, 493–509, 1978.

Mari, X.: Carbon content and C:N ratio of transparent exopolymeric particles (TEP) produced by bubbling exudates of diatoms, Marine Ecology Progress Series, 183, 59–71, https://doi.org/10.3354/meps183059, 1999.

Mari, X., Passow, U., Migon, C., Burd, A. B., and Legendre, L.: Transparent exopolymer particles: Effects on carbon cycling in the ocean, Progress in Oceanography, 151, 13–37, https://doi.org/10.1016/j.pocean.2016.11.002, 2017.

Marinov, I., Follows, M. J., Gnanadesikan, A., Sarmiento, J. L., and Slater, R. D.: How does ocean biology affect atmospheric pCO2? Theory and models, Journal of Geophysical Research: Oceans, 113, C07 032, https://doi.org/10.1029/2007JC004598, 2008.

Marra, J. F., Barber, R. T., Barber, E., Bidigare, R. R., Chamberlin, W. S., Goericke, R., Hargreaves, B. R., Hiscock, M. R., Iturriaga, R., Johnson, Z. I., Kiefer, D. A., Kinkade, C., Knudson, C., Lance, V., Langdon, C., Lee, Z.-P., Perry, M. J., Smith, W. O., Vaillancourt, R., and Zoffoli, L.: A database of ocean primary productivity from the 14C method, Limnology and Oceanography Letters, 6, 107–111, https://doi.org/10.1002/lol2.10175, 2020.

Marsay, C. M.: Particulate trace metals, carbon and nitrogen in the mesopelagic, Doctor of philosophy, University of Southampton, 2012.

Marsay, C. M., Sanders, R. J., Henson, S. A., Pabortsava, K., and Achterberg, E. P.: Attenuation of sinking particulate organic carbon flux through the mesopelagic ocean, Proceedings of the National Academy of Sciences, 112, 1089–1094, https://doi.org/10.1073/pnas.1415311112, 2015.

Martin, J. H.: Glacial-interglacial CO2 change: The iron hypothesis, Paleoceanography, 5, 1–13, https://doi.org/10.1029/PA005i001p00001, 1990.

Martin, J. H. and Fitzwater, S. E.: Iron deficiency limits phytoplankton growth in the north-east Pacific subarctic, Nature, 331, 31–33, 1988.

Martin, J. H., Knauer, G. A., Karl, D. M., and Broenkow, W. W.: VERTEX: Carbon cycling in the northeast Pacific, Deep-Sea Research, 34, 267–285, https://doi.org/10.1016/0198-0149(87)90086-0, 1987.

Martin, W., Baross, J., Kelley, D., and Russell, M. J.: Hydrothermal vents and the origin of life, Nature Reviews Microbiology, 6, 805–814, https://doi.org/10.1038/nrmicro1991, 2008.

Mayor, D. J., Sanders, R. J., Giering, S. L. C., and Anderson, T. R.: Microbial gardening in the ocean's twilight zone: Detritivorous metazoans benefit from fragmenting, rather than ingesting, sinking detritus, Bioessays, 36, 1132–1137, https://doi.org/10.1002/bies.201400100, 2014.

McCave, I. N.: Size spectra and aggregation of suspended particles in the deep ocean, Deep-Sea Research, 31, 329–352, https://doi.org/10.1016/0198-0149(84)90088-8, 1984.

McDonnell, A. M. P. and Buesseler, K. O.: Variability in the average sinking velocity of marine particles, Limnology and Oceanography, 55, 2085–2096, https://doi.org/10.4319/lo.2010.55.5.2085, 2010.

McDonnell, A. M. P. and Buesseler, K. O.: A new method for the estimation of sinking particle fluxes from measurements of the particle size distribution, average sinking velocity, and carbon content, Limnology and Oceanography: Methods, 10, 329–346, https://doi.org/10.4319/lom.2012.10.329, 2012.

Meakin, P.: Fractals, scaling and growth far from equilibrium, Cambridge Non-linear Science Series 5, Cambridge University Press, Cambridge, 1998.

Meng, S. and Liu, Y.: New insights into transparent exopolymer particles (TEP) formation from precursor materials at various Na+/Ca2+ ratios, Scientific Reports, 6, 19 747, https://doi.org/10.1038/srep19747, 2016.

Miklasz, K. A. and Denny, M. W.: Diatom sinking speeds: Improved predictions and insight from a modified Stokes' law, Limnology and Oceanography, 55, 2513–2525, https://doi.org/10.4319/lo.2010.55.6.2513, 2010.

Millero, F. J.: Thermodynamics of the carbon dioxide system in the oceans, Geochimica et Cosmochimica Acta, 59, 661–677, https://doi.org/10.1016/0016-7037(94)00354-O, 1995.

Milutinovic, S. and Bertino, L.: Assessment and propagation of uncertainties in input terms through an ocean-color-based model of primary productivity, Remote Sensing of Environment, 115, 1906–1917, https://doi.org/10.1016/j.rse.2011.03.013, 2011.

Moore, J. K., Doney, S. C., Kleypas, J. A., Glover, D. M., and Fung, I. Y.: An intermediate complexity marine ecosystem model for the global domain, Deep-Sea Research II, 49, 403–462, https://doi.org/10.1016/S0967-0645(01)00108-4, 2001.

Moore, J. K., Fu, W., Primeau, F. W., Britten, G. L., Lindsay, K., Long, M. C., Doney, S. C., Mahowald, N. M., Hoffman, F., and Randerson, J. T.: Sustained climate warming drives declining marine biological productivity, Science, 359, 1139–1143, https://doi.org/10.1126/science.aao6379, 2018.

Morel, F. M. M., Rueter, J. G., and Price, N. M.: Iron nutrition of phytoplankton and its possible importance in the ecology of ocean regions with high nutrient and low biomass, Oceanography, 4, 56–61, https://doi.org/10.5670/oceanog.1991.03, 1991.

Mouw, C. B., Barnett, A., McKinley, G. A., Gloege, L., and Pilcher, D. J.: Global ocean particulate organic carbon flux merged with satellite parameters, Earth System Science Data, 8, 531–541, https://doi.org/10.5194/essd-8-531-2016, 2016a.

Mouw, C. B., Ciochetto, A., Mckinley, G. A., Gloege, L., and Pilcher, D. J.: Phytoplankton size impact on export flux in the global ocean, Global Biogeochemical Cycles, 30, 1542–1562, https://doi.org/10.1002/2015GB005355, 2016b.

Murray, J. W., Young, J., Newton, J., Dunne, J. P., Chapin, T., Paul, B., and McCarthy, J. J.: Export flux of particulate organic carbon from the central equatorial pacific determined using a combined drifting trap-234Th approach, Deep-Sea Research II, 43, 1095–1132, https://doi.org/10.1016/0967-0645(96)00036-7, 1996.

Najjar, R. G., Jin, X., Louanchi, F., Aumont, O., Caldeira, K., Doney, S. C., Dutay, J., Follows, M. J., Gruber, N., Joos, F., Lindsay, K., Maier-Reimer, E., Matear, R. J., Matsumoto, K., Monfray, P., Mouchet, A., Orr, J., Plattner, G.-K., Sarmiento, J. L., Schlitzer, R., Slater, R. D., Weirig, M.-F., Yamanaka, Y., and Yool, A.: 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, GB3007, https://doi.org/10.1029/2006GB002857, 2007.

Nellemann, C., Corcoran, E., Duarte, C. M., Valdés, L., De Young, C., Fonseca, L., and Grimsditch, G.: Blue carbon: A rapid response assessment, Tech. rep., United Nations Environment Programme, GRID-Arendal, 2009.

Nielsen, T. G. and Thomas, K.: Regulation of zooplankton biomass and production in a temperate, coastal ecosystem. 2. Ciliates, Limnology and Oceanography, 39, 493–507, https://doi.org/10.4319/lo.1994.39.3.0493, 1994.

Niemeyer, D., Kriest, I., and Oschlies, A.: The effect of marine aggregate parameterisations on nutrients and oxygen minimum zones in a global biogeochemical model, Biogeosciences, 16, 3095–3111, https://doi.org/10.5194/bg-16-3095-2019, 2019.

Oguz, T. and Merico, A.: Factors controlling the summer Emiliania huxleyi bloom in the Black Sea sea: A modeling study, Journal of Marine Systems, 59, 173–188, https://doi.org/10.1016/j.jmarsys.2005.08.002, 2006.

Ohman, M. D. and Romagnan, J. B.: Nonlinear effects of body size and optical attenuation on Diel Vertical Migration by zooplankton, Limnology and Oceanography, 61, 765–770, https://doi.org/10.1002/lno.10251, 2016.

Oliver, S.: User Guide: Applying Optimisation Algorithms to Climate Models with OptClimSO, Tech. rep., University of Oxford, 2019.

Oliver, S., Cartis, C., Kriest, I., Tett, S., and Khatiwala, S.: A derivative-free optimisation method for global ocean biogeochemical models, Geoscientific Model Development, In review, 1–24, https://doi.org/10.5194/gmd-2021-175, 2021.

Omand, M. M., D'Asaro, E. A., Lee, C. M., Perry, M. J., Briggs, N., Cetinić, I., and Mahadevan, A.: Eddy-driven subduction exports particulate organic carbon from the spring bloom, Science, 348, 222–225, https://doi.org/10.1126/science.1260062, 2015.

Omand, M. M., Govindarajan, R., He, J., and Mahadevan, A.: Sinking flux of particulate organic matter in the oceans: Sensitivity to particle characteristics, Scientific Reports, 10, 5582, https://doi.org/10.1038/s41598-020-60424-5, 2020.

Owens, S. A., Buesseler, K. O., Lamborg, C. H., Valdes, J., Lomas, M. W., Johnson, R. J., Steinberg, D. K., and Siegel, D. A.: A new time series of particle export from neutrally buoyant sediments traps at the Bermuda Atlantic Time-series Study site, Deep-Sea Research I, 72, 34–47, https://doi.org/10.1016/j.dsr.2012.10.011, 2013.

Pace, M. L., Knauer, G. A., Karl, D. M., and Martin, J. H.: Primary production, new production and vertical flux in the eastern Pacific Ocean, Nature, 325, 803–804, https://doi.org/10.1038/325803a0, 1987.

Parekh, P., Dutkiewicz, S., Follows, M. J., and Ito, T.: Atmospheric carbon dioxide in a less dusty world, Geophysical Research Letters, 33, L03 610, https://doi.org/10.1029/2005GL025098, 2006.

Paris, C. B., Helgers, J., Van Sebille, E., and Srinivasan, A.: Connectivity modeling system: A probabilistic modeling tool for the multi-scale tracking of biotic and abiotic variability in the ocean, Environmental Modelling and Software, 42, 47–54, https://doi.org/10.1016/j.envsoft.2012.12.006, 2013.

Passow, U.: Transparent exopolymer particles (TEP) in aquatic environments, Progress in Oceanography, 55, 287–333, https://doi.org/10.1016/S0079-6611(02)00138-6, 2002.

Passow, U. and Carlson, C. A.: The biological pump in a high CO2 world, Marine Ecology Progress Series, 470, 249–271, https://doi.org/10.3354/meps09985, 2012.

Passow, U., Sweet, J., Francis, S., Xu, C., Dissanayake, A. L., Lin, Y. Y., Santschi, P. H., and Quigg, A.: Incorporation of oil into diatom aggregates, Marine Ecology Progress Series, 612, 65–86, https://doi.org/10.3354/meps12881, 2019.

Pearson, B. and Fox-Kemper, B.: Log-Normal turbulence dissipation in global ocean models, Physical Review Letters, 120, https://doi.org/10.1103/PhysRevLett.120.094501, 2018.

Pincus, R., Batstone, C. P., Hofmann, R. J. P., Taylor, K. E., and Glecker, P. J.: Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models, Journal of Geophysical Research, 113, D14 209, https://doi.org/10.1029/2007JD009334, 2008.

Pinti, J., Kiørboe, T., Thygesen, U. H., and Visser, A. W.: Trophic interactions drive the emergence of diel vertical migration patterns: A game-theoretic model of copepod communities, Proceedings of the Royal Society B, 286, 20191 645, https://doi.org/10.1098/rspb.2019.1645, 2019.

Platt, T., Sathyendranath, S., and Ravindran, P.: Primary production by phytoplankton: Analytic solutions for daily rates per unit area of water surface, Proceedings of the Royal Society B, 241, 101–111, https://doi.org/10.1098/rspb.1990.0072, 1990.

Ploug, H.: Small-scale oxygen fluxes and remineralization in sinking aggregates, Limnology and Oceanography, 46, 1624–1631, https://doi.org/10.4319/lo.2001.46.7.1624, 2001.

Ploug, H., Iversen, M. H., and Fischer, G.: Ballast, sinking velocity, and apparent diffusivity within marine snow and zooplankton fecal pellets: Implications for substrate turnover by attached bacteria, Limnology and Oceanography, 53, 1878–1886, https://doi.org/10.4319/lo.2008.53.5.1878, 2008.

Prowe, A. E. F., Visser, A. W., Andersen, K. H., Chiba, S., and Kiørboe, T.: Biogeography of zooplankton feeding strategy, Limnology and Oceanography, 64, 661–678, https://doi.org/10.1002/lno.11067, 2019.

Raven, J. A. and Geider, R. J.: Temperature and algal growth, New Phytologist, 110, 441–461, https://doi.org/10.1111/j.1469-8137.1988.tb00282.x, 1988.

Reed, R. K.: On estimating Insolation over the ocean, Journal of Physical Oceanography, 7, 482–485, 1977.

Riebesell, U. and Wolf-Gladrow, D. A.: The relationship between physical aggregation of phytoplankton and particle flux: A numerical model, Deep-Sea Research, 39, 1085–1102, https://doi.org/10.1016/0198-0149(92)90058-2, 1992.

Roberts, L.: DFO-LS Documentation, Release 1.0.1, 2018.

Roca-Martí, M., Puigcorbé, V., Van der Loeff, M. R., Katlein, C., Fernández-Méndez, M., Peeken, I., and Masqué, P.: Carbon export fluxes and export efficiency in the central Arctic during the record sea-ice minimum in 2012: A joint 234Th/238U and 210Po/210Pb study, Journal of Geophysical Research: Oceans, 121, 5030–5049, https://doi.org/10.1002/2016JC011816, 2016.

Roca-Martí, M., Benitez-Nelson, C. R., Umhau, B. P., Wyatt, A. M., Clevenger, S. J., Pike, S., Horner, T. J., Estapa, M. L., Resplandy, L., and Buesseler, K. O.: Concentrations, ratios, and sinking fluxes of major bioelements at Ocean Station Papa, Elementa: Science of the Anthropocene, 9, 00 166, https://doi.org/10.1525/elementa.2020.00166, 2021.

Rosati, A. and Miyakoda, K.: A general circulation model for upper ocean simulation, Journal of Physical Oceanography, 18, 1601–1626, 1988.

Saiz, E. and Kiørboe, T.: Predatory and suspension feeding of the copepod Acartia tonsa in turbulent environments, Marine Ecology Progress Series, 122, 147–158, https://doi.org/10.3354/meps122147, 1995.

Sanders, R. J., Henson, S. A., Koski, M., De La Rocha, C. L., Painter, S. C., Poulton, A. J., Riley, J. S., Salihoglu, B., Visser, A. W., Yool, A., Bellerby, R., and Martin, A. P.: The biological carbon pump in the North Atlantic, Progress in Oceanography, 129, 200–218, https://doi.org/10.1016/j.pocean.2014.05.005, 2014.

Sanders, R. J., Henson, S. A., Martin, A. P., Anderson, T. R., Bernardello, R., Enderlein, P., Fielding, S., Giering, S. L. C., Hartmann, M., Iversen, M. H., Khatiwala, S., Lam, P., Lampitt, R. S., Mayor, D. J., Moore, C. M., Murphy, E., Painter, S. C., Poulton, A. J., Saw, K., Stowasser, G., Tarling, G. A., Torres-Valdés, S., Trimmer, M., Wolff, G. A., Yool, A., and Zubkov, M. V.: Controls over ocean mesopelagic interior carbon storage (COMICS): Fieldwork, synthesis, and modeling efforts, Frontiers in Marine Science, 3, 136, https://doi.org/10.3389/fmars.2016.00136, 2016.

Sarmiento, J. L. and Gruber, N.: Ocean biogeochemical dynamics, Princeton University Press, 2006.

Sarmiento, J. L. and Toggweiler, J. R.: A new model for the role of the oceans in determining atmospheric pCO2, Nature, 308, 621–624, https://doi.org/10.1038/308621a0, 1984.

Sarmiento, J. L., Dunne, J. P., Gnanadesikan, A., Key, R. M., Matsumoto, K., and Slater, R. D.: A new estimate of the CaCO3 to organic carbon export ratio, Global Biogeochemical Cycles, 16, 1107, https://doi.org/10.1029/2002gb001919, 2002.

Sarthou, G., Timmermans, K. R., Blain, S., and Tréguer, P.: Growth physiology and fate of diatoms in the ocean: A review, Journal of Sea Research, 53, 25–42, https://doi.org/10.1016/j.seares.2004.01.007, 2005.

Schartau, M. and Oschlies, A.: Simultaneous data-based optimization of a 1D-ecosystem model at three locations in the North Atlantic. Part I — Method and parameter estimates, Journal of Marine Research, 61, 765–793, https://doi.org/10.1357/002224003322981147, 2003.

Schartau, M., Wallhead, P., Hemmings, J., Löptien, U., Kriest, I., Krishna, S., Ward, B. A., Slawig, T., and Oschlies, A.: Reviews and syntheses: Parameter identification in marine planktonic ecosystem modelling, Biogeosciences, 14, 1647–1701, https://doi.org/10.5194/bg-14-1647-2017, 2017.

Schulz, M., Prospero, J. M., Baker, A. R., Dentener, F., Ickes, L., Liss, P. S., Mahowald, N. M., Nickovic, S., García-Pando, C. P., Rodríguez, S., Sarin, M., Tegen, I., and Duce, R. A.: Atmospheric transport and deposition of mineral dust to the ocean: Implications for research needs, Environmental Science and Technology, 46, 10 390–10 404, https://doi.org/10.1021/es300073u, 2012.

Séférian, R., Bopp, L., Gehlen, M., Orr, J., Ethé, C., Cadule, P., Aumont, O., Salas-Mélia, D., Voldoire, A., and Madec, G.: Skill assessment of three earth system models with common marine biogeochemistry, Climate Dynamics, 40, 2549–2573, https://doi.org/10.1007/s00382-012-1362-8, 2013.

Séférian, R., Berthet, S., Yool, A., Palmiéri, J., Bopp, L., Tagliabue, A., Kwiatkowski, L., Aumont, O., Christian, J. R., Dunne, J. P., Gehlen, M., Ilyina, T., John, J. G., Li, H., Long, M. C., Luo, J. Y., Nakano, H., Romanou, A., Schwinger, J., Stock, C., Santana-Falcón, Y., Takano, Y., Tjiputra, J., Tsujino, H., Watanabe, M., Wu, T., Wu, F., and Yamamoto, A.: Tracking Improvement in simulated marine biogeochemistry between CMIP5 and CMIP6, Current Climate Change Reports, 6, 95–119, https://doi.org/10.1007/s40641-020-00160-0, 2020.

Sheldon, R. W., Prakash, A., and Sutcliffe, W. H.: The size distribution of particles in the ocean, Limnology and Oceanography, 17, 327–340, https://doi.org/10.4319/lo.1972.17.3.0327, 1972.

Sherman, E., Moore, J. K., Primeau, F. W., and Tanouye, D.: Temperature influence on phytoplankton community growth rates, Global Biogeochemical Cycles, 30, 550–559, https://doi.org/10.1002/2015GB005272, 2016.

Shima, S. I., Kusano, K., Kawano, A., Sugiyama, T., and Kawahara, S.: The super-droplet method for the numerical simulation of clouds and precipitation: A particle-based and probabilistic microphysics model coupled with a non-hydrostatic model, Quarterly Journal of the Royal Meteorological Society, 135, 1307–1320, https://doi.org/10.1002/qj.441, 2009.

Siegel, D. A. and Deuser, W. G.: Trajectories of sinking particles in the Sargasso Sea: Modeling of statistical funnels above deep-ocean sediment traps, Deep-Sea Research I, 44, 1519–1541, https://doi.org/10.1016/S0967-0637(97)00028-9, 1997.

Siegel, D. A., Buesseler, K. O., Doney, S. C., Sailley, S., Behrenfeld, M., and Boyd, P. W.: Global assessment of ocean carbon export by combining satellite observations and food-web models, Global Biogeochemical Cycles, 28, 181–196, https://doi.org/10.1002/2013GB004743, 2014.

Siegel, D. A., Buesseler, K. O., Behrenfeld, M., Benitez-Nelson, C. R., Boss, E. S., Brzezinski, M. A., Burd, A. B., Carlson, C. A., D'Asaro, E. A., Doney, S. C., Perry, M. J., Stanley, R. H. R., and Steinberg, D. K.: Prediction of the export and fate of global ocean net primary production: The EXPORTS science plan, Frontiers in Marine Science, 3, 22, https://doi.org/10.3389/fmars.2016.00022, 2016.

Sigman, D. M. and Boyle, E. A.: Glacial/interglacial variations in atmospheric carbon dioxide, Nature, 407, 859–869, https://doi.org/10.1038/35038000, 2000.

Sigman, D. M., Hain, M. P., and Haug, G. H.: The polar ocean and glacial cycles in atmospheric CO2 concentration, Nature, 466, 47–55, https://doi.org/10.1038/nature09149, 2010.

Silsbe, G. M., Behrenfeld, M., Halsey, K. H., Milligan, A. J., and Westberry, T.: The CAFE model: A net production model for global ocean phytoplankton, Global Biogeochemical Cycles, 30, 1756–1777, https://doi.org/10.1002/2016GB005521, 2016.

Sinerchia, M., Vallerga, S., and Woods, J. D.: The Lagrangian Ensemble Recruitment Model (LERM), Tech. rep., Virtual Plankton Ecology Technical Report 8, 2008.

Sinerchia, M., Field, A. J., Woods, J. D., Vallerga, S., and Hinsley, W. R.: Using an individual-based model with four trophic levels to model the effect of predation and competition on squid recruitment, ICES Journal of Marine Science, 69, 439–447, https://doi.org/10.1093/icesjms/fsr190, 2012.

Smayda, T. J.: Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea, Limnology and Oceanography, 42, 1137–1153, https://doi.org/10.4319/lo.1997.42.5_part_2.1137, 1997.

Smith, E. L.: Photosynthesis in relation to light and carbon dioxide, Proceedings of the National Academy of Sciences, 22, 504–511, https://doi.org/10.1073/pnas.22.8.504, 1936.

Smoluchowski, M.: Drei vortrage uber diffusion, Brownsche molekularbewegung und koagulation von kolloidteilchen, Physikalische Zeitschrift, 17, 557–571, 1916.

Soltwedel, T., Bauerfeind, E., Bergmann, M., Budaeva, N., Hoste, E., Jaeckisch, N., Von Juterzenka, K., Matthiessen, J., Mokievsky, V., Nöthig, E.-M., Quéric, N.-V., Sablotny, B., Sauter, E., Schewe, I., Urban-Malinga, B., Wegner, J., Wlodarska-Kowalczuk, M., and Klages, M.: HAUSGARTEN: Multidisciplinary Investigations at a Deep-Sea, Long-Term Observatory in the Arctic Ocean, Oceanography, 18, 46–61, https://doi.org/10.5670/oceanog.2005.24, 2005.

Stamieszkin, K., Pershing, A. J., Record, N. R., Pilskaln, C. H., Dam, H. G., and Feinberg, L. R.: Size as the master trait in modeled copepod fecal pellet carbon flux, Limnology and Oceanography, 60, 2090–2107, https://doi.org/10.1002/lno.10156, 2015.

Steinberg, D. K. and Landry, M. R.: Zooplankton and the ocean carbon cycle, Annual Review of Marine Science, 9, 413–444, https://doi.org/10.1146/annurev-marine-010814-015924, 2017.

Steinberg, D. K., Van Mooy, B. A. S., Buesseler, K. O., Boyd, P. W., Kobari, T., and Karl, D. M.: Bacterial vs. zooplankton control of sinking particle flux in the ocean's twilight zone, Limnology and Oceanography, 53, 1327–1338, https://doi.org/10.4319/lo.2008.53.4.1327, 2008.

Stemmann, L. and Boss, E. S.: Plankton and particle size and packaging: From determining optical properties to driving the biological pump, Annual Review of Marine Science, 4, 263–290, https://doi.org/10.1146/annurev-marine-120710-100853, 2012.

Stemmann, L., Jackson, G. A., and Ianson, D.: A vertical model of particle size distributions and fluxes in the midwater column that includes biological and physical processes. Part I: Model formulation, Deep-Sea Research I, 51, 865–884, https://doi.org/10.1016/j.dsr.2004.03.001, 2004.

Stemmann, L., Eloire, D., Sciandra, A., Jackson, G. A., Guidi, L., Picheral, M., and Gorsky, G.: Volume distribution for particles between 3.5 to 2000 μm in the upper 200 m region of the South Pacific Gyre, Biogeosciences, 5, 299–310, https://doi.org/10.5194/bg-5-299-2008, 2008.

Stephens, B. B. and Keeling, R. F.: The influence of antarctic sea ice on glacial-interglacial CO2 variations, Nature, 404, 171–174, https://doi.org/10.1038/35004556, 2000.

Stewart, G., Moran, S. B., Lomas, M. W., and Kelly, R. P.: Direct comparison of 210Po, 234Th and POC particle-size distributions and export fluxes at the Bermuda Atlantic Time-series Study (BATS) site, Journal of Environmental Radioactivity, 102, 479–489, https://doi.org/10.1016/j.jenvrad.2010.09.011, 2011.

Stock, C. A., Dunne, J. P., Fan, S., Ginoux, P., John, J. G., Krasting, J. P., Laufkötter, C., Paulot, F., and Zadeh, N.: Ocean biogeochemistry in GFDL's Earth System Model 4.1 and its response to increasing atmospheric CO2, Journal of Advances in Modeling Earth Systems, 12, e2019MS002 043, https://doi.org/10.1029/2019MS002043, 2020.

Suess, E.: Particulate organic carbon flux in the oceans: Surface productivity and oxygen utilization, Nature, 288, 260–263, https://doi.org/10.1038/288260a0, 1980.

Thornton, D. C.: Dissolved organic matter (DOM) release by phytoplankton in the contemporary and future ocean, European Journal of Phycology, 49, 20–46, https://doi.org/10.1080/09670262.2013.875596, 2014.

Tilstone, G., Smyth, T., Poulton, A. J., and Hutson, R.: Measured and remotely sensed estimates of primary production in the Atlantic Ocean from 1998 to 2005, Deep-Sea Research II, 56, 918–930, https://doi.org/10.1016/j.dsr2.2008.10.034, 2009.

Tilstone, G., Taylor, B. H., Blondeau-Patissier, D., Powell, T., Groom, S. B., Rees, A. P., and Lucas, M. I.: Comparison of new and primary production models using SeaWiFS data in contrasting hydrographic zones of the northern North Atlantic, Remote Sensing of Environment, 156, 473–489, https://doi.org/10.1016/j.rse.2014.10.013, 2015.

Timothy, D. A., Wong, C. S., Page, J., White, L. A., and Macdonald, R. W.: Climatology of sediment flux and composition in the subarctic Northeast Pacific Ocean with biogeochemical implications, Progress in Oceanography, 116, 95–129, https://doi.org/10.1016/j.pocean.2013.06.017, 2013.

Toggweiler, J. R.: Variation of atmospheric CO2 by ventilation of the ocean's deepest water, Paleoceanography, 14, 571–588, https://doi.org/10.1029/1999PA900033, 1999.

Totterdell, I. J.: Description and evaluation of the Diat-HadOCC model v1. 0: The ocean biogeochemical component of HadGEM2-ES, Geoscientific Model Development, 12, 4497–4549, https://doi.org/10.5194/gmd-12-4497-2019, 2019.

Turner, J. T.: Zooplankton fecal pellets, marine snow, phytodetritus and the ocean's biological pump, Progress in Oceanography, 130, 205–248, https://doi.org/10.1016/j.pocean.2014.08.005, 2015.

Van Cappellen, P., Dixit, S., and Van Beusekom, J.: Biogenic silica dissolution in the oceans: Reconciling experimental and field-based dissolution rates, Global Biogeochemical Cycles, 16, https://doi.org/10.1029/2001GB001431, 2002.

Van Heuven, S., Pierrot, D., Rae, J. W. B., Lewis, E., and Wallace, D. W. R.: MATLAB program developed for CO2 system calculations, vol. ORNL/CDIAC, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee, https://doi.org/10.3334/CDIAC/otg.CO2SYSS_MATLAB_v1.1, 2011.

Van Sebille, E., Scussolini, P., Durgadoo, J. V., Peeters, F. J., Biastoch, A., Weijer, W., Turney, C., Paris, C. B., and Zahn, R.: Ocean currents generate large footprints in marine palaeoclimate proxies, Nature Communications, 6, 6521, https://doi.org/10.1038/ncomms7521, 2015.

Van Sebille, E., Aliani, S., Law, K. L., Maximenko, N., Alsina, J. M., Bagaev, A., Bergmann, M., Chapron, B., Chubarenko, I., Cózar, A., Delandmeter, P., Egger, M., Fox-Kemper, B., Garaba, S. P., Goddijn-Murphy, L., Hardesty, B. D., Hoffman, M. J., Isobe, A., Jongedijk, C. E., Kaandorp, M. L., Khatmullina, L., Koelmans, A. A., Kukulka, T., Laufkötter, C., Lebreton, L., Lobelle, D., Maes, C., Martínez-Vicente, V., Morales-Maqueda, M. Á., Poulain-Zarcos, M., Rodríguez, E., Ryan, P. G., Shanks, A. L., Shim, W. J., Suaria, G., Thiel, M., Van Den Bremer, T. S., and Wichmann, D.: The physical oceanography of the transport of floating marine debris, Environmental Research Letters, 15, 023 003, https://doi.org/10.1088/1748-9326/ab6d7d, 2020.

Villa-Alfageme, M., De Soto, F., Ceballos-Romero, E., Giering, S. L. C., Le Moigne, F. A. C., Henson, S. A., Mas, J. L., and Sanders, R. J.: Geographical, seasonal, and depth variation in sinking particle speeds in the North Atlantic, Geophysical Research Letters, 43, 8609–8616, https://doi.org/10.1002/2016GL069233, 2016.

Visser, A. W.: Motility of zooplankton: Fitness, foraging and predation, Journal of Plankton Research, 29, 447–461, https://doi.org/10.1093/plankt/fbm029, 2007.

Visser, A. W. and Fiksen, Ø.: Optimal foraging in marine ecosystem models: Selectivity, profitability and switching, Marine Ecology Progress Series, 473, 91–101, https://doi.org/10.3354/meps10079, 2013.

Volk, T. and Hoffert, M. I.: Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO2 changes, in: The Carbon Cycle and Atmospheric CO2: Natural Variations Archean to Present, edited by Broecker, W. S. and Sundquist, E. T., vol. 32, pp. 99–110, American Geophysical Union (AGU), https://doi.org/10.1029/GM032p0099, 1985.

Ward, B. A., Friedrichs, M. A. M., Anderson, T. R., and Oschlies, A.: Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models, Journal of Marine Systems, 81, 34–43, https://doi.org/10.1016/j.jmarsys.2009.12.005, 2010.

Ward, B. A., Schartau, M., Oschlies, A., Martin, A. P., Follows, M. J., and Anderson, T. R.: When is a biogeochemical model too complex? Objective model reduction and selection for North Atlantic time-series sites, Progress in Oceanography, 116, 49–65, https://doi.org/10.1016/j.pocean.2013.06.002, 2013.

Ward, B. A., Dutkiewicz, S., and Follows, M. J.: Modelling spatial and temporal patterns in size-structured marine plankton communities: Top-down and bottom-up controls, Journal of Plankton Research, 36, 31–47, https://doi.org/10.1093/plankt/fbt097, 2014.

Weber, T., Cram, J. A., Leung, S. W., DeVries, T., and Deutsch, C.: Deep ocean nutrients imply large latitudinal variation in particle transfer efficiency, Proceedings of the National Academy of Sciences, 113, 8606–8611, https://doi.org/10.1073/pnas.1604414113, 2016.

Westberry, T., Behrenfeld, M. J., Siegel, D. A., and Boss, E. S.: Carbon-based primary productivity modeling with vertically resolved photoacclimation, Global Biogeochemical Cycles, 22, GB2024, https://doi.org/10.1029/2007GB003078, 2008.

Wetherill, G. W.: Comparison of analytical and physical modeling of planetesimal accumulation, Icarus, 88, 336–354, https://doi.org/10.1016/0019-1035(90)90086-O, 1990.

White, F. M.: Viscous fluid flow, McGraw Hill, New York, 1974.

Wilson, J. D., Barker, S., Edwards, N. R., Holden, P. B., and Ridgwell, A.: Sensitivity of atmospheric CO2 to regional variability in particulate organic matter remineralization depths, Biogeosciences, 16, 2923–2936, https://doi.org/10.5194/bg-16-2923-2019, 2019.

Wilson, R. J., Speirs, D. C., and Heath, M. R.: On the surprising lack of differences between two congeneric calanoid copepod species, Calanus finmarchicus and C. helgolandicus, Progress in Oceanography, 134, 413–431, https://doi.org/10.1016/j.pocean.2014.12.008, 2015.

Wong, C. S., Waser, N. A. D., Whitney, F. A., Johnson, W. K., and Page, J.: Time-series study of the biogeochemistry of the North East subarctic Pacific: Reconciliation of the C org / N remineralization and uptake ratios with the Redfield ratios, Deep-Sea Research II, 49, 5717–5738, https://doi.org/10.1016/S0967-0645(02)00211-4, 2002.

Woods, J. D.: The Lagrangian Ensemble metamodel for simulating plankton ecosystems, Progress in Oceanography, 67, 84–159, https://doi.org/10.1016/j.pocean.2005.04.003, 2005.

Woods, J. D. and Onken, R.: Diurnal variation and primary production in the ocean preliminary results of a Lagrangian ensemble model, Journal of Plankton Research, 4, 735–756, https://doi.org/10.1093/plankt/4.3.735, 1982.

Xue, P., Schwab, D. J., Zhou, X., Huang, C., Kibler, R., and Ye, X.: A hybrid Lagrangian-Eulerian particle model for ecosystem simulation, Journal of Marine Science and Engineering, 6, 109, https://doi.org/10.3390/jmse6040109, 2018.

Yool, A., Popova, E. E., and Anderson, T. R.: MEDUSA-2.0: An intermediate complexity biogeochemical model of the marine carbon cycle for climate change and ocean acidification studies, Geoscientific Model Development, 6, 1767–1811, https://doi.org/10.5194/gmd-6-1767-2013, 2013.

Zahnow, J. C., Maerz, J., and Feudel, U.: Particle-based modeling of aggregation and fragmentation processes: Fractal-like aggregates, Physica D, 240, 882–893, https://doi.org/10.1016/j.physd.2011.01.003, 2011.

Zwirglmaier, K., Jardillier, L., Ostrowski, M., Mazard, S., Garczarek, L., Vaulot, D., Not, F., Massana, R., Ulloa, O., and Scanlan, D. J.: Global phylogeography of marine Synechococcus and Prochlorococcus reveals a distinct partitioning of lineages among oceanic biomes, Environmental Microbiology, 10, 147–161, https://doi.org/10.1111/j.1462-2920.2007.01440.x, 2008.